PROCEEDINGS

DMS 2008

The 14th International Conference on Distributed Multimedia Systems

Sponsored by

Knowledge Systems Institute Graduate School, USA

Technical Program

September 4 - 6, 2008

Hyatt Harborside Hotel, Boston, Massachusetts, USA

Organized by

Knowledge Systems Institute Graduate School

ISBN 1-891706-23-3 (paper)

Additional Copies can be ordered from:
Knowledge Systems Institute Graduate School
3420 Main Street
Skokie, IL 60076, USA
Tel:+1-847-679-3135
Fax:+1-847-679-3166
Email:office@ksi.edu
http://www.ksi.edu

Proceedings preparation, editing and printing are sponsored by
Knowledge Systems Institute Graduate School

Printed by Knowledge Systems Institute Graduate School

Foreword

Welcome to Boston. The International Conference on Distributed Multimedia Systems has entered its fourteenth year. For the past thirteen years, the Conference on Distributed Multimedia Systems has covered a wide spectrum of paper presentations, technical discussions and demonstrations in the field of distributed multimedia systems. The main topics of the DMS2008 conference are: network and systems, emergency management and security, situated computing, multimedia software engineering, multimedia and geo-systems, distributed multimedia computing, multimedia information retrieval, digital home and health care. This year's technical program also includes two workshops: International Workshop on Distance Education Technologies (DET2008) and International Workshop on Visual Languages and Computing (VLC'2008) with a special track on sketch computing.

The DMS2008 Program Committee selected papers for publication in the proceedings and presentation at the Conference based upon a rigorous review process. This year, authors from nineteen countries will present papers at the conference: Brazil, Canada, France, Germany, India, Iran, Italy, Japan, Malta, Netherlands, Portugal, Romania, South Korea, Spain, Sweden, Taiwan, United Kingdom, United States and Venezuela.

We appreciate having had the opportunity to serve as the program co-chairs for this conference, and are very grateful for the outstanding efforts provided by the organizers and program committee members of the above mentioned workshops and special sessions and tracks. The Program Committee members and reviewers provided excellent support in promptly reviewing the manuscripts. We are grateful to the authors and sessions chairs for their time and efforts to make DMS2008 a success. The support of the Computer Science Department of the University of Venice in the organization of the special session on multimedia and geosystems is gratefully acknowledged, as well as the support of the University of Salerno in the organization of several special sessions on visual languages, gesture computing and others, and other universities' support of similar efforts. As always, Dr. S. K. Chang of the Knowledge Systems Institute, USA, provided excellent guidance throughout the effort. Last but not least, we all owe a debt of gratitude to the heroic efforts of Mr. Daniel Li, as well as other staff members of Knowledge Systems Institute.

Erland Jungert and Masahito Hirakawa
DMS2008 Program Co-Chairs

The 14th International Conference on Distributed Multimedia Systems (DMS 2008)

September 4-6, 2008
Hyatt Harborside Hotel, Boston, Massachusetts, USA

Organizers & Committees

Steering Committee Chair

Shi-Kuo Chang, *University of Pittsburgh, USA*

Conference Chair

Son T. Vuong, *University of British Columbia, Canada*

Program Co-Chairs

Masahito Hirakawa, *Shimane University, Japan*
Erland Jungert, *FOA, Sweden*

Program Committee

Vasu Alagar, *Concordia University, Canada*
Frederic Andres, *National Institute of Informatics, Japan*
Arvind K. Bansal, *Kent State University, USA*
Ioan Marius Bilasco, *Laboratoire d'Informatique de Grenoble (LIG), France*
Augusto Celentano, *Universita Ca Foscari di Venezia, Italy*
Yeim-Kuan Chang, *National Cheng Kung University, Taiwan*
Ing-Ray Chen, *Virginia Tech (VPI&SU), USA*
Shu-Ching Chen, *Florida International University, USA*
Cheng-Chung Chu, *Tunghai University, Taiwan*
Chih-Ping Chu, *National Cheng Kong University, Taiwan*
Gennaro Costagliola, *Univ of Salerno, Italy*

Joseph E. Urban, *Arizona State University, USA*
Mario Vacca, *University of Salerno, Italy*
Athena Vakali, *Aristotle University, Greece*
Ellen Walker, *Hiram College, USA*
Chi-Lu Yang, *Institute for Information Industry, Taiwan*
Kang Zhang, *University of Texas at Dallas, USA*

Proceedings Cover Design

Gabriel Smith, *Knowledge Systems Institute Graduate School, USA*

Conference Secretariat

Judy Pan, *Chair, Knowledge Systems Institute Graduate School, USA*
Jessica Braunstein, *Knowledge Systems Institute Graduate School, USA*
Grace Lisboa, *Knowledge Systems Institute Graduate School, USA*
David Huang, *Knowledge Systems Institute Graduate School, USA*
Daniel Li, *Knowledge Systems Institute Graduate School, USA*

International Workshop on Distance Education Technologies (DET 2008)

September 4-6, 2008
Hyatt Harborside Hotel, Boston, Massachusetts, USA

Organizers & Committees

Workshop Chair

Paolo Maresca, *University Federico II, Napoli, Italy*

Program Co-Chairs

H. S. Chen, *National Taiwan University, Taiwan*
Tim Arndt, *Cleveland State University, USA*

Program Committee

Alfio Andronico, *University of Siena, Italy*
Giovanni Adorni, *University of Genova, Genova*
Heng-Shuen Chen, *National Taiwan University, Taiwan*
Ching-Yang Chou, *National Kaohsiung Normal University, Taiwan*
Yuan-Sun Chu, *National Chung Cheng University, Taiwan*
Luigi Colazzo, *University di Trento, Italy*
Rita Francese, *University of Salerno, Italy*
Wu He, *Old Dominion University, USA*
Kao-Shing Hwang, *National Chung Cheng University, Taiwan*
Pedro Isaias, *Portuguese Open University Portugal*
Jer-Junn Luh, *National Taiwan University, Taiwan*
Paolo Maresca, *University "Federico II" , Napoli, Italy*
Teresa Roselli, *University of Bari, Italy*
Giuseppe Scanniello, *University of Salerno, Italy*
Hui-Kai Su, Nanhua *University, Taiwan*
Goran Trajkovski, *South University, USA*

International Workshop on Visual Languages and Computing (VLC 2008)

September 4-6, 2008
Hyatt Harborside Hotel, Boston, Massachusetts, USA

Organizers & Committees

Workshop Co-Chairs

Kang Zhang, *University of Texas at Dallas, USA*
Paul Fishwick, *University of Florida, USA*

Program Co-Chairs

Giuseppe Polese, *University of Salerno, Italy*
Giuliana Vitiello, *University of Salerno, Italy*

Program Committee

Paolo Buono, *University of Bari, Italy*
Kendra Cooper, *University of Texas at Dallas, USA*
Philip Cox, *Dalhousie University, Canada*
Vincenzo Deufemia, *University of Salerno, Italy*
Stephan Diehl, *University of Trier, Germany*
Jing Dong, *University of Texas at Dallas, USA*
Filomena Ferrucci, *University of Salerno, Italy*
Andrew Fish, *University of Brighton, UK*
Corin Gurr, *University of Reading, UK*
Chris Hundhausen, *Washington State University, USA*
Erland Jungert, *Swedish Defence Research Establishment, Sweden*
Jun Kong, *North Dokota State University, USA*
Zenon Kulpa, *Institute of Fundamental Technological Research*
Robert Laurini, *University of Lyon, France*
Benjamin Lok, *University of Florida, USA*

Table of Contents

DMS-2 Distributed Multimedia Systems II

DMS-3 Distributed Multimedia Computing

DMS-4 Distributed Multimedia Systems III

DMS-5 Digital Home and Healthcare

DMS-6 Multimedia Software Engineering I

DMS-7 Multimedia Software Engineering II

DMS-8 Multimedia Software Engineering III

DMS-9 Emegrency Management and Security

DMS-10 Situated Computing

DMS-11 Multimedia and Geosystems

DMS-12 Content-based Multimedia Systems and Models

VLC Workshop

VLC-1 Visualization Techniques for Environmental and Cultural Heritage

VLC-2 Sketch Computing

VLC-3 Sketch Computing

VLC-4 Visual Models and Metaphors for Computational Process Visualization

DET Workshop

DET-1 Eclipse Open Source Ecosystem and E-learning

DET-2 Intelligent Resources for E-learning

DET-3 Distance Learning for Culture and Arts

DET-4 Educational Technologies Practical and New Learning Models

Note: (S) means short paper.

Keynote I:
On-Demand P2P Media Streaming Over the Internet

Son T. Vuong

Abstract

On-demand media streaming has gained considerable attention in recent years owing to its promising usage in a rich set of Internet-based services such as movies on demand and distance learning. However, the current solutions for on-demand media streaming, primarily based on the client-server model, are not economical and not scalable due to the high cost in deploying the servers and the heavy media traffic overload on the servers.

We propose a scalable cost-effective P2P solution called **Bit-Vampire** that aggregates the peers' storage and bandwidth to facilitate on-demand media streaming. In this talk, we will discuss the state of the art of P2P on-demand video streaming and in particular present a novel category overlay P2P architecture that enables efficient content search called **COOL** (CategOry OverLay) search, and the concept of Bit-Vampire that makes use of COOL search and a scheduling algorithm for efficient concurrent media streaming from multiple peers. Other related practical issues such as efficient pre-release distribution for flash crowd congestion avoidance are also discussed. A movie streaming demo is also presented to illustrate the application of Bit-Vampire that will enable the "Internet DVD player" technology in the future.

About Dr. Son T. Vuong

Prof. Dr. Son T. Vuong received his Ph.D. in Computer Science from the University of Waterloo, Canada where he was a lecturer and assistant professor in 1981-1982. Since 1983, he has been a professor in the Computer Science Department at the University of British Columbia in Vancouver, Canada. He was a founder of the Distributed System Research Group and is now Director of the Laboratory for Networks and Internet Computing (NICLab). He is an internationally renowned researcher on protocol engineering, distributed multimedia systems, and collaborative computing. His areas of research interest in recent years also include semantic search, p2p video streaming, mobile agent technology and applications, network security, and eLearning. He has co-authored a US patent, written over 180 papers and co-edited three books, including the book on "Recent Advances in Distributed Multimedia Systems" published in 1999. Dr. Vuong has served on many conference program committees and was co-chair and co-organizer of eight international conferences (Multimedia'08, DMS'08, NOMS'06, DMS'97, ICDCS'95, PSTV'94, FORTE'89, IWPTS'88).

Keynote II:
Visual Summaries of Geographic Databases

Robert Laurini

Abstract

For several applications, it is interesting to have an idea of database contents. For instance in geo-marketing, a decision-maker could be interested in knowing the places where his products are not well sold. In other words, the summary of a database must include only the important items.

The goal of this paper will be to present an international project aiming at generating visual summaries of geographic databases based on chorems. For years, chorems – as schematized representations of territories invented by a geographer named Brunet – are used to represent the more salient aspects of a territory. But in this methodology, the designer is supposed to know everything and to propose a schema. In contrast, in our project, the starting point is a geographic database representing information over a territory for which one wants to generate a drawing emphasizing only the salient aspects called patterns. So, two main problems exist, first to extract the salient patterns from spatial data mining, and the second visualizing the results after having done both a geometric generalization of the territory boundaries, and a semantic generalization of the contents.

This paper will first present the problems of discovering patterns and selecting the more important ones (5 ± 2). Then, we will explain how to make the geometric and semantic generalizations. After having presented a model of patterns, we will present the ChorML language used for both for storing and visualizing them.

LAURINI R., MILLERET-RAFFORT F., LOPEZ K. (2006) "A Primer of Geographic Databases Based on Chorems". In proceedings of the SebGIS Conference, Montpellier, Published by Springer Verlag LNCS 4278, pp. 1693-1702.

DEL FATTO V., LAURINI R., LOPEZ K., LORETO R., MILLERET-RAFFORT F., SEBILLO M., SOL-MARTINEZ D., VITIELLO G. (2007) "Potentialities of Chorems as Visual Summaries of Spatial Databases Contents", VISUAL 2007, 9th International Conference on Visual Information Systems, Shanghai, China, 28-29 June 2007. Edited by Qiu, G., Leung, C., Xue, X.-Y., Laurini, R., Springer Verlag LNCS, Volume 4781 "Advances in Visual Information Systems", pp. 537-548.

DEL FATTO V., LAURINI R., LOPEZ K., SEBILLO M. TORTORA G., TUCCI M., VITIELLO G. (2007) " Chorem Editing - Visual Summary of Spatial Database Content". In Proceedings of the 13th "International Conference on Distributed Multimedia Systems - Workshop on Visual Languages and Computing (VLC2007)" San Francisco USA, 6-8 Sept. 2007, pp.256-259.

About Dr. Robert Laurini

Born in 1947, Prof. Robert LAURINI is a highly distinguished professor at the National Institute for Applied Sciences (INSA-Lyon), and more particularly at the Computing Department and at the LIRIS Laboratory. He speaks fluent French, English, Italian and Spanish.

Previously, he carried out research at the University of Cambridge, UK (1976-1977) and at the University of Maryland, USA (1986-1987). For five years, he was head of the Laboratory for Information System Engineering (LISI), shared by the National Institute for Applied Sciences (INSA) and Claude Bernard University of Lyon, France. Until recently, he was deputy head of the LIRIS Research Center for Images and Information Systems from January 2003 to December 2006.

He is also member of ACM and of IEEE Computer Society. In November 1998, he received an ACM award for Technical Contributions and Leadership. He is also:

- Vice-president of UDMS (Urban Data Management Society), a European association whose goal is the promotion of urban information systems.
- Associate editor of "Journal of Visual Computing and Languages" and of "Computers, Environment and Urban Systems",
- Member of scientific committees of several other journals including "GeoInformatica".

His main interests are spatial and visual information systems with applications to urban and environmental planning. He has written more than 200 scientific peer-reviewed papers, and advised 40 PhD’s.

He has co-authored 5 books, out of which the most important is "Fundamentals of Spatial Information Systems", co-authored with D. Thompson. His last book titled "Information Systems for Urban Planning, A Hypermedia Cooperative Approach" was issued in February 2001.

He also teaches at the University of Venice (IUAV), Italy.

Proceedings
International Conference on
Distributed Multimedia Systems
(DMS 2008)

Co-Editors

Masahito Hirakawa, *Shimane University, Japan*
Erland Jungert, *FOA, Sweden*

A Recommendation System for Segmented Video-Streaming Content utilizing Categorized Events and Stimuli

Billy Chi-Hong Cheung
Department of Computer Science
University of British Columbia
Vancouver, BC, Canada, V6T 1Z4
bccheung@cs.ubc.ca

Son T. Vuong
Son T. Vuong
Department of Computer Science
University of British Columbia
Vancouver, BC, Canada, V6T 1Z4
vuong@cs.ubc.ca

ABSTRACT

Despite the rise of video-on-demand systems and in particular, those that rely on peer-to-peer(p2p) technology, the role of the actual multimedia content being transferred remains remarkably stagnant. Though it is immediately obvious that the primary purpose of the multimedia is to transfer information from the provider to the consumer, this relationship is quickly becoming reciprocal. While traditional methods of ascertaining user preferences and subsequently providing recommendations have been limited to general content genre and user selections and user feedbacks, the emergence of more interactive and pervasive forms of media, such as aforementioned p2p VOD systems, have given rise to new avenues for keeping track of and subsequently meeting user needs. Based on the idea of contextual priming, we introduce a new framework, SUE, that essentially changes the role of media content from a static point of reference for advertisements to an active determinant in the decision making process by taking advantage of the intimate level of user profiling afforded by the Internet as well as the linear and segmented nature of p2p technology to determine a user's exact on-screen experience at any given time interval. Our design differs from other existing systems in two ways: (a) the level of granularity it can support, accommodating factors from both the user's on-screen and physical environment in making its recommendations and (b) in addressing some of the shortcomings seen in current applications, such as those imposed by coarse user profiling and faulty associations.

1. Introduction

Peer-to-peer (p2p) systems are traditionally considered as only a means of data sharing, often overlooked as also a medium in which end-users spend a lot of time interacting with. Currently, these systems are moving away from being just a passive tool to facilitate file sharing [1][2](i.e. using the system to download a file found on a tracker, and then using another program, such as a media player, to access the file) to a more interactive agent, where the entire process of seeking, retrieving and accessing file is available in the same application.[3][4] This is particularly true in the case of the increasingly prevalent market of online video (i.e. VoD) systems.

Such systems often maintain information about user activity and can offer an intimate and previously unattainable level of customization through which a user's choices and preferences can be used to further supply them with advertisements tailored specifically for their indicated needs/taste. This becomes even more significant when one compares the level of granularity possible with this type of targeting to that available in other forms of media of comparable scale. (i.e magazines and web page advertisements are relevant only to their immediate content, TV ads likewise are based on target audiences of the genre of the show being broadcasted and/or the channel being broadcasted on, and e-mail spam is done indiscriminately at best, and usually offers completely unnecessary products.) But we have yet to see anything that takes full advantage of this new medium beyond content-based recommendations relying on data from user selection of channels or movies or through the use of collaborative filtering and user profiles.

In this paper, we aim to address this oversight through the introduction of a new framework that allows for a much finer level of granularity with which to track and make use of users' activities to provide them with more suitable recommendations, using a new paradigm to evaluate the user state by revising the role of a media's content in gauging user preference. The crux of our framework's design relies on the conveniently sequential nature of video viewing(as opposed to, for example, text browsing, in which a user's focus can jump rapidly across any particular area of text), which allows us to determine the exact on-screen content a user is being exposed to at any given time. Coupled with user profile information that maintains (a) what the user has been experiencing in the immediate past and the stimuli these past events generate, (b) the user's distant past history and inclinations, and (c) the physical environment the user is in, intimately relevant recommendations can be made.

Instead of relying strictly on what the user thinks(such as use of user ratings) and what they state to be their preferences (such as the channels they subscribe to or the movies they select) we essentially focus on what users do and what/how they feel at any given moment to make recommendations. In short, we introduce the framework for a system utilizing empathy, or S.U.E. We base the theoretical viability of our framework on the idea of contextual priming.

2. Background and Design

Since our framework relies heavily on the concept of contextual priming, we begin with some basic background information concerning contextual information and their applications with regards to advertisements as well as how we apply this knowledge into the realm of videos for the purpose of our framework.

2.1 Contextual Priming

The priming principle[7] is a concept in advertising which states that different types of advertisements are perceived by consumers with varying intensity based on the context in which it is experienced. The principle asserts that the susceptibility of the consumer to a particular type of advertisement depends on the context that is serving as a primer. Context, as defined by [10] is the characteristics of the content of the medium in which the advertisement is inserted. The mood congruency-accessibility hypothesis[9][14] therefore suggests that advertisements that are relevant for or congruent with the mood of a subject at the moment may be more easily accessed and processed, with the implication that depending on the type of advertisement and the context, certain desires/needs can be made more salient.

We refer to the various factors that could act as a primer and thus affect user preference and desire as the stimuli that can influence user behaviour. Therefore, short term interests that arise in a user are constructed, at least in part, based on the various stimuli they have been exposed to in the immediate past while any long term preferences they have can be viewed as their perceived susceptibility to any given stimuli. Applying this notion of stimuli to a user viewing a video, we denote two different sources of stimuli affecting a user, their on-screen experience and their off-screen experience.

2.1.1 Events and Stimuli

As users expose themselves to video content, they are subjected to the influence of the stimuli within. We refer to the various time intervals within a video where one or more stimuli's influences have especially strong impact as events within the video. We can consider an event then as simply a collection of stimuli and their corresponding strengths. Since the user will only be experiencing a particular scene when it is being displayed by the system, we are able to effectively determine what stimuli they will be exposed to from their onscreen content at any given time.

Likewise, the physical environment in which a user is operating in also contributes to their immediate interests and desires. Environmental psychology states that an individual's environment(such as weather, temperature and other such meteorological events) can have a significant effect on the mood and inclination of an individual[18][19][20]. In [18], Tai and Fung explored the influences of environmental cues in stimulating a user's pleasure and thus willingness to make purchases. Furthermore, the experience and mood of the user has a direct correlation, as seen [21], in the willingness of the user to become receptive to advertisements and to carry through and even shop, indicating the importance of being able to both ascertain and keep track of a user's environmental data.

We can consider the various aspects of their physical surroundings, such as temperature, location, time of access, barometric changes and weather as the environmental factors which will influence user behaviour and preference through their effects on the perceived strength of any particular stimuli. We can then treat each environmental factor as a collection of stimuli and weights. These weights are then applied to any stimuli a user experiences to determine the effective strength of the stimulation.

2.2 System Design

In this section, we examine the design of our system. First, we will look into the way our media, in this case segmented videos, are handled. Then, we will examine the architecture of our framework.

2.2.1 Multimedia markup

As we indicated in the previous section, since our system only cares about the various events within a video, the additional space overhead needed to implement our system over an existing one is minimal. By representing each video(or more likely, their individual segments) as a list of <time, Event> pairs and treating this information as metadata that must be retrieved along with their corresponding video content, the appropriate stimuli can be applied. This information could easily be added into the metadata information that most existing video-on-demand systems already maintain on their content. (See Figure 1)

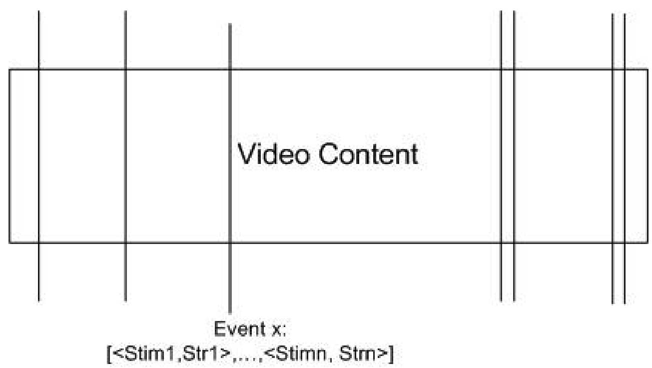


Figure 1. Video events, occurring at sporadic time intervals, as represented by their constituent stimulus/strength pairs

2.2.2 System Architecture

We divide our recommendation framework into three distinct components; Advertisement Decision, Advertisement Retrieval and Advertisement Presentation. As the intention of this paper is to demonstrate the viability of such a framework in exploiting the nature of the media in VoD systems, we will focus on the Advertisement Decision component and how it makes use of the aforementioned stimuli data, briefly touching on certain considerations for how advertisements can be retrieved and stored when the system is operating over various underlying networking structures. Further information on the overall framework and details about its corresponding modules can be found in [13].

The advertisement decision component refers to the aspect of the framework responsible for determining what type of advertisements would be most suitable/well-received by the user at any given time. Although the specific flow of events(such as Figure 2) is of course up to the specific underlying implementation , we propose that there are three main factors that influences the final recommendation being made to a user. Specifically, a recommendation is made through the amalgamation of the user's past/usual preferences, the immediate stimulus/stimuli they are being exposed to, the nature of the content they are being exposed to and the physical environment that the user is in. We refer to these as the user's General Interest, Immediate Interest, and Environmental Factors respectively.

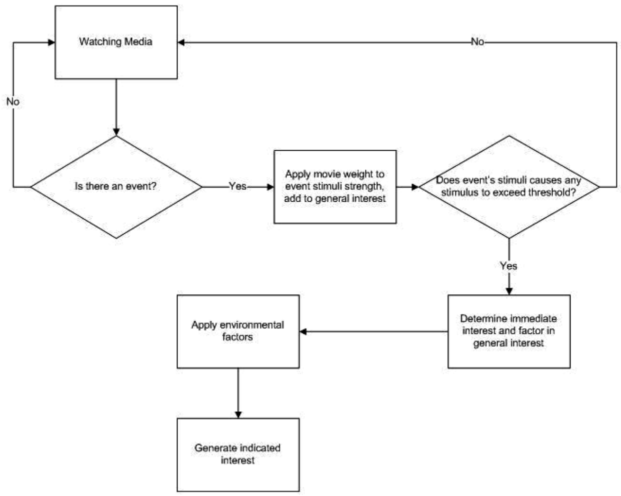


Figure 2. Sample flow of events

From Figure 2, we see that a user's preferences, which refers to the stimulus/stimuli that they are being aroused by, is continuously affected by their interaction with the media, in this case a video. As such, the immediate interest of the user represents what type of stimulus the user is most susceptible to at any given time during their use of the application. This interest is defined when one or more stimuli's value, which continuously increased or decreases based on user experience, exceeds a certain given threshold. This represents when the user has been influenced by their experiences to a degree sufficient to become susceptible to a specific type of information/advertisement[10]. The threshold can be anything, ranging from an arbitrary value to some product of the user's past history as well as the nature of the content they are interacting with, dependent on the underlying implementation.

Once threshold has been exceeded, the user's past preferences, which has already been taken into account when they were being exposed to the various stimuli within a content's events to correctly reflect the level of perceived impact each stimulus has on the user based in individual stimulus susceptibility, is factored in. Since each user has their own specific taste and thus are susceptible to different stimuli at varying degrees, the use of their past preferences, which is composed of information such as previously selected advertisements and favoured stimuli, allows us to gauge the likely effectiveness of each stimuli on the user and select accordingly. Another factor in determining stimuli to push to the user is the aforementioned environmental factors. The use of environmental factors serve as a catch-all filter allows us to not only remove irrelevant or potentially offensive recommendations, but also push advertisements that are more immediately relevant. Environmental information will let us be able to, for example, avoid pushing generic advertisements for products to individuals who are found in countries where such products are illegal or otherwise frowned upon.

After a stimulus or stimuli have been chosen by the system, it will, through the use of the underlying Advertisement Retrieval component, seek out advertisements or recommendations within the network that matches the desired stimulus/stimuli and the advertisement will be retrieved and subsequently displayed by the Advertisement Presentation component.

It is worth noting that there are certain considerations regarding system implementation that needs to be taken into account depending on the type of network our system will be running over. Here, we will briefly examine the changes we will need to make to system design for each of the more prevalent network models available (client/server, unstructured p2p and structured p2p(i.e. DHTs) systems) to accommodate our framework, discussing briefly their pros and cons.

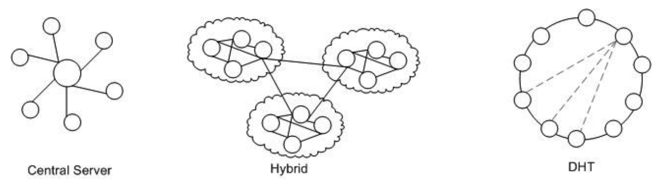


Figure 3. Types of networks

In a central server model, the storage and retrieval of advertisements is trivial. The server would contain a database of advertisements and searching for an appropriate advertisement would then be a simple matter of submitting standard SQL queries to the server to match the weight/stimuli combination desired. The use of a central server would allow for a very fine fidelity of matching to be done for the advertisements to be retrieved. After the appropriate advertisement is found, the server returns the file. Likewise, centralized p2p networks would operate in a similar manner except it would instead just provide the querying node with a list of nodes that possess an appropriate advertisement and its corresponding advertisement id. Of course, the obvious shortcoming of this type of network is the introduction of a bottleneck and single point of failure.

For unstructured p2p systems where superpeers/seedpeers are deployed, they can be used as a means of processing the queries while the advertisement objects will be spread across the system based on user activity and existing clusters/domains. Beyond that, the difference between the central server and hybrid model is minimal. However, one will need to have some way of ensuring that a search will be complete in terms of advertisement searching/availability since there is no central system to rely on. To ensure that the network has at least one copy of each valid advertisement at any given time, one can make use of superpeers as seedpeers as an option. By forming a backbone/overlay between the superpeers, clusters can communicate easily and the overlays be given the responsibility of handling the different categories.

Over structured p2p networks(i.e those that make use of DHTs), we introduce the problem of advertisement matching. Since we will need to perform range matching over a host of criteria to identify the appropriate advertisement to retrieve, we will need a way to distribute storage of advertisements with similar or even identical properties in different nodes to ensure system health. To overcome this apparent conflict to the keyword-based nature of DHT, solutions as the content-based pub/sub systems, such as that presented by [23], is an option.

3. Related Works

In this section, we will examine some existing works related to our research and discuss briefly the potential benefits of using SUE in comparison to existing methods of recommendation/ advertisement generation.

3.1 Psychologically Targeted Persuasive Advertising

In [17], Saari et al. explored the viability for a framework to "induce desired emotion and attention related status." [17], the idea being that what the users experience will affect their subsequent willingness or inclination towards other behaviours. Their work pushed for the idea of the use of context, placing advertisements at 'well recognized' onscreen locations and even inducing endearment of the user towards the system in question, but did not go into detail on how they plan on doing that. While they briefly touch on many of the ideas we also make use of, their goal is a much more general framework than ours, which targets videos since they allow for a much greater level of granularity. Furthermore, their proposed system actively aims to impose emotions on a user, while ours simply offers the ability to better gauge user state, making our system far less intrusive.

3.2 Collaborative and Content-Based Recommendation Systems

Recommendation systems provide users with suggestions on items that might be of interest to them. They typically employ two types of filtering methods in order to determine user preference: collaborative filtering[11] and content-based filtering[12]. Collaborative filtering offers recommendations to users based on the selections of other users with similar profiles/interests, and therefore assumed to have similar taste. Content filtering on the other hand provides recommendations based on the intrinsic properties of the item selected.

The most indicative use of collaborative filtering can be seen in commercial sites such as [5] and [6]. Such systems make use of information such as user history and user feedback ratings in order to build up user profiles. These profiles are then compared, offering a user selections made by other users. Problems commonly associated with this filtering technique include the need for rating information from other users [12] in order to provide recommendations, a problem with the accuracy of matches if there is insufficient overlap between profiles [11], requiring user feedback and ignoring the potentially useful properties of the content being [19].

Content-based recommendations, on the other hand, make extensive use of the attributes of the content in order to make suggestions[16]. Attributes and keywords associated with the content are given strength ratings, the recommendations made based on matching said ratings through text categorization[15]. However, current content-based recommendation systems can suffer from the problem of being unable to classify or define a user's profile or give accurate predictions on possible future user interests.

Many video-on-demand(VoD) systems, such as[14], also make use of a hybrid approach, using the movies or channels that the user subscribes to as well as user provided ratings in order to decide what kind of recommendations are suitable for the user while matching their selections against other profiles. Others, such as [8], assign categories to advertisements and entire videos in order to make matches. As well, they make use of data from explicit user interactions with past advertisements in order to determine which advertisements to push.

In this regard, the main difference between our framework and the others is that we don't particularly care about what the user states as their desired channel or preferences or which movie they select. Instead we base our advertisement suggestions on what the user actually watches. For example, if a user only watches a particular segment of a film repeatedly(say, a particularly good chase scene), then we only extract user preference information from that particular segment. This yields significantly different data than if we were to use traditional methods of profiling based on movie selection[5] or even channel registered[4], which is even more problematic since channel contents are not necessarily homogenous(i.e shows of Fox don't all share the same characteristics). Considering that similar content can be found on different and possibly philosophically opposite channels, the unreliability of using channel selection as a criterion for user profiling becomes even more evident.

4. Conclusion and Future works

While there have been a lot of work done in both the commercial sector and academia regarding recommendation systems and ways of gauging and predicting user interest, they focus on the user's selections and choices as static points of reference, which in the case of the increasingly dynamic nature of the content users are being exposed to, is inadequate. With the emergence of the internet and video-on-demand systems, it is now viable to know exactly what a user is experiencing on-screen at any given time as well as the kind of user watching it. With the ability to break video content down into their constituent events, and these events into the constituent stimuli that they excite in a user, there is no reason why this information should not be used to help determine what type of advertisements would appeal most to a user at any given time.

In this paper, we introduced SUE, a recommendation framework that utilizes the content of a media to help determine what sort of advertisements a user would be more agreeable to. By allowing the system to keep track of user interest via their activity and the type of stimuli they have been exposed to in both the recent and distant past, a profile can be built that determines a user's susceptibility to a particular stimulus. Coupled with the use of meteorological data on the user's general location, which is easily accessible, it allows us to determine and provide targeted advertisements that will be received much more saliently. In terms of its difference over existing systems, SUE's paradigm offers both a much finer level of granularity and, more importantly, the use of additional, previously unexplored sources of information from which a system can make more relevant recommendations.

Through the markup of media content and the strength with which they impose their corresponding stimuli's effects on a user, SUE can allow an advertiser to know, at any given time, in what ways have the user's viewing experience likely affected their shopping or personal preferences. By incorporating environmental data into the decision making process, far more relevant recommendations can be made to users based on information such as user location, time of access, local temperature, barometric changes, and weather. From a list of suggestions for local Italian restaurants with porches during a particularly sunny day after having detected that the user has been watching Iron Chef Mario Batali cooking for the last 5 hours to warnings of natural disasters in the local area, SUE provides the necessary infrastructure to make recommendations on a level of intimacy current systems simply cannot.

In terms of future work, several obstacles remain before SUE can be fully realized. The chief among these is the need for real data. In our simulations [13], all of our stimuli-related data were randomly generated. In order to truly verify the validity of SUE as a design, we need to acquire more realistic information on the effects of weather and environmental conditions on user preference. Along that line of inquiry, we also need to find a way to accurately mark up a video content's constituent event elements and determine both the stimuli they appeal to as well as the corresponding strength of each their effects. Only then can we perform any meaningful user study and experimentation on gauge the effectiveness and accuracy of our system. Beyond the need for data, another avenue of exploration is the format and presentation of recommendations once they have been determined and retrieved.

Obviously, the nature and strength of the retrieved advertisement, as well as the current on-screen content, will be factors in determining how the ad should be presented to the user (i.e. We do not want to interrupt an action sequence with a pop-up). This issue of timing is also constrained by the need to deliver the targeted advertisements before the indicated interest expires. A viable avenue is the addition of even more metadata to mark time intervals within segments where it is more acceptable from the user's perspective for 'commercial breaks' to occur. Not only would this go to alleviate user annoyance, but it could also be incorporated into a business model where the frequency and timing of advertisements is dependent on user hierarchy.

Another area of exploration that is readily available and more functionally applicable is in content searching and classification. Traditional systems offer searching and matching based on cast, genre, etc. However, by making use of the existing event stimuli lists that are already present in the meta- data for each segment of a movie, we can easily allow for users to search for specific content(i.e. a particular romance scene) or even emotional experiences through a database of movies.

5. References

[1] BitComet – www.bitcomet.com

[2] Gnutella – www.gnutella.com

[3] Hulu – www.hulu.com

[4] Joost – www.joost.com

[5] Netflix – www.netflix.com

[6] eBay – www.ebay.com

[7] P. M. Herr. Priming price: Prior knowledge and context effects. *The Journal of consumer research*, 16(1):67, 1989.

[8] Moonka; Rajas ; et al. Using viewing signals in targeted video advertising, March 2008. United States Patent Application 20,080,066,107.

[9] M. E. Goldberg and G. J. Gorn. Happy and sad tv programs: How they affect reactions to commercials. *The Journal of consumer research,* 14(3):387, 1987.

[10] Wim Janssens and Patrick de Pelsmacke. Advertising for new and existing brands: The impact of media context and type of advertisement. Journal *of Marketing Communications,* 11(2):113, 2005.

[11] Xin Jin, Yanzan Zhou, and Bamshad Mobasher. A maximum entropy web recommendation system: combining collaborative and content features. In *KDD '05: Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining*, pages 612-617, New York, NY, USA, 2005. ACM.

[12] Junzo Kamahara, Tomofumi Asakawa, Shinji Shimojo, and Hideo Miyahara. A community-based recommendation system to reveal unexpected interests. In *MMM '05: Proceedings of the 11th International Multimedia Modelling Conference*, pages 433-438, Washington, DC, USA, 2005. IEEE Computer Society.Sadas

[13] Billy Chi Hong Cheung. SUE : An Advertisement Recommendation Framework utilizing Categorized Events and Stimuli. Master's thesis, Department of Computer Science, The University of British Columbia, 2008.

[14] SD Perry, SA Jenzowsky, CM King, H. Yi, JB Hester, and J. Gartenschlaeger. Using humorous programs as a vehicle for humorous commercials. *Journal of Communication*, 47(1):20-39, 1997.

[15] P. Melville, R. Mooney, and R. Nagarajan. Content-boosted collaborative filtering for improved recommendations, 2002.

[16] Raymond J. Mooney and Loriene Roy. Content-based book recommending using learning for text categorization. In DL '00: *Proceedings of the fifth ACM conference on Digital libraries*, pages 195-204, New York, NY, USA, 2000. ACM.

[17] Timo Saari, Niklas Ravaja, Jari Laarni, Marko Turpeinen, and Kari Kallinen. Psychologically targeted persuasive advertising and product information in e-commerce. In *ICEC '04: Proceedings of the 6th international conference on Electronic commerce*, pages 245-254, New York, NY, USA, 2004. ACM.

[18] Susan H. C. Tai and Agnes M. C. Fung. Application of an environmental psychology model to in-store buying behaviour. *The International review of retail, distribution and consumer research*, 7(4):311, 1997.

[19] L. P. Chiu. Do weather, day of the week, and address affect the rate of attempted suicide in hong kong? Social psychiatry and psychiatric epidemiology, 23(4):229, 1988.

[20] E. A. Deisenhammer. Weather and suicide: the present state of knowledge on the association of meteorological factors with suicidal behaviour. Acta psychiatrica Scandinavica, 108(6):402, 2003.

[21] William Swinyard. The effects of mood, involvement, and quality of store experience on shopping intentions. The Journal of consumer research, 20(2):271, 1993.

[22] Kirk L. Wakefield and Julie Baker. Excitement at the mall: Determinants and effects on shopping response. *Journal of retailing*,74(4):515, 1998

[23] Wei Li. A Content-based Publish/Subscribe Framework over Structured Peer-to-Peer Networks. Master's thesis, Department of Computer Science, University of British Columbia, 2008.

A Comparison of Distributed Data Parallel Multimedia Computing over Conventional and Optical Wide-Area Networks*

Fangbin Liu
ISLA, Informatics Institute
University of Amsterdam, Kruislaan 403
1098 SJ Amsterdam, The Netherlands
fliu@science.uva.nl

Frank J. Seinstra
Department of Computer Science
Vrije Universiteit, De Boelelaan 1081A
1081 HV Amsterdam, The Netherlands
fjseins@cs.vu.nl

Abstract

The research area of Multimedia Content Analysis (MMCA) considers all aspects of the automated extraction of knowledge from multimedia data streams and archives. As individual compute clusters can not satisfy the increasing computational demands of emerging MMCA problems, distributed supercomputing on collections of compute clusters is rapidly becoming indispensable. A well-known manner of obtaining speedups in MMCA is to apply data parallel approaches, in which commonly used data structures (e.g. video frames) are being scattered among the available compute nodes. Such approaches work well for individual compute clusters, but - due to the inherently large wide-area communication overheads - these are generally not applied in distributed cluster systems. Given the increasing availability of low-latency, high-bandwidth optical wide-area networks, however, wide-area data parallel execution may now become a feasible acceleration approach. This paper discusses the wide-area data parallel execution of a realistic MMCA problem. It presents experimental results obtained on real distributed systems, and provides a feasibility analysis of the applied parallelization approach.

1 Introduction

Image and video data is rapidly gaining importance along with recent deployment of publicly accessible digital television archives, surveillance cameras in public locations, and automatic comparison of forensic video evidence. In a few years, analyzing the content of such data will be a problem of phenomenal proportions, as digital video may produce high data rates, and multimedia archives steadily run into Petabytes of storage space. Consequently, for urgent problems in multimedia content analysis, distributed supercomputing on large collections of clusters (Grids) is rapidly becoming indispensable.

It is well-known that data parallel approaches in image, video, and multimedia computing can provide efficient acceleration. In such approaches images and video frames are scattered among the available compute nodes, such that each node calculates over a partial structure only. Inter-node dependencies are then resolved by communication between the nodes. In the past 15 years numerous multimedia applications have been executed successfully in this manner, in particular using compute clusters [5], [8], [11].

For fine grain data parallel execution, inter-node communication is the major hurdle for obtaining efficient speedups. Because communication *within* a cluster is generally an order of magnitude faster than communication *between* clusters, data parallel execution *using multiple cluster systems* is generally believed to be inefficient. However, with the advent of low-latency, high-bandwidth optical networks, wide-area data parallel execution may now become a feasible acceleration approach. Investigation of this approach is specifically relevant for applications that work on very large data structures (e.g. hyperspectral images).

In this paper we apply wide-area data parallelism to a low-level problem in multimedia content analysis, i.e. curvilinear structure detection. For the development of the application we use a multimedia computing library (Parallel-Horus [11]) that allows programmers to implement data parallel applications as fully sequential programs. The application is tested on two distributed multi-cluster systems. One system is equipped with a dedicated wide-area optical interconnect, while the other uses conventional Internet links. We present experimental results, and provide a feasibility analysis of the applied parallelization approach.

This paper is organized as follows. Section 2 introduces the Parallel-Horus library. Section 3 describes our two experiemntal setups. Section 4 describes the low level image processing problem as applied in our experiments. Subsequently, Section 5 gives an evaluation of the performance and speedup results obtained on the two distributed platforms. Concluding remarks are given in Section 6.

2 Parallel-Horus

Parallel-Horus [11] is a cluster programming framework, implemented in C++ and MPI, that allows programmers to implement data parallel multimedia applications as fully sequential programs. The library's API is made identical to that of an existing sequential library: Horus [4]. Similar to other frameworks [7], Horus recognizes that a small set of *algorithmic patterns* can be identified that covers the bulk of all commonly applied functionality.

Parallel-Horus includes patterns for functionality such as unary and binary pixel operations, global reduction, neighborhood operation, generalized convolution, and geometric transformations. Current developments include patterns for operations on large datasets, as well as patterns on increasingly important derived data structures, such as feature vectors. For reasons of efficiency, all Parallel-Horus operations are capable of adapting to the performance characteristics of a parallel machine at hand, i.e. by being flexible in the partitioning of data structures. Moreover, it was realized that it is not sufficient to consider parallelization of library operations *in isolation*. Therefore, the library was extended with a run-time approach for communication minimization (called *lazy parallelization*) that automatically parallelizes a fully sequential program at runtime by inserting communication primitives and additional memory management operations whenever necessary [9].

Results for realistic multimedia applications have shown the feasibility of the Parallel-Horus approach, with data parallel performance (obtained on individual cluster systems) consistently being found to be optimal with respect to the abstraction level of message passing programs [11]. Notably, Parallel-Horus was applied in the 2004, 2005, and 2006 NIST TRECVID benchmark evaluations for content-based video retrieval, and played a crucial role in achieving top-ranking results in a field of strong international competitors [11], [12]. Moreover, recent extensions to Parallel-Horus, that allow for services-based multimedia Grid computing, have been applied successfully in large-scale distributed systems, involving hundreds of massively communicating compute resources covering our entire globe [11]. Clearly, Parallel-Horus is a system that serves well in bringing the benefits of high-performance computing to the multimedia community, but we are constantly working on further improvements, as exemplified in the following sections.

3 Experimentation Platforms

In the next sections we evaluate the obtained performance gains for distributed data parallel execution of a fine-grained Parallel-Horus program. To this end, the next section provides a description of the multimedia application, and the applied parallelization approaches. This section discusses the hardware platforms applied in the experiments.

3.1 DAS-2

The first platform is the (old) Distributed ASCI Supercomputer 2 (DAS-2), a 200-node system located at five different universities in The Netherlands: Vrije Universiteit Amsterdam (72 nodes), Leiden University, University of Amsterdam, Delft University of Technology, and University of Utrecht (32 nodes each). All nodes consist of two 1-Ghz Pentium-III CPUs with up to 2 GByte of RAM, and are connected by a Myrinet-2000 network. At the time of measurement, the nodes ran RedHat Enterprise Linux AS 3.2.3. The cluster at the University of Amsterdam has been dismantled in 2007, leaving a four-cluster system of 168 nodes.

3.2 DAS-3 and StarPlane

The second platform is the recently installed follow-up to DAS-2, i.e. the DAS-3, see http://www.cs.vu.nl/das3/. Similar to DAS-2, DAS-3 is a five-cluster system, with the following specifications: Vrije Universiteit Amsterdam (85x4 cores, 2.4 GHz), Leiden University (32x2 cores 2.6 GHz), University of Amsterdam (1 cluster of 40x4 cores, 2.2 Ghz; 1 cluster of 46x2 cores, 2.4 GHz), and Delft University of Technology (68x2 cores, 2.4 GHz). Besides the ubiquitous 1 and 10 GBit/s Ethernet networks, DAS-3 incorporates the recent high-speed Myri-10G interconnect technology from Myricom as an internal high-speed interconnect.

DAS-3 employs a novel 10Gbit/s wide-area interconnect based on light paths (StarPlane, see www.starplane.org). The rationale of the StarPlane optical interconnect is to allow part of the photonic network infrastructure of the Dutch SURFnet6 network to be manipulated by Grid applications to optimize performance. The novelty of StarPlane is that it does give flexibility directly to applications by allowing them to choose one of multiple logical network topologies in real time, ultimately with subsecond switching times.

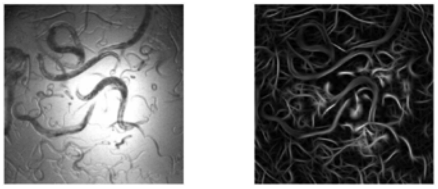
4 Line Detection

As discussed in [1], the important problem of detecting lines and linear structures in images is solved by considering the second order directional derivative in the gradient direction, for each possible line direction. This is achieved by applying anisotropic Gaussian filters, parameterized by orientation θ, smoothing scale σ_u in the line direction, and differentiation scale σ_v perpendicular to the line, given by

$$r''(x, y, \sigma_u, \sigma_v, \theta) = \sigma_u \sigma_v \left| f_{vv}^{\sigma_u, \sigma_v, \theta} \right| \frac{1}{b^{\sigma_u, \sigma_v, \theta}}, \qquad (1)$$

with b the line brightness. When the filter is correctly aligned with a line in the image, and σ_u, σ_v are optimally tuned to capture the line, filter response is maximal. Hence, the per pixel maximum line contrast over the filter parameters yields line detection:

$$R(x, y) = \arg \max_{\sigma_u, \sigma_v, \theta} r''(x, y, \sigma_u, \sigma_v, \theta). \qquad (2)$$



(a) (b)

Figure 1. Detection of *C. Elegans* **worms (courtesy of Janssen Pharmaceuticals, Belgium).**

Figure 1(a) gives a typical example of an image used as input to this algorithm. Results obtained for a reasonably large subspace of $(\sigma_u, \sigma_v, \theta)$ are shown in Figure 1(b).

4.1 Sequential Implementations

The line detection problem can be implemented sequentially in many ways. In the remainder of this section we consider two possible approaches. First, for each orientation θ it is possible to create a new filter based on σ_u and σ_v. In effect, this yields a rotation of the filters, while the orientation of the input image remains fixed. Hence, a sequential implementation based on this approach (which we refer to as ***Conv2D***) implies full 2-dimensional convolution for each filter. A second approach (called ***ConvRot***) is to keep the orientation of the filters fixed, and to rotate the input image instead. The filtering now proceeds in a two-stage separable Gaussian, applied along the x- and y-direction.

Pseudo code for the ***ConvRot*** algorithm is given in Listing 1. The program starts by rotating the input image for a given orientation θ. Then, for all (σ_u, σ_v) combinations the filtering is performed by xy-separable Gaussian filters. For each orientation step the maximum response is combined in a single contrast image structure. Finally, the temporary contrast image is rotated back to match the orientation of the input image, and the maximum response image is obtained.

For the ***Conv2D*** algorithm the pseudo code is given in Listing 2. Filtering is performed in the inner loop by a full two-dimensional convolution. On a state-of-the-art sequential machine either version of our program may take from a few minutes up to several hours to complete, depending on the size of the input image and the extent of the chosen parameter subspace. Consequently, for the directional filtering problem parallel execution is highly desired.

```
FOR all orientations θ DO
  RotatedIm = GeometricOp(OriginalIm, "rotate", θ);
  ContrastIm = UnPixOp(ContrastIm, "set", 0);
  FOR all smoothing scales σu DO
    FOR all differentiation scales σv DO
      FiltIm1 = GenConvOp(RotatedIm, "gaussXY", σu, σv, 2, 0);
      FiltIm2 = GenConvOp(RotatedIm, "gaussXY", σu, σv, 0, 0);
      DetectedIm = BinPixOp(FiltIm1, "absdiv", FiltIm2);
      DetectedIm = BinPixOp(DetectedIm, "mul", σu ∗ σv);
      ContrastIm = BinPixOp(ContrastIm, "max", DetectedIm);
    OD
  OD
  BackRotatedIm = GeometricOp(ContrastIm, "rotate", −θ);
  ResultIm = BinPixOp(ResultIm, "max", BackRotatedIm);
OD
```

Listing 1: Pseudo code for the *ConvRot* algorithm.

```
FOR all orientations θ DO
  FOR all smoothing scales σu DO
    FOR all differentiation scales σv DO
      FiltIm1 = GenConvOp(OriginalIm, "gauss2D", σu, σv , 2, 0);
      FiltIm2 = GenConvOp(OriginalIm, "gauss2D", σu, σv , 0, 0);
      ContrastIm = BinPixOp(FiltIm1, "absdiv", FiltIm2);
      ContrastIm = BinPixOp(ContrastIm, "mul", σu ∗ σv);
      ResultIm = BinPixOp(ResultIm, "max", ContrastIm);
```

Listing 2: Pseudo code for the *Conv2D* algorithms.

4.2 Parallel Execution

Automatic parallelization of the ***ConvRot*** program has resulted in an optimal schedule for this application [10]. The `OriginalIm` structure is broadcast to all nodes before each calculates its partial `RotatedIm` structure. The broadcast takes place only once, as `OriginalIm` is not updated in any operation. Subsequently, all operations in the innermost loop are executed locally on partial image data. The only need for communication is in the exchange of image borders in the Gaussian convolution operations.

The two final operations in the outermost loop are excuted in a data parallel manner as well. As this requires the distributed image `ContrastIm` to be available in full at each node [10], a gather-to-all operation is performed. Finally, a partial maximum response image `ResultIm` is calculated on each node, which requires a final gather operation to be executed just before termination of the program.

The schedule generated for the ***Conv2D*** program is more straightforward. First, the `OriginalIm` structure is scattered such that each node obtains an equal-sized partial image. Next, all operations are performed in parallel, with border exchange communication required in the convolution operations only. Finally, before termination of the program `ResultIm` is gathered to a single node.

4.2.1 Lazy vs. Naive Parallelization

The parallelization described above is based on the Parallel-Horus approach of 'lazy parallelization' that removes redundant communication steps at application run-time, with close to zero overhead. While this approach is essential for obtaining highest performance, we can decide not to apply lazy parallelization at all. If we do so, applications will communicate much more than needed, but will still produce the same result. Here, we refer to this latter approach as 'naive parallelization'. In the following, our 'naive' parallel runs are included only to enhance our understanding of the impact of communication on the obtained results.

5 Evaluation

To run our experiments on the DAS-2 system, we have compiled our programs with the MPICH-G2 library [3], which provides the correct communication setup for MPI programs consisting of multiple components that need to run on multiple sites. For highest performance, we have initialized the communication such that, apart from IP-based communication between cluster sites, we use the local Myrinet-2000 network for intra-cluster communication. To start our multi-cluster MPI-jobs on DAS-2, we have used the DRunner job submission tool, as part of the KOALA system [6], that allows for the simultaneous allocation of resources in multiple clusters.

As the KOALA system has not yet been put in place on the DAS-3 system, we have applied the OpenMPI library instead [2]. In our case, the relevant feature of OpenMPI is its support for heterogeneous networks. This is provided in a manner which is fully transparent to the application developers, whilst delivering very high performance. For our experiments we have initialized OpenMPI such that we use the optical StarPlane links between the clusters, and the local Myrinet-10G network for intra-cluster communication.

To avoid confusing results, we have refrained from using multiple cores per CPU, or multiple CPUs per node. Also, single cluster runs always have been performed at the Vrije Universiteit (VU), using a maximum of 64 nodes (on DAS-2 and DAS-3, respectively). Dual-cluster runs always have been performed using the VU and Leiden (LU) clusters simultaneously, with an equal number of nodes on both clusters. Because it was not possible to use more than 24 nodes on either LU cluster, dual-cluster runs have been performed using a maximum of 48 nodes in total. Four-cluster runs have been performed in a similar manner, resulting in runs using a maximum of 96 nodes in total.

5.1 Performance and Speedup

From the description of the two versions of our application it is clear that the ***ConvRot*** version is more difficult to parallelize efficiently. This is due to the data dependencies in the applied algorithm (i.e., the repeated image rotations), and not due to the capabilities of Parallel-Horus. Hence, the ***ConvRot*** program is expected to have speedup characteristics that are not as good as those of the ***Conv2D***program. However, ***Conv2D*** is expected to be the slower sequential implementation, due to the excessive accessing of image pixels in the 2-dimensional convolution operations.

Table 1, depicting the performance results for the two versions given a realistic orientation scale-space, shows that these expectations indeed are correct. On one node ***ConvRot*** is about 7 to 8 times faster than ***Conv2D*** on both DAS-2 and DAS-3. For 64 CPUs on a single cluster, and using our fast 'lazy parallelization' approach, this factor has dropped to 1.7 on DAS2, and 5.3 on DAS-3, respectively.

	DAS-2: # Clusters			DAS-3: # Clusters		
# CPUs	1	2	4	1	2	4
1	172.5	-	-	65.08	-	-
2	92.55	100.3	-	32.92	33.33	-
4	50.13	61.70	64.86	16.73	17.43	20.40
8	26.55	38.41	44.35	8.625	9.402	11.45
16	16.02	27.41	32.91	4.628	5.601	6.883
32	13.41	22.11	28.33	2.665	3.794	4.787
48	14.71	25.25	31.23	2.107	3.635	4.562
64	15.96	-	26.66	1.770	-	3.792
80	-	-	36.46	-	-	4.279
96	-	-	31.78	-	-	3.813

(a) ConvRot with Lazy Parallelization (results in seconds)

	DAS-2: # Clusters			DAS-3: # Clusters		
# CPUs	1	2	4	1	2	4
1	219.4	-	-	73.44	-	-
2	147.1	420.6	-	44.09	66.07	-
4	128.2	378.1	508.8	23.61	50.14	86.20
8	82.85	361.2	505.3	15.16	42.68	78.81
16	68.37	351.0	492.2	12.20	39.63	75.90
32	66.72	349.4	483.5	11.24	37.84	75.62
48	80.82	304.2	479.2	11.73	37.20	72.30
64	84.45	-	484.5	12.20	-	73.16
80	-	-	459.6	-	-	76.03
96	-	-	430.2	-	-	69.36

(b) ConvRot with Naive Parallelization (results in seconds)

	DAS-2: # Clusters			DAS-3: # Clusters		
# CPUs	1	2	4	1	2	4
1	1377.2	-	-	566.6	-	-
2	695.9	703.7	-	283.7	283.9	-
4	349.7	359.6	361.2	142.3	142.4	156.2
8	176.1	185.9	187.3	71.28	71.31	78.51
16	89.95	95.65	96.11	35.90	35.91	39.75
32	46.98	50.52	50.95	18.11	18.20	20.45
48	34.97	37.62	37.67	12.66	12.73	14.42
64	27.66	-	29.76	9.297	-	10.82
80	-	-	28.91	-	-	8.402
96	-	-	26.28	-	-	7.392

(c) Conv2D with Lazy Parallelization (results in seconds)

	DAS-2: # Clusters			DAS-3: # Clusters		
# CPUs	1	2	4	1	2	4
1	1433.4	-	-	572.9	-	-
2	748.1	954.8	-	290.2	366.0	-
4	409.3	616.4	713.2	150.9	169.9	209.7
8	241.1	435.6	524.7	76.74	97.71	129.9
16	152.7	344.5	460.2	41.59	44.07	89.67
32	111.4	299.3	446.7	27.00	43.94	70.29
48	92.84	243.3	347.0	20.10	36.93	64.34
64	87.86	-	369.0	17.19	-	61.66
80	-	-	349.4	-	-	60.71
96	-	-	372.6	-	-	56.04

(d) Conv2D with Naive Parallelization (results in seconds)

Table 1. Performance of ConvRot (a and b) and Conv2D (c and d) with 'lazy' and 'naive' parallelization. Results for computing an orientation scale-space at $5°$ angular resolution (i.e., 36 orientations) and 8 (σ_u, σ_v) combinations. Scales computed are $\sigma_u \in \{3, 5, 7\}$ and $\sigma_v \in \{1, 2, 3\}$, ignoring the isotropic case $\sigma_{u,v} = \{3, 3\}$. Image: 512×512 (4-byte) pixels.

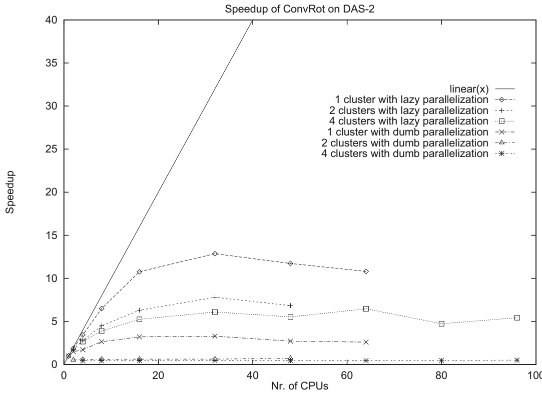


Figure 2. Speedup of ConvRot on DAS-2.

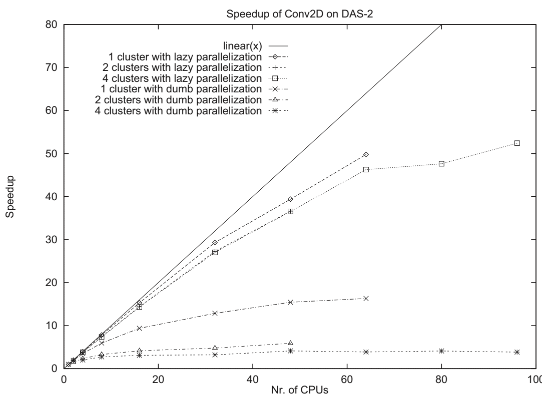


Figure 3. Speedup of Conv2D on DAS-2.

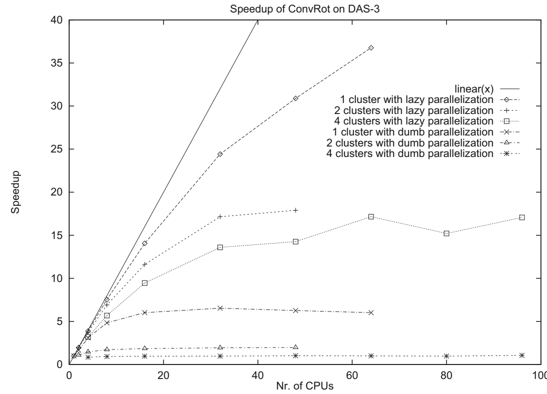


Figure 4. Speedup of ConvRot on DAS-3.

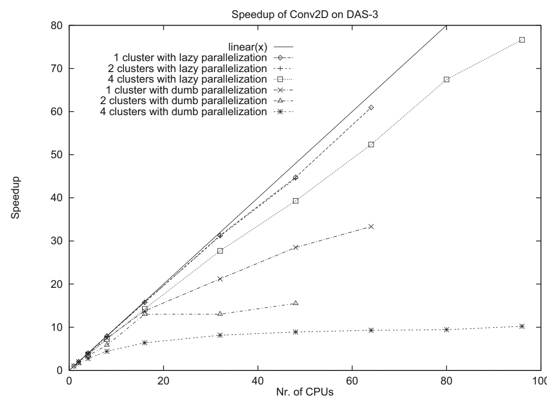


Figure 5. Speedup of Conv2D on DAS-3.

When comparing the multi-cluster runs in Table 1(a), it is clear that ***ConvRot*** indeed can benefit from an increased availability of nodes — in particular on DAS-3 — despite the fact that 64-node single-cluster runs always provide fastest execution. ***Conv2D*** shows the best multi-cluster results, with four-cluster results even outperforming the 64-node single-cluster runs.

From this performance comparison we conclude that, for all versions of our line detection problem, one can obtain increased performance from using nodes at multiple sites, in particular on DAS-3. As expected, a performance drop is introduced when moving from one cluster to multiple clusters, but such a move is still worthwhile in case the available number of nodes at one cluster is limited, or if the potential for parallelization of a problem at hand is (much) larger than can be exploited by a single cluster. In particular, for applications with good parallelization characteristics, such as our ***Conv2D*** program, the benefits can be significant.

Before we present our speedup analysis, we need to indicate that there is an apparent discrepancy between the 'lazy' and 'naive' single-node, single-cluster results. As these runs essentially represent sequential execution in both cases, pairwise lazy/naive parallel runs on a single node of the same cluster should give identical performance results. This is not the case, however, as the presented results are obtained from parallel runs performed on a single node. In the lazy parallel case, the single-node parallelization overhead is such, that a factual sequential run can be said to be identical to a a single-node lazy parallel run. In contrast, single-node naive parallel runs do show a significant performance drop, because of excessive creation, copying, and removal of internal parallelization structures.

When considering the speedup graphs of Figures 2-5, it is immediately apparent that DAS-3 provides much better speedup characteristics than DAS-2 — for single-cluster and multi-cluster runs alike. This is particularly remarkable given the fact that the sequential runs on DAS-3 are about 3 times faster than on DAS-2, thus reducing the theoretical potential for obtaining good speedups on DAS-3 relative to DAS-2. Our results are explained by the fact that the communication speeds relative to the computation speed have improved significantly on the new DAS-3 sys-

Program/Cluster	VU	LU	UvA	MN
ConvRot	65.08	60.66	70.95	65.65
Conv2D	566.6	523.4	617.7	566.7

Table 2. Single-cluster, single-node performance on each of the applied DAS-3 clusters.

tem. Importantly, this result indicates that our optical StarPlane network improves communication speeds over conventional Internet links by a similar (or even higher) relative amount as our local Myrinet-10G links improve communication speeds over the old Myrinet-2000 local interconnect.

One further significant result is found in Figure 5: the two-cluster lazy parallel runs provide speedup results that are *identical* to the single-cluster lazy parallel runs. Clearly, for applications with a good parallelization potential, multi-cluster runs constitute a realistic, and attractive alternative. In this respect one may wonder why the related four-cluster lazy parallel speedup results are somewhat lagging behind. This is explained by the fact that the four-cluster runs also incorporates the DAS-3 cluster at the University of Amsterdam, which has significantly slower compute nodes. As a result, in our current parallelization strategy (which results in *bulk synchronous* parallel execution) the faster compute nodes repeatedly have to wait for the slower UvA nodes. Table 2 shows the significance of the difference (i.e., even up to 18%) in the computation speeds among the different DAS-3 clusters applied in our experiments. Note that this effect, which is certainly not caused by any form of increased wide-area communication overhead, plays a role in all four-cluster results obtained on DAS-3.

Given the above results, we have clearly shown that — in contrast to common belief — fine-grained distributed data parallelism indeed can be an viable acceleration approach. Also, we conclude that DAS-3, with its optical interconnect, is a magnificent system for obtaining significant speedups, for single-cluster runs and multi-cluster runs alike.

6 Conclusions

In this paper we have applied a wide-area data parallelization approach to two different implementations of a well-known problem in multimedia content analysis. For the development of the applications we have used the Parallel-Horus multimedia computing library that allows programmers to implement data parallel applications as fully sequential programs. The applications have been tested on two different distributed systems, both consisting of multiple compute clusters located at different universities in The Netherlands. The DAS-3 system is equipped with a dedicated wide-area optical interconnect, called StarPlane, while DAS-2 uses conventional Internet links. We have presented experimental results obtained on both systems, and have provided a feasibility analysis of the applied parallelization approaches.

The most significant conclusion is that, while the compute nodes of DAS-3 are about 3 times faster than those of DAS-2, DAS-3 was still capable of providing much better speedup characteristics — even for larger numbers of CPUs, and for single-cluster and multi-cluster runs alike. This result is explained by the fact that the innovative local and wide-area networking technologies applied in DAS-3 (Myrinet-10G and StarPlane) provide speed improvements that are similar to, or even outweigh, the speed improvements obtained by the compute elements of DAS-3. In particular, given our results for the ***Conv2D*** version of our line detection program, we conclude that, for applications that have a good parallelization potential, multi-cluster runs constitute a realistic, and attractive approach. This result indicates that, in combination with the increasing adoption of optical interconnects in wide-area computing, there is a growing potential for scalable techniques for multimedia data analysis.

References

[1] J. Geusebroek et al. A Minimum Cost Approach for Segmenting Networks of Lines. *Int. J. Comp. Vis.*, 43(2):99–111, July 2001.

[2] R. Graham et al. Open MPI: A High-Performance, Heterogeneous MPI. In *Proceedings of HeteroPar'06*, Barcelona, Spain, Sep. 2006.

[3] N. Karonis et al. MPICH-G2: A Grid-enabled Implementation of the Message Passing Interface. *J. Par. Dist. Comp.*, 63(5):551–563, May 2003.

[4] D. Koelma et al. Horus C++ Reference. Technical report, Univ. Amsterdam, The Netherlands, Jan. 2002.

[5] J. Lebak et al. Parallel VSIPL++: An Open Standard Software Library for High-Performance Signal Processing. *Proc. IEEE*, 93(2):313–330, Feb. 2005.

[6] H. Mohamed and D. Epema. Experiences with the KOALA Co-Allocating Scheduler in Multiclusters. In *CCGrid2005*, Cardiff, UK, May 2005.

[7] P. Morrow et al. Efficient Implementation of a Portable Parallel Programming Model for Image Pro-cessing. *Concur.: Pract. Exp.*, 11:671–685, Sep. 1999.

[8] A. Plaza et al. Commodity Cluster-based Parallel Processing of Hyperspectral Imagery. *J. Par. Dist. Comp.*, 66(3):345–358, Mar. 2006.

[9] F. Seinstra et al. Finite State Machine-Based Optimization of Data Parallel Regular Domain Problems Applied in Low-Level Image Processing. *IEEE Trans. Par. Dist. Syst.*, 15(10):865–877, Oct 2004.

[10] F. Seinstra et al. User Transparency: A Fully Sequential Programming Model for Efficient Data Parallel Image Processing. *Concur. Comp.: Pract. Exp.*, 16(6):611–644, May 2004.

[11] F. Seinstra et al. High-Performance Distributed Video Content Analysis with Parallel-Horus. *IEEE Multimedia*, 14(4):64–75, Oct.-Dec. 2007.

[12] C. Snoek et al. The Semantic Pathfinder: Using an Authoring Metaphor for Generic Multimedia Indexing. *IEEE Trans. Pat. Anal. Mach. Intel.*, 28(10):1678–89, Oct. 2006.

A Lightweight Augmented Reality System

Yugo Hamakawa*[1], Takuya Kawase+[2], and Masahito Hirakawa+
*Interdisciplinary Graduate School of Science and Engineering
Shimane University, Japan
+Interdisciplinary Faculty of Science and Engineering
Shimane University, Japan
hirakawa@cis.shimane-u.ac.jp

Abstract

Sophisticated multimodal interfaces are the key to success in future computer development and many ideas have been proposed so far. Augmented reality (AR) is such an idea. Despite of its conceptual novelty, augmented reality systems are still at the laboratory level.

This paper presents a lightweight AR system toward its wide use in a practical living environment. In order to enable the system run on conventional tools and equipment, the system discounts a registration task by achieving 2D tracking of real world objects instead of recognizing their 3D structures. Here we propose integrating RFID (Radio Frequency Identification) and vision-based methods for the registration. An interactive and adaptive user-interface of the system is also presented.

1. Introduction

One view of the future of computing that has captured our interest is augmented reality (AR), in which everyday objects gain electronic properties without losing their familiar physical properties [1]. Though almost twenty years have passed since the idea was first published, AR systems are still at the research level.

The reasons would be twofold. Firstly, traditional AR systems often require highly specialized tools and equipment such as wearable head mounted displays (HMD), digital gloves, and magnetic/ultrasonic sensors. They are rather expensive. More seriously, wearing such devices is bothersome for the user and impediment to his/her work.

The other reason lies in difficulties of world modeling and registration [2], [3]. Applications such as surgery, military, and architecture demand accurate registration based on a detailed 3D model of the real world so that computer-generated objects can be superimposed over real world objects without any error and they are inseparable at all by our vision. But, there are some other applications where such accurate registration is not mandatory, which include shopping, navigation, and personal information management.

Consider, for example, that you are in a shop and just looking around while waiting for your friend. The computer courteously informs you of the existence of the same (type of) item which you checked on a website few days ago. In another situation, the computer gives you a notice that your parent's birthday is approaching when you scan the room with the system and turn to the photo frame in which his/her photo is. It is noted that you don't necessarily have a specific goal (i.e., target objects) in mind and the system advises you if there is any information which is worth telling you. Object-based reference through such unconscious interface would be promising.

From a technical viewpoint, one of the keys to creating breakthroughs for AR research is the idea of lightweightness. Development of lightweight AR systems by conventional tools and equipment is getting active in recent years [4]-[6]. While the systems have limited capabilities, they make use of technologies and devices that people are likely to be already using for other purposes. This can help AR systems spread out in our daily life environment.

This paper presents a lightweight AR system in which a registration task is discounted by focusing attention to 2D positions of real world objects.

A combination of RFID (Radio Frequency Identification) and vision-based object recognition methods is applied to registration processing. While visual illusory capabilities of the system are limited, target real world objects over which computer-generated objects are superimposed are not limited to a small number of objects as assumed in existing AR systems, but cover almost all objects in our surroundings in the expectation that RFID tags will be attached commonly to the objects.

Meanwhile, presenting computer-generated annotations to all the real world objects the system notices may not be

[1] Y. Hamakawa is now at Seiko Epson Corp., Japan.
[2] T. Kawase is now at Japan Advanced Institute of Science and Technology, Japan.

adequate since the display can be filled with them. It is thus needed to control the amount of information to be presented to the user. The proposing system is capable of showing the information adaptively depending on, for example, user's focus and object properties.

The rest of the paper is organized as follows. After discussing related work in Section 2, we will explain in Section 3 our solution to the lightweight AR system development. Section 4 gives details of the implementation for recognizing and tracking real world objects. Interface issues explaining how actually the user interacts with the system are presented in Section 5. Finally, in Section 6, we conclude the paper.

2. Related Work

Registration is a technique to properly align computer-generated objects with real world objects as the user moves his/her viewpoint, and is one of the key issues in AR systems. Many registration techniques have been proposed so far, which include sensor-based registration and vision-based registration.

In sensor-based registration, some sensors such as magnetic, mechanical, ultrasonic, and optic ones are used to measure the positions of real world objects [7], [8]. While those are advantageous in realtimeness, extra expenses are necessary for bulky equipments. Furthermore, user's movable area is limited.

Vision-based registration, on the other hand, doesn't require such costly hardware sensors. The 3D positions of real world objects and their motion are estimated through analyzing a sequence of images taken from a camera. Here there are two types of approaches: marker-based [5], [9], [10] and natural feature-based [11]. The vision-based approaches work nicely, but some constraints/assumptions are imposed to lighten the processing cost and then make the methods feasible, which include attachment of specially designed visual markers, specification of control features/points, and necessity of a processing stage for world modeling in advance.

Meanwhile, development of lightweight AR systems by means of conventional tools and equipment has been active. To make the system run in reasonable performance, ideas of discounting registration tasks are adopted.

3D M-Ads system [5] adopts the emerging smart phones technology as an effective booster of the utility of AR. For registration it uses colored tags which are specially designed for low end products, i.e., smart phones.

[4] presents a discount registration method which is intended for outdoor applications. By knowing the fact that target objects are far away from the user, the system tracks user's position and movement simply by a GPS and a 3DOF angular tracking sensor.

Another discounted AR solution appears in [6]. The user manually places a virtual object and changes its scale and orientation to visually match it with the scene. After the placement, simple 2D tracking is applied to keep the object glued to the scenery. The system runs on such low-end camera phones as in [5].

3. Overview of the System

3.1 System design

We propose developing a lightweight AR system which adopts a discount registration facility. While the system has limited capabilities in registration, it makes use of devices that people are likely to be already using for other purposes. This can help the system be spread out widely in our daily life environment. Basic ideas of the system are explained as follows (see Fig. 1).

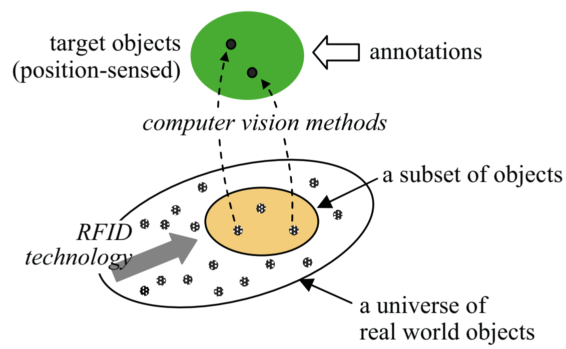


Figure 1 Conceptual idea of the system

Firstly, the system catches what objects exist in surroundings (hopefully, in front) of the user by means of RFID technology under the expectation that it is likely to have tags attached to almost all the objects in the near future. In theory, it is not impossible to identify their positions. Studies of localization of RFID tagged objects have been carried out [12], [13]. But the preciseness is not sufficient enough for AR applications.

This "objects are around here" information is used to remove unnecessary processing for vision-based object recognition which follows the tag-based object recognition, and helps to improve vision-based recognition performance. A similar approach appears in [14] but is designed for a mobile robot. It is noted here that the tag-sensed objects are just candidates and may not appear in the area that the user sees actually.

Finally, annotations are superimposed over the real world objects in video frames to help the user get better ideas about them. The font size and content of the annotations change depending on the user's focus and interest to save the screen space.

3.2 System architecture

A prototype system comprises five hardware components as depicted in Fig.2. They are RFID tags, a tag reader, a video camera, a mobile computer, and a database server.

The user can see real world objects through a display of the mobile computer (e.g., laptop PC, PDA, and cell phone). At the same time, the tag reader reads tags which are attached to the objects. Reference to the database server enables the system to get the information which is associated with the captured tag IDs, including annotations to be superimposed and image properties to be used for performing vision-based object recognition. The vision system determines the fine position of the possible real world objects. Context-sensitive annotations are then presented to the user. Details of the object tracking and user-interface will be explained in the subsequent sections.

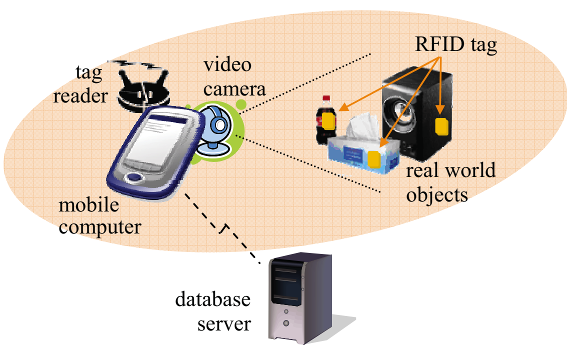


Figure 2 An organization of the system

Figure 3 shows an actual setup of the system. Here there are two approaches in setting up tags and the receiver. In one approach, the tag reader is attached to the mobile computer. This seems a reasonable approach though the size of the tag receiver is a factor to be considered. On the other hand, the receiver can be placed on the side of real world objects. In this case, a tag should be attached to the mobile computer as well to have the system know that the user approaches to the tag reader (i.e., objects in the sensing area). It is also needed to prepare one receiver at each object area.

Figure 3 System in use

4. Recognition of Real World Objects

4.1 Tag-based object localization

RFID tags that we use in our experiment are of an 300MHz active type by NTT Advanced Technologies, as shown in Fig. 4. One of the reasons why active tags are chosen is their operating range. Though passive tags are cheaper and smaller than those active ones, a receiver must be closer to tags in order to catch their signals. In practice, our active tags can have operational ranges of around 50 meters. Here the signal is attenuated so that tags are sensible just within a few meters, considering that the system is applied to indoor applications at, for example, shops, libraries, and home.

Figure 4 RFID tags and a receiver

Computer-generated annotations are given to the user when the user turns the mobile computer towards the target object(s) he/she is interested in. Here it is desirable that the receiver catches only the signals coming from the front. But this task is rather difficult due to the complexity of wireless signal wave characteristics. Furthermore, the operating range varies depending on the orientation of a tag as well as its position (i.e., distance between the tag and the receiver). That's why we use RFID technology as a means of detecting what objects exist in user's surrounding and vision-based techniques are provided to recognize where exactly the objects are located. Details of the vision-based techniques are explained in the next subsection.

4.2 Vision-based object localization

After getting IDs of the real world objects in user's surrounding, the system performs object recognition to find where exactly they are located. Two object recognition methods are adopted: One is a feature-based method and the other a color histogram-based method.

As the feature-based method, we use the Haar-like object detector provided in the OpenCV library, whose idea was presented initially in [15], [16]. Though it requires a huge amount of time at a training stage, object detection works in real-time.

In our experiments, classifiers are prepared for three objects: a duck toy, a box of facial tissues, and a cup-type instant noodle. 24x24-sized 7,000 positive sample images and 4,000 negatives are used for their training. The size of the resultant classifier xml files ranges from 11KB to 193KB.

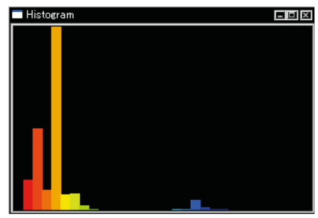
Figure 5 shows results of the object detection. The Haar-like technique is robust to intensity changes and object hiding as seen in the figure.

Figure 5 Results of applying Haar-like object classifiers

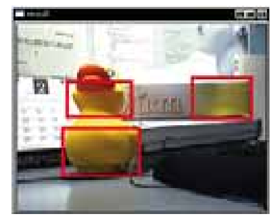
Meanwhile, since the Haar-like technique concerns feature patterns and doesn't care about colors, we also apply CAMSHIFT [17] which is a fast, simple color-based tracking algorithm provided in OpenCV.

To run the CAMSHIFT algorithm, a search window must be assigned in advance of the processing. The system first accesses the database server and obtains a color histogram of each of the objects which have been detected by means of RFID tags. The region(s) whose color is similar to the dominant one of the target object is determined as the search window.

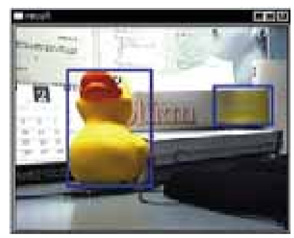
Suppose, for example, the system detects a tag for the duck toy. Figure 6(a) is the histogram of the duck toy. Knowing that yellow is the dominant color, three regions having the yellow color are extracted as in Fig. 6(b). The CAMSHIFT algorithm is applied to each of the regions to get more reliable object regions. Figure 6(c) shows the result. Here it may happen that some regions become identical. Actually, two possible regions for a duck toy are merged into one. Finally, the system compares the color histogram of every tuned region with that of the searching object, i.e., duck toy, and the one having the highest score is chosen as the target.



(a) color histogram of duck toy



(b) possible regions



(c) tuning of object regions

Figure 6 CamShift-based object detection

Once the object region is determined, the CAMSHIFT algorithm is applied successively to track the object. Meanwhile, if the system senses multiple tags for the same object type, a same number of regions having higher similarities are chosen as targets.

Meanwhile, we adopt two object tracking methods, Haar-like and CAMSHIFT. Here it is noted that Haar-like method is applied prior to CAMSHIFT method. A scheme of combining those two is explained as follows.

While the Haar-like method is robust and reliable, it may fail to catch target objects correctly. To compensate for lack of color properties, the system analyzes a color histogram of the object region which is extracted by the Haar-like method and evaluates how the accuracy of object estimation is. If the score of color histogram matching is below a threshold, the output by the Haar-like method is discarded. If no output is obtained by the Haar-like method, the system counts on the CAMSHIFT method.

One more issue to be considered is how Haar-like feature classifier files and color histograms of real world objects are provided. We expect that manufacturers will prepare those feature data when they throw the products into the market, as they prepare their bar codes.

5. Interaction with the System

Computer-generated annotations are drawn over real world objects in real time. Since it is expected RFID tags will be attached to almost all objects in our surroundings, the display can be occupied by those data. Facilities to avoid the display space overflow are essentially required. In general, this is known as information visualization research.

Many trials toward sophisticated visualization of information have been conducted so far. One key idea is the focus+context principle which allows users to see one or more selected focus regions in full detail, while information in the periphery is compressed to occupy less screen space [18]. In our system this focus+context principle is applied to the presentation of computer-generated annotations.

Figure 7 shows a snapshot of the display screen. Object regions are indicated by rectangular frames, and textual annotations are allocated beside them as shown in the figure. As is seen, the font size of those annotations changes depending on their position on the display. A bigger font is adopted as the associated object comes closer to the center of the display.

Short annotations may not be enough in some cases. If the user places a cursor over a certain real world object (or touches the object with a pen in the case of PDA/smartphone), a longer scrolling message is presented at the bottom of the display, as is also shown in Fig. 7.

Figure 7 Interactive visualization based on focus+context principle

In addition, the system provides a facility of selectively choosing objects to which computer-generated annotations are attached in order to save more display space [19]. We apply the idea of collaborative filtering which relies on opinions of possible partners having a similar preference to help the user receive the information most likely to be interesting to him/her or relevant to his/her needs. Success in Amazon.com has shown the effectiveness of the collaborative filtering technique. Recommended objects will appear with an appealing color frame. Conversely, annotations may not be provided for non-recommended objects.

6. Conclusion

This paper presented a lightweight AR system having the discount registration approach, which enables the user to interact with a variety of objects in our living environment. We proposed integrating RFID and vision-based object recognition techniques for implementation of the discount registration. Furthermore, the system provides an interactive interface for allowing the user to have computer-generated annotations adaptively depending on the user's focus and object properties.

Much work still remains. One concerns identification of objects. Active tags are large and expensive to be used widely in our environment. Use of passive tags is a challenge to explore more practical solutions. Improving the performance of the vision-based object recognition algorithms is another research issue. In addition, evaluation of recognition performance and usefulness of the system in an actual application should be investigated through experiments.

References

[1] P. Wellner, W. Mackay, and R. Gold, "Computer-Augmented Environments: Back to the Real World," CACM, Vol.36, No.7, pp.24-26, July 1993.

[2] D. Schmalstieg, G. Schall, D. Wagner, I. Barakonyi, G. Reitmayr, J. Newman, and F. Ledermann, "Managing Complex Augmented Reality Models," IEEE Computer Graphics and Applications, Vol.27, No.4, pp.48-57, July/Aug. 2007.

[3] R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, B. MacIntyre, "Recent Advances in Augmented Reality," IEEE Computer Graphics and Applications, Vol.21, No.6., pp.34-47, Nov./Dec. 2001.

[4] B. Wilkinson and P. Calder, "Augmented Reality for the Real World," Proc., IEEE Int'l Conf. on Computer Graphics, Imaging and Visualization, pp.452-457, 2006.

[5] F. Chehimi, P. Coulton, and R. Edwards, "Augmented Reality 3D Interactive Advertisements on Smartphones," Proc., IEEE Int'l Conf. on Management of Mobile Business, 2007.

[6] P. Honkamaa, J. Jappinen, and C. Woodward, "A Lightweight Approach for Augmented Reality on Camera Phones using 2D Images to Simulate 3D," Proc., ACM Int'l Conf. on Mobile and Ubiquitous Multimedia, pp.155-159, 2007.

[7] B. Jiang, U. Neumann, and S. You, “A Robust Hybrid Tracking System for Outdoor Augmented Reality,” Proc., IEEE Virtual Reality Conf., pp.3-10, 2004.

[8] M. Kourogi, N. Sakata, T. Okuma, and T. Kurata, “Indoor/Outdoor Pedestrian Navigation with an Embedded GPS/RFID/Self-contained Sensor System,” Proc., Int’l Conf. on Artificial Reality and Telexistence, pp.1310-1321, 2006.

[9] K. Tateno, I. Kitahara, and Y. Ohta, “A Nested Marker for Augmented Reality,” Proc., IEEE Virtual Reality Conf., pp.259-262, 2007.

[10] F. Farbiz, A.D. Cheok, W. Liu, Z.Y. Zhou, K. Xu, S. Prince, M. Billinghurst, and H. Kato, “Live Three-dimensional Content for Augmented Reality,” IEEE Transactions on Multimedia, Vol. 7, Issue 3, pp. 514 – 523, June 2005.

[11] M. L. Yuan, S.K. Ong, and A.Y.C. Nee, “Registration using Natural Features for Augmented Reality Systems,” IEEE Trans. on Visualization and Computer Graphics, Vol.12, No.4, pp.569-580, July/Aug. 2006.

[12] L.M. Ni, Y. Liu, Y. C. Lau, and A. P. Patil, “LANDMARC: Indoor Location Sensing Using Active RFID,” Wireless Networks, Vol.10, Issue 6, pp.701-710, 2004.

[13] A. Bekkali, H. Sanson, and M. Matsumoto, “RFID Indoor Positioning based on Probabilistic RFID Map and Kalman Filtering,” Proc., IEEE Conf. on Wireless and Mobile Computing, Networking and Communications, 2007.

[14] H. Chae and S. Na, “Mobile Robot Navigation using RFID and Vision Sensor,” Proc., Symp. on Information Technology Convergence, pp.52-57, 2007.

[15] P. Viola and J. Michael, “Rapid Object Detection using Boosted Cascade of Simple Features,” Proc. IEEE Conf. on Computer Vision and Pattern Recognition, pp.511-518, 2001.

[16] R. Lienhart and M. Jochen, “An Extended Set of Haar-like Features for Rapid Object Detection,” Proc., IEEE Int’l Conf. on Image Processing, pp.900-903, 2002.

[17] G. R. Bradski, “Real Time Face and Object Tracking as a Component of a Perceptual User Interface,” Proc., IEEE Workshop on Applications of Computer Vision, pp.214-219, 1998.

[18] C. Chen, “Information Visualization: Beyond the Horizon (2nd edition),” Springer-Verlag, London, 2006.

[19] M. Hirakawa, “Going Beyond Completeness in Information Retrieval,” in Databases in Networked Information Systems, Lecture Notes in Computer Science, Springer Press, pp.70-80, 2007.

Measuring Reliability Component in Business Process Automation

Cihan Varol[1] and Coskun Bayrak[2]
[1]Applied Science Department---- [2]Computer Science Department
University of Arkansas at Little Rock
2801 S. University Ave.
Little Rock, Arkansas, 72212
{Email: cxvarol | cxbayrak@ualr.edu}

***Abstract*:** *Quality of Service (QoS) of the Workflow Process Automation is limited in the workflow generation part of the systems. However, impact of the spelling errors in the customer expectation file or unrecognized input records heavily affects the outcome of the workflow process automation. In this study, a custom spelling correction method is implemented using phonetic strategies, pattern matching techniques and statistical results in order to fix the input data. A natural language parser is adapted to understand the customer's expectation from the task. The definition of a reliability parameter and its affect to workflow are discussed for both the goal state and input records supplied by the customer.*

***Index Terms*—Business Process Automation, Reliability, QoS, Workflow**

Introduction

Like so many other information technology improvements, process automation in office environment achieved some of its vital promises, such as the elimination of paper use which has proliferated. However, office tasks such as faxing and word processing have been automated from the concept of workflow, which has evolved from the notion of process in manufacturing and the office [1]. A workflow may describe business process tasks at a conceptual level necessary for understanding, evaluating, and redesigning the business process. On the other hand, workflows may capture information process tasks at a level that describe the process requirements for information system functionality and human skills [1]. The definition of the workflow can be made with using the both statements given above or simply the process automation can be defined as the passing of information from one participating system to another and the application of appropriate rules in order to achieve a business objective [2]. However, it is important to emphasize that a workflow refers to the automation of processes that involve the processing of cases and execution of tasks in a particular order, by particular resources so that some objective is met [3].

Some of the researchers have identified workflows as the computing model that enable a standard method of building Web service applications and process to connect and exchange information over the Web [4, 5, 6]. Because the new developments in e-service applications and the higher expectations from the results, they set new requirements and challenges for workflow systems. One of the most important missing requirements is the management and applying quality of service dimensions to the system. Organizations operating such as e-commerce activities and Web service interactions are required to provide QoS to their costumers.

Being able to characterize workflows based on QoS has some advantages.

- Monitoring the system from a QoS perspective: While fulfilling consumer expectations from a business process automation tool, workflows and adjacent tools must be constantly monitored throughout their life cycles to assure both initial QoS requirements and targeted objectives are met. When undesired metrics are identified or when threshold values are reached, the QoS monitoring allows for adaptation of new strategies or stop the process in order to check the workflow design.
- Design the system based on QoS: Because the business process can be designed according to QoS metrics, it allows organizations to translate their vision into their business process more efficiently.
- Selection and Execution based on QoS: It allows for the selection and execution of workflows based on their QoS, to better fulfill customer expectations.

The lack of not having QoS in Workflow automation may result in inconsistent outputs, system failures, and inefficient use of system. Therefore, in this study we present the affects of spelling errors in personal names and the outcome of not recognized costumer expectations in workflow automation. In Section 2, some relevant definitions and background of the system are introduced. The methodology to overcome the problem is discussed in Section 3. At the end, the paper is concluded by summarizing the test result and assessing them in light of related work.

Background

Although Qos has been widely discussed in the areas of real time applications [7], networking [8, 9], just few research teams have concentrated on applying the efforts to the field of Workflow or Business Process Automation. So far, most of the research carried out to extend the functionality of workflow systems and QoS has only been done in time and reliability dimension. The industry has a major interest on the QoS of workflows and fully-automated business processes. The Crossflow [10] and Meteor [11] are the leading projects in the field. Not only time dimension is considered, but also cost associated metrics was designed in Crossflow project. In Crossflow, the information about past workflow execution (which are collected in a log) used to derive a continuous-time Markov chain. Since Markov chains are not directly supporting the concept of parallel executions introduced by the and-split/and-join structures, the power set of parallel activities which can reach millions of states. On the other hand, METEOR prototype exploits CORBA for inter-component communications. It is a fully distributed implementation of the workflow server functionality [11]. METEOR WFMS include a comprehensive toolkit for building workflows (map/data/task design), supporting high-level process modeling, and detailed workflow specification in the health field while using time and reliability parameters for QoS.

Although these research projects applied some of the fundamental QoS parameters to their workflow generation, they lack the representation and formulation of their approaches in the field. Moreover, their applications are restricted to only workflow generation part of the automation. However, impact of the spelling errors in the customer expectation file or unrecognized input records heavily affects the outcome of the workflow process automation. Therefore, the QoS attributes for each part of the system varies and supports enhanced business process automation.

In this paper we will present the reliability parameter of the QoS in Business Process Automation. The parameter is applied to:

- *Understanding Business Language*: The goal statement from the customer.
- *Recognizing Input Records*: The input records supplied by the customer.

Reliability of Business Process Automation

A. Natural Language Parser and Spelling Correction

As allowing the user to use natural language to describe their job requirement, it will remove the burden of data administrators to pick up the goal state fields, ease the complexity of job description, minimize the special training to use the workflow system, and greatly expand the system usability. However, extracting the most relevant information is a complex process, because of free formatted and voluminous data. Moreover, understanding the costumer expectation from ill-defined data (misspelled or mistyped data) is more challenging. Most of the time, the use of Data administrators or a tool that has limited capabilities to correct the mistyped information can cause enormous number of problems. Same as in the case of workflow, the more accurate the selected word is, the more useful the information that is retrieved or the better workflow is provided.

The ill-defined data has different types and definitions. The main source of errors, known as *isolated-word error,* is a spelling error that can be captured simply because it is mistyped or misspelled [12]. As the name suggests, isolated-word errors are invalid strings, properly identified and isolated as incorrect representations of a valid word [13]. *Typographic errors,* also known as “fat fingering”, have been made by an accidental keying of a letter in place of another. These errors are made assuming that the writer or typist knows how to spell the word, but may have typed the word hastily resulting in an error [14]. *Cognitive errors* refer to errors that have been made by a lack of knowledge from the writer or typist [14].

Phonetic errors can be seen to be a subset of cognitive errors. These errors are made when the writer substitutes letters they believe sound correct into a word, which in fact leads to a misspelling [15]. Once a potential word has been detected, then the correction of that word is another issue in the spelling correction process. There are many isolated-word error correction applications, and these techniques entail the problem to three sub-problems: treat detection of an error, generation of candidate corrections and ranking of candidate corrections as a separate process in sequence [14].

Our *Personal Name Recognizing Strategy* (PNRS) is based on the results of number of strategies that are combined together in order to provide the closest match. Before applying any techniques to suggest a valid word for particular field, the information in the proper place needs to be free of non-ASCII characters. Our approach for removing the non-ASCII characters initially requires the original blank spaces to be kept in the data and later used as delimiters; then removes the non-ASCII characters from the records; and finally consolidates the partitioned word pieces. After removing the non-ASCII characters from the input data, prediction algorithm gets executed to provide a suggestion list.

Among the other approaches, the *near miss strategy* is the most fairly simple and widely used way to generate suggestions. In this approach, two words are considered near, if they can be made identical by inserting a blank space, interchanging two adjacent letters, changing one letter, deleting one letter or adding one letter [16]. If a valid word is generated using these techniques, then it is added to the suggestion list. However, the near miss strategy doesn't provide the best list of suggestions when a word is truly misspelled. That is where the phonetic strategy becomes attractive. A phonetic code is a rough approximation of how the word sounds [14]. The English written language is a truly phonetic code, meaning that each sound in a word is represented by a symbol or sound picture**.** The phonetic strategy is comparing the phonetic code of the misspelled word to all the words in the word list. If the phonetic codes match, then the word is added to the suggestion list.

The input data involving names and possible international scope make the phonetic equivalents of certain letters difficult to standardize. Therefore, the algorithm that is used provides results of near miss strategy and phonetic strategy in a suggestion pool at the same time and equal weight [17]. Once we have a list of suggestions (based on the data type) an edit-distance algorithm is either used to rank the results in the pool or to suggest a word for mistyped information. If the input data is a goal state, we specifically use edit distance to rank the results from the pool. The edit distance is defined as the smallest number of insertions, deletions, and substitutions required for changing one string into another [12]. In order to provide meaningful suggestions, the threshold value, t, is defined as 2. In case the first and last characters of a word don't match, we modified our approach to include an extra edit distance. The main idea behind this is that people generally can get the first character and last character correct when trying to spell a word.

On the other hand, if the input records are being used in correction stage, edit-distance is used only to rank the results. Different and better decision mechanism is needed to suggest the best possible solution for individual names. The rational behind this is that at the final stage there is a possibility to see several possible candidate names for a mistyped one which has one or two edit distance from the original word. Relying on edit distance doesn't often provide the desired result. Therefore, we designed our decision mechanism based on the content of the input information. The systems uses the U.S Census Bureau decision mechanism, which is compiled a list of popular first and last names which are scored based on the frequency of those names within the United States [18]. This allows the tool to choose a "best fit" suggestion to eliminate the need for user interaction. The strategies are applied to a custom dictionary which is designed particularly for the workflow automation.

After the correction process is finished for the input and goal state, natural language parser is used to understand the user expectations (goal state). A *parser* for the natural language (English, Portuguese, etc.) is a program that *diagrams* sentences of that language [19]. Most of the natural language processing methods [20] break the process down into several successive stages, using several levels of representation: morphologic, syntactic, or semantic then pragmatic analysis. These conceptual representations and understandable models of texts depend on contextual knowledge: each stage brings in some specific knowledge about the text being analyzed, and that information can be saved in a database [21]. However, there is an inability to draw common-sense reasoning on real world knowledge. For this reason, the Shallow Parsing method was the most suitable for our needs. Shallow Parsing is a form of Information Extraction (IE) which only tries to distinguish a finite number of topics instead of full text comprehension. It looks for named entities, phrases or keywords. By looking at the sentences syntactically instead of semantically, it will reduce the complexity of the NLP. Tagger and Chunker are chosen to be used in this study for one main reason: First high accuracy, the average accuracy that is reported for state-of–the-art data-driven Part-Of-Speech (POS) Taggers lies between 95% and 98% [22].

As shown in high level functional model diagram (Figure 1), the system is composed of a Natural Language Parser (NLPr), a token analyzer (PhraseGroup), a Statistical Analysis Tool (SATool), and a dictionary to provide the list of output fields. Logical natural language sentences are given to NLPr for tagging each word then assembling each tagged word into a Noun Phrase Chunk. The token analyzer compares the gathered information from Natural Language Parser with the designed dictionary and business SATool to provide the list of selected output fields for workflow automation.

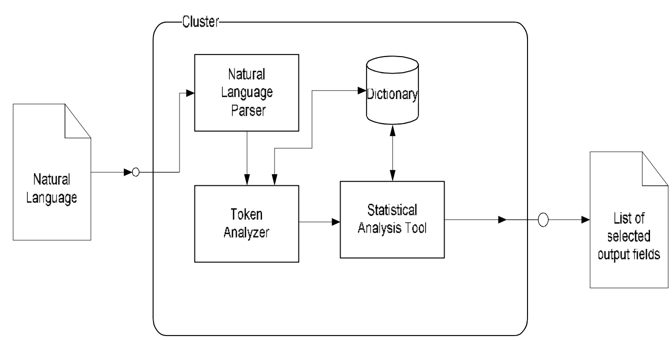


Figure 1. High Level Functional Model of NLP

B. Input File Correction and Result Analysis: A case study

Since input files can consist of any personal or contact information for variety of people, we intend to use a simple task (see Table 1) to make this case study more understandable. Let us assume that the following records are given by a customer, along with the expectation (goal state), to see if "the customer's file to be flagged for dirty words."

Table 1. Input Records Given by Customer

Jas?n Richardson, 123 Main St., Memphis, TN
Jesen Kay, 1269 High St., Santa Monica, CA
Mi?e Barkley, 453 Wilson Ave., Morgantown, WV 26505
Mike Anderson, 1645 Blake Ave., Los Angeles, CA
Herrry Jacobsen, 1385 Broadway #2106, New York
The White House, 1600 Pennsylvania Ave., Washington, DC
Maly Huntington, 567 Main St., Little Rock, AR
icardinh Prince, 2824 3rd St., Dallas, TX

Step1: After removing the question marks and consolidating the remaining pieces, the data reshaped as in Table 2.

Table 2. Input Records Recovered from Non-ASCII Characters

Jasn Richardson, 123 Main St., Memphis, TN
Jesen Kay, 1269 High St., Santa Monica, CA
Mie Barkley, 453 Wilson Ave., Morgantown, WV 26505
Mike Anderson, 1645 Blake Ave., Los Angeles, CA
Herrry Jacobsen, 1385 Broadway #2106, New York
The White House, 1600 Pennsylvania Ave., Washington, DC
Maly Huntington, 567 Main St., Little Rock, AR
icardinh Prince, 2824 3rd St., Dallas, TX

Step2: After applying the PNRS to the file, several candidate names are suggested for some of the misspelled ones. Such as, Jason, Jan, Jayson are possible individual alternatives for Jasn. This is where Census Bureau records are used to make final decision. In this case, the name of Jason much more appeared than the other two at the Census Records. Same algorithm is applied to other records as well and the correction tool provided following suggestions for the misspelled names:

Table 3. Corrected Input Records

Jason Richardson, 123 Main St., Memphis, TN
Jason Kay, 1269 High St., Santa Monica, CA
Mike Barkley, 453 Wilson Ave., Morgantown, WV 26505
Mike Anderson, 1645 Blake Ave., Los Angeles, CA
Harry Jacobsen, 1385 Broadway #2106, New York
The White House, 1600 Pennsylvania Ave., Washington, DC
Mary Huntington, 567 Main St., Little Rock, AR
icardinh Prince, 2824 3rd St., Dallas, TX

As it is illustrated in Table 3, the last record is not corrected with the tool. Although a possible candidate "Ricardinho" is only two edit distance away from this misspelled name, the edit distance is calculated as three and the tool marked this word as out of boundary. The rational behind three edit-distances is the result of both first and last characters are not matching. However, this problem can be solved by applying Longest Common Substring [23] algorithm when t is exceeded two.

Although, our algorithm is unique from the other existing once, there is a need to compare results each other to reflect the effectiveness of the tools. The result provided through the use of PNRS is much more promising than arguable largely known ASPELL [24], JSpell HTML [25] and Ajax Spell Checker [26] (Table 4).

Table 4. Comparison of the Spelling Correction Algorithm

Before	Best Shot with (VBSC)	ASPELL	JSpell HTML	Ajax Spell Checker
Jas?n	Jason	N/A	Jags	Jason
Jasn	Jason	Jason	Jason	Jason
Jesen	Jason	Jason	Jensen	Jasen
Mi?e	Mike	N/A	Eh	No misspelling found
Mie	Mike	Mie	Mime	Mei
Herrry	Harry	Harry	Harry	Harry
Maly	Mary	Mary	Malay	Malay
icardinh	Edit-Distance too high	Saccharin	Circadian	Acton

C. Impact of Input File Reliability on the Business Process Automation

Task reliability (R) models can be considered the most important class of workflow failures, task failures. Task Reliability can be organized into two main classes: workflow system reliability based on user goal state and process reliability of the user input file.

1) *Workflow System Reliability based on User Goal State:* This task consist of business rules exceptions in workflow systems which lead to irrelevant flow for a given goal state. Although, spelling correction algorithms and then NLP are applied to the original user goal state information, some terms are not recognized or provide inconsistent output for a particular case. Therefore, we have calculated the System failure rate as:

$$SF_r = \frac{\lambda}{\tau}$$

where, SF_r, is the ratio between the numbers of time a task did not perform for its users, λ, and the number of times the task was called for execution, τ. Similarly the, Workflow System reliability rate, WRr, is based on the following definition.

$$WR_r = 1 - SF_r$$

2) *Process Reliability of the User Input File:* This part of the study consists of business process exceptions in spelling correction tool which lead to an anomalous termination of a name correction task when fail or abort stages is activated. In the user input file spelling correction tool, task structure has an initial state, an execution state, and two distinct terminating states. The model used to represent each task indicates that only one starting point exists when performing a task, but two different states can be reached upon its execution as show in Figure 2. For example, a name which has more question marks than the letters fails because of the unrecognizable input record, when the task enters the aborted state. On the other hand, the correction attempt reaches the failed state, when edit distance is more than two.

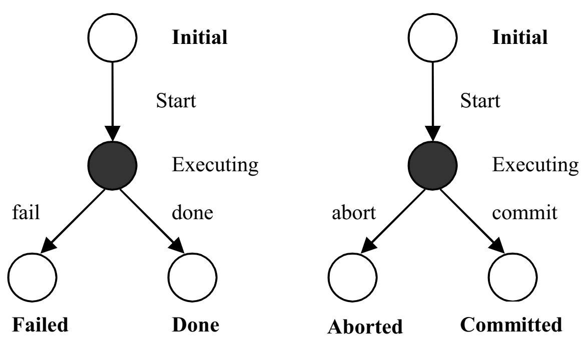


Figure 2. Two Task Structures

The behavior of the process can be defined as:

$$PF_r = \frac{A}{A+B}$$

where PF_r denotes the process failure rate, B represents the success, whereas A is the incomplete process. While, PF_r, provides information concerning the relationship between the number of times the state done/committed is reached, *B*, and the number of times the failed/aborted state, *A*, is reached after the execution of a task, Process Reliability Rate, PR_r can be defined as;

$$PR_r = 1 - PF_r$$

Test Results

Success of the spelling correction and the dictionary for understanding the business language directly impacts the workflow. Correctly recognized business expectations and error-free input record files lead to a generation of a reliable workflow. However, voluminous of the data and especially the international scope of the names reflect as unexpected results. Twelve real life samples applied for both tests. As shown in Figure 3, workflow system reliability based on the natural language processing of the user goal statement ranged from 0.69 to 1. In other words, the Natural Language Parser provides an average success rate of 0.88 correctly designed workflow.

On the other hand, process reliability (Figure 4) which is based on the percentage of correction the names in record fields succeeded from 0.45 to 0.96 with an average of 0.8 in twelve experiments.

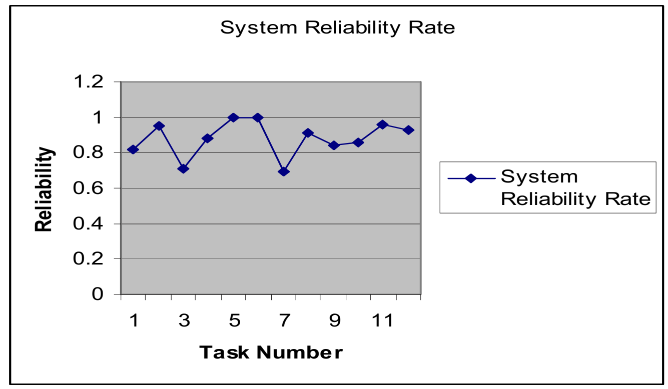


Figure 3. Workflow System Reliability Rate Based on User Goal State

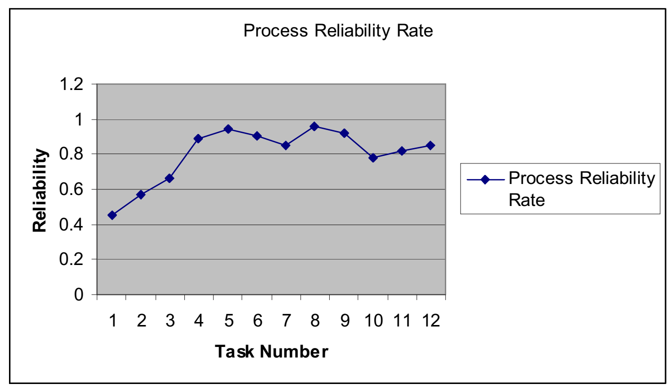


Figure 4. Process Reliability Rate for the Correction of User Input File

Conclusion

In this study, a knowledge-based Natural Language Parser and a spelling correction are briefly presented as an enhancement to workflow automation engine. Based on the experiments conducted, not only the reliability of the workflow engine, but also the impact of user goal state and input file are important as well. Although, the reliability of the input file is calculated, its impact to the workflow is yet to be formulized. Moreover, we are also aiming to put efforts on calculating the error propagation in the workflow.

References

[1] Brouns, G.A.J.F., "Featuring Workflow Management- An overview of the distinctive features

of workflow processes and their consequences for workflow management", *SPOR-Report* 2000-05, Eindhoven University of Technology, 2000

[2] Becker, J., von Uthmann, C., zur Mühlen, M., Rosemann, M., "Identifying the Workflow Potential of Business Processes" *In: Sprague, Ralph Jr.: Proceedings of the 32nd Hawaii International Conference on System Sciences (HICSS 1999)*, Wailea (HI), January 5-8, 1999. Los Alamitos, 1999

[3] Black, D., "Workflow Software: A Layman's Handbook, Part I," *INFORM*, April 1994

[4] Fendel, D., and Bussler, C., "The Web Service Modeling Framework" 2002, *Vrije Universiteit Amsterdam (VU) and Oracle Corporation.* http://www.cs.vu.nl/~dieter/ftp/paper/wsmf.pdf

[5] Shegalov, G., Gillmann, M., Weikum D., "XML-enabled workflow management for e-services across heterogeneous platforms", *The VLDB Journal,* 2001, 10:p.91-93

[6] Chen, Q., et al. "Dynamic-Agents, Workflow and XML for E-Commerce Automation" *EC-Web,* 2000. p.314-323

[7] Clark, D., Shenker, S., and Zhang, L., (1992). "Supporting Real-Time Applications in an Integrated Services Packet Network: Architecture and Mechanism" *Proceedings of ACM SIGCOMM.* pp. 14-26.

[8] Cruz, R. L. (1995). "Quality of service guarantees in virtual circuit switched networks." *IEEE J. Select. Areas Communication.* 13(6): 1048-1056.

[9] Georgiadis, L., Guerin, R., Peris, V., and Sivarajan, K., (1996). "Efficient Network QoS Provisioning Based on Per Node Traffic Shaping." *IEEE ACM Transactions on Networking*. 4(4): 482-501.

[10] Klingemann, J., Wäsch, J., and Aberer, K., (1999). „Deriving Service Models in Cross-Organizational Workflows" *Proceedings of RIDE - Information Technology for Virtual Enterprises (RIDE-VE '99),* Sydney, Australia. pp. 100-107.

[11] Miller, J. A., Fan, M., Wu, S., Arpinar, I. B., Sheth, A. P., and Kochut, K. J., "Security for the Meteor workflow management system", *UGA-CS-LDIS Technical Report*, University of Georgia, 1999.

[12] Levenshtein Edit-Distance Algorithm, http://www.nist.gov/dads/HTML/Levenshtein.html

[13] Becchetti C., Ricotti L. P., 1999. "Speech Recognition: Theory and C++ Implementation". *John Wiley & Sons*

[14] Kukich, K., "Techniques for Automatically Correcting Words in Text", *ACM Computing Surveys*, vol. 24, No. 4, 1992.

[15] ASPELL, http://aspell.net/metaphone/

[16] Near Miss Strategy, http://www.codeproject.com/csharp/NetSpell.asp?print=true

[17] Varol, C., Kulaga, J., Robinette, H., Bayrak, C., "Application of Near Miss Strategy and Edit Distance to Handle Dirty Data", *ALAR 2006*, March 3, 2006, Conway, Arkansas.

[18] Census Bureau Home Page, www.cencus.gov

[19] Black, E., "Evaluation of Broad Coverage Natural Language Parsers" http://cslu.cse.ogi.edu/HLTsurvey/ch13node6.html

[20] Fuchs, C., Danlos, L., Lacheret-Dujour, A., Luzzati, D., Victorri, B. "Linguistique et traitement automatique des langues" Hachette, Paris-France, 1993.

[21] Desclés, J-P., Cartier, E., Jackiewicz, A., Minel, J. L., "Textual Processing and Contextual Exploration Method", *CONTEXT'97*, 189- 197, Rio de Janeiro, Brazil, 1997.

[22] Megyesi, B., "Shallow Parsing with PoS Taggers and Linguistic Features", *Journal of Machine Learning Research 2* (2002) 639–668

[23] Friedman C., Sideli, R., "Tolerating spelling errors during patient validation" *Computers and Biomedical Research*, 25"486-509, 1992

[24] ASPELL, http://aspell.net/metaphone/

[25] Jspell HTML, http://www.thesolutioncafe.com/html-spell-checker.html

[26] AJAX Spell Checker, http://www.broken-notebook.com/spell_checker/

Optimizing the Architecture of Adaptive Complex Applications Using Genetic Programming

Ilham Benyahia & Vincent Talbot
Université du Québec en Outaouais, 101 St-Jean-Bosco St., Box 1250, Station B,
Gatineau, Québec, Canada J8X 3X7
ilham.benyahia@uqo.ca

Abstract- *Complex distributed applications for control and monitoring based on high-speed networking and quality of service (QoS) constraints are found in energy, telecommunication and transportation management such as traffic control based on video transmission. These applications are comprised of agents characterized by large-scale component-based library architecture. Since they interact with unpredictable environments, they run a high risk of performance degradation. In order to deal with environmental uncertainties, there is a requirement for adaptive computing during the system run time. Thus, an automatic evolution of the agent's software through architecture changes is needed. To maintain the agents' QoS, the architecture configuration must be optimized according to the environment behavior. For instance, a complex process for video coding is used if the deadlines are not tight, while simpler processing is necessary when the deadline is tight. Genetic programming techniques are employed to dynamically achieve the optimal architecture configuration. Experiments using our test bed are reported in this paper. Numerical results confirm performance gain following adaptation of the configuration of complex agent architecture to each significant environment behavior*

Keywords- Adaptive complex applications, genetic programming, quality of service, component-based software engineering, frameworks and systems, software architecture optimization.

1. INTRODUCTION

New technologies in computing and telecommunications based on high-speed networking have brought valuable new dimensions to the development of complex applications. Such applications can be found in the control of telecommunication, energy networks and transportation. We consider traffic control in the transportation domain as a multimedia application. The states of traffic and routes are continuously transmitted by videos to a control station where all the information is visualized and analyzed. Then, control data are transmitted to different locations within a specified deadline to avoid route congestion and other consequences after the occurrence of abnormal situations. Such temporal constraints pose serious problems for existing software development methodologies. Current software performance analysis is achieved at the design phase where the architecture decision is taken. Complex applications considered here must be adaptive in order to meet the required quality of service (QoS) for any environment change.

To make processing components adaptive, our approach is based on the design and development of a test bed which will associate the best architecture configuration with each significant environment behavior. This test bed incorporates a genetic programming (GP) framework to find the best fit between an environment behavior and a combination of components.

This paper is organized as follows. Section 2 illustrates complex applications and their characteristics and describes our development framework. Section 3 defines a generic test bed designed to find the best complex agent architecture configuration for each significant environment behavior according to an environment behavior generator. Section 4 presents experiments using our test bed, and a conclusion follows in Section 5.

2. COMPLEX APPLICATION DEVELOPMENT CHARACTERISTICS

A complex system is distributed and interacts with its environment, any change in which implies a flow of events to be processed within a deadline. The main difficulty in managing such systems relates to performance degradation due to the non-foreseen behavior of their associated environment.

2.1 The complex applications environment

We consider two operation modes or regimes for complex applications, the stable regime and the perturbed regime. The stable regime is an environment that behaves in accordance with a foreseen specification based on the inter-arrival of events according to a probability distribution. A system is said to operate in a perturbed regime when its environment is faulty and has a non-foreseen behavior. An example of a disturbed regime is a power system where various failures may occur due to faulty event correlation [10]. Cascading faults can also affect telecommunication networks using immature technologies such as SONET (Synchronous Optical Network) and ATM (Asynchronous Transfer Mode) [5].

2.2 Complex applications development framework

In previous work, we have presented a generic framework [1] to develop complex applications in different domains such as telecommunication [3] and *transportation* [2]. Our framework architecture is illustrated in Fig. 1.

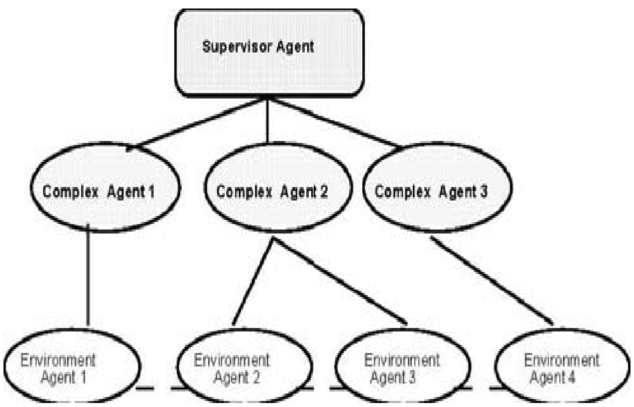


Fig. 1. Architecture of a complex system.

Complex agents are responsible for control and monitoring of the environment *components* called e*nvironment agents* which communicate *events,* and their associated complex agents trigger suitable actions. In a traffic control application, cars represent environment agents that communicate events after any route incident. The *supervisor agent* is responsible for the collaboration of complex agents to improve performance depending on their processing load. A complex agent is characterized by the following sequence of event processing, in which it:

- Perceives significant environment events (reactivity),
- Filters new events when the environment is perturbed,
- Delegates events through the supervisor agent to other complex agents when the complex agent is *saturated* (unable to process events before deadline),
- Considers a global context to evaluate actions to be triggered for accepted events,
- Meets the event deadlines using *scheduling techniques,*
- Executes selected actions.

A complex agent has an architecture made up of different *components,* each responsible for carrying out the above phases using one or several *sub-components.*

2.3 Library of complex applications components

The architecture of each complex agent is based on a library of components. A component can have several sub-components with the same role as scheduling techniques [1]. For instance, in the telecommunication domain, different strategies are presented in the form of components to control the congestion in ATM networks [3]. An architecture component is defined by parameters which can be either generic or application-dependent. The functional type of a component may remain generic. However, parameters such as degree of accuracy and execution time are application- and environment-dependent. For instance, the processing for noise filtering of video transmissions in traffic control can have different degrees of accuracy according to its complexity. In this paper, we consider only two parameters in our component notations, the functional type of the component and the CPU execution time. Thus, a component *c* is defined as follows: *ca_b,* where *a* represents the component function and *b* its execution duration. The selection of complex agent architecture components among those with similar roles can greatly affect the performance of the system. Related work on performance analysis integrated with software design can be found in distributed, concurrent and real-time applications and software-oriented component libraries.

2.4 Software performance analysis

Performance analysis for distributed, concurrent and real-time applications is based on sequence diagrams and scheduling theory [6], [8]. In these approaches, systems are structured in tasks and a time budget (CPU time) is associated with each task. The time budget is used to evaluate the schedulability of these tasks which is defined in terms of a CPU utilization that must be less than or equal to the specified system response time (deadline). Such performance analysis is not suitable for complex applications such as traffic control due to its dynamic environment.

The importance of performance specification becomes painfully evident when its absence leads to harmful consequences as has been reported for component library systems oriented for real time [11]. The performance evaluation approaches for component software performance evaluation which are now gaining the most popularity are based on an execution graph [12]. However, most of them have little relevance for complex applications because they consider limited components.

Complex applications require a design methodology that integrates performance analysis to enable adaptive component interconnections. A complex agent may operate with different combinations of components to process a particular event according to its own state and its environment behavior. To define our design methodology, we consider two main requirements: first, how to define the best combination of complex agent architecture components when existing repositories for large-scale applications contain thousands of different components: and second, how to formulate the quality of an architecture configuration which must be environment-dependent. To deal with combinatorial problems and find a solution that satisfies a required quality, we use a genetic programming technique to generate the best solution.

3. DEVELOPMENT AND OPERATION PARAMETERS OF THE EXPERIMENTAL TEST BED

We aim to generate architecture component interconnections that provide the best QoS according to the current environment behavior.

Our approach is empirical using our experimental test bed which has three main subsystems: an environment generator, a genetic programming framework called ECJ10 [4] and a simulator. The results of experiments are expressed as rules that associate each environment behavior with the best architecture configuration. We present a general overview of genetic programming.

3.1 An overview of genetic programming

Genetic programming with its population of "programs" provides a way to perform program induction through the following steps. These programs are constructed from a predefined set of functions F and terminals T, and are encoded as trees (as opposed to the linear structures, known as chromosomes, found in genetic algorithms) [9].

Step 1. *Initialization.* Create an initial random population of P programs made up of all possible combinations of functions.

Step 2. *Selection*. Select P programs in the current population. The selection *process* is probabilistically biased in favor of the best programs.

Step 3. *Modifications.* Apply operations such as *crossover* and *mutation* to the selected programs to generate new programs.

Step 4. *Evaluation*. Evaluate the *fitness* of each program in the new population by applying this program to a set of fitness cases (examples), and by measuring the correspondence between the desired responses and the responses produced by the program.

Step 5. Repeat Steps 2, 3 and 4 for a pre-specified number of generations or until the system shows no further improvement.

3.2 The test bed structure

The test bed is made up of three subsystems (Fig 2).

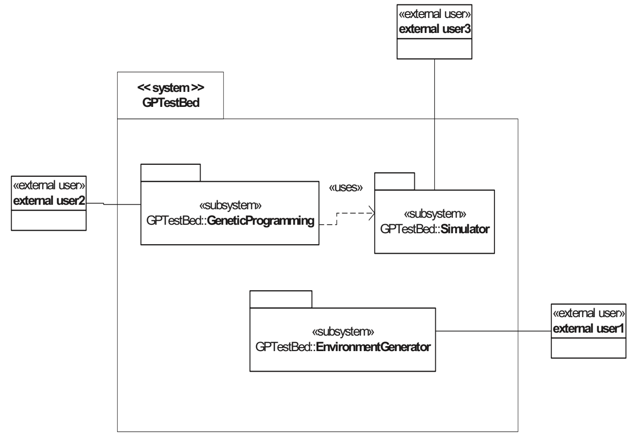


Fig. 2. Architecture of the test bed.

The first subsystem, the *Environment Generator,* generates simulation models representing environment behaviors characterized by events inter-arrival. It produces an Environment object that models environment behavior which is based on distribution laws such as the Normal distribution. Once generated, *Environment* objects can be tested by the *Simulation* subsystem, which simulates the execution of an architecture component interconnection.

The *Genetic Programming* subsystem is an object-oriented framework called ECJ10 [4]. The main class is GP-Problem which defines the genetic problem to be solved and creates and evaluates a population of individuals. To create the individuals of the initial population, GP-Problem defines the GP-Nodes of trees with operators in the form of temporal relations (before, after) and terminals represented by architecture components.

The Evaluation class defined in Step 4 implements the calculation of the individual's fitness as follows. First, it assigns the worst results to the non-validated individuals to be eliminated. Then it calculates the objective function defined by two parameters, quality and performance.

The subsystem responsible for simulating the execution of complex agent architecture is presented as a graph (tree). It relies on the Environment class to provide an environment context under which the simulated architecture configuration executes.

The formula for calculating the individual's fitness value is:

Individual fitness = Individual quality + Individual performance

where

Individual quality = 1- (Qmax – Qtotal) / Qmax (1) – in order to consider that quality globally (considering the size of the population),

Individual performance = EVTlost* / EVTreceived** (2) is the rate of individuals that finish the processing before their specified deadline

Qmax: the maximum Qtotal in the configuration space

Qtotal: the sum of component qualities of a configuration

****EVTlost***: the number of events that cannot be processed before their deadlines

*****EVTreceived***: the number of events received during the simulation

4. EXPERIMENTS USING THE TEST BED

Experiments were conducted using the test bed to optimize the required QoS of complex systems. The objective is to produce accurate results in the form of associations between environment behaviors and the systems architecture configuration. The first step of our experiments is calibration of the test bed parameters.

4.1 Experimental parameter calibration

To optimize the simulation duration, we undertook a series of trial runs to find the minimum duration that would produce the desired results. Based on a maximum task processing time of 3 time units, we generated stable and perturbed Environment objects, and set the minimum duration of a simulation to 300 time units.

The main genetic programming parameters we used for our initial trials are based on GP parameter settings as defined in [4] (see Table 1).

Number of generations	10
Population size	100
Objective function type	0 vs. infinity
Ideal individuals	Fitness = 0.0
Stop criteria	Result convergence
Terminal and operators	9 components and "before" operator
Reproduction method	Tournament : 7
Mutation method	Tournament: 7
Crossover	Tournament: 7
Maximal depth	17

Table1. Genetic programming parameters.

To define a stable regime, we consider the Normal probability distribution. The event inter-arrival is set to three time units with a minimum value of component execution duration set to three time units and up. To model a perturbed environment, we use event inter-arrival time of less than three time units to increase the frequency of events. The perturbed behavior presented in Fig. 4 is obtained by using Weibul distribution to generate inter-arrival times of perturbations, Triangular distribution to generate the duration of perturbations and Normal distribution to generate the amplitude of perturbations. To validate the compliance of the defined environment model, we apply the concept of *moving average* used in statistics [7], [13].

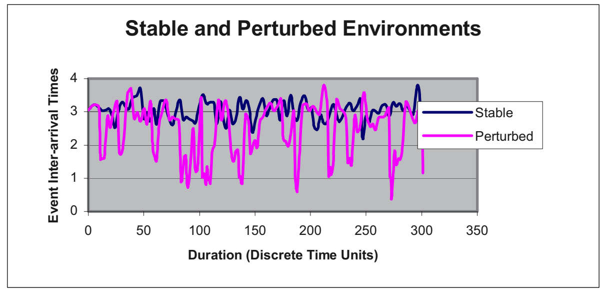


Fig. 3. Environment behavior modeling.

4.2 Experimental results

We carried out two phases of experiments. The first used simulation to generate exhaustive combinations of architecture components, while the second employed simulation based on genetic programming to evaluate the performance of architecture configurations in graph form.

The first phase produced results showing associations between architecture configurations and environment behavior following an exhaustive process of component combinations.

The second phase produced two main results. The first underscores the valuable contribution of genetic programming since there is a rapid convergence to the best solution compared to the exhaustive search. The second phase demonstrates the optimization of the QoS formulated as a GP fitness for a complex application operating in both the stable and perturbed regimes. Table 2 presents examples of numerical results in the form of QoS gain achieved by finding the best configurations. We also measure the potential loss for systems with perturbed environment behavior operating with the best configuration found for the stable environment such as the individual 2_1,1_2,2_3.

The individual 2_1,1_2,2_3 used for the perturbed environment will be responsible for the loss of 75 events representing 54% which is enormous. Thus, processing adaptations are necessary to maintain the required QoS.

Deadline	Best configurations	Occurred events	Lost events	% loss	Fitness
6	1_2,2_1	100	0	0%	0.31
7	2_1,1_2,1_3	100	4	4%	0.18
8	2_1,1_2,2_3	100	1	1%	0.01
9	2_1,1_2,2_3	100	0	0%	0.00
10	2_1,1_2,2_3	100	0	0%	0.00
11	2_1,1_2,2_3	100	2	2%	0.02
12	2_1,1_2,2_3	100	0	0%	0.00
13	2_1,1_2,2_3	100	0	0%	0.00
14	2_1,1_2,2_3	100	0	0%	0.00

Table 2. Examples of gains in QoS for best configurations

5. CONCLUSION

This paper deals with complex applications that interact with unpredictable environments and are based on large libraries of components. Such applications must meet high performance requirements even when operating in extreme situations.

Existing approaches to software design with performance considerations are limited either to typical operating scenarios or to individual component descriptions. Since complex applications interact with continuously changing environments, they must adapt their architecture configuration to each significant environment behavior. However, the challenge is to solve the NP-hard problem to combine the suitable processing components and find the best architecture configuration.

The main contribution of this paper is the use of genetic programming to identify the best solution from a large range of possible processing solutions defined from a library of components. Numerical results obtained from our experiments show the superior efficiency of genetic programming over an exhaustive search for solutions based on simulation. Our experiments also demonstrate the benefits of architecture reconfigurations when complex applications operate in extreme situations. Thus for traffic control applications, we will be able to dynamically change the required components for video processing as noise filtering and visualization to take the best decisions within a specified time.

ACKNOWLEDGMENT

Financial support is provided by the Natural Sciences and Engineering Research Council of Canada (NSERC).

REFERENCES

[1] I. Benyahia and M. Hilali, “An adaptive framework for distributed complex applications development,” IEEE TOOLS USA 2000 (Int. Conf. on Technology of Object-Oriented Languages and Systems), Santa Barbara, California, August 2000, pp. 339-349.

[2] I. Benyahia and J.Y. Potvin, “A framework for information management in dynamic dispatching,” IEEE HICSS-34 (Hawaii Int. Conf. on System Science), 2001.

[3] I. Benyahia and P. Kouamé, “Design methodology for adaptive ATM switch,” in: Proc. 7th IFAC Symposium on Cost-oriented Automation, 2004, pp. 301-304.

[4] http://cs.gmu.edu/~eclab/projects/ecj/

[5] R. Doverspike, “Trends in layered network management of ATM, SONET, and WDM technologies for network survivability and fault management,” J. Network and Systems Management 5, 2 (1997).

[6] H. Gomaa, Designing Concurrent, Distributed, and Real-time Applications with UML, Addison-Wesley, 2000.

[7] K. Goebel and A. Agogino, “An architecture for fuzzy sensor validation and fusion for vehicle following in automated highways,” 29th Int. Symposium on Automotive Technology and Automation (ISATA), June 1996.

[8] D. Jin and D. C. Levy, “An approach to schedulability analysis of UML-based real-time systems design,” in: Proc. 3rd International Workshop on Software and Performance, ACM Press, 2002, pp. 243-250.

[9] R. J. Koza, Genetic Programming: On the Programming of Computers by Means of Natural Selection, Cambridge, Massachusetts: MIT Press, 1992.

[10] S. Pertet and P. Narasimhan, “Handling cascading failures: The case for topology-aware fault-tolerance,” First Workshop on Hot Topics in System Dependability, IEEE, June 30, 2005.

[11] A. Rasche and A. Polze, “Dynamic reconfiguration of component-based real-time software,” in: Proc. Workshop on Object-oriented Dependable Real-time Systems (WORDS 2005), Sedona, Arizona, February 2005.

[12] J. T. Russell and M. F. Jacome, “Architecture-level performance evaluation of components-based embedded systems,” in: Proc. 40th Annual ACM IEEE Design Automation Conference, 2003, pp. 396-401.

[13] http://lorien.ncl.ac.uk/ming/filter/filewma.htm

IVJ: A Soft Real Time Scheduling Algorithm

Sedigheh Asiaban
Se.asiaban@mail.sbu.ac.ir

Mohsen Ebrahimi Moghaddam
m_moghadam@sbu.ac.ir

Electrical and computer engineering department, Shahid Beheshti University; GC, Tehran, Iran

Abstract

The problem of scheduling real-time tasks with minimum jitter is particularly important in many multimedia applications; however, there are no much studies in this field of research. Here, a new algorithm that is called Improving Variance of Jitter (IVJ) is proposed to schedule non-preemptive tasks efficiently. The goal of this algorithm is to improve the jitter of the well known Earliest Deadline First (EDF) algorithm, particularly in overloaded conditions. Proposed algorithm is based on a new parameter that is added to each process and it is called Start time. Start time is calculated by proposed method dynamically. Simulation results showed that the IVJ algorithm could efficiently reduce the jitter and it had a better miss balancing.

Key Words- jitter, Non-preemptive, Real time scheduling, Earliest Deadline First (EDF)

1. Introduction

Multimedia applications are used widely on computers nowadays. These applications have some real time requirements. The real time properties which are required to play back multimedia applications acceptably are often identified as Quality of Service (QoS) parameters. These parameters include miss rate, minimum and maximum delay (that bound the jitter), average bandwidth available and peak bandwidth available [1]. The guarantee of quality of service (QoS) has become an important topic in the design of real time operating systems. One of the most important issues to guarantee QoS is scheduling algorithms that are employed at real time operating systems. Most of the recent researches have concentrated on variants of well known Earliest Deadline First (EDF) algorithm to present scheduling algorithms to decrease miss rate, fix jitter and etc [2].

However, in multimedia applications, it is not necessary for every instance of a periodic task to meet its deadline [3] and it depends on the application, e.g., the ear is more sensitive from the eye, so a variance of even a few milliseconds in delivery times is acceptable, however the delay time cannot exceed a specified bound. As a result, delivery rate in multimedia applications is very important and should be strictly bounded to receive good performance [2].

Some researches have previously been done on the topic of jitter control. Stankovic and Di Natale proposed a scheduling approach for distributed static systems that develops innovative ways to specialize simulated annealing to find solutions which contain feasible schedules with minimum jitter [4]; also they have studied jitter-minimization in the context of providing end-to-end timing guarantees in the distributed real-time systems [5]. S. Baruah and G. Buttazzo have studied the output jitter in the context of the preemptive uniprocessor scheduling of periodic real-time tasks, and have proposed two scheduling algorithms for minimizing jitter [6]. The one is polynomial-time while the other is pseudo-polynomial time in the size of the problem, but provides better jitter bounds. Their algorithms achieve very good results when the system is underloaded [6]. S. Baruah and D. Chen have addressed the problem of accommodating jitter in situations that there is little or no control over the occurrence of jitter [7]. K. Lin and A. Herkert have discussed the jitter control issue in time-triggered real time systems using the DC (Distance Constrained) scheduling model to reduce jitters in time-triggered systems [8].

Researchers in network communications have also devoted considerable attention to the phenomenon of jitter. E.g, A. Dua and N. Bambos have studied low-jitter scheduler design for deadline-aware crossbar packet switches, within a DP (Dynamic Programing) framework [9]. Y. Mansour and B. Patt-Shamir have proposed on-line algorithms that controls jitter [10]. Also Jitter-control schemes have been proposed by Ferrari [11] to minimize the end-to-end jitter in circuit-switched networks. Jorg Liebeherr and Erhan Yilmaz have presented a modified version of EDF, Earliness-based EDF (EEDF), which is a compromise between work conserving and non-work conserving disciplines. It attempts to consolidate the advantages and drawbacks of the work conserving and non-work conserving versions. They demonstrated that EEDF schedulers can balance buffer requirements and average end-to-end delay [12]. Dinesh C. Verma and Hui Zhang have described a method for guaranteeing delay jitter for real-time channels in a packet-switched store-and-forward wide-area network. This method is similar to proposed method but they are not same. They proposed that for each node (n), local delay bound ($d_{i,n}$) and local jitter bound ($j_{i,n}$) are determined based on source-to-destination delay bound (D) and source-to-destination delay jitter bound (J). So a packet arriving at node n at time t_0 will usually be assigned a node deadline equal to $t_0 + d_{i,n}$, and a node eligibility-time equal to $t_{0+} d_{i,n} - j_{i,n}$. The packet is ineligible for transmission until its eligibility-time, which ensures that the minimum delay requirements for channel i are met in node n. Note that performance parameters such as D and J must be specified by clients [13].

In this paper, a new scheduling algorithm is proposed, that is called Improving Variance of Jitter (IVJ). This algorithm decreases the jitter of the EDF scheduling, particularly in overloaded conditions. Overload conditions usually occurs in multimedia operating systems when many clients try to watch videos. An example of such a system is near video on demand systems. The simulation results show that proposed algorithm always has lower jitter variance and more miss balancing in overloaded conditions than EDF. The miss ratio of EDF algorithm is lower than the proposed algorithm in underloaded conditions, but when the load of system and number of tasks increase, the obtained results of both algorithms show similar data. The time complexity of IVJ is the same as EDF.

The reset of the paper is organized as follows. In Section 2, a formal real time system model is presented and EDF algorithm is described. Proposed IVJ scheduling algorithm and simulation results are shown in sections 3 and 4, respectively and the conclusion is given in Section 5.

2. Background

2.1. Classic model of a real time system

Real-time applications generally have a sequence of temporally dependent computation units executing some application functions. In real time applications, computation units can be periodic or sporadic and have constant or variable processing time [1]. Here, we focused on Periodic Variable Processing Time (PVPT) real-time applications.

A computation unit with timing constraints is defined as a job $j = (r, c, d)$, which r is its release time (or its arrival time); c is its computation time and d is its deadline, and a sequence of dependent jobs is called as a task $T = (J, P)$, which J is a set of jobs and P defines the task's periodicity [1]. In this model a real time system is a set of tasks $\tau = \{T_1, T_2, T_3, \dots, T_N\}$.

2.1. EDF

The EDF scheduling algorithm is a priority driven algorithm in which higher priority is assigned to the request that has earlier deadline, and a higher priority request always preempts a lower priority one. In a set of preemptive tasks (periodic, aperiodic, or sporadic), EDF will find a schedule if it is possible [6, 8, 14, 15]. EDF scheduling algorithm achieves very good results when the system load is underloaded for both preemptive and non-preemptive tasks. But, when the system is overloaded, it has been shown that the EDF approach leads to very poor performance and also, it does not guarantee the regular jitter [14]. Therefore, a new algorithm should be provided for these conditions.

Jitter refers to the variation between the inter-completion times of successive jobs of the same task [6]. More formally;

$$D_i^{min} = \min_{k \geq 0} \{f_i^{k+1} - f_i^k\} \quad (1)$$

$$D_i^{max} = \max_{k \geq 0} \{f_i^{k+1} - f_i^k\} \quad (2)$$

When D_i^{min} and D_i^{max} denote the minimum and maximum delay between successive completions jobs of T_i. f_i^k and f_i^{k+1} show the finish time of k and (k+1) job of T_i.

A jitter-free schedule is one that $D_i^{min} = D_i^{max} = P_i$, where P_i is the period of T_i. Schedules that have as their secondary objective the minimization of output jitter attempt to minimize the variation of D_i^{min} and D_i^{max} such that they tend toward P_i [6].

In EDF algorithm, when j_i^k (k^{th} jobs of T_i) does not complete sooner than c_i (the computation time of each jobs of T_i) unit time after its $j_i^k.r$ (release time of j_i^k) and also it does not complete later than c_i unit time than $j_i^k.d$ (deadline of j_i^k), then we can conclude that c_i and $2P_i - c_i$ are bounds on the maximum and minimum inter-completion times of T_i, respectively [6].

3. IVJ Algorithm

In proposed IVJ algorithm a new parameter is added to each job that is called Start time, which is dynamically set in the run time. This parameter determines start time of each job, such that each job cannot start its execution before the Start time. The Start time value is identified by IVJ algorithm dynamically. Table 1 presents the notations that are used to describe IVJ algorithm.

Table 1. Description of notations used in IVJ

Notation	Description
j_i^k	k^{th} job of T_i
$j_i^k.s$	Start time of j_i^k
$j_i^k.r$	Release time of j_i^k
STN_i	Start time of next job of T_i
$Q_{Release}$	The queue of jobs are released up to now
TC	The Current time
Q_{Ready}	The queue of jobs are ready up to now ($j_i^k.s \le TC$)
P_i	Period of T_i
$j_i^k.f$	Finish time of j_i^k
$j_i^k.c$	Computation time of j_i^k
$j_i^k.d$	Deadline of j_i^k

IVJ algorithm

Step 1. If a new job (j_i^k) arrived, set its start time parameter as follows:

1-1 If j_i^k is the first job of task T_i ($k = 1$) then $j_i^k.s = j_i^k.r$

1-2 Else $j_i^k.s = STN_i$

1-3 Insert j_i^k in $Q_{Release}$.

Step 2. Update Q_{Ready}: each job (j_i^k) from $Q_{Release}$ that $j_i^k.s \le TC$ should be removed from $Q_{Release}$ and add to Q_{Ready}.

Step 3. Select a job with earliest deadline from Q_{Ready} and remove it from this queue.

Step 4. Execute selected job and set STN_i as follow:

4-1 If selected job (j_i^k) can be completed successfully then $STN_i = j_i^k.f + P_i - j_i^k.c$

4-2 Else $STN_i = j_i^k.d + P_i$

Step 5. Go to step 1.

According to IVJ definition, the bounds on the maximum and minimum inter-completion times of T_i are P_i and $2P_i - c_i$, respectively.

Note that the single processor system is assumed here.

4. Simulation results

To compare IVJ and EDF algorithms, several numerical tests are simulated. MATLAB is used to generate tasks using the random probability distributions. For each chosen set of parameters, each experiment has been repeated 20 times and the average of miss rate, jitter variance mean, and miss variance are computed. Each experiment was terminated when the predetermined simulation time (5000 unit time) has expired.

In the following, the results are showed and the sensitivity of IVJ is analyzed with regards to the various parameters including the total system load, and the number of the tasks. The non-preemptive task model is used for evaluation.

The following relations were used to evaluate algorithm results (most of the parameters of these equations are defined in Table.1):

The total load of system is defined as

$$\theta(\tau) = \sum_{i=1}^{N} \theta(T_i) \qquad (3)$$

Where N is the number of the tasks, τ is a set of task and

$$\theta(T_i) = \frac{1}{P_i} \sum_{k=0}^{n} \frac{j_i^k.c}{n} \qquad (4)$$

Is load of task T_i. The average of jitter variance of total system is defined as:

$$J_{out}^{avg}(\tau) = \frac{1}{N} \sum_{i=1}^{N} J_{out}^{var}(T_i) \qquad (5)$$

Where $J_{out}^{var}(T_i)$ is the jitter variance of T_i and defined as below:

$$J_{out}^{var}(T_i) = \frac{1}{N_{succ}(T_i)} \sum_{k=1}^{N_{succ}(T_i)} (j_i^k.f - J_{out}^{mean}(T_i))^2 \qquad (6)$$

$$J_{out}^{mean}(T_i) = \frac{1}{N_{succ}(T_i)} \sum_{k=1}^{N_{succ}(T_i)-1} (j_i^{k+1}.f - j_i^k.f) \qquad (7)$$

Equation (7) shows the mean of output jitter of task T_i and $N_{succ}(T_i)$ is the number of jobs of task T_i that are finished successfully.

The miss variance of the total system is defined as

$$var_{miss}(\tau) = \frac{1}{N} \sum_{i=1}^{N} (N_{miss}(T_i) - mean_{miss}(\tau))^2 \qquad (8)$$

Where

$$mean_{miss}(\tau) = \frac{1}{N} \sum_{i=1}^{N} N_{miss}(T_i) \qquad (9)$$

Equation (9) shows the miss mean of the total system, and $N_{miss}(T_i)$ is the num of jobs of task T_i that are missed and finally MR, the ratio of num of jobs that are missed with regards to num of total jobs, is showed by

$$MR = \sum_{i=1}^{N} \frac{N_{miss}(T_i)}{N_{succ}(T_i) + N_{miss}(T_i)} \qquad (10)$$

Figures 1-9 show the simulation results of IVJ algorithm in compare with EDF algorithm.

4.1. The effect of system load and the number of the tasks on jitter variance

In figure 1, 2 and 3 $J_{out}^{avg}(\tau)$ is defined as the average of jitter variance of total system, $\theta(\tau)$ is total load of system, and N shows the number of the tasks.

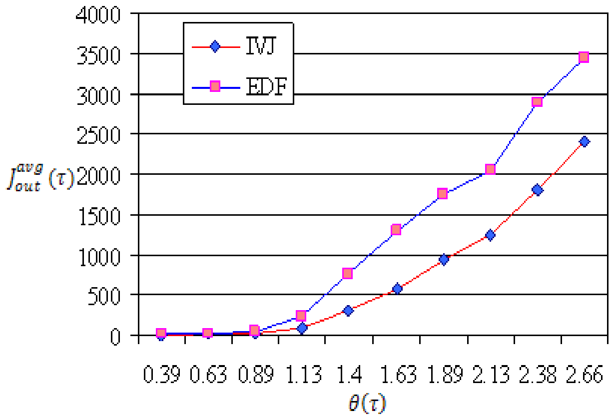


Figure 1. Average of jitter variance of total system v.s the total load of system which is received by simulation when $N = 10$

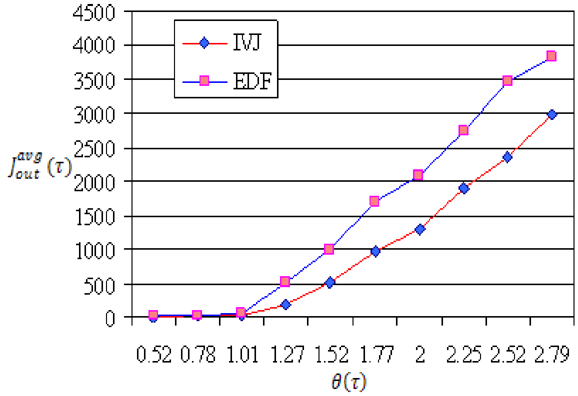


Figure 2. Average of jitter variance of total system v.s the total load of system which is received by simulation when $N = 20$

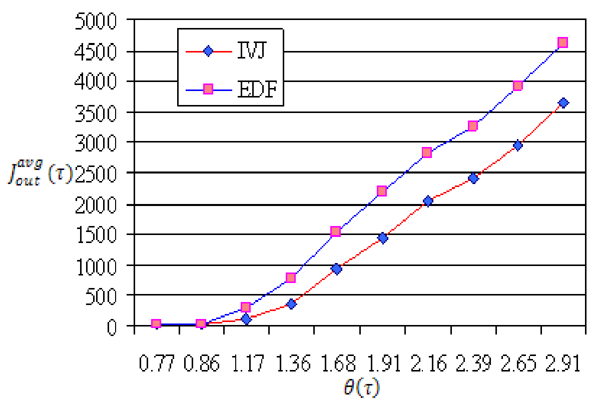


Figure 3. Average of jitter variance of total system v.s the total load of system which is received by simulation when $N = 30$

As it is presented in the figures 1, 2, and 3, jitter variance in IVJ algorithm is always lower than EDF algorithm. When the system is overloaded ($\theta(\tau) \geq 1$), the jitter variance difference between two algorithm is remarkable. This gap is preserved when the system load and the number of the task increased.

4.2. The effect of system load and the number of the tasks on miss variance

In the graphs of figures 4, 5, and 6 $\text{var}_{\text{miss}}(\tau)$ is a parameter that shows miss variance of the total system, $\theta(\tau)$ is total load of system, and N is related to the number of the tasks.

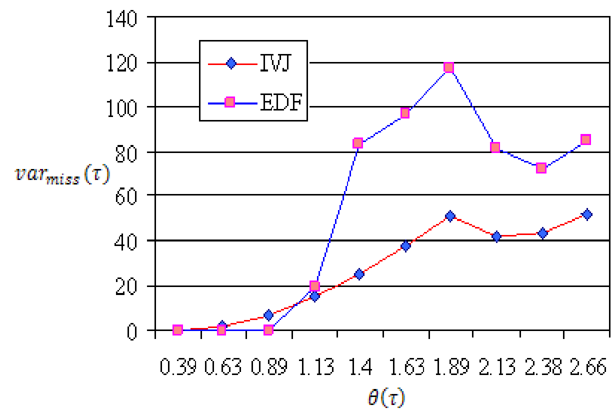


Figure 4. Miss variance of total system v.s the total load of system which is received by simulation when $N = 10$

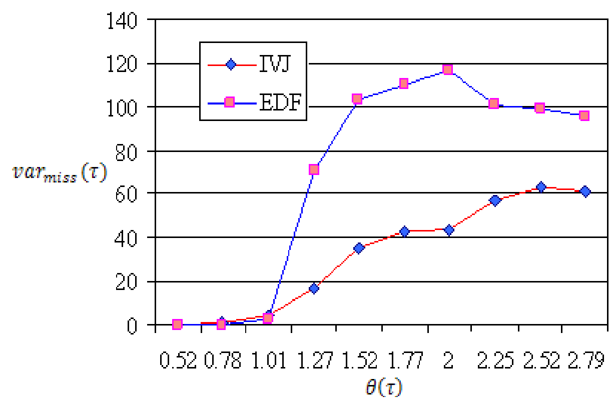


Figure 5. Miss variance of total system v.s the total load of system which is received by simulation when $N = 20$

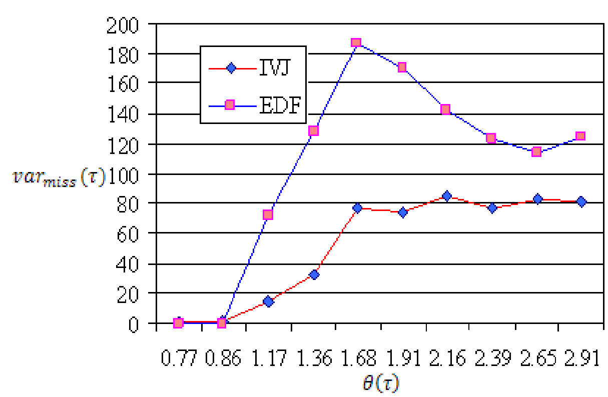


Figure 6. Miss variance of total system v.s the total load of system which is received by simulation when $N = 30$

As it is shown in figure 4, 5, and 6, miss variance in IVJ algorithm is lower than EDF algorithm when the system is overloaded ($\theta(\tau) \geq 1$) and the miss variance difference between two algorithm is remarkable. This gap is preserved when the system load and the number of the task increased.

4.3. The effect of system load and the number of the tasks on miss ratio

In the figures 7, 8, and 9, MR is miss ratio of the total system and θ (τ) and N defenitions are same as the last figures.

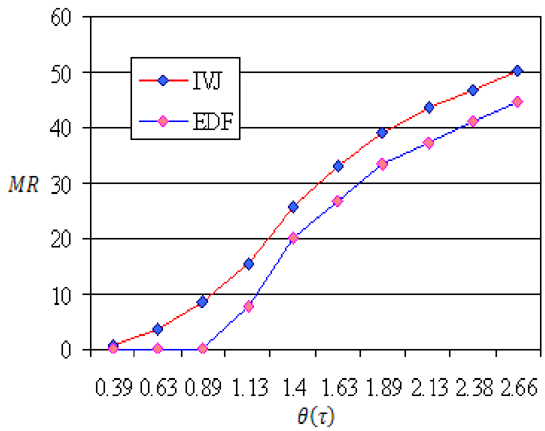


Figure 7. Miss ratio of total system v.s the total load of system which is received by simulation when $N = 10$

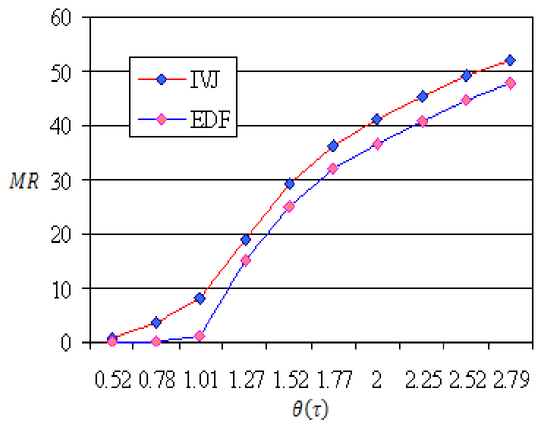


Figure 8. Miss ratio of total system v.s the total load of system which is received by simulation when $N = 20$

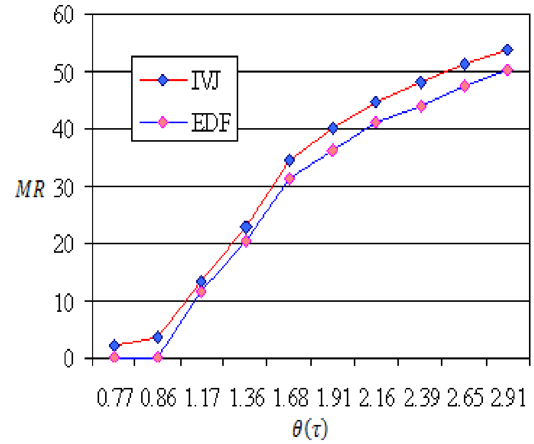


Figure 9. Miss ratio of total system v.s the total load of system which is received by simulation when $N = 30$

As it is shown in these figures, miss ratio in EDF algorithm is always lower than IVJ algorithm. However, when the number of the task increased, their difference is not so remarkable and they are near.

5. Conclusions and future work

In this paper, we proposed a new scheduling algorithm, to improve jitter of the well known EDF algorithm, particularly in overloaded conditions. The overload condition usually occurs in multimedia systems. Therefore, the proposed method is suitable for these kinds of systems. The proposed method is called IVJ. In IVJ algorithm a new parameter is added to each job as Start time, which is dynamically set in run time. The simulation results show IVJ algorithm has lower jitter variance and more miss balancing in comparison with EDF, especially when the load on system is very high.

In future work, we plan to evaluate IVJ algorithm for real work loads and find out the impact of IVJ on other QoS parameters.

Acknowledgements

We are appreciate to thanks the Iran Telecommunication Research Center (ITRC) for their fund support of this research.

References

[1]. Andrew S. Tanenbaum, *Modern operating systems*, 2nd edition, Prentice Hall, 1999.

[2]. Wenming Li, Krishna Kavi, and Robert Akl, “A non-preemptive scheduling algorithm for soft real-time systems,” *Computers & Electrical*

Engineering, Volume 33, Issue 1, January 2007, Pages 12-29.

[3]. Yoo, M.R, Ahn B, Lee D.H, and Kim H.C, "A new real-time scheduling algorithm for continuous media tasks," *Communications, Computers and signal Processing,* IEEE Pacific Rim Conference, 2001, Volume 2, Issue, pp. 417 - 420.

[4]. Di Natale M,Stankovic J.A, "Scheduling Distributed Real-Time Tasks with Minimum Jitter," *Transactions on Computer*s, Apr 2000, Volume 49, Issue 4, pp. 303 – 316.

[5]. Di Natale M, and Stankovic J.A, "Dynamic end-to-end guarantees in distributed real-time systems," *Real-Time Systems Symposium*, 1994, Proceedings, pp.216 – 227.

[6]. Baruah S, Buttazzo G, Gorinsky S, and Lipari G, "Scheduling periodic task systems to minimize output jitter," *Real-Time Computing Systems and Applications*, 1999, Sixth International Conference, pp. 62 – 69.

[7]. Sanjoy Baruah, Deji Chen, and Aloysius Mok, "Jitter concerns in periodic task systems," *In Proceedings of the Eighteenth Real-Time Systems Symposium*, December 1997.

[8]. K.J. Lin and A. Herkert, "Jitter control in time-triggered systems," *In Proceedings of the 29th Hawaii International Conference on System Sciences*, Maui, Hawaii, January 1996.

[9]. Dua Aditya and Bambos Nicholas, "Low-Jitter Scheduling Algorithms for Deadline-Aware Packet Switches," *Global Telecommunications Conference*, 2006, pp. 1 – 5.

[10]. Mansour Y, and Patt-Shamir B, "Jitter control in QoS networks," Networking IEEE/ACM Transactions, Aug 2001, Volume 9, Issue 4, pp 492 – 502.

[11]. D. Ferrari, "Client requirements for real-time communications services," *IEEE communications*, November 1990, Volume 28, Issue 11.

[12]. Jorg Liebeherr, Erhan Yilmaz, "Workconserving vs. Non-workconserving Packet Scheduling: An Issue Revisited," *In Proc. IEEE/IFIP IWQoS*, May 1999.

[13]. D.Verma, H. Zhang, and Domenico Ferrari, "Delay jitter control for real-time communication in a packet switching network," *In Proceedings of the IEEE Conference on Communications Software: Communications for Distributed Applications and Systems*, 1991.

[14]. Thomas Gustafsson, and Mehdi Amirijoo, "Which one sucks: EDF or RM?," Technical Report, Department of Computer Science, George Mason University, January 20, 2006.

[15]. Yuan, Nahrstedt K, and Kihun Kim, "R-EDF: A Reservation-Based EDF Scheduling Algorithm for Multiple Multimedia Task Classes," *Real-Time Technology and Applications Symposium*, Proceedings, Seventh IEEE, 2001, pp 149 – 154.

Videoconference Support System Using Printed and Handwritten Keywords

Kikuo Asai[1,2], Hideaki Kobayashi[2], Kimio Kondo[1,2], and Noritaka Osawa[1,2]
[1] *National Institute of Multimedia Education*
[2] *The Graduate University for Advanced Studies*
{asai, hidekun, kkondo, osawa}@nime.ac.jp

Abstract

We developed a videoconference support system using keywords for distance lectures between outdoors and indoors. The presenter holds to show keyword cards so that the presentation can be done outdoors without any complicated gadgets. The audiences at remote sites see additional information with the keywords. Although keyword captions were originally used in newscasts for supporting comprehension, handheld keyword cards may have the potential for improving the quality of telecommunications. We describe design and implementation of our prototype system that recognizes printed and handwritten keywords in video images and presents their relevant information. We then show a preliminary experiment to investigate the properties of keyword presentation by comparing handheld cards and stationary captions, and obtained the results that the handheld cards are better than the stationary captions in following talks and keeping concentration.

1. Introduction

Audiovisual telecommunication is one of the best means as alternatives to face-to-face communication when the groups are in locations apart. Commercial videoconferencing systems and videoconferencing software applications have been equipped with additional functions such as material and image sharing, chatting, file transfer, and voice calling.

Although the additional functions are convenient in remote lectures and business meetings, they are not sufficient for interactive telecommunications between outdoor fields and indoor classrooms. In fieldwork, for example, it is difficult for a speaker to make a presentation with typical presentation materials such as PowerPoint (Microsoft), because of inconvenience of holding a laptop PC. A palm-on gadget was considered to be one of the solutions for the problem in carrying such heavy equipment. However, manipulating the gadget makes the speaker look down during the talk. This environment makes it difficult to keep eye contact among the participants.

We have therefore proposed an approach of keyword based telecommunications [1]. In this approach, a presenter shows keywords instead of presentation materials, and the audiences at remote sites see additional information with the keywords. We here developed a videoconference support system using handheld cards of printed and handwritten keywords, so that a speaker can make presentations outdoors without handling complicated gadgets.

Keyword captions have been used as redundant information for supporting audiences' comprehension in newscasts. It has been demonstrated experimentally that there is no competition for input information presented simultaneously in visual and verbal media [2, 3]. It was reported that the synchronous dual media presentations were an efficient way to improve learning. This is explained as dual coding theory that hypothesizes representational connections of sensory input and response output systems independently and cooperatively to mediate visual and verbal behavior [4].

Keyword captions were then applied to video materials for learning second languages [5, 6]. The captions are expected to have two main reasons why they are preferred to full-text captions; (1) people have limited ability to process information within a certain time [7] and (2) text information tends to be looked at in video images [8].

We here took a handheld-card-based approach to smooth presentations at an outdoor environment. A presenter holds to show keyword cards in which keywords are printed beforehand or handwritten on site when required during a session. Although the handheld-card-based approach is expected to produce the merit of attention awareness due to movements of the keyword cards, there is no empirical study on presentations using handheld cards. We investigated the properties of the handheld-card-based approach with a preliminary experiment in terms of presentations.

2. Related Work

Using keywords at a videoconference has not really been attempted, even though there are many systems that use text information as a method of supporting telecommunications. The text chat systems provide us with an online communication environment on the Internet [9], where users communicate with one another by typing a line of text. The text communications have been extended to annotations and whiteboards in videoconferencing [e. g., 10, 11], sharing text comments and drawings among remote participants.

A multimedia distance learning system has been developed as a virtual classroom that broadcasts a live class with presenter's video, presentation slides, and handwritten text on a whiteboard [12]. In the system, both the presenter's appearance and the text information have been transmitted as videos separately. The audience can see the presenter and the text information simultaneously in the different windows.

Text information in presentations was also used for language translation. Automatic simultaneous translation of presentation slides was realized by linking the presenter's slide with the PowerPoint file, being used as an input to a PowerPoint translation package [13]. Although these systems would be convenient in videoconferencing, they did not provide any mechanism for cognitively supporting collaboration among remote people.

Nonverbal cues such as gaze and gesture play an important role in catching what a speaker intends to say in telecommunications as well as in face-to-face communications. Gaze awareness was supported in multipoint videoconferencing by providing camera/display surrogates that the participant looked at during a session [14]. Eye contact was faithfully maintained by placing cameras behind the image of the eyes using a semi-transparent screen [15]. Multiple participants were allowed to keep eye contact in a group-to-group meeting by narrowing a viewing direction with a multiple layer screen [16]. Although these systems have supported nonverbal cues preserving spatial faithfulness, they did not introduce any approach for improving presentations to attract attention and create awareness that would make telecommunications clearer in videoconferencing.

3. Scenario

We here describe a scenario in which we suppose how the handheld-card-based telecommunications are used for distance education and the infrastructure where we try to apply our prototype system to videoconferencing.

Figure 1 shows an overview of telecommunications sharing keywords and their relevant information. A presenter speaks outdoors holding some cards of printed or handwritten keywords. The printed keywords and their relevant information are prepared beforehand, and the handwritten keywords are written on the spot. The audiences indoors see the keywords with the speaker's face in a video-image window and the additional information related to the keywords in a different window.

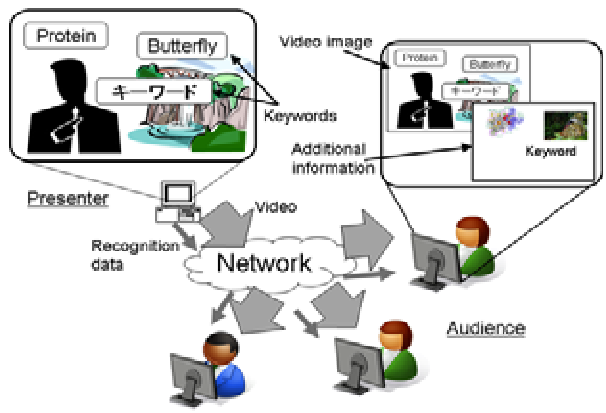


Figure 1: Telecommunications using keywords.

Another scenario for the handheld-card-based telecommunications is intercultural exchange that requires dealing with a wide range of topics and changing content flexibly responding to audiences. The presentation style using handheld cards enables a speaker to facilitate modification on topics or order of the topics during a session. In this scenario, a speaker presents a series of keywords holding cards. The audiences including some foreigners at remote sites see the keywords and the translated words simultaneously.

Let us consider a basic setting of remote lectures between an outdoor field and indoor classrooms. We here suppose that a multipoint videoconferencing system, called Space Collaboration System (SCS) [17], is used for the remote lectures. SCS has been operated for educational exchanges among universities using satellite communications network in Japan. It includes 150 stations at universities, a transportable earth station, and a hub station.

When a session is held among several universities, all the stations cannot always transmit audiovisual signals due to limitation of satellite channels. The basic assignment of the satellite channels is one-channel transmission and two-channel reception for each station, and allows only two stations to transmit simultaneously in a session group. In SCS, a chair station is registered in advance, and participants at the chair station can manage the transmitting stations based on transmission requirements from the other stations.

4. System Design

Figure 2 shows an example of the keyword presentation in our prototype system, assuming telecommunications at intercultural exchange. A user presents a handwritten keyword in the image window, and the translated words as information related to the keyword are displayed in the content window. The prototype system enables the user to present annotations, images, 3D models and sounds as well as translations as the content of the relevant information.

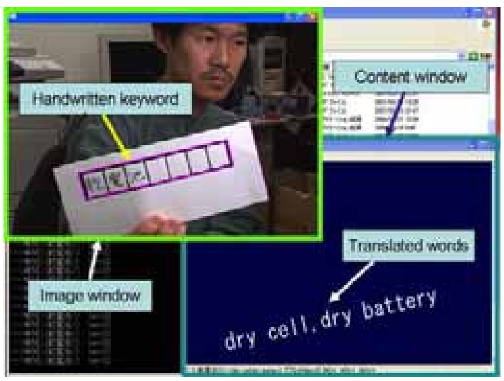


Figure 2: Keyword presentation.

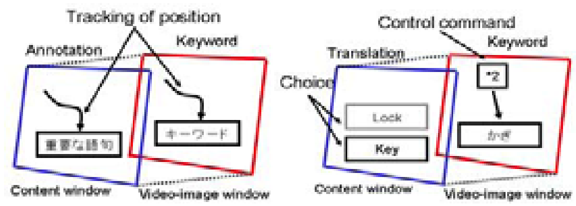
4.1. Functions

The content presentation was equipped with two main functions for visually relating the content to the keyword and intuitively controlling the content.

4.1.1. Layout of the content. A speaker sometimes prefers to present multiple keywords in videoconferencing. The relevant information is presented for all keywords, but the audiences may be confused about which keyword corresponds to the relevant information. We used a visual strategy to maintain the correspondence between them.

The content of the relevant information is collocated to the position in the content window, based on the position of the keyword cards in the video-image window, as shown in Figure 3 (a). That is, the annotation word is moved as if tracking the position of the keyword card. This makes it easy to understand the spatial relationship between keywords and annotations when multiple keywords are presented simultaneously and dynamically change in position.

4.1.2. Control command. There are some cases that a user wants to control the content of the information related to the keywords. For example, one of the translated words should be selected with the speaker's intention when the keyword has synonyms. An additional marker works as a control command to select the translation, as shown in Figure 3 (b). Besides the selection, the control commands change the font color of the translated words and change the content from the translations to the others such as annotations, images, and 3D models.



(a) Collocation of the relevant information. (b) Selection of the translated words.

Figure 3: Functions.

4.2. Architecture

Figure 4 outlines the architecture for our prototype system. The system is mainly composed of image processing and content presentation components. These two components have a server-client relationship, and they exchange data using socket communications. UDP/IP is used as the protocol to support multiple clients on the network.

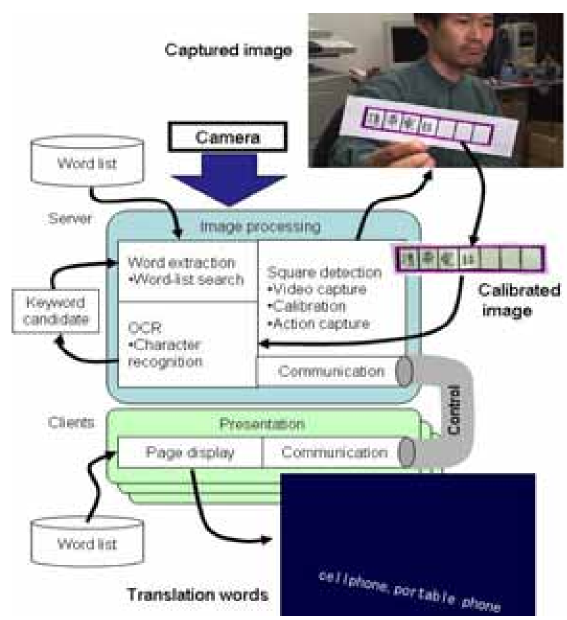


Figure 4: System architecture

4.2.1. Image processing. Video images are captured by a camera, and are fed into a server. A square frame of the keyword card is detected in the video images. A matrix of the position, orientation, and size is obtained for the square frame, and is used for the layout of the content in the content window being transferred to clients.

The pixels inside the square frame are clipped out as an area where keyword characters have been printed or handwritten in the keyword card. The pixel image is calibrated to a rectangle for trapezoidal deformation of the square frame before recognition process.

The recognition process begins with character recognition for the printed or handwritten keywords using optical character recognition (OCR) engines. The recognition result is output as a possible keyword. Unfortunately, the OCR engines sometimes make mistakes in recognizing the characters. Therefore, a word list has been introduced in order to check whether the possible keyword matches the words in the word list, occasionally correcting incorrect ones. The keyword is transferred to clients with the data of the position, orientation, and size.

When the OCR engines do not output any recognition results, the process proceeds to check whether the pixel image matches the image data of the pre-registered markers for the control commands. The detection of the control command is noticed to the clients, and the command controls the presentation of the content in the content window.

4.2.2. Content presentation. The content is displayed on an HTML browser at the clients, based on the keyword-recognition data transferred from the server. The sort of content is initially set at a parameter file at the server, and is changed on the content window at the clients to the other that the user prefers.

The content is presented at the layout formed by an HTML format. The layout of the content in the content window is set based on the position, orientation, and size of the keyword card in the video-image window. When the keyword cards shift, rotate, and zoom-in or zoom-out, the content also shifts to the corresponding position, rotates to the corresponding orientation, and magnifies or reduces the corresponding size.

When the keyword-recognition data changes, a new HTML file is created and reloaded on the HTML browser. A page is renewed by recreating and reloading the new file when receiving the different recognition data. The change of the layout is controlled smoothly without renewing the page by using Javascript functions.

4.2.3. Keyword card and word list. A square frame is used as an identification marker to detect the position of a keyword card. A keyword is printed in a rectangular frame and handwritten in a frame composed of grids, as shown in Figure 5 (a) and (b), respectively. It is required to define the grid layout as a character field beforehand for handwritten character recognition. The prototype system gives the frame seven grids placed in horizontal where each grid has the aspect ratio of 2.5 to 2.

(a) Rectangular frame of the printed keyword card

(b) Grid layout of the handwritten keyword card

Figure 5: Two kinds of frames for keyword cards.

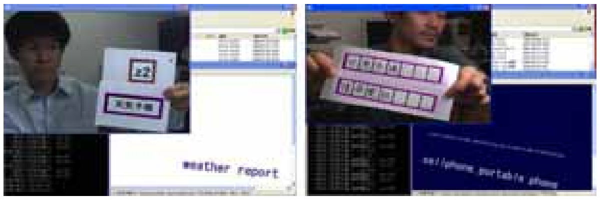
A word list is used as assistant to reliable character recognition, as described in the previous section. The word list is also used for language translation instead of a machine-translation engine. It works as a kind of database to designate the multimedia content such as images, 3D models, sounds as well as the translations to the relevant information. We originally created the word list based on daily conversations, which included roughly 6,000 words in Japanese, English and Thai, respectively.

5. Implementation

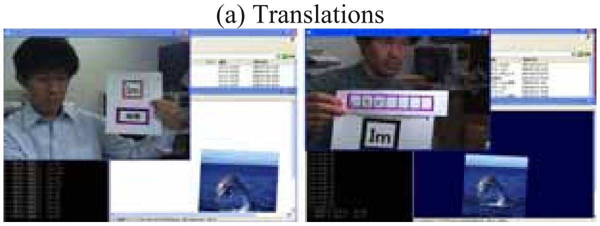
The prototype system was implemented on a desktop PC with a 2.4-GHz Pentium IV processor for both the image processing and content presentation components for demonstration. A DV camera (DCR-HC1000, Sony) was used to capture video images.

OCR middleware tools, Yonde!! KOKO (A. I. Soft) and Color OCR Library (Panasonic Solution Technologies), were applied as the recognition engines for printed and handwritten characters, respectively. ARToolkit [18], a set of open-source libraries, was used as the image-processing engine to detect the square frames in which characters were printed or handwritten, to derive the position, orientation, and size from the frame shape, and to recognize the pre-registered control commands.

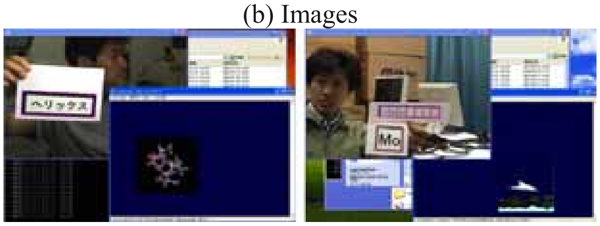
Figure 6 shows screenshots of displaying the content of the relevant information for the keyword presentations; translations in (a), images in (b), and 3D models in (c). The left and right screenshots indicate examples of the printed and handwritten keyword cards, respectively. Plug-in tools, Chime [19] and Cortona [20] were implemented for viewing the protein data from the Protein Data Bank and the VRML data of the 3D animations, respectively.



(a) Translations



(b) Images



(c) 3D models

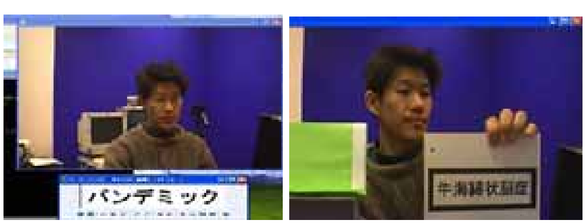
Figure 6: Example of the screenshots

6. Preliminary Experiment

The motivation for using the handheld keyword cards was to facilitate presentations at outdoor environments. Although the handheld cards would be useful for a speaker, it is not known how the handheld cards affect the comprehension of the audiences. We then conducted a preliminary experiment to investigate the properties of the keyword presentation styles, comparing stationary keyword captions and handheld keyword cards.

6.1. Method

The two presentation styles are compared in the experimental lectures through videoconferencing. Figure 7 shows an overview of each condition at the experiment.



(a) Keyword caption (b) Handheld card

Figure 7: An overview of conditions at the experiment.

6.1.1. Participants. Twelve students (10 males and 2 females) participated in the experiment, forming four groups in which three participants were included for each. They have had experience of a debate in class but haven't any videoconference.

6.1.2. Task. Participants were required to listen to talks about the two topics: (1) Pandemic (bird flu) and (2) BSE (mad cow disease). The talks on Pandemic and BSE were given using the stationary keyword captions and the handheld keyword cards, respectively. Each topic was presented for roughly twenty minutes, and was randomly selected at the first place.

6.1.3. Measurements. After the talks, the participants answered the questions at a preference test, as listed in Table 1. The questions mainly ask impressions at each presentation style, and also try to measure the difficulty of comprehension for each topic. Open-ended comments were obtained as complementary opinions. The appearance of the experiment was recorded on a video.

Table 1: Questions at the preference test.

No.	Question item
1	The talk is intelligible.
2	It is easy to memorize the outline.
3	It is easy to follow the talk.
4	The talk is excursive.
5	The keyword helps comprehension of the talk.
6	You keep feeling tension.
7	You keep concentration.
8	You grasp the key points of the talk.

6.2. Results

Figure 8 shows the average score for each question at the preference test. The white and black bars indicate the stationary keyword captions and the handheld keyword cards, respectively. The error bar is the standard deviation.

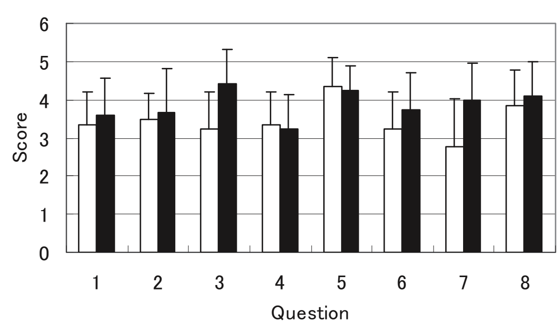


Figure 8: Score results of the preference test (white: keyword caption, black: handheld card).

There was no significant difference in the questions 1, 2, 4, 5, 6, and 8, but there was significant difference in the questions 3 and 7 ($p<0.05$). The style of handheld keyword cards had better scores than that of stationary keyword captions in both the questions 3 and 7 that asked traceability of the talk and durability of concentration during the talk, respectively.

6.3. Discussion

We here discuss the reasons why the handheld keyword cards were preferred to the stationary keyword captions in the traceability of the talk and the durability of concentration.

The major difference between the two styles is the position movements of the keywords in the display. The movements would attract visual attention to the keywords, working as an appropriate cue that the participants followed the talk at the handheld keyword cards. Some participants stated, “I was not aware that the keywords were changed at the stationary keyword captions”, which means that just putting stationary keyword captions may be insufficient as a cue for attention awareness.

Another difference between the two styles is the visual separation of the keywords from the speaker’s face. The visual separation was fairly small in the handheld keyword cards, which possibly allowed the participants to see the speaker and the keywords simultaneously. This may work toward keeping of eye contact, resulting in the retainment of concentration. Some participants stated, “I felt direct link between the facial expressions and the keywords”.

7. Conclusion and Future Work

In this paper, we described a videoconference support system using printed and handwritten keywords, in which a presenter holds to show keyword cards and the audiences at remote sites see the additional information with the keywords. A preliminary experiment was then performed to investigate the properties of keyword presentation by comparing the styles of stationary captions and handheld cards. The results showed that the handheld-card style was preferred to the stationary-caption style.

No demonstration was shown for distance lectures between outdoors and indoors. We plan to make them using our prototype system in practical situations, and will evaluate handheld-card-based telecommunications and usability of the system.

8. References

[1] K. Asai, and F. Saito, Support system for telecommunication using keywords, Proc. International Conference on Information, Communications, and Signal Processing, pp.1084-1088, 2005

[2] P. Baggett, and A. Ehrenfeucht, Encoding and retaining information in the visuals and verbals of an educational movie, Educational Communication and Technology Journal, vol. 31, pp.23-32, 1983

[3] S. D. Reese, Visual-verbal redundancy effects on television news learning, Journal of Broadcasting, vol. 28, pp. 79-87, 1984

[4] J. M. Clark, and A. Paivio, Dual coding theory and education, Educational Psychology Review, vol. 3, pp. 149-210, 1991

[5] C. E. Kirkland, E. M. Byrom, M. A. MacDougall, and M. D. Corcoran, The effectiveness of television captioning on comprehension and preference, Proc. Annual Meeting of the American Educational Research Association, pp.1-15, 1995

[6] H. G. Guillory, The effects of keyword captions to authentic French video on learner comprehension, Journal of the Computer Assisted Language Instruction Consortium, vol.15, pp.89-108, 1998

[7] J. R. Anderson, Cognitive psychology and its implications, Worth Publishers, 2005

[8] H. Itoh, An analysis of eye movements while watching educational TV programs, Bulletin of the National Institute of Multimedia Education, vol.5, pp.147-162, 1991

[9] Z. L. Berge, and M. P. Collins, Computer-Mediated Communications and the Online Classroom, Hampton Press Inc., 1995

[10] E. D. Mynatt, T. Igarashi, W. K. Edwards, and A. LaMarca, Flatland: new dimensions in office whiteboards, Proc. CHI 1999, pp.346-353, 1999

[11] K. Hiraki, T. Yonekura, and S. Shibusawa, “Web-Com”: interactive browser for web-based education, IEICE Trans. on Information & Systems, vol. E88-D, pp.912-918, 2005

[12] S. G. Deshpande, and J.-N. K. Hwang, A real-time interactive virtual classroom multimedia distance learning system, IEEE Trans. Multimedia, vol.3, pp.432-444, 2001

[13] B. Erol, J. J. Hull, and D.-S. Lee, Linking multimedia presentation with their symbolic source documents: algorithm and applications, Proc. ACM International Conference on Multimedia, pp.498-507, 2003

[14] A. Sellen, W. Buxton, and J. Arnott, Using spatial cues to improve videoconferencing, Proc. CHI 1992, pp.651-652, 1992.

[15] K. Okada, F. Maeda, Y. Ichikawa, and Y. Matsushita, Multiparty videoconferencing at virtual social distance: MAJIC design, Proc. CSCW 1994, pp.385-393, 1994

[16] D. Nguyen, and J. Canny, MultiView: spatially faithful group video conferencing, Proc. CHI 2005, pp.799-808, 2005.

[17] K. Tanaka, and K. Kondo, Configuration of inter-university satellite network (Space Collaboration System), IEICE Trans., J82-D-I, pp.581-588, 1999 (in Japanese)

[18] ARToolkit, HITLab, University of Wasington, http://www.hitl.washington.edu/artoolkit/

[19] Chime, MDL, http://www.mdl.com/

[20] Cortona 3D Viewer, Parallel Graphics, http://www.cortona3d.com/cortona/

Reporting Events in a multimedia content distribution and consumption system

Víctor Torres[1], Eva Rodríguez[2], Jaime Delgado[2]
[1]*Universitat Pompeu Fabra, Passeig de Circumval·lació, 8, 08003, Barcelona, Spain*
[2]*Universitat Politècnica de Catalunya, Jordi Girona, 1-3, 08034, Barcelona, Spain*
victor.torres@upf.edu, evar@ac.upc.edu, jaime.delgado@ac.upc.edu

Abstract

In multimedia content protection and management systems, it is important to keep track of the operations performed by the different actors involved in the value chain, from content creators to final users. These operations involve from the issuance of a specific digital license that restricts the actions that can be performed over a digital object, to the consumption of resources at the end user side. The collection and storage of reporting messages generated by the different actors of the multimedia value chain, or even by the system itself, can be used for tracking, statistics and even accounting purposes. In order to enable interoperability with other systems, it is important to follow, when possible, standard initiatives. In this paper, we are going to analyse the applicability of MPEG-21 Event Reporting to see if is suitable to deal with the requirements of a digital rights management (DRM) architecture, the DMAG-MIPAMS system. We will also briefly discuss if the standard specification is sufficient to provide integrity, privacy and non-repudiation for such reporting messages.

1. Introduction

Reporting Events in a multimedia content distribution and consumption scenario is a key issue because it enables a wide range of functionalities and provides additional value to the system, something that would not be possible without collecting and storing the reporting messages used to describe the events. By enabling Event Reporting functionality in a system, we can track the performed operations, compute statistics over them and even use them for accounting purposes or to prosecute the illegal usage of contents.

In this paper we will analyse the applicability of the Event Reporting [1] part of the MPEG-21 standard [2] and discuss its implications.

With this aim, the paper is organised as follows. First, we present some relevant parts of the MPEG-21 standard, to see how digital objects can be structured and to describe the standard definition of reporting messages in the MPEG-21 context. Then, we describe a distributed application environment where we need a reporting system and explain how to tackle the reporting messages structure from the MPEG-21 approach. After this, we analyse the deficiencies of the MPEG-21 standard in this sense and present our proposals. Next, we provide different usage scenarios to illustrate which kind of information would the reporting messages convey, depending on the circumstances, and we determine whether the MPEG-21 Event Reporting specification is suitable for the selected environment. Finally, we discuss some security aspects involving the reporting messages generation and storage and we describe our implementations in different contexts.

2. MPEG-21

The MPEG-21 standard [2] is divided into several parts, which deal with different aspects of multimedia information management. In the MPEG-21 context, the information is structured in Digital Items, which are the fundamental unit of distribution and transaction. Digital Items are digital documents written in XML according to a XML Schema [3]. A Digital Item is defined in [4] as a structured digital object, including a standard representation and identification, and metadata within the MPEG-21 framework.

2.2. Event Reporting

Event Reporting [1] is required within the MPEG-21 Multimedia Framework to provide a standardised means for sharing information about Events amongst Peers and Users. An Event, which can be defined as the occurrence of a reportable activity, is related to Digital Items and/or Peers that interact with them. In MPEG-21, the messages that report information about different aspects of media usage are called Event Reports.

Event Reporting could be useful when monitoring the usage of copyrighted material. The provider offering Digital Items for download would specify in an Event Report Request that, whenever a Resource within a Digital Item is rendered (e.g. played), he would receive an Event Report enabling him to manage his royalties. Upon rendering, the Peer would generate an MPEG-21 Event Report which would be delivered to the rights holder specified in an Event Report Request, containing information about the Digital Item, the Resource, and the conditions under which it would have been rendered.

Fundamentally, Event Reporting facilitates interoperability between consumers and creators, thereby enabling multimedia usage information to be both requested and represented in a normalised way. Examples where Event Reports may be requested include usage reports, copyright reports, financial reports and technical reports.

The basic model of Event Reporting indicates that the Events that need to be reported may be specified by interested parties through the use of an Event Report Request (ERR). An ERR is used to define the conditions under which an Event is deemed to have occurred. Events defined by ERRs trigger the creation of an associated Event Report (ER), which contains information describing the Event, as specified in the associated ERR.

The ER purpose is to indicate which Peer created it, define the data items that need to be included in such Event Reports, provide a reference to the originating ERR and provide status information regarding its completion and creation, along with a freeform description.

Although the MPEG-21 standard specifies the ERR format, it is worth noting that it is not normative that an ER is only created as the result of the processing of an ERR. This means that applications may create Event Reports which are normative on their own initiative.

Event Reports contain three main Elements. They are used to provide description of the Event Report, to contain the Event Report's payload and to optionally contain an Embedded Event Report Request.

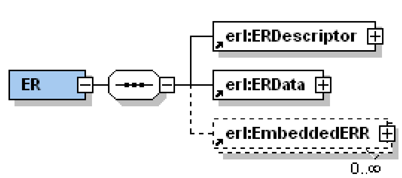


Figure 1. MPEG-21 Event Report

These three Elements, which are depicted in Figure 1 are, in more detail, the following:

• ERDescriptor: It contains information on whether the Peer was able to compliantly generate the Event Report as well as information regarding the creation of the Event Report, such as who has modified the ER, the ER completion status, the recipient of the ER and the source of the ER.

• ERData: It contains the "payload" data of the Event Report, which describes the performed action or operation, such as the peer that created the ER, the user that was using the peer, the time when it was generated, the operation performed by the user, metadata items related to the Digital Item, domain-specific data, etc.

• EmbeddedERR: It contains an ERR that has been included within and is associated with the ER.

Next, we are going to briefly describe the fields that are standardized for ERDescriptor and ERData fields, whereas in subsequent chapters we will see which are the implications of its usage in a multimedia content distribution and consumption environment.

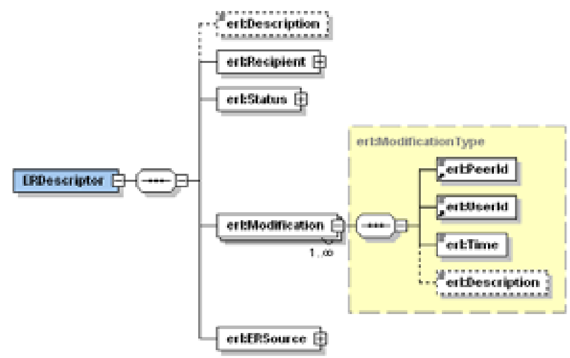


Figure 2. MPEG-21 Event Report Descriptor

3. Application Environment

In [5], we described an architecture to manage multimedia information taking into account digital rights management (DRM) and protection. The architecture, called DMAG Multimedia Information Protection and Management System (DMAG-MIPAMS), whose name is after our group acronym DMAG [6], consists of several modules or services, where each of them provides a subset of the whole system functionality needed for managing and protecting multimedia content. DMAG-MIPAMS is a service-oriented DRM platform and all its modules have been devised to be implemented using the web services approach, which provides flexibility and enables an easy deployment of the modules in a distributed environment, while keeping the functionality independent from the programming language and enabling interoperability.

All the modules provide a standard WSDL [7] description of the services they provide, enabling other authenticated components to use its services. The communication between modules is performed by means of SOAP messages [8].

DMAG-MIPAMS encompasses an important part of the whole content value chain, from content creation and distribution to its consumption by final users. The DMAG-MIPAMS architecture is depicted in Figure 5.

Next we are going to have a quick overview of its main functionalities to understand how Event Reporting can be used to satisfy the reporting requirements of the architecture. The following list briefly describes the modules of the architecture:

Content server. It provides the content that final users may request. It enables users to browse/select content, provides the content that final users may request to user applications, encodes and adds metadata to raw contents received from providers, and registers the objects describing resources (metadata), which can be stored independently from the resource itself.

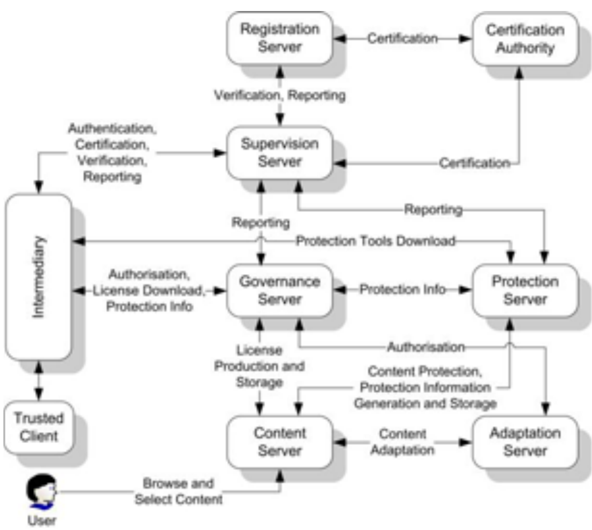


Figure 5. DMAG MIPAMS architecture.

Adaptation server. It performs the adaptation of content and its associated metadata, depending on transmission, storage and consumption constraints.

Protection server. It is responsible for protecting the content or digital objects, which become protected objects (for content server as well as for client production tools), describing the protection tools used to protect content and making the tools available for download. It also generates and optionally stores protection keys.

Governance server. The governance server includes several functionalities: license generation, license storage, authorisation and translation support. The authorisation functionality consists on searching in the license database for any license that authorises a user every time the user tries to perform an action over a resource, in order to see if the user is granted to do what he intends to do.

Certification server. It involves different functionalities: Certification Authority (CA), user registration, tool registration, authentication, certification of users and tools by calling the CA and verification of tools and devices.

Supervision server. It receives reporting messages that include information about different aspects of media usage or other events occurred in the system, and it stores them for future access.

More in detail, the reports collected by Supervision server are stored in a server database, available to be accessed by specific applications to keep track of what happens in the system by generating statistics and traces, or even for accounting purposes.

Within the depicted architecture, all the generated reporting messages are automatically created by the involved application, and are not the result of a report request, as explained in the MPEG-21 context. This solution is fully MPEG-21 compatible and simplifies the generation of this kind of reports.

In next section we explain the limitations of MPEG-21 Event Reporting [1] for being applied in the MIPAMS application environment and describe the proposals we have made in order to overcome them.

4. Distributed Generation of Events

MPEG-21 Event Reporting standard specification has defined the necessary mechanisms to identify the peers that have created and/or modified an Event Report. An Event Report can be generated by different parties located in different places of a system, as for example, on the client and server part. A typical example is the case where an end user asks for authorisation to a server. In this case, a possible solution is that the client application generates an initial Event Report where it states the action the user intends to do. In this case, the Governance server, responsible for the authorisation of the requested operation may modify the received client Event Report to add some information that can be determined only after the authorisation, as the authorisation result and/or the license identifier used to get that result.

For this purpose, the Modification element of the MPEG-21 Event Reports, which maintains the history of modifications of the ER, has been defined. This element consists of the PeerId, UserId, Time and Description elements. The PeerId element identifies the Peer that has created or modified the ER; the UserId element identifies the User that has created or modified the ER; the Time element specifies the date and time at which the ER was created of modified; and the Description element is a free form field to provide additional information. On the other hand, the reported information is a placeholder for inclusion of the reported data into an ER.

Therefore, in an Event Report, as defined in [1], we can specify the Peers that have created and modified the ER by means of different Modification elements, but we can not specify the fields or the data that each Peer has reported or modified in a structured way.

In this sense, we made a contribution to MPEG-21 ER, where we proposed three different solutions [9] to allow the specification of the data that each Peer has reported. Those three different solutions were evaluated at the 80th MPEG-21 meeting and one of them was selected for producing a corrigendum [10] of the MPEG-21 ER part, for which one of the authors is the co-editor.

The selected proposal consists on enabling the use of more than one ERData in an ER. In this way, it is possible to associate the Modification information, as the Peer or User that has created or modified the ER, with the specific data that this Peer or User has reported. For this purpose, we have defined the idData attribute for the ERData element to uniquely identify this element in an ER; and the idRefData attribute for the Modification element to reference to the data that a Peer or User has reported. Moreover, an important aspect of our solution is that it enables Users and Peers to digitally sign the specific data they report.

In next section we describe how the reporting messages can be protected to ensure confidentiality, integrity and non repudiation, while in section 5 we provide an example of Event Report which has been generated by different parties and digitally signed by each of them.

5. Protection of Events

An important issue when notifying events in Digital Rights Management systems is to ensure the integrity and authenticity of both the requested and reported data. On the other hand, in non-trusted systems it could also be necessary to protect fully or partially the data.

MPEG-21 Event Reporting standard specification [1] in its current version does not provide a way to ensure integrity and authenticity neither to Event Reports nor to the Event Report Requests. Moreover, the standard does not specify any mechanisms to encrypt partially or totally the requested and reported data although, initially, security requirements were defined for MPEG-21 Event Reporting. The security requirements for protecting data fields for ERs and ERRs and ensuring their integrity, authentication and confidentiality were defined in the MPEG-21 Event Reporting requirements document [11].

In [12] the authors propose a solution that was contributed to MPEG-21, which has been adopted to produce the Event Report in Figure 7.

6. Usage Scenarios

6.1. Content Consumption

The content consumption scenario refers to the case where an end user performs an action over an object that is governed, that is, which includes a digital license expressing the rights and conditions of usage of the related content.

A typical content consumption scenario could involve a user that wants to play a song or watch a movie. This action attempt would be done by using a user-side tool which includes a trusted module or intermediary. The trusted module could be an integral part of the tool or otherwise a plug-in for an already existing tool, which enables the consumption of digital objects that use the specific content and protection format of the system.

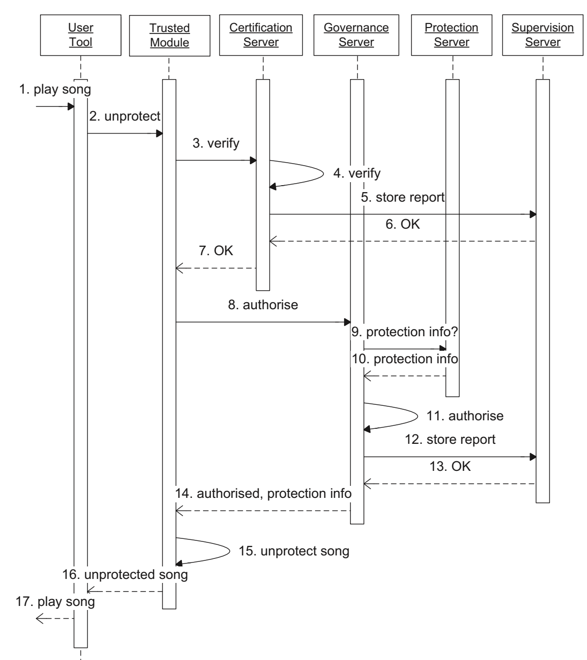


Figure 6. Content Consumption sequence diagram

The sequence diagram would be as depicted in Figure 6. In step 12 an event report is automatically generated and sent to the Supervisor Server, which collects and stores it in a specific database. The event report contents for this particular case would be the following:

ERDescriptor. The status field will be false before the operation is authorized in the Governance Server and true after the same server fills the LicenseId field.

ERData. DIOperation field will be filled with the action the user has performed over the content (e.g. “play”), whereas LicenseID field will contain the end User License identifier used to authorise the user to perform the requested operation.

Figure 7 shows a typical ER created during content consumption, which includes the User and the Governance server signatures, which could be used to prove that the ER was not modified after its generation and also to prove identity of each of the parties that have modified it.


```
<?xml version="1.0" encoding="UTF-8"?>
<ER xmlns="urn:mpeg:mpeg21:2005:01-ERL-NS" xmlns:mpeg7="…"
xmlns:dsig="…" xmlns:xsi="…" xsi:schemaLocation="…">
  <ERDescriptor>
    <!-- Recipient Identification: Supervision server ID -->
    <Recipient><PeerId>urn:mipams:SID:1l4e28ba-…</PeerId></Recipient>
    <Status value="true"/>
    <Modification idDataRef=”D001”>
      <!-- End User modification-->
      <PeerId>urn:mipams:TID:4bla0c2-…/PeerId>
      <UserId>urn:mipams:UID:3nkr14jg-…</UserId>
      <Time>2006-11-20T12:22:30</Time>
      <Description>PeerId, UserId, DII, DIOperation, Location …
      </Description>
    </Modification>
    <Modification idDataRef=”D002”>
      <!-- Governance server modification-->
      <PeerId>urn:mipams:GID:1jvrt5ba-…</PeerId>
      <UserId>urn:mipams:UID:3nkdsfjg-2idfs9-…</UserId>
      <Time>2006-11-20T12:22:32</Time>
      <Description>LicenseID</Description>
    </Modification>
    <!-- Source of the ER: User Tool Identifier -->
    <ERSource>
      <OtherSource>urn:mipams:TID:4bla0c2-ce3q-…</OtherSource>
    </ERSource>
  </ERDescriptor>
  <ERData idData="D001">
    <!-- Who, where and when-->
    <PeerId>urn:mipams:TID:4bla0c2-ce3q-…</PeerId>
    <UserId>urn:mipams:UID:3nkr14jg-2ih9-…</UserId>
    <Time>2006-11-20T12:22:30</Time>
    <Location><mpeg7:Region>es</mpeg7:Region></Location>
    <!-- Involved Object, Operation and others-->
    <DII>urn:mipams:OBJ:8b4fkt5a-2hq4-1gnm-8a2f-r9gnj157i8fb</DII>
    <DIOperation>REL:mx:Play</DIOperation>
    <ReportedDomainData>
      <Name>urn:mipams:DOM:19h6k2ba-…</Name>
      <ToolFingerprint>FhRuD1iGkUbej0fwBzT92Q==</ToolFingerprint>
    </ReportedDomainData>
    <ReportedDIMetadata>
      <CollectingSocietyId>urn:mipams:CSID:…</CollectingSocietyId>
      <CreatorId>urn:mipams:CID:6d4e28ga-2fai-...</CreatorId>
      <WorkId>urn:mipams:WID:8j4e2fca-…</WorkId>
      <DistributorId>urn:mipams:DID:…</DistributorId>
    </ReportedDIMetadata>
  <!-- User Digital Signature-->
  <dsig:Signature>
    <dsig:SignedInfo>
      <dsig:CanonicalizationMethod Algorithm="…"/>
      <dsig:SignatureMethod Algorithm="…"/>
      <dsig:Reference>
        <dsig:Transforms>
          <dsig:Transform Algorithm="urn:mpeg:mpeg21:2007:01-ER-
erTransform"/>
          <dsig:Transform Algorithm="urn:uddi-org:schemaCentricC14N:
2002-07-10"/>
        </dsig:Transforms>
        <dsig:DigestMethod Algorithm="..."/>
        <!-- ER Digest -->
        <dsig:DigestValue>...i4y39gfcDFgov7H=</dsig:DigestValue>
      </dsig:Reference>
    </dsig:SignedInfo>
    <dsig:SignatureValue>...91hII11btcIcmo8M=</dsig:SignatureValue>
    <dsig:KeyInfo>
      <dsig:KeyValue>
        <dsig:RSAKeyValue>
          <dsig:Modulus>0xR9lZdUEF0ThO4w==</dsig:Modulus>
          <dsig:Exponent>AQABAA==</dsig:Exponent>
        </dsig:RSAKeyValue></dsig:KeyValue>
     </dsig:KeyInfo>
</dsig:Signature>
  </ERData>
  <ERData idData="D002">
    <ReportedDomainData>
      <!-- End User license identifier used in the authorisation -->
      <LicenseId>urn:mipams:LID:2c4eg8sa-…</LicenseId>
      <!—Auth. success: TRUE = authorised; FALSE = not authorised -->
      <AuthorisationSuccess>TRUE</AuthorisationSuccess>
    </ReportedDomainData >
  <!-- Governance server digital signature-->
  <dsig:Signature>
    <dsig:SignedInfo>
      <dsig:CanonicalizationMethod Algorithm="…"/>
      <dsig:SignatureMethod Algorithm="…"/>
      <dsig:Reference>
        <dsig:Transforms>
          <dsig:Transform Algorithm="urn:mpeg:mpeg21:2007:01-ER-..."/>
          <dsig:Transform Algorithm="urn:uddi-org:schemaCentricC1..."/>
        </dsig:Transforms>
        <dsig:DigestMethod Algorithm="..."/>
        <!-- ER Digest -->
        <dsig:DigestValue>Q6bNcmGachtoi3z65gcfFDsaf6X=
        </dsig:DigestValue>
      </dsig:Reference>
    </dsig:SignedInfo>
    <dsig:SignatureValue>pRj0rxmx11E1hIIA1btcIcmo81=
    </dsig:SignatureValue>
     <dsig:KeyInfo>
      <dsig:KeyValue>
        <dsig:RSAKeyValue>
          <dsig:Modulus>0xQ8lXdARF0ThO4w==</dsig:Modulus>
          <dsig:Exponent>AQEBCC==</dsig:Exponent>
        </dsig:RSAKeyValue></dsig:KeyValue>
     </dsig:KeyInfo>
</dsig:Signature></ERData></ER>
```


Figure 7. ER Contents in Content Consumption scenario

6.2. End User License Generation

The end user license generation scenario refers to the case where a content distributor wants to create a license for an end user. An end user license grants an end user to consume a particular content according to some conditions. In order to be able create this kind of licenses, distributors need to have a distribution license that grants them to issue end user licenses.

A typical distribution license creation scenario could involve a content distributor that wants to sell a song to an end user. After verifying the license production tool, an authorisation would be performed on Governance server to determine if the user who requests the license is authorised by a distribution license to do so. If the authorisation is successful, the license is generated and a report is generated and sent to Supervision server. The Event Report content for this particular case would have the following features:

ERData. In this scenario DII, CreatorId, WorkId and DistributorId fields cannot be determined, as we are dealing only with licenses and not with Digital Items or Digital Objects. On the one hand, the DIOperation field will always be set to "issue", as this is the operation that is being performed. The corresponding LicenseId will refer to the distribution license that grants the requesting user to issue the end user license, whereas the CollectingSocietyId will be that of the Collecting Society involved in the license issuance process.

6.3. Other scenarios

As we have explained, the DMAG-MIPAMS architecture enables the certification of the user tools used in the client side in order to be able to work in the system, as well as their verification to ensure that they are still trusted during their whole life operation.

To keep track of successful and unsuccessful certifications and verifications of tools, reporting messages are also used.

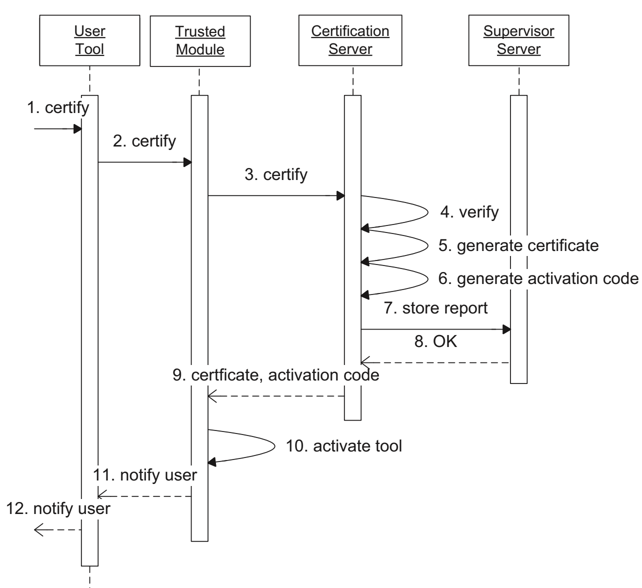


Figure 8. Tool Certification sequence diagram

The sequence diagram for the verification scenario is included in previous section (see step 5 in Figure 6). Regarding the tool certification case, the diagram would be as depicted in Figure 8. The Event Report content for this particular case would have the following features:

ERData. Regarding the DIOperation field, it will be set to something like "certification OK", "certification NOK", "verification OK" or "verification NOK", as this is the operation that is being performed. For the unsuccessful cases, an additional field could be then added to describe the reasons why the certification or verification of the tool failed.

7. Implementation

Some of the mentioned reporting mechanisms have been implemented in the context of the AXMEDIS project [13]. AXMEDIS (Automating Production of Cross Media Content for Multi-channel Distribution) is an FP6 IST integrated project co-funded by the European Commission. It started in September 2004 and is planned for 4 years.

The AXMEDIS project has defined and implemented an architecture that enables the creation, distribution and consumption of multimedia protected objects which are governed by DRM licenses. The AXMEDIS architecture has several common points with the MIPAMS architecture. It is also service-oriented and includes a client and server part. More details about its implementation can be found in [5].

Within the AXMEDIS architecture, the Event Reporting functionality is provided by the so-called Action Logs. Action Logs are the reporting messages that convey the information regarding content usage and other events in the AXMEDIS system, but customised with its particular names. Current implementation includes a reporting functionality provided as a call to a secure web service, with both client and server authentication, responsible for the collection of such reports. The service parameter consists of a data structure that includes each of the report fields, which are received inside a SOAP message. The report information is stored into a structured relational database used for accounting as well as for usage statistics purposes. Current implementation does not consider the signature or protection of Event Reports, as this was something not devised in the initial composition of the Event Reporting part of the MPEG-21 standard and because AXMEDIS is a trusted environment. However, as this is a desirable functionality, it is something that is being considered to be added to the architecture.

The AXMEDIS project has already developed several demonstrators for different delivery channels (PC distribution, satellite distribution, mobile distribution), which are being publicly presented and which have been very useful to validate the operation of the proposed modular architecture as well as the reporting functionalities within the framework. Specific events are organised periodically to make a public performance and present the AXMEDIS technologies.

On the other hand, we have developed an API and a demonstration software application devised to provide some of the security features we have described in this paper. The software consists of a java application with a friendly GUI. The application enables the user to load an existing Event Report and secure it in different manners. The first option is the ER signature, which can be performed over the full ER or only over the ERData part of the ER. The second available option is the ER encryption, whose result can be selected to be compatible with MPEG-21 IPMP or with XML Encryption recommendation [14]. The private key used to sign the ER is extracted from an external keystore, which can be provided by the user. The symmetric key used to encrypt the ER is randomly generated and protected with the public key extracted from the keystore. The next step will be to extend the software to enable the signature of ERs that have been generated in a distributed manner and need the signatures of each of the involved parties.

This software, which has been developed in the context of VISNET-II Network of Excellence [15] has been contributed to the 81st MPEG-21 meeting [16].

8. Conclusions

In this paper we have described how we can apply the MPEG-21 Event Reporting format to fulfil the reporting requirements of a multimedia information protection and management system. We have analysed which will be the content conveyed in such reporting messages in different scenarios, such as content consumption, license generation and other particular operations that are related with the system operation.

Up to this point, we have determined that the MPEG-21 Event Reporting standard specification is suitable up to a certain point to be used for those purposes, as we have shown that it is not devised to enable protection mechanisms. In this sense, we have provided a solution which was presented to the MPEG-21 standard.

Finally, we have introduced the AXMEDIS project, where an implementation of a part of the reporting functionalities has been developed, but where the authenticity, integrity and encryption functionalities would be desirable. We have also described the Event Reporting software we have developed and which has been contributed to the 81st MPEG-21 meeting [16]. The next steps to be tackled refer to implementation of the security measures in a distributed environment, as we have proposed in this paper.

Acknowledgements. This work has been partly supported by the Spanish administration (DRM-MM project, TSI 2005-05277) and is being developed within AXMEDIS [6], a European Integrated Project funded under the European Commission IST FP6 program, and VISNET-II [24], a European Network of Excellence in the same program.

9. REFERENCES

[1] ISO/IEC 21000-15:2006 – Event Reporting.
[2] MPEG 21, http://www.chiariglione.org/mpeg/standards/mpeg-21/mpeg-21.htm.
[3] W3C XML Schema, http://www.w3.org/XML/Schema.
[4] ISO/IEC 21000-2:2005 – Digital Item Declaration.
[5] Torres, V., Delgado, J., Llorente, S. An implementation of a trusted and secure DRM architecture. In On the Move to Meaningful Internet Systems 2006: OTM 2006 Workshops (IS'06) (Montpellier, France, Oct. 30 – Nov. 1, 2006). LNCS 4277. Springer-Verlag, Berlin Heidelberg New York, 2006, 312-321.
[6] Distributed Multimedia Applications Group (DMAG), http://dmag.upf.edu, http://research.ac.upc.edu/dmag
[7] W3C Web Services Description Language (WSDL) 1.1, http://www.w3.org/TR/wsdl/.
[8] W3C Simple Object Access Protocol (SOAP) 1.1, http://www.w3.org/TR/soap/.
[9] Rodríguez, E., Delgado, J., Torres, V. Some issues on the generation and modification of Event Reports in the MPEG-21 Event Reporting. ISO/IEC JTC1/SC29/WG11 MPEG2007/M14508. 2007.
[10] Timmerer, C., Delgado, J. ISO/IEC 21000-15:2006/DCOR 1 MPEG-21 Event Reporting. ISO/IEC JTC1/SC29/WG11 MPEG2007/N9118. 2007.
[11] Requirements for Event Reporting. ISO/IEC JTC1/SC29/WG11 MPEG2003/N6279. 2003.
[12] Rodríguez, E., Llorente, S., Delgado, J. Protecting Notification of Events in Multimedia Systems. In Proceedings of the 4th International Workshop on Security in Information Systems (WOSIS 2006) (Paphos, Cyprus, May 23-24, 2006). Insticc Press, 2006, 153-162.
[13] Automatic Production of Cross Media Content for Multichannel Distribution (AXMEDIS), IST 2004 511299, http://www.axmedis.org.
[14] XML Encryption, http://www.w3.org/TR/2002/REC-xmlenc-core-20021210/
[15] NETworked audioVISual media technologies (VISNET II), FP6-2005-IST-41, http://www.visnet-noe.org/.
[16] Rodríguez, E., Delgado, J., Sesmero, J. Software Implementation of Security in Event Reporting. ISO/IEC JTC1/SC29/WG11MPEG2005/M14623. 2007.

Distributed Interactive Multimedia for Technology-Enhanced Learning and Automated Content Production and Distribution

Bee Ong,[1] Kia Ng[1] and Pierfrancesco Bellini[2]
[1] *ICSRiM – University of Leeds, School of Computing & School of Music, Leeds LS2 9JT, UK*
kia@icsrim.org.uk, www.kcng.org, www.icsrim.org.uk
[2] *DISIT-DSI – University of Florence, Via S. Marta 3, 50139 Firenze, Italy*
pbellini@dsi.unifi.it, www.disit.dsi.unifi.it

Abstract

This paper brings together two major collaborative European projects involving several main themes of this special track of the conference, namely distributed multimedia systems, tools and systems for e-education and e-entertainment, and distribution methods and solutions for complex multimedia content. It briefly introduces the projects involves and provide a number of example results and application scenarios on the applications of distributed. The AXMEDIS project develops a framework and tools for automated cross-media content production and multi-channel distribution, and the I-MAESTRO project focuses on technology-enhanced learning for music with interactive multimedia and cooperative working environments.

1. Introduction

Digital media content related activities, e.g. digitisation, creation, production, distribution and many more, have major impacts to our daily life at all levels; works, education, infotainment, heritage preservation and others [3]. Increasingly, our requirements and applications of digital media content demand for amalgamations of many different medias that are interlinked and correlated in a form of multimodal representations and analysis to enable a further level of understanding and new possibilities and new applications.

This paper presents two multi-national collaborative projects that are partly supported by the EC under the 6th Framework Programme (FP6):

- AXMEDIS (www.axmedis.org) [1-2, 4, 6-7, 9] and
- I-MAESTRO (www.i-maestro.org) [5, 8, 10-11],

to bring together a number of important themes of this special track focusing on distributed multimedia systems and applications.

1.1. AXMEDIS EC IST Project

The AXMEDIS project (Automating Production of Cross Media Content for Multi-channel Distribution) creates the AXMEDIS framework which brings together innovative methods and tools to speed up and optimise content production and distribution, for leisure, entertainment and digital content valorisation and exploitation.

The AXMEDIS format can include any other digital formats [1]; it can exploit and enhance other formats such as MPEG-4, MPEG-7, MPEG-21, as well as other *de facto* standards. The AXMEDIS format offers the functionality for the combination of content components, and secure distribution, etc., supporting a large variety of DRM rules and models according to concepts of interoperability among DRM models.

1.2. I-MAESTRO EC IST Project

The I-MAESTRO [11] is a research and development project co- supported by the EC under the IST 6th Framework Programme. The I-MAESTRO project develops and utilises interactive multimedia environments for technology-enhanced music education. The project explores novel solutions for music training in both theory and performance.

Music performance is not simply playing the right note at the right time. The I-MAESTRO project studies and explores many aspects of music making in order to produce methods and tools for music education with innovative pedagogical paradigms, taking into account key factors such as expressivity, interactivity, gesture controllability and cooperative-work among participants.

2. AXMEDIS Framework and Tools

The AXMEDIS framework can be dichotomized into two main areas: the *Production area,* also referred to as AXMEDIS Factory, and the *Distribution area* (see Figure 1).

The AXMEDIS Factory [1] is responsible for the whole spectrum of content production such as Content Processing, Database, Editors and Viewers, and many other tools and functionalities. The *production* of AXMEDIS objects and content components is in connection with the AXMEDIS P2P tool (AXEPTool) for B2B distribution that supports DRM with a certification authority (AXMEDIS Certifier and Supervisor).

The Distribution area includes the distributors who deliver digital contents via different channels (e.g. internet, mobile, and satellite), the AXMEDIS P2P tool (AXEPTool) for B2B distribution, the AXMEDIS Players for end users, and so on. Both the production and the distribution areas are connected to the Protection and Supervising Tools that support digital rights management.

In AXMEDIS, the channel distributors can continue to use the same mechanisms for reaching the final users. In AXMEDIS, the content is distributed at B2B level by using the P2P tool (AXEPTool). Each AXMEDIS object may contain a single or a collection of digital resources to be delivered or shared, such as MPEG-4, MPEG formats, PDF, HTML, SVG, images, documents, videos, etc. This content can be adapted by using AXMEDIS compliant tools to reach some specific editorial formats and to satisfy the needs of the end user device and channel.

2.2. Summary

The AXMEDIS framework is capable of supporting data collection and translation from accessible CMSs, and automatically transforms legacy digital content into AXMEDIS objects. The framework consists of a wide range of processing modules for the production of the cross-media contents with the functionalities to preserve the security level along the whole value chain.

The AXMEDIS solution is mainly based on MPEG-21 model. It aims to stimulate the applications and exploitations of the new features by creating many AXMEDIS compliant tools and solutions and to allow these core aspects and solutions accessible in the form of an integrated AXMEDIS Framework.

AXMEDIS technologies is easily accessible offering a higher level of integrative capability with other industrial and standard allowing their exploitation in different production and distribution chains.

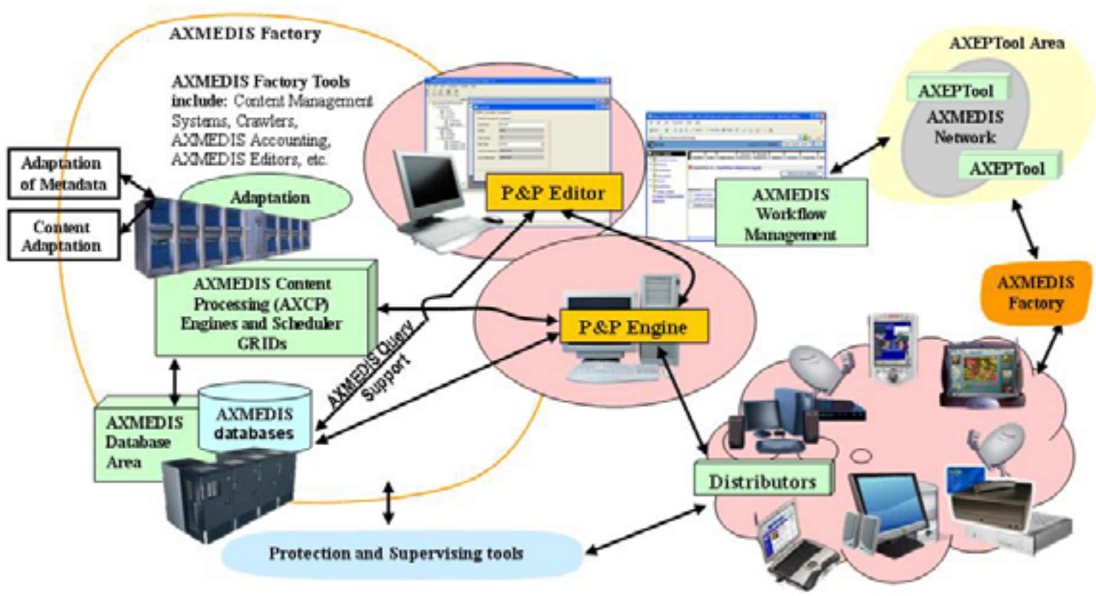


Figure 1. An overview of the AXMEDIS framework and tools

3. I-MAESTRO Framework and Tools

A framework such as I-MAESTRO can assist teachers in following the progress of each student and in evaluating their achievements with the support of innovative methods. Using the I-MAESTRO tools, the teachers will be able to create training material corresponding to the level of a student or of a group of students in a simple way, thus helping to personalise the tuition and supporting student's progress monitoring. In this context, I-MAESTRO provides for the creation of an almost continuous interaction between students and teachers, improving teaching continuity. The I-MAESTRO general architecture includes:

- I-MAESTRO School Server: an area for Lesson distribution, pedagogical material, history of students, training material, tools for the teachers, historical data collection and navigation, music school training management for the different scenarios.
- I-MAESTRO Client Tools: all I-MAESTRO Client Tools have a similar structure. Several different Client Tools will be studied and developed for students, teachers, conductors, impaired students, assessment experts, etc. They will support different devices: PC, Tablet and simpler PDA tools for sharing information and Lessons.
- I-MAESTRO Production Tools: for content production and organisation, include different tools like a Music Exercise Authoring Tool, Music Exercise Generators, etc.
- I-MAESTRO P2P & Cooperative Work Support: for sharing content, supporting the cooperative work in classroom and among students, stimulating collaboration and common experiences and setting up of groups of study, practice or theory training, ensemble training and playing, etc.

3.1. An I-MAESTRO Multimodal Tool for Gesture Support

This section presents one of the I-MAESTRO multimodal tools which focus on playing gesture support with interactive multimedia technologies involving 3D motion capture, audio, video, visualization and sonification.

Musicians often use mirrors to study themselves practicing. More recently we have seen the increased use of video recording in music education as a practice and teaching tool. For the purpose of analysing a performance these tools are not ideal since they offer a limited 2D perspective view of the performance.

Figure 2: The i-Maestro 3D Augmented Mirror.

Using 3D Motion capture technology it is possible to overcome this limitation by visualising the instrument and the performer in a 3D environment. Visualisation such as this can serve as a "3D Augmented Mirror" (AMIR) [8, 10-11] to help students improve their technique and develop self awareness. It assists teachers to identify and explain problems to students (see Figure 2).

By analysing and processing 3D motion capture data it is possible to obtain comprehensive information about a performance. For example, in the case of a string player, we can analyse the different characteristics of bow movements and the relationships between the bow, the instrument and the body. This information can be extremely useful for both student and teacher. It can be used in real time to provide instantaneous feedback about the performance and may also be used to make recordings for in-depth study after the performance has taken place.

3.2. Summary and Next Steps

This section discussed the I-MAESTRO project. With a general overview of the I-MAESTRO project and framework, we introduced one of the gesture-based analysis tools, to facilitate string instrument learning and teaching.

The outcomes of the project are being validated through a set of tools and methodologies including (i) tools and methods for music courseware production; (ii) interactive and creative tools for instrumental training in different environments such as classrooms and cooperative work, on different devices such as PC and Tablet-PC; and (iii) assessment and pedagogical models for music training.

4. Conclusions

In this paper, we have discussed the AXMEDIS framework designed to address key challenges of the digital domain, in the cross-media market, and the I-MAESTRO framework with one of the interactive multimedia tools which ultilised interactive multimedia technologies for learning and teaching for music.

This year, both the AXMEDIS and I-MAESTRO projects are reaching their completion year. Currently both the projects are actively working on validations to further improve and optimise their framework and tools and exploring potential exploitation activities for a set of products and services.

It is foreseeable that both these distributed multimedia projects can be developed further with common areas in content packaging, protection and distribution, both in term of automated processing, transcoding and also intelligent media evolutions, automated repurposing, and automated adaptations.

Latest developments and information about the projects, including upcoming events and activities, can be found online at the project website at www.axmedis.org for the AXMEDIS project and www.i-maestro.org for the I-MAESTRO project.

5. Acknowledgements

The authors would like to thank the EC IST FP6 for the partial funding of the AXMEDIS project, and to express gratitude to the project partners internal and external to the project including the Expert User Group and all affiliated members, for their interests, support and collaborations.

This paper also included researches supported in part by the European Commission under Contract IST-026883 I-MAESTRO. The authors would like to acknowledge the EC IST FP6 for the partial funding of the I-MAESTRO project (www.i-maestro.org), and to express gratitude to all I-MAESTRO project partners and participants, for their interests, contributions and collaborations.

6. References

[1] Bellini, P. and P. Nesi, An Architecture of Automating Production of Cross Media Content for Multi-channel Distribution, in Proc. of AXMEDIS 2005, Florence, Italy, 2005. pp. 123-133, ISBN 0-7695-2348-X, IEEE Computer Society.

[2] Delgado, J., Kia Ng, Paolo Nesi, Pierfrancesco Bellini (eds), Proc. of the 3rd Int Conf on Automating Production of Cross Media Content for Multi-channel Distribution Conference, Spain, IEEE Computer Society Press, ISBN 0-7695-3030-3, 2007.

[3] Ng, K. and Nesi, P. (eds), Interactive Multimedia Music Technologies , ISBN: 978-1-59904-150-6 (hardcover) 978-1-59904-152-0 (ebook), 394 pages, IGI Global, Information Science Reference, Library of Congress 2007023452, 2008.

[4] Ng, K., A. Badii, and Paolo Nesi (eds), Proc. of the 2nd Int Conf on Automating Production of Cross Media Content for Multi-channel Distribution Conference (AXMEDIS 2006), 13-15 December 2006, Leeds, UK, IEEE Computer Society Press, ISBN: 0-7695-2625-X, 2006.

[5] Ng, K., A. Badii, and Paolo Nesi (eds), Proc. of the 2nd International Conference on Automating Production of Cross Media Content for Multi-channel Distribution Conference (AXMEDIS 2006), Volume for Workshops, Tutorials, Applications and Industrial, 13-15 December 2006, Leeds, UK, Firenze University Press, ISBN: 88-8453-526-3

[6] Ng, K., A. Badii, Maulik Sailor, Bee Ong, Royce Neagle and Garry Quested, AXMEDIS Programme and Publication Tools Integration with Workflow-enabled Communication and Process Control, in Proc. of AXMEDIS 2006, Leeds, UK, IEEE Computer Society Press, Dec 2006.

[7] Ng, K., B. Ong, R. Neagle, P. Ebinger, M. Schmucker, I. Bruno, and P. Nesi, AXMEDIS Framework for Programme and Publication and On-demand Production, in Proc. of AXMEDIS2005, Florence, Italy, 2005. pp.247-250, ISBN 0-7695-2348-X, IEEE Computer Society.

[8] Ng, K., Oliver Larkin, Thijs Koerselman, Bee Ong, Diemo Schwarz, Frederic Bevilacqua, The 3D Augmented Mirror: Motion Analysis for String Practice Training, p. 53-56, in Proc. of ICMC 2007, Vol.II, pp. 27-31 August 2007, Copenhagen, Denmark, ISBN: 0-9713192-5-1

[9] Ng, K., Roberta Campo, Minh Thang Dang, Royce Neagle, and Bee Ong, AXMEDIS Programme and Publication Tools: A Case Study with Satellite Distribution, in Proc. of AXMEDIS2007, pp. 11-14, DOI 10.1109/AXMEDIS.2007, IEEE Computer Society Press, ISBN 0-7695-3030-3, Spain, 28-30 November 2007.

[10] Ng, K., Tillman Weyde, Oliver Larkin, Kerstin Neubarth, Thijs Koerselman, and Bee Ong, 3D Augmented Mirror: A Multimodal Interface for String Instrument Learning and Teaching with Gesture Support, in Proc. of the 9th Int Conf on Multimodal interfaces, Nagoya, Japan, pp. 339-345, ISBN: 978-1-59593-817-6, ACM, SIGCHI, DOI: http://doi.acm.org/10.1145/1322192.1322252, 2007

[11] Ng, K. (ed), Proceedings of the 4th i-Maestro Workshop on Technology-Enhanced Music Education, co-located with the 8th International Conference on New Interfaces for Musical Expression (NIME 2008), Genova, Italy, ISBN: 978 0 85316 269 8, 4 June 2008. www.i-maestro.org/workshop/

Performance evaluation and statistical analysis of a BitTorrent tracker

Pierfrancesco Bellini, Ivan Bruno, Daniele Cenni, Paolo Nesi
Distributed Systems and Internet Technology Lab,
Department of Systems and Informatics, University of Florence, Florence, Italy
{nesi, cenni, pbellini, ivanb}@dsi.unifi.it, http://www.axmedis.org

Abstract

The requests of efficient solutions for sharing and distributing digital content are constantly increasing. In this paper, a performance evaluation and analysis of a BitTorrent tracker are presented. The analysis has been performed by using a BitTorrent protocol and architecture developed for content sharing among business actors and consumers. The solution improved the BitTorrent technology obtaining a system with high efficiency, scalability and fault tolerance useful to serve P2P requests, and that can be used to implement a large set of different strategies of content distribution. The work has been developed for the AXMEDIS Integrated Project IST FP6 (Automating Production of Cross Media Content for Multi-channel Distribution) of the European Commission.

1. Introduction

The P2P-based content distribution is rapidly increasing its diffusion. The success of P2P solutions is mainly due to the efficiency in sharing files (see Napster, Kazaa, Gnutella, Emule, etc.). P2P networks can be centralized or decentralized.

Among the many P2P system solutions proposed, the BitTorrent is one of the most efficient for content download. For these reasons is the most attractive for exchanging large files. The work presented in this paper reports the method to perform the performance analysis of a BitTorrent solution. The analysis reported in this paper has been performed on the BitTorrent P2P solution elaborated for AXMEDIS (Automating Production of Cross Media Content for Multi-channel Distribution) Integrated Project IST FP6 of the European Commission (http://www.axmedis.org) [1], [2]. The AXMEDIS consortium consists of European digital content producers, integrators, aggregators, and distributors, together with information technology companies and research groups. The AXMEDIS P2P is a scalable, fault tolerant and fast BitTorrent based solution for the distribution of cross media content, MPEG-21 with DRM, and of any files. AXMEDIS P2P integrates BitTorrent and GRID technologies to obtain a solution that can be used for implementing a large set of different strategies for content distribution, publishing, promoting, monitoring, etc.

The paper is organized as follows. Section 2 reports the overview of AXMEDIS P2P network. In Section 3, the performance analysis model used and the results are reported. The analysis has permitted to identify the limits of the AXTrack and of the clients tools used as control nodes of the AXMEDIS P2P BitTorrent network. Conclusions are drawn in section 4.

2. AXMEDIS P2P Architecture

The AXMEDIS P2P architecture is based on BitTorrent protocol and can be used for setting up P2P networks among Business and Consumers actors or for a mix of them. The classical BitTorrent solutions do not address a set of functional requirements which are addressed by the AXMEDIS solution:

- integration of content protection support and DRM, so that to control the distribution and sharing of non certified content;
- querying/indexing of content on the basis of the metadata and related object/file cataloguing and querying;
- control of the P2P network allowing:
    - removal of content/files from the network, removing them at least from the Tracker;
    - notification of new files and thus controlled seeding files among the network;
    - automating the publishing and downloading files by integration of the P2P network facilities with the content production facilities;
- monitoring activities and user behavior on Nodes and on Tracker.

The AXMEDIS P2P solution is based on the following elements, as depicted in Figure 1.

AXTrack: is a modified BitTorrent Tracker that manages the AXMEDIS P2P network. Main features are: tracking of BitTorrent files; catalog of tracked files; creating statistics. In the global geographic system, many AXTracks may provide services and share each other information regarding nodes and content files, so that to distribute the workload and making it fault tolerant. Thanks to the AXTrack, the AXMEDIS P2P clients (AXEPTool/AXMEDIA) are able to know where they can retrieve the segments constituting a file.

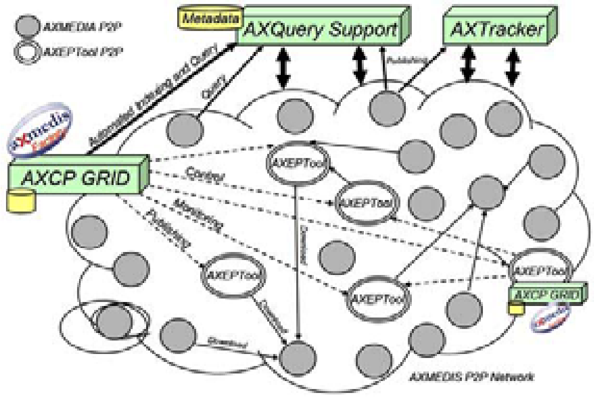


Fig. 1 – AXMEDIS P2P Architecture

AXEPTool is a P2P BitTorrent Client Node suitable for producers, distributors, integrators, etc. It accepts automated requests of publishing, download, monitoring, etc., from AXMEDIS GRID nodes in the AXMEDIS Content Production/Distribution Factory [3]. These facilities permit the P2P network controllability. In addition, the AXEPTool accepts also manual commands for publishing, downloading and monitoring via a graphical user interface.

AXMEDIA is a P2P BitTorrent Client Node for final users content sharing.

AXCP GRID is an instance of the GRID tools [3] to automate publishing, downloading, monitoring and control the P2P Network calling one or several the AXEPTools via their Web Services.

AXQuery Support is a server on which the user and the AXCP GRID may perform queries. Queries can be in a complex format and include classification, identification and legal aspects. Once a new content is published on the network by a P2P Network node or by an AXCP GRID, the metadata are passed to the AXQuery Support.

3. Performance analysis of BitTorrent

In this section, a method for measuring the performance of BitTorrent trackers and thus of a P2P networks is presented. The main goal of the performance assessment of P2P solutions is to identify its limits in real conditions.

In this context, a BitTorrent P2P network is a set of nodes working with a Tracker identified by its unique URL. Behind the unique Tracker URL, a cluster of servers can be arranged to share the workload. Some P2P clients/nodes may refer to multiple trackers but the capabilities of a BitTorrent P2P network depends on the capabilities of the Tracker behind the unique URL. In fact, the size (in terms of nodes, users and files) and thus the costs of the BitTorrent P2P network depend on the specific implementation and capabilities of the single Tracker server, and on the network bandwidth. The work performed by the trackers consists in responding to periodic requests coming from each active file (called peer, using the standard BitTorrent terminology it means to have a peer for each open instance of a digital resource/file opened on a client) on each client tool in the P2P network. The tracker receives/provides from/to the P2P client updated information about the swarm and the file seeding status, and keeps overall statistics about the torrent. Each client contacts the tracker with a GET request through HTTP sendin a query string with some parameters (info_hash, peer_id, port, uploaded, downloaded, left, event etc…). The information provided to the clients (tracker request parameters) consists of a dictionary with several key fields (failure reason, warning message, interval, min interval, tracker id, complete etc…). The tracker response to the clients consists of text/plain document consisting of a bencoded dictionary with several key fields (failure reason, warning message, interval, tracker id, peers etc…).

3.1 Performance Estimation Scenario

The performances reported in this paper have been estimated by using: an AXTrack P2P tracker server, HP Proliant ML310 G4, Linux Server, a set of AXEPTool P2P clients, Intel® Pentium® D. The publication and download of the tests performed were automated using the AXCP Rules [3]. The AXCP Rule Scheduler scheduled the processes executed by AXCP nodes to create the P2P traffic, via the AXEPTools control nodes of the P2P network. Each of them managed in seeding and/or download some thousands of files. The AXEPTool settings used to test the tracker efficiency and performance include an announce time of 60 s.

3.2 Maximum workload per AXTrack

The goal of this work was to stress the AXTrack to find its current performance limits that may guarantee the responsiveness with respect to download requests. To this end, a specific BitTorrent network has been set up and loaded for collecting statistical information about its status in different conditions.

During the experiments, data were taken every 60 s, to reduce at minimum the standard deviation of collected results. All the tests were performed around stable workload conditions in terms of peers on the network and objects into the database, then 1 hour of data collecting for each test around that point. During the measures the maximum load reached by the AXTrack machine was about 160000 peers. The AXTrack catalogue was loaded with 19000 objects. The idea of the performance analysis is to estimate the limit of a single tracker to estimate the number of trackers needed to set up large networks of millions of peers.

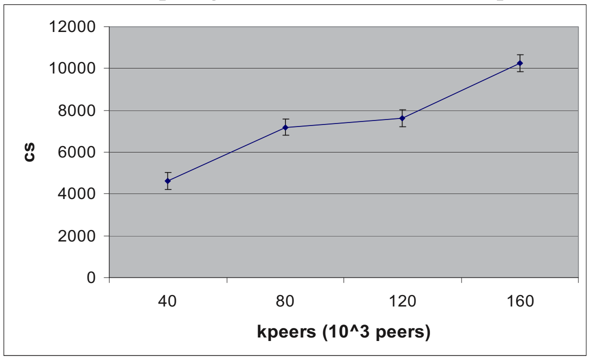


Fig. 2 – AXTrack context switches per second

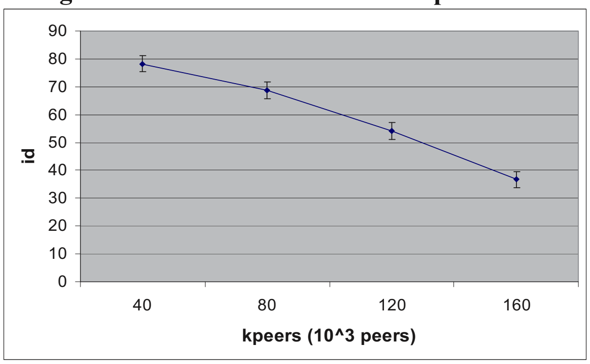


Fig. 3 – AXTrack time spent idle

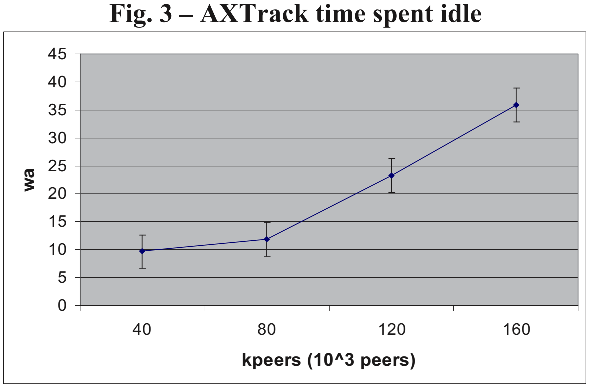


Fig. 4 – AXTrack time spent waiting for I/O

With a limit of 160000, a network of 1,6 Millions of peers can be reached with at a cluster of about 10 trackers. The scalability is not perfectly linear but the models allow estimating the magnitude with a very good approximation. According to the statistical analysis, the tracker under observation reached its limits at about 160000 since its CPU idle time (see Fig.3) was close to the limits and the time spent in waiting was rapidly growing (see Fig. 4), as the number of context switches per second (see Fig 2). Around 160 Kpeers, the AXMEDIS P2P network and AXTrack presented still a good responsiveness. In Table 1, a summary of the main AXTrack system values is presented; for each of them some statistical indicators are provided (average, standard deviation, median and confidence interval at 95%), demonstrating the congruence and system stability. The average memory consumption on the AXTrack machine was about 966 MB.

Field	average	std.dev.	median	Con.95%
r	6.2	12.1079	0	3.063688
b	1.5833	0.90743	2	0.22961
swpd	132	0	132	-
free	143691	5090.71	145644	1288.105
buff	60589.3	132.872	60568	33.6208
cache	3162157	4998.648	3160320	1264.809
in	1502.76	64.27943	1495	16.2635
cs	10263.6	659.8095	10211	166.9518
us	21.95	1.731317	22	0.438076
sy	5.36667	0.485961	5	0.122963
idle	36.6667	3.559026	37	0.900541
wait	35.8667	3.217888	35	0.814223

Table 1 – AXTrack system values, 160000 peers

Where:

- *r*, number of processes waiting for run time;
- *b*, number of processes in uninterruptible sleep;
- *swpd*, the amount of virtual memory used;
- *free*, the amount of idle memory;
- *buff*, the amount of memory used as buffers;
- *cache*, the amount of memory used as cache;
- *in*, the number of interrupts per second, including the clock;
- *cs*, the number of context switches per second;
- *us,* the time spent running non-kernel code. (user time, including nice time);
- *sy,* the time spent running kernel code. (system time);
- *idle*, time spent idle;
- *wait*, time spent waiting for I/O;

After that point (e.g., at 180000 Peers) the values of responsiveness of data collected were unacceptably high and caused the AXTrack's instability with a variance (>15%). At 180000 peers the idle time was under 25% with a standard deviation of about 20%.

3.4 Maximum workload per AXEPTool

During the experiments each AXEPTool client was loaded with max 10000 objects. The average CPU load was about 50% on a P4, with about 850 MB of RAM used by the JVM. The AXEPTool responsiveness was tested during the experiments, by putting a new object in download, and thus measuring the corresponding time to start the new transfer (see Fig.5).

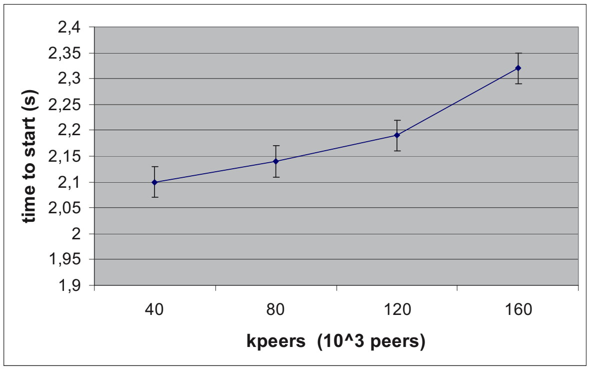


Fig. 5 – AXEPTool time to start a download (s)

The responsiveness of the AXTrack and the download speed of the clients (see Fig.6) had no significant changes in different conditions of the load. The time to start at 180 Kpeer dramatically goes to values greater than 1000 times leading the system to condition of no service. For this reason, we decided to limit the workload to 160 Kpeer that is about 10% under the limit. Given the measured variance a more accurate estimation of the workload limits have no sense.

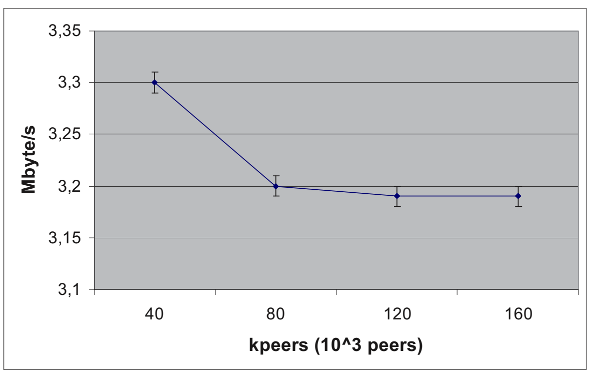


Fig. 6 – AXEPTool download speed (Mbyte/s)

On the basis of the above data, we can move some tuning parameters. By increasing the announce time by a factor of 3.125 would mean to reach a 500000 peers limit. On the other hand, increasing too much the timeout would mean a higher latency in updating the BitTorrent swarm information. The largest BitTorrent tracker (e.g., Piratebay) features about 8 Millions peers, so a cluster of 16 servers would give to the AXMEDIS P2P Network quite the same strength as the most popular BitTorrent tracker.

3.5 Network performances

The average inbound traffic on the Ethernet interface of the AXTrack was somehow limited at each sampled point, varying from 1.52 Mbit/s to 2.89 Mbps. The average outbound traffic was in the range 1.18-2.29 Mbps.

4. Conclusions and future tasks

In this paper, a performance evaluation and statistical analysis of BitTorrent solutions has been described. The estimation model has been used to asses the AXMEDIS P2P network and tracker: AXTrack. The presented results demonstrate the AXTrack efficiency, robustness and scalability, and also the network controllability. On the basis of this analysis it is possible to define the size of a cluster of AXTracks on the basis of the P2P workload. AXMEDIS solution has been obtained by improving the BitTorrent protocol and tools, maintaining compatibility; integrating the P2P networking with those of the AXMEDIS Content Processing facilities, to allow setting up a large range of strategies for content distribution, seeding, publishing, promoting, etc. The main benefits of the proposed solution and architecture are scalability, content sharing control, efficient rules scheduling, fast object publishing and dissemination, and in general network controllability without constraining the actors of the network.

The AXMEDIS P2P network has been set up since the April 2007, and can be joined installing an AXEPTool Node which can be freely downloaded from http://www.axmedis.org.

5. Acknowledgments

The authors would like to thank all the AXMEDIS project partners (BBC, TISCALI, XIM, TEO, ANSC, AFI, EUTELSAT, Giunti ILABS, HP Italy, MBI, FHGIGD, DIPITA, ACIT, CPR, EXITECH, University of Leeds, University of Cataluña, University of Reading, etc.), the Expert-User-Group and all affiliated members for their contributions, supports and collaborations. The authors would also like to express their thanks to the European Commission IST DG-INFSO FP6 for partially funding of AXMEDIS.

6. References

[1] AXMEDIS, "Framework and Tools Specifications", http://www.axmedis.org

[2] Bellini, P., Bruno, I., Cenni, D., Nesi, P., Rogai, D., "P2P Architecture for Automated B2B Cross Media Content Distribution", Automated Production of Cross Media Content for Multi-Channel Distribution, 28-30 Nov. 2007

[3] Bellini P., Nesi P., "An architecture of Automating Production of Cross Media Content for Multi-channel Distribution", Proc. of the first Int. Conf. on Automated Production of Cross Media Content for Multi-channel Distrib., 30/11-2/12 2005, Florence, Italy, IEEE press.

High Resolution Digital Media Personalisation for All Delivery Context Device Agents

Atta Badii*, Ali Khan*, Adedayo Adetoye* and David Fuschi*
* Intelligent Media Systems and Services Research Centre
School of Systems Engineering
University of Reading
Reading RG6 6AY UK
atta.badii@reading.ac.uk; http://www.imss.reading.ac.uk

Abstract— **Multimedia content distribution, whether on-demand or scheduled, becomes complicated when performed on a mass-scale involving various channels with distinct and dynamic network characteristics, and a variety of terminal devices offering a wide range of capabilities. It is practically impossible and uneconomical to pre-package various static versions of the same content to match all the varying demand parameters of clients for various contexts. In this paper we present a profile-based media adaptation and transcoding approach for on-demand as well as scheduled dynamically personalised media content delivery integrated with the AXMEDIS Framework. The client profiles comprise the representation of User, Terminal, Network and Natural Environment of content delivery based on MPEG-21:DIA. The framework described here is based on Gecode [24], a Generic Constraint Development Environment, along with complementing AXMEDIS (www.axmedis.org) Content Processing Engine (AXCP) adaptation rules. The framework is deployed for distribution, by the AXMEDIS Programme and Publication module, through various channels e.g. satellite and Internet, to a range of client terminals e.g. desktops, kiosks, IPtv etc.**

I. Introduction

Personalisation in terms of exploration of the possibility on the part of the service provider to make efforts for achieving (ideally) absolute fulfillment of the customer or user's personal preferences is strongly related to customer satisfaction of (e-)services. Recent advances in profiling and personalisation techniques, e.g. as deployed in collaborative filtering and recommender systems, have contributed to the capability of online media distribution systems to provide personalised digital contents to suit various user's needs each on an individualised basis.

Modern content providers have to retrieve, assemble and deliver information from a variety of sources with respect to specific user's needs onto a variety of devices [1]. Thus the user can request any content, in any format across any distribution channel. In order to provide highly personalised services, such a framework should have the means of collecting and storing information about 1) the user's content-specific preferences, 2) the standard range of capabilities of the user's device on which the content will be delivered, 3) the user's customisation preferences of the standard capabilities of the target user's device, 4) the characteristics of the data network and/or distribution channel that will be used to deliver the content, 5) the tools that are to be used to personalise the content etc.

In this paper we address the issue of dynamic profile management and its use in content adaptation for personalised multimedia delivery within the AXMEDIS Framework. The ontological engineering of the profiling schemas and dynamic profile value selection and resolution in AXMEDIS have been detailed in our previous papers [15, 16]. Briefly, for each subset of the profile we need to provide: the profiling "maintenance overhead" for the on-demand, or other scenarios as follow: 1) (Near) real-time (synchronous and asynchronous, direct/indirect) means of collection of each profile element value, 2) Storage of each profile element value, 3) Access to / Dynamic update of each profile element value, and 4) Dynamic Resolution of the profile element value.

II. Adaptation in AXMEDIS

The AXMEDIS platform [5] supports cross-channel distribution, including that involving Mobiles, Satellite, Internet and kiosks, for distribution of content in both push (e.g. multicast) and pull (on-demand) modes [9]. Hence the AXMEDIS Distribution channels must enable dynamic access to, structuring, update, storage, and integration of personal, device and delivery context profile element values, using the AXMEDIS Content Processing Engine (AXCP) [4]. Personalisation requirements can be seen in practice to demand a capability for profiling management, coupled with a requirement for dynamic collection and evaluation of personal, default and closure values of profile elements depending on the profile element values availability - Profiling Resolution.

The seventh part of ISO/IEC 21000 (MPEG- 21) [5] specifies tools for the adaptation of Digital Items. More specifically, it proposes a set of normalised tools describing the usage environment of a digital item to command adaptation tools. According to such precedence, AXMEDIS-compliant devices should support content adaptation conformant with a number of MPEG-21 usage environment tools. In order to fulfill these dynamic media adaptation requirements, we have envisaged the requirements for five different kinds of profile elements derived from MPEG-21 DIA [5], which when deployed in combination enable enhanced personalised services. These profiles can be broadly categorised as follows:

1. User Profiles: This profile type represents the salient personal information about the user and his preferences,

e.g. presentation preferences, auditory or visual impairment, etc.

2. Device Profiles: This profile represents the salient information about the user's device(s) including both hardware and software specific information, e.g. codecs, formats, I/O, etc.

3. Network Profiles: This profile represents the salient information about the distribution channel that the service provider must use for the delivery of the content, e.g. QoS related elements, bandwidth, error rate, security etc.

4. Context Profiles: This profile represents the salient delivery context information about the user and his usage environment.

5. Business User Profiles: This profile is the Distributor's or prosumer's (producer-consumer) profile to support B2B media transaction scenarios, e.g. licenses acquired, etc. This paper focuses on the key profile types i.e. the first three profiles as they are derived from MPEG21-DIA. The elements for Context Profile are covered in the Users Profile and Network Profile, while the Business User Profile can be seen as subsume-able within the User Profile in the generalised sense of the Users as including all buyer/seller or "prosumers" of digital media.

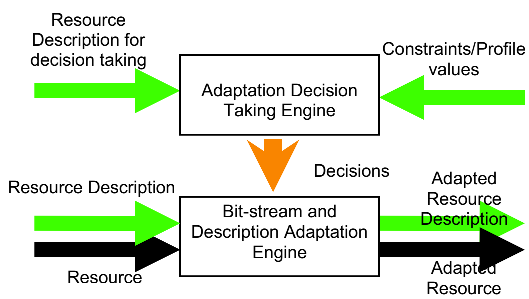


Figure 1: Constituents of Adaptation Engine (reproduced from [18])

III. Media Selection, Bundling, Transcoding & Delivery

The requested content is transcoded based on the information contained in the Device Profile and the Network Profile. For Transcoding, the user's device specific information for processing the content has to be considered, e.g. supported codecs, display resolution etc. The transcoded content is adapted to match the play-display quality capabilities of the actual target device and its operational needs as expressed in the device profile. Once the adaptation has been achieved the new content can be sent to the end-user with the related usage license. All adaptation should be able to use a managed mix of cached and on-the-fly adaptation responsively depending on the requirements for the particular device, the mix of business models (offer types), and, distribution regimes (on-demand, and/or multi-cast) that may be operating. Finally the transcoded product is then dispatched to the device using an appropriate channel distribution server.

A. Personalisation Resolution

We distinguish three personalisation related modalities in terms of challenges and complexity for real-time delivery:

1. Recommendation Systems
2. Mass Personalisation Systems
3. High resolution individualised Personalisation

1) Recommendation Systems

Media Selection for personalised matchmaking requires the delivery context information as available plus relevant promotional offers, current media popularity charts etc. For this we can deploy a collaborative filtering (CF), or Recommender System (RS) based on social networks profiling that derives recommendations using the behaviour of a group of users, especially those that have displayed similar tastes and interests in the past [10].

Most of the Recommender Systems to-date have focused mainly on algorithms that drive the system and not so much on the design issues from the user's perspective for example lack of disposable quality time and attention for choice-making [11]. Such information as group preferences can be used to modify the menu flow in order to ensure that the user's navigation and product selection is more optimised. The Recommender can promote content to the users and/or suggest a narrow cross-section of content to accelerate navigation and enhance choice making based on the client capabilities, users' history and that of other users from the same interest groups. Hence the Recommender Systems that are deployed by distributors should prioritise design features that support accelerated navigation focused on matchmaking between presently available content and what the users need to purchase immediately; but also allow users to explore and evolve their tastes for the future [12]. User privacy has to be considered for some mechanisms [13]. For more information on state-of-the-art in CF and RS refer to [18, 19, 20]. For security-privacy context resolution deploying virtualisation of user transaction contexts so as to distinguish between the different privacy and security concerns of the user relevant to different transaction domains, the reader is referred to the MobiPETS domain security-privacy contexts virtualisation paradigm [17].

2) Mass Personalistaion Systems

This can be a more generalised deployment of Recommender Systems whereby delivery of media and/or services is adapted to whole social classes, regions, and nationalities based on assumed shared preferences/values e.g. linguistic, cultural, socio-economic grouping, professional, gender, age, dietary etc, rather than only on the basis of relatively narrow interests that characterise the preferences of the members of some online social networks.

3) High Resolution Individualised Personalistaion

The needs and preferences of a user that may arise from the user's context or environment, the technical requirements of the user's device, the tools available (e.g., assistive technologies), the user's background, or a disability in the traditional sense, need to be considered while making decisions for matchmaking of media delivery to the user and dynamic media adaptation for individualised accessibility, in general. Such highly individualised matchmaking has to rely heavily on the delivery context information specified in the user, terminal (device), network and natural environment profiles.

Thus individualised media selection and adaptation follows the explicit needs and preferences of the particular user. The profile processing Decision Engine adjusts the user interface or configuration of the learning environment, locates needed resources and adapts the resources to match them to the needs and preferences of a user. Optimal accessibility requires optimal transcoding which in turn requires both the parameters of the source data as well as the characteristics set out in the profiles of the user and his/her natural environment, the network and the user client device to which it is to be delivered. All this must be know-able otherwise informed choice making about media type and appropriate decision rules regarding correct conversion (transcoding) for delivery to the client will be impossible: 1) Relevant prototypical profiles (for user's media adaptation needs and preferences, client device and network), 2) Target Media Resource Parameters (information about resource e.g. size, font etc.), 3) Target delivery-context-sensitive Adaptation Guidelines, 4) Fully operational Transcoding plug-ins for various media types e.g. audio, text, video etc., and 5) Situational Delivery Context Over-rides (environmentally-triggered exceptions).

In case of AXMEDIS, the Adaptation Engine takes as input: an AXMEDIS object (constituent resources), resource descriptions and constraints i.e. profile values. The adaptation process performs necessary activities to arrive at adapted constituent resources (collectively an adapted AXMEDIS object) and the adapted resources descriptions. The input sets and output sets of the DIA Decision Engine can be further characterised as

$$X = f\,(\text{resource descriptor values, UP, NP, DP}).$$

Where UP is User Profile, NP is Network Profile, DP is Device Profile and X is the target set of adaptation-actionable (aa) sub-sets of the respective profile values; as

$$X = (\text{aa-UP, aa-NP, aa-DP}).$$

Essentially the adaptation decision making needs to identify the intersection of all the relevant constraint sets through identifying those elements of the UP, DP and NP that have to be the decision driver, i.e. logically the prevailing criterion in deciding the value of each aspect of the way the content has to be rendered on the target screen. This amounts to an attempt to find the intersection of four sets namely Resource Description, UP, DP, NP and presentation template (for compound SMIL objects) as illustrated in figure 2.

IV. AXMEDIS MPEG-21 DIA Profile Management

We consider the problem of extracting an adaptation strategy from the client profiles for optimal and contextual content adaptation in the AXMEDIS framework and propose a technique to its automation. In a typical scenario, we are presented with a digital resource which must be adapted to a target AXMEDIS object that must conform to a set of AXMEDIS profile specifications. These specifications can be considered as constraints guiding the adaptation process. Hence, the systematic selection of a solution space of feasible adaptations is amenable to the standard artificial intelligence technique of constraint satisfaction [25, 26]. The problem can be cast to a general Constraint Satisfaction Problem (CSP), which provides us with generic tools to automate the discovery of solution spaces.

A. Constraint Modelling in AXMEDIS

The key step in transforming the AXMEDIS resource adaptation problem to a form which can be fed to, and automated by a CSP framework is the development of a model of the AXMEDIS world as a constraint satisfaction problem. The AXMEDIS object models and profiles are specified as XML schemas which circumscribe the variables X that are used within the AXMEDIS framework. The domains D of these variables may also be extracted from these schemas. So the first stage of the analysis involves the selection from the XML schema specifications, the relevant variables that will be needed for our analyses as well as defining their domains of definition. Once the modelling has been completed, rules and constraints specifying relationship among the identified variables can be specified. These can continue to be added incrementally to the system as more requirements arise and better heuristics are discovered. These rules may also be removed to relax constraints. Solution to the CSP problem is given as output that describes sets of target/adapted resource valid for the given AXMEDIS resource and profiles.

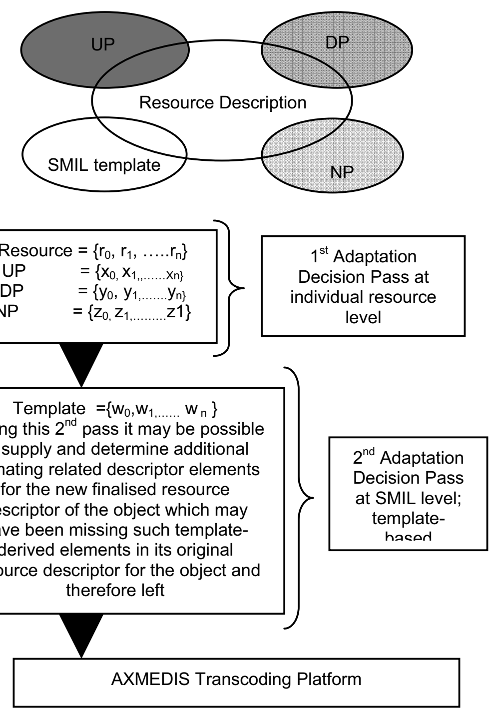


Figure 2: AXMEDIS DIA Engine Decision Making

B. Illustration

Suppose an uncropped image is to be displayed on mobile phone screen with size $l \times w$ (measured in inches). This device can display images at different resolutions r taken from the set {r1,...,rn} in pixels per inch (assuming that the horizontal and vertical resolutions are the same) and supports several image formats such as JPEG, GIF, BMP; and supports data connectivity modes over WIFI, Satellite, GPRS. These information are gathered from the AXMEDIS device profile. Furthermore, depending on the network used, let us assume that the cost to this user varies because the various network service providers charge based on the available bandwidth and duration of data transfer. Thus, a goal might be to minimise delivery time and consumer cost for this service with minimal compromise on image quality by choosing an optimal network and image format combination. This may also involve the selection of an image format with efficient compression capability to minimise the size of the image to be transferred across the network. All these are constraints to the solution space, which must guide the selection of an appropriate adaptation strategy.

1) Modelling the example

The necessary parameters to our constraint satisfaction problem include the target screen dimensions ($l \times w$), the resolution r in $\{r_1,...,r_n\}$ to use, the image format f in {JPEG, BMP, GIF ...}, the network type n {WIFI, Satellite, GPRS …} supported by the target device, for example. Let us assume that the file size (s) of the adapted image needs to be calculated. For simplicity, we assume that there is a constant, k_f, associated with each image format f describing the compression capability of that format; that is, a factor which translates the number of pixels to be encoded to a file size in that format. Hence, we compute the file size of the image (under the format f) as follows:

$$s_f = l \times w \times r^2 \times k_f$$

Since the network service providers charge by download size and the duration of download, we need to compute, in addition to the size of the image to be sent over the network, the duration (d_n) that it takes to send the file over the network n. This, again for simplicity, we assume is a linear function of the file size calculated as

$$d_n = (d'_n \times s) / b_n$$

where d'_n is the network-related data transfer overheads for sending file of size s over network n and b_n is the associated data-rate (that is, "bandwidth") of the network. The parameters d'_n and b_n are derived or extracted from the profile of network. Finally, let us assume that the cost c_n of transferring a file of size s within a duration of d_n over n is computed as

$$c_n = c'_n \times s \times d_n,$$

where c'_n is some pricing factor for the network n.

TABLE I.
A SAMPLE MODEL AND ITS SOLUTION

```
typedef DIM[1:10]; l, w : DIM;
bandwidth[1:10, 20:30]; size[1:50]; resolution[1:10];
data_rate[3:5, 20]; cost[100:500];
size = l * w * resolution * resolution * 1;
data_rate * bandwidth = 3 * size;
cost = 3*size*data_rate;


Solution:
No.  ({l, w, bandw, size, resolution, cost, data_rate})
1.      {1, 1, 25, 25, 5, 225, 3}
…….
{5, 3, 9, 15, 1, 225, 5}
```

In this example, we are just interested in checking what combination of parameters are possible, that is, the solution to this system without any additional constraints. As highlighted earlier, the goal might be to minimise or maximise the parameters identified above for cost efficiency or quality of service purposes. This could then be added incrementally to the system as additional constraints and is supported by our framework.

The script shown in Table 1 below is a model of the example presented earlier where various constants have been assigned values and domains have been defined and assigned to the various model parameters. This illustrates the flexibility of the framework and the declarative nature

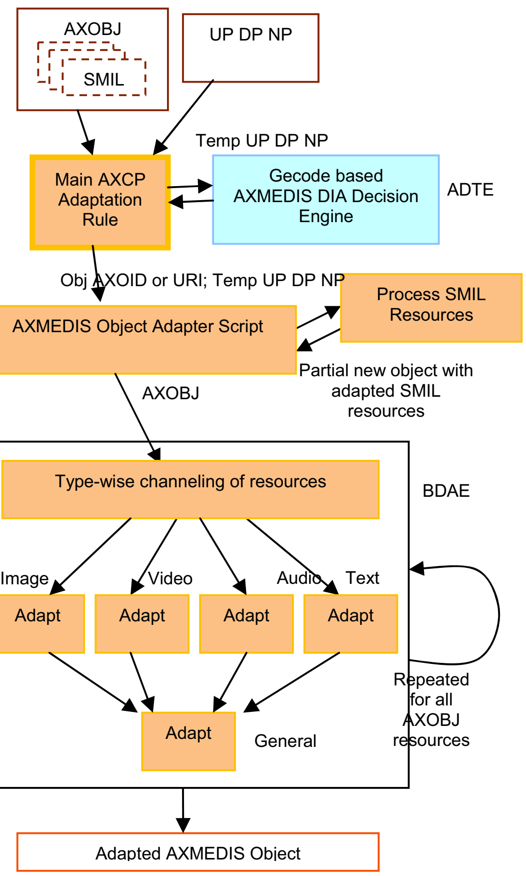


Figure 3: AXMEDIS Adaptation rule at work

of the solution.

C. *The AXMEDIS Constraint Solving Engine*

We have developed a powerful constraint solving system for decision-making in the AXMEDIS adaptation process. This engine uses Gecode [24] internally for solving constraints. The key aspect of this is a modelling language, which we have developed for defining domains and expressing constraints in the AXMEDIS system. The key idea is that parameters in our adaptation strategy such as user preferences, device profiles and possible target adaptations of an AXMEDIS resource can be used in the development of a model which is passed to the engine in a fairly high-level language. This model is automatically solved and the solutions, if any, form the basis of our decision for a particular resource adaptation strategy.

The engine does semi-symbolic computation, based on a model written in a very flexible and expressive AXMEDIS modelling language to find the set of solutions to an AXMEDIS constraint problem. The result is a solution object which can be queried for possible solutions to be used in the adaptation stage. This interface to the engine can accept modelling data from arbitrary input streams such as files, URLs and so on.

The AXMEDIS constraint solving engine performs sub-symbolic analysis by using a model specification written in a fairly high-level modelling language that we have developed. This model is semantically analysed to check whether it conforms to the grammar of the AXMEDIS modelling language and the Gecode object representations that are generated provide appropriate expression for the problem to be consequently solved. The AXMEDIS model Lexical and Grammar analysers are based respectively on GNU Flex and Bison.

V. AXMEDIS MPEG-21 DIA ADAPTATION ENGINE

The media adaptation engine comprises of various scripts, each with the assigned role of dealing with a specific media type. They invoke the relevant transcoding algorithms for adaptation of the AXMEDIS object's constituent resources using the adaptation strategy provided by the AXMEDIS Constraint Solving Engine to arrive at an adapted AXMEDIS object as output.

The AXMEDIS MPEG-21 DIA Adaptation Engine adapts media types namely image, video, audio and text as well SMIL. For all these media type, the DIA Adaptation Decision Engine incorporates the general media type independent component of adaptation as well as the media type specific component of the adaptation.

In the following sub-sections, we outline the features and the modus operandi of the Adaptation Decision Engine in respect of each modality.

There are two main constituents of the Adaptation Engine [18:

1. ADTE: Adaptation Decision Taking Engine
2. BDAE: Bit-stream and Description Adaptation Engine

In AXMEDIS, we are able to extract resources from objects and treat them in the adaptation engine one by one. This allows seamless adaptation of complex objects that may comprise of multiple resources of the types image, video, text etc. integrated and presentably formatted using SMIL templates.

In case of video and image resources, some adaptation features of the AXMEDIS Decision Engine include: 1) Acquisition and analysis of user preferences pertaining to brightness, saturation and contrast, in addition to the user's information regarding visual impairments (if any), as well as resource information, 2) Adaptation of resource for optimal presentation on target device, 3) Screen size – resizing the image if necessary, 4) Visual impairments – factored in adaptation process, 5) Device capability – colour capable, image capable etc.

In case of Audio resources, these steps include 1) Resource Information Acquisition e.g., sampling frequency, bit rate, number of channels etc, 2) Acquisition of user preferences and terminal capabilities to ascertain optimal audio adaptation parametric values e.g. adapting stereo content to mono if target device can only support mono, 3) Change the audio sampling rate, bit rate, number of channels.

For Text resource, the following steps are taken: 1) Resource Information Acquisition, 2) Convert to format compatible (if not already) with the receiving terminal, 3) Converting text resources to image using ImageMagick [22] OR MikTeX [23].

SMIL resources in the AXMEDIS object require necessary regional and positional adaptation. The Adaptation Engine carries out such modifications by 1) Adapting region, text, image and video features e.g. size and position, 2) Adapt text clickable area size and position, 3) Adapt image clickable area size and position, etc. The Adaptation Engine also determines if device and network limitations in relation to the object overall size would have adverse effects on the quality of service, and whether delivery is feasible.

The AXMEDIS DIA Decision Engine has been tested with various complex AXMEDIS Objects using device profiles for mobile devices such as Samsung i320, Motorola Z3 and imate JASJAR. Some results are shown in figures 4 and 5.

Figure 4: AXOBJECT with image, audio and text resources

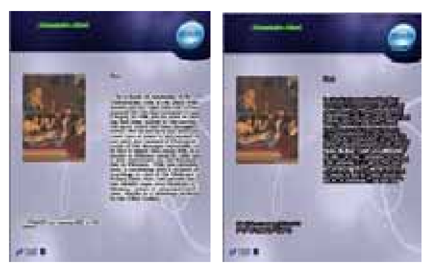


Figure 5: Object adapted for Motorola Z3. a) Text to image conversion using MikTeX, b) Text to image conversion using ImageMagick

VI. CONCLUSION

We have discussed the main aspects of dynamic profiling management and content adaptation for personalised multimedia distribution as deployed in AXMEDIS. In order to manage the complexity of delivering personalised content on-demand, we employed the usage of profiles for Users, User's Devices Capabilities and the Network Capabilities. These profiles are based on the MPEG-21 DIA which meets the requirements for content distribution, in order to make available the media files on various devices across various channels. A decision engine based on the Gecode [24] constraint satisfaction library has been developed to develop an adaptation strategy in the context of these profiles. This is complemented by an adaptation engine which employs the various AXMEDIS adaptation plug-ins to automatically transcode the content as and when required.

ACKNOWLEDGMENT

The authors would like to thank the EC IST FP6 for the partial funding of the AXMEDIS project [3], and all AXMEDIS project partners including the Expert User Group and all affiliated members, for their interest, support and collaboration.

REFERENCES

[1] W. Kießling, W. T. Balke, M. Wagner, Personalized Content Syndication in a Preference World. In Proc. EnCKompass 2001), Eindhoven, 06/01.

[2] AXMEDIS Framework for All, http://www.axmedis.org/tiki/index.php, April, 2006.

[3] The AXMEDIS Consortium http://www.axmedis.org/ last accessed April 2006.

[4] P. Bellini, I. Bruno, P. Nesi, A Distributed Environment for Automatic Multimedia Content Production based on GRID, in Proc. AXMEDIS2005

[5] ISO MPEG-21, Part 7 - Digital Item Adaptation, ISO/IEC JTC1/SC29/WG11/N5231, (Oct 2002)

[6] W3C, CC/PP Information Page http://www.w3.org/Mobile/CCPP, accessed, April06

[7] UAProf Schema and Information available at http://www.openmobilealliance.org/tech/profiles/index.html

[8] Resource Description Framework, http://www.w3.org/RDF/, accessed April 2006.

[9] K. Ng, B. Ong, R. Neagle, P. Ebinger, M. Schmucker, I. Bruno, P. Nesi, AXMEDIS Framework for Programme and Publication and On- Demand Production, in Proc. AXMEDIS2005.

[10] D. Pemberton, T. Rodden, and R. Procter., Groupmark: A WWW recommender system combining collaborative and information filtering in Proc. 6th ERCIM Workshop on "UserInterfaces for All", ERCIM October2000, no 12 in, page 13.

[11] Badii A, Context, Context, Wherefore Art Thou? "C-Assuring" shared sense-making and meta-scaffolding for Collaborative Creative Communities Doing Society with IS, Proc.UKAIS-AIS 2000.

[12] K. Swearingen and S. Rashmi, Interaction Design for Recommender Systems, in Designing Interactive Systems 2002. ACM, 2002.

[13] J. Canny, Collaborative filtering with privacy via factor analysis. Proc. 25th Int. ACM SIGIR Conference Information Retrieval, 2002.

[14] Badii A et al Profiling Management for Personalised Multimedia Delivery On-Demand within the AXMEDIS Framework, Proceedings of AXMEDIS 2006 Conference, Dec 2006, University of Leeds, UK.

[15] Kia Ng, Badii A, et al Programme and Publication Tools Integration with workflow-enabled communication and process control Proceedings of AXMEDIS 2006 Conference, Dec 2006, Leeds, UK.

[16] Badii A, Hoffman M, Heider J, MobiPETS-GRID, context-aware mobile service provisioning framework deploying enhanced Personalisation, Privacy and Security Technologies (.PETS), to appear at the 1st IEEE international conference on Communication Systems Software and Middleware (COMSWARE), New Dehli, Jan 2006

[17] D Mukherjee (2004) MPEG-21 DIA: Objectives and Concepts, HP Labs. Lecture slides as part of ECE 289J – Multimedia Networking at UC Davis taught by Prof. Mihaela van der Schaar

[18] MPEG-21 DIA http://www.chiariglione.org/mpeg/tutorials/technologies/mp21-dia/index.htm

[19] WY Lum, FCM Lau (2002) A Context-Aware Decision Engine for Content Adaptation, 2002 IEEE, University of Hong Kong

[20] Mukherjee, D., Kuo, G., Hsiang, S., Liu, S, Said, A. (2004). Format-Independent Scalable Bit-Stream Adaptation Using MPEG-21 DIA. International Conference on Image Processing 2004. 0-7803-8554-3/04 02004 IEEE.

[21] ImageMagick. Available online. WWW:< http://www.imagemagick.org/script/index.php>

[22] MikTeX. Available online. WWW:< http://miktex.org/>

[23] Gecode. Generic Constraint Development Environment. Available online. WWW:<www.gecode.org>

[24] R. Dechter. Constraint Processing. Morgan Kaufmann, 2003.

[25] F. Rossi, P. van Beek, and T. Walsh, editors. The Handbook of Constraint Programming. Elsevier, 2006.

Unifying Quality Standards to Capture Architectural Knowledge for Web Services Domain

Francisca Losavio, Alfredo Matteo, Roberto Rahamut
MoST Laboratory, School of Computer Science, Faculty of Science,
Universidad Central de Venezuela, Caracas, Venezuela
flosav, almatteo, rrahamut {@cantv.net}

Abstract

The development of applications based on WS (Web Services) considers client requirements, where elicitation, analysis and specification are main concerns to ensure the WS quality as a software component and the overall quality of the application using the WS. Specification of requirements, quality requirements and their translation into architectural elements, are still open research issues. The main goal of this paper is to unify three standard frameworks (WSA (Web Services Architecture) base-line of the W3C (World Wide Web Consortium), the standards ISO/IEC 9126-1 to specify the domain quality properties and ISO/IEC 13236 to specify the QoS metrics), to capture architectural knowledge for the WS domain. Besides these frameworks, a WS categorization based on functionality is also considered. Since these standards have been formulated independently, unified guidelines are established to integrate them for practical use within a domain characterization process. A major benefit of this domain characterization based on the unified knowledge on quality standards, is the automatic generation of standard contractual bases between WS clients and providers. Moreover the guidelines provided for traceability among the standards contribute to the establishment of a common language for the WS community. Finally, the applicability of our approach is illustrated for transactional WS.

1. Introduction

Web Services (WS) are reusable software components accessible over the standard Internet Protocol (IP), allowing *interoperability* and *portability* in Web applications, according to a Service Oriented Architecture (SOA) style. A WS is considered a software component offering a service, i.e. providing a kind of functionality, characterized by its functional and non functional requirements. Different types of WS can be grouped on the basis of the functionality or service they provide. WS-based software development considers component-based programming platforms, providing modular and reusable structures, independent from the technological details, such as operating system and data format [14, 23]; however, within this scheme, contractual issues have to be established and respected. In what follows, the main concepts on the *Web Services Architecture* (WSA) [11], ISO/IEC 9126-1 [21] and ISO/IEC 13236 [22] commercial and international standard frameworks respectively, that will be integrated to characterize most of the architectural knowledge for WS domain, are presented.

1.1 The WSA reference architecture

The architectural styles underlying WS applications are a layered SOA style instance for components, and a message-passing style, following a Peer-to-Peer pattern [12], for connectors. A reference architecture has been defined by the Web Services Architecture Working Group (WSARCH) to guarantee the interoperability of WS applications. It will be called the WSA base-line architecture in this context. According to the W3C (WWW consortium), a WS is a software system designed to support interoperable machine-to-machine interaction over a network. It has an interface described in a machine-processable format, specifically WSDL (Web Service Description Language). Other systems interact with the WS in a manner prescribed by its interface description using SOAP (Simple Object Access Protocol) messages, typically conveyed by HTTP with an XML serialization in conjunction with other Web-related standards. The main principles to construct WS are based on three aspects or mechanisms, *location*, *description* and *call*, and follow the XML universal data format [15, 23]. The WSA requirements document of the W3C describes seven critical top-level quality goals (see Table 1) that are the minimal set of requirements for a common architecture that a WS application should comply [11].

Table 1.W3C requirements for WSA Reference Architecture.

WSA Requirements	Description
Interoperability	The capability of WS to interoperate within different environments.
Reliability	The availability and stability of WS.
WWW Integration	WS must be consistent with the WWW evolution.
Security	WS must provide a reliable environment to perform its online processes.
Scalability and Extensibility	WS must allow flexible applications in the sense of adaptation to changing volume of information or to the addition/remotion of new components.
Team Goals	WS must meet the needs of the user community.
Management and Provisioning	WS must provide maintenance facility, allowing the adaptability to a changing environment.

1.2 The ISO/IEC 9126-1 standard

The achievement of quality software products is the goal of software engineering. The quality properties related to functional and non functional requirements are called quality requirements and are specified in the standard as quality characteristics and sub-characteristics (see Table 2). They are classified into three views of quality for a software artifact or product: *external quality* (considered within a testing environment), *internal quality* (considered within the development process) or *in use* *quality* (considered by the final user). ISO/IEC 9126-1 defines a hierarchical quality model based on six main high level quality characteristics. They are refined into sub-characteristics, continuing with the refinement until attaining the attributes or measurable elements. A sub-characteristic can then be associated to several attributes and metrics, units and/or counting rules can be associated to each attribute; this process is part of the so called measurement model.

Table 2. The ISI/IEC 9126-1 quality model for external/internal views: quality characteristics and sub-characteristics [21].

Quality Characteristics	**Quality Sub-characteristics**
Functionality	Suitability, Accuracy, Interoperability, Security, Compliance
Reliability	Maturity, Fault tolerance, Recoverability, Compliance Availability is a combination of the other first three sub-characteristics.
Usability	Understandability, Learnability, Operability, Attractiveness, Compliance.
Efficiency	Time behavior, Resource utilization, Compliance.
Maintainability	Analyzability, Changeability, Stability, Testability, Compliance .
Portability	Adaptability, Installability, Co-existence, Replaceability, Compliance.

1.3 The ISO/IEC 13236 standard

The goal of the ISO/IEC 13236 standard is to assist in the specification and design of technology-based software systems; it describes how to characterize, specify and manage requirements related with the *quality of the service* (OoS). It provides a common language to services, clients and providers. According to International Telecommunication Union (ITU) [7], *QoS* is defined as a set of quality requirements present in the collective behavior of one or more objects. Note also that a requirement originates from a client entity that uses a service and it is translated to different QoS requirements, expressed as *parameters*. A *mechanism* is realized by the entity to satisfy one or more QoS parameters. Mechanisms are part of the *management functions* and parameters are part of the *context* of a QoS. The range of values for the attributes (value of the QoS parameters) is established by the metrics for the WS quality requirements external view.

The main goal of this work is to unify the mentioned market and international quality-based standards to be reused as architectural knowledge in the characterization of families of WS applications, in the sense of Berard [16]. Domain analysis is a usual practice in the product-line approach of architecture design, favoring reuse, and the use of standards will ease communication and understanding. A process is defined which on one hand establishes a correspondence between the WSA top-level critical goals of the W3C and the ISO/IEC 9126-1 [21] standard for software product quality, to facilitate at a high abstraction level, common understanding of the architectural properties among the stakeholders. This unified standard can be easily adapted to specify architectural properties, considering the architecture as an intermediate product of the software development process [2, 20, 28]. Moreover, the quality requirements associated with the functionality of each WS type, defining sub-families of applications, are also specified using the ISO/IEC 9126-1 quality model [21]. On the other hand, the ISO/IEC 13236 standard [22] for the QoS, is used to define a measurement model by specifying the measurable attributes, according to general QoS metrics. In consequence, a standard quality model, incorporating all these standards, is defined for the WS application domain.

This paper is structured as follows, besides the introduction and conclusion: section 2 presents the related works. Section 3 describes the process to characterize the WS application domain. Finally, in section 4 the process is applied to a transactional WS case study.

2. Related Works

In general, standards are used in manufacturing, but not much has been done at early stages of software development. Moreover, the use of standards by practitioners is not easy, since they often lack of guidelines or rationale. We have successfully used the ISO 9126-1 standard to specify architectural properties and found it a useful tool, even if lack of flexibility is claimed. We favor the fact that is important to speak a common language [8, 16, 17]. The goal-oriented approaches to build architectures from requirements could easily use such standards, however they do not [1, 27]. To characterize early the crosscutting concerns, some attempts have been made that use a similar approach [19, 24]. The characterization of architectural knowledge for a family of applications is part of the domain analysis discipline in domain engineering [17, 20, 28], being crucial for architectural design and in particular for product line architectures, Moreover, the handling of general domain knowledge is recently gaining importance to reduce the gap between the modeling of business and system requirements; it is the starting point to define a first rough architecture for the system, where the specification of business requirements is taken into account. The ISO/IEC 9126-1 quality model has demonstrated to be a useful tool at this stage. In the context of families of web applications, the architectural domain knowledge is used to automatically generate quality contractual basis between WS clients and providers. In this sense, several languages for specifying the SLA (Service Level Agreement) have been proposed, most notably, a toolkit to generate the contract file in XML format is publicly available on the IBM alphaWorks™ Web site [5]. This utility is part of the IBM Emerging Technologies Tool Kit and serves as a guideline to learn how to specify the SLA using Web Service Level Agreement (WSLA) language. This toolkit helps to define the contractual agreement, but only semantically, it lacks to provide the standard default QoS value or range for a specific WS . SLAng [29] is another approach, which is an XML-based language that describes QoS properties to include in SLAs. SLAng does not focus only on web service interactions, but also to specify SLAs for hosting service provisioning (between container and component providers), communication service provisioning (between container and network service providers) and so on. Although SLAng is expressive enough to represent the QoS parameters included in SLA, more work is needed on the definition of its semantics. Web Service Offering Language (WSOL) [30] focuses on web service interactions. Another proposal is the Web Services Agreement Specification [31], which is also an XML-based document containing descriptions of the functional and non functional properties of a WS application. It consists of two main components that are the agreement Context and the agreement Terms. WS-Agreement [31] is an industry based protocol for the establishment of service level agreements and is being adopted widely, but, it lacks a precise definition of the meaning of its constructs, because it does not support explicitly the negotiation of the agreement, there is no monitoring of how close a term is to being violated at execution time, and, the breaking of one single term of a running agreement results in termination, while a more graceful degradation would be desirable.

3. CHARACTERIZATION OF THE WS DOMAIN

For a standard quality-based characterization of the domain of WS applications, the following process is proposed:

Input: problem statement, requirements taxonomy including business rules

(i) Define functionality. Establish a classification of WSs, according to functionality.

(ii) *Define Quality Model*. Specify the quality requirements for the families of WS applications, considering the ISO/IEC 9126-1 standard, as follows:

(a) *Specify architectural quality*. Use the WSA critical goals [11], establishing a correspondence with the ISO/IEC 9126-1 sub-characteristics and sub-sub-characteristics, to obtain a standard quality model for WSA.

(b) *Specify functional quality*. For each type of WS, quality properties are assigned to the functional requirements to express the goals that must be fulfilled. Standard quality requirements are obtained for each type of WS. New sub-characteristics or sub-sub-characteristics can be added, if needed.

(iii) *Define Measurement Model*: For each type of WS, *refine the quality model* obtained in step (ii), by specifying the attributes and metrics for each type of WS, obtaining a characterization of a sub-family of the domain, as follows:

(a) *Specify quality attributes and metrics*. For each sub-characteristic or sub-sub-characteristic of the quality model.

(b) *Assign the attributes (QoS characteristics) and the metrics,* according to the ISO/IEC 13236 standard.

Output: architectural knowledge for WS application domain

The process described provides guidelines that can be practically applied to facilitate traceability among the standards, which in the literature appear separate, and to settle the basis for the automatic generation of a standard quality contractual specification between WS clients and providers. The guidelines can be used as a starting point for a common language for the WS community. Each step is detailed in what follows.

3.1 STEP 1. DEFINE FUNCTIONALITY. CLASSIFICATION OF WS.

The classification presented in Table 3 shows the functional requirements for WS [15, 25]. Metrics must be provided as part of the contractual issue to establish the extent to which the WS fulfills its functionality [18, 23] (see Step 3). Note that in this taxonomy, the types of WS are not necessarily disjoint, since WS can be *composed*. For example, a Security WS can be used by a Transactional WS. This aspect is known as *orchestration* [13], where a central service controls the other services of the composition, to accomplish the required functionality.

Table 3. Functionality based WS classification.

WS type	Functional requirements
Information and collaborative environments	Data Base operations
Transactional	E-commerce operations, encrypting
Workflow	Process monitoring operations
Web Portal	E-search and e-communication
Security	Access control, encrypting

3.2 STEP 2. DEFINE A QUALITY MODEL

The quality model specifies a minimal set of properties characterizing applications; all the applications within this domain will share these properties. The quality requirements for the family of WS applications are specified using the ISO/IEC 9126-1 standard. A correspondence with the WSA critical goals [11] is established to specify architectural quality in terms of sub-characteristics and sub-sub-characteristics. Hence, the standard quality model is customized to the WSA goals.

Table 4. Quality Model for WS application domain, showing traceability among ISO/IEC 9126-1 and WSA. The codes of the WSA goals are taken from [11].

ISO/IEC9126-1 characteristics	Correspondence between ISO/IEC9126-1 sub-characteristics [6] and WSA critical goals [7]		
Functionality	**Interoperability**		Semantics is similar. **Goal: high**
	ISO/IEC	Interoperability	
	WSA	Interoperability (AG001)	
	Security		Semantics is similar. **Goal: high**
	ISO/IEC	Security	
	WSA	Security (AG004)	
	Suitability		Semantics is similar. **Goal: medium**
	ISO/IEC	Suitability	
	WSA	Team Goals (AG006)	
Reliability	**Availability**		Semantics is similar. **Goal: high**
	ISO/IEC	Availability	
	WSA	Reliability (AG006)	
Maintainability	**Extensibility**		The WSA *extensibility* goal is considered as sub-sub-characteristic of changeability. ***Goal: high.***
	ISO/IEC	Changeability	
	WSA	Scalability and Extensibility (AG006)	
	Management and provisioning		The WSA management and provisioning goal is considered a sub-sub-characteristic of changeability **Goal: high.**
	ISO/IEC	Changeability	
	WSA	Management and Provisioning (AG007)	
	Integration		The WSA integration goal is considered a sub-sub-characteristic of changeability. **Goal: medium**
	ISO/IEC	Changeability	
	WSA	Integration (AG003)	
Portability	**Scalability**		The WSA scalability goal will be considered a sub-sub-characteristic of adaptability. **Goal: medium**
	ISO/IEC	Adaptability	
	WSA	Scalability and extensibility (AG006)	

The quality model shows the minimal characteristics that providers must comply to guarantee *user satisfaction*. These properties are part of the contractual agreement for the control and measure procedures. In consequence, a WS must satisfy some of the quality properties indicated in Table 4. According also to this table, a WSA compliant service is now also compliant with the ISO/IEC 9126-1 standard characteristics (sub-characteristics and sub-sub-characteristics) for internal/external product quality, to which a high, medium or low goal ranking has been assigned by consensus by an expert group. Usability and efficiency are ranked low because they are not present as WSA critical goals [11]; they are not shown in the Table 4. In consequence, *the quality model for the WS domain considering relevant architectural properties* is conformed in this case by characteristics ranked *high* or *medium*, which are the following:

- functionality *(interoperability, security, suitability)*
- reliability *(availability),*
- maintainability *(Changeability (extensibility, management and provision, integration))*
- portability *(adaptability (scalability)).*

Table 5. Quality requirements for each WS type.

WS Type	**Quality requirements for sub-families of WS-based applications - characteristics and sub-characteristics according to ISO/IEC 9126-1**				
	Functionality	Reliability	Maintainability	Portability	**Efficiency**
Information and collaborative environments	**-accuracy**	-availability	-changeability		
Transactional	-security (integrity) **-accuracy**	-availability			**-time behavior -resource utilization**
Workflow	-suitability				
Web Portals				-adaptability: scalability	**-time behavior -resource utilization**
Security	-security				

The WSA quality model, enriched with the qualities related to WS functionality, constitutes the *standard quality model for the WS domain*. It expresses the overall *architectural quality* for WS applications. The enrichment is obtained considering all the quality characteristics shown before for WSA and adding the quality characteristics (shown in boldface), derived from each WS functionality, see Table 5: **efficiency** (*time behavior, resource utilization*) is required for some of the WS and the sub-characteristic **compliance to standards and regulations** is required to achieve interoperability, in order that the service conforms to standards like SOAP, UDI, WSDL in their respective versions. We assume that this characteristic is present for all WS types and so it is not specified in Table 5. Sub-characteristic **accuracy** has been included for data transactions indicating the precision of an event, set of events, condition or data [22]. Notice that often the term *integrity* is used in WS transactions to denote the fact of maintaining the correction of the transaction with respect to the source, which we are considering in the model as security.

3.3 STEP 3. DEFINE MEASUREMENT MODEL.

In this work, the QoS, which are quantifiable aspects or parameters, are considered *attributes of the sub-characteristics* of the WS domain quality model. Observe that traceability between the standards ISO/IEC 13236 [22] and ISO/IEC 9126-1 [21] is not explicitly provided by the standards, making difficult their practical usage. This work is a contribution towards the fulfillment of this gap. It is clear that the metrics presented are quite general and should be customized to establish the contractual part when using the service in a particular application. In what follows, the quality model is further refined for each WS type of Table 5. The refinement consists in finding the attributes or measurable elements and their metrics, for each WS quality property. These attributes and metrics correspond to the QoS characteristics considered in the ISO/IEC 13236 standard [22]. Table 6 shows the refinement for the Transactional WS category, since this is a complex WS type. The other refinements can be obtained in a similar way and they will not be shown here to ease the presentation, they are detailed in [30]. Notice also that Table 6 only shows some of the attributes, to facilitate legibility. They must be used depending on the application requiring the WS and on what is to be measured. For more detailed information on attributes and metrics the ISO/IEC13236 [4, 22, 26] standard document should be consulted.

Table 6. Measurement Model: QoS metrics for Transactional WS.

WS		**Quality characteristics and sub-characteristics according to ISO/IEC 9126-1**	**Attributes (QoS characteristics, according to ISO/IEC 13236)**	**Metrics, according to ISO/IEC 13236**
Transactional WS	Functionality	Compliance with standards and regulations for Interoperability	Depending on the regulations	Depending on the regulations
		Security	protection	Probability
			access control	Value or level derived from an access control policy.
			data protection	Value or level derived from the data integrity policy.
			confidentiality	Value or level derived from the data confidentiality policy.
			authenticity	Value or level derived from the data authentication policy.
	Functio	Accuracy	accuracy: {addressing, delivery, transfer, transfer integrity, allowable, release establishment} error	Probability
	Reliability	Availability	fault-tolerance	MTBF[a] =MTTF[b] + MTTR[c]
			fault-containment	Probability
			resilience, recovery error	Probability
			Agreed service time (channel, connection, processing)	A=MTBF/(MTBF+MTTR) when maintainability is involved, 0≤A≤1
	Efficiency	Time behavior	date/time	Any unit of time
			time delay: transit, request/reply, request/confirm	TD=T2- T1
			lifetime	Any unit of time
			remaining lifetime	Any unit of time
			freshness (or age of data)	Any unit of time
			capacity	Any unit of time
		Resource utilization	throughput (communication capacity)	Units depend on the resource type
			processing capacity	Rate (bits/seconds, bytes/seconds)
			operation loading	Instructions/seconds - Relation between used and available capacity

Note: Table notes.[a] MTBF: mean time between failures, [b] MTTR: mean time to replace, [c] MTTF: Mean Time to Failure.

4. CASE STUDY: VOICE PORTAL FOR AIRLINE COMPANIES

A Voice Portal (VP) outsourcer, AIRPORTAL, offers a suite of WS based on a self service platform which, combined with IP Telephony and Open Standards, like VoiceXML, provides a powerful speech and touch-tone solutions for airline companies [6, 9, 10]. AIRPORTAL enables integration through interoperability with its standardized approach to voice-based applications and common airline systems, enriching and refining a caller's experience using simultaneous voice and data interactions. Functionalities like phone-based booking for travel (make reservations, cancellation, blocking and confirmation transactions), user registration, online ticketing (purchasing, payment, status), query flights information, speech automation of airline-related call center routines, millage account balance, among others, can be offered by AIRPORTAL, through Voice Portal Web Services (VP-WS), to give their customers the highest levels of self-service, whether they are using a telephone, a computer or a mobile phone. In order to characterize the domain for VP-WS, the process proposed on Section 3 is applied. Based on the WS functionality, AIRPORTAL is a provider of Transactional WS and uses Web Portal WS (see Table 3). Only SLA Requirements for Transactional WS will be considered in this study.

AIRLINE is a regular AIRPORTAL's customer. AIRPORTAL, rather than acquiring and owning the infrastructure for hosting these WS, seeks out a provider of computational services called AIRSERVICE. It supplies to AIRPORTAL the computing resources needed to host the VP-WS, which include VP applications hosting and auditing (monitoring, calculation, notification), call center agents on-line reporting, storage systems, data base hosting, server-rental, networking components, and Internet connectivity [5, 18]. Finally the SLA is checked and the information requested is send to the AIRLINE. A contract between AIRPORTAL and AIRLINE is then established, this SLA (Service Level Agreement) specifies a minimal set of properties and characteristics of the provided voice portal services, such as: average conversation time, volume of abandoned calls, call center agents occupation, transactions successfully executed, among others [6, 9, 18]. The main functional and non functional requirements are:

- *Reservation Transactions*: At least 95% of the reservation requests shall always provide real time information about flight availability information, 80% of the confirmation volume is accepted, a lower value is considered a fault and the call must be delivered directly to an agent. Those requests should have a server-side response time of 10 seconds.
- *Service Factor*: Maintain an 80/20 services level of attendance, which implies to assure 80% of total calls to be answered before 20 seconds, this involves an online monitoring of agents in the call center and the calls volume. The VP should be available 99.9% of the time.
- *Navigation Time*: Represent the VP application conversation time and the average of talk time shouldn´t be higher than 5 minutes.

The general information for SLA for Transactional WS is contained in the Measurement Model for WS domain (see Table 6). Using the Quality Model of the domain for WS (See Table 5), the characteristics that providers must comply to guarantee *user satisfaction* can be identified.

Then, instantiating the Measurement Model for the VP-WS, the attributes or measurable elements and their metrics are defined. Table 7 shows the SLA requirements for VP-WS offered by AIRPORTAL and required by AIRLINE [4, 22, 26] and Table 8 shows how they are going to be measured and presented in the SLA contract.

Table 7. Contractual features defined between AIRLINE and AIRPORTAL

WS	**SLA Requirements**	**Quality Characteristic**	**Quality sub-characteristic and attributes (QoS)**	
			ISO/IEC 9126-1	**ISO/IEC 13236**
VP-WS	Reservation Transactions	Functionality	Accuracy	Transfer integrity error
		Efficiency	Time behavior	Time delay
	Service Factor	Reliability	Availability	Agreed processing time
				Fault-tolerance
	Navigation Time	Efficiency	Time behavior	Time delay

Table 8. Contractual measures and conditions/actions defined between AIRLINE and AIRPORTAL

WS	**SLA Require-ments**	**Measurement**	**Conditions/ Actions**	**WSLA Variable Label**
VP-WS	Reservation Transactions	A Probability. It refers to the precision of the information obtained by the user from the service. Levels Definition: 95% is Required. 80% is considered Acceptable. Any value >80% is a Fault.	If Total Errors >80% call must be delivered directly to an agent.	RT _F_AC
		The elapsed time between two general events, in this case start when the agent initiates a request. Levels Definition: 10 Seconds per Transaction is Requested. Any value upper than 10 is considered as a Fault.	Notify if average of Time Response <10 seconds.	RT_ E_TB
	Service Factor	Proportion of agreed processing time for which satisfactory service is available. $A=MTBF^{a}/(MTBF+MTTR)^{b}$ when maintainability is involved, $0 \le A \le 0.8$	If availability is under 80% more agents must be logged-in.	SF_R_AV1
		Average time it takes for the system to fail plus the average it takes to recover. $MTBF = MTTF + MTTR^{c}$	If availability < 99.9% use mainframe system.	SF_R_AV1
	Navigation Time	The elapsed time between two general events, in this case start when the customer is attended by the voice portal application. Period of Calculation: Minutely. Levels Definition: 3 Minutes is Expected. Any value upper than 5 is considered as an Alarm.	Notify the average if the Voice portal conversation time >5 min.	NT _E_TB

The SLA expressed in WSLA [5, 6, 9, 18], has been generated automatically from Tables 7 and 8. The code can be accessed at: *http://www.informationsys.com/WSLA_VP-WS.xml*

5. CONCLUSION AND FUTURE WORK

The architectural knowledge for the domain of WS applications has been characterized in this work, unifying three quality standards to capture architectural knowledge. The standard specification of the quality properties related with a WS has been emphasized. To establish this characterization, a process has been proposed: first, take into account the WSA critical goals [11, 12]. The WSA market reference establishes the minimal set of requirements that a family of WS applications must hold. A quality model is established for this domain, according to ISO/IEC 9126-1 [21]. Secondly, the WS functionality is identified; for each functionality, quality properties and goals are established. The initial quality model is enriched, to characterize each family of WS, focusing also on the quality properties inherent to the functionality. Finally, this quality model is instantiated for families of WS, considering the attributes or QoS metrics, according to ISO/IEC 13236 [22]. In this way, three separate standards have been related and put into a practical use. Common understanding among different stakeholders has been set by this correspondence. As a case study, a family of Transactional WS has been considered, to show the final quality model with metrics; the process is similar for any type of WS. The quality model established is a reusable artifact and can be customized to any WS application and family of WS. From this the SLA is automatically generated in WSLA.

Ongoing research works are:

- On one hand the definition of a specification pattern based on WSLA; this pattern will facilitate the automatic identification of requirements that must be considered, as well as the metrics that must be calculated to guarantee the commitment. The contractual specification based on quality issues allows guaranteeing the functionality of the services, since it can be monitored by notifying mechanisms that take actions accordingly, if required.
- On the other hand, we are working further on the domain characterization using a standard quality model, often product of the combination of different standards, as a specification tool for quality concerns. This model is used as a common language between different stakeholders and can be used to monitor quality through the development process, considering the different quality views. In an architectural design approach [3] it is used to define the properties that the initial system architecture must comply.

REFERENCE

[1]. D. Gross, E. Yu. (2001). “From Non-Functional Requirements to Design Through Patterns”, Requirements Engineering Journal, vol. 6, pp. 18-36

[2]. F. Losavio, L. Chirinos, A. Matteo, N. Lévy, A. Ramdane-Cherif, (2004) “ISO quality standards to measure architectures”, Journal of Systems and Software, Elsevier Science, vol. 72, pp. 209-223

[3]. F. Losavio, N. Lévy, P. Davari, F. Colonna. (2005), “Pattern-Based Architectural Design Driven by Quality Properties: A Platform for Scientific Calculatiom, 2nd. European Workshop, EWSA, Pisa, Italy, June 2005, LNCS 3527, pp. 94-112.

[4]. L. Chirinos. F. Losavio, J. Boegh. (2005). “Characterizing a Data Model for Software Measurement”. (2004). Journal of Software Systems, Elesevier Science, vol. 74, pp. 207-226.

[5]. A. Dan, D. Davis, R. Kearney, A. Keller, R. King, D. Kuebler, H. Ludwig, M. Polan, M. Spreitzer, A. Youssef. Web services on demand: WSLA-driven automated management. 0018-8670/04/$5.00 © 2004 IBM. IBM SYSTEMS JOURNAL, VOL 43, NO 1, 2004. http://www.research.ibm.com/journal/

[6]. Avaya. Avaya Voice Portal. © 2005 Avaya Inc. All Rights Reserved Avaya. Printed in the U.S.A. 07/05. GCC2784. http://www.avaya.com

[7]. Framework for QoS routing and related traffic engineering methods for IP-, ATM-, and TDM-based multiservice networks. E.360.1 (2002). http://www.itu.int/home/

[8]. F. Losavo, L. Chirinos, N. Lévy, A. Ramdane-Cherif, (2003). “Quality Characteristics for Software Architecture”, Journal of Object Technology (JOT), March/April 2003. vol 2, no. 2, pp.133-150. http://www.jot.fm/issues/issue_2003_03/article2

[9]. Genesys Telecommunications Laboratories, Inc. www.genesyslab.com. Genesys Voice Platform: Enterprise Edition 7.2 Deployment Guide. Copyright ©2002–2006. Genesys Telecommunications Laboratories, Inc. All rights reserved. Document Version: 72gvp-ee_dep_02-2006_v7.2.001.00. http://www.genesyslab.com

[10]. World Wide Web Consortium. Voice Extensible Markup Language (VoiceXML) Version 2.0. W3C Recommendation 16 March 2004. Copyright © 2004 W3C® (MIT, ERCIM, Keio), All Rights Reserved W3C. http://www.w3.org/TR/voicexml20/

[11]. World Wide Web Consortium.. Web Services Architecture Requirements. W3C Working Group Note 11 February (2004). Copyright © 2004 W3C® (MIT, ERCIM, Keio). http://www.w3.org/TR/wsa-reqs

[12]. World Wide Web Consortium. Web Services Description Requirements. W3C Working Draft 28 October (2002). Copyright © 2002 W3C® (MIT, ERCIM, Keio). http://www.w3.org/TR/ws-desc-reqs

[13]. World Wide Web Consortium. Web Service Modeling Ontology. W3C Member Submission 3 June 2005. W3C Working Draft 28 October (2002). Copyright © 2005 W3C®. http://www.w3.org/Submission/WSMO/

[14]. A. Mani, A. Nagarajan. “Understanding quality of service for Web services”. (2002) IBM Software Labs, India.

[15]. A. Thomas Manes. “Web Services. A Manager’s Guide”. (2003). Addison Wesley.

[16]. Berard E. (1992). “Essays in Object-Oriented Software Engineering”, Prentice Hall, 1992.

[17]. De Champeaux D., D. Lea, P. Faure. (1993). “Object-Oriented System Development”, Addison-Wesley

[18]. D. Menascé, V. Almeida. (2000). “Cappacity Planning and Performance Modeling”, Prentice Hall.

[19]. E. Navarro, P. Letelier I. Ramos, (2004). “Goals and Quality Characteristics: Separating Concerns”, Spanish CICT, DYNAMICA project, www.um.es/giisw/i-oficial/proyectos.htm

[20]. I. Sommerville, P. Sawyer. (2000). “Requirements Engineering, A Good Practice Guide”.

[21]. ISO/IEC 9126-1, 2 “Quality characteristics and guidelines for their use (1, 2), (1999). International Organisation for Standardisation / International Electrotechnical Commission.

[22]. ISO/IEC 13236. Quality of Service: Framework. Version 1. (1999). International Organisation for Standardisation / International Electrotechnical Commission.

[23]. P. Caudwell, R. Chaawla, V. Chopra, Gary Damschen, C. Dix, T. Hong, F. Norton, U. Ogbuji, G. Olander, M. Richman, K. Saunders, Z. Zaev. “Servicios Web XML”. (2001). Ediciones ANAYA.

[24]. A. Moreira, J. Araujo, I. Brito (2002) “Crosscutting Quality Attributes for Requirements Engineering”, SEKE, pp. 167-174

[25]. A. Ginige, S. Murugesan. (2001). “Web Engineering: An Introduction”. University of Western Sydney, Australia. IEEE Multimedia, vol. 8, no. 1, Jan-March, 2001, pp. 14-47.

[26]. F. Losavio, A. Matteo, R. Rahamut. CARACTERIZACION DE APLICACIONES CON SERVICIOS WEB BASADA EN CALIDAD. CIAWI 2006 – IADIS. Conferencia Ibero Americana WWW/Internet 2006. Murcia, España., October 2006, pp. 90-95.

[27]. M. Weiss. (2003). “Pattern-driven Design of Agent Systems: approach and case stud”, proceedings caise03.

[28]. Thayer, R., Dorfman, M. (2000). “Software Requirements Engineering”, IEEE Computer Society.

[29]. D. Lamanna, J. Skene, and W. Emmerich, “SLAng: A Language for Defining Service Level Agreements,” in Proceedings of the 9th IEEE Workshop on Future Trends of Distributed Computing Systems, San Juan, Puerto Rico, May 2003.

[30]. C. Molina-Jimenez, S. Shrivastava, E. Solaiman, and J. Warne, “Contract Representation for Run-time Monitoring and Enforcement,” in Proceedings of the IEEE International Conference on E-Commerce Technology, Newport Beach, California, USA, June 2003.

[31]. A. Andrieux, K. Czajkowski, A. Dan, K. Keahey, H. Ludwig, T. Nakata, J. Pruyne, J. Rofrano, S. Tuecke, and M. Xu, “Web Services Agreement Specification (WS-Agreement),” Grid Resource Allocation and Agreement Protocol (GRAAP) WG, September 2005, http://www.ggf.org/ Public Comment Docs/Documents/Oct-2005/WS-AgreementSpecificationDraft050920.pdf

Tassonomy and Review of Complex Content Models

Pierfrancesco Bellini, Ivan Bruno, Paolo Nesi, Michele Spighi
Distributed Systems and Internet Technology Lab,
Department of Systems and Informatics, University of Florence, Florence, Italy
{nesi, pbellini, ivanb}@dsi.unifi.it

Abstract

The content technology needs to conquer new forms with more intelligence, more flexibility and more complete features than those which are present on the market or proposed by standards. Innovative content formats have been proposed while they have difficulties to conquer the market for the lack of general consensus and tools. In this paper, an analysis of the state of the art about complex content models is presented. The analysis has been developed in the context of European Commission projects, in the intelligent content area. The analysis allowed us to identify a number of topics and features on which the models should evolve and the major lacks in the most diffuse models and formats.

1. Introduction

Presently there is a large number of content formats that ranges from the simple files basic digital resources (documents, videos, images, audio, multimedia, etc.) to integrated content models such MPEG-21 [1], [2], MXF, MHP [12], AXMEDIS [3], [4], [5], NewsML, SCORM/IMS [6], MPEG-4 [7], and intelligent content proposals [8], and proprietary formats such as Macromedia, Adobe, etc. Most of these formats try to wrap different kinds of digital resources/files in a container/package with their related information (e.g., content metadata and descriptors, relationships among resources, etc.) and they make such resources ready for delivery (streaming and/or downloading), in plain (clear-text) and/or protected forms. The presentation level and user interaction are formalized with specific formats such as: SMIL, HTML, MHP, MHEG, Laser, Java, SVG, BIFS, etc. The content behavior is in few cases formalized in Java and/or Javascript. The metadata are frequently defined together with content ID codes. Among the metadata: Dublin Core, TVAnyTime, MPEG-7, etc. and among the ID codes, ISBN, ISAN, ISRC, ISMN, etc.

MPEG-21 is focused on the standardization of the content description related to DRM, digital rights management aspects [9], [10]. MPEG-21 was initially not designed to be used as package for cross media content with internal presentation layers. More recently the AXMEDIS format, extended MPEG-21 [11] to cope with cross media content packaging, with HTML and SMIL presentation and layers for user interaction. AXMEDIS response provoked the starting into MPEG Forum of an action to evolve the MPEG-21, in that direction. SCORM/IMS has been defined to formalize the learning content and packages. SCORM is a comprehensive standard for the organization and delivery of learning packages, and has defined a run-time environment for web-browsers with several available implementations. MHEG is a standard for the formalization of the presentation/interaction aspects for broadcasting. Differently from AXMEDIS format the MHEG has not been designed for content distribution and packaging. The MHP [12] enables the reception and execution of interactive, Java-based applications on a TV-set. Interactive MHP applications can be broadcast together with audio and video streams. The applications can be on information services, games, interactive voting, e-mail, SMS or shopping. MXF has been designed as an exchange format to address a number of problems of non-professional formats. MXF has full timecode and metadata support, and it is intended as a platform-agnostic stable standard for future professional video and audio applications. SMIL is similar to an HTML-like language designed to produce interactive presentations, and may have links to other SMIL presentations and graphical elements to allow user interaction. SMIL provides features like transitions, animations, etc., is one of the underlying technologies used by HD DVD and MMS for advanced interactivity. SVG is an XML specification and file format for describing vector graphics, both static and animated. The W3C explicitly recommends SMIL as the standard for animation in SVG, however it is more common to find SVG animated with ECMAScript (javascript). BIFS is MPEG-4 Part 11, it is a binary format for two and three-dimensional audiovisual content, with graphic rendering and interaction. BIFS is MPEG-4 scene description protocol to compose, describe interactions and animate MPEG-4 objects. NewsML and SportML belong to a family of news formats that presently demonstrate their inefficiency to cope with the complexity of the novaday news.

Among the most relevant European Commission research and development projects in the field, we have posed the attention on the most focused on intelligent content such as: ACEMEDIA [13], X-MEDIA [17], AXMEDIS [3], SALERO [15], ICONS [16], and LIVE [14]. ACEMEDIA [13] defined a new format of content that may be created by a commercial content provider, to enable personalized self-announcement and automatic content collections, or may be created in a personal content system in order to make summaries of personal content, or automatically create personal albums of linked content. However, some important features such as the capabilities of evolving that are the evolutional aspects are not considered. X-MEDIA project and content model are mainly focused on semantic aspects in content that can be managed by ontologies and RDF. X-Media is mainly oriented towards knowledge management and sharing with limited application to text and image contents and has related content objects with very limited autonomy of work that are not proactive with the user. SALERO aimed mainly to cross media-production for games, movies and broadcasting. In SALERO, the object evolution management and collaboration during the whole objects lifecycle are not provided, thus the content is not proactive. LIVE project content is focused on semantic based description of live content for streaming and broadcasting. While it is very far from the concept of intelligent content for the lack of autonomy.

Other object containers are (i) EMMOs (Enhanced Multimedia Meta Objects) [18] which encapsulates relationships among multimedia objects and maps them into a navigable structure. An EMMO contains media objects, semantic aspect, associations, conceptual graphs, functional aspect; (ii) KCO Knowledge Content Objects which is not a package [19], and it is based on the DOLCE foundational ontology and have semantic aspects to describe the properties of KCOs, including raw content or media item, metadata and knowledge specific to the content object and knowledge about the topics of the content (its meaning). The semantic information in a KCO includes: content description; propositional description (Semantic Description and Content Classification); presentational description; community description (the purpose); business description (the trade, price…); trust and security description, self description (the structure).

These last models are less powerful with respect to the above mentioned models such as AXMEDIS and MPEG-21, while they present descriptors that may be used for more powerful classifications of these objects.

The paper is organized as follows. Section 2 reports the model for the analysis of complex content models and thus the tassonomy for comparing them. The analysis has permitted to identify the limits of many presently diffuse formats and the future needs for future models and improvement. Conclusions are drawn in section 3.

2. Assessing Content Model

In order to assess content models and packages against their capabilities of autonomy and intelligence a number of factors have been identified. Some of them are present in state of the art formats while others are innovative aspects. In Table 1, a taxonomy of the relevant features which allowed us to assess the above mentioned formats is reported. The table does not report all format and models considered but only the most representative among the different categories mentioned before: standards, proprietary format, research projects, presentation models, etc.

In the next paragraphs, a short description of the most relevant factors that have been taken into account to perform the analysis is reported.

Structural aspects: how an object is composed in terms of digital essences/files, metadata, other objects; in which relationships the elements are arranged; how it is possible to link the internal and/or external elements of the content, etc. One object elements may have references/links to other objects and elements. The structural complexity is a problem for the content distribution and usage since multiple paths and non linear stories are not simple to be streamed and accessed in real time when played. Regarding structural aspects, cross media formats are those that may contains other essences, for example AXMEDIS, SCORM, MXF, etc. In some cases, they can be also managed as groups of files glued by presentation models, such as HTML, SMIL.

Behavioral aspects have to do with synchronizations, animations, functional aspects such as the coding of procedural aspects into content to be executed, procedures associated with multimodal user actions, the business logics of the content. This aspect is frequently modeled with simple languages (java, javascript) that could include communications capabilities, event managements, collaborative and coordinated activities, processing capabilities, etc. In formats such as MXF this aspects is missing, while in MPEG-21 this aspect is delegated to DIP/DIM parts which are not fully defined in terms of semantics in the standard. AXMEDIS proposed a large extension to these capabilities.

Annotation aspects have to do with the aim of adding descriptors and additional information to content elements and structure. Thus creating additional information and non linear paths, that can be created by a specific user and/or shared with others: multiuser annotations. Semantic annotations are fundamental in order to coordinate different tools and for defining intelligent behavior. It should be possible to perform annotations of multimedia content such as audio, video, images, 3D, etc., with multimedia content. For example, a video can be annotated in a given point (time and space of a scene) by an audio, a text, a video, etc.; annotation of images, audio, video, documents, 3D, animations, etc.; activation of behavior from annotations, etc. Profiling can be also dynamically adjusted-tuned on the bases of the user behavior and interaction.

Metadata and descriptors: the object model has to include structured knowledge and representation for an efficient data search and indexing. In order to generate metadata and semantic descriptors, automatic and manual technologies are used. In addition, the model has to support the extension of metadata so as to host any kind of metadata and to convert them in other models. Metadata has to include the basic elements such as identification codes and classifications models. The object behavior has to be capable to take into account metadata and descriptors in its evolution. Descriptors included into content objects such as in SALERO, KCO, AXMEDIS, may be used for creating powerful indexing and classification models and thus for sophisticated query search of content on the basis of semantics. Descriptors can be also freely added into AXMEDIS, MPEG-21 or other packages. So that the indexing and the semantics search can be performed also on these content formats and models. Descriptors may be formalized in MPEG-7 of in direct XML. Ontologies and RDF formats are used as well.

Processing aspects to cope with capabilities which are needed to process content and data, take decisions, evolving according to the status, for example (i) to make internal content processing, transcoding when an adapted version of a given content element has to be produced, (ii) to elaborate data when those accessible are not conformant to those needed, (iii) to convert formats of video, audio, and documents, (iv) to make an internal searching even considering user added content, (v) to take into account about profiling and status conditions of the content, etc. Thus reporting in the processing all the evolutionary capabilities of the model and format. The capabilities related to take decisions on the basis of the current profiles and content status implies a certain level of autonomy in taking decisions directly enforced into the content object. This is possible only when the content itself has some capability of processing logic instructions via some autonomous engine. Frequently this possibility is delegate to Java, Javascrip, and/or to some rule based models.

Interaction aspects can be managed from the server side, such as the usage of URL in HTML pages, or from the client side. For example, when all the interactive possibilities are pre-encoded such as in MPEG-4, MHP, etc. A mix, when client side interaction may lead to server side links. Recently, multimodal user interactions are requested, supporting different devices (mouse, remote control, mic, gesture, joystick, etc.), external events and actuators, interaction with 3D represented environments, etc. The presence of processing and behavior aspects may help in creating more powerful interactivity on client side.

Presentational aspects: presentation aspects, audio and visual aspects and perception, screen layout format and arrangement, styling, rendering of digital essences, overlaying on videos, over-positioning on video, user interaction integration with presentation level. Among these aspect also the activation of presentational aspects from behavioral, synchronization, insertion of 3D animations interacting with the users and with the other issues of the scene are very relevant for the new kind of content. MHP and HTML have the capabilities of activating javascript/java functional aspects so that the presentation is integrated with behavior, while they do not present satisfactory integration with the metadata and with the structural and processing capabilities.

Profiling aspects: data profiling related (e.g., of user, device, context, network, etc.) that can be used to take decisions about behavior and rendering on the player side. Profiling can be also dynamically adjusted-tuned on the bases of the user behavior and interaction. Profiling aspects (user, device, network and context) are also related to behavioral aspects since the object behavior should be based on the profiling. CCPP and MPEG-21 standards can be used for modeling profiling.

Communication aspects: most of the present formats are very limited in terms of communications. They can only communicate with the server side of the application and the content itself cannot take communication control among different content elements. An example is HTML that has good capabilities of communicating via HTTP protocol while limitations are present when communication with other users is needed. MPEG-21 has no support on these aspects.

	MPEG-4	MPEG–21	SCORM/IMS	NewsML	MHEG	MHP	SM	HTML	MXF	BIFS	SVG	Lase r	F l ash	AXMEDIS	ACEMEDIA	SALERO	X-MEDIA	LIVE	ICONS
Structural																			
Cross media nesting levels	P	Y	Y	Y					P					Y	P				
Any kind of content files		Y	Y				P	A						Y	A				Y
Link and navigation	P			P			Y	Y			Y	Y	Y	Y	P	Y			P
Behavioural																			
Synchronization	Y				Y	Y	Y		Y	Y	Y	Y	Y	P		Y		Y	
Internal functional parts	Y	Y	Y		Y	Y		A		Y	Y	Y	Y	Y	P				
Annotation capabilities																			
Basic annotation	P	Y	Y										Y	Y	Y	Y	Y	Y	Y
Multimedia annotation		P							P					P		Y	Y		
Metadata and descriptors																			
Basic Metadata	Y	Y	Y	Y	Y	Y	Y	Y	Y		Y	Y		Y	P	Y	Y	Y	Y
Extendible Metadata		Y	Y	Y					Y		Y			Y	P	Y	Y	Y	Y
Descriptors and semantics	A													P	P	Y	Y	Y	P
Processing definition																			
Processing capabilities	Y	Y			Y	Y		A		Y			P	Y	P				
Internal searching		Y						Y						Y					
Interaction data/info																			
Interactive aspects	Y	P			Y	Y	Y	Y		Y	Y	Y	Y	Y	P	Y		P	
Multimodal interaction												P	P		P	Y			
Presentation data																			
Integrated presentation layers	Y				Y	Y	Y	Y		Y	Y	Y	Y	Y	P	Y		Y	Y
Animations	Y					Y	Y	A		Y	Y	Y	Y	Y	P	Y			
3D aspects	Y							A		Y	Y	Y	Y	A	A	Y			
Profiling																			
User profiling		Y												Y		Y		Y	
Device profiling					Y			P		Y		Y		Y	P	Y		Y	
Other profiling														P					
Communication																			
General Communication		Y	Y	Y			Y	Y				Y	Y	P			Y		Y
Collaboration														P			Y		Y
Model Management																			
Self versioning/evolution/merging																			
Undo, selective undo																			
File format																			
XML and/or Text format	Y	Y	Y	Y	Y	Y	Y	Y		Y	Y	Y		Y	Y	Y	Y		Y
Binary File format	Y	Y							Y	Y	Y	Y	Y	Y				Y	
Packaging	Y	Y	Y						Y		Y	Y	Y	Y					
Distribution																			
Download	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Streaming/Progressive download	Y	P			Y	Y		P	Y	Y		Y	P	P				Y	
DRM and protection																			
IPR management	Y	Y												Y					
Protection models	Y	Y												Y					
Tools																			
Graphic Authoring					Y	Y	Y	Y			P		Y	Y					
Players: PC, Mobiles, PDA, STB	Y	P	Y	P	Y	Y	Y	Y	P	P	Y	P	Y	Y	P	P	P	P	P

Keys: Y = supported, P = partially supported, A = supported by plug-in/external extensions

File format aspects: presently most of the open formats are XML based (e.g., MPEG, SCORM, HTML) while most of the proprietary are mainly binary (e.g., Flash, macromedia). The formats are in large parts covered by patents. The openness of the format has to be guaranteed, and XML based formats have to be integrated with those of binary formats. Some formats may support both XML and Binary such as AXMEDIS. The advantages of binary formats are mainly in the performance and file size. Packaged formats that integrate several essences and data are a way to reduce complexity as perceived by the user (e.g., MXF, AXMEDIS). When complex packages are distributed the binary model may limits the capabilities of distribution for the complexity of the model, see the section on Distribution and in general MPEG-21 Digital Item Streaming, AXMEDIS [3] and LIVE [14].

Model management aspects: how the content object may change its status, issues of versioning and merge. The versioning has also to do with activities of linear, non linear undo and selective undo [20]. In addition, these aspects are even more complex to be managed when the content can be produced in collaborative manner [21]. The versioning and selective undo are strongly related to the control and protection of the intellectual property: that are the actions performed and the contributions provided by a given user in the community, into the complex content. The versioning can be totally regarded as an authoring tool feature. On the other hand, versioning and undo are features requested by the final users and thus the content has to be saved with the history information inside, including information about who has performed the model changes in the collaborative environment.

Distribution aspects: how the object can be distributed. Most of the formats are mainly focused on download, while only a part of them can be streamed or progressively downloaded. When the file format is structurally complex, containing several essences, the streaming becomes too complex or impossible. A balance is needed to have both capabilities without losing generalities and capabilities. Typically, formats that can be streamed contain simple resources such as single audio visual. In some cases, multiple video streams can be arranged by losing the concept of the package. The distribution is strongly related to the protection and digital rights management aspects. Who is interested in protecting the content has to enforce into the content model the support for distributing the content in protected packages or via protected channels.

Digital rights management, DRM, aspects have to cope with the management of the IPR (intellectually property rights) at different levels. The technologies of DRM and protection integrated together should keep under control the consumption of rights according to rights owners wishes. The DRM may support different business models such as pay per play, subscription, etc. Some of the models may be based on sophisticated REL (rights expression languages) such as that of MPEG-21 REL and OMA. Simple solutions are based on CAS, conditional access systems, which only control the access to the content from the users. Typical broadcasting distributions of MHP content (e.g., DVB-T in the present implementations in Europe) and of MPEG-4 such as in OpenSKY are based on simple CAS, and thus have limitations in the number of business models that they may support and implement. Gently or educative DRM may provide information to what the user cannot do or can do without limiting the user actions.

Tools for complex content may range from simple authoring tools which are capable to cover only some aspects (see for example the MPEG-4 authoring tools) to complete authoring tools endowed of graphic user interface and help to support powerful products such as Flash, HTML, SMIL. Among the tools, it is quite diffuse to distribute simple players that allow playing content without any or with strongly reduced capabilities of authoring – e.g., flash player, SMIL, AXMEDIS player. Content tools have to be integrated with all the above mentioned aspects to be considered functional to the model. Players may be provided for multiple platforms such as: PC, PDA, mobiles and STB.

2.1. Comments

The analysis performed was very complex since each line of the table implies several aspects to be assessed, tested and verified. Thus a large number of models and formats have been reviewed and for each of them we have given a full Yes only when the corresponding feature was almost satisfactory covering a large part of the issues mentioned in the description of the table line. On the contrary, the table cell has been left empty. In some cases, the features considered have been identified as Partially covered by the solution provided, so that in those cases, a P has been assigned. In some cases, when the features are covered by additional tools and A has been placed.

The most challenging features of the next future content models would be the integration of: content packaging and semantic aspects with versioning, multimodal user interaction and collaboration capabilities. These features will provoke a large revolution on the content formats since structural changes in the models are needed to supports these features and the present formats are not ready.

3. Conclusions and future tasks

In this paper, a model and tassonomy for assessing and comparing complex content models has been reported. The analysis has been developed in the context of European Commission project AXMEDIS in the intelligent content area. The analysis allowed us to identify a number of topics on which the future models should evolve and the major lacks in other very diffuse models and formats. According to the analysis, collaboration and versioning aspects are most innovative features to be integrated together with semantics aspects. The complexity of managing the versioning, semantics aspects and multiple users and the related authoring tools are the most relevant challenges that have to be solved in the next years in the area of complex intelligent content. The analysis performed allowed us to draw the design and implementation of the AXMEDIS model and tools. Performing now the analysis again, at the end of the AXMEDIS project, we realized that AXMEDIS resulted among the models considered one of the most complete and powerful. Additional future work for AXMEDIS is mainly in the direction of enforcing capabilities of multimedia annotation and versioning in the AXMEDIS/MPEG-21 model. MPEG-21 has basic support for annotation, while complex structures and descriptors can be added only by additional modeling work. The work already started in those directions.

4. Acknowledgments

The authors would like to thank all the AXMEDIS project partners (BBC, TISCALI, XIM, TEO, ANSC, AFI, EUTELSAT, Giunti ILABS, HP Italy, MBI, FHGIGD, DIPITA, ACIT, CPR, EXITECH, University of Leeds, University of Cataluña, University of Reading, etc.), the Expert-User-Group and all affiliated members for their contributions, supports and collaborations. The authors would also like to express their thanks to the European Commission IST DG-INFSO FP6 for partially funding of AXMEDIS.

5. References

[1] Burnett, I.; Van de Walle, R.; Hill, K.; Bormans, J.; Pereira, F., "MPEG-21: goals and achievements", IEEE Multimedia, Vol.10, Issue 4, pp.60-70, Oct-Dec 2003.

[2] Burnett, I.S., Davis, S.J., Drury, G. M., "MPEG-21 digital item declaration and Identification-principles and compression", IEEE Transactions on Multimedia, Vol.7, N.3, pp.400-407, 2005.

[3] AXMEDIS, "Framework and Tools Specifications", http://www.axmedis.org

[4] Bellini P., P. Nesi, D. Rogai, and A. Vallotti, "AXMEDIS Tool Core for MPEG-21 Authoring /Playing," presented at International Conference on Automating Production of Cross media content for multichannel distribution AXMEDIS2005, Florence, 30 Nov – 2 Dec 2005.

[5] Bellini P., P. Nesi, D. Rogai. "Exploiting MPEG-21 File Format for Cross Media Content". 13th International Conference on Distributed Multimedia Systems (DMS), San Francisco Bay, USA, 6-8 September 2007.

[6] Mourad, M., Hnaley, G.L., Sperling, B.B., Gunther, J., "Toward an Electronic Marketplace for Higher Education", Computer of IEEE, pp.58-67, June 2005.

[7] Pereira F. and Ebrahimi T., eds. (2002). The MPEG-4 Book. IMSC Press, Prentice Hall.

[8] Bordegoni M., G Faconti, S Feiner, M Maybury, T Rist, S Ruggieri, P Trahanias, and M Wilson. A standard reference model for intelligent multimedia presentation systems. Computer Standards and Interfaces, vol.18, pp.477-- 496, 1997.

[9] Lee, J., Hwang, S.O., Jeong, S.-W., Yoon, K.S., Park, C.S., Ryou, J.-C., "A DRM Framework for Distributing Digital Contents Through the Internet", ETRI Journal, Vol.25, N.6, pp.423-435, December 2003.

[10] Nesi, P.; Rogai, D.; Vallotti, A., "A Protection Processor for MPEG-21 Players". In Multimedia and Expo, 2006 IEEE International Conference. Page 1357—1360. 9-12 July 2006, Toronto, Canada.

[11] Bellini P., P. Nesi, L. Ortimini, D. Rogai, A. Vallotti, "Model and usage of a core module for AXMEDIS/MPEG21 content manipulation tools", IEEE International Conference on Multimedia & Expo (ICME 2006), Pages 577--580 Toronto, Canada, 9-12 July, 2006.

[12] MHP: http://www.mhp.org,

[13] ACEMEDIA: http://www.acemedia.org/aceMedia

[14] LIVE: http://www.ist-live.org

[15] SALERO: http://www.salero.info/en/project/index.html

[16] ICONS: http://www.icons.umd.edu/

[17] X-MEDIA: http://www.x-media-project.org

[18] Schellner Karin, Westermann Gerd Utz, Zillner Sonja and Klas Wolfgang CULTOS: Towards a World-wide Digital Collection of Exchangeable Units of Multimedia Content for Intertextual Studies. In Conference of Distributed Multimedia Systems (DMS2003), 2003.

[19] Behrendt, W. Gangemi, A., Maass, W., and Westenthaler, R. (2005): Towards an Ontology-Based Distributed Architecture for Paid Content. In: A. Gómez-Pérez and J. Euzenat (Eds.): ESWC 2005, LNCS 3532, pp. 257–271, 2005. Springer-Verlag Berlin Heidelberg 2005

[20] Berlage, T. 1994. "A selective undo mechanism for graphical user interfaces based on command objects", ACM Transactions on Computer-Human Interaction, Vol.1, Issue 3, 269-294.

[21] Bellini P., P. Nesi, and M. Spinu, "Cooperative Visual Manipulation of Music Notation", ACM Transactions on Computer-Human Interaction, ACM TOCHI, Vol.9, N.3, pp.194-237, ISSN 1073-0516, Sept. 2002.

TESTING MULTI-AGENT SYSTEMS FOR DEADLOCK DETECTION BASED ON UML MODELS

Nariman Mani *Vahid Garousi* *Behrouz H. Far*

Department of Electrical and Computer Engineering
Schulich School of Engineering, University of Calgary, Canada
{nmani, vgarousi, far}@ucalgary.ca

ABSTRACT

There is a growing demand for Multi-Agent Systems (MAS) in the software industry. The autonomous agent interaction in a dynamic software environment can potentially lead to runtime behavioral failures including deadlock. In order to bring MAS to the main stream of commercial software development, the behavior of MAS must be tested and monitored against the risk of unwanted emergent behaviors including deadlocks. In this paper, (1) we introduce a method for preparing test requirements for testing MAS; and (2) we deploy a MAS monitoring method for deadlock detection in MAS under test. The first method helps create test requirements using a resource requirement table from the MAS analysis and design. The second method monitors the MAS behavior to detect deadlocks at the run-time. Also, as model based software techniques such as Multi-agent Software Engineering (MaSE) are gaining more popularity; these model based approaches can help MAS developers to reduce the risk of having unwanted emergent behaviors such as deadlock in MAS.

Index Terms— Multi-agent system, Software testing, Deadlock detection, UML.

1. INTRODUCTION

Increasing demand for applications which can communicate and exchange information to solve problems collaboratively has led to the growth of distributed software architecture consisting of several interoperable software systems. One of the main difficulties of interoperable software systems is heterogeneity. Heterogeneity reflects the fact that the services offered by constructed components are independent from the designers and the design methodology [1]. Different programs written in different languages by different programmers must operate in a dynamic software environment. Agent based software engineering is one of the approaches devised to handle collaboration and interoperability. An autonomous agent is a computational entity that can perceive, reason, act, and communicate [2]. Multi-Agent Systems (MAS) consists of autonomous agents that try to achieve their goals by interacting with each other by means of high level protocols and languages [1]. However, the agent interaction can potentially lead to runtime behavioral failures including deadlock. Thus, testing and monitoring MAS to eliminate the risk of unwanted emergent behaviors, such as deadlock, is an essential precondition for bringing MAS to the main stream of commercial software. Also, as model-based software development practices are gaining more popularity [3], more and more MAS are developed using model-based practices such as the Multi-agent Software Engineering (MaSE)[4]. Thus, model-based testing techniques for deadlock detection for MAS can be useful since they can help MAS engineers to eliminate the risks of deadlocks in the MAS development.

In this paper we focus on proposing a methodology for testing MAS by preparing test requirements for deadlock detection. The artifacts used are the models prepared during the analysis and design stages of a MAS using the MaSE methodology[4] . Figure 1 illustrates the approach. Using the procedure explained in Section 5.1, resource requirement table is constructed based on Control Flow Paths (CFP) extracted from the MaSE task diagrams. The resource requirement table is used for searching for potential deadlocks (Section 5.2). Test requirements for testing MAS are prepared based on the potential deadlocks. The test requirements are used to generate the test cases. For deadlock detection on MAS under test we deploy our MAS monitoring methodology in [5]. Using the procedure explained in Section 4, a MAS behavioral model, consists of UML sequence diagrams, is constructed using MaSE analysis and design artifacts such as “role sequence diagram”, “agent class diagram” and “task diagram”. Two deadlock detection techniques, introduced in Section 6, are instrumented into the MAS under test’s source code. Test driver executes the test cases on MAS under test and runtime deadlocks are detected using the MAS behavioral model [5].

The remainder of this paper is structured as follows. The related works and background are described in Section 2. The MAS metamodel is introduced in Section 3. Constructing MAS behavioral model based on the MaSE is discussed in Section 4. Test requirement preparation is described in Section 5. MAS monitoring for deadlock detection is explained in Section 6. Finally conclusions and future work are given in Section 7. An illustrated example is used to explain the methodology in the subsequent sections.

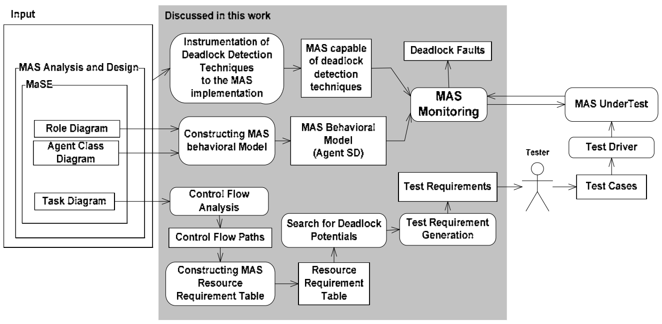


Figure 1 – An overview of our approach

2. RELATED WORKS AND BACKGROUND

2.1 MAS Verification and Monitoring

Existing works on MAS verification are categorized into axiomatic and model checking approaches [2]. In [6], axiomatic verification is applied to the Beliefs, Desires and Intentions (BDI) model of MAS using a concurrent temporal logic programming language. However, it was noticed that this kind of verification cannot be applied when the BDI principles are implemented with non-logic based languages [2]. Also in design by contract [7] pre- and post-conditions and invariants for the methods or procedures of the code are defined and verified in runtime. Violating any of them raises an exception. But as it is also claimed in [2] the problem is that this technique does not check program correctness, it just informs that a contract has been violated at runtime.

Model checking approaches seem to be more acceptable by industry, because of less complexity and better traceability as compared to axiomatic. Automatic verification of multi-agent conversations [8] and model checking MAS with MABLE programming language [9] are a few examples of model checking approaches that both use SPIN model checker [10], a verification system for detection of faults in the design models of software systems.

2.2 Deadlock Detection Techniques

Resource and communication deadlocks models are considered in message communication systems. Most deadlock models in distributed systems are resource models [11-13]. In these models, the competition is on acquiring required resources and deadlock happens whenever an entity is waiting permanently for a resource which is held by another. As indicated in [13], the communication deadlock model is general and can be applied to any message communication system. The communication model is an abstract description of a network of entities which communicate via messages. A deadlock detection mechanism based on the communication model deadlock for distributes systems and operating systems is provided in [13]. In literature a deadlock situation is usually defined as "A set of processes is deadlocked if each process in the set is waiting for an event that only another process in the set can cause" [14]. There are four conditions that are required for a deadlock to occur [14]. They are (1) "Mutual Exclusion" which means each resource can only be assigned to exactly one process; (2) "Hold and Wait" in which processes can hold resources and request more; (3) "No Preemption" which means resources cannot be forcibly removed from a process; and (4) "Circular Wait" which means there must be a circular chain of processes, each waiting for a resource held by the next member in the chain [14]. Similar to other types of the faults there are four techniques commonly used to deal with deadlock problem: ignorance, detection, prevention, and avoidance [14].

2.3 Agent Based Development Methodology: MaSE

MaSE uses several models and diagrams driven from the standard Unified Modeling Language (UML) to describe the architecture-independent structure of agents and their interactions [4]. In MaSE a MAS is viewed as a high level abstraction of object oriented design of software where the agents are specialized objects that cooperate with each other via conversation instead of calling methods and procedures. There are two major phases in MaSE: analysis and design (Table 1). In analysis phase, there are three steps which are capturing goals, applying use cases and refining goals. In the design phase, there are four steps which are creating agent classes, constructing conversations, assembling agent classes and system design[4].

Table 1- MaSE methodology phases and steps [4]

MaSE Phases and Steps	*Associated Models*
1. Analysis Phase	
a. Capturing Goals	Goal Hierarchy
b. Applying Use Cases	Use Cases, Sequence Diagrams
c. Refining Roles	Concurrent task, Role Diagram
2. Design Phase	
a. Creating Agent Classes	Agent Class Diagrams
b. Constructing Conversations	Conversation Diagrams
c. Assembling Agent Classes	Agent Architecture Diagrams
d. System Design	Deployment Diagrams

3. MAS METAMODEL

Figure 2 shows a metamodel for the MAS structure. In this figure, each MAS can be presented by MAS behavioral model in terms of sequence diagrams which shows the conversations of several agents and the message exchanging among them. The way of constructing such kind of behavioral model from MaSE design and analysis diagrams is introduced in Section 4. Each MAS consists of several agents whose roles are the building blocks used to define agent's classes and capture system goals during the design phase. Associated with each role are several tasks and each task can be presented by MaSE task diagram [4]. A task diagram in MaSE is a UML state machine diagram which details how the goal is accomplished in MAS and can be represented by a Control Flow Graph (CFG) [15, 16]. A CFG is a static representation of a program that represents all alternatives of control flow. For example, a cycle in a CFG implies iteration. In a CFG, control flow paths (CFPs), show the different paths a program may follow during its execution.

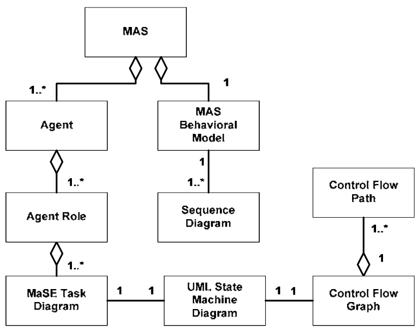


Figure 2- metamodel for MAS

4. CONSTRUCTING MAS BEHAVIORAL MODEL

Agents of a MAS communicate by exchanging messages. The sequence of messages is useful for understating the situation during faults detection conversation. A common type of interaction diagrams in UML is a sequence diagram in which each agent or role is represented by a lifeline in sequence diagram.

We deploy a method for transforming the conversations of agents from MaSE to UML sequence diagrams. These sequence diagrams are used in MAS monitoring method for deadlock detection in MAS under test [5]. The MAS sequence diagrams is not provided by MaSE per se and must be constructed using information provided by the MaSE artifacts such as role diagram and agent class diagrams [2].

The role sequence diagram in “Applying Use Cases” step in analysis phase of MaSE shows the conversations between roles assigned to each agent [4]. The agent class diagram is created in the “Constructing Agent Classes” step of MaSE represents the complete agent system organization consisting of agent classes and the high-level relationships among them. An agent class is a template for a type of agent with the system roles it plays. Multiple assignments of roles to an agent demonstrate the ability of agent to play assigned roles concurrently or sequentially. The agent class diagram in MaSE is similar to agent class diagram in object oriented design but the difference is that the agent classes are defined by roles, not by attributes and operations. Furthermore, relationships are conversations between agents [4]. Figure 3 shows examples of MaSE role sequence and agent class diagram.

The approach for constructing sequence diagrams based on the two above mentioned MaSE diagrams is defined as follow [5]. Each role sequence diagram is searched for the roles which are listed in the same agent class in the agent class diagram. Then, all of the roles in each role sequence diagram are categorized based on the agent which they belong to. Therefore, each category corresponds to an agent class in agent class diagram and the messages which it exchanges with other categories are recognizable. On the other hand, a new agent sequence diagram can be generated from agent class diagram which the lifelines are agents’ types. The recognized messages between each two categories are entered into agent sequence diagram as a new conversation. For example, in Figure 3, the role sequence diagram 1 is categorized into three different categories, the first one consists of Role 1 and Role 2 and the second one consists of Role 3 and Role 4 and the last one consists of Role 5. The first one corresponds to agent class 1, the second one corresponds into agent class 2, and the third one corresponds to agent class 3. The constructed agent sequence diagrams from role sequence diagram 1, 2 and 3 and agent class diagram in Figure 3 are shown in Figure 4.

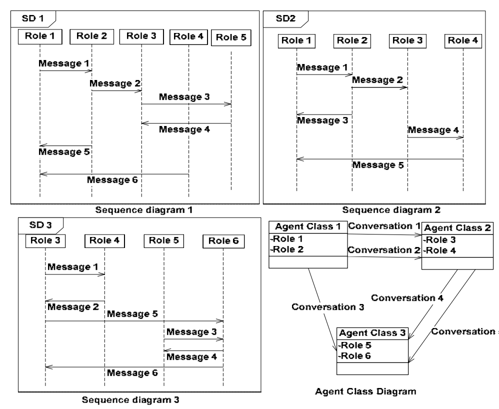


Figure 3- MaSE role sequence and agent class diagrams

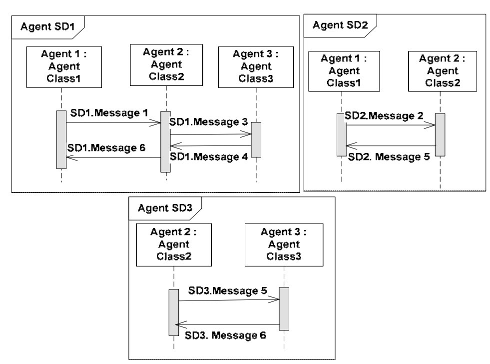


Figure 4- Constructed agent sequence diagrams

UML provides ways to model the behavior of an object oriented system using different types of diagrams such as state machine diagram. UML’s state machine diagram is based on finite state machines (FSM) augmented with the concepts such as hierarchical and concurrent structure on states and the communications mechanism through events transitions[3]. UML's state machine diagram is commonly used to describe the behavior of an object by specifying its response to the events triggered by the object itself or its external environment. State machine diagram has long been used as a basis for generating test data [15-17]. In MaSE [4], roles are the building blocks used to define agent's classes and capture system goals during the design phase. Every goal is associated with a role and every role is played by an agent class. Role definitions are captured in a role model diagram which includes information on communications between roles, the goals associated with each role, the set of tasks associated with each role, and interactions between role tasks. In MaSE, a task is a structured set of communications and activities, represented by a state machine diagram [4]. The MaSE's task diagram is then converted to UML's state machine diagram by converting some MaSE's task diagram notations such as the protocol transition, choices, and junctions to the UML notation. Using the state machine diagram, the CFG and its associated CFPs can be identified [15, 16].

5. TEST REQUIREMENT PREPARATION

In this section we focus on proposing a methodology for testing MAS by preparing test requirements for deadlock detection. Test requirements are generated using resource requirement table defined in Section 5.1.The resource requirement table is used in search for deadlock potentials (Section 5.2). The results from search for deadlock potentials are used for test requirement generation (Section 5.3). The test requirements are used by testers to generate the test cases for deadlock detection in MAS.

5.1 Resource Requirement Table for Agents

As discussed in Section 4, the behavior of each agent can be presented by the several MaSE task diagrams each reflecting a task assigned to a specific role of an agent. Each task consists of several CFPs that represent the different runs of the MaSE task diagram represented by UML state machine diagram. During the execution of each CFP, several resources are held and acquired by an agent. We define resource requirement table for each agent which shows the resource requirement for different tasks which are assigned to different roles of an agent (see Figure 5). Each row in resource requirement table shows the Required Resource Set (RS_{ij}) during execution steps of a specific CFP_i. If the required resources needed by a particular CFP_i are changed during its execution, a new set of the required resources on that stage is added to the resource requirement table for that CFP. Each column in resource requirement table represents the Sequence of Required resource Sets (SRS_i) by one CFP_i . We present the SRS formal definition as below:

$$SRS_i = < RS_{ij} | RS_{ij} \text{ is the jth required resource set of the } CFP_i >$$

And

$$RS_{ij} = \{ R_p | R_p \text{ is a required resource by } CFP_i \}$$

The metamodel in Figure 6 depicts the definition of resource requirement tables and its elements.

	Task 1					...	Task n			
	$CFP_{1,1}$	$CFP_{1,2}$	$CFP_{1,3}$	...	$CFP_{1,n}$	...	$CFP_{k,1}$	$CFP_{k,2}$	...	$CFP_{k,n}$
$Agent_i$	$\{R_1\}$									
	$\{R_1,R_2\}$									
	$\{R_3\}$									
	$\{R_4\}$									
	$\{R_4,R_5\}$									
	⋮									
	$\{R_3,R_{10}\}$									

Figure 5- An example for Resource Requirement Table for Agent

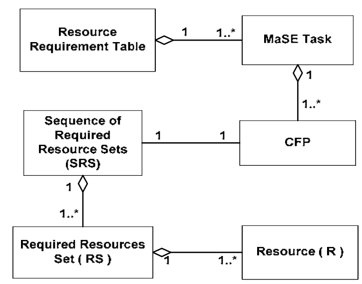


Figure 6- Resource requirement table metamodel

5.2 Search for Potential Deadlocks

In order to prepare the test requirement for deadlock detection between the CFPs, we first describe a scenario in which a deadlock happens. Figure 7 shows an example of resource allocations and resource requests (wait-for graph [18]) in deadlock situation.

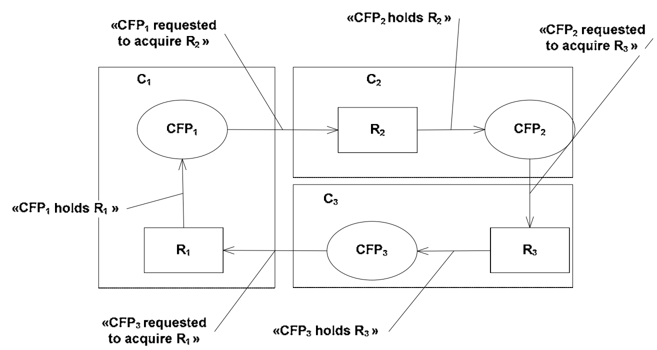


Figure 7- Resource allocations and requirements in deadlock situation (wait-for graph)

For explanation simplicity, we consider a situation that the RS set for each CFP has only one member .Each CFP holds one and request for acquiring the next required resource. The required resource may have already been acquired by another resource and the requestor has to wait for that resource. In resource model, CFP_1 is said to be dependent on another CFP_k if there exists a sequence of CFPs such as $CFP_1, CFP_2, \ldots, CFP_k$ where each CFP in sequence is idle and each CFP in sequence except the first one holds a resource for which the previous CFP is waiting.

If CFP_1 is dependent on CFP_k, then CFP_1 has to remain in idle status as long as CFP_k is idle. CFP_1 is deadlocked if it is dependent on itself or on a CFP which is dependent on itself. The deadlock can be extended to a cycle of idle CFPs, each dependent on the next in the cycle. Therefore, the deadlock detection approach is to declare the existence of that cycle.

The information for each SRS_i for each CFP_i can be retrieved from the resource requirement table defined in Section 5.1. In the Figure 7 wait-for graph [18], each resource set (each resource set has just one member in this example) in the cycle will be in sequence of required resource sets (SRS) of two CFPs. One CFP is holding the resource set and the other one requesting for acquiring it. As an example, $\{R_1\}$ is required by both CFP_1 and CFP_3 as it is shown below:

$$\{R_1\} \in SRS_1 =< \{R_1\},\{R_2\} > \ and\ SRS_3 =< \{R_3\},\{R_1\} >$$

The procedure of finding potential deadlocks in the behavioral model of MAS is defined as follow. The sequence of required resources set by a CFP_i (SRS_i) is retrieved for the all the CFPs in the MAS from the resource requirement table (Section 5.1). For each CFP, we assume that it is holding one of its required resource sets (RS_{ij}). $RS_{i,j}$ represents the jth required resource set of the SRS_i. Then, the next required resource set by that CFP, $Next_i$ is identified using the SRS_i for CFP_i.We search inside the SRSs for other CFPs that require at least one resource from $Next_i$ and assume in the worst case, they are holding it. We repeat this procedure until we find one CFP requiring a resource which is held by the CFP that we have already traversed by our procedure. In this case a deadlock cycle is detected. We consider this cycle as a potential deadlock cycle.

We explain the procedure with an example shown in Figure 8. For explanation simplicity, we consider a situation that the RS set for each CFP has only one member .We start the procedure from CFP_1 and assume that it holds its first resource set $\{R_1\}$. The next required resource set by CFP_1 is $\{R_2\}$. We assign the next required resource set by CFP_1 , $Next_1$ as $\{R_2\}$ and search the CFPs which has at least one resource from $Next_1$ (in this case just $\{R_2\}$) as the required resource in their sequence of required resources sets SRS_i. CFP_2 is found and it is assumed that in the worst situation it holds R_2. So, if CFP_2 holds R_2, the next resource set required by CFP_2 is $\{R_3\}$ according to the SRS_2. The search is started again for finding the CFPs which require R_3 as the required resource. CFP_3 is found and it is assumed that it holds R_3 . So, the next resource set required by CFP_3 if it holds R_6 is $\{R_1\}$. We find out that R_1 has been already assumed to be held by CFP_1 when we wanted to start the procedure. Therefore, a deadlock potential cycle is detected.

The pseudocode of searching for the potential deadlock cycles is shown in Figure 10. In finding the potential cycles function we define potential deadlock cycle data structure $PotentialDeadlockCycle_r$ which illustrates r-th potential deadlock in the MAS as below:

$$PotentialDeadlockCycle_r = \{(CFP_i, RS_p, RS_q) \mid CFP_i \text{ is holding required resource set } RS_p \text{ and requesting for acquiring required resource set } RS_q\}$$

So, bases on the explained procedure the potential deadlock cycle, $PotentialDeadlockCycle_1$ for the example provided in Figure 8 is created as follow:

$$PotentialDeadlockCycle_1 = \{(CFP_1, \{R_1\}, \{R_2\}), (CFP_2, \{R_2\}, \{R_3\}), (CFP_3, \{R_3\}, \{R_1\})\}$$

These potential deadlocks which are found by the explained procedure are used for test requirement generation in the Section 5.3.

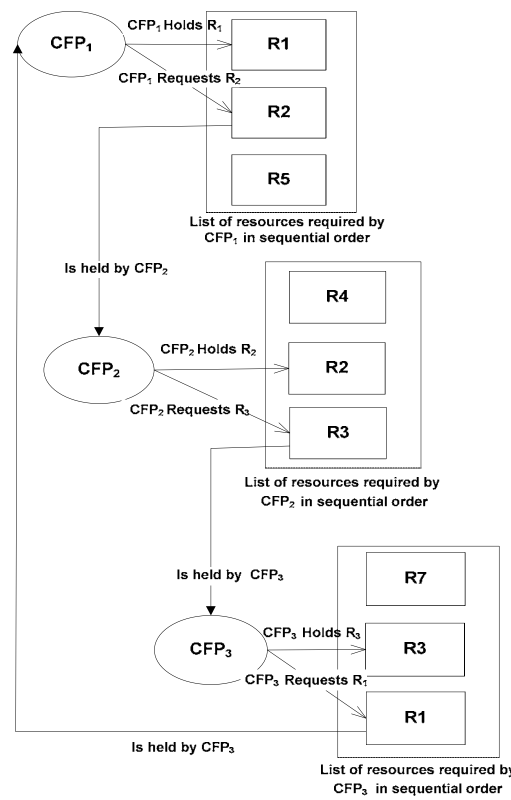


Figure 8- An example of finding deadlock potentials in the behavioral model of MAS

5.3 Test Requirement

We define the test requirement metamodel for testing MAS for deadlock detection and for each deadlock cycle in Figure 9. As it can be seen in Figure 9 test requirement for deadlock detection for a single deadlock cycle is divided into two parts. The first part is the Hold Set (HS) which represents the resource holdings in the MAS and is defined as below:

Hold Set: HS= $\{(CFP_i, RS_{p,})|\ CFP_i$ is holding required resource set $RS_{p,}\}$

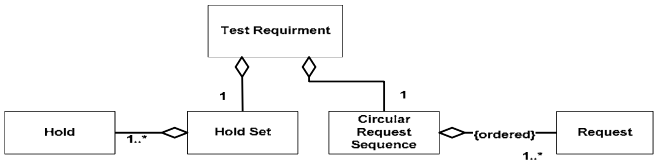


Figure 9 - Test requirement metamodel

The second part is the Circular Request Sequence (CRS) which shows the sequence of resource requests in the MAS and since it represents a cycle of requests we call it Circular Requests Sequence. It is defined as follow:

Circular Requests Sequence: CRS = < $Req_1, Req_2, \ldots, Req_n$>
$Req_i = (CFP_i, RS_q)|CFP_i$ requests for acquiring required resource set RS_q

In Figure 8 example, the test requirement based on the potential deadlock cycle ($PotentialDeadlockCycle_1$) found in the Section 5.2 is prepared as below:

HS= $\{(CFP_{1,}, \{R_1\}), (CFP_{2,}, \{R_2\}), (CFP_{3,}, \{R_3\})\}$
CRS = <$(CFP_{1,}, \{R_2\}), (CFP_{2,}, \{R_3\}), (CFP_{3,}, \{R_1\})$>

SoFarTraversed= $\{(CFP_{i,}, RS_{p,}RS_q)$ | CFP_i is holding required resource set RS_p and requesting for acquiring required resource set $RS_q\}$
FindPotentialDeadlocks(SRS_i, RS_{ij})
{
$Next_i$= NextRequiredResourceSet (SRS_i, RS_{ij})
Add ($CFP_{i,}, RS_{ij,} Next_i$) to SoFarTraversed list
Add (CFP_i, RS_{ij}) to the HS set
If there is any CFP_j other than SoFarTraversed in MAS $Next_i$ is in its SRS_j
Then For each CFP_j in MAS that $Next_i$ is in its SRS_j
- Assume that each CFP_j holds $Next_i$
- Find the next required resource set for each of them ($Next_j$)
- Search in HS to see if the $Next_j$ that they require has been already assumed to be held
- If $Next_j$ exists in HS
  Then so add SoFarTraversed+($CFP_{j,}, Next_i, Next_j$) as a new $PotentialDeadlockCycle_r$, Return
  Else
  Call FindPotentialDeadlock (SRS_j, $Next_j$)
  End if
End For
Else
Return
End if
}
Main ()
{
For all SRS_i in MAS
For all $RS_{ij,}$ in each SRS_i
FindPotentialDeadlocks (SRS_i, RS_{ij})
}

Figure 10- A pseudo-code for searching potential deadlocks

5.4 Testing MAS for Deadlock Detection

The test requirement prepared in Section 5.3 is used by a tester to generate the test cases for deadlock detection. In each test case the hold (HS in test requirement) and request (CRS test requirement) situations should be created and the system is tested to check if deadlock happens. Generated test cases are executed using the test driver. In this step, a deadlock detection methodology for executing system at runtime is required to detect the deadlocks and report them as fault. Our monitoring method for deadlock detection [5] can be used as the deadlock detection methodology to monitor the system behavior at runtime to detect deadlocks. This methodology focuses on model based deadlock detection by checking MAS communication for existence of deadlock. During the next section (Section 6) the method and its application to our approach in this paper is presented.

6. MAS MONITORING FOR DEADLOCK DETECTION

MAS monitoring for deadlock detection [5] is a model based deadlock detection which checks MAS communication for existence of deadlocks. The artifacts used are the models prepared during the analysis and design stages of a MAS using the MaSE methodology[4]. An overview of monitoring approach is also illustrated Figure 1. In the MAS monitoring the source code of the system is instrumented with two deadlock detection techniques discussed in this section to enable runtime deadlock detection in MAS under test.

6.1 Deadlock Detection in Resource Deadlock Model

Resource model of MAS consists of agents, resources and controller agents. A controller agent is associated with a set of agents and a set of resources which it manages [5]. Agents request for acquiring resources from their controller. Also, the controller can communicate with other controllers in case of requesting resources from other controllers. In [5] a gateway agent is proposed as a translator of the controllers' communications. In, a communication protocol is defined for controller agents to communicate, acquire resources and handle behavioral faults such as deadlock.

Whenever agent A_a in controller C_i needs to acquire a resource R_i associated to another controller C_j, it sends its request to its controller C_i. C_i communicates with controller C_j regarding the requested resource R_i. If the required resource is available, C_j provides that resource for agent A_a in controller C_i .But if it is hold by another agent A_b, C_i provides the identification of agent A_b to controller C_i . So, each controller agent has information about the internal resource allocation inside its set and the external resources that each agent in its set has already acquired or wants to acquire.

In order to determine for an idle agent A_a whether it is in deadlock state or not, a communication is initiated by its controller agent. In deadlock detection communication, controller agents send an *investigator* message to each other. An *investigator* message is of the form

$Investigator(n, m, a, b, r, c)$ denoting that it is initiated by controller of agent A_a for process P_n and transaction T_m regarding agent A_b which requested to acquire resource R_r that it is currently held by A_c . It follows that if C_i receives $Investigator(n, m, a, b, r, a)$ from another controller for any possible b and r and if R_r is one of the resources which is held by A_a, a circular wait is detected and C_i declare A_a as deadlocked.

Figure 11 shows the message communication between controllers for deadlock detection for the wait-for graph scenario discussed in Figure 7 (Section 5). Agent A_1 is holding resource R_1 associated to its controller C_1 and requested to acquire R_2 associated to controller C_2. A_2 in C_2 is holding R_2 and requested to acquire R_3 from C_3. A_3 in C_3 is holding R_3 and requested to acquire to acquire R_3 which is held by A_1 . According to our assumptions for resource deadlock model, three of four deadlock conditions are true in this example which are (1) "Mutual Exclusion"; (2) "Hold and Wait"; (3) "No Preemption". Also as it can be seen in Figure 11 the fourth condition circular wait can be detected after receiving $Investigator(n, m, A_1, A_3, R_3, A_1)$ by C_1 and identifying that R_3 requested by A_3 and is held by A_1. So all the four condition of deadlock is true and C_1 can declare A_1 as deadlocked.

6.2 Deadlock Detection in Communication Deadlock Model

In the communication deadlock model of MAS there is no controller agent or resources. Associated with each idle agent is a set of dependent agents which is called its *dependency set*. Each agent in that set can change its state about one particular task from idle to executing upon receiving a message from one of the members of its dependency set regarding that task. We define a nonempty set of agents as deadlocked if all agents in that set are permanently idle regarding a special task. An agent is called permanently idle, if it never receive a message from its dependency set to change its state. In the more precise definition, a none empty set of agents *S* is called deadlock if and only if all agents in *S* are in idle state, the dependency set of every agent in *S* is a subset of *S* and there are no active conversation between agents in *S*. Based on this definition, all agents in the set can be called permanently idle, if the dependency set of each agent such as A_i is in *S*, so they are all in idle state, and also if there is not any trigger message in transit between agents just because there is no active conversation.

An idle agent can determine if it is deadlocked by initiating a deadlock detection conversation with its dependency set when it enters to idle state. An agent A_i is deadlock if and only if it initiates a query conversation to all the agents in its dependency set and receives reply to every query that it sent. The dependency set for each agent is identified using the MAS behavioral model constructed in Section 4. The purpose of initiating this query is to find out if the agent A_i belongs to a deadlock set S with the mentioned conditions above. On receiving a query by an idle agent in dependency set, it should forward the query to its dependency set if it has not done already.

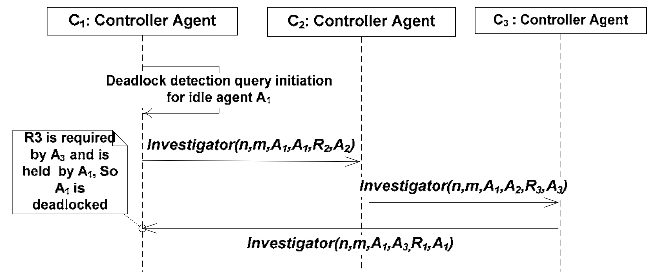


Figure 11 - deadlock detection example in resource deadlock model

Each query has the form $DeadlockDetectionQuery(n, m, i, a, b)$ denoting this message belongs to mth deadlock detection communication initiated for nth communication in MAS by agent A_i which is sent from agent A_a to agent A_b .Each agent A_k keeps the latest deadlock detection communication which it has been participated in it by $LatestDeadlockDetectComm(i)$ denoting the latest deadlock detection communication number that A_k was participated in it and initiated by A_i. The state of each agent (idle/executing) also is stored by $State(n, m, i)$ denoting the state of agent A_k for mth deadlock detection communication initiated by A_i for nth communication in MAS. Also the number of replies which should be received by an agent for mth deadlock detection communication initiated by A_i for nth communication in MAS is stored in $NumOfReplies(i)$.

We present the following scenario as an example for deadlock detection in the communication deadlock model of a hypothetical MAS with the sequence diagrams shown in Figure 4 (Section 4). In this scenario, agent A_1 is executing and has not received any message in one of its communications (in this case the communication in Agent SD 1) for a defined time *T* from its dependency set {A_2, A_3}. This is since A_2 and A_3 are both in waiting state and A_1 is not aware of it. After the defined time *T* , A_1 identifies itself as an idle agent and initiates a deadlock detection conversation with each agent in its dependency set. An agent A_1 declare itself as deadlocked if and only if it initiates a query conversation to all the agents in its dependency set and receives reply to every query that it had sent.

The complete deadlock detection scenario for the mentioned scenario is shown as a sequence diagram in Figure 12. A_1 initiates two query conversations with its dependency set which are A_2 and A_3 . A_2 and A_3 propagate the queries to their own dependency set which are{ A_1, A_3 } for A_2 and { A_1 , A_2} for A_3 . Respectively A_2 and A_3 receive reply from their own dependency sets. Thus they both replies to A_1 which is the initiator of deadlock queries in this scenario. So, A_1 receives reply for all quires which it had initiated to its dependency set. So, is declares itself as deadlocked.

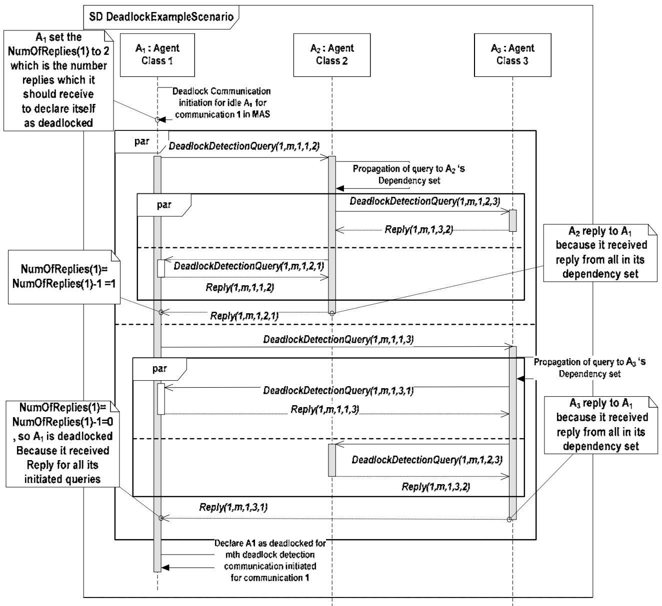


Figure 12 – A deadlock detection scenario in communication deadlock model using our technique

7. CONCLUSION AND FUTURE WORK

In this paper, we presented a methodology for test requirement generation and monitoring the behavior of MAS. The methodology is used to detect deadlocks as one of the unwanted emergent behaviors. Test requirements for deadlock detection are prepared using a resource requirement table. The test requirements are used by a tester in test case generation process. Test cases are executed using a test driver on the MAS under test. For deadlock detection in the MAS under test, a MAS monitoring methodology is proposed using our work in [5]. The source code of the system is instrumented with specific instructions in terms of deadlock detection techniques to enable runtime deadlock detection. As the future work, we plan to automate the test case generation process based on the test requirement for deadlock detection in MAS. Also, our methodology will be applied to a few more MAS case studies to evaluate its coverage, effectiveness and efficiency.

8 . REFERENCES

[1] M. R. Genesereth and P. K. Ketchpel, "Software agents " *Commun. ACM* vol. 37 pp. 48-53, 1994.

[2] F. Bergenti, M.P.Gleizes, and F. Zambonelli, *Methodologies and Software Engineering for Agent System* vol. 11. New York: Kluwer Academic Publishers, 2004.

[3] Object Management Group (OMG), "UML 2.1.1 Superstructure Specification," 2007.

[4] S. A. DeLoach, "The MaSE Methodology," in *Methodologies and Software Eng. for Agent System*. vol. 11, F. Bergenti, M.P.Gleizes, and F. Zambonelli, Eds. New York: Kluwer Academic Publishers, 2004, pp. 107-147.

[5] N. Mani, V. Garousi, and B. Far, "Monitoring Multi-agent Systems for Deadlock Detection Based on UML Models," in *IEEE CCECE 08 - Computer Systems and Applications*, 2008.

[6] M. J. Wooldridge and P. Ciancarini, "Agent-Oriented Software Engineering: The State of the Art," in *Proc. of the Workshop on Agent-Oriented Soft. Eng.*, 2000, pp. 1-28.

[7] B. Meyer, "Applying Design by Contract," *IEEE Computer,* vol. 25, pp. 40–51, 1992.

[8] H. L. Timothy and S. A. DeLoach, "Automatic Verification of Multiagent Conversations," in *the Annual Midwest Artificial Intelligence and Cognitive Science* Fayetteville, 2000.

[9] M. J. Wooldridge, M. Fisher, M. Huget, and S. Parsons, "Model Checking Multi-Agent Systems with MABLE," in *Proc. of the Int. Joint Conf. on Autonomous Agents and Multiagent Systems*, 2002, pp. 952–959.

[10] G. J. Holzmann, "The Model Checker Spin," *IEEE Trans. on Soft. Eng.,* vol. 23, pp. 279–295, 1997.

[11] V. Gligor and S. Shattuck, "Deadlock detection in distributed systems," *IEEE Trans. Soft. Eng.,* pp. 435-440, 1980.

[12] B. Goldman, "Deadlock detection in computer networks," Massachusetts Institute of Technology, Cambridge, Mass., Tech. Rep. MIT-LCS-TR185, 1977.

[13] K. M. Chandy, J. Misra, and L. M. Haas, "Distributed Deadlock Detection," *ACM Trans. on Computer Systems,* vol. 1, pp. 144-156, 1983.

[14] A. Tanenbaum, *Modern Operating Systems*. Englewood Cliffs: Prentice Hall Inc., 1992.

[15] T. S. Chow, "Testing Software Design Modeled by Finite-State Machines," *IEEE Trans. on Software Eng.,* vol. 4, pp. 178-187 1978.

[16] H. S. Hong, Y. G. Kim, S. D. Cha, D. H. Bae, and H. Ural, "A test sequence selection method for statecharts," *Software Testing, Verification and Reliability,* vol. 10, pp. 203 - 227, 2000.

[17] L. C. Briand, Y. Labiche, and Q. Lin, "Improving Statechart Testing Criteria Using Data Flow Information," in *Proc. of the 16th IEEE Int. Symposium on Software Reliability Engineering*, Washington, DC, USA, 2005, pp. 95 - 104

[18] D.A. Menasce and R.R. Muntz, "Locking and Deadlock Detection in Distributed Data Bases," *IEEE Transactions on Software Eng.,* vol. 5, pp. 195-202, 1979.

IP-Surveillance System Using Interactive Dual Camera Hand-Off Control with FOV Boundary Detection*

Chu-Sing Yang[1], Ren-Hao Chen[1], Chin-Shun Hsu[2], Chen-Wei Lee[3], Eugene Yeh[4]

[1]Institute of Computer and Communication Engineering, National Cheng Kung University, Tainan 70101, Taiwan, R.O.C

[2]Institute For Information Industry, Networks & Multimedia Institute Southern Innovation Center, Tainan 70101, Taiwan, R.O.C

[3]Department of Computer Science and Engineering, National Sun Yat-Sen University, Kaohsiung 804, Taiwan R.O.C

[4]National Center for High-performance Computing, Taiwan R.O.C

ABSTRACT

In this paper, an IP-surveillance system using interactive dual camera hand-off control with FOV boundary detection, which is called IIDCHC-FBD, is developed. The proposed camera hand-off control is used to assist client applications in changing a monitoring connection automatically when tracking people move between dual cameras in an enclosed rectangular surveillance environment, e.g. an ATM lobby or a museum. According to experimental results, IIDCHC-FBD would be a feasible solution that provides more reliable and scalable surveillance services in the next-generation IP-surveillance system.

1. INTRODUCTION

Nowadays, most surveillance systems still need considerable human input to monitor potential threats. Thus, it may cause the key frames to be lost, or need too many operators to control. In contrast, this system can capture key frames automatically without too many operators anymore. Otherwise, there is some research that uses a large number of multiple sensors: floor pressure sensors [1], infra-red sensors [2], RFID (radio frequency identification) tag systems [3] and cameras, to detect the position of users and track their movement. In addition, a joystick can be used to control the direction of movement of a PTZ (pan-tilt-zoom) camera to track an intruder. But, it is subject to human error.

Furthermore, recently some research has proposed dual-camera frameworks. Collins et al. [4] describes a master-slave architecture designed to acquire biometric imagery of humans at distance with a stationary wide field of view master camera is used to monitor an environment at distance and a narrow field of view slave camera is commanded to acquire the target human. Liu et al. [5] reports the design and develops a hybrid video camera system that combines the best aspects of PTZ cameras and a panoramic camera. Nevertheless, in all research, both cameras are placed facing the same direction, hence, frequently reducing the possibilities of capturing good quality frontal images of the monitored people. For example, when intruder moving and back to the camera will cause these systems to miss information about the intruder's face. The high quality frontal images can be useful for further automated processing e.g. the tracking of an identified object, face finding, face recognition etc. Therefore, acquiring information of the face is an extremely important task. In general, the intruder faces toward the direction in which he advances. This property is utilized to design an interactive dual camera hand-off control approach to decide automatically that which camera is the best suited for obtaining images of the face of the intruder.

The aim of IIDCHC-FBD is to develop an intelligent video-based surveillance system for tracking the intruder automatically with the best suited camera.

The characteristics of IIDCHC-FBD are threefold:

1. Cameras can track the intruder automatically without human assistance.
2. The system is capable of selecting the camera best suited to achieve surveillance task.
3. The system can procure optimum surveillance purpose while deploying the least number of cameras.

The organization of the remainder of the paper is as follows. In section 2, the detailed description of our framework and a scenario in a simulated environment is provided. In section 3 the details of the practical implementation for the scenario are described. In section 4 the work done and the contributions made are summarized.

2. PROPOSED FRAMEWORK

The proposed framework uses two cameras to undertake the task of monitoring the door to detect any new intruders while tracking a current intruder. A diagrammatic representation of the surveyed premises used for experiments in our framework has been shown in figure 1.

The framework decides which camera is better suited for tracking the intruder. The other camera automatically takes over the task of monitoring the door. In our framework, a minimum of two cameras are required so as to provide images of the face in both directions.

2.1 The algorithmic approach for the proposed framework (Interactive dual camera hand-off control algorithm)

The algorithmic approach for the proposed framework has been illustrated in figure 2. The triggering event for the

* *This work was supported in part by project grant NSC 96-2218-E-006-287 from the National Science Council, Taiwan, ROC.*

dual camera system is the entry of an intruder into the room. By object segmentation (section 2.2), the centroid of the intruder can be ascertained. At present, the masterCam tracks the intruder and the slaveCam monitors the entrance. According to our Camera moving algorithm (section 2.3), the system can decide the direction of movement of the masterCam.

Subsequently, the system can select the best suited camera to track the intruder by way of Camera selecting algorithm (section 2.4). In hand-off, the slaveCam passes information about the current intruder's position to the masterCam. The masterCam continues to track the intruder until he exits the surveyed premises.

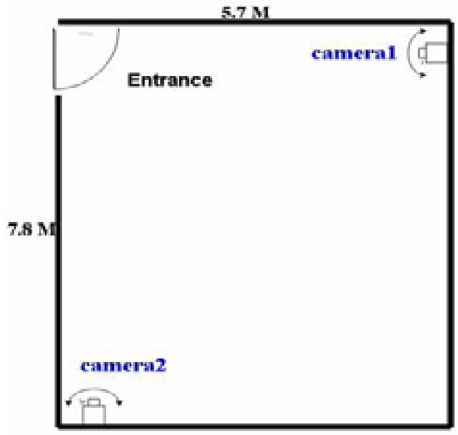


Fig.1 The surveyed premises

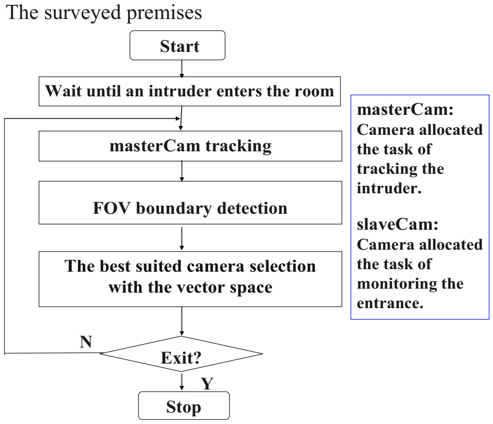


Fig. 2 Interactive dual camera hand-off control algorithm

2.2 Object segmentation

In order to obtain the centroid of the intruder, a Background subtraction algorithm has been adopted to identify the alien image of the intruder in the scene. The background image is stored as the reference image, in which there is no movement. Then, in every subsequent frame subtract the reference image to extract the alien image of the intruder. Having ascertained an image of the intruder, a bounding-box can be applied to him. Finally, the centroid of the intruder can be found based on the bounding-box.

2.3 FOV boundary detection

The first definition at time t, the coordinate of the FOV (field of view) is $(X_{F(t)}, Y_{F(t)})$ which is called FOV(t), and FOV(t) is a global coordinate. Every FOV has an unique coordinate respectively.

There is the position of the centroid of the intruder which is (χ_c, y_c) on the inside of the every FOV; χ_{Br} is the x-axis position of the right dotted line (right boundary); χ_{Bl} is the x-axis position of the left dotted line (left boundary); y_{Bu} is the y-axis position of the above dotted line (up boundary); y_{Bd} is the y-axis position of the under dotted line (down boundary); (χ_c, y_c), χ_{Br}, χ_{Bl}, y_{Bu} and y_{Bd} are local coordinates (Figure 3).

Here, a Camera moving algorithm is proposed that will decide the direction of movement of the masterCam according to the position of the centroid in the current FOV.

Camera moving algorithm:

View by camera2:

$(X_{F(t+1)},Y_{F(t+1)})=(X_{F(t)}+1,Y_{F(t)}+1)$, if $(\chi_c>\chi_{Br})\&\&(y_c<y_{Bu})$;
$(X_{F(t+1)},Y_{F(t+1)})=(X_{F(t)}-1,Y_{F(t)}+1)$, if $(\chi_c<\chi_{Br})\&\&(y_c<y_{Bu})$;
$(X_{F(t+1)},Y_{F(t+1)})=(X_{F(t)}-1,Y_{F(t)}-1)$, if $(\chi_c<\chi_{Bl})\&\&(y_c> y_{Bd})$;
$(X_{F(t+1)},Y_{F(t+1)})=(X_{F(t)}+1,Y_{F(t)}-1)$, if $(\chi_c>\chi_{Br})\&\&(y_c> y_{Bd})$;
$(X_{F(t+1)},Y_{F(t+1)})=(X_{F(t)}+1,Y_{F(t)})$, if $(\chi_c>\chi_{Br})$;
$(X_{F(t+1)},Y_{F(t+1)})=(X_{F(t)}-1,Y_{F(t)})$, if $(\chi_c<\chi_{Bl})$;
$(X_{F(t+1)},Y_{F(t+1)})=(X_{F(t)},Y_{F(t)}+1)$, if $(y_c<y_{Bu})$;
$(X_{F(t+1)},Y_{F(t+1)})=(X_{F(t)},Y_{F(t)}-1)$, if $(y_c> y_{Bd})$;
Do nothing, otherwise;

View by camera1:

$(X_{F(t+1)},Y_{F(t+1)})=(X_{F(t)}-1,Y_{F(t)}-1)$, if $(\chi_c>\chi_{Br})\&\&(y_c<y_{Bu})$;
$(X_{F(t+1)},Y_{F(t+1)})=(X_{F(t)}+1,Y_{F(t)}-1)$, if $(\chi_c<\chi_{Br})\&\&(y_c<y_{Bu})$;
$(X_{F(t+1)},Y_{F(t+1)})=(X_{F(t)}+1,Y_{F(t)}+1)$,if $(\chi_c<\chi_{Bl})\&\&(y_c> y_{Bd})$;
$(X_{F(t+1)},Y_{F(t+1)})=(X_{F(t)}-1,Y_{F(t)}+1)$, if $(\chi_c>\chi_{Br})\&\&(y_c> y_{Bd})$;
$(X_{F(t+1)},Y_{F(t+1)})=(X_{F(t)} -1,Y_{F(t)})$, if $(\chi_c>\chi_{Br})$;
$(X_{F(t+1)},Y_{F(t+1)})=(X_{F(t)} +1,Y_{F(t)})$, if $(\chi_c<\chi_{Bl})$;
$(X_{F(t+1)},Y_{F(t+1)})=(X_{F(t)},Y_{F(t)} -1)$, if $(y_c<y_{Bu})$;
$(X_{F(t+1)},Y_{F(t+1)})=(X_{F(t)},Y_{F(t)} +1)$, if $(y_c> y_{Bd})$;
Do nothing, otherwise;

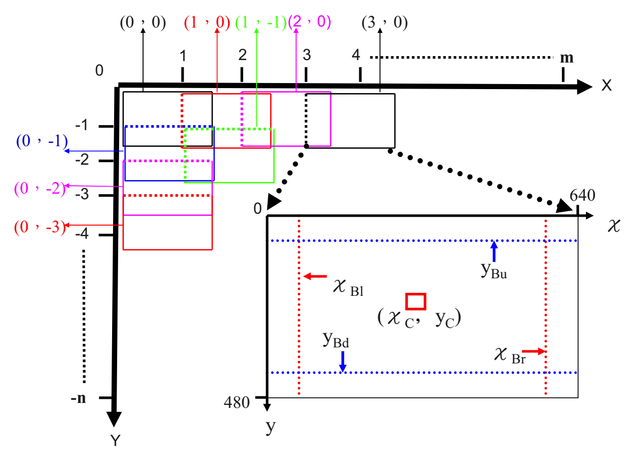


Fig.3

2.4 The best suited camera selection with the vector space

In this framework, the vector space can express which camera is the best suited to track the intruder (Figure 4). If $-45°\leqq\theta<135°$, the best suited camera is camera1. If $135°\leqq\theta<315°$, the best suited camera is camera2.

Based on two contiguous FOVs, a Camera selecting algorithm to get a motion vector and an included angle θ between the motion vector and the x-axis. Combining θ with the vector space will select the camera best suited to track the intruder.

For example, FOV(t)= ($X_{F(t)}$, $Y_{F(t)}$) at time t and FOV(t+1)= ($X_{F(t+1)}$, $Y_{F(t+1)}$) at time t+1. Based on the difference of the coordinates between FOV(t) and FOV(t+1) to get a motion vector = <Vx, Vy> and an included angle θ= $\tan^{-1}$ (Vy / Vx). Hence, the current best suited camera can be selected by the vector space (Figure 5).

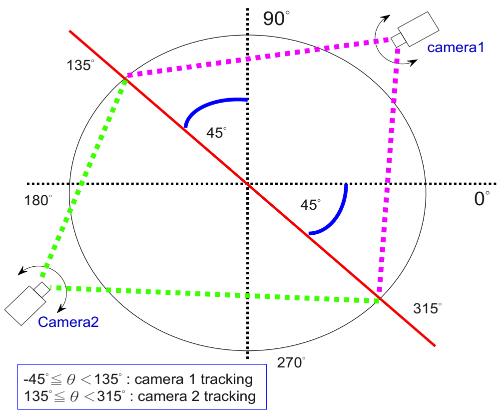


Fig. 4 represents diagrammatically the vector space allowing the best suited camera to be selected according to the angle θ.

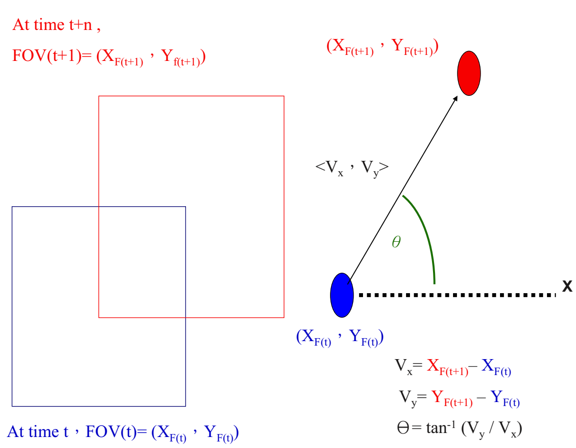


Fig. 5 explains diagrammatically how a motion vector and an included angle θ (between the motion vector and the x-axis) are obtained.

2.5 Scenario

The intruder moves from the left to the right, then turns around and goes back to the origin (Figure 6). When the intruder is moving, the system always selects the camera best suited to track him. Without the loss of generality, camera1 is the slaveCam (monitoring the entrance) and camera2 is the masterCam (tracking the intruder). The top-left rectangle is the current FOV of camera1. The top-middle rectangle is the current FOV of camera2. The small rectangle is the position of the centroid of the intruder. The dotted line is the right boundary of the current FOV of the masterCam (Figure 7).

The intruder moves right and the position of the centroid of the intruder surpasses the right boundary of the

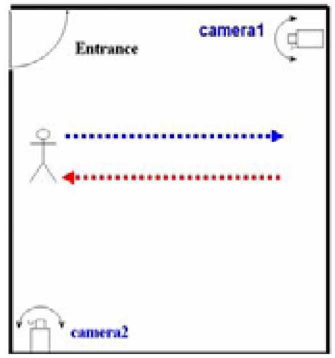


Fig. 6 The path of the scenario.

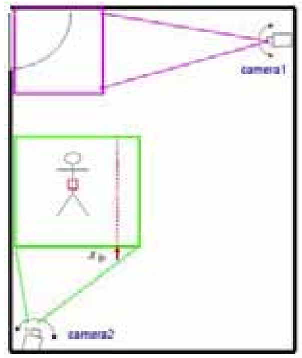


Fig. 7

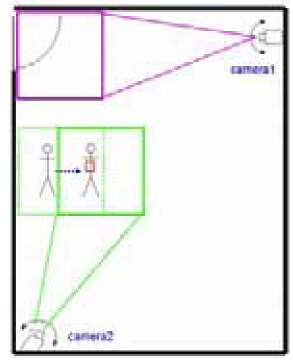


Fig. 8 The masterCam is camera2 just now

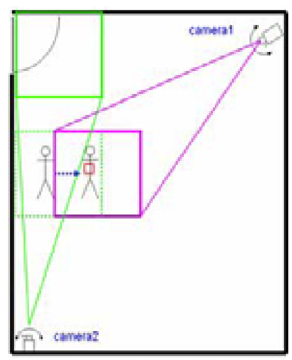


Fig. 9 Hand-off realized

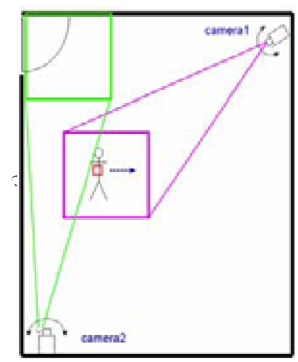


Fig. 10 The masterCam is camera1 just now

current FOV of the masterCam. The masterCam is panned right to the next FOV (Figure 8). Based on the difference of the coordinates between the previous and current FOV, a motion vector and an included angle θ=0° **(**between the motion vector and x-axis**)** are obtained. Due to θ=0°, thecamera best suited is selected by the vector space, which is camera1. Hence hands-off control is achieved (Figure9). Camera2 passes the current FOV's coordinate to camera1 and camera2 changes from being the masterCam to being the slaveCam to monitor any new intruders. At the same time, camera1 changes from being the slaveCam to being the masterCam and turns to the position of the intruder immediately (Figure 10).

3. EXPERIMENT RESULTS

Sample results from the fully automated system that tracks the intruder across dual cameras in section 2.5 are presented here (Figure 11). The input image size is 640x480 and the frame rate is 30 fps using two AXIS PTZ-215 (Pan Tilt Zoom) IP cameras.

Without the loss of generality, camera1 is the slaveCam and camera2 is the masterCam (a). The intruder moves right and the intruder's position of the centroid surpasses the right boundary of the current FOV with the masterCam (b). The masterCam is panned right to the next FOV (c). Based on the difference of the coordinates between the previous and current FOV, a motion vector and an included angle θ=0° between the motion vector and x-axis are acquired. Due to θ=0°, the camera best suited to tracking the intruder is selected by the vector space, which is camera1. Hence hand-off control is realized. Camera2 passes the current FOV's coordinate to camera1 and camera2 changes from being the masterCam to being the slaveCam which monitors any new intruders. Likewise, camera1 changes from being the slaveCam to being the masterCam and turns to the position of the intruder immediately (d). The masterCam tracks the intruder continuously and the slaveCam monitors the entrance continuously if hands-off control fails to happen between (e) and (j).

Until the intruder turns (k), the intruder moves right and the intruder's position of the centroid surpasses the right boundary of the current FOV (l) and the masterCam is panned right to the next FOV (m). A motion vector and an included angle θ=180° between the motion vector and x-axis are obtained. Due to θ=180°, the camera best suited to tracking the intruder is selected by the vector space, which is camera2. Hence hands-off control is achieved. Camera1 passes the current FOV's coordinate to camera2 and camera1 changes from being the masterCam to being the slaveCam which monitors any new intruders. Likewise, camera2 changes from being the slaveCam to being the masterCam and turns to the position of the intruder immediately (n). The masterCam continuously tracks the intruder and the slaveCam continuously monitors the entrance since hand-off control has not been activated between (o) and (y) until the intruder comes back to the origin(z).

4. CONCLUSIONS

In this paper, the key components of a fully automated system that can track the intruder in a dual-camera indoor surveillance scenario is presented. Experimental results have demonstrated that the system can automatically select the most appropriate camera to perform tracking while also deploying the least number of cameras.

5. REFERENCES

[1] de Silva, G.C.; Anavi, S.; Yamasaki, T.; Aizawa, K." A Sensor Network for Event Retrieval in a Home Like Ubiquitous Environment", Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on Volume 4, 15-20 April 2007 Page(s):IV-1393 - IV-1396

[2] Yamazaki, T. "Ubiquitous home: real-life testbed for home context-aware service", Testbeds and Research Infrastructures for the Development of Networks and Communities, 2005. Tridentcom 2005. First International Conference on 23-25 Feb. 2005 Page(s):54 - 59

[3] Yamazaki, Tatsuya; Kayama, Kentaro; Igi, Seiji."Detection of People and Tracking Test Bed Using Multiple Sensors", Multimedia and Ubiquitous Engineering, 2007. MUE '07. International Conference on April 2007 Page(s):135 - 140

[4] Zhou, X., Collins, R.,Kanade, T.,Metes, P."A master-slave system to acquire biometric imagery of humans at distance", ACM International Workshop on Video Surveillance, pp. 113–120. Berkley,USA (2003)

[5] Liu, Q.,Kimber, D., Foote, J.,Wilcox, L., Boreczky, J." FLYSPEC: a multi-user video camera system with hybrid human and automatic control", ACM Conference on Multimedia, pp. 484–492. New York, USA (2002)

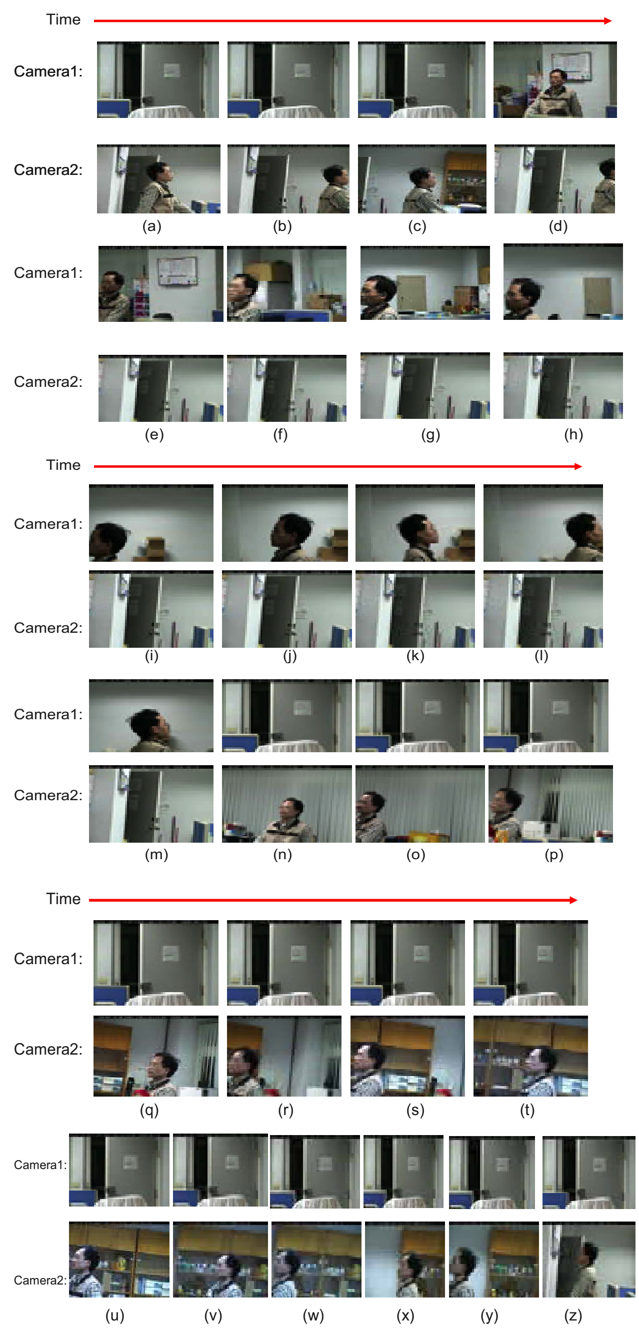


Fig. 11 Sample results of the scenario.

Action Patterns Probing for Dynamic Service Composition in Home Network

Sheng-Tzong Cheng[1], Mingzoo Wu[1,2], and Shih-An Weng[1]
[1]Department of Computer Science and Information Engineering, National Cheng Kung University,
[2]Networks and Multimedia Institute, Institute for Information Industry,
[1,2]Tainan, Taiwan R.O.C.
E-mail: stcheng@mail.ncku.edu.tw, wu.mingzoo@gmail.com, shihan@csie.ncku.edu.tw

Abstract The UPnP architecture defines the communication protocol for networked control point and devices. Exploiting UPnP technologies, home users can easily control intelligent devices through control points. However, these devices lack a composition mechanism to complete a novel application or value-added service. In this paper, we propose a dynamic service composition system to coordinate primitive UPnP services. Action patterns probing (APP) algorithm is introduced to predict the action and the data flow with satisfactory accuracy. Initially, we define data type ontology for UPnP devices to describe their service interfaces. Then service graph construction is used to describe which services can be composed together. After analyzing user's action patterns, we can find the devices which can be composed and worked together in common use. These devices can be composed dynamically by user's habits and can be automated by our mechanism.

Keywords: UPnP, action patterns probing, semantic ontology, service composition, action prediction

1. Introduction

With the growing of home networked devices, all kinds of devices could be discovered and controlled by UPnP [1] architecture. UPnP is architected for pervasive peer-to-peer network connectivity of intelligent appliances, wireless devices, and PCs of all form factors. UPnP enables discovery, event notification, and control of devices on a network, independent of operating system, programming language, or physical network connection. The purpose of UPnP architecture is to provide the communication protocols for control point and devices. UPnP is a XML-based protocol and uses common standards which are independent of the underlying physical media and transports. It promises to be followed and extended the control protocols as needed by UPnP vendors. Many technologies and research initiatives have tried to develop the communication protocol between control point and devices such as Java for Intelligent Network Interface (Jini), Open Services Gateway Initiative (OSGi) [2], Digital Living Network Alliance (DLNA) and Home Audio and Video Interoperability (HAVi) [3]. However, up to now, there are few researches about composing the primitive services to create complex value-added services.

The user-device interacting information can be recorded by capturing the UPnP data streams. Furthermore, these records can be mined as patterns, which can be used by intelligent agent to automate device interactions. Predicting user's actions is an important step for providing intelligent interactions between users and devices at home. However, this knowledge discovery problem is complicated by several challenges, such as excessive noise in the data, data that does not naturally exist as transactions, a need to operate in real time, and a domain where frequency may not be the best discriminator. Current prediction algorithms are usually based on the order of actions to be taken. However, the prediction accuracy of those algorithms is not satisfied because actions can be divided into groups, and actions in the same group are usually taken arbitrarily. These groups should be discovered before the prediction of actions. A series of actions constitute an action patterns. Those algorithms can't dynamically probe the subtle changes either through the time and place.

In this paper, we design a dynamic service composition system with capability of predicting user's actions. We propose an Action Patterns Probing (APP) algorithm, which is mainly based on the information of action patterns and the construction of service graph. The data type ontology is proposed to define the communication interface of a service. The semantic description is used to facilitate the dynamic service composition system. The service graph will be constructed by matching the interface. We collect the information of action patterns and utilize the service graph to produce the rules for prediction. Finally, we predict actions by the rules and make them interact by the service graph.

The rest of the paper is organized as follows. Section 2 reviews the related work of UPnP and action prediction. Section 3 describes the design of our action patterns probing system. In Section 4, we describe the simulation of the APP algorithm. Finally, conclusion remarks are drawn in Section 5.

2. Related Work

UPnP architecture [1] enables the communication between control point and device. Each device specifies its own device description and service description. These XML-based descriptions are provided by UPnP device vendors. Device description specifies the device type, name, ID, manufacturer information, etc. Service description describes the service type, service name, controlling URL, eventing URL, etc.

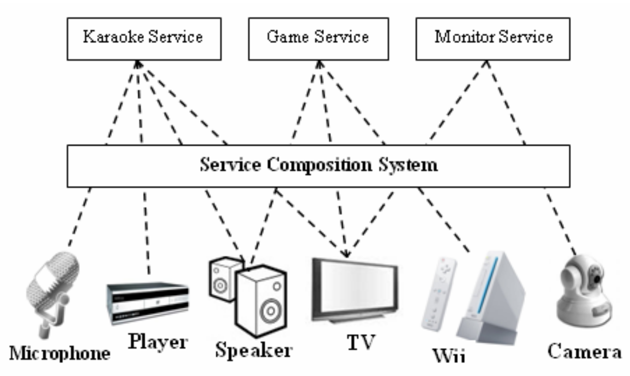


Fig. 1: The concept of dynamic service composition

Common Internet standards are employed by UPnP, such as: UDP, SSDP, SOAP, GENA, HTTP, and XML. Non-IP devices can be accessed by UPnP bridging technology. It is designed to bring easy-to-use, flexible, standards-based connectivity to ad-hoc or unmanaged networks. The UPnP architecture is a distributed, open networking architecture that leverages TCP/IP and the Web to enable seamless proximity networking in addition to control and data transfer among networked devices. It provides the service oriented architecture, which is the base of emerging Web service technology as well and has many advantages such as automatic discovery of services.

The UPnP networking is accomplished through six steps: addressing, discovery, description, control, eventing, and presentation [1].

Online networked services are getting popular in our daily life. These Web services are convenient for easy living. There are some technologies for Web services such as Universal Description, Discovery, and Integration (UDDI) for discovery, Web Services Description Languages (WSDL) for description and Simple Object Access Protocol (SOAP) for communication. With the popularity of Web services, composition of Web services becomes more and more important. Although dynamic service composition is a relatively new topic, several systems that implement dynamic service composition have been proposed. We classify the existing systems into two categories, namely rule-based systems and interface-based systems. Many technologies for Web service composition are suggested and implemented such as Business Process Execution Language for Web Services (BPEL4WS) [4] for process-oriented service composition, CoSMoS [5] using semantic information and eFlow [6] for template-based service composition. With the development of the Web environment, there are more and more applications for Web services.

Some researches have focused on the discovery of significant groups from an action patterns for home automation. Heierman and Cook proposed a data-mining algorithm, called Episode Discovery (ED) [7][8]. They also proposed a Smart Home Inhabitant Prediction (SHIP) [9] algorithm to predict the most likely next action. SHIP matches the most recent actions with a collected action sequence. But the predictive accuracy for data set 1 is 33.3%, and for data set 2 is 53.4% in SHIP. Liao, Lai, and Hsiao proposed an action group discovery (AGD) [10] algorithm. The AGD algorithm is based on the 1-order Markov model and a reverse 1-order Markov model. The GDRate decreases rapidly as the interlace count increases. Therefore, the GDRate approaches zero when the interlace count is greater than two. In real life, if more than two users take the actions at the same time, the AGD algorithm will discovery the action groups incorrectly.

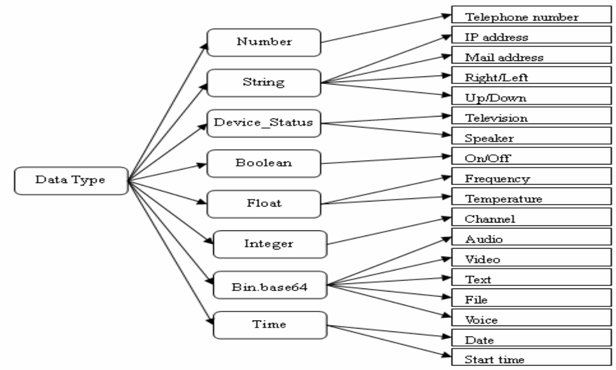


Fig. 2: Data type ontology

3. System Architecture

In this section, we present data type ontology and interface matching method for home appliances to accomplish the action patterns probing. We also present how to construct the service graph and use it to find the interaction rules for virtual devices. Fig. 1 presents the concept of dynamic service composition. These useful services can collaborate with each other and create a virtual device or a novel application. We collect the information of actions including action numbers (unique) and action execution time. We save the information into an AET (Action Execution Table) and transform it into an ARDB (Action Record Database). We also present an action patterns probing (APP) algorithm for processing these information in ARDB.

3.1 Ontology Classification

Merely using data type definition is not sufficient to complete the work of dynamic service composition. Hence the semantic tag is patched for assistance. Each service has its own input and output interface. Each action argument are specified by name, data type, and semantic. And each interface has two variables which are classified according to data type at first and semantics afterward. Each data type is subdivided into one or more parts according to the semantics (as shown in Fig. 2).

Control point may retrieve device and service information with data type and semantics to provide accessible services. If the data type ontology share and publish the same underlying ontology, the control point can extract and aggregate the data type and semantic information to do service matching. The benefit of using data type ontology is easy for developers to design service interfaces and for users to understand.

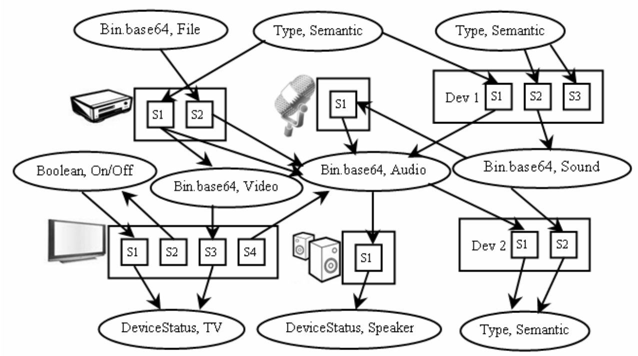


Fig. 3: Directed service graph

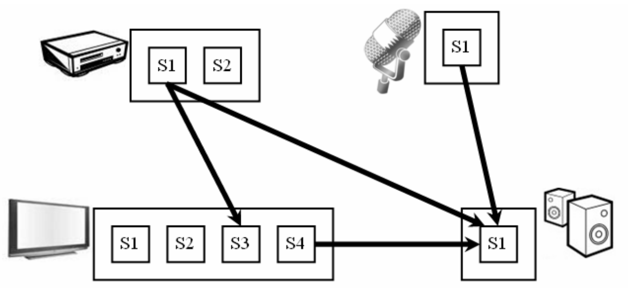


Fig. 4: Reduced service graph

3.2 Service Interface Matching

Service descriptions define all provided actions. These actions have arguments, state variables, data type, range, and event characteristics. A service interface specifies methods that can be performed on the service. Service's interfaces are public for an external use. A service may keep two types of interfaces: input and output.

With the support of data type ontology, data type and semantics of a service can be defined extensively without qualifications. So we extend the service interface with semantics for service interface matching. If service A's output interface is exactly the same with service B's input interface. Then service A and service B can be composed together. With the service interface matching, we can easily tell which services can be composed together for a collaborative task.

3.3 Service Graph

Service graph represents all possible combinations of services. There are two types of node defined in a service graph: Data Semantic Node (DSN) and Service Information Node (SIN). The links between these nodes represents their associations.

- DSN: Let ordered pair (Data type, Semantic) denote a DSN. Each DSN represents a pair of data type and semantics defined by the data type ontology to represents the service's interface.
- SIN: Let ordered pair (Device name, Service name) denote an SIN. One device may have several services, and each service may have input/output interfaces. We take down the device and its service information in SIN. Each SIN represents a pair of the device name and its service name.

Table 1: Definition of Action Patterns Probing Algorithm

Terminology	Description
Action Execution Table	The information of each invoked action will be recorded to the action execution table, including the action number and the execution time.
Action Record Database	Actions with neighboring execution time are grouped as an action block.
Actionset	A set of actions
Occurrence	The occurrence number of actionset X recorded in action record DB.
Frequency	The ratio of the occurrence X to the amount of blocks.
X→Y	Actionset X is the max connected component of X in service graph, so does actionset Y. If there is a link between X and Y in service graph, we say X→Y is established.
Reliance	(Frequency of X→Y)/(Frequency of X)
Min_Freq	Minimum frequency
Min_Reli	Minimum reliance
Candidate Actionset	All actionsets conform to X→Y
Large Actionset	Actionset whose frequency is bigger than Min_Freq in candidate actionsets.

To construct a directed service graph, we add links between DSNs and SINs by interface matching. When a device is discovered by control point, our system checks its data type and semantics of the service interface. If the data type and semantics in the DSN are the same as in the service input interface, a link is produced and directed from DSN to SIN. Otherwise, if the data type and semantics in the DSN are the same as in the service output interface, a link is produced and directed from SIN to DSN.

Once a device is discovered by the control point, our system would create links between DSN and SIN in the service graph. Fig. 3 illustrates an example of directed service graph.

3.4 Action Patterns Probing

We propose an Action Patterns Probing (APP) to find the rules for prediction and the paths of the composite services for execution. We divide APP into three steps: service graph maintenance, action patterns probing, and virtual service creation.

The first step is to maintain the service graph. We show how to construct service graph in the previous section. Before probing action patterns, we have to transform the service graph into a reduced service graph by ignoring the DSN. Fig. 4 shows a reduced service graph from Fig. 3.

The second step is probing action patterns. Table 1 lists terminology used in the APP algorithm. The pseudo code of the Action Patterns Probing Algorithm is shown in Fig. 5.

Fig. 5 shows Action Patterns Probing Algorithm. After getting the reduced service graph, it is beginning to record the action information in Action Execution Table with time going by. Based on the execution time of each action by Action Execution Table, we catch the action pattern of

Action Patterns Probing Algorithm:

Step 1: L_2 = All Large 2-Actionset
Step 2: for ($k = 3$; $L_{k-1} \neq \varnothing$; k++) do begin
Step 3: C_k = Candidate_gen (L_{k-1}) ;
Step 4: for each Action Block t
Step 5: count the Occurrence in C_k
Step 6: L_k = actionset satisfies Min_Freq in Candidate Actionset
Step 7: end
Step 8: return L = all Large Actionset

Candidate_gen Procedure:

Step 1: for each Actionset $X_1 \in L_{k-1}$
/* $G_1(X_1, E_1)$ presents the maximum connected component of X_1 in service graph */
/* Service Graph = $G_2(X_2, E_2)$ */
/* $G1(X_1, E_1) \subset G_2(X_2, E_2)$ */
Step 2: for each Action m $\in X_2$ and $\notin X_1$
Step 3: if ($G_1(X_1, E_1)$ and Action m exist a link $\in E_2 - E_1$)
Step 4: then { $c = X_1 \cup$ m };
Step 5: if $c \notin Ck$
Step 6: then join c to C_k ;
Step 7: else delete c ; }
Step 8: return C_k = all Candidate k-Actionset

Fig. 5: The APP algorithm

Table 2: Simulation parameters

Parameters / Experiment	Length of action pattern	No. of different actions	Avg. confidence	Min_Reli
A	60, 120, …, 600	10, 20, 30, 40, 50	0.8	0.7
B	80, 160, …, 800	30	0.2, 0.4, …, 1.0	0.7
C	500	10, 20, 30, 40, 50	0.1, 0.2, …, 1.0	0.7
D	500	10, 20, 30, 40, 50	0.8	0.1, 0.2, …, 1.0

some time ranges into Action Record DB in order to get the action pattern information to understand what other actions execute before or after the service execute. Then we count the reliance of every probable action pattern. After obtaining the reliance information, we can finally get some reliable action patterns.

4. Simulation

To assess the APP algorithm for action prediction, we calculate the prediction accuracy by the rules generated by APP. Four experiments were designed to measure the prediction accuracy of the APP algorithm influenced by the following parameters. The parameter settings of the four experiments are listed in Table 2. These settings were based on a user's possible behavior in a home.

For each experiment, ten different source data were generated for each combination of parameter values. The final accuracy was the average of the prediction accuracy of the ten source data. The experimental results are shown in the following figures.

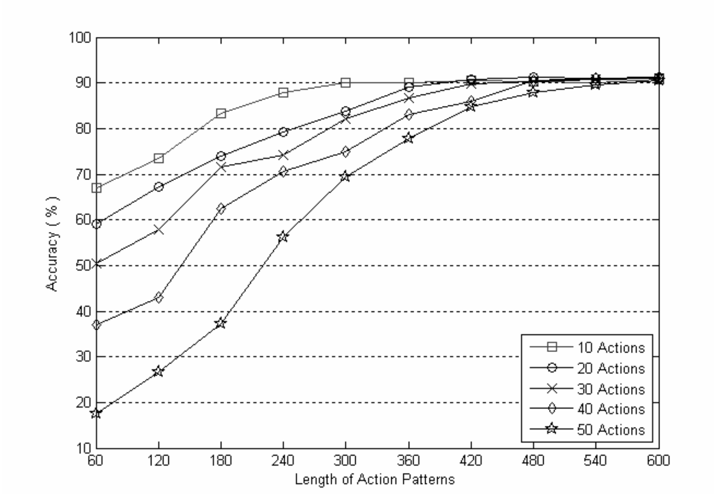


Fig. 6: Experiment A

Fig. 6 shows the result of experiment A. As we can see, the prediction accuracy approaches 90% as the length of action patterns is greater than 500. Although the number of different actions is distinct, the APP algorithm can still predict most of the virtual services with a satisfied value.

Fig. 7 shows the result of experiment B. The prediction accuracy increases with the length of action patterns. The prediction accuracy also depends on the confidence of source data. The higher the confidence, the higher prediction accuracy is obtained.

Fig. 8 and 9 show the results of experiment C and D, respectively. According to Fig. 8, the prediction accuracy approaches 90% when the confidence of source data is greater than 0.8. For various numbers of different actions, the APP can still predict most of the virtual services with a satisfied value.

Fig. 9 tells us that the prediction accuracy approaches 90% when the Min_Reli is greater than 0.7. Although the number of different actions is distinct, the APP algorithm can still predict most of the virtual services with a satisfied value.

The AGD algorithm [10] is based on the 1-order Markov model and a reverse 1-order Markov model. The GDRate decreases rapidly as the interlace count increases. Therefore, the GDRate approaches zero when the interlace count is greater than two. In real life, if more than two users take the actions at the same time, the AGD algorithm will discovery the action groups incorrectly. Fig. 10 presents a comparative graph. According to Fig. 10, the prediction accuracy increases with the length of action patterns. But the prediction accuracy of APP is always greater than AGD. Because in real life, the actions from different groups maybe interlace with more than two users. When this occurs, the prediction accuracy will decrease rapidly. But in APP, we define a time slot to overcome this problem. So the APP algorithm has greater prediction accuracy than the AGD.

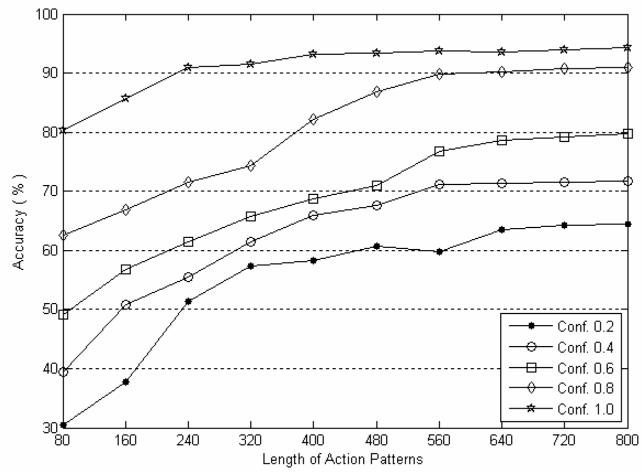


Fig. 7: Experiment B

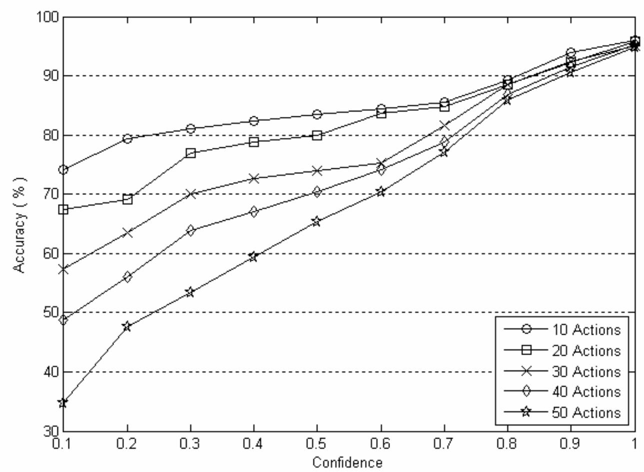


Fig. 8: Experiment C

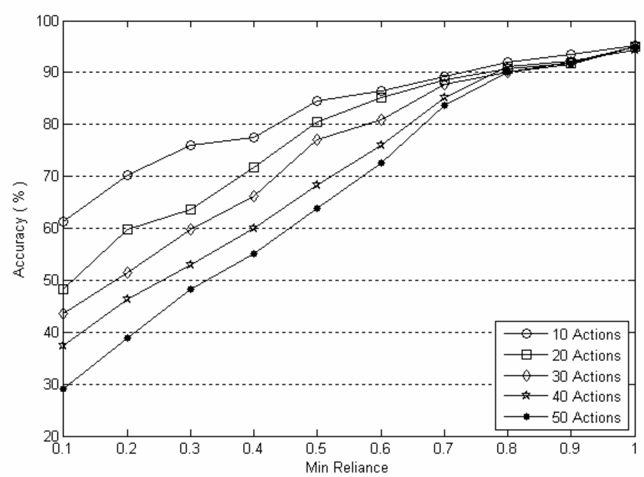


Fig. 9: Experiment D

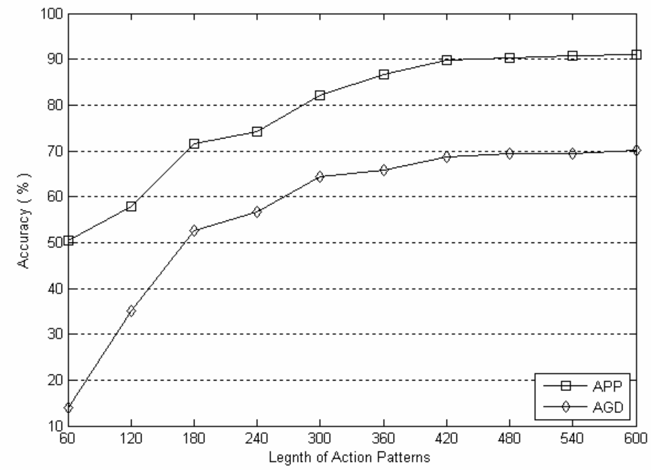


Fig. 10: Performance comparison

5. Conclusion

In this paper, we present how to use semantic descriptions to facilitate the dynamic service composition of home UPnP services. We propose a service interface matching with data type and semantic information. Data type and semantic ontology is easy to design service interface for developers and is easy to understand for users. Service interfaces are public for an external use. With interface matching method, we could know which the service's output can be fed into the next service's input. We also present service graph and APP to find regularly composite execution paths of virtual service. We analyze the record of user's actions by using the service graph. Finally, we can find the devices which can be composed and worked together in common use. These devices can be composed dynamically by user's habits and can be automated by our mechanism. Applications are no longer restricted to designer's imagination and will be automatic in the future.

Acknowledgment

This study is conducted under the "Applied Information Services Development & Integration project" of the Institute for Information Industry which is subsidized by the Ministry of Economy Affairs of the Republic of China.

Reference

[1] Microsoft Corporation, Universal Plug and Play Device Architecture, version 1.0, Jan. 2000.

[2] D. Marples and P. Kriens, "The Open Services Gateway Initiative: An Introductory Overview," IEEE Communications Magazine, vol. 39, no. 12, pp. 110-114, Dec. 2001.

[3] HAVi, "HAVi, the A/V digital network revolution," White Paper, 1999. http://www.havi.org/havi.pdf.

[4] F. Curbera, Y. Goland, J. Klein, F. Leymann, D. Roller, S. Thatte, S. Weerawarana, "Business Process Execution Language for Web Services, v1.0," Jul. 2001.

[5] K. Fujii, and T. Suda, "Dynamic Service Composition Using Semantic Information," International Conference on Service Oriented Computing, Nov. 2004.
[6] F. Casati, S. Ilnichi, L. Jin, V. Krishnamoorthy, and M.-C. Shan, "Adaptive and Dynamic Service Composition in eFlow," International Conference on Advanced Information Systems Engineering, Mar. 2000.
[7] D.J. Cook, M. Youngblood, and E.O. Heierman, "MavHome: An Agent-Based Smart Home," IEEE Int'l Conf. on Pervasive Comp. and Comm., Mar. 2003.
[8] E.O. Heierman and D.J. Cook, "Improving Home Automation by Discovering Regularly Occurring Device Usage Patterns," IEEE Int'l Conf. on Data Mining, Nov. 2003.
[9] S.K. Das, D.J. Cook, et al., "The Role of Prediction Algorithm in the MavHome Smart Home Architecture," IEEE Wireless Communication, vol. 9, no. 6, pp. 77-84, Dec. 2002.
[10] H.-C. Liao, B.-Y. Lai, S.-C. Hsiao, "The Discovery of Action Groups for Smart Home," Journal of Computers, vol. 18, no. 2, Jul. 2007.
[11] S.-T. Cheng, C.-Y. Wang, M. Wu, W.-T. Ho, and C.-M. Chen, "Dynamic Service Composition for Virtual Devices Creation in Home Network Environment," Int'l Conf. on Web Information and Technologies, Mar. 2007.

A Flexibile Contxt-aware Middleware for Developing Context-aware Applications

Po-Cheng Huang
Computer Science & Information Engineering
National Cheng Kung University
huangbc@ismp.csie.ncku.edu.tw

Yau-Hwang Kuo
Computer Science & Information Engineering
National Cheng Kung University
kuoyh@ismp.csie.ncku.edu.tw

***Abstract*—With the appearance of mobile devices and digital appliances, a context-aware middleware becomes increasingly popular in recent years. An efficient context-aware application in context-aware middleware must adapt to variation in context information, such as a user's location, environmental temperature and network bandwidth, etc. Therefore, a context-aware application has to control hardware devices based on context information to achieve the user's requirements. Developing a context-aware application is time-consuming. For resolving this problem, a context-aware middleware is proposed to assist developer in designing a context-aware application. In this paper, a context-aware middleware is proposed to gather context information form sensing device and provide a convenient interface to control the service device. Based on the context-aware middleware, the developer can alleviate a burden to program a context-aware application.**

***Keywords*—context-aware middleware, context-aware application**

I. INTRODUCTION

The emergence of pervasive computing technologies provides "anytime, anywhere" computing by decoupling users from devices and viewing applications as entities that perform tasks on behalf of users [1]. Developing a context-aware application is complex and time-consuming due to lack of an appropriate context-aware middleware. In order to alleviate a burden of developing context-aware application, many researchers devote to the development of context-aware middleware.

Due to the mentioned reasons, a context-aware middleware has become a key role in developing a context-aware application. The important factors, when designing a context-aware middleware, are simplicity, flexibility, extensibility and expressiveness [2]. Therefore, the processing of context-aware middleware on context information should be as efficient as possible. Furthermore, the context-aware middleware should supply a flexible and extensible way to acquire the context information and verify the context information. For resolving this problem, in this paper, we attempt to construct the ontology-based context model to assist context information storage, analysis and utilization.

The remainder of this paper is organized as follows. In Section II, the related works of context-aware middleware and the context model are illustrated. In Section III, we present the architecture of the feasible context-aware middleware. In Section IV, the ontology-based context model for smart home is proposed. In Section V, the implement process of context-aware application is illustrated. Finally, we will conclude the result in section VI.

II. RLEATED WORK

In the past year, there were lots of theoretical and empirical works presented to demonstrate the premium of context-aware applications. Service-Oriented Context-Aware Middleware[3] (SOCAM) is an architecture with a central server called Context Interpreter, which gains context information through both external and internal Context Providers and offers it in mostly processed form to the clients. The context-aware mobile services in SOCAM adapt various environments according to context information.

Context-Awareness Sub-Structure [4] (CASS) as same as SOCAM is an extensible centralized middleware composed of Interpreter, Context Retriever, Rule Engine, and Sensor Listener. Sensor Listener listens for context updates from sensors nodes and stores the gathered information in the database. Context Retriever is responsible for retrieving stored context information. Through Rule Engine, context-aware service can adapt various environments.

Gaia [9] brings the functionality of an operating system to physical spaces. The extended concept forms GaiaOS to supports mobile, user-centric active space applications. The kernel consists of a component management core for component creation, destruction and uploads, with currently seven services built on top of it. Application Framework uses a MPCC (model, presentation, controller, and coordinator) pattern, which is an extension of the traditional MVC (model, view, and controller) pattern.

Context Broker Architecture (CoBrA) [6] is an agent based context-aware system in intelligent space. Context Broker is the core functional entity which stores context information and performs context reasoning. Agents in intelligent space play several roles, including applications, hardware services (e.g., projector service, light controller, and temperature controller), and Web services (e.g., services keeping track of users and items).

Context Toolkit[5] proposed by Salber, Dey and Abowd is a peer-to-peer architecture with centralized discoverer where distributed widgets (sensor units), interpreters, and aggregators are registered in order to be found by client applications. The toolkit provides object-oriented API for context-aware application developers to create their own components.

Semantic Space [10] is a pervasive computing infrastructure that exploits Semantic Web technologies to support explicit representation, expressive querying, and flexible reasoning of contexts in smart spaces. The context infrastructure consists of several collaborating components: wrappers, an aggregator, a knowledge base, a query engine, and a reasoner. Context wrappers use the UPnP general event notification architecture (GENA) to publish context changes as events to which consumers can subscribe. Context aggregator is implemented as an UPnP control point that inherits the capability to discover context wrappers and subscribe to context events. Context-aware applications use RDQL (RDF Data Query Language) to query context information stored in Context knowledge base.

Most of the context-aware middlewares [3][4][9] require one or more centralized components to enrich the system functionalities (e.g., resource discovery, historical context data, etc.). On the other hand, some peer-to-peer context-aware systems [6][5][10] gather context information from distributed devices by using the context sharing protocol, which complicates the system design and implementation. However, peer-to-peer context-aware system can avoid the failure caused by the breakdown of key components. In order to increase the efficiency of context sharing and reduce the burden of small sensors or actuators, our context-aware middleware adopts distributed agent architecture with some centralized components that give assistance in sharing context information.

III. ARCHITECTURE OF CONTEXT-AWARE MIDDLEWARE

Since building a context-aware middleware requires integrating a large amount of resource and information distributed in the network, we propose a multi-agent middleware named Context-Aware Digital Home Application Framework (CADHAF)[7][8] to ease the burden of software application developers. CADHAF in Fig.1 is divided into six layers: Device Layer, Component Service Layer, Service Abstraction Layer, Service Script Layer, Kernel Layer, and Personal Agent Layer. In this section, we explain these six layers in detail.

A. Device Layer

Device Layer is the bottom layer of the system. It contains two types of hardware devices: sensors and actuators. Sensors are used to detect the environment status (e.g., temperature, lightness, and the signal strength of wireless access point, etc) and then dispatch these data to Component Services. Furthermore, actuators are used to provide users with system feedbacks including sound, light control, and multimedia output, etc. In our context-aware middleware, each device at least connects to one Component Service and acts according to the commands from Component Service.

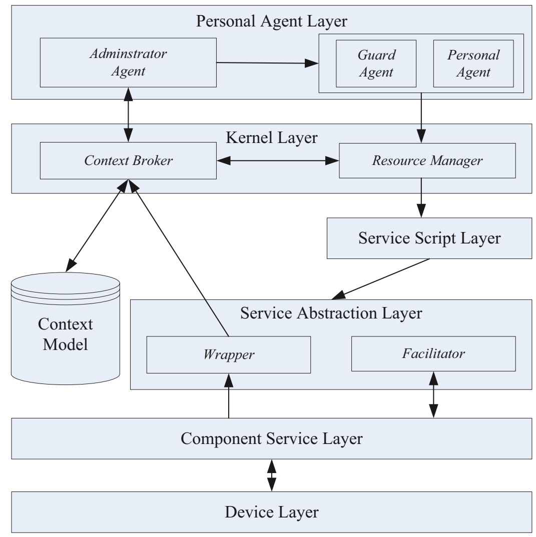


Figure 1 Architecture of CADHAF

B. Component Service Layer

A component in Component Service Layer is the tiniest entity to provides basic functionalities such as text to speech synthesis, automatic speech recognition service, and object tracking service, etc. From the viewpoint of their behaviors, Component Services are divided into two categories. One actively receives sensor data from sensors and transforms them into useful context information. The other passively accepts the requests of other components and exploits appropriate actuators to complete the requests.

C. Service Abstraction Layer

Service Abstraction Layer is composed of two kinds of agents: Wrappers and Facilitators. This layer provides a concept of abstraction to decouple application development from context acquisition or actuator control.

1) Wrapper

Wrappers connecting Component Services to gather contexts (e.g., temperature at 32° C, or wireless access point signal strength at -20dBm), process them if necessary, and then insert data into context knowledge base for the convenience of other agents requiring necessary information.

2) Facilitator

Facilitators partially realize Service Scripts to accomplish user's service. After receiving the command issued by Service Scripts, they connect to appropriate

Component Services to do the corresponding actions which Service Scripts require.

D. Service Script Layer

The major functionality of Service Script Layer is to assist in context-aware application designs. A designed context-aware application describes the scenario goal we want to achieve (i.e., defines the expected environment status and the hardware reaction to users after application execution). Service Scripts order a set of proper Facilitators to complete composed services via Agent Communication Language (ACL). All Service Scripts are independent agents which are able to be executed simultaneously. Application developers can design their Service Scripts separately.

E. Kernel Layer

Kernel Layer comprises two central entities including Context Broker and Resource Manager to accommodate and facilitate the interoperation between components.

1) Context Broker

Context Broker mainly provides the interface to insert, query, and reason the related information about users and environment (e.g., user profile, location, activity, and temperature). The contexts stored in the context knowledge base via Context Broker are able to be accessed via the same unique interface. To reason about high-level context information, we use Web Ontology Language (OWL) to construct context model. With this context model and the reasoning engine Jena2, Context Broker has the ability to infer high-level contexts based on first-order logic.

2) Resource Manager

Resource Manager is responsible for the resource management of the system. When a Service Script asks for resources, Resource Manager judges from the utilities of resources to check whether there are enough resources to complete the service. If the resources are not sufficient, Resource Manager rejects the request of the Service Script.

F. PA Layer

1) PA

In the proposed framework, PA(PA) is assigned to serve each user individually. PA decides which Service Scrip to invoke according to user's expectation, user profiles and environment status from context knowledge base. When PA detects the trigger event of the specific Service Script needed by the user, PA sends a resource request to Resource Manager to obtain the execution permission. Different users own different access rights to different available Service Scripts. For example, the children are not allowed to enjoy movies in the midnight.

2) Administrator

Administrator is a PA for a special purpose but not a general PA serving a home member. Administrator owns a special Service Script named Registry Scenario which is activated when a new home member wants to register to the system. Registry Scenario applies TTS Facilitator, ASR Facilitator, and Registry UI Facilitator to help the user input the necessary information during registration. Once the user completes registration, user profile is stored in context knowledge base for later use. Administrator sets up a PA for every user according to his role and preference. Administrator also creates a virtual Guard Agent introduced in next paragraph for each room.

3) Guard Agent

Besides PA serving for home members and Administrator responsible for the creation of PAs, Guard Agent is also one of the special PAs in this layer. Guard Agent plays the role of the guard watching a room. For instance, the light will be turned off automatically by the Guard Agent when nobody is in the room. In addition, Guard Agent provides the functionalities of speaker identification and intruder alarm.

IV. CONTEXT MODEL

In order to provide an adaptive context-aware application, there are many researches working on the context model of context-aware middleware. The context model in context-aware middleware is proposed to assist the developer in gathering the context information. In our context-aware middleware, a flexible context model is designed with ontology architecture composed of context entities. The ontology-based context model is shown in Fig.2. Conceptually, the context model is designed for storing the environment situation and inferring the context information. Therefore, context-aware applications can access context information which they want and react appropriately to the different situation relying on the different stimulations of context information.

A. Description of Sensing Data and Context Information

In our architecture, the sensing data provided by sensing devices can be described as $<P,R>$. R means the sensing device. P means the value of sensing data which provide by sensing device. Furthermore, the context information is described as four-tuples $<O,P,R,T>$ including object, context provider and timestamp of context information. The detail structure is described below.

Notations:

O: Object which context information describe about

P: Property of context information

R: Context provider which provide context information

T: Timestamp of context information

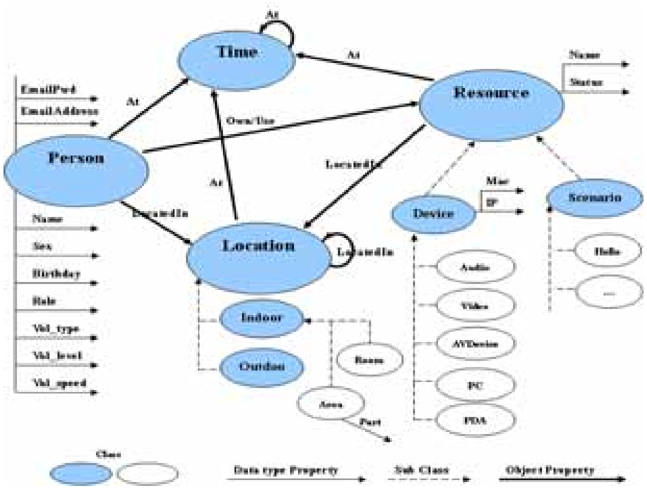


Figure 2 Context Model

The context information of context model is divided into two different categories. The first category of context information which is provided directly by sensor device or actuators devices is named low-level context information. Low-level context information means that the object O is equal to the context provider R. On the other hand, the other context category of context information which is inferred according to the low-level context information is named the high-level context information. It means that the object O is different from the context provider R. The context information of our context-aware middleware are transferred to four-tuples and stored in database. When context-aware application needs specific context information, it can query the context information through context broker.

V. Implementation of Context-Aware Application

In this section, we implement a context-aware application of music service to a specific person according to his location information. The context-aware application uses the location wrapper to collect the location information which provide by PDA. The location wrapper component transfers the sensing data to context information and store the location information in the smart home ontology through the context broker.

The administrator agent who got the personal information through the context broker products the PA according to the personal profile. The PA checks the event which product by location wrapper. If there were the speaker around the personal location, the PA will trigger the Roaming Media Scenario.

According to the Roaming Media Scenario, we measure the time cost of every single step during the conversation between users and CADHAF. The time cost of the interaction between CADHAF and the user are described as follows. After the user issues the command to enjoy the music, the automatic speech recognition subsystem takes 1.2s to identify the user's command which is inserted into Smart Home Ontology through Context Broker. The Roaming Media Scenario activated by PA applies TtsAgent to synthesize and play speech sounds "Which song do you like?" which takes 3s including 0.1s used by Text To Speech Server, 0.9s spent by File Sender, and 2.0s consumed by Sound Player. Roaming Media Scenario also exploits AsrAgent to recognize the title of the music the user requests. This step takes 10s including recording (3.8s) and recognition (6.2s) phase. Again, TtsAgent takes 3s to inform the user the classical music named Swan Lake begins to play. Media Player takes additional 2s to stream the multimedia. Totally, it takes 20.1s from the command issued to the service provided, as shown in Table 1.

Table 1 Process of Roaming Media Scenario

Step	Actor	Action	Time cost (sec)
1	User	[Music!]	0.8
2	ASR	Speech recognition	1.2
2.1		Single keywords recognition	1.2
3	TTS	[Which song do you like?]	3.0
3.1		Speech synthesis	0.1
3.2		Audio file transmission	0.9
3.3		Audio file output	2.0
4	User	[Swan Lake.]	3.8
5	ASR	Speech recognition	6.2
5.1		Audio file transmission	1.1
5.2		Multiple keywords recognition	5.1
6	TTS	[Here comes Swan Lake.]	3.0
6.1		Speech synthesis	0.1
6.2		Audio file transmission	0.9
6.3		Audio file output	2.0
7	Media Player	Multimedia streaming	2.1
			Total: 20.1

The result shows that most time is spent in speech recognition and speech synthesis.

VI. Conclusions

In this paper, a context-aware middleware (CADHAF) is proposed to assist the development process of context-aware applications. CADHAF is contributive in collecting context information form sensing device and provide a convenient interface to control the service device. Based on CADHS, the developer can alleviate a burden to program a context-aware application. Besides, a context model is also beneficial in improving the flexibility of context information.

Although CADHAF helps the effective burden reduction of developers in constructing context-aware applications, the issue of privacy protection mechanism has not been explored. To deal with this issue, we can attach additional attribute called context owner to contexts. We will attempt to define a set of access rules to protect the personal privacy in the future.

ACKNOWLEGMENT

This paper is based partially on work supported by the National Science Council (NSC) of Taiwan, R.O.C., under grant No. NSC96-2221-E-006-017.

REFERENCES

[1] K.H., J.I., and A.R., “Infrastructure for Pervasive computing: Challenges,” Workshop on Pervasive Computing Informatik 2001, Vienna, Sept. 25-28, 2001

[2] M. B. and S. D., “A Survey on Context-aware system,” Technical University of Vienna Information System Institute Distributed System Group, Nov. 2004

[3] T. Gu, H. K. Pung and D. Q. Zhang, “A Middleware for Building Context-Aware Mobile Services,” in proceeding of Vehicular Technology Conference, May 17-19, 2004

[4] P. Fahy and S. Clarke, “CASS - Middleware for Mobile Context-Aware Applications,” Workshop on Context Awareness at MobiSys, June 2004

[5] D. Salber, A. K. Dey and G. D. Abowd, “The Context Toolkit: Aiding the Development of Context-Enabled Applications,” in proceeding of CHI’99, May 15-20, 1999

[6] H. Chen, T. Finin, and A. Joshi, “An Intelligent Broker for ContextAware Systems,” in proceeding of Ubiquitous Computing 2003,

[7] H. T. Cho, P. C. Huang, R. H. Luo and Y. H. Kuo, “A New Context-aware Application Validation Method Based on Quality-driven Petri Net Models,” Second International Conference on Innovative Computing, Informatio and Control (ICICIC 2007), pp. 200, Sept. 2007

[8] P. C. Huang, Y. H. Ye, M. F. Horng and Y. H. Kuo, “An Ontology-based Context Model in Smart Living Space,” 3rd International Conference on Soft Computing and Intelligent Systems and 7th International Symposium on advanced Intelligent Systems, Nov. 2006

[9] R. M., H. C.K., C. R., R. A., C. R.H. and N. K., “Gaia: A Middleware Infrastructure to Enable Active Spaces,” IEEE Pervasive Computing, 1(4), Oct.-Dec. 2002

[10] X.H. Wang, J.S. Dong, C.Y. Chin, S.R. Hettiarachchi and D.Q. Zhang, "Semantic Space: An Infrastructure for Smart Spaces, " Pervasive Computing, IEEE Vol. 3, Issue 3 , pp.32-39, July-Sept. 2004

ARCATS - Tackling State Explosion in a Multi-Phased, Compositional Way

Yung-Pin Cheng   Chun-Yu Lee

Dept. of Computer Science and Information Engineering
National Taiwan Normal University
Taipei,TAIWAN
ypc@csie.ntnu.edu.tw

Abstract

Automatic verification techniques, which analyze all processes at once, typically do not scale well for large, complex concurrent software systems because of the theoretic barrier – PSPACE *hard complexity in worst case. In this paper, we present our tool named* ARCATS *(Architecture Refactoring and Compositional Analysis Tool Suite).* ARCATS *consists a set of tools to combat state explosion in a divide-and-conquer, hierarchical fashion. These tools can be applied in a multi-phased manner until a balance between intractability and feasibility is reached. We build these tools to seek out perfect combinations for analyzing large-scale, complex software system with state space growth carefully controlled.*

1. Introduction

Automatic verification techniques such as model checking have been viewed as a promising method to ensure the quality of complicated systems, particularly complex software systems. Many hard-to-detect errors, such as deadlocks, can be manifested by these techniques. In past decades, considerable progress has been made in these techniques. Several prototype tools, such as SPIN[8][9], SMV[10], have been built and applied to many software examples.

Model checking techniques typically deal with software that involves concurrency. So, the beginning steps of verification tools often involve in a parallel composition to explore the reachable states of target systems. This computation, unfortunately, has PSPACE lower bound in worst case. As a result, model checking tools which analyze all processes/threads at once, typically do not scale well. Excessive reachable states soon exhausts the memory resources, which is called *state explosion*. Some reduction methods such as partial order [6] and symmetry can alleviate the state explosion for particular class of problems but are not effective in general.

Because of the theoretic lower bound, we believe approaches to increasing the size of system that can be accommodated in a single analysis step must eventually be combined with effective compositional techniques [12, 4, 7] that divide a large system into smaller subsystems, analyze each subsystem, and combine the results of these analysis to verify the full system. So, the major focus of this paper is to present a tool suite which facilitates compositional analysis of concurrent software. The goal of ARCATS (Architecture Refactoring and Compositional Analysis Tool Suites) is to provide an alternative to global analysis tools like SPIN. When global analysis tools fail to scale on large and complicated systems, ARCATS provides a series of compositional methods to continue the battle.

In this paper, we present a tool suite called *ARCATS*. ARCATS provide combined methods of compositional analysis and simulation in a multi-phase manner to seek out a balance between intractability and feasibility. These phases are:

1. Compositional analysis
2. Compositional analysis + refactoring [3]
3. Simulation
4. Simulation + compositional analysis

When verification tools like SPIN fail to provide feasible answers, ARCATS provides a set of tools to facilitate compositional analysis. However, some systems' as-built structures may not be amenable to compositional analysis. In this cases, refactoring tools [3] can be used to change the system structures into ones that enable compositional analysis.

If compositional analysis with refactoring fails again, it is time to compromise with the intractability. A simulation tool is provided in ARCATS, which allows randomly walking in the state space to detect violation of safety properties. The random walks in state space may not necessarily detect the violations but if more time is spent, more confidence is gained.

If the degree of confidence needed is high, results of the compositional analysis can be used to enhance the simulation coverage. In this phase, ARCATS's simulation tool can walk from individual processes (in a form of control flow graph) into the interface processes which represent the observational behaviors of a subsystem. In principle, the walk into the states of an interface process is equivalent to the walk of states in a subsystem. In a subsystem, many internal states are of no interests to the subsystem's environment. As a result, the depths or steps to find the violations can be reduced. Much more confidence can be gained from the simulation.

In this paper, the techniques in each phase will be described respectively.

2. Architecture and Meta Data Structure of ARCATS

ARCATS consists of a set of tools. These tools read and write files to communicate and cooperate. Fig. 1 shows the overall architecture of ARCATS. ARCATS starts with input languages on the top. Currently, ARCATS support Promela, SPIN's modeling language, as input language. In the future, it is planned to support Java-like modeling language. The boxes in gray color are the tools of ARCATS.

The parser to parse a Promela file is called a ADT-promela parser. ADT stands for Abstract Data Type. More higher level language constructs such as queue and set are supported by ADT-Promela to address the automatic refactoring problem[2].

2.1 Control Flow Graph

The most important files generated by the language parsers are the control flow graph (CFG) files. One CFG file will be generated for each process/thread specified in input files. Each directed edge is attached with an abstract syntax tree (AST) of the statement that makes the transition. The statement AST may contain variables or constant. Their types or values can be referred by the symbol tables, which are also generated by the parsers.

The CFG is meant to be executable. Tools like simulation engine may set up a symbol table and work as an interpreter to run the CFG. When a CFG edge is executed, the statement's AST is evaluated and variables in the symbol table are modified. A statement in CFG can be a regular statement or a communication/synchronization statement. Statements such as assignment statements are regular

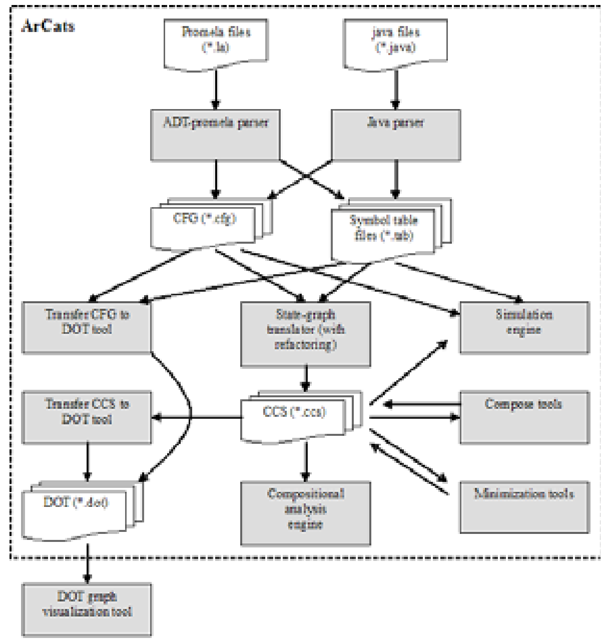


Figure 1. The architecture of ARCATS.

statements. Their executions only change the values of variables in symbol table. A communication/synchronization statement involves interactions across processes/threads. In Promela, a communication/synchronization statement is either *"ch?msg"* or *"ch!msg"*, where *'ch'* is a channel name, *'msg'* is the message to be transmitted, '?' stands for receiving end, and '!' stands for sending end. When these statements are evaluated, the simulator establishes a channel data structure in the symbol table and shares it between processes.

This statement AST, unfortunately is language dependent, but the CFG execution algorithms should be language independent. Some object-oriented efforts have been made to address the extension and evolution problem.

2.2 CCS State Graph

Control flow graph can not be used as a formalism for verification. A control flow graph is still equivalent to a program. The behaviors are still parameterized by variables. Verification relies on some forms of communication finite state machines such as CCS [11] or CSP. In ARCATS, Minler's CCS (with two-way rendezvous synchronization) is adopted.

In a CCS state graph, an edge label is a constant label called *action* which is paired with another constant label in other processes. Two paired labels are denoted by a and $\bar{a}$ in CCS. When two paired actions are fired in two processes, they communicate. CCS's synchronization is equivalent to channel of size zero. To model a channel of size greater than 1, an individual process modeling the channel behaviors should be added to the system. In ARCATS, Promela's syntax is followed, *"ch?msg*" and *"ch!msg"* are treated as paired actions if *'ch'* and *'msg'* are the same.

A tool called *cfg2ccs* in ARCATS is used to translate a CFG into a CCS. During this translation, each variable value is symbolically expanded and any transition in CFG which does not involve communication is hidden or modeled as a τ event (an internal action) in CCS graph. In CCS state graph, the edge label is a constant string, where variables are replaced by symbolic instances.

Once a process is translated into a CCS graph, they can be used for computing Cartesian product to construct the reachability graph. They can also be used for compositional analysis, which will be described later. In ARCATS, all the graphs such as CFG and CCS can be translated into a DOT format to be visualized by the *graphviz*[5].

3. Multi-Phase Solutions to State Explosion

The state explosion problem is an old problem but remains as a major barrier for verication tools. In ARCATS, different levels of techniques are implmented to combat the problem. When state explosion occurs, ARCATS provides a series of methods to tackle the problem. In this section, we describe these methods consecutively.

3.1 Global Analysis

In global analysis, each state of a process/thread will participate the parallel composition of the global state space, i.e., the Cartesian product of all processes. As a result, two independent transitions which do not have synchronizations between them appear as interleaving in the global state space. ARCATS can compute pure Cartesian products for the given processes but has not supported well-known analysis algorithms such as LTL. In principle, some well-known reduction method – partial order reduction can be used to reduce the reachability states generated. When the system size is small, well-known tools such as SPIN can complete the job. The strength of ARCATS is not at global analysis.

3.2 Compositional Analysis

In principle, parallel composition operators are commutative, such as "|" in CCS. Suppose a system is composed of *(A|B|C)*. Instead of composing three processes at once, we can compose the system gradually as *((A|B)|C)*. However, doing so does not have particular merits in saving memory resources. The computation of *(A|B)* without the context of C can sometimes generate more states than *(A|B|C)* [7]. Eventually, each state and transition in $(A|B)$ will participate the parallel composition with C. The result is the same as *(A|B|C)*.

In $(A|B)$, there are transitions which no longer wait for the communications from C. These transitions are called *internal actions* (labeled as τ in CCS) with respect to $(A|B)$. On the other hand, there are transitions in $(A|B)$ which wait for the transitions in C to synchronize. We call these *external interfaces* of $(A|B)$. The goal of compositional analysis is to replace $(A|B)$ in the parallel composition of $(A|B|C)$ with a simpler one. Let the simplified one be $s(A|B)$, which is obtained from $(A|B)$ via some transformation s. $s(A|B)$ typically has fewer states than $(A|B)$. As a result, in compositional analysis, we compute $(s(A|B)|C)$ to avoid state explosion in a divide-and-conquer manner.

In principle, s can be some arbitrary computation, as long as the property of interest is preserved. In practice, s is state space minimization methods like bisimulation minimization. Weak bisimulation minimization (WBM) and branching bisimulation minimization (BBM) [1] are two methods we use for s. In WBM or BBM, states that are functionally equivalent are merged into a single state. The result of $s(A|B)$ often contains only behaviors that are observable outside, i.e., the interfaces. So, we call it the *interface process* of $(A|B)$.

In a compositional analysis, we often group a set of processes into a subsystem (or a module). There are two basic criteria of a "good" subsystem. First, the processes inside the subsystem must not generate excessive state space. Second, the subsystem's state space must be able to be replaced by a much simpler interface process to represent the subsystem's state space. So, simple interface is the key to a "good" subsystem. In other words, an effective subsystem should be *loosely coupled* to its environment so that the chance of having a simple interface process to replace it in compositional analysis is higher. At last, "good" subsystems and processes must produce another larger "good" subsystem in the composition hierarchy until the whole system is analyzed.

The power of divide-and-conquer can be enormous if the structure of a system is suitable for such approach. For example, compositional analysis can analyze up to thousands of dining philosophers because they synchronize in a ring structure.

3.3 Compositional Analysis + Refactoring

In practice, compositional techniques are inapplicable to many systems (particularly the large and complex ones) because their as-built structures may not be suitable for compositional analysis. A structure suitable for compositional analysis must contain loosely coupled components so that every component can be replaced by a simple interface process in the incremental analysis. Besides, the processes in a component must not yield intractable analysis. Otherwise, we need to recursively divide the component into smaller loosely coupled components until every subsystem in the composition hierarchy can be analyzed. However, an ideal structure like that seldom exists in practice. Designers often structure their systems to meet other requirements with higher priority. It is impractical to ask designers to structure a design in the beginning for the purpose of obtaining correctness.

If it is difficult to prove the correctness of a program under its original design, one may need to prove the correctness on a transformed, equivalent version of the program. This is a notion known as program transformation, which has been widely studied in the area of functional and logic languages. In object-oriented design, the "refactoring" is a technique which gradually change the system structure through refactoring steps but without destroying original program behaviors.

Here, we apply the idea to transform finite-state models to aid automated finite-state verification. In general, the purpose of our transformations is for obtaining, starting from a model P, a semantically equivalent one, which is "more amenable to compositional analysis" than P. It consists in building a sequence of equivalent models, each obtained by the preceding ones by means of the application of a rule. The rules are aimed for restructuring the as-built structures which are not suitable for compositional techniques. The goal is to obtain a transformed model whose structure contains loosely coupled components, where processes in each component do not yield state explosion.

In this subsection, we use a simple example to briefly show how refactoring is generally work. In Fig. 2(a) and Fig. 2(b), we show the state graphs of three example processes *X,Y,* and *S* in CCS semantics [11] (where synchronization actions are matched in pairs) and their synchronization structure. Such kind of structure, a star-shape structure, appears very often in practice, for example, a stateful server which communicates with clients via separate (or private) channels. Many systems can even have structures of multiple stars.

We say *S* is *tightly coupled* to its environment (which consists of *X* and *Y*) because it has complicated interfaces to its environment. Suppose *S* is a server and *X,Y* are clients. Image the number of clients is increased to a larger number. Any attempt to include *S* as a subsystem is bound to fail because of the complicated interfaces to its environment. That is, no feasible subsystems and composing hierarchies exist in this structure, particularly when client number is large.

The key transformations are to decompose centralized, complicated behaviors of a process into several small new processes while behavioral equivalence is preserved. In [3],

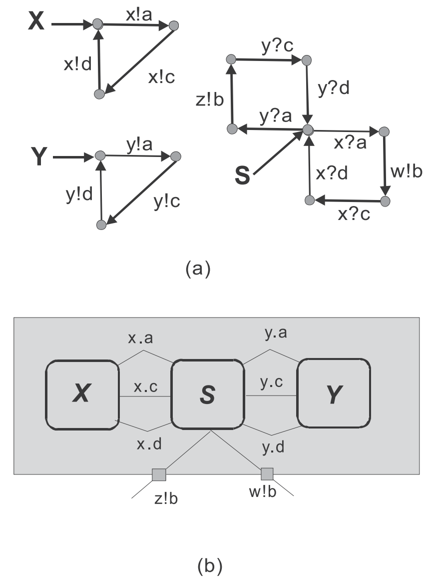


Figure 2. (a) A simple example with 3 processes X, Y, and S. (b) The synchronization structure of the example.

we described the basic tool support[1] for refactoring and showed that a refactored elevator system can be analyzed up to hundreds of elevators but global analysis and compositional analysis (without refactoring) can only analyze up to 4 elevators. A refactored Chiron user interface system can be analyzed up to 14 artists (17 processes in total), whereas non-refactored one can only be analyzed up to 2 artists (5 processes in total).

For instance, we show the refactored *X,Y,* and *S* in Fig. 3(a) and the new synchronization structure in Fig. 3(b). In Fig. 3(a), the behaviors related to channel *x* (or to process *X*) is removed and wrapped into a new process *Sx*. Similarly, the behaviors related to channel *y* is removed and wrapped into a new process *Sy*. So, the rendezvous of *x!a, x!c,* and *x!d* are now redirected to *Sx*. However, *Sx* and *Sy* are now two individual processes which can execute concurrently, but their original joint behaviors in *S* can not. So, extra synchronizations (*e!lock and e!release*) are inserted to maintain behavioral equivalence; that is, before invoking *x!a* and *y!a, X* and *Y* are forced to invoke *e!lock* first. Then, at the end of *Sx* and *Sy*, *e!release* is used to free *S*.

The idea of refactoring equivalence is easy to explain. Let's image the modified processes (*X,Y*, and *S*) are contained in a black box. Image you are an external observer of the black box. The external behaviors of the black box are defined by *z!b* and *w!b*. In Fig. 2(b), the black box (which

[1]The tool support can successfully refactor many systems in an automated fashion, particularly the behavioral patterns which do not involve complicated data structures.

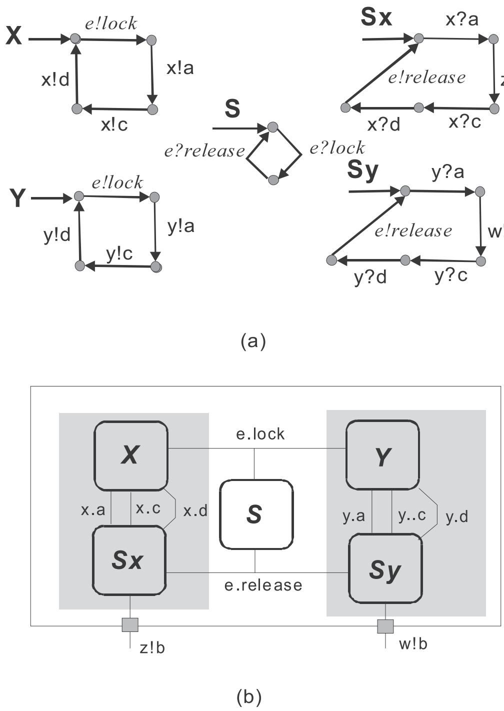


Figure 3. The refactored example system.

we call it B1) is implemented by 3 processes. The black box (we call it B2) in Fig. 3(b), on the other hand, is implemented by 5 processes. The external behaviors are also defined by *x!b* and *y!b*. Our refactoring must ensure the external behaviors are equivalent before and after a transformation. Intuitively, B1's external behaviors can be viewed as an specification. Then, we choose to implement the specification with 5 processes. Since we use 5 processes to do the same work which was originally done by 3 processes, extra communications for process coordination are inevitable. As long as the extra synchronizations are restricted inside the black box, the two black boxes behave equivalently to an external observer.

3.4 Simulation

When the above methods fail to assure properties of interest, simulation may still be able to detect property violations, such as safety errors like deadlocks. The power of simulation is equivalent to testing. It cannot guarantee properties like absence of deadlocks, but it can detect deadlock if there is one. Compared with testing, it is much easier to select execution paths or interleaving choices in simulation. Unlike testing, a complete system is not a requirement for simulation. When verifcation fails to provide answers for the critical properties, simulation in plenty of time can provide higher confidence in these critical properties. Currently, ARCATS provides random simulation to find deadlock. Other safety properties, such as assertion, is planned for the near future.

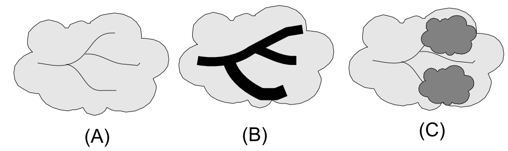


Figure 4. The concept of combined verification and simulation.

3.5 Simulation + Compositional Analysis

Simulation can be viewed as an expedient approach to address the infeasibility of verification methods. The random walk among giantic state space makes most people unconfident. In principle, any attempt to guide the simulation will result in the increase of memory resources. For example, to prevent simulation from walking into the paths that were already visited requires storage of visited states. Essentially, such an attempt is equivalent to explore the reachable states, only at a different extent. The line between verification and simulation will become blurred once any guidence strategy is added to the simulation.

In the cases where properties violation occurs in large depth, simulation can be easily lost in the gigantic state space. To tackle this problem, ARCATS provides a combined approach of simulation and verification to alleviate this problem. In Fig. 4, a conceptual diagram is shown. Fig. 4(A) is the conceptual illustration of simulation, where the cloud represents the state space. Simulation seeks to walk in the state space to find any properties of violation. Fig. 4(B) shows the an improved simulation, called symbolic simulation. A walk in the state space is equivalent to cover a wide range of states. Fig. 4(C) shows the concept of our approach. In this approach, subsystems are composed as in compositional analysis and are replaced by simpler interface processes. Safety properties such as deadlocks are preserved in this process. When simulation walks into the states of interface process, it is equivalent to walk the states of a subsystem, in which many internal states are not interested by its environment.

In Fig. 5, an experiment to show the power of combined approach is shown. The example in the experiment is the well-know dinning philosopher problem (DPP). The deadlock occurs when all the philosophers simultaneously pickup the chopstick at right hand or left hand. When the number of philosophers are increased, the depth of the deadlock is increased. However, a random simulation can not be smart enough to reach the deadlock straightforwardly. Most of the time, simulation roam around the state space before

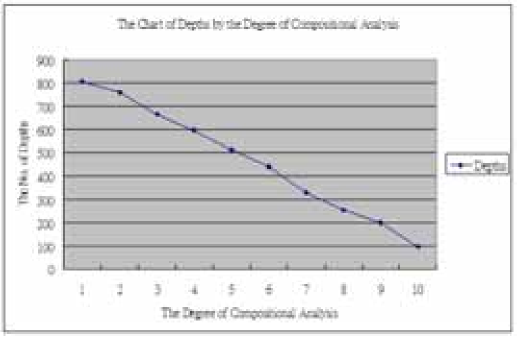


Figure 5. The average depth of deadlock detection using simulation and compositional analysis.

visiting the deadlock states. When a deadlock state is visited, the steps (depth) of this walk is recorded. We let the simulation run for 10 times and then compute the average depth.

In Fig. 5, the example has 100 philosophers. The Y-Axis is the average depth for the simulation to visit the deadlock states. The X-axis is the degree of compositional analysis. For example, degree is 5 means 50% of the philosophers and chopsticks are composed into a subsystem. When there are no compositional analysis, average depths in 10 times run is close to 800. When the degree of compositional analysis is increased, the average depths of random walk decreased significantly. As a result, the time spent to detect the deadlock is also reduced proportional to this declining rate.

3.6 Simulation + Compositional Analysis + Refactoring

When the as-built architecture of a system does not contain good subsystems to be used for the combined approach, refactoring again can come into play. However, refactoring modify the communication structure of the system at the CCS levels but CFG remains in the as-built architecture. Many problems need to be solved. This is not yet a supported phase. It is a future research.

4. Discussions and Conclusions

The worst-case complexity of reachability analysis is *P-SPACE* hard. In general, any approaches that attempt to analyze the whole system at once (a.k.a., global analysis) may only work for systems of small scale or of a particular class. To increase the scalability of verification techniques for industrial usage, pragmatic approaches such as the multi-phased solutions combined with compositional analysis and simulation have been proposed. From one phase to next phase, however, properties of interest may be compromised. For example, absence of deadlock can be compromised into high confidence in absence of deadlock. This confidence can be increased by investing more time in simulation, which is a practical and commonly used approach in hardware industry.

In the future, ARCATS plans to automate the transition from one phase to another and produce warnings to explain how properties of interest are compromised so that a tool user does not need to take care the manual steps between phases. To achieve this goal, much work remains to be done. Hopefully, a practical verification tool can be produced from ARCATS's research results in the foreseeable future.

References

[1] A. Bouali and R. de Simone. Symbolic bisimulation minimization. In *Fourth Workshop on Computer-Aided Verification, Volumne 663 of LNCS*, pages 96–108. Springer-Verlag, 1992.

[2] Y.-P. Cheng. Crafting a promela front-end with abstract data types to mitigate the sensitivity of (compositional) analysis to implementation choices. In *LNCS*, volume 3639, pages 139–153, 2005.

[3] Y.-P. Cheng, M. Young, C.-L. Huang, and C.-Y. Pan. Towards scalable compositional analysis by refactoring design models. In *Proceedings of the ACM SIGSOFT 2003 Symposium on the Foundations of Software Engineering*, pages 247–256, 2003.

[4] S. C. Cheung and J. Kramer. Context constraints for compositional reachability analysis. 5(4):334–377, October 1996.

[5] E. R. Gansner, E. Koutsofios, S. C. North, and K.-P. Vo. A technique for drawing directed graphs. *IEEE Transactions on Software Engineering*, 19(3):214–230, 1993.

[6] P. Godefroid, D. Peled, and M. Staskauskas. Using partial-order methods in the formal validation of industrial concurrent programs. In *Proceedings of 1996 International Symposium on Software Testing and Analysis*, pages 261–269, San Diego, California, USA, January 1996.

[7] S. Graf and B. Steffen. Compositional minimization of finite state systems. In *Proceedings of the 2nd International Conference of Computer-Aided Verification*, pages 186–204, 1990.

[8] G. J. Holzmann. *Design and Validation of Computer Protocols*. Prentice-Hall, Englewood Cliffs, NJ 07632, 1991.

[9] G. J. Holzmann. The model checker SPIN. *Software Engineering*, 23(5):279–295, 1997.

[10] K. L. McMillan. *Symbolic model checking*. Kluwer Academic Publishers, Massachusetts, 1993.

[11] R. Milner. *A Calculus of Communicating Systems*, volume 92. New York, 1980.

[12] W. J. Yeh and M. Young. Compositional reachability analysis using process algebra. pages 49–59, Oct. 1991.

Optimization of Multilevel Power Adjustment in Wireless Sensor Networks

Sheng-Tzong Cheng[1] and Mingzoo Wu[1,2]
[1]Department of Computer Science and Information Engineering, National Cheng Kung University,
[2]Networks and Multimedia Institute, Institute for Information Industry,
[1,2]Tainan, Taiwan R.O.C.
E-mail: stcheng@mail.ncku.edu.tw, wu.mingzoo@gmail.com

Abstract Restricted to the limited battery power of nodes, energy conservation becomes a critical design issue in wireless sensor networks (WSNs). Transmission with excess power not only reduces the lifetime of sensor nodes, but also introduces immoderate interference in the shared radio channel. It is ideal to transmit packets with just enough power. In this paper, we propose a multilevel power adjustment (MLPA) mechanism for WSNs to prolong the individual node lifetime and the overall network lifetime. The energy conservation is achieved by reducing the average transmission power. The analytical model is built for the MLPA mechanism enabled with k distinct power levels (k-LPA). Under a free space loss (FSL) model, the closed-form expression of optimal power setting is derived and the average transmission power can be minimized as $(k + 1)/2k$ of original fixed power.

Keywords: Wireless sensor network, energy conservation, power control

1. Introduction

Wireless sensor network (WSN) is a large scale wireless ad hoc network with special-purpose applications, such as environmental monitoring, homeland security, healthcare system, etc. A WSN comprises a large number of tiny sensor nodes equipped with sensing, computing, and communication capabilities [1], [2]. Usually, the nodes of a WSN are randomly deployed and share the same communication medium. In most cases, sensor nodes are battery-powered and left unattended after the initial deployment. A crucial issue of WSNs is the limited source of energy supplied by batteries coming with sensor nodes. In many scenarios, it seems impracticable to replace or recharge batteries of sensor nodes. For this reason, energy conservation becomes a critical design issue to prolong the lifetime of WSNs.

Both network connectivity and network lifetime are crucial issues that should be considered in WSNs. The connectivity level of a WSN depends on the transmission power of the nodes. If the transmission power is too small, the network might be disconnected. Consequently the network coverage will be diminished. However, as mentioned earlier, the energy is the scarcest resource of WSN nodes, and it determines the lifetime of a WSN. If a sensor node transmits with excessive high power to its neighboring nodes, the node lifetime will be reduced. Furthermore, an excess of transmission power generates redundant interference in the shared radio channel. Ideally and intuitively, the transmission power of each node should be adjusted on a point-to-point basis to maximize the possible energy savings and to preserve the required network connectivity [3]-[7].

In literature, there are several studies which consider a common transmission power (or a common transmission range) used in WSNs [8]-[12]. Panichpapiboon et al. [8] investigate the minimum transmission power used by all nodes to preserve the network connectivity of a WSN. In [8], only the closed-form expression of common optimal transmission power is derived for a square grid network topology. For networks with random topology, the common optimal transmission power is obtained from simulations. In this case, Bettstetter el al. [10] presents an analytical investigation of the connectivity of wireless ad hoc networks with randomly uniformly distributed nodes and homogeneous transmission range assignment. Tseng et al. [11] analyze the optimum transmission range to provide a power efficient connectedness control for wireless ad hoc networks. In [12], the minimum uniform transmission power is investigated in a wireless ad hoc network to maintain the network connectivity.

Apparently, the transmission power can be minimized by employing a common transmission power in a square grid network topology. However, in a random network topology, the transmission power can not be minimized just by using a common transmission power. In addition, using the optimal common transmission power predetermined in the design phase may not achieve the expected performance in the runtime environment.

In this paper, we propose a multilevel power adjustment (MLPA) mechanism for WSNs to prolong the individual node lifetime. The network connectivity is preserved as well. Each sensor node is equipped with power-adjustable transmitter. The energy conservation is achieved by reducing the average transmission power. Different transmission power is employed according to the distance of receiving node. The main contribution of this paper is that we build an analytical model to study the relationship between the MLPA and the degree of power conservation. Under the free space model, the closed-from expressions are derived for the minimum average transmission power and the optimal power configuration.

The rest of this paper is organized as follows. Section 2 introduces the proposed MLPA mechanism. Section 3 delivers an objective function indicating the average power

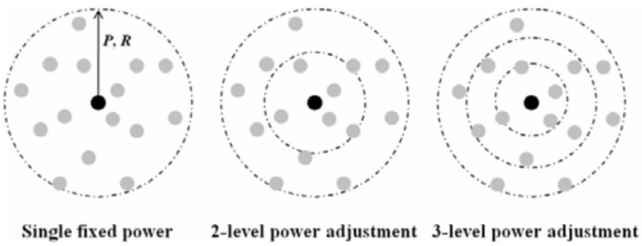


Fig. 1: Concept of multilevel power adjustment

Table 1: CC2420 output power configuration

PA Level	Output Power (dBm)	Output Power (mW)
31	0	1
27	-1	0.794
23	-3	0.501
19	-5	0.316
15	-7	0.2
11	-10	0.1
7	-15	0.032
3	-25	0.003

consumption of sensor nodes. In Section 4, we analyze the optimal power configuration for nodes enabled with *k* distinct power levels (*k*-LPA) in a free space loss (FSL) environment. Finally, conclusion and remark are drawn in Section 5.

2. Multilevel Power Adjustment

The concept of MLPA is depicted as shown in Fig. 1. Sensor nodes equipped with adjustable transmission power is not a brand-new concept. In literature, Lin et al. [13] present an adaptive transmission power control (ATPC) for WSNs, where each node builds a model for each of its neighboring nodes, describing the correlation between transmission power and link quality. A feedback-based transmission power control algorithm is employed to dynamically maintain individual link quality over time. In the real world, many industrial radio products are augmented with multiple transmission powers, e.g. CC2420 and CC2430 [14], [15]. The data sheet of CC2420 enumerates eight programmable transmission levels ranging from -25 to 0 dBm (see Table 1).

To the best of our knowledge, there are two fundamental problems which have not been investigated for WSNs enabled with adjustable transmission power:

Q1: How does the MLPA affect the degree power conservation? In other words, what's the relationship between the number of power levels and the degree of power conservation?

Q2: As the number of power levels is determined, how to configure the transmission power of each level, such that the maximum energy saving can be carried out?

In order to study problems Q1 and Q2, an objective function (or cost function) is formulated intended to minimize the average transmission power of each sensor node.

2.1. The MLPA Mechanism

In the proposed MLPA mechanism, each sensor node is equipped with a power-adjustable transmitter. The key feature of MLPA is to employ just enough power level to communicate with neighboring nodes. A lower power level is employed for closer nodes. On the other hand, a larger power is employed for farther nodes. The MLPA mechanism consists of the following three phases:

- **Phase I: Neighbor discovering and network topology construction**
  Initially, each node broadcasts a beacon to discover all possible neighboring nodes. A full transmission power is used to maximize the network connectivity level.
- **Phase II: Transmission power negotiation and adjustment**
  The energy conservation is achieved by reducing the average transmission power. It is unnecessary to transmit with a full power to a nearby neighbor. In phase II, a sensor node negotiates with its neighboring nodes, and determines which power level to be used for each neighbor. Take negotiation overhead into account, the piggyback mechanism can be used to reduce the overhead from power negotiation.
- **Phase III: Runtime maintenance**
  The wireless communication environment may change dynamically after the node deployment. The network topology may be changed due to the node movement. Sensor nodes have to repeat phase I and II to update the network topology and to preserve the optimal power assignment for each neighbor. In this paper, we consider a stationary WSN. The maintenance overhead is not taken into consideration at this time.

3. Power Objective Function

We formulate the objective function (or cost function) to evaluate the average transmission power of a sensor node enabled with *k* distinct power levels (*k*-LPA). Without loss of generality, we make the following assumptions:

A1: We consider a scenario where N sensor nodes are randomly deployed over a surface with finite area of A. The node spatial density (ρ) is defined as the number of nodes per unit area and is denoted as $\rho = N/A$.

A2: For an arbitrary node, the probabilities of transmission to each of its neighboring nodes are equally.

A3: Assume that a signal is attenuated with a distance d raised to the power γ, where γ is the path loss exponent (PLE). The power of the intended signal from the transmitter as observed at the receiver can be written as [16]

$$P_r = \frac{\alpha P_t}{d^\gamma}, \quad (1)$$

where

$$\alpha = \frac{G_t G_r \lambda^2}{(4\pi)^2} . \quad (2)$$

P_t stands for the transmission power at transmitting antenna. P_r denotes the signal power measured at receiving antenna. G_t and G_r are the transmitter and receiver antenna gains. λ denotes the wavelength of carrier. The PLE depends on the wireless environment and typically varies from 2 to 4. Initially, we consider the case of free space loss (FSL) model, i.e. PLE $\gamma = 2$.

In the proposed MLPA system, each node has an adjustable transmission power and an omni-directional antenna. Let P be the maximum transmission power and R be the maximum transmission range, correspondingly. Without any power adjustment, sensor nodes always transmit packet with the maximum power P. For a sensor node enabled with k distinct power levels, we have k different transmission ranges. Let R_i with an integer subscript $1 \le i \le k$ denote the k distinct transmission ranges. The R_i is described as

$$R_i = r_i R, \quad (3)$$

where $0 < r_1 < r_2 < \ldots < r_{k-1} < r_k = 1$.

3.2. MLPA under FSL Model

Under the FSL model, the transmission power is proportional to the square of corresponding transmission range. We use P_i with an integer subscript $1 \le i \le k$ to denote the transmission power of i-th level. P_1 indicates the smallest power level, and P_k indicates the largest power level using power P. The P_i is described as

$$P_i = r_i^2 P . \quad (4)$$

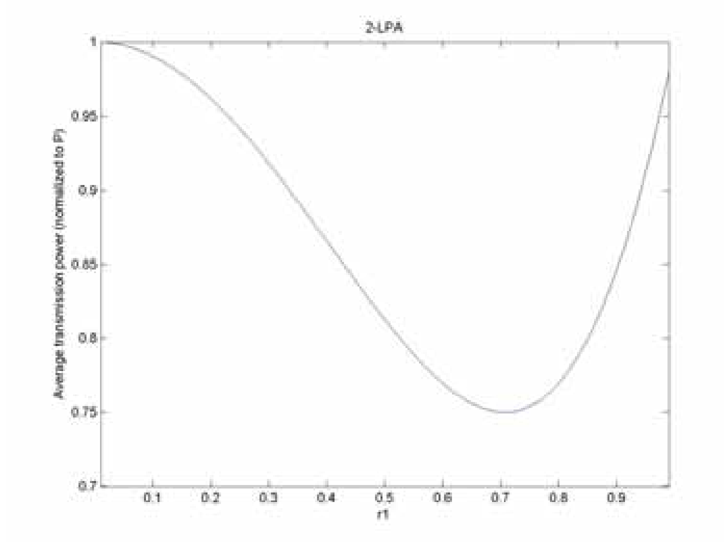
With adjustable transmission power, a sensor node employs transmission power P_i for data transmission to nodes which have distance longer than R_{i-1} and shorter than R_i. Let n_i denote the number of neighboring nodes served by power P_i, we have

$$n_i = \begin{cases} \rho\pi R_1^2 & i = 1 \\ \rho\pi(R_i^2 - R_{i-1}^2) & 2 \le i \le k \end{cases} \quad (5)$$

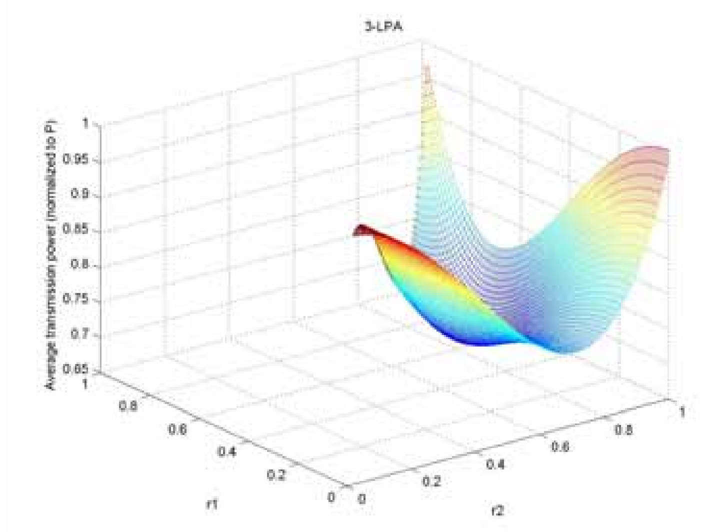
Under the assumption that the probabilities of transmission to each neighboring node are equally (A2), the average transmission power of k-LPA can be formulated as

$$P_{avg}(k) = \frac{\sum_{i=1}^{k} P_i n_i}{\sum_{i=1}^{k} n_i} . \quad (6)$$

Without MLPA, the transmission power is fixed as P regardless of the distance of receiving node. With adjustable transmission power, sensor nodes can employ appropriate power for transmission to neighboring nodes with different distance. Consequently, the average transmission power can be reduced, i.e. some energy can be saved. Fig. 2 shows the average transmission power versus different configurations for 2-LPA and 3-LPA. As we can see, the average transmission power depends on the configuration of r_i. Fig. 2a shows that the average transmission power is minimized to $0.75P$ when r_1 is close to 0.7. The power configuration of each level becomes an assignment problem or an optimization problem of objective function as shown in Eq. 6. In next section, we investigate how to configure each power level such that the average transmission power can be minimized, in other words, the maximum possible power saving can be carried out.



(a) 2-LPA, $0 < r_1 < r_2 = 1$



(b) 3-LPA, $0 < r_1 < r_2 < r_3 = 1$

Fig. 2: MLPA configurations

4. Optimal Power Configuration

Let $P_{avg}^{\min}(k)$ denote the minimum average transmission power of $P_{avg}(k)$ subject to the constraint $0 < r_1 < r_2 < \ldots < r_k = 1$.

$$P_{avg}^{\min}(k) = \underset{0<r_1<r_2<\ldots<r_k=1}{\arg\min} P_{avg}(k) . \quad (7)$$

For MLPA with few power levels, e.g. 2-LPA and 3-LPA, the minimum average transmission power ($P_{avg}^{\min}(k)$) and

Table 2: Optimal configuration for *k*-LPA

	P_1, R_1	P_2, R_2	P_3, R_3	P_4, R_4	P_5, R_5	$P_{avg}^{min}(k)$
$k=2$	$\frac{1}{2}P,\ \frac{\sqrt{2}}{2}R$	P, R	—	—	—	$\frac{3}{4}P$
$k=3$	$\frac{1}{3}P,\ \frac{\sqrt{3}}{3}R$	$\frac{2}{3}P,\ \frac{\sqrt{6}}{3}R$	P, R	—	—	$\frac{2}{3}P$
$k=4$	$\frac{1}{4}P,\ \frac{1}{2}R$	$\frac{1}{2}P,\ \frac{\sqrt{2}}{2}R$	$\frac{3}{4}P,\ \frac{\sqrt{3}}{2}R$	P, R	—	$\frac{5}{8}P$
$k=5$	$\frac{1}{5}P,\ \frac{\sqrt{5}}{5}R$	$\frac{2}{5}P,\ \frac{\sqrt{10}}{5}R$	$\frac{3}{5}P,\ \frac{\sqrt{15}}{5}R$	$\frac{4}{5}P,\ \frac{2\sqrt{5}}{5}R$	P, R	$\frac{3}{5}P$

the optimal power configuration (P_i) can be obtained by solving differential equations.

4.1. Power Optimization for 2-LPA

In the case of 2-LPA mechanism, each node has two distinct transmission powers: $P_1 = r_1^2 P$ and $P_2 = r_2^2 P$, where $0 < r_1 < r_2 = 1$. Accordingly, there are two different transmission ranges: $R_1 = r_1 R$ and $R_2 = r_2 R$.

The number of neighboring nodes served by P_1 and P_2 are $n_1 = \rho\pi(r_1 R)^2$ and $n_2 = \rho\pi(r_2^2 - r_1^2)R^2$, respectively. Substituting into equation (6), the average transmission power by using 2-LPA mechanism can be described as

$$P_{avg}(2) = \frac{r_1^4 - r_1^2 r_2^2 + r_2^4}{r_2^2} P . \quad (8)$$

The average transmission power $P_{avg}(2)$ depends on the value of r_1 (note that r_2 is fixed as 1). Since $0 < r_1 < r_2$, we set $r_1 = q r_2$, where $0 < q < 1$. Then $P_{avg}(2)$ can be alternatively expressed as

$$P_{avg}(2) = (q^4 - q^2 + 1) r_2^2 P . \quad (9)$$

Let function $f(q) = q^4 - q^2 + 1$. To obtain the minimum average transmission power of 2-LPA, we turn to minimize the function $f(q)$ subject to the constraint $0 < q < 1$. Utilizing the first derivative test, we have $P_{avg}^{\min}(2) = 0.75P$ at $q = \sqrt{2}/2$ (≈0.707). The optimal configuration for 2-LPA in a FSL environment is given as follows.

$$\begin{cases} P_1 = (1/2)P \\ P_2 = P \end{cases}, \begin{cases} R_1 = (\sqrt{2}/2)R \\ R_2 = R \end{cases} \quad (10)$$

4.2. Power Optimization for 3-LPA

Similarly, for sensor nodes equipped with 3-LPA mechanism, there are three different usable transmission powers: $P_1 = r_1^2 P$, $P_2 = r_2^2 P$, and $P_3 = r_3^2 P$, where $0 < r_1 < r_2 < r_3 = 1$. The corresponding transmission ranges are: $R_1 = r_1 R$, $R_2 = r_2 R$, and $R_3 = r_3 R$. The number of neighboring nodes served by P_1, P_2, and P_3 are $n_1 = \rho\pi(r_1 R)^2$, $n_2 = \rho\pi(r_2^2 - r_1^2)R^2$, and $n_3 = \rho\pi(r_3^2 - r_2^2)R^2$, respectively. Substituting into (6), the average transmission power by using 3-LPA mechanism can be described as

$$P_{avg}(3) = \frac{r_1^4 + (r_2^2 - r_1^2) r_2^2 + (r_3^2 - r_2^2) r_3^2}{r_3^2} P . \quad (11)$$

Since $0 < r_1 < r_2 < r_3$, we set $r_1 = q_1 r_3$ and $r_2 = q_2 r_3$, where $0 < q_1 < q_2 < 1$. The average transmission power of 3-LPA can be alternatively expressed as

$$P_{avg}(3) = \{q_1^4 + (q_2^2 - q_1^2) q_2^2 + (1 - q_2^2)\} r_3^2 P . \quad (12)$$

Let function $f(q_1, q_2) = q_1^4 + (q_2^2 - q_1^2) q_2^2 + (1 - q_2^2)$. To minimize the average transmission power, we turn to minimize function $f(q_1, q_2)$ subject to the constraint $0 < q_1 < q_2 < 1$.

$$\begin{cases} \dfrac{\partial f(q_1, q_2)}{\partial q_1} = 0 \\ \dfrac{\partial f(q_1, q_2)}{\partial q_2} = 0 \end{cases} \Rightarrow \begin{cases} 4q_1^3 - 2q_1 q_2^2 = 0 \\ 4q_2^3 - 2q_1^2 q_2 - 2q_2 = 0 \end{cases} \quad (13)$$

Solving the system of partial differential equations (13), we have $P_{avg}^{\min}(3) = (2/3)P$ at $q_1 = \sqrt{3}/3$ and $q_2 = \sqrt{6}/3$. The minimum average transmission power is $(2/3)P$ for 3-LPA in a FSL environment. And the optimal configuration is given as follows.

$$\begin{cases} P_1 = (1/3)P \\ P_2 = (2/3)P \\ P_3 = P \end{cases}, \begin{cases} R_1 = (\sqrt{3}/3)R \\ R_2 = (\sqrt{6}/3)R \\ R_3 = R \end{cases} \quad (14)$$

4.3. Optimal Configuration for *k*-LPA

For higher degree of MLPA, it is hard to acquire the optimal configuration by using pencil and paper. We resort to the built-in function "Minimize" of Mathematica [17], a powerful mathematical software, to minimize $P_{avg}(k)$ subject to the constraint $0 < r_1 < r_2 < \ldots < r_k = 1$. Table 2 lists the optimal configuration and the minimum average transmission power for *k*-LPA, where *k* varies from 2 to 5.

Looking into Table 2, we could conjecture the closed-form expression of the optimal configuration for *k*-LPA, where *k* is a positive integer. In the optimal power configuration of *k*-LPA, the average transmission power is minimized, the power of the *i*-th level ($1 \le i \le k$) is conjectured to be

$$P^{opt}(k, i) = \frac{i}{k} P . \quad (15)$$

Consequently, the transmission range of the *i*-th power level is described as

$$R^{opt}(k, i) = \sqrt{\frac{i}{k}} R . \quad (16)$$

And the number of neighboring nodes served by the *i*-th power level is described as

$$n^{opt}(k, i) = \frac{\rho\pi R^2}{k} , \quad (17)$$

which is independent of the index of power level. By employing *k*-LPA, the average transmission power can be reduced and minimized to

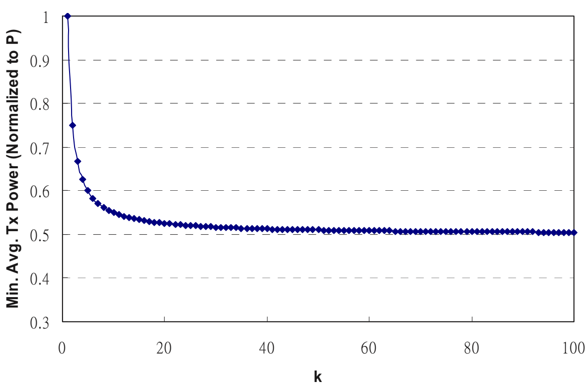


Fig. 3: Minimized average transmission power

$$P_{avg}^{\min}(k) = \frac{k+1}{2k} P . \quad (18)$$

The closed-form of optimal configuration can be proved by using techniques of mathematical induction.

4.4. Analytic Results

Fig. 3 shows the minimum average transmission power versus the degree of MLPA (*k*). As we can see in Fig. 3, the minimum average transmission power is a monotonically decreasing function of *k*. Eq. 18 tells us that the higher degree of MLPA is the more transmission power can be saved. The limit of $P_{avg}^{\min}(k)$ exists as the *k* approaches infinity.

$$\lim_{k \to \infty} P_{avg}^{\min}(k) = \frac{P}{2} . \quad (19)$$

Ideally, half of transmission power can be saved as the number of power levels approaches infinity, where the just enough power is used for transmission to each neighboring node. But it is infeasible and unnecessary to provide an MLPA mechanism with infinite power levels. The power negotiation overhead should be considered as well. For *k* = 10, there are at most 45% of transmission power can be saved, which reaches 90% of the maximum possible power saving in the FSL model. Generally, by using optimized *k*-LPA, there are {100(k – 1)/2k}% of transmission power can be saved, which reaches {100(k – 1)/k}% of the maximum possible power saving in the FSL model.

5. Conclusion

In this paper, the MLPA mechanism is proposed for WSNs to prolong the individual node lifetime and the network lifetime. In stead of using single fixed transmission power, the main feature of MLAP is to employ appropriate power level for each neighboring node. We focus on analyzing the relationship between the MLPA and the degree of power conservation, and proposing the optimal power configuration for MLPA.

Initially, we focus on analyzing the transmission power consumption of *k*-LPA under a FSL environment. The closed-form expressions of optimal configuration for *k*-LPA are given in Section 4. The average transmission power of *k*-LPA under FSL can be minimized as $\{(k + 1)/2k\}P$. More transmission power can be saved as the number of power levels increased. About 50% of transmission power can be saved as the *k* approaches infinity.

Appendix: Proof of Optimal Configuration

Statement *S*: $P_{avg}^{\min}(k) = \frac{k+1}{2k} r_k^2 P$, where

$$r_i = \begin{cases} 1 & i = k \\ \sqrt{\frac{i}{k}}\, r_k & i = 1, 2, \ldots, k-1 \end{cases}$$

We use the technique of mathematical induction to prove statement *S* holds for all positive integer *k*.

The basis: Showing that statement *S* holds for *k* = 1. For *k* = 1, the transmission power is fixed as *P*, and $r_1 = 1$. $P_{avg}^{\min}(1) = P$. The statement *S* is true for *k* = 1.

The inductive step: showing that if statement *S* holds for *k* = *m*, then the same statement also holds for *k* = *m* + 1. Assume that statement *S* is true for *k* = *m*, i.e.

$$P_{avg}^{\min}(m) = \frac{m+1}{2m} r_m^2 P \text{, where } r_i = \begin{cases} 1 & i = m \\ \sqrt{\frac{i}{m}}\, r_m & i = 1, 2, \ldots, m-1 \end{cases}$$

For *m* = *k* + 1.

$$\begin{aligned} P_{avg}(m+1) &= \frac{\sum_{i=1}^{m+1} P_i n_i}{\sum_{i=1}^{m+1} n_i} \\ &= \frac{r_1^4 + \sum_{i=2}^{m+1} (r_i^2 - r_{i-1}^2) r_i^2}{r_{m+1}^2} P \\ &= \frac{r_1^4 + \sum_{i=2}^{m} (r_i^2 - r_{i-1}^2) r_i^2}{r_m^2} \times \frac{r_m^2}{r_{m+1}^2} P + \frac{(r_{m+1}^2 - r_m^2) r_{m+1}^2}{r_{m+1}^2} P \\ &= \frac{r_m^2}{r_{m+1}^2} P_{avg}(m) + (r_{m+1}^2 - r_m^2) P \end{aligned}$$

Minimizing $P_{avg}(m+1)$, we have

$$\begin{aligned} P_{avg}^{\min}(m+1) &= \arg\min_{0 < r_1 < r_2 < \ldots < r_{m+1} = 1} P_{avg}(m+1) \\ &= \arg\min_{0 < r_1 < r_2 < \ldots < r_{m+1} = 1} \{ \frac{r_m^2}{r_{m+1}^2} P_{avg}(m) + (r_{m+1}^2 - r_m^2) P \} \\ &= \arg\min_{0 < r_1 < r_2 < \ldots < r_{m+1} = 1} \{ \frac{r_m^2}{r_{m+1}^2} P_{avg}^{\min}(m) + (r_{m+1}^2 - r_m^2) P \} \\ &= \arg\min_{0 < r_m < r_{m+1} = 1} \{ \frac{r_m^2}{r_{m+1}^2} \times \frac{m+1}{2m} r_m^2 P + (r_{m+1}^2 - r_m^2) P \} \end{aligned}$$

Let $r_m = q r_{m+1}$, where $0 < q < 1$.

$$P_{avg}^{\min}(m+1) = \arg\min_{0<r_m<r_{m+1}=1}\{\frac{r_m^2}{r_{m+1}^2}\times\frac{m+1}{2m}r_m^2P+(r_{m+1}^2-r_m^2)P\}$$
$$= \arg\min_{0<q<1}\{(\frac{m+1}{2m}q^4-q^2+1)r_{m+1}^2P\}$$

Let function $f(q)=\frac{m+1}{2m}q^4-q^2+1$. Utilizing the first derivative test, the $f(q)$ is minimized as $\frac{m+2}{2(m+1)}$ at $q=\sqrt{\frac{m}{m+1}}$. Then we have

$$P_{avg}^{\min}(m+1) = \arg\min_{0<q<1}\{(\frac{m+1}{2m}q^4-q^2+1)r_{m+1}^2P\}$$
$$= \frac{(m+1)+1}{2(m+1)}r_{m+1}^2P$$

where

$$r_i=\begin{cases}1 & i=m+1\\ \sqrt{\frac{m}{m+1}}r_{m+1} & i=m\\ \sqrt{\frac{i}{m}}r_m=\sqrt{\frac{i}{m}}\times\sqrt{\frac{m}{m+1}}r_{m+1}=\sqrt{\frac{i}{m+1}}r_{m+1} & i=1,2,...,m-1\end{cases}$$

We have shown that if statement *S* holds for $k = m$ ($m \geq 1$), then statement *S* also holds for $k=m+1$. We conclude that statement *S* holds for all positive integer *k*. Furthermore, by deducing from statement *S*, we can obtain the closed-form expressions shown in Section 4.

Acknowledgement

This study is conducted under the "Applied Information Service Development & Integration project" of the Institute for Information Industry which is subsidized by the Ministry of Economy Affairs of the Republic of China.

Reference

[1] D. Estrin, "Embedded Networked Sensing Research: Emerging Systems Challenges," NSF Workshop on Distributed Communications and Signal Processing, Northwestern University, Dec. 2002.
[2] D. Estrin, R. Govindan, J.S. Heidemann, and S. Kumar, "Next Century Challenges: Scalable Coordination in Sensor Networks," Proc. ACM/IEEE MobiCom, Seattle, WA, pp. 263-270, Aug. 1999.
[3] S. Agarwal, R.H. Katz, S.V. Krishnamurthy, and S.K. Dao, "Distributed Power Control in Ad-Hoc Wireless Networks," Proc. IEEE PIMRC 2001, vol. 2, pp. 59-66, Sep. 2001.
[4] R. Ramanathan and R. Rosales-Hain, "Topology Control of Multihop Wireless Networks Using Transmit Power Adjustment," Proc. IEEE INFOCOM, vol. 2, pp. 404-413, Mar. 2000.
[5] T. Elbatt and A. Ephremides, "Joint Scheduling and Power Control for Wireless Ad Hoc Networks," IEEE Trans. Wireless Comm., vol. 3, no. 1, pp. 74-85, Jan. 2004.
[6] R.L. Cruz and A.V. Santhanam, "Optimal Routing, Link Scheduling and Power Control in Multi-Hop Wireless Networks," Proc. IEEE INFOCOM, vol. 1, pp. 702-711, Apr. 2003.
[7] V. Rodoplu and T.H. Meng, "Minimum Energy Mobile Wireless Networks," IEEE J. Selected Areas in Comm., vol. 17, no. 8, pp. 1333-1344, Aug. 1999.
[8] S. Panichpapiboon, G. Ferrari, and O.K. Tonguz, "Optimal Transmit Power in Wireless Sensor Networks," IEEE Trans. Mobile Computing, vol. 5, no. 10, pp. 1432-1447, Oct. 2006.
[9] S. Narayanaswamy, V. Kawadia, R.S. Sreenivas, and P.R. Kumar, "Power Control in Ad-Hoc Networks: Theory, Architecture, Algorithm and Implementation of the COMPOW Protocol," Proc. European Wireless 2002 Next Generation Wireless Networks: Technologies, Protocols, Services, and Applications, pp. 156-162, Florence, Italy, Feb. 2002.
[10] C. Bettstetter and J. Zangl, "How to Achieve a Connected Ad Hoc Network with Homogeneous Range Assignment: An Analytical Study with Consideration of Border Effects," Proc. IEEE Int'l Workshop Mobile and Wireless Comm. Network, pp. 125-129, Sep. 2002.
[11] C.-C. Tseng and K.-C. Chen, "Power Efficient Topology Control in Wireless Ad Hoc Networks," Proc. IEEE WCNC, vol. 1, pp. 610-615, Mar. 2004.
[12] Q. Dai and J. Wu, "Computation of Minimal Uniform Transmission Power in Ad Hoc Wireless Networks," Proc. IEEE ICDCS 2003, pp. 680-684, May 2003.
[13] S. Lin, J. Zang, G. Zhou, L. Gu, T. He, and J. A. Stankovic, "ATPC: Adaptive Transmission Power Control for Wireless Sensor Networks," ACM SenSys, Nov. 2006.
[14] Chipcon AS, "CC2420 2.4 GHz IEEE 802.15.4 / ZigBee-ready RF Transceiver," Data Sheet, Rev. 1.4, 2006.
[15] Chipcon AS, "A True System-on-Chip solution for 2.4 GHz IEEE 802.15.4 / ZigBee," Data Sheet, Rev. 2.01, 2006.
[16] T.S. Rappaport, "Wireless Communications: Principles and Practice," Prentice Hall, 2002.
[17] Wolfram Research, Mathematica. http://www.worlfram.com/.

Projects and goals for the Eclipse italian community

P. Maresca*

*Dipartimento di Informatica e Sistemistica, Università di Napoli "Federico II"
Via Claudio 21, Napoli
paolo.maresca@unina.it

Abstract

ECLIPSE is a development community with the only statutory "to contribute to the development, evolution, promotion and support of a platform (ECLIPSE) taking into account the participating open source communities, their competencies and services". This effort is based upon projects and subprojects, ECLIPSE itself is, in fact, known as the ECLIPSE project, and JDT (Java Development Tools) is an ECLIPSE subproject. ECLIPSE is not only an extendible platform through plug-ins, but is a "modus operandi" as well, and a working philosophy aimed at increasing the competencies and abilities of the underlying community, obtained by using standards, tools, the sharing of information and the reuse of any sort of artefact. The communities are the projects' engine, as they provide information on projects, resources and events.

The need to set up an Eclipse Italian community, involving both universities and companies, is based on many considerations that will be discussed following. Eclipse Italian community was born in December 2005 during a workshop at IBM-Tivoli Lab Rome. This paper shows the state of the art of this community viewed thorough their projects, sustained from Eclipse ecosystem. Moreover we will show progress and goals for the Eclipse Italian community for the next future.

1. Introduction

ECLIPSE [1] is a software development community with approximately 162 member companies, and 70 major Open Source projects made up with a lot of subprojects.

Many people say that ECLIPSE is *quite close to the ideal software* due to several characteristics: (i) the supported operative systems (Linux, HP-UX, AIX, Solaris, QNX, Mac OS X and Windows), (ii) the supported methodologies, (iii) its internationalization, accessibility and maturity at company level. What is more important above all, is that you do not need to attend any courses to learn how to use it, and is as flexible as you wish it may be. This is why it is so widely adopted in university courses, even at the very first study levels. In one word, and *extendible environment useful for everything, and nothing in particular*: *just like a Swiss pocket knife!*

The need to set up an Italian ECLIPSE community [2], involving both universities and companies, is based on the following considerations:

- many groups in Italian Universities currently work with ECLIPSE;
- there is the need to spread the open source culture and "modus operandi" in Italian Universities, for both cultural and economic reasons;
- some Italian companies, like IBM-Tivoli lab, have open source, and ECLIPSE in particular, as their core businesses;
- this community could be an excellent opportunity to interlace relationships between universities and companies, investing through the compulsory internship recently adopted by Italian law. According to which, the students, as third actors, may contribute to the growth of this community with their training, thus improving the quality of these internships, and strengthening the relationships *Academia-Companies*;
- there is also the need to internationalize the relationships between the students, and this can be achieved by letting them participate to projects that require interactions with communities of different cultural environments;
- the fact that CNIPA committee fosters the development and growth of open source communities. In particular, CNIPA, following the Directive issued by the Ministry for Innovation and Information Technology dated December 19th, 2003, has launched an Open Source Observatory [3], whose main objectives are:
  - to favour the *launch of initiatives aimed at spreading the open source experiences already matured, or under development, at national universities or research centres*;
  - *the promotion and the exchange of experiences with similar EU based units*.

Based on the above-mentioned considerations, we now aim to outline a certain number of projects that we think will interest the ECLIPSE community. Some of these have also been pointed out and discussed during the first Italian Eclipse community conference call, kick-offing its birth on January 26th, 2006 and presented also in the first international conference Eclipse-IT 07 on November 4-5.

2. The Eclipse renaissance craftsman's studio model

One of the most important issue when approaching to a new practice community is to study their organization. The Eclipse ecosystem is the best existing model for multiple corporations to create innovation networks, enabling cooperation on the development of product-ready open source software. Eclipse is a hierarchic ecosystem, it is a community of communities where the governance model is replicated inside several levels of communities to support organisational performance, interaction and networking and capturing best practices. The Eclipse governance model is re-creating a craft studio model of governance where everyone learns from each other. Pete McBreen depicted this kind of learning [4], but Eclipse adds one more feature to it one which ***fosters innovation***: ***the Eclipse model is similar to the renaissance craft man studio.***

Every Free Software Project has two characteristic phases: the ***cathedral phase*** and the ***bazaar phase***. The most important change to be made during the transition between these two phases is in the project management. In the cathedral phase, the project author or the core team makes all decisions on their own. In the bazaar phase, instead, the project author has to open up the development process and allow others to participate in the project. Other developers have to be encouraged and their contributions have to be accepted and rewarded. The management spectrum of projects in the bazaar phase varies between rigid management style and loose management style. Management style is related to the characteristics of project too: e.g. Linux kernel maintenance needs a relatively rigid control, in fact it is vital to avoid feature creep and bloat in kernel and to maintain a clear separation between kernel and user space, while KDE project needs less rigid control because it is a feature-rich environment where feature improvement and addition are desirable in many different areas.

We believe that the renaissance craft man studio model sets a management style between the previous two extreme ends and shifts the coordination challenges from planned and centralized governance to distributed yet catalyzed governance. Relations between innovative capacity and collaboration are notorious [5]. Collaborative research is the source of the main successful innovations and some firms designed very efficient methods as CISCO, in the frame of its *Emergent Technologies* with its internal entrepreneurship-minded start-up teams with external assigned persons. We believe that the renaissance craftsman studio model increases the ability of learning inside the project and allows an organization of work consistent with typology of projects. It stimulates the creative production too; in fact it inspires shared mental models leading to deep collaboration. The two kinds of innovation identified so far: ***continuous innovation*** and ***sudden innovation***, both rely upon useful perception-action learning to speed innovation. In Eclipse communities we have both little changes made by a lateral community and sudden changes, what a leader of the community gathers from his/her network. We can imagine on X-axis representation of the innovation made by number of reuses and little changes of products, while the Y-axis representation shows the strong innovation where it is assimilated to a severe change in product. As we can see, Eclipse retains both types of innovation. We believe also that innovation is always a matter of ideas coming from the environment that will materialize in the mind of the innovator even if it looks a stroke of genius [17]. From this point of view, innovation is strongly related to knowledge management and it is seen under a dynamic angle not of collection and storage, but of circulation and exchanges. The Eclipse governance model fosters the gathering and exchange of well selected *tacit knowledge* due to sponsor firms that act as catalysers of good quality results. In this process of distributed knowledge creation and diffusion and of collaboration reinforcement, the local communities benefit of roles of Industrial Catalyst (ICs) and Local Catalysts (LCs) . In fact, Industrial Catalysts and Local Catalyst are the terms by which the Eclipse project could define global and local actors that, as in a chemical reaction, get the role of facilitating and speeding up the process of Eclipse ecosystem implementation. The choice of recognizing one local organization (LC) firmly rooted in the regional system as responsible of the Eclipse local implementation has been a successful one and it's replicated every time and experimented in new realities by Industrial Catalyst. Using the metaphor of renaissance, sponsors acts as kings or nobles that sponsor the activity of renaissance craftsmen studios.

But what does the LC do? Local Catalysts work to create a climate of reciprocal trust in order to promote the Eclipse approach local exploitation, to sustain the communities and to establish the concrete usage of the technological environment. It is the LC that engages the developers and users and plug-in developers, coordinating their training programs in order to support their technical systems and monitor their activities. Again in the metaphor of renaissance, LCs acts as a Master that sponsors the activity of renaissance studio craftsman.

LCs also has four main goals:

1. engage local individuals and organizations for communities
2. sustain Eclipse technological development and customization to local needs
3. attract policy makers' interest
4. increase the Eclipse communities' network.

In order for individuals and organizations to make the decision to engage with Eclipse, collaborating with others, share strategic functions as well as sensitive information, it is essential that they feel they can count on intermediaries able to facilitate the process of network building and knowledge transfer. The concept of *trust* is a fundamental one, the projects experience confirmed literature results about innovation adoption: beside the technological quality of an innovation the trust towards who is promoting such innovation is a central variable in the adoption process.

ICs act as Sponsors of the Eclipse, and LCs as Community Leaders. We have here nobles and masters working together to create innovation.

With the term 'sponsor' we mean the capacity of ICs and LCs to create consensus among Eclipse and develop an atmosphere of trust around it; using this trust background, LCs work as gatekeepers for the territorial community and are able to attract other innovation leaders as additional Eclipse sponsors (large enterprises, SMEs, research centres, academics intermediate actor such as chamber of commerce, development agencies, entrepreneurial association, etc.). The activation of social local networks facilitates the decision of a single individual or organization to get involved. As in medieval ages cathedral were built by hundreds of masters and craftsmen, today products of communities' work are driven by new masters. In the past the relations between sponsor and craftsman studio were regulated on practical basis of work done and on how much it made the stars of sponsors shine. The masters of the Renaissance were really able in choosing the best available partners and to elaborate and customize with their geniality each person's contribution.

Renaissance craftsman studio is the realization of the distributed authority and heterarchy that involves relations of interdependence; both are characterized by distributed intelligence and organization of diversity [7]. In fact, when a renaissance craftsman studio had problems in a project to find all necessary knowledge to accomplish stakeholders needs, it would seek help from other renaissance craftsmen studios. This behaviour is typical of Eclipse communities, when they experiment trouble in knowledge and skills. Putnam [8] states that networks have value, first, to the people who are in the networks. For example, economic sociologists have repeatedly shown that labour markets are thoroughly permeated by networks, so that most of us are as likely to get our jobs through whom we know as through what we know. Diversity has a value for societies too, in the long run immigration and diversity as social problems are likely to have important cultural, economic, fiscal, and developmental benefits. In the short run, however, immigration and ethnic diversity tend to reduce social solidarity and social capital. We guess an Eclipse community regulates its diversity growth to accomplish those phases of diversity introduction for its own characteristics, (e.g. catalysers) but this intuition has to be verified on field.

The network of the renaissance craftsmen studios was territorially distributed, as there was not only the studio that already reached success, but even others that grouped able boys, etc. Each renaissance craftsman studio could be very easily reached. How is it possible today to coordinate all the studios? We need a meta-renaissance studio: Eclipse.org. It is the renaissance studio of studios. It mixes influences of sponsors and communities, giving a unified view of the targets and the expected results from the ICs and LCs networks.

However, will this model of governance take roots somewhere?

2. The Eclipse Italian community projects

Since their constitution Eclipse Italian community have had only one international conference in Napoli (IT) [9]. The scientific material relating to this conference is available in [2] . Here we are planning to show the content of the more recent projects we are developing in conjunction with other eclipse academic group and/or industrial partner. For each one we will provide a brief presentation and goals.

2.1 The Eclipse Italian community with Second Life projects (Eclipse-IT-SL)

The Eclipse Italian Community has a web development team always looking for new interaction forms, able to improve the communication qualities between members, allowing people to exchange their knowledge in a more informal way, realizing a more pleasant and less alienating communication. So Eclipse Italian Community couldn't sit by one of biggest phenomenon of internet-based interaction technology of last two years: Second Life

Second-Life [10] is the biggest virtual world ever conceived; a product distributed in 2003 by Linden Labs, it joins recreational aspects of a massive-multi-player videogame with social and economics ones. In this big virtual world, each user can create a personal avatar, a kind of virtual alias for himself, with which he can explore lands and interact with other people. Composed of islands, Second Life allows, using teleport, to move from a place to another and eventually to create own house or company office after buying a piece of land. SL in fact recognizes copyrights on created objects, so a user can sell his creation receiving payment in a virtual coin, called Linden Dollar (L$) that can be converted in real US Dollars. The great peculiarity of this virtual world is that it can be expanded by the same users, buying new islands and building on them.

The potentialities of SL in the didactical area and in the diffusion of knowledge are evident: using a virtual world, in which each person can identify himself in a virtual avatar, joins the entertainment and the possibility of extend his own knowledge. In the same time it can involve new generations, very skilled in using complex gaming systems, allowing people to interact in a more informal way, realizing a more pleasant and less alienating communication.

So the web development team of Eclipse Italian Community decided to follow the example of many American Universities and several Organizations, that saw in Second Life an investment, and decided to establish an headquarter of our organization in Second Life where members can organize meetings and show the community developed products. This development project has been named Eclipse-IT-SL.

Building a headquarter for Eclipse Italian Community in Second Life, implies the choice of a building that represents the organization and, in the meantime, that is functional to all community activities. Bearing in mind that SL is a videogame, is simple imagine that it allows realizing every kind of abstract building.

We looked for a kind of building, that, immersed in the territory, was of big visual impact and that represented, as well, the soul and aims of Eclipse Italian Community. The building must be erected on "IBM Eternal City" island, in which are already present very beautiful buildings. So we have chosen a lighthouse.

Allegorically a lighthouse can be intended as a point of reference that with his huge structure is erected sovereign and stationary in the surrounding chaos, place of certainties and order. The lighthouse (Fig. 1), in fact, is built by men pro men, and, in our case, hides inside a very comfortable place, where each one can find answers to his own questions and offer his know-how to the community. Supreme result of human hard-working, it well symbolizes the spirit of sharing knowledge and of the common growth of Eclipse Italian Community.

The lighthouse base is composed by two rooms placed on two different floors:

- the hall is a wide square room, destined to host the acceptance and the info point of the community.
- The other room is a bitter smaller than the hall and hides inside a smaller area that will be used as conference room (Fig. 2).

2.2 The Unified-Based Scenario Design (Eclipse-USBD)

The USBD Project is conducted in conjunction with IBM-Tivoli lab Italia and Eclipse Italian community. The projects aims to develop an highest design level editor for designing business process. The developing team is composed of 4 people from Eclipse community and 3 people from IBM-Tivoli lab.

The innovation in developing is the use of jazz platform for cooperating and developing the project (see fig.3).

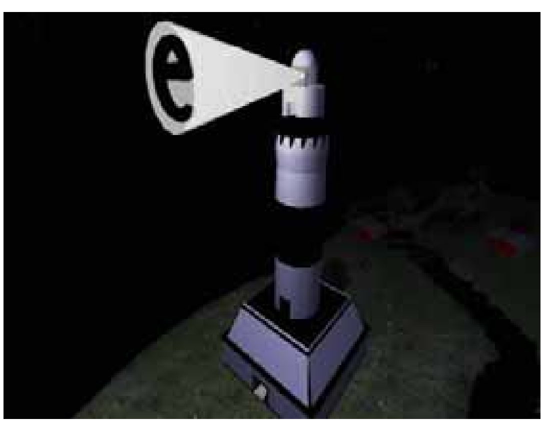


Figure 1. The Lighthouse

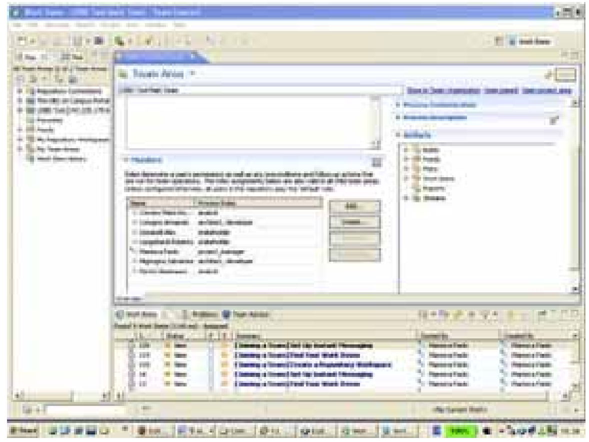


Figure 3 Jazz- team concert environment for USBD development

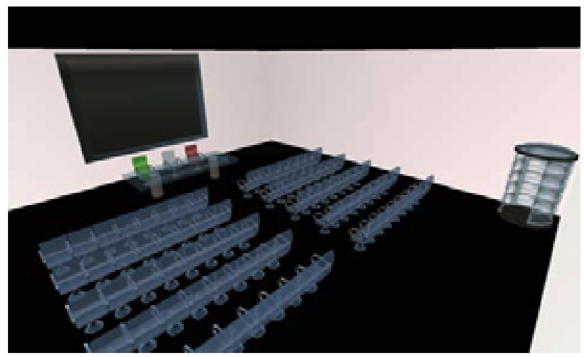


Figure 2. The SL conference room of Eclipse Italian Community lighthouse

2.3 The Eclipse – Web Services Project (Eclipse-WS)

Web Services are the most promising SOC-based technology. They use the Internet as the communication framework and open Internet-based standards, including Simple Object Access Protocol (SOAP) for transmitting data, the Web Services Description Language (WSDL) for defining services, a lot of visual tools for representing web services (ex. Protège), and the Business process Execution Language for Web Services (BPEL4WS) for orchestrating services. The Eclipse platform is the ideal platform to sustain a WS approach given that SOC group together many intricately concepts, protocols and technologies that originate in a wide range of disciplines.

Many Eclipse projects are ongoing actually in the Eclipse platform, but one of most interesting challenges is represented by the logical composition of web services. This Eclipse-WS Project that aims to compose web services in logical manner. Fig. 4 shows the tool developed by Eclipse Italian community at Federico II

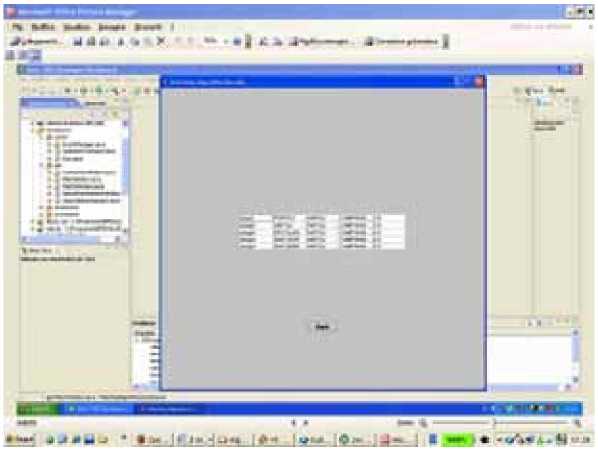


Figura 4 Eclipse-WS environment

2.4 The Eclipse – Learning and Cooperative environment (Eclipse-LCE)

Another project is Eclipse LCE its aims to enable student to cooperate when they are developing common project. The architecture is complex and for the sake of the brevity we omitted it. It is the case to outline that is based on Eclipse server and students can access it using the same development environment they usually use to compile a program (in different languages) or to learn the courses content. This is one of the didactical project of the Eclipse Italian community. A snapshot is showed in fig. 5

2.5 The Multimedia Knowledge Eclipse Environment (MKEE)

MKEE project is one of the oldest projects in the Eclipse Italian community. Actually there exists almost three release of the same tool and the projects is conducted jointly from University of Napoli Federico II and University of Pittsburg (OH) U.S.A. A latest version is appearing in [11].

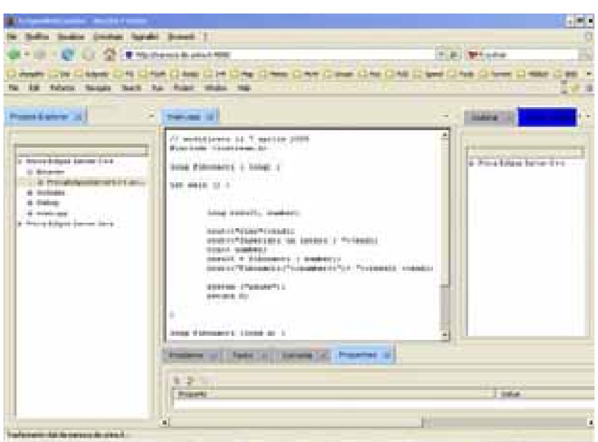


Figura 5 Eclipse Learning & Cooperative Environment

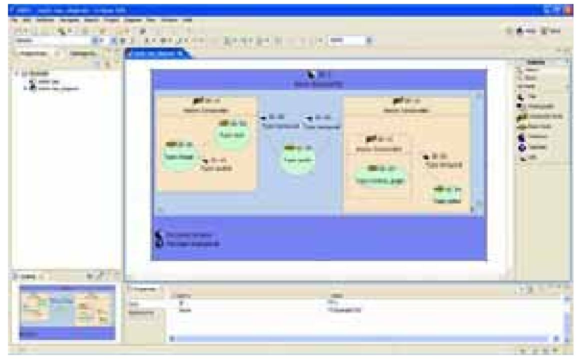


Figura 6 Multimedia Knowledge Eclipse Environment

In one of these release we want to demonstrate the Multimedia Knowledge Eclipse Environment (MKEE) as a web engineering tool able to realize multimedia system used jointly from the university of Naples Federico II and the university of Pittsburg (USA). The use of an Eclipse based development environment, allows us to treat the multimedia system as a new Tao Project, that is represented by a folder in the project explorer view of Eclipse environment.

MKEE allows to separately manage the static and dynamics part of the system, dominating so the intrinsic complexity related to the double structure. For such a reason two different editors based on Eclipse GMF technology has been used.

The Multimedia Knowledge Eclipse Environment for Tao (MKEE4TAO), illustrated following, allows the multimedia designer to define the whole system static structure, in terms of TeleAction Objects, in simple and intuitive way. all it takes is selecting in the tool palette the elements that is wanted to insert and to position them on the Canvas, through drag and drop Operation. It is necessary, besides, with the purpose to give consistence to the system, to plan the ownerships of every inserted element, using the Property view, which allows a practical insertion of the values using a tabular structure. It is possible to sail the whole Canvas using the Outline view in the angle in low to the left (see fig. 5). In analogous way it is possible to use the Multimedia Knowledge Eclipse Environment for Index Cells (MKEE4IC), to draw the index cells that constitute the system knowledge structure.

2.6 The Db2 on competition project

The db2 Competition Campus project is involved Eclipse Italian community three university and IBM in the activity to develop to a system and open source for the competition of students and inherent thematic on the date base and in particular Db2.

The db2 Competition Campus project, is born from the intuition to create a virtual community of practical that beyond finding the skill on a specific argument can exercise a dissemination activity and a good practice of learning through the competition characterizes them or to squares. The project is born from the collaboration of various international units of search. Sponsored from IBM it sees been involved the unit of Naples Federico II for the development of the system, the university of Cordoba in Argentine for the realization of the ER model and several other developers scattered in the world. The role carried out from the Italian community Eclipse open source has been crucial for the organization of the development and the constitution of the practice community developers. In particular its task has been that one to coordinate the phase of definition of the detailed lists and formalization of the same ones being been involved 15 teams of development for a total of approximately 70 students of according to year of the course of bachelor in engineering electronic in the within of the plan within the course of programming I. Following the fig. 6 shows the project architecture up to now we are in advanced developing phase

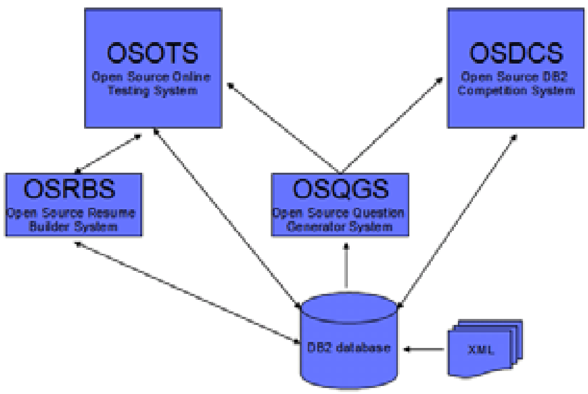


Figura 2 Db2 campus competition portal project

2.5 The other Eclipse projects

Recently other Eclipse Italian projects started. Here we can only spend some words in order to describe them since the same projects will be described in a more detailed way in a future papers.

Opera 21 and Eclipse Italian community are working together into web services field in order to manage some kind of transformation inside web services process representation. Another project is started with the aims to create some kind of documentation and formation for Eclipse evangelism courses to use inside university and/or industries belonging to the community, the project is named *Eclipse documentation*. The web site project is ever alive at [2] together with the wiki that hosts a lot of working projects (such as USBD, DB2 on competition, etc.) and the other services that assist user to acquire more info aboutopen source community (Eclipse- it XX conference free proceedings, forums, Eclipse fre documentation and so on)

3. Conclusions and future development

The Eclipse Italian community is about three years old. The community received numerous recognition and awards. Particularly the Federico II community has been received in December 2005 the best plug-in development (MKEE) in the workshop at IBM-Tivoli Lab Rome (December 2005). In the 2006 a member of community, prof. F. Lanubile from university of Bari, received an IBM award for the best innovation plug-in developed. In 2006 and 2007, the Federico II community coordination received an IBM country project award-sponsorship for the first international Eclipse conference-Eclipse-IT 07 (Napoli (IT) and a server for the community. Moreover we measure the community progress also in terms of other community involved in our mission. Particularly the community involved are: Rational User Group Italia (RUGI), DB2 User Group Italia (DUGI), Lotus User group Italia, Java Italian User Group, Project Management Institute (PMI)- south chapter. Actually we have many group working in Eclipse Italian Community, other than Napoli, coming from Bologna, Trento, Salerno, Bergamo. Together we are developing commons project and the next workshop will be held in Bari next November 2008. The industries involved in Eclipse Italian community are IBM-Tivoli lab Roma, IBM Milano, Alcatel, Opera21, CRIAI.

Bibliography

[1] Eclipse Community, http://www.eclipse.org

[2] Eclipse Italian Community, http://eclipse.dis.unina.it/

[3] Open Source Observatory, http://www.cnipa.gov.it/

[4] Mc Breen, P., *Software Craftsmanship: The New Imperative*, N.Y., Addison Wesley, 2001.

[5] Benchimol, G., *A Strategy to Innovate*, 2007. Cogprints http://cogprints.org/5551/

[6] www.ottantaventi.it/perle/teamgenioartistico1.htm

[7] Stark D., *For a Sociology of Worth*, October 2000. Working Paper Series, Center on Organizational Innovation, Columbia University. Available online at *http://www.coi.columbia.edu/pdf/stark_fsw.pdf*.

[8] Putnam R., *E Pluribus Unum*: Diversity and Community in the Twenty-first Century ,The 2006 Johan Skytte Prize Lecture

[9] *P. Maresca*, First International conference on eclipse technologies Eclipse-IT 07", Eclipse: a Great Opportunity for Industry and Universities in Italy, Paolo Maresca Editor, Cuzzolin Editore, pp. 1-131 , Napoli, 4-5 Ottobre 2007.

[10] Second Life, http://www.secondlife.com

[11] *Shi-Kuo Chang, Zhoulan Zhang[1], Colin J. Ihri Paolo Maresca and Valentina Ternelli*, Transformations for Rapid Prototyping of Time-Critical Applications, SEKE 2008, to be published.

Acknowledgements

The Eclipse Italian Community thanks IBM Italy to have hosted the test project Eclipse-IT-SL on IBM Eternal City Island. The Eclipse Italian community also thanks Eclipse.org and IBM for their support in the Eclipse Italian community ecosystem growing.

Enhancing Rationale Management with Second Life Meetings

Andrea De Lucia, Fausto Fasano, Rita Francese, and Ignazio Passero

Dipartimento di Matematica e Informatica, Università degli Studi di Salerno, via Ponte don Melillo 1, Fisciano (SA), Italy
adelucia@unisa.it, ffasano@unisa.it, francese@unisa.it, ipassero@unisa.it

Abstract

In this paper we investigate how Second Life meeting can be used to discuss software decisions and capture rationale information during synchronous meetings, especially among geographically distributed software engineering teams. To this aim, a system supporting the management of collaborative activities in Second Life, named SLMeeting, has been enhanced with rationale management features and integrated with ADAMS, a software project and artefact management system. We also present the Rationale Management subsystem of ADAMS, enabling to capture, update and retrieve rationale information. This integration supports the interchange of rationale information among the communication infrastructure, the issue base, and the project management system.

***Keywords**: Rationale Management, Second Life, Meeting Management, Collaborative Virtual Environment*

1. Introduction

Many companies manage projects that involve people from different teams and other companies around the world. Meetings are the only mechanism enabling the effective resolution of issues and the building of consensus [1]. Decisions for resolving issues are usually taken during meeting. However, project documentation describes only the results of decisions. The recording of rationale information captures trade-offs and discarded alternatives that would not be considered in the traditional documentation. Recording the rationale of each decision is an investiment: we record a decision during a phase of the process which will be revised later.

Managing rationale is an complex and time consuming activity that encompasses documenting rationale, capturing and maintaining rationale models [9], managing communicating about rationale, negotiating issues, and resolving arising conflicts [1][2][7]. As a consequence, tools are strongly needed to support the software engineer during each phase of this process.

Since its beginnings in the 1980s, Computer Supported Collaborative Work (CSCW) emerged as a multidisciplinary field, enabling people to meet their colleagues [10]. In particular, CSCW scope ranges from managing collocated meetings to redefining the way people work across time and space boundaries, from enhancing the work of small groups to supporting large groups of people. Along the way it has always been a field where computer and social scientists investigate how using synchronous and asynchronous computer-based tools to cooperate.

Asynchronous modality concerns the team cooperation and management. In particular, it is focused on the monitoring of the activity status, documentation management, scheduling, work breackdown structure management. Configuration management tools help in solving coordination problem, controlling the access to the artefacts during their lifecycle.

Synchronous modality aims at supporting the meetings in the resolution of their main drawback: their cost and the difficulty of their management.

Several research effort has been devoted to support meeting. In particular, 3D interfaces are a very popular base for groupware. Indeed, the metaphor of meeting rooms is adopted by several tools, such as [3], [15], [21], [23], [24]. These tools represent human participants with avatars. Generally they do not offer particular features for the meeting control and set-up. Anyway, it is not an easy task to create such an environment: it is necessary to create a world model, animate avatars which have to temporally and spatially share the environment, and to implement communication facilities.

The goals is to exploit the opportunity offered by the technological revolution in 3D virtual worlds [8] to obtain a simple, low-cost setup that is affordable for everyday use and integrates synchronous and asynchronous team and rationale management.

In this paper we investigate how Second Life meetings can be used to discuss software decisions and synchronously capture rationale information, especially among geographically distributed software engineering teams.

2. Related work

Documentation of rationale behind decisions was already a diffused practice more than 35 years ago, typically to manage design of building and cities. This practice has been proposed to manage software engineering design decision in the 1980s and, recently, it has been adopted to manage activities other than design. Most of the available tools for rationale management adopt an argumentative approach. The first approach proposed is IBIS (Issue-Based Information System), a way of modelling argumentation [16] aimed at addressing wicked design problems. IBIS has been implemented by gIBIS [4], a hypertext system. PHI (Procedural Hierarchy of Issues) is an extension of IBIS that introduces subissues to manage hierarchies of issues, where the resolution of an issue depends on the resolution of another. Several hypertext systems were created to support PHI [11][20] [19].

Another approach is represented by QOC (Questions, Options, and Criteria) [17] that is similar to IBIS, even if questions, diversely from issues, are intended to deal exclusively with features of the artefact being designed. QOC is more expressive than IBIS, introducing and making explicit the assessed criteria within the argumentation process.

DRL (Decision Representation Language) [12] can be considered an extension of IBIS similar to QOC. With respect to these approaches, DRL introduces new types of relations between rationale elements and allows the creation of hierarchies, similarly to PHI. The authors of DRL also implemented SYBIL [13], a knowledge-based hypertext system to support collaborative use of DRL.

In [7] Dotoit and Paech proposed an integrated process with dedicated guidance for capturing requirements and their rationale during meetings using a web-based interface. The Rationale model they proposed is similar to ours. In [2] Boulila et Al. describe an environment supporting synchronous collaborative object-oriented software design process. They enables developers to collaboratively create diagrams and synchronously attach to diagrams rationale information.

3. Rationale Management in ADAMS

In this section, after a brief description of the functionalities offered by ADAMS, we present the adopted Rationale Model and the Rationale management features offered by the newly developed ADAMS Rationale Management subsystem.

3.1 ADAMS overview

ADAMS enables the definition of a process in terms of the artefacts to be produced and the relations among them [5]. A great emphasis on the artefact life cycle is posed by associating software engineers with the different operations they can perform on each artefact. This, together with the resource permissions definition and management, represents a first process support level and allows the Project Manager to focus on practical problems involved in the process and avoid getting lost in the complexity of process modelling, like in workflow management systems.

ADAMS also supports quality management by associating each artefact type with a standard template and an inspection checklist, according to the quality manual of the organisation. Each template can be customised for specific projects or artefacts according to the quality plan. As well as the artefact standard template, standard checklist can be associated with each artefact type and used during the review phase of the artefact life cycle. Software engineers are also supported in the creation and storing of traceability links between artefacts.

The support for cooperation is provided through typical configuration management features. In fact, ADAMS enables groups of people to work on the same artefact, depending on the required roles. Different software engineers can access the same artefact according to a lock-based policy or concurrently, if branch versions of the artefact are allowed. A fine-grained management of software artefact enables to manage each single element of the document as a different entity and versioning is provided for both the atomic entity changed and the entities it is contained in. This approach allows the definition of more detailed traceability links.

The system has been enriched with features to deal with some of the most common problems faced by cooperative environments, in particular context awareness and communication among software engineers. A first context-awareness level is given by the possibility to see at any time the people who are working on an artefact. Context awareness is also supported through event notifications: software engineers working on an artefact are notified whenever relevant events happen to the artefacts they are interested in.

3.2 Rationale Capturing, Storing, and Navigation

ADAMS rationale management tool has been implemented by extending the artefact management functionalities of the system. Rationale elements are treated as artefacts, while relations among them are defined in terms of traceability links. The ADAMS

Rationale Management subsystem manages the following items:

- the ***issues*** that are addressed. An issue is a concrete problem which does not have a single solution;
- the ***options*** or *alternatives* considered as answers to an issue;
- the ***criteria*** used to guide decisions;
- the ***arguments*** that are opinion expressed by a person who agrees or disagrees with an option;
- the final ***decision*** resolving the issues.

Special stereotypes are used to manage specific relations between the rationale elements (e.g., meets / fails between options and criteria or addresses / resolves between options and issues). Rationale elements can also be related to standard artefacts developed within a project (e.g., to the functional requirements raising the issue or the design decision description). To this aim the software engineer can select the type of artefact and use a filter in order to reduce the number of links proposed by the system.

The elements (standard artefacts and/or rationale elements) recovered according to the specified parameters are used to propose new links. The software engineer can uses stereotypes to specify the relation type. Rationale elements can be also visualised in the graph-based view of ADAMS, by selecting the checkbox Rationale Items. Figure 1 shows some rationale elements which are visualised in the graph-based view.

The rationale management tool has a form based interface. Figure 2 shows, as an example, the interface which enables the software engineers to address open issues that need to be investigated by the team. The form based interface enables to classify the rationale items and to set up their relationships.

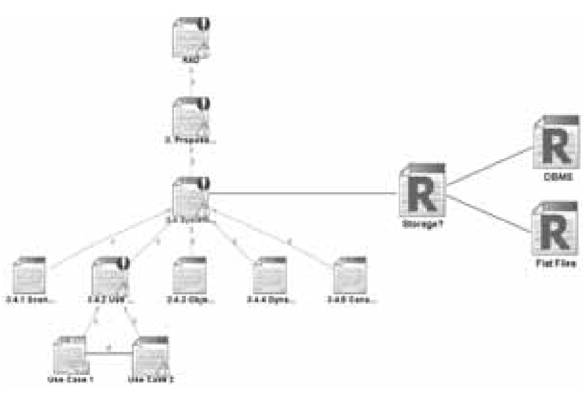


Figure 1. Rationale Items in the graph-based view

Options can be proposed to address issues and arguments can be linked to the options. Once a decision is reached the issue is closed and can be accessed in the close issue view. During the issue resolution process, project criteria already defined can be linked to the rationale elements and new criteria can be defined using the web interface and the relationships among rationale elements can be graphically resumed as depicted in Figure 1.

Figure 2. Issue management in ADAMS

4. Enhancing Rationale Capturing with SLMeetings

In this section we describe how we have enhanced Second Life to be not only a means for social interaction support, but also (beyond that) a tool for CSCW and for synchronous rationale capturing.

4.1 SLMeeting

The participants of a meeting should be able to express their ideas and make decisions. They should dispose of the needed material, such as documents, meeting minutes, presentations. Additional features supporting coordination and awareness services are required to coordinate and to solve possible conflicts between collaborative entities involved in the session.

To this aim we proposed a system, named SLMeeting [6], which enhances the communication facilities of Second Life by supporting the management of collaborative activities which can be organized as conference or Job meeting and later replayed, queried, analyzed and visualized.

Once in the environment, people have a first person perspective, they participate, do not only watch. Situational awareness, “who is there”, is well supported, as well as awareness on “what is going”. The adoption of this technology enables group awareness, i.e. a participant exactly knows who is present to the meeting and what is happening [6].

It is also an interesting feature offered by Second Life the possibility of seeing when a user does not actively participate: in case he/she does not follow the discussion the avatar visibly sleeps. SLMeeting enhances Second Life to support workspace awareness ,"How did it happen?". To this aim, the information exchanged by the meeting participants is recorded for later consultation.

The meeting organization and management functionalities are performed by ad-hoc objects created using the development framework offered by Second Life. The counterpart Web 1.0 of the meeting management is a web site, which communicates with the SL objects and automatically records the meeting minute and all the information concerning the event.

As shown in Figure 3, an SLMeeting environment is equipped with two types of objects: *boards* and *gesture bars*.

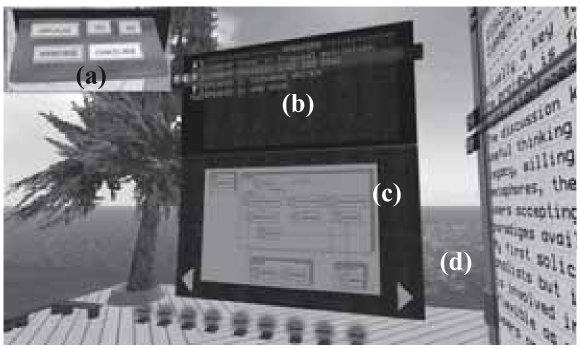


Figure 3. Some SLMeeting objects

Boards are useful to display the information needed for making decision. In particular, we equipped the environment with:

- an *Agenda*, a panel showing the title of the meeting and lists items and the corresponding speakers, shown in Figure 3(b). The meeting facilitator clicks on this object and selects the next item action. Then the item is highlighted and the assigned speaker can start his/her talk.
- The *Meeting Chat board*, reproducing the text typed in the chat by the speaker. It is saved in the Web site. This board is equipped with a button for going backward and forward in the chat. A detail is shown in Figure 3 (d)

To support floor control two kinds of gesture bar are available: the *facilitator command* bar, used by the meeting facilitator for manage the meeting and to coordinate the participant interventions, while the *participant gesture bar* offers the features of *applause*, *yes*, *no*, and *hand rising*, as shown in Figure 3(a). When the hand rise button is pressed, the participant is added to the booking list. Pressing the cancel button it is possible to erase the booking. Only the facilitator has the permission to manage the information displayed on the boards.

A *slider viewer* object enables the projection of slides, shown in Figure 3 (c). The speaker can change the slide by clicking on this component. Only the speaker and the facilitator have this permission.

Let us note that animations and gestures are offered by Second Life to augment face to face communication. As a drawback, gestures are somewhat counter intuitive to use and may even cause disruptions to conversations when a user searches for the appropriate gesture or animation to augment his/her communication [14]. Thus, we introduced gesture bars not only to support floor control, but also to improve user presence sensation. Indeed, a user is facilitated in control his/her avatar actions and is less distracted on concentrate himself on how to perform the action. In this way the sensation of "being there" is augmented.

4.2 Integrating SLMeetings and the ADAMS Rationale Management Subsystem

In Figure 4 an example of Meeting setting equipped for capturing rationale is shown. Let us note that we selected an outdoor setting, because after the construction of a typical meeting room, we noticed that the avatar and camera movements and the potentiality of flying were strongly bounded by such a working environment.

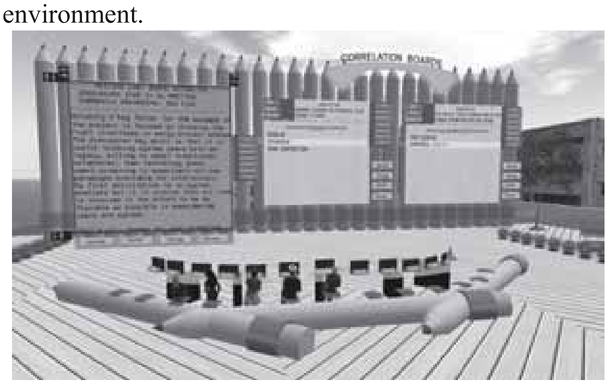


Figure 4. Rationale SLMeeting Setting

A meeting participant, maybe the facilitator, is responsible for selecting rationale items and creating the relationship between them. To this aim the Chat Board, shown in Figure 5 has been equipped to capture rationale as follows: a text selector is available to identify the rationale item between the words or sentence displayed on the board. The rationale item is selected on the chat board acting on the arrows shown in Figure 5. The selected text is classified as a rationale item using the rationale buttons. In particular**,** *six rationale buttons* are available, namely issues, proposals, alternative, criteria, arguments, and resolution.

Figure 5. SLMeeting Chat board

For examining the relationships among the selected items, two boards are proposed, named *correlation boards*, a detail is depicted in Figure 6. Each board is equipped with the rationale buttons. When a button is pressed, all the items of that type which have been previously classified during the meeting are shown. As an example, to relate the issue “Interface” to the proposal “Graphic Display” it is necessary to select issue on the first board and proposal on the second, to highlight “Interface” on the first board and “Graphic Display” on the second and press the *link* button on the first board. It is also possible to show the rational item object diagram automatically generated by ADAMS.

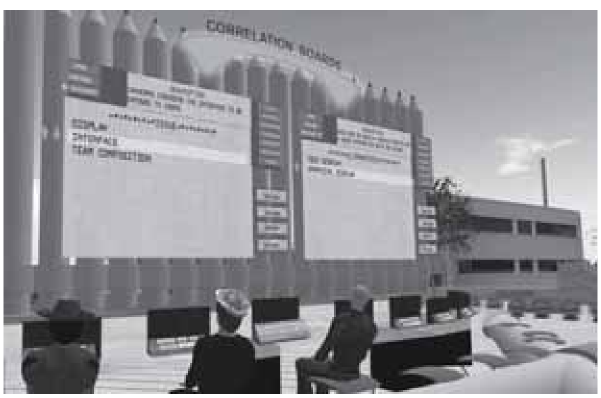


Figure 6. SLMeeting Rationale Correlation boards

To support workplace awareness, conversations and decisions are automatically saved for later references. To this aim, SLMeeting Objects in Second Life send data to the SLMeeting server outside SL using HTTP requests to PHP pages, as shown in Figure 7, where the overall SLMeeting architecture is depicted. The server then accesses the database and provides the required information back to the objects which handle it by LSL scripts.

The Rationale information is sent to the ADAMS repository following the structure depicted in Figure 7 for later retrieval.

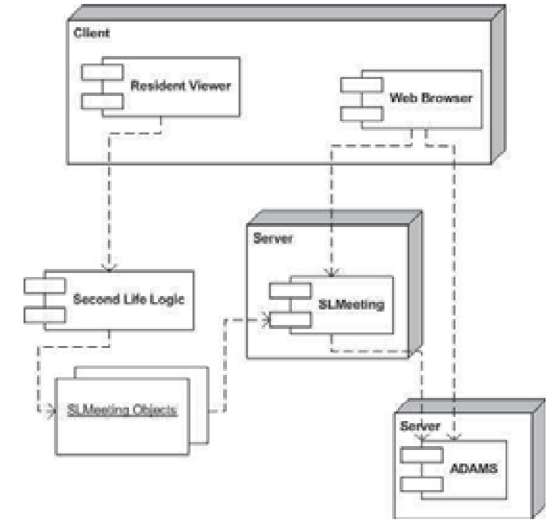


Figure 7. The proposed architecture

All the scheduled participants are invited to adhere to the meeting through ADAMS. Participants have to communicate their adhesion and their SL identities to enable access control to the meeting area and to schedule and control their interventions. To reach the meeting, participants can use a link on ADAMS to directly teleport themselves to the Second Life meeting room, or access directly from SL.

5. Conclusion and future work

In this paper we have presented a system supporting the management of rationale through the integration of an artefact-based process support system, managing asynchronous cooperative work, and a system managing synchronous meeting. As a result, knowledge management processes, such as the development of shared understanding and organizational memory can be easily managed and retrieved, versioning of this information can be maintained, and traceability links can be used to relate the knowledge arose during the discussion to the actual artefacts implementing it. At the present we are extending SLMeeting to support the retrieval and display of Rationale information stored in ADAMS for consultation while a meeting occurs.

Acknowledgment

This work is supported by the project METAMORPHOS, funded by MiUR (Ministero dell’Università e della Ricerca) under grant PRIN-2006-2006098097.

References

[1]. Bruegge B., Dutoit A. H., Object-Oriented Software Engineering: Using UML, Patterns and Java, 2nd Edition Publisher: Prentice Hall, Upper Saddle River, NJ, 2003.

[2]. Boulila N., Dutoit A., H., Bruegge, B., SCOOP: A frame work for supporting Sybchronous Collaborative Object-Oriented Software Design Process (2004). Oesterreichische Computer Gesellschaft.

[3]. Carlsson, C., Hagsand, O., "DIVE - a Multi-User Virtual Reality System", In Proceedings of IEEE Virtual Reality Annual International Symposium (VRAIS'93), Seattle, WA, September1 993, pp. 394-400.

[4]. Conklin, J. and Begeman, M. L. 1988. gIBIS: a hypertext tool for exploratory policy discussion. In Proceedings of the 1988 ACM Conference on Computer-Supported Cooperative Work (Portland, Oregon, United States, September 26 - 28, 1988). CSCW '88. ACM, New York, NY, 140-152. DOI= http://doi.acm.org/10.1145/62266.62278

[5]. A. De Lucia, F. Fasano, R. Francese, G. Tortora, "ADAMS: an Artefact-based Process Support System", in Proceedings of 16th International Conference on Software Engineering and Knowledge Engineering, Banff, Alberta, Canada, 2004, pp. 31-36.

[6]. De Lucia, A., Francese, R., Passero, I., Tortora, G. (2008). SLMeeting: Supporting Collaborative Work in Second Life. (2008) To appear in the Proceedings of Advanced Visual Interface AVI08.

[7]. Dutoit, A., H. and Paech, B., Rationale-Based Use Case Specification. Requirements Eng (2002) Springer-Verlag London Limited7:3–19.

[8]. Edwards, C., Another World. IEEE Engineering & Technology, December 2006.

[9]. Geyer, W., Richter, H., Fuchs, L., Frauenhofer, T., Daijavad, S., and Poltrock, S., Team collaboration space supporting capture and access of virtual meetings. In Proceedings of the 2001. International ACM SIGGROUP Conference on Supporting Group Work (Group 2001), ACM Press, New York, NY, 2001, 188 – 196.

[10].Jacovi, M., Soroka, V., Gail Gilboa-Freedman, Sigalit Ur, Elad Shahar, Natalia Marmasse, The Chasms of CSCW:A Citation Graph Analysis of the CSCW Conference. In the ACM Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work, 2006.

[11].Fischer, G., McCall, R., Morch, A., "JANUS: Integrating Hypertext with a Knowledge-Based Design Environment", Proceedings of Hypertext'89 (Pittsburgh, PA), ACM, New York, November 1989, pp. 105-117.

[12].Lee, J., Extending the Potts and Bruns model for recording design rationale, Proceedings of the 13th international conference on Software engineering, p.114-125, May 13-17, 1991, Austin, Texas, United States

[13].Lee, J. 1990. SIBYL: a tool for managing group design rationale. In Proceedings of the 1990 ACM Conference on Computer-Supported Cooperative Work (Los Angeles, California, United States, October 07 - 10, 1990). CSCW '90. ACM, New York, NY, 79-92.

[14].Lockner, M. and Winroth, U., Usability and Social Interaction in a Collaborative Virtual Environment. http://www.eurescom.de/~public-webspace/P800-series/P807/results/Usability/R7/D2-T3-Usability-R7-SocialInteraction.pdf

[15].Light, J., and Miller, J., D. Miramar: a 3D workplace. In the IEEE Proc. IPCC, 2002., 271-282.

[16].Kunz, W., Rittel, H., Issues as Elements of Information Systems, Working Paper 131,. Center for Urban and Regional Development., Univ. Calif. at Berkeley, 1970.

[17].MacLean, Allan, Richard M. Young, Victoria M. E. Bellotti, Thomas P. Moran. Questions, Options, and Criteria: Elements of Design Space Analysis. Moran, Carrol (eds.), Design Rationale - Concepts, Techniques and Use. Lorence Erlbaum Publishers, N.J. pp.295-321, 1996.

[18].McCall, R.J. (1991). "PHI: A Conceptual Foundation for Design Hypermedia". Design Studies 12 (1): 30-41.

[19].R. McCall, P. Bennett, P. d'Oronzio, J. Ostwald, F. Shipman, N. Wallace, "PHIDIAS: A PHI-Based Design Environment Integrating CAD Graphics into Dynamic Hypertext", Proceedings of the European Conference on Hypertext (ECHT'90), 1990

[20].McCall, R., Mistrik, I., Schuler, W., "An Integrated Information and Communication System for Problem Solving". Proceedings of the Seventh International CODATA Conference, Pergamon, London, 1981.

[21].Nijholt, A., Zwiers, J., Peciva, J., The Distributed Virtual Meeting Room Exercise, In the Proceedings of the Workshop on Multimodal multiparty meeting processing (ICMI 2005), Trento, Italy.

[22].Nunamaker, J.F. et al. Electronic meeting systems to support group work, Communications of the ACM, 34(7):40-61, July 1991.

[23].Rosenman, M., Merrick, K., Maher, M., L., and Marchant D., Designworld: A Multidisciplinary Collaborative Design Environment Using Agents In a Virtual World. Design Computing And Cognition, Springer Netherlands (2006).

[24].Yankelovich, N. et al. Meeting Central: Making Distributed Meetings More Effective. In the Proc. of ACM Conference on Computer Supported Cooperative Work (CSCW 2004), November 6-10, 2004, Chicago, Illinois, USA.

A Method to Diagnose Self-weaknesses for Software Development Organizations

Chi-Lu Yang[1, 2], Yeim-Kuan Chang[1], Chih-Ping Chu[1]
[1] Department of Computer Science and Information Engineering, National Cheng Kung University
[2] Networks and Multimedia Institute, Institute for Information Industry
[1,2]Tainan, Taiwan R.O.C.
stwin@iii.org.tw, ykchang@mail.ncku.edu.tw, chucp@csie.ncku.edu.tw

Abstract

A root cause is a source of a defect such that if it is removed, the defect is subsequently decreased or removed. By analyzing the root causes of defects of our software projects, we will be able to determine the weaknesses of our software development teams. We could thus decide on how much effort to be invested on specific actions to improve the weaknesses of the teams. In this paper, we first described how defects were objectively collected during project development. Second, the root causes were defined and categorized into six groups. Then we focused on analyzing defects to find out their root causes. Based on statistical results, the weaknesses of the project teams were determined. The results showed that the disturbing defects in our projects were mainly injected in the design phase, especially in the detail design phase. Moreover, we should invest considerable effort on enhancing our detail design skills, such as designing components, algorithms and interfaces, and so on. Some corrective actions and prevention proposals would correspondingly be acted upon and planned, respectively. Overall, we believe that our experiences and methods are worthy of sharing.

Keywords: cause analysis and resolution, root causes, defect detection, defect prevention, test cases generation

1. Introduction

Defects in software could have wide and varying effects with several levels of inconvenience to users. Some defects only result in a slight effect on the functionality. Thus, they may be undetectable for a long period of time. On the other hand, more serious defects may cause the system to crash or freeze. We should try our best to stop these defects from reaching to our customers. Defects might be discovered in different situations during project development [1], such as (1) Reports of project control with corrective actions for problems. (2) Defective reports from the customers. (3) Defective reports from end users. (4) Defects found in peer reviews. (5) Defects found in testing. (6) Process capability problems, etc. If defects are discovered, they will be fixed with specific actions at an extra cost.

A root cause is a source of a defect such that if it is removed, the defect is decreased or removed. By analyzing the root causes of defects in our projects, we will be able to determine the weaknesses of our software development team. This is useful for similar projects in the future. Based on this analysis, we could iteratively decide to invest how much effort is required on improving the weaknesses of the teams in the next project generation. To prevent defects through enhancing weaknesses of project teams is an ideal method for software development organizations.

2. Related works

A defect, synonymous with fault, is a deviation between the specification and the implementation. A defect implies a problem discovered after the product has been released to customers or end-users (or to another phase in the software life cycle) [2]. Consequently, defects reasonably make software people look bad. As software engineers, we hope to detect and prevent as many defects as possible before the customer and/or end-users encounter them.

A defect would naturally amplify in the next phase during the software life cycle, resulting in higher project cost. To detect or prevent defects injected into a product is an important task during software development. Meng [3] proposed a general framework to prevent defects. This framework consists of organization structure, defect definition, defect prevention process, and quality culture establishment. C.-P. Chang [4] proposed an action-based approach to prevent defect in software processes. Action-Based Defect Prevention (ABDP) approach applies the classification and Feature Subset Selection (FSS) technologies to project data during execution. Van Moll [5] discussed the importance of life cycle modeling for defect detection and prevention, proven methods that can be used in an efficient way. Marek [6] reported a retrospective root cause defect analysis study of the defect Modification Requests (MR) discovered. Through classifying root causes and defects, he tried to detect the defects at the earliest possible point during software development. Khaled [7] inquired the causal analysis of changes for a large system. The findings from Khaled served as input for process improvement within organizations.

The most significant challenge for causal analysis is to identify the causes of defects among large amounts of defect records where the cause-effect diagram and control chart can be utilized to support the analysis process [8]. The defect tendency is difficult to investigate when the root cause analysis schema is complicated and the sample size of the defect is small [9]. To solve this problem, the historic data on multiple releases of products can be utilized to discover the defect patterns, and be used to predict the possible defects. To decrease the effort involved in causal analysis, defect distribution can be applied to show the metrics of defects and classify them in terms of their causes [10].

For defect prevention and root causes analysis, we would propose a defect analyzing method to diagnose the weaknesses of the project teams. In this paper, we would first describe how defects were objectively collected in distinct situations during project development. Second, root causes are defined and categorized by inducting our experiences. Then we will focus on analyzing defects to determine their root causes. When we could be aware of how and when defects are injected into our projects, we would be able to prevent them as early as possible. Since weaknesses of the team could be identified, corrective actions and prevention proposals could then be enacted and planned, respectively.

3. Project Information and Test policy

The objectivity for the detection and correction of defects is greatly related to project organization. Testers and developers should have equal relationships in the project position. Both of them report to the project leader in their own channels. Accordingly, the leader coordinates or arbitrates their arguments when they have issues. Testers will be able to detect bugs in an objective manner in this hierarchy. Furthermore, they will perform document review and collect information on defects through testing based on testing plan in phases. Project Organization is showed in figure 1. Other stakeholders' definitions in the project are described as follows.

3.1 Objectively Defects Acquirement

The project leader is responsible for the project. He should plan the project and monitor its progress. His group members cooperate with each others to accomplish the project. There is a special Project Analytic Group (PAG) between the members of the project and the stakeholders of customers. PAG is led by the project leader and is made up by sub-leaders of developers and testers. PAG is organized to draw out the requirements and business models from customers. It is also responsible for high level design, which is also called architecture design in the project life cycle.

Customers provide requirements and special needs to the project. In our project, the customers also work with us to plan innovative business models. The requirement analysis is considered as one of the most important techniques in software engineering. In real projects, certain defects are incurred in this phase. We spent about 30% of the project time in analyzing requirements. Another special role in the project is the Project Quality Assurance. His mission is to make sure that the project processes are correctly performed.

The testers' main missions are to plan and perform the test plan. In the test plan, there are test strategies, test items, test cases and simulate environment, etc. After the test plan has been confirmed, the testers in different phases will perform the unit test, integration test, system test, and B-test to collect test results. One special task with testers in the project is that their representative should join the requirement and design document review. Through this early stage of participation, they would be able to understand the key requirements.

In this project, there are three developing sub-groups, which are named homebox, bio-server, and gateway. Developers in the sub-groups will design detailed functions of modules and components, which are initially peer reviewed. In addition, the developers are responsible for developing and integrating modules and components. Fixing bugs is one of their most significant tasks.

The data analyzer is responsible for analyzing and implementing system data. When there are data demands existing in the systems, the data analyzer will design table schemas in database, system configuration by XML or data flowchart between modules, etc.

The project life cycle is divided into phases, which are system analysis (SA), system design (SD), coding, testing, demo, deployment, and so on. There are different missions performed by stakeholders during the project life cycle. These missions are defined in table 1. In different phases, stakeholders perform specific missions for the project. These are accordingly shown in table 2.

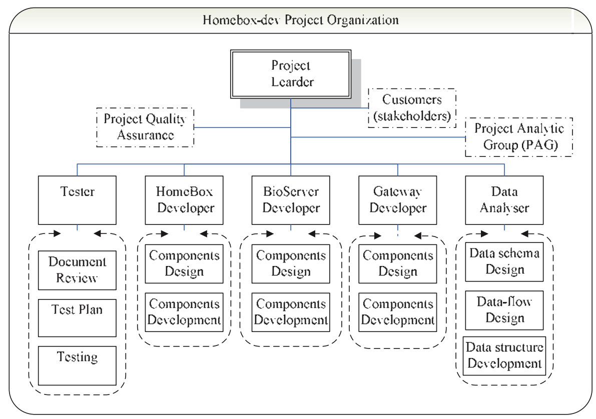


Fig. 1. Project Organization of Homebox-devices

Table 1. Specific Missions in a Project

Missions	Descriptions
MM	To Manage and coordinate work items and members of the project
DR	To Review Documents, including requirement, design, and testing documents
PR	To Plan the Requirement document and analyze requirement
PD	To Plan the Design document, high level design, and detail design
PT	To Plan the Testing document, test strategies, test plan, and test cases
AF	To Analyze the data scheme and to design the data Flowchart. (high level data analysis in design document)
AI	To Analyze and Implement details of the data scheme. (detail data analysis in design document)
RD	To Research and Develop algorithms for functions
Dbug	To fix bugs which are detected in all phases
A-test	To test functions of software and system in the Lab
B-test	To test the system in real environment

Table 2. Roles vs. Specific Missions

Phases / Roles	SA	SD	Coding	Integration & Testing	Demo / Deploy-ment
Leader	PR	DR	DR	MM	MM
Customer	DR	--	--	---	B-test
Tester	DR	DR	PT	A-test	B-test
Homebox Developer	DR	PD	RD	Dbug	Dbug
Bio-server Developer	DR	PD	RD	Dbug	Dbug
Gateway Developer	DR	PD	RD	Dbug	Dbug
Data Analyzer	DR	AF	AI	Dbug	Dbug

Defects might be injected during development and would be discovered in later phases. Such defects would be hidden in documents or systems. The document defects could be detected by peer review, regular review or milestone review. Engineers with more experience would be able to detect more precisely based from the documents. However, system defects would be detected through code review, test in lab, training course to customers or bugs report from end-users, and so on. The test policy in our projects is described in the following sub-section.

3.2 Test Case Generation and Execution

Our healthcare projects are iteratively developed in the past three years. The life cycle of the projects are divided into phases, which are customers' needs and business modeling, requirement analysis, system design, implement and integration, and deployment. As the phases progress, defects are not only discovered by reviewing documents but also detected by executing test cases, which are generated based on documents in the early phases. The life cycle progress and test cases generation are showed in figure 2.

In figure 2, there are three defined types of test cases: unit test case, integrated scenario test case, and system test case. The objective of unit test cases is to verify the correctness of unit functions, which were built from design documents. On the other hand, the objective of integrated scenario test cases is to verify requirements or special goals, which were combined from several relative unit functions. Lastly, the objective of system test cases is to verify the full executing system. Unit test cases are generated at the design phase, while integrated scenario test cases are created by requirement documents. System test cases are used to validate original customers' needs.

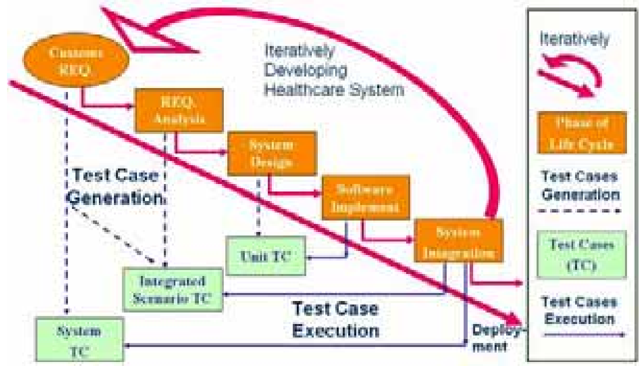


Fig. 2. Generating & Executing Test Cases to uncover defects

Every test case includes the following columns: test case ID, testing item, test step (each step contains pre-defined input data/frame, pre-defined output (pass/fail criteria), practice output, etc.), environment requirement (software and hardware), testing results, causes generally analysis, special needs of testing process, pre-test case ID, tester signature, tester/developer manager signature, remark, and so on.

When defects however are detected in later phases such as the deployment phase, even if it is discovered by engineers or customers, it would still require more cost to trace the root causes (when and how) of the defects. Therefore, the project cost would extremely increase.

4. Cause Analysis and Resolution

4.1 Root Causes Definition

The existing defects or problems in the project result from certain root causes during project development. After the defects or problems are injected, they would be detected or discovered in later phases. Engineers would spend considerable effort to detect and fix them. The later the defects are detected, the more project cost will be incurred. In order to decrease the project cost and prevent future defects, we should try to address when and how defects are injected into the projects. Our major objective is to prevent and fix defects as early as possible. We will define twenty-six frequent engineering defective types by inducing our defects' root causes. The descriptions of root causes are shown in table 3.

4.2 Grouping Root Causes

The root causes of defects are aggregated through the characteristics along the phases. They are categorized into six groups, such as business model, requirement, design, test, deployment, and hardware. Each root cause exactly belongs to one group. These groups are shown in the figure 3.

Table 3. Root Causes Types

ID_RC	Root Causes Types	The Descriptions of Root Causes
B1	Missing Address in Business	In early phases, requirements or constraints which customers didn't mention, and the analyzers also missed to address them.
B2	Changing Functional Spec.	After the engineers implemented functions, customers suggest changing functional requirements or business models.
B3	Changing Non-Functional Spec.	After the engineers implemented functions, customer suggests changing non-functional requirements or business models.

R1	Missing Functional Spec.	In the analysis phase, engineers missed functions which customers had mentioned or implied in early phases.
R2	Faulty Functional Spec.	In the analysis phase, engineers planned wrong functions which customers had mentioned or implied in early phases.
R3	Faulty Interface Spec.	In the analysis phase, engineers planned the wrong interface which customers had mentioned or implied in early phases. Interface types are categorized into users interface and systems interface.
R4	Ambiguous Non-functional Spec.	Non-functional specs were lost in analysis phase. For example, performance constraint, response time, transmission rate, and services capacity.
D1	Missing Design Spec.	In the design phase, engineers missed designs of functional details which had been recorded or implied in early phases.
D2	Faulty Design Spec.	In the design phase, engineers designed wrong functional details which had been recorded or implied in early phases.
D3	Missing Exception Handler	In the design phase, engineers missed exception handlers of functional details which are either necessary or implied in early phases. It is necessary to check the user's input format.
D4	Faulty Data Schema Design	In the design phase, analyzers designed inapplicable data schema or structure. Data models do not match with customers' requests.
D5	Faulty Data-flow Design	In the design phase, analyzers designed inapplicable data-flow for the schema. Data models do not match with customers' requests.
D6	Faulty Algorithm Design	In the design phase, researchers designed inapplicable algorithms for the functions mentioned in earlier phases. Algorithms don't achieve customers' requests or requirement.
IM1	Erroneous Implementation	In the implement phase, programmers make erroneous components or modules recommended in earlier phases.
T1	Undetected Unit Test	Un-detected Unit Test
T2	Undetected Integration Test	Un-detected Integration Test
T3	Undetected System Test	Un-detected System Test
P1	Real Network Mistake	In the real network, the product works abnormally as a result of other network devices.
P2	Wrong Version Control	Defects arose from someone releasing wrong software version.
P3	Wrong Configuration Setting	Defects arose from someone operating to setup the wrong configurations in the homebox or bio-server.
P4	Improper User Operation	Defects occurred as a result of users operating the devices improperly.
H1	Server Hardware Failure	Bio-server's exceptions result in system's failures.
H2	Peripherals Failure	Peripherals' exceptions result in system's failures.
H3	OS Failure	Operation System's exceptions result in system failures.
H4	Homebox Hardware Failure	Defects arose from homebox hardware
O1	Other Undefined Types	Undefined defective types. Certain defects which did not originate from engineering development processes.

5. Experiment Results and Discussion

By reviewing documents and testing systems, defects were formally collected during the years 2005 and 2007. Along the product life cycle, certain defects were detected in the lab, while some were reported by customers. There were 256 defects which were formally recorded as shown in figure 4. The phase-distribution of the discovered defects is showed in figure 5. Here, we can observe that defects are almost normally distributed in all phases.

Subsequently, defects were also traced backwards to determine their root causes. Since all specifications are recorded in documents, we could then trace their developing progress to determine their root causes. The distribution of root causes of all defects is shown in figure 6. Here, the situation was generally surprising. Previously, we misapprehended that the disturbing defects came mainly from the requirement analysis. However, the 43% in figure 6 have shown that the disturbing defects were mainly injected in the design phase, especially in the detail design phase. This was rather interesting considering that we frequently hear from engineers that bugs usually come from careless requirement analysis. However, more bugs came indeed from thoughtless design. Although we always have to pay more attention on our testing skills for verification, we must, moreover, spend substantial efforts in enhancing our detail design skills, such as designing components, algorithms, interfaces, and so on.

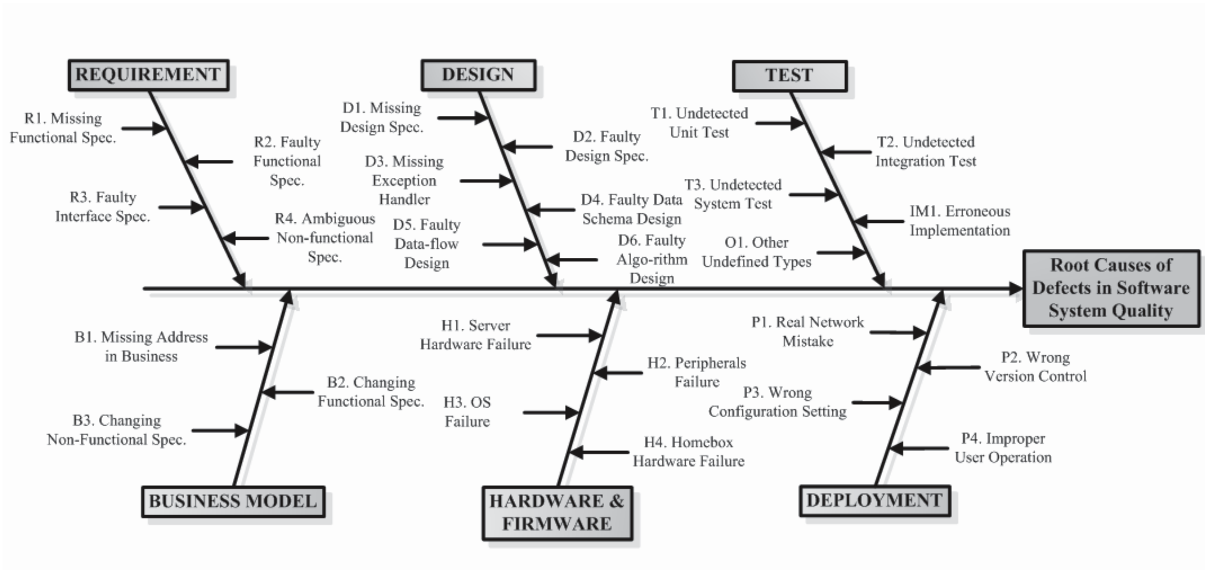


Fig. 3. Cause-and-effect (fishbone) diagrams

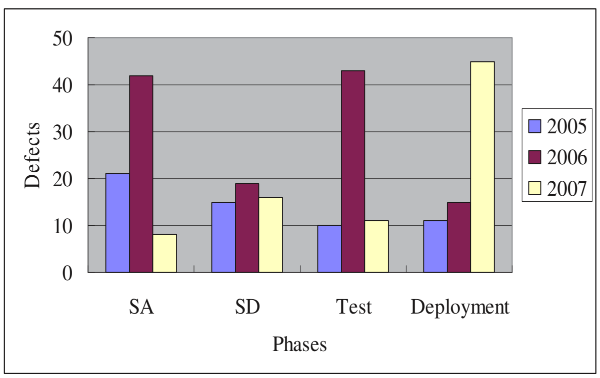


Fig. 4. Defects Collection by phases

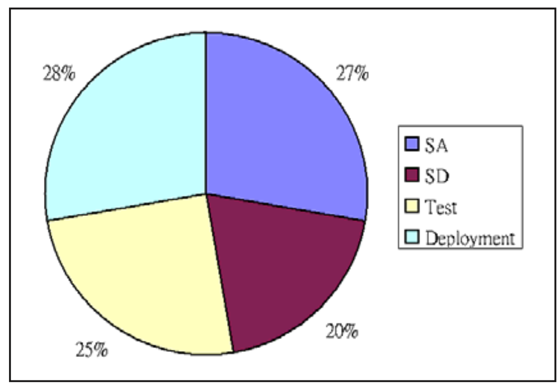


Fig. 5. The Phase-Distribution of the Discovered Defects

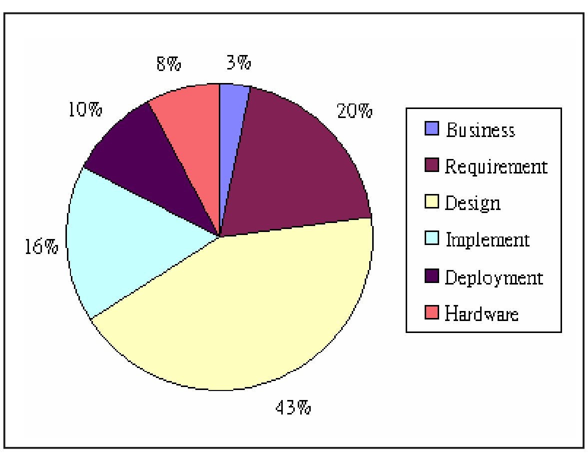


Fig. 6. The Distribution of Root Causes

In Figure 7 below are some interesting findings of our study. Figure 7 (a) shows that increasing efforts spent by engineers in SA and SD would lead to less defects detected in the test and deployment. Figure 7 (b) shows that as the efforts spent by engineers in SA increases, there would be lesser defects detected in the deployment. Consequently, less effort spent by engineers in SD would lead to more defects detected in the test. In figure 7 (c), when engineers spend less effort in SA, SD and the test, much more defects would be detected in the deployment phase.

6. Conclusion

Defects in software could have wide and varying effects with several levels of inconvenience to users. If defects are discovered, they will be fixed through specific actions with extra costs. In order to bring this cost down, preventing defects by enhancing the weaknesses of project teams is an ideal method for software development organizations. For defect prevention, we proposed a defects analysis method to diagnose the weaknesses of the project teams. We described how the defects were objectively collected during project development. The root causes were defined and categorized into six groups. Afterwards, we focused on analyzing the defects to determine their root causes. Our study has pinpointed how and when defects are injected into our projects. The weaknesses of the teams could be identified, so corrective actions and prevention proposals could then be enacted and planned, respectively. Statistical results have shown that the disturbing defects in our projects were mainly injected in the design phase, especially in the detail design phase. Moreover, considerable effort should be invested on enhancing our detail design skills, such as designing components, algorithms and interfaces, and so on.

Acknowledgements

This research was supported by the Applied Information Services Development and Integration project of the Institute for Information Industry (III) and sponsored by MOEA, Taiwan R.O.C.

References

[1] CMMI Product Team, Capability Maturity Model Integration V. 1.2, Carnegie Mellon University: SEI, Pittsburgh, USA, Aug. 2006.

[2] Pressman, Roger S., Software Engineering: A Practitioner's Approach, 6th Ed., New York: McGraw-Hill, 2005.

[3] Meng Li., He X. and Ashok S., "Defect Prevention: A General Framework and Its Application," in the Proceedings of Sixth International Conference on Quality Software (QSIC'06), Beijing, China, Oct. 2006.

[4] C.-P. Chang and C.-P. Chu, "Defect prevention in software process: An action-based approach," The Journal of Systems and Software (80), p.559-570, 2007.

[5] J.H. Van Moll, J.C. Jacobs, B. Freimut and J.J.M. Trienekens, "The Importance of Life Cycle Modeling to Defect Detection and Prevention," In the Proceedings of the 10th International Workshop on Software Technology and Engineering Practice, Washington, DC, USA, 2002

[6] Marek L., Dewayne E. P. and Eieter S., "A Case Study in Root Cause Defect Analysis," In the Proceedings of the 22nd international conference on Software engineering, p.428-437, Limerick, Ireland, 2000.

[7] Khaled E.E., Drik H. and Nazim H. M., "Causal Analysis of the Requirements Change Process for a Large System," In the Proceedings of the International Conference on Software Maintenance, Washington, DC, USA, 1997

[8] Card, D.N., "Learning from our mistakes with defect causal analysis," IEEE Software 15 (1), p56-63, 1998.

[9] Leszak, M., Perry, D.E., and Stoll, D., "Classification and evaluation of defects in a project retrospective," The Journal of System and Software (61), p173-187, 2002.

[10] Pooley, R., Senior, D., and Christie, D., "Collecting and analyzing web-based project metrics," IEEE Software 19 (1), p52-58, 2002.

[11] Chi-Lu Yang and Richard Weng, "A Method for Verifying Usability and Performance of a Multi-user Healthcare Embedded System," in the Proceedings of the Thirteenth International Conference on Distributed Multimedia Systems (DMS'07), p75-80, San Francisco, USA, Sep. 2007.

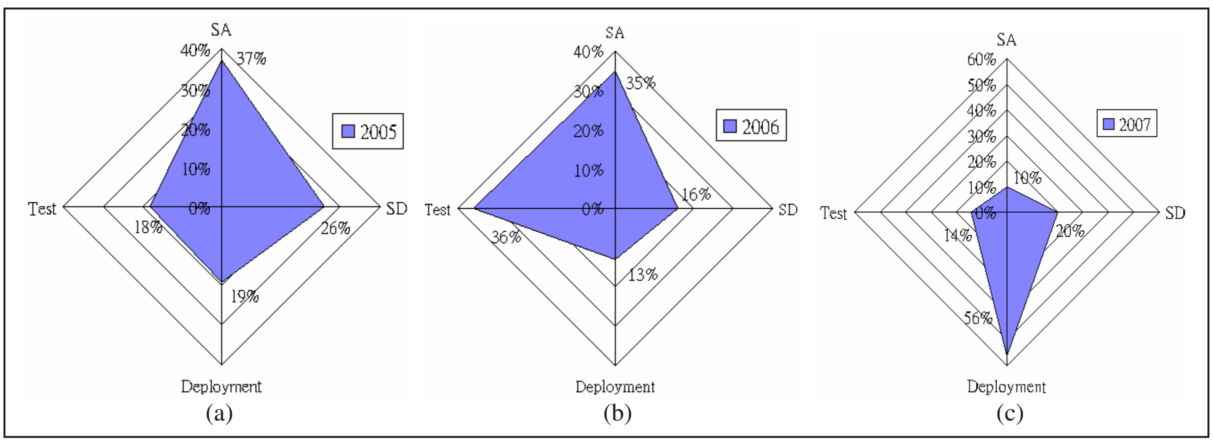


Fig. 7. Variance of Defects in Phases

A Distributed System for Continuous Integration with JINI

Y. C. Cheng, P.-H. Ou, C.-T. Chen and T.-S. Hsu
Software Systems Lab
Department of Computer Science and Information Engineering
National Taipei University of Technology
Email: {yccheng,t5598036,s1669021,s4599002}@ntut.edu.tw

Abstract

The present work extends a previously developed continuous integration system to support platform dependent builds and to reduce turnaround time by incorporating distributed computing capabilities. The two objectives are achieved by adopting the Replicated-Worker Pattern as the distributed architecture and using JINI distributed technology as the means of implementation. Experimental results in support of the claims are presented.

1. Introduction

Continuous Integration (CI) is a software development best practice to ensure software quality and maintain system stability [10]. Typical tasks performed in CI include compilation, unit tests, integration tests, and static code analysis. By integrating different software components periodically and frequently, CI allows developers to obtain feedbacks on the latest integration build for early discovery and correction of bugs. In practice, continuous integration requires tool support. Many tools have been developed, such as CruiseControl by ThoughWorks [6], Apache Gump [2] and Continuum [3], and our own Java Continuous Integration System (JCIS) [13]. Besides the functional requirements, CI tools must meet a number of *quality requirements*, including *extensibility* (adding new integration tasks), *availability* (providing integration service reliably), *maintainability* (managing platform dependent builds), *performance* (reducing turn-around time), and so on.

In this paper, we seek to achieve maintainability and performance for JCIS. In the maintainability aspect, existing CI tools perform integration builds on a single host. This means that *platform dependent builds* (e.g., builds of a Java application that runs on Windows and Linux, respectively) require the CI software to be installed on multiple platforms. An immediate consequence is that platform dependent builds become unrelated builds that must be managed separately, which creates a maintainability problem. In the performance aspect, builds that take a long time to complete constitute bottlenecks in the use of existing CI tools;the long waiting time hampers the developer's willingness to run continuous integration frequently. In our present work, both requirements are viewed as *heterogeneous distributed computing* problems, which are solved by applying the Replicated-Worker Pattern [7] and the JINI technology [9].

2. Related Works

We survey several well-known open source CI projects, including CruiseControl [6], Anthill OS [1], Gump [2] and Continuum [3]. These projects are briefly overviewed and compared with JCIS in the following list.

- CruiseControl supports both Ant and Maven. CruiseControl lets user customize the building process. Multiple hosts can be used to perform distributed builds to cut the turn-around time provided that each of the hosts has a copy of CruiseControl installed. Reports on builds are collected to a central dashboard to show overall result. CruiseControl is not based on distributed architecture.
- AntHill is based on Ant and is easy to install and configure. The open source version of AntHill does not support distributed builds.
- Gump supports both Ant and Maven. It allows dependency to be specified among several projects. Gump does not support distributed builds.
- Continuum supports Ant and Maven and is easy to install and use. It provides a web portal for managing integration builds. Continuum does not support distributed builds.

Our survey shows that these open source CI tools are not yet sufficient in achieving both the maintainability and performance requirements. This motivates us to extend JCIS achieve both requirements using an appropriate distributed computing technology.

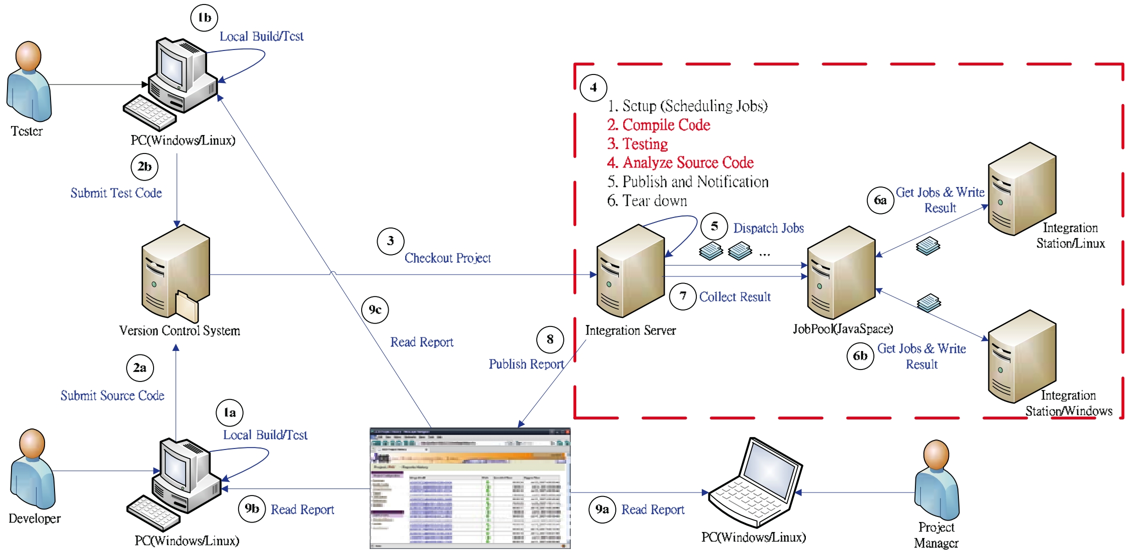


Figure 1: the process of the distributed CI system

3. System Architecture and Design

From our previous work, JCIS was developed to provide automatic continuous integration. JCIS is extensible through its IBuilder interface. For example, both C++ and Java program are supported by adding the corresponding compilers to JCIS through the IBuilder interface. Figure 1 extends JCIS integration scenario by including distributed computing:

1a/1b. Developers write, compile and test the program on their own computers./Testers check out the current version of the program from the version control system, write and execute tests, e.g., integration tests and functional tests.

2a./2b. Developers commit their programs to the version control system after passing all unit tests./Testers commit the test code to the version control system after pass all the test cases.

3. Integration Server checks out the current version of the program and test code automatically, and starts the integration builds.
4. Customized integration builds are performed according to project specification. The basic flow includes compiling and testing. Additional builds such as various static analyses are available from JCIS as well.

5~7. Integration Server dispatches the building jobs to different integration stations. After all of the integration stations finish their jobs, Integration Server aggregates the building results to produce an integration report.

8. Integration Server publishes all the artifacts, including the object code, documents and the integration report to the portal.
9. Project Manager, Developers and Testers view the integration report from the portal with a browser.

The Replicated-Worker Pattern [7] is applied to provide the distributed computing capability (Figure 2). This pattern defines three roles: *Master*, *Worker*, and *Space*. The master works with several workers. It is responsible for dividing the integration task to several smaller sub-tasks and for putting the sub-tasks to the space. The workers then take the sub-tasks from the space and execute the sub-tasks. Upon finishing, a worker writes the execution result back to the space. Lastly, when all sub-tasks are completed, the master collects and aggregates these execution results.

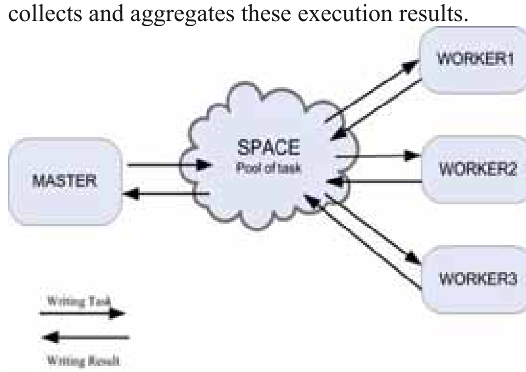


Figure 2: Replicated-Worker Pattern

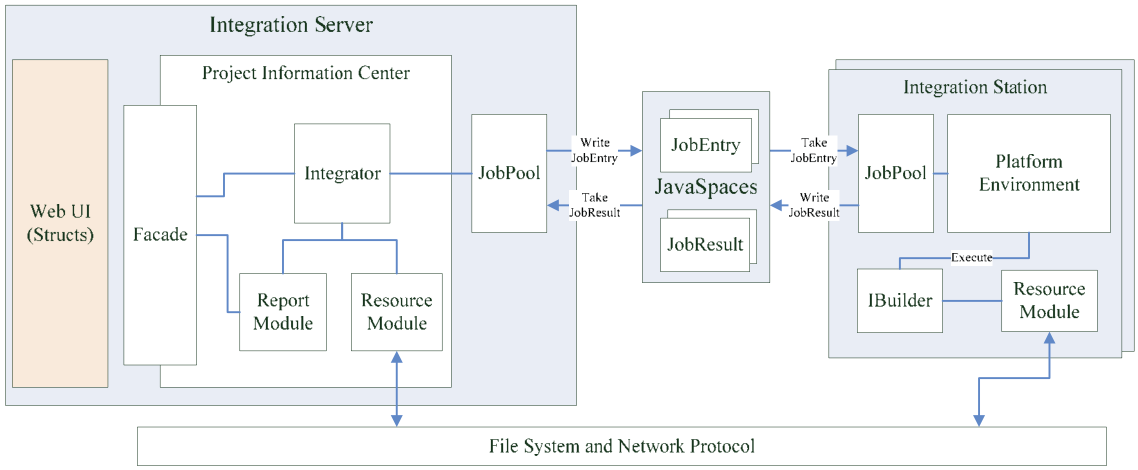


Figure 3: DJCIS System Architecture

We use the JINI distributed technology [9] to implement the Replicated-Worker pattern. Figure 3 shows the DJCIS architecture with JINI. The Integrator plays the Master role; the JavaSpaces plays the Space role; and the Integration Station plays the Worker role. First, the Integrator divides an integration task into a number of sub-tasks by builder types. For example, an integration build involves compiling and unit testing is divided into two sub-tasks of the corresponding types. Metadata (e.g., URL of resources such as source code, object code and executable of a builder) for a sub-task are packed into a JobEntry object, which is then written into the JavaSpaces through the JobPool object. The integration makes the resource available on a web server. An Integration Station retrieves a JobEntry from JavaSpaces, unpacks it, retrieves the resources and performs the build. Upon completion, the Integration Station writes the build result back into JavaSpaces as a JobResult object. The actual build results are compressed and uploaded to the web server. Figure 4 shows the collaborations that take place during a build.

In DJCIS, a builder (e.g., GCCcompile builder, StatSVN builder, etc.) performs a single integration task. All builders implement the IBuilder interface. The design allows the Integration Station to launch the builders without knowing its specific implementation. Therefore, even if a new builder is added into DJCIS, an Integration Stations can simply retrieve the corresponding builder using the information provided in the JobEntry.

Sequencing of the job entries is controlled by the Integrator. Sequencing is required when a build makes reference to other build results. To this end, the job entries are grouped by dependency relationships into *parallel* and *sequential* groups. Job entries inside the parallel group are put into JavaSpaces simultaneously for parallel execution; Job entries inside the sequential group are handled one by one. In DJCIS, the dependencies are specified with a XML configuration file.

The proposed design of DJCIS achieves maintainability. A JINI-facilitated Integration Station can run on Windows or Linux. Thus, platform dependent builds can be defined as an integration task that contains the same builds that run on different platforms. The proposed design of DJCIS also achieves performance: turnaround time can be cut by adding more Integration Stations. This is further explored with three experiments.

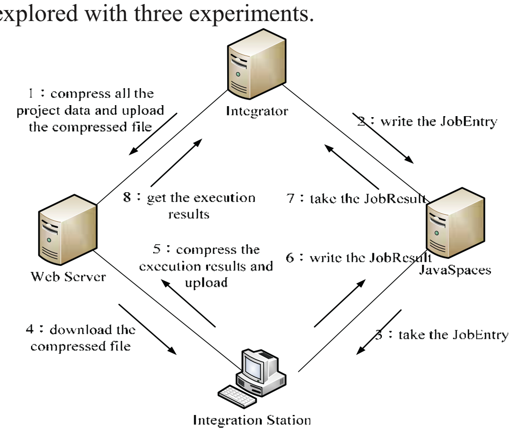


Figure 4: Collaborations inside integration

4. Experiments on performance

We design three experiments to illustrate the performance improvement achieved by DJCIS.

As shown in Figure 4, DJCIS incurs many overheads in going distributed, including the time of putting the jobs into JavaSpaces, the time of compressing and uncompressing project resources, and the time of downloading and uploading compressed files. The present experiments will ignore the time of putting the job entry or job result to JavaSpaces since it is relatively small and hard to measure.

The development case under study is a WiMAX system developed by the Software Development Research Center of the National Taipei University of Technology [14]. The WiMAX project contains six sub-projects code-named SYS, VIDEO, MAC, PHY1, PHY2 and CHN, which covers the complete WiMAX protocol stack from application to channel. The development language is C/C++. The target platform is Linux. Project size information is listed in Table 1. Since the sub-project teams operate mostly independently, a check-in to the version control system by any sub-project team can break the entire system. DJCIS is employed to perform daily builds for discovering such problems.

Table 1: WiMAX sub-projects information

Project Name	The Number of Files	Size(KB)	LOC(lines)
SYS	188	545	6568
VIDEO	839	8386	15839
MAC	210	1444	16472
PHY1	77	7188	3437
PHY2	89	319	4112
CHN	66	365	4622

Figure 5 shows the integration flow of the WiMAX system. For each sub-project, DJCIS performs three builds: compilation with GCC [8], unit testing with CppUnit [5] and source code size statistics with StatSVN [12]. The compilation builders and the StatSVN builders are executed in parallel. Notably, the compilation builder for SYS makes use of the object code produced by the other five sub-projects and must wait for results to become available. CppUnit builder of SYS then performs the unit testing contributed by all sub-projects.

The specification of the hosts for the Integration Server and Integration Station are as follows:

- Integration Server
  OS: Windows XP
  CPU: Intel Core2 T7200 2.00GHz
  RAM: 2048MB
- Integration Station
  OS: Fedora 8
  CPU: Intel Core2 6320 1.86GHz
  RAM: 2048 MB

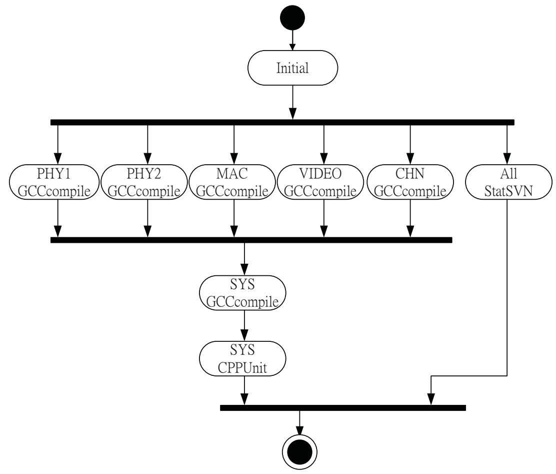


Figure 5: The integration flow of the WiMAX Project

Table 2 shows the execution time of compression and decompression performed by DJCIS for each sub-project. We found the compressing and uncompressing time overheads are relatively small. The compressing ratios are between 11%~44%.

Table 2: Compression and decompression time

Project Name	Compressing Time/s	Decompressing Time/s	Compressing Size/KB (Ratio)
SYS	0.1	0.2	226 (41%)
VIDEO	1	2	2109 (25%)
MAC	0.2	0.25	391 (27%)
PHY1	0.75	0.25	801 (11%)
PHY2	0.1	0.1	108 (33%)
CHN	0.1	0.1	162 (44%)

To provide a baseline for measuring performance speedup, Table 3 shows the average execution time (taken from 100 integration builds) of the WiMAX integration executed on sequential JCIS, the release without JINI technology. Note that total execution time the integrated project is more than sum of all the sub-projects, because extra time is spent by JCIS to collect and aggregate the execution results for generating the integration report.

Table 3: Average integration time of WiMAX project on sequential JCIS

Project Name	GCC Ccompile/s	StatSVN/s	CppUnit/s	Sub-Project Time/s
CHN	3	21	X	24
MAC	45	23	X	68
PHY1	4	20	X	24
PHY2	15	22	X	37
VIDEO	86	23	X	109
SYS	46	24	60	130
Total Time/s	422			

Experiment 1:

The integration environment contains one Integration Server and one Integration Station. The measurement of the experiment 1 is listed in table 4. We found the total time is 100 seconds (or 23%) more than that of sequential JCIS. The extra time comes from the overheads of compression, decompression, and transmission time for sending the metadata data and resource back and forth.

Table 4: Average integration time of WiMAX project on distributed JCIS with one Integration Station

Project Name	GCC Ccompile/s	StatSVN/s	CppUnit/s	Sub-Project Time/s
CHN	6	24	X	30
MAC	49	27	X	76
PHY1	7	23	X	30
PHY2	21	25	X	46
VIDEO	106	25	X	131
SYS	57	27	68	152
Total Time	522			

Experiment 2:

The integration environment contains one Integration Server and two Integration Stations. Figure 6 shows the best job allocation of the WiMAX integrated project, and total time is about 285 seconds. However, since these two Integration Stations are free in taking jobs, it can happen that one is made to wait for the other one in case there is job dependency. Figure 7 shows an example where the total time exceed 370 seconds.

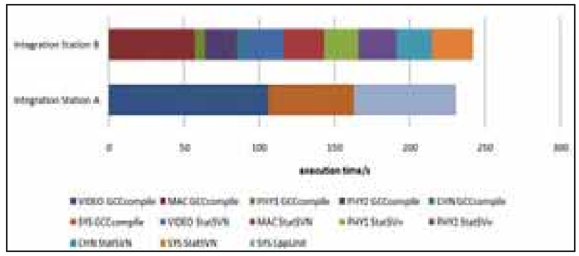


Figure 6: A best case of job allocation with two Integration Stations

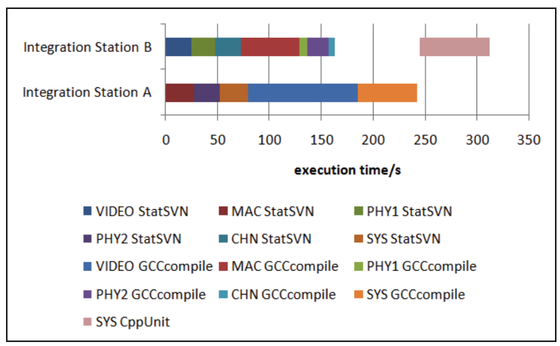


Figure 7: A worse case of job allocation with two Integration Stations

Figure 8 show the total execution time of the over 100 measurements. The average is about 280 seconds, which is about 66% of the sequential JCIS time, a significant performance improvement.

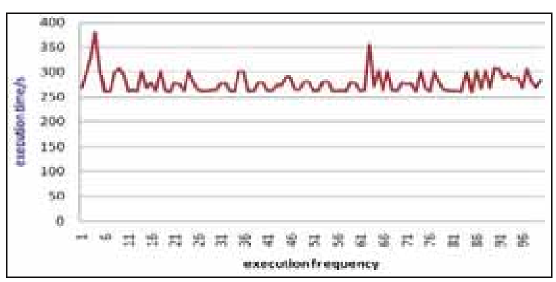


Figure 8: Total integration time over 100 measurements in Experiment 2

Experiment 3:

The integration environment contains one Integration Server and three Integration Stations. Figure 9 (respectively Figure 10) shows a best (respectively, worst) case of job allocation, which takes about 270 seconds (respectively, 330 seconds) to complete.

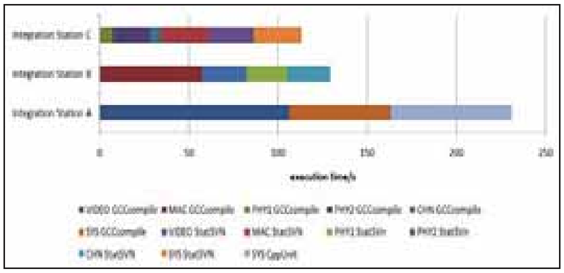


Figure 9: A best case of job allocation with three Integration Stations

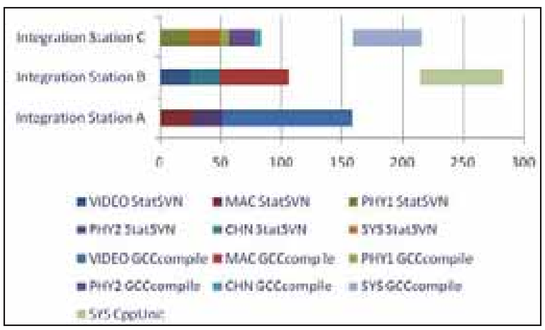


Figure 10: A worse case of job allocation with three Integration Stations

Figure 11 show the total time over 100 measurements. The average is about 281 seconds. The value is almost the same with experiment 2: the sequential group of GCC Compilation and CppUnit in SYS constitutes a bottleneck.

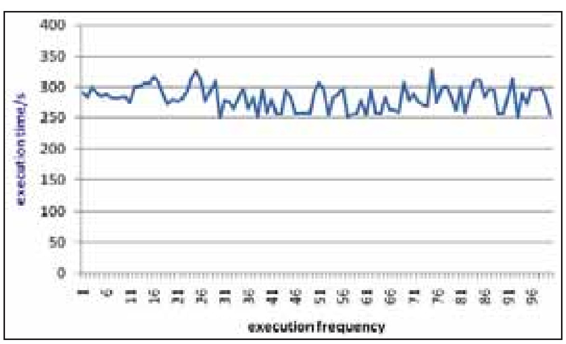


Figure 11: Total integration time over 100 measurements in Experiment 3

5. Conclusion

In this paper, we extend the JCIS architecture to incorporate distributed computing capabilities. Platform dependent builds are more easily managed and performance gains are possible due to parallel execution of the integration builds. A case study with three experiments to demonstrate speedup with distributed DJCIS has been presented. We found that the total integration time is reduced when two or more two Integration Stations are used. However, the speedup is constrained by the dependencies between integration builds.

6. Acknowledgements

This research is supported by the National Science Council of Taiwan under the grants NSC95-2221-E-027-046-MY3 and NSC96-2218-E-027-017.

7. References

[1] Anthill, http://www.urbancode.com/projects/anthill/default.jsp.

[2] Apache Gump, http://gump.apache.org/.

[3] Apache Continuum, http://maven.apache.org/continuum/.

[4] Apache Ant, http://ant.apache.org/.

[5] CppUnit, http://sourceforge.net/projects/cppunit.

[6] CruiseControl, http://cruisecontrol.sourceforge.net/.

[7] E. Freeman, S. Hupfer, and K. Arnold, JavaSpeces Principles, Patterns, and Practice, Addison Wesley, 1999.

[8] GCC, the GNU Compiler Collection, http://gcc.gnu.org/

[9] Jini Network Technology, http://java.sun.com/developer/products/jini/index.jsp.

[10] Kent Beck, *Extreme Programming Explained: Embrace Change*, Addison-Wesley, 1999.

[11] Maven, http://maven.apache.org/.

[12] StatSVN, http://www.statsvn.org/.

[13] Chia-Hao Wu, Tsui-Chen She, Chien-Tsun Chen, and Y. C. Cheng, JCIS: Supporting Platform Dependent Builds, Proceedings of the Third Taiwan Conference on Software Engineering, Taichung, June 8-9, 2007. (in Chinese)

[14] W.K. Chen et al., WiMAX System Software and Tool Development, National Science Council (Taiwan) Open Source Project NSC95-2218-027-009.

Transforming UML Sequence Diagrams to Class Diagrams

Yen-Chieh Huang[1,2], Chih-Ping Chu[1]
[1]*Department of Computer Science and Information Engineering,*
National Cheng-Kung University, Tainan, Taiwan
[2]*Department of Information Management*
Meiho Institute of Technology, Pingtung, Taiwan
E-mail :p7894121@mail.ncku.edu.tw

ABSTRACT

Using Unified Modeling Language (UML) to represent Object-Oriented Analysis/Design has become a standard for information system development. UML defines many different models to present information systems. In these models, there are some interdependent relations and overlapping parts. Many approaches and CASE tools have been introduced to facilitate the transformation between diagrams, but they do not define and describe detailed mapping steps. In this study, we propose a procedure which transforms sequence diagrams into class diagrams by using a metamodel. We propose a model transformation system that uses an XML specification to describe UML models. First, we map the sequence diagram into its metamodel, and then we use transformation rules, XSL file format, that can transform between different metamodels. Finally, we map the output metamodel file to its graphical view, and then transform it to a class diagram. This research can help software engineers for faster developing system.

Keywords: UML, XML, CASE, Transformation

1: INTRODUCTION

UML [2, 4, 7] as been accepted as an industrial standard for specifying, visualizing, understanding, and documenting object oriented software systems. UML provides different diagram types supporting the development process from requirement specifications to implementations. OOSA&D [1] generally begin from requirement analysis to system analysis. Later, it distinguishes two design methodologies, one is to implement static structure modeling, the method draws class diagrams first and follows with interaction diagrams; the other is to implement dynamic structure modeling, the second method draws interaction diagrams first and follows with class diagrams. To build all diagrams might take much time. The most usage UML diagrams are class diagrams, sequence diagram, use case diagrams. In this paper we study the relationships of sequence diagrams and class diagrams. We propose a supporting method to transform sequence diagram into class diagrams by metamodel and transformation rules.

2: Literature Review

Over the past few years a considerable number of studies have been made on UML research, research work related to single transformation operations between different diagrams types has been carried out by several authors [5, 9, 10]. Schonberger et al. proposed some algorithm steps about collaboration diagrams transfer to state diagrams [8]. Selonen et al. [10] used the metamodel to define their transform operation. All UML diagrams can be described to the metamodel. They proposed theoretical transformations between UML diagrams, but they do not define and describe detailed mapping steps. This paper extends their methodology and implements how to transform from sequence diagrams to class diagrams.

Tetsuro Katayama [5] proposed a supporting method of diagrams generation, which used the transformation rules in UML to draw diagrams eliminated contradiction in a short time by using common elements which other diagrams have. He proposed class diagrams can be transformed from other diagrams and implemented use case diagrams transformed to class diagrams in the new system. In his research, he compared the number of elements in diagrams. A percentage inside the parenthesis in the column of total shows the rate of elements by transforming. The same method in this paper, we take the case study of class diagrams generated with transformation rules from sequence diagrams. This paper proposes a model transformation system that uses XML specification to describe UML models and verify the transformation procedure proposed can work currently.

A metamodel is simply a model of a modelling language. It defines the structure, semantics, and constraints for a family of models. Selonen et al. used the metamodel to define their transform operation. All UML diagrams can be described to the metamodel.

XMI is a model driven XML Integration framework for defining, interchanging, manipulating and integrating XML data and objects. Peltier et al. [6] used XMI specification to represent all models and metamodels. XSLT could be used to transform models, which are represented in the XML formalism [3, 6].They used XSLT to transform models. The core of our method is the definition of transformation rules between the metamodel of diagrams.

So this paper use metamodel to define UML diagrams. The source diagram can save to metamodel schema with XML, then XMI can transform XML metamodel to another target metamodel by ourselves defined XSLT language, then the target diagram can restore by some CASE tools.

3: Techniques for UML Diagram Transformation Operations

3.1: Transformation Processes

Many CASE tools can use XML to save UML diagrams. But every company has its own XML format; they cannot interchange with each other. Only the same CASE tool of some UML diagrams can interchange with each other.

This research used the CASE tool Violet to draw UML diagrams. Because Violet is a free and open source UML CAES tool, it can save UML diagrams to XML documental format. In this paper, we design the XSLT for transforming from the metamodel of sequence diagrams to the metamodel of class diagrams. Then Violet opens the XML output file, and class diagrams can be automatically generated. The XSLT process flow chart is shown in Figure 1.

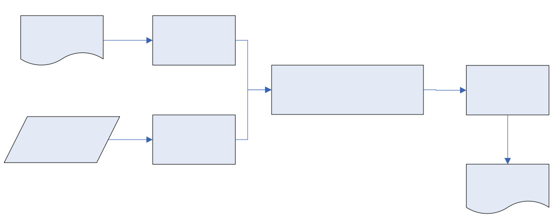


Figure 1 XSLT process flowchart

This paper adopt Tetsuro Katayama's (2002) method to design the table from sequence diagrams transform to class diagrams, some elements can be transformed from sequence diagrams to class diagrams, but there are many other elements which cannot be transformed. For example, attribute is an element of class diagrams, but sequence diagram has no detailed attributes in it. So attributes must be added to class diagrams after the sequence diagrams transformation. The attributes can be found from other information, such as Use Cases description, Drawing, Data Glossary and so forth.

3.2: Transformation Rules

Selonen et al. proposed two approaches to find an interpretation of a transformation. One is *push approach*, the other is *pull approach*. In this paper, we follow his rules and add one rule to transform. These approaches are the following:

1. Direct mapping (push)

Every element of source diagram could be mapped to an element of destination diagram. The transformation rules using push approach is listed in Table 1.

Table 1 Transformation rules using push approach

Sequence Diagrams	Class Diagrams
Object	Class name
Message	Relationship between classes
Call message name	Operation in accepted class
Return value	Attribute in accepted class
Condition or constraint	Notes

2. Indirect mapping (pull)

Every element of target diagram is transformed from some one element of source diagram. The transformation rules using pull approach are listed in Table 2.

Table 2 Transformation rules using pull approach

Sequence Diagram	Class Diagram
An object sends messages to many objects that have the same type	One to many multiplicity
Many objects send messages to the same object	Many to one multiplicity
Many objects send messages to many objects that have the same type	Many to many multiplicity
An object sends a created messages and a destroyed messages to the other object	Composition relationship
An object sends a created messages to the other object	Aggregation relationship
Some objects have the same attributes	Generalization relationship

3. The other mapping

The elements can not be push or pull from source diagram, which can be obtained from other information. For example, the attributes of class can be found from Drawing or Data Glossary or Use Cases description, they can be summarized to the third mapping.

3.3: Implement Transformation

To compare the sequence diagrams and class diagrams of XML format, this paper developed a XSL transformation program. Sequence diagrams is opened by the CASE tool XMLSPY, then import the XSL file to generate the output file, another CASE tool Violet open the output file, class diagrams is transformed to a template. Class name and operations might automatic generate, but attributes must be added manually.

4: Case Studies

4.1: Study Environment

To illustrate the feasibility of the proposed methodology, a web-based book ordered system was developed. In this case study, we focus on the transformations from sequence diagrams to class diagrams. On the whole, a sequence diagram can be transformed to a class diagram. For example, while we develop the web-based ordered system, which has a use case diagram and has five use cases in it. The five use cases can draw to five sequence diagrams; then every sequence diagram can be transformed to a class diagram, and then gathered the five class diagrams to become a system class diagram.

Figure 2 shows a sequence diagram of 'add new order item' which was drawn with Violet, then the file is opened by XMLSPY, we imported the designed XSL file (SequencetoClassDiagram.xsl), XMLSPY can transform to an output file, we saved as the file become a XML file, then it is opened by Violet. The result shows in Figure 3. This is a prototype of class diagram, because it has class name and operations, but no attribute in it. You can find the attributes only "product List", "shoppingCart List" in Figure 3; they are only a substitute for finding the correct detailed attributes. The correct attributes of class can be found from Use Cases description, Drawing or Data Glossary and so forth.

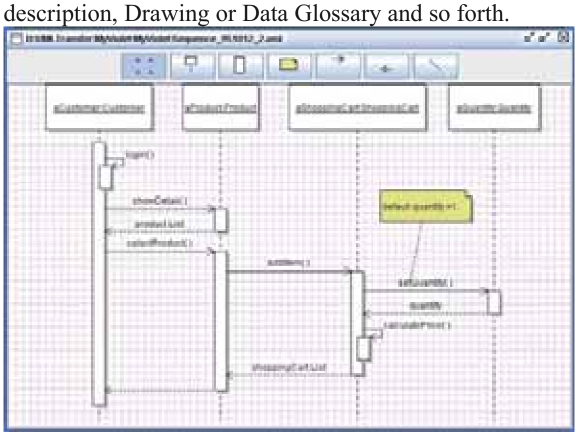


Figure 2 A sequence diagram showing the add new order item

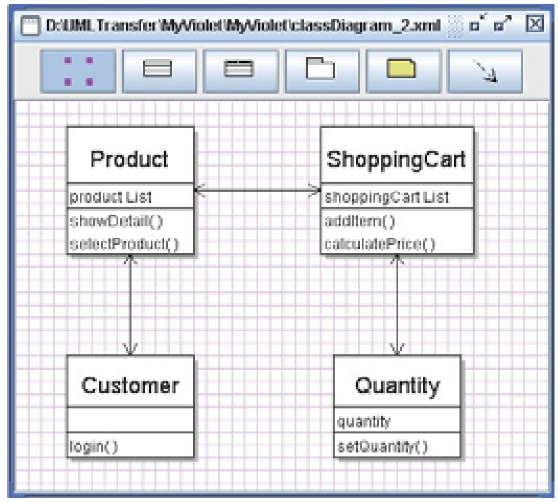


Figure 3 Add new order item class diagram transformed from sequence diagram in Figure 2

One use case can be drawn to one sequence diagram, so that the example of use case diagram can be drawn to five sequence diagrams. Every sequence diagram can be transformed to one class diagram by the designed rules, and then all class diagrams can gather to become a system class diagram. It shows in Figure 4.

The Class of Customer and Order must add their attributes are shown in Figure 5. It is a completed class diagram after adding all attributes.

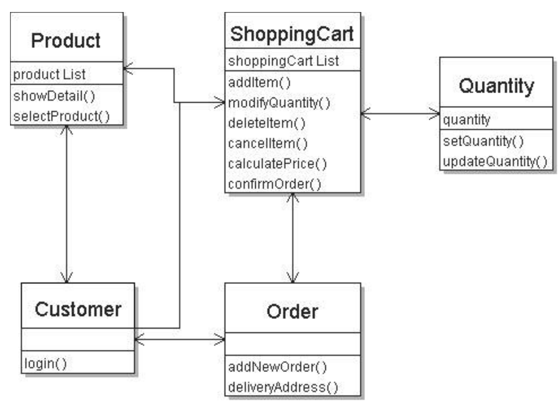


Figure 4 The system class diagram

4.2: Result Analysis

To compare the transformation result and count the number of elements between the two diagrams shows in Table 3. The column A is the number of elements automatically transformed from a sequence diagram to a class diagram, the column B is the number of elements in completed of system class diagram, the column C is the number of elements added to the class diagram manually. It shows the automatic transformation rate is sixty-five percentages; the system class diagram must be added thirteen attributes manually. Using the research of automatic transformation processes from sequence diagrams to class diagrams, following these transformation rules can save the time when system analyzers design the class diagrams.

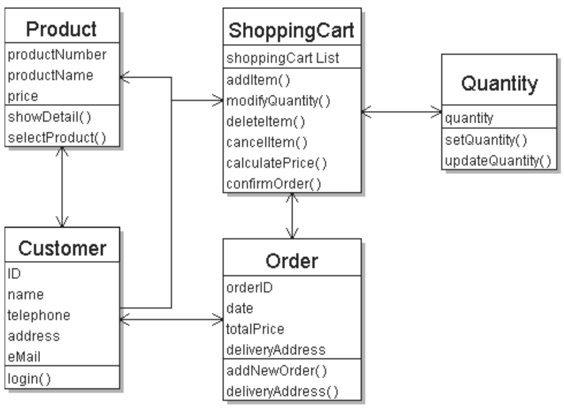


Figure 5 Completed system class diagram after adding attributes

Table 3 The number of elements before and after sequence diagram transformation

metamodel	**A**	**B**	**C**
Class	5	5	0
Attribute	0	13	13
Operation	13	13	0
Association	6	6	0
Generalization	0	0	0
total	24(65%)	37	13

A: the number of elements in sequence diagrams transformed to the class diagrams
B: the number of elements in completed class diagrams
C: the number of elements needed to add in class diagrams (C=B-A)

5: Conclusions and Future Work

The paper proposes a method to transform from sequence diagrams to class diagrams. The method defines a XSL transformation program and implement by XML format. When supporting by appropriate tools, such operations can be transformed automatically. But some elements must be added after transformation, because sequence diagrams have no attributes in it.

This paper has a demo of case study to implement this approach. There are sixty-five percent of elements can be transformed, but the class diagram is only a prototype diagram. In addition, the multiplicity must be added manually, it is not automatically transformed in this case study. There are some contributions in this research are listed as follows:

1. We can depict a new class diagram from an existing sequence diagram. If the system we are modeling is very complex, the transformation can speed up the development process and avoid or reduce the possibility of human error.
2. If somebody is drawing class diagrams before sequence diagrams. We can use the new class diagram to compare with the existing class diagrams in order to check the consistency of diagrams.
3. Using the CASE tools Violet, XMLSPY and so forth, they can help us to design the modeling diagrams.

If the CASE tools support more functions, we can show the UML notation more correct. In the future, someone might enhance the functions of the CASE tool Violet or develop another CASE tool. If new CASE tools can support more UML diagrams, we can directly transform more information between different diagrams. These issues will be topic of our future research.

REFERENCES

[1] Bennett, S., McRobb, S. and Farmer, R., *Object-Oriented Systems Analysis And Design Using UML Second Edition*, McGraw-Hill, 2002.

[2] Booch, G., Rumbaugh, J. and Jacobson, I., *The Unified Modeling Language User Guide*, Addison-Wesley, 1999.

[3] Elliotte, R. H., and Scott, W., M., *XML In a Nutshell*, Third Edition, O'RELLY, 2004.

[4] Fowler, M., and Scott K., *UML Distilled: A Brief Guide to the Standard object modeling Language*, Third Edition, Massachusetts, Addison Wesley, 2003.

[5] Katayama, T., Proposal of a Supporting Method for Diagrams Generation with the Transformation Rules in UML, Proceedings of the Ninth Asia-Pacific Software Engineering Conference, 2002.

[6] Peltier, M., Bézivin, J., and Guillaume, G., MTRANS: A general framework, based on XSLT, for model transformations, Workshop on Transformations in UML, Genova, Italy, pp. 93-97, April, 2001.

[7] Rumbaugh, J., Jacoboson, I., and Booch, G., *The Unified Modeling Language Reference Manual*, Addision Wesly, Massachusetts, 1999.

[8] Schonberger, S., Keller, R., and Khriss, I., "Algorithmic Support for Model Transformation in Object-Oriented Software Development", Concurrency and Computation: Practice and Experience, Volume 13, Issue 5, pp. 351–383, 2001.

[9] Selonen, P., Koskimies, K., and Sakkinen, M., How to Make Apples from Oranges in UML, Proceedings of the 34th Annual Hawaii International Conference on System Sciences (HICSS-34), Volume 3, 2001

[10] Selonen, P., Koskimies, K., and Sakkinen, M., Transformations between UML diagrams, Journal of Database Management, 14(3), pp. 37-55, Jul-Sep, 2003.

Goal-Driven Design Transformation by Applying Design Patterns

Yong-Yi FanJiang[1] Wen-Ting Lee[2] Jonathan Lee[2]

[1] Department of Computer Science and Information Engineering, Fu Jen Catholic University
[2] Department of Computer Science and Information Engineering, National Central University

Abstract

Recently, the growing popularity of model driven frameworks and methodologies has implied an increasing focus on model transformation. Meanwhile, the impact of design patterns on software quality has attracted a gradual attention since design patterns encapsulate valuable knowledge to resolve design problems and improve design quality. As an attempt towards the investigation of applying goals and design patterns to realize the design refactoring, we propose, in this paper, a goal-driven design refactoring by applying design patterns to transform an analysis model into its corresponding design model with an emphasis on the non-functional requirements. The use of goals makes it easier to transform the functional and non-functional requirements into the software models, and derives the candidate design patterns to help satisfy non-functional requirements for resolving the design problems and improving software quality.

Keywords: Design patterns, model transformation, design quality, nonfunctional requirements.

1. Introduction

Recently, the growing popularity of model driven framework and methodologies, as well as the Model Driven Architecture (MDA) [8] initiated by Object Management Group (OMG) has implied an increasing focus on model transformation. Model transformation takes a set of model fragments, which conform to the source metamodel, as input and applies a set of rules to them, resulting in a set of models, which conform to the target metamodel [20]. Meanwhile, the software quality has long been recognized as an important topic since the early days of software engineering. In the past, researchers and practitioners alike have examined how systems can meet specific software quality requirements. Therefore, a growing number of practitioners have shown great interests in using design patterns towards high-quality software, since design patterns represent high-level abstractions that reflect the experiences of no other than skilled practitioners themselves. Design patterns have become a popular means to encapsulate object-oriented design knowledge. They capture successful solutions to recurring problems that arise when building software systems.

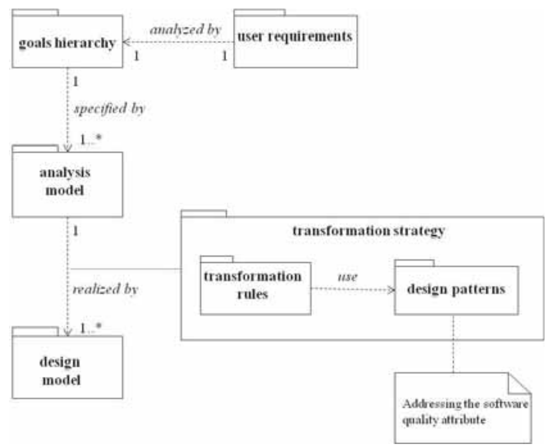


Figure 1: An overview of our proposed approach

Generally, a transformational approach has two major steps. First, a formal specification is derived from user requirements. Then, a set of transformations is applied to these formal specifications to obtain an implementation system. In this paper, we adopt our pervious work on the goal-driven use case (GDUC) [16, 17] and fuzzy object-oriented modeling (FOOM) [18] served as the formal specifications to capture and specify imprecise requirements, and provide a set of pattern-based transformation rules to deal with the software quality issues. As an attempt towards the investigation of applying goals, design patterns, and model transformation approach, a goal-driven model transformation by applying design patterns approach has been proposed (see Fig. 1 for an overview). In Fig. 1, the designer constructs the goals hierarchy based on goals analysis and goals interaction. Before applying the design pattern transformation rules, the designer selects a suitable design pattern from the candidate design patterns through matching the non-functional goals and intents of design patterns, and then transforms the analysis model into design model by applying the transformation rules incrementally.

Our proposed approach has the following features:

- Goal-driven use cases and goals hierarchy are built according to the goals and use cases analyzing. The analysis model is constructed based on the goals hierarchy by using the FOOM notations and semantics to specify various alternative models.
- Design patterns serve as a supporting design knowledge base to help the development of a system. A goal-driven method is provided to facilitating the selection of candidate design patterns from the analysis model, and the design pattern transformation rules are used to helping the enhancement of non-functional requirements and improve the software quality.

The organization of this paper is as follows. We first briefly describe the background knowledge about model transformation and applying design patterns to model transformation in the next section. The relationship between design patterns and non-functional issues is depicted in section 3, and the section 4 provides the transformation rules. A meeting scheduling system is provided as an example to illustrate our approach in section 5, and some concluding remarks are given in section 6.

2. Background

A number of researches in the area of model transformation and pattern based model transformation have made their marks on our goal-driven pattern-based model transformation method.

2.1. Model transformation approaches

Graph transformation. Varró et al [22] describe a system for model transformation based on Graph Transformations. This style of transformation is based on the application of an ordered set of rules. Operators available include transitive closure, and repeated application. Rules identify sub-graphs which define before and after states, and may refer to source and target model elements and introduce associations between these two models.

Relational transformation. The basic idea is to state the source and target element type of a relation and specify it using constraints. In its pure form, such specification is non-executable. However, declarative constraints can be given executable semantic, such like in the case of logic programming. In fact, logic programming with its unification-based matching, search, and backtracking seems a natural choice to implement the relational approach, where predicates can be used to describe the relations. Gerber et al. [11] explore the application of logic programming to implement transformations.

Model transformation through high-level programming language. Implementing a model transformation can be carried out by using a general programming language with a specific API. Indeed, nowadays, repositories such as dMOF [4] can be used to save models and metamodels that are MOF compliant. These tools allow an API to be generated and its basic implementation for each contained metamodels. Thereafter, programmers can use these APIs to create, manipulate and modify models that conform to a metamodel contained by the repository.

Model transformation based on transformation metamodel. Engineers build models to better understand the systems that are being developed. In a similar way, to understand existing models we may provide models of these as well. This activity is called meta-modeling. K. Duddy et al. [5] defined a transformation metamodel of which instances will be input parameters of a generator, which transforms a model into a program that implements the transformation process described by this model. These generated programs are based on the dMOF repository.

Table 1 Comparison of researches on design patterns

	objective	Engineering	Pattern selection	Pattern application	Pattern usage	Abstraction level
Budinsky *et al*. 1996 [1]	auto. Pattern application	forward eng.	N/a	script language	N/a	coding level
Eden *et al*. 1997 [7]	auto. pattern application	forward eng.	N/a	meta-programming	N/a	coding level
Jeon *et al*. 2002 [14]	auto. pattern selection	forward eng.	structure-based (hot-spot)	rule-based transformation	N/a	coding level
O'Cinneide and Nixon 2000 [3]	re-restructure legacy system to provide flexibility	re-eng.	evolution-based (precursor)	layered transformation	improve flexibility	coding level
Tahvildari and Kontogiannis 2002 [20]	re-restructure legacy system to improve quality	re-eng.	requirement-based	layered transformation	improve quality	coding level
Khriss *et al*. 2000 [15]	provide traceability by using patterns	forward eng.	N/a	layered transformation	refine model	design level
Huston 2001 [13]	examine the compatibility between design patterns and metrics	N/a	N/a	N/a	improve quality	N/a
Gross and E. Yu 2000 [10]	provide a development methodology	forward eng.	requirement based	N/a	deal with non-functional reqt.	requirement and design level
Chung *et al*. 2003 [2]	provide a methodology for developing adaptable system	forward eng.	requirement based	N/a	Improve adaptability	architecture level

2.2. Pattern-based model transformation approaches

Work in a number of fields has made their marks on the pattern-based approach. These researches are organized into following categories: automatic application of design patterns, quality improvement by design patterns, pattern-based development methodology, and others.

Table 1 makes a comparison of these researches from the perspectives of their objectives, engineering direction, pattern selection, pattern application and pattern usage. The first category of researches [1, 7, 14] focus on automatic pattern application and ignore the usage of design patterns. The second category of researches [3, 13, 15, 21] focus on how to restructure a legacy system to be more qualified for the assistance of pattern-based transformation skills. The third category [2, 10] attempts to providing a methodology to serve as a bridge between users' requirement and design pattern technology. They put efforts on analysis of user needs rather than automation of pattern application.

3. Design patterns and non-functional requirements

Non-functional requirements are not easy to handle because they are subjective (they can be interpreted differently by different people), relative (their importance is depending on the system being considered) and interacting (their achievement may impair or help that of others) [6, 18]. Design patterns [9] provide a possible way to deal with non-functional requirements since they provide solutions to satisfy functional requirements as well as better solutions to meet non-functional requirements. In particular, besides providing a basic, functional solution to a problem, a design pattern offers a qualified, non-functional improvement to that solution. For example, considering the original intent described in *Observer* design pattern:

Define a one-to-many dependency between objects so that when one object changes state, all its dependents are notified and updated automatically. By elaborating the intent, we can understand the design pattern is designed to resolve the communication between a subject objects and its related observer objects. Viewing from the functional aspect, it requires the subject to notify all observers when the subject changes its state. Viewing from the non-functional aspect, it requires the notification should work automatically without knowing types of observers. In other words, Observer design pattern has a FR-intent (functional requirement intent) to address functional problems and an NFR-intent (nonfunctional requirement intent) to improve non-functional quality requirements. We thus transform a design pattern's intent into functional-intent (FR-intent) and non-functional intent (NFR-intent) to highlight the quality contribution of the pattern [12]. The FR-intent describes what does the pattern do, and the NFR-intent concentrates on the extension to the FR-intent to describe what quality attribute the pattern can achieve, such as reusability, maintenance, or extensibility.

The structure distributes the object model to achieve the goal described in the intent property. With respective to the FR-intent and NFR-intent, FR-structure and NFR-structure are developed respectively. The major difference of NFR-structure to FR-structure is it applies object techniques to resolve problems. These techniques, such as polymorphism, abstraction, encapsulation, delegation, dynamic binding and inheritance are keys to make object-oriented system more reusable, extensible and maintainable. Figure 2 illustrates the NFR-structure and its corresponding FR-structure of Abstract factory design pattern. Essentially, NFR-structure is an extension of FR-structure to satisfy the associated NFR-intent. In our approach, the extension plays an important guideline to help us transfer a basic model to an extension model that is compatible with the NFR-structure.

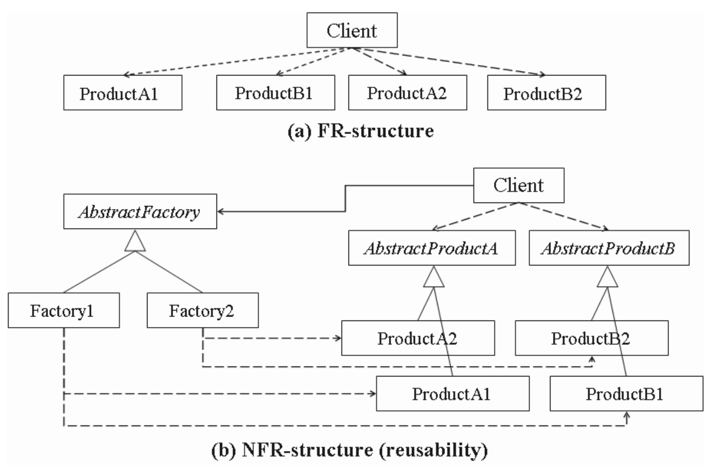


Fig. 2: The FR-structure and NFR-structure of *Abstract Factory* design pattern.

4. Transformation Rule Schema

Each refinement process is based on the application of a design pattern. A transformation is described graphically by a schema called refactoring rule schema. A transformation rule schema is parameterized by model elements to be specified by the designer, and is composed of two compartments. The first compartment describes the source model of the design while the second compartment shows its corresponding target model after application of a design pattern. Fig. 3 shows the transformation rule schema that we have defined for the Observer pattern.

Rule-name: Observer-001
Purpose: The subject object can notify the interested objects without specifying their concrete classes
Category: Non-functional requirements (extensibility, consistency)
Using pattern: Observer
Preconditions:
When a change to one object requires changing others; or when an object should be able to notify other objects without having to know the identify of these objects.
Transformation actions:
1. (*Abstraction*) For each subject and observer class creates a new superclass (abstract class)
2. (*Inheritance*) Build the inheritance relation between the superclass and its concrete class.
3. (*Operation addition*) Insert the attach() and detach() operations into the subject's super class.
4. (*Association change*) Change the association between subject and observers
   4.1 Make an association between subject's superclass to observer's superclass
   4.2 Delete the association between subject to observer

Fig. 3: Transformation rule schema of *Observer* pattern

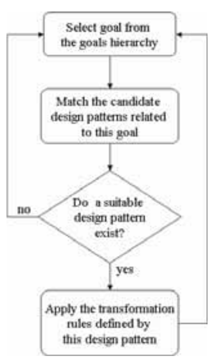


Fig. 4: Procedure of applying design patterns

The procedure of applying design patterns to transform the analysis model into design model through transformation rules is depicted in Fig. 4. After establishing the goals hierarchy for obtaining alternative models and constructing stable kernel and alternatives, the designer must specify the analysis model based on the goals hierarchy in an incremental fashion. Initially a goal in goals hierarchy is chosen that will serve as a start point for the design pattern transformation applying under our proposed approach.

According to the functional and non-functional aspects of the chosen goal, we can match one or more design patterns to deal with this goal's non-functional requirement. Designer can choose one of these design patterns suitable to match the non-functional requirement, and apply transformation rules defined with this design pattern to satisfy the non-functional goal.

5. Meeting scheduling example

In this section, we use an example - meeting scheduler system to demonstrate the idea of our approach. Goals are identified based on the proposed verb-based classification scheme. In this example, we have identified five goals and formed a goal-driven use case model (see Fig. 5).

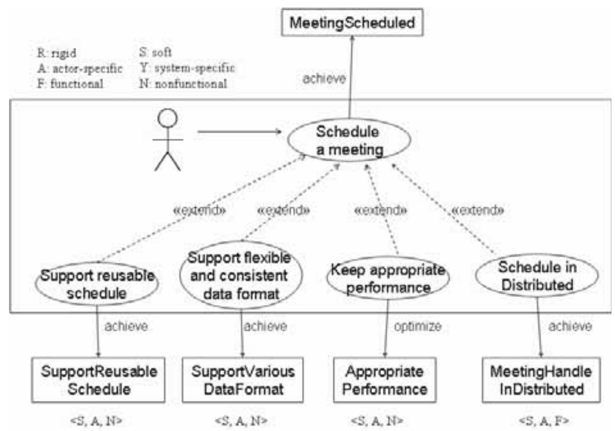


Fig. 5: Goal-driven use case model for the meeting scheduling system

- Schedule: [initiator, meeting date and location, initiator, participants, ϕ, rigid]
- Handle: [system, plan meetings, initiator, ϕ, in distributed, rigid]
- Support: [initiator, meeting date and location, initiator, participants, reusable schedule, soft]
- Support: [system, meeting date, initiator, participants, support flexible and consistent date format , soft]
- Provide: [system, performance, ϕ, ϕ, an appropriate level, soft]

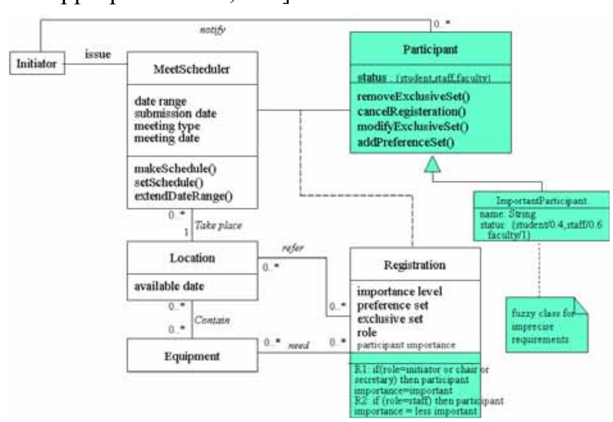


Fig. 6: Goals considered: $\{G_{MS}\}$

An initial analysis model is obtained from the goals hierarchy (see Fig. 6) and the construction of the design model is proceeded based on stable kernel in an incremental fashion from G_{MS} to G_{SFC}.

Since the G_{MHID} represent the system can be scheduled and handled in distributed manner, the designer choose the remote proxy pattern to support the distributed management (see Fig. 7).

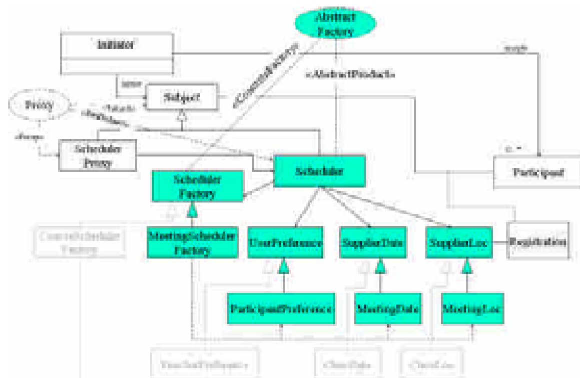


Fig. 7: Goals considered: {G_{MS}, G_{MHID}} after applying *Proxy* pattern

By considering the G_{SRS} and G_{SFC} to realize the goal *SupportReusableSchedule* and the goal *SupportVariousDataFormat* by applying the abstract factory and the observer patterns are shown in Fig. 8 and Fig. 9. Fig. 8 illustrates the resulting model after applying *Abstract Factory* pattern in the meeting scheduler system to enhance *ReusableSchedule* defined by G_{SRS}. Instead of creating the product objects directly, the *Scheduler* uses *MeetingSchedulerFactory* for creating the product objects of *ParticipantPreference*, *MeetingDate* and *MeetingLoc*. The scheduling strategy can be simply reused in a course scheduling system in the way of designing a *CourseSchedulerFactory* which creates objects of *TeacherPreference*, *ClassDate* and *ClassLoc*.

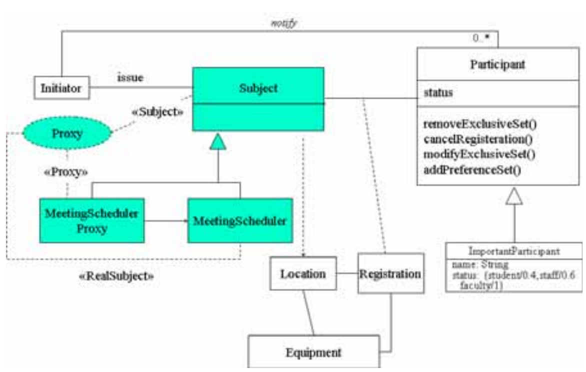


Fig. 8: Goals considered: {G_{MS}, G_{MHID}, G_{SRS} } after applying *Abstract Factory* pattern

The G_{SFC} represents the meeting scheduler system could support various data format and keep consistence between different data views, i.e., a change on one data view must make the same change on the others. To resolve the inconsistency problem, the Observer design pattern is recommended and used in Fig. 9. *Observer* pattern is used to resolve the inconsistency problem between a set of objects (said observers) which have a common interest on subject. By requiring the observers to register on the subject before operating, observers can be notified for keeping consistency whenever the subject changes its status.

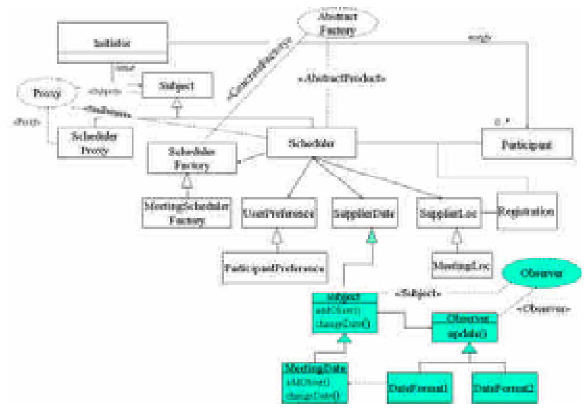


Fig. 9: Goals considered: {G_{MS}, G_{MHID}, G_{SRS}, G_{SFC}} after applying *Observer* pattern

6. Conclusion

In this paper, we have proposed an approach to providing a goal-driven method to transform the analysis model into design model for dealing with the software quality issues based on design patterns. With the aid of design patterns, in our approach, a goal can be achieved by a well-specified transformation rules, recovered by a proven solution, and enhanced by a known skill. Our approach offers two benefits: (1) deriving candidate design patterns from analysis model based on the goals hierarchy; and (2) transforming analysis model into design model incrementally based on the design pattern transformation rules for dealing with the non-functional requirements.

Acknowledge

This work was partially supported by the National Science Council (Taiwan) under grant NSC 95-2221-E-030-024

References

[1] F. Budinsky, M.Finie, J.Vlisdes, and P.Yu. Automatic code genration from design patterns. *Object Technology*, 35(2):172-191, 1996.

[2] L. Chung, K. Cooper, and A. Yi. Developing adaptable software architectures using design patterns: an NFR approach. Computer Standards and Interfaces, 23:253-260, 2003.

[3] M.O. Cinneide. Automated refactoring to introduce design patterns. In *Proceedings of the International Conference on Software Engineering*, 2000.

[4] dMOF user guide (1.1 release). Distributed Systems Technology Center (DSTC), Australia, July 2000.

[5] K. Duddy, A. Gerber, M. Lawley, K. Raymond, and J. Steel. Declarative transformation for object-oriented models. *Transformation of Knowledge*, Information and Data: Theory and Application, Idea Group, 2004.

[6] C. Ebert. Putting requirement management into praxis: dealing with nonfunctional requirements. *Information and Software Technology*, 40:175-185, 1998.

[7] A. Eden, J. Gil, and A. Yehudai. Precise specification and automatic application of design patterns. In *proceedings of the 12th IEEE International Automated Software Engineering Conference (ASE 1997)*, November 1997.

[8] D. S. Frankel. Model Driven Architecture: Applying MDA to Enterprise Computing. John Wiley & Sons, 2003.

[9] E. Gamma, R. Helm, R. Johnson, and J. Vlissides. Design Patterns: Elements of reusable object-oriented software, Reading, MA: Addison-Wesley, 1994.

[10] E. Gross and E. Yu. From non-functional requirements to design through patterns. *Requirements Engineering*, 6:18-36, 2001.

[11] A. Gerber, M. Lawley, K. Raymond, J. Steel, A. Wood. Transformation: The Missing Link of MDA, In A. Corradini, H. Ehrig, H.-J. Kreowski, G. Rozenberg (Eds.): *Graph Transformation: First International Conference (ICGT 2002)*, 2002.

[12] N.-L. Hsueh and W.-H. Shen. Handling Nonfunctional and Conflicting Requirements with Design Patterns. In *the proceedings of the 11th Asia-Pacific Software Engineering Conference*, Page(s):608 – 615, 2004.

[13] B. Huston. The efforts of design pattern application on metric scores. *The Journal of Systems and Software*, 58:261-269, 2001.

[14] S. Jeon, J. Lee, and D. Bae. An automated refactoring approach to design pattern-based program transformations in Java programs. In *the proceedings of the Ninth Asia Pacific Conference of Software Engineering*, pages 337-345, 2002.

[15] L. Khriss, R.K. Keller, and I. Hamid. Pattern-based refinement schemas for design knowledge transfer. *Knowledge-Based Systems*, 13:403-415, 2000.

[16] J. Lee and G.H. Hsu. April 2002. Modeling software architectures with goals in virtual university environment. *Information and Software Technology*, 44(6): 361-380.

[17] J. Lee and N.L Xue. July/August 1999. Analyzing user requirements by use cases: A goal-driven approach. *IEEE Software*, 16 (4):92-101.

[18] J. Lee, N.L. Xue, K.H. Hsu and S.J.H. Yang. 1999. Modeling imprecise requirements with fuzzy objects. *Information Sciences*, 118:101-119.

[19] J. Mylopulos, L. Chung, and B. Nixon. Representing and using nonfunctional requirements: A process-oriented approach. *IEEE Transactions on Software Engineering*, 18(6):483-497, 1992.

[20] S. Sendall. Source element selection in model transformation. In *UML'03 workshop in Software Model Engineering (WiSME)*, 2003.

[21] L. Tahvildari and K. Kontogiannis. A software transformation framework for quality-driven object-oriented re-engineering. In *Proceedings of the International Conference on Software Maintenance (ICSM02)*, 2002.

[22] D.V.G. Varró and A. Pataricza. Designing the automatic transformation of visual languages. *Journal of Science and Computer Programming*, 44: 205-227, 2002.

A Real-Time Scheduler Designed for Virtual Machine to Support Multiple Service Classes

Chieh-Ping Tsai[1], Mong-Fong Horng[2], Yau-Hwang Kuo[1], Jui-Yi Kuo[3], Jang-Hao Jiang[4]

Dept. of Computer Science and Information Engineering[1]
National Cheng Kung University, Tainan, Taiwan
Dept. of Electronics Engineering[2]
Institute of Information Industry, Tainan, Taiwan[3]
National Kaohsiung University of Applied Science, Kaohsiung, Taiwan
Advanced Multimedia Internet Technology, Tainan, Taiwan[4]

Abstract

In this paper, a real-time scheduler designed for virtual machine (VM) is proposed. In traditional VMs, low throughout and long response time are the typical problems encountered due to the sharing of processor time and I/O peripherals among VMs. Instead of classic first-in-first-out task scheduling, a priority-based multiple queuing design is proposed to ensure the executions of real-time tasks. The details of the proposed scheduler is presented, including scheduler modeling, operations, analysis and performance evaluation. The experimental results demonstrate valuable features of the proposed scheduler, including (1) the great improvement of the scheduling-induced delay (SI-Delay) for real-time tasks, (2) fast response time (3) feasible software solution. The developed scheme is the key to prompt the VM efficiency, in particular, for the applications with high real-time demand, such as multimedia services.

1. Introduction

Recently, Virtual Machine (VM) technology has attracted lots of interests from industrial and academic. VM is a technology to realize multiple operating system instances on a single hardware platform. As the great improvements of computation hardware technology, the computation power of processors is far beyond the need of ordinary application software. In most applications, the software execution only consumes a small portion of processor resource. Thus, in most time, processor is in the idle state and most computational capacity of processor is wasted. How to efficiently and effectively utilize the computation resource of processors has become one of significant and valuable issues in computation technology.

Traditionally, there are two types [1] of VMM designs as shown in Fig. 1. The significant difference between them is the existence of host operating system (HOS). In 1960s, Type 1 VMM (T1VMM) was developed to realize time sharing system in mainframe [2]. T1VMM operates on hardware directly to support VMs with various operating systems, called as guest OS (GOS). HOS and GOS are also denoted as executable domains, *Di*, for *i =0, 1, 2, …* . Due to the indispensable overhead of time-sharing operation between domains, T1VMM only were applied on high-power mainframe. For personal computers, T1VMM was not applicable to realize. Afterward, T1VMM was emerged by operating systems as the basic of Multi-programming [3].

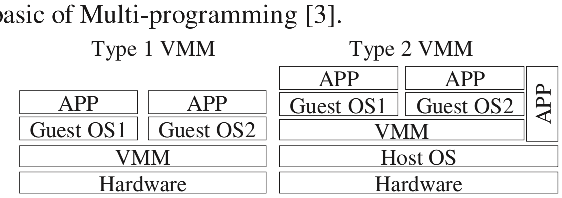


Fig. 1 Types of VMM

Since 1990s, the introduction of multi-core and multithread technology in processors conducts the development of multiple OS on a single machine. However the low performance of multithreading, Type 2 VMM (T2VMM) was proposed to operate on a HOS to offer the execution environment of multiple GOSs. There were lots of T2VMM software presented, for examples, VMware Workstation, VirtualBOX, QEMU, etc. T2VMM, as an application, is implemented on a classical operating system. The advantages of T2VMM are easy to setup, use, and no requirement of special hardware support. But since HOS acts as an additional layer of architecture, the performance overhead of T2VMM is worse than T1VMM. Therefore, T1VMM recently has attracted more and more lots of interest in academic and industrial, such as VMware ESX and Xen [4].

The key component of VM system is Virtual Machine Monitor (VMM). VMM, also called hypervisor, is responsible for the monitoring, isolation and resource allocation for VMs. Thus, the design of VMM directly affects system effectiveness and efficiency of VM system. In this paper, the kernel of VMM, called as VMM scheduler, is investigated. VMM scheduler is responsible to determine the order and duration of processor-sharing for operating VMs. A high-quality VMM scheduler will improve the response time of VMs and to meet the requirements of real-time systems, such as multimedia service systems.

However, in typical VMM scheduler design, the processor time is equally shared by all existed VMs, regardless of the task characteristics of individual VMs. Thus, typical VMM schedules lack of guarantee of real-time services. Actually, in heterogeneous VM system, various VM services demand exclusive service quality. For examples, for a VM of multimedia service, conversational service requires short response time; high throughput is more important to FTP service than less response time.

In this paper, a real-time scheduler to support multiple VMs service is proposed to meet the requirements of various VMs. The details of the proposed scheduler are presented, including scheduler modeling, operations, analysis and performance evaluation. The experimental results demonstrate valuable features of the proposed schedule, including (1) the great improvement of the scheduling-induced delay (SI-Delay) of real-time tasks, (2) fast response time (3) feasible software solution.

The rest of this paper is organized as follows. In Section 2, related works are presented. In Section 3, a real-time scheduler based on Xen VMM is introduced and illustrated about its operations. Experimental results and analysis are given in Section 4. Finally, we conclude this work in Section 5.

2. Related Works

2.1 Previous works on VMM schedulers

Scheduling-induced delay (SI-Delay) [5] induced by the VMM scheduler is the waiting time of GOS for processor. Because the time sharing of various GOSs, GOS can not react on the receipt of interrupts. Thus SI-Delay will slow down the handle time of interrupts and lead to poor I/O performance. In current scheduler design, the processor loading affects SI-Delay, and this make SI-Delay unpredictable. For networking, SI-Delay directly affects the delay time and jitter of network packets, also implicitly affects the bandwidth of TCP flow, and this makes the network servers hard to provide high quality services. This is why current network servers consolidated in the VM system are most Web servers and databases, because they have high tolerance of packets delay and jitter.

There are several methods to reduce SI-Delay. Govindan *et. al* [5] gather statistics of I/O operations to find out the most SI-Delay of GOSs, Accordingly, the scheduler adjusts the execution order of VM in the next scheduling round to improve average delay. Umeno *et. al* [6] create a new state called self-wait for idle GOSs. GOSs in self-idle state have higher priority to handle interrupts. This method is based on a fact that the GOS running I/O-Bound processes consumes little processor resource, and is idle in most time. Ongaro *et. al* [7], also based on this fact, assign high priority to idle GOS when an interrupt is enabled. However, none of methods, so far, has an ability to provide stable I/O performance for GOS.

2.2 Introduction to Xen

This paper chooses Xen as the target platform because Xen provides the minimum performance overhead [4] in all VMMs by using paravirtualization technology. Paravirtualization provides the necessary interfaces to support running GOS instead of virtualizing the whole hardware devices. Even though, SI-Delay still impacts I/O performance seriously.

2.2.1 Architecture of Xen

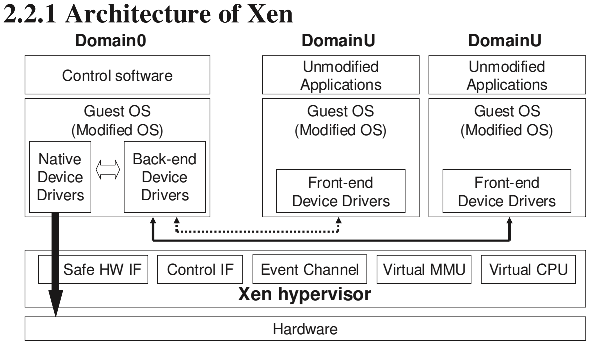


Fig. 2 The architecture of Xen with 2 DomainUs [8]

The architecture of Xen is shown in Fig 2. There are three elements: Xen hypervisor, Domain0, DomainU. Xen hypervisor manages hardware resources directly, and constructs an abstract hardware layer. Based this abstraction, Xen can offer unified platform for various GOSs. The GOS on Xen is formally called as domain. Domain0, a privileged domain, has the ability to control other domains via control interface. Other unprivileged domains are called as DomainU. Domains can communicate with each other by sending an event called as virtual interrupt through event channel. Physical interrupts are also transformed to virtual interrupts before sending to domains.

Xen uses the split driver model to virtualize I/O devices. Domain0 has another privileged ability to drive almost all the devices via native device drivers. Native device drivers are the ones used on existing operating systems. Front-end device drivers are responsible to provide I/O services to DomainU, and communicate with back-end ones through shared memory and event channel. Back-end device drivers are responsible to forward the I/O requests of front-end device drivers to native device drivers and forward the results of I/O operations from native to front-end.

2.2.2 Credit Scheduler

Currently, the default Credit Scheduler employed by Xen uses credits and debit system to distribute processor resource among domains. There are two domain states defined according to the domain credit as illustrated in Fig. 3.

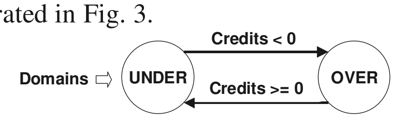


Fig.3 State diagram of domains

Domains in UNDER state have positive quota of processor resource. In contrast, domains in OVER state have run out of the quota and must wait for sufficient quota of processor resource.

Credit Scheduler maintains a run queue for each processor. Domains are inserted to it by their states as illustrated in Fig. 4.

Fig. 4 The run queue

In run queue, domains in UNDER state are prior to the domains in OVER state. The domains in the same state are scheduled in FIFO. In every scheduling round, Credit Scheduler serves the domain in the head of run queue. In each scheduling round, each domain has one slice (default 30ms) to run unless it is preempted. When the domain is scheduled out, it is reinserted to run queue according to its state. Credit Scheduler sets a timer for each processor to debit the credits of the domain running on the processor. This action, called TICK, is invoked every 10ms, and debits 100 credits from the running domain. In every three ticks, Credit Scheduler judges whether redistributing credits to all domains, changing domain states, and resorting run queues. This mechanism ensures the fairness of sharing processor resource between domains.

For improving I/O performance, Credit Scheduler adds a new state: BOOST. Domains in BOOST sate have higher priority over those in the states of UNDER and OVER. When an idle domain receives an interrupt, it will move to BOOST state. Domains running I/O-bound processes should be idle frequently since they require little processor resource. Therefore, the priority mechanism reduces SI-Delay and improves I/O performance.

2.3 The impact of SI-Delay on I/O performance

The impact of SI-Delay on Credit Scheduler is serious. Consider the case of 9 domains operating on the testing platform, denoting as D_0-D_8. D_0 and D_8 stay in idle. D_7 runs a TCP bandwidth benchmark. D_1-D_6 may run infinite loop, named as BURN, or stay idle. We test the ping latency of D_8 and the bandwidth of D_7 to estimate the response time and throughput.

In this case, the more domains compete for processor resource, the worse response time and throughput will be. The increasing of ping latency variance implies that I/O performance becomes much unstable. The details of this case are shown in Table 1.

Table 1 Relationship between the impact of SI-Delay and the processor loading

Types of D_1-D_6	Avg. Ping Latency (ms)	MAX Ping Latency (ms)	VAR of Ping Latency	Bandwidth (Mb/s)
IDLE*6	9.503	11.200	2.810	95.580
BURN*1 IDLE*5	10.551	39.300	12.056	92.050
BURN*2 IDLE*4	12.265	96.200	98.977	86.210
BURN*3 IDLE*3	46.869	414.000	5265.806	71.160
BURN*4 IDLE*2	100.489	679.000	18719.834	51.180
BURN*5 IDLE*1	218.042	1242.000	78243.644	30.990
BURN*6	510.785	2159.000	229389.029	11.170

2.4 Problem formulation

The SI-Delay arises from that processor resource is not allocated to domains in the right time. Thus, the I/O operations suffer from unexpected SI-Delay. However, if we can allocate processor resource to domains when they receive an interrupt, the SI-Delay can be eliminated. Besides, the allocated processor time is also important. The insufficient allocated processor time leads to the pending of some interrupt handlings. Thus, the pending time certainly degrade SI-Delay. Consequently, the proper allocation of processor resource in the right time is the challenging issue to investigate in this paper.

3. An real-time scheduler based on Xen VMM

3.1 Task category

In this work, there are a variety of network servers considered to deploy on VM systems, such as Conversational server, Game server, Web server and so on, to collaborate. To satisfy the SI-Delay demands of VMs, we classify VMs according to their SI-Delay requirements, including Minimum, Fixed-Scope, Upper-Bound and Best-Effort as depicted in Table 2.

Table 2 Requirements of SI-Delay

SI-Delay Requirements	Example	Tolerance of SI-Delay	
		Delay	Jitter
Minimum	Conversational Server	Very Low	Very Low
Fixed-Scope	Timer	Don't care	Low
Upper-Bound	Game Server	Medium	Medium
	Web Server	High	High
Best-Effort	Mail/FTP server	Very High	Very High

Minimum class is applicable for the services such as Conversational server to have minimum response time. Fixed-Scope class is specially designed for the virtual timer in VMM system to keep the timer from SI-Delay and improve its accuracy. In Fixed-Scope class, SI-Delay of domains is restricted to be a constant. The restricted SI-Delay contributes the elimination of timer error.

Upper-Bound class is designed for the services which are both CPU-Bound and I/O-bound such as Game server and Web server. Upper-Bound class restricts the maximum SI-Delay of domains. When the experienced SI-Delay exceeds the bound time, this domain will be redirected to Minimum class in this slice. Best-effort class is for other services, regardless of SI-Delay, like Mail server and FTP server.

These classes can fulfill different requirements for heterogeneous multimedia servers, and provide acceptable and stable I/O performance for them.

3.2 Architecture of real-time scheduler

As illustrated in Fig. 5, there are two phases in real-time scheduler. In the first phase, called classification phase, domains are classified and inserted into different queues accordingly. In the second phase, i.e. scheduling phase, the scheduler determines the execution order of in-queue domains. In run queue, Credit mechanism is applied, and FIFO is applied to the other queues. In every scheduling round, the scheduler searches queues in the order of Minimum, Fixed-Scope, and run queue. The domains in former queue are always prior to those in latter queue.

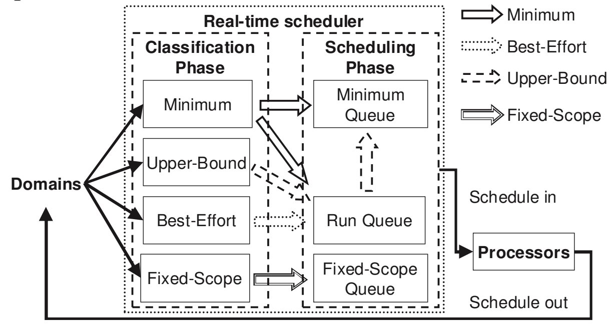


Fig. 5 Architecture of real-time scheduler

3.3 Modified Event Channel

In legacy scheduler of Xen, the events, including I/O, only can be sensed when the domain is idle. This causes that busy domains lacks of significant I/O notification to complete real-time I/O operations. Therefore, we modify event channel to notify the scheduler on event transmissions. But, the scheduler does not react to all notifies. Most of them are just discarded, so the performance overhead of this modification is small enough to neglect.

3.4 State diagram and algorithm

3.4.1 Minimum class design

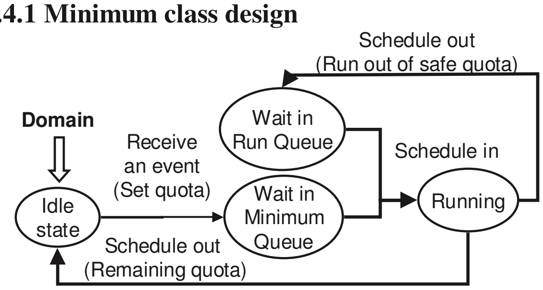


Fig. 6 State diagram of Minimum class design

The concept of Minimum class design is also based on that the domain running I/O-Bound processes consumes small portion of processor resource, and is idle in most time. But we give domain a permitting time to handle interrupts. The domain running on processor without exceeding the permitting time is still regarded as idle. We use a safe quota to count the permitting time. The quota can be adapted to different environments, but should be large enough to reach maximum I/O performance and avoid SI-Delay.

The overall algorithm is as follows: When an idle domain receives an event, it is inserted to the Minimum queue with a safe quota. When the domain is scheduled in processor, the quota is decreased periodically. If the domain is scheduled out with no remaining quota, it is treated as running CPU-bound processes. Therefore, the domain is reinserted to run queue, and scheduled by a Credit mechanism to avoid consuming exceeding quota of processor resource. Otherwise, the domain can continue to have the high scheduling priority.

3.4.2 Fixed-Scope class design

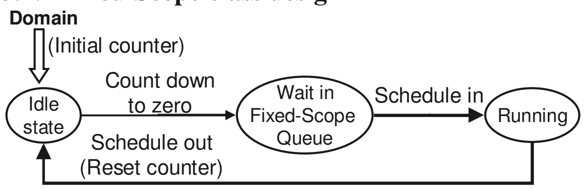


Fig. 7 State diagram of Fixed-Scope class design

To constrain the SI-Delay to Fixed-Scope, we schedule domains in processor periodically. The Fixed-Scope domain is put to idle state with a counter in default. The scheduler periodically counts down the counters. When a domain counter reaches zero, the domain is inserted to Fixed-Scope queue and wait to be scheduled in processor. Since the priority of Fixed-Scope queue is lower than Minimum queue, the Fixed-Scope domain has no interference with Minimum domains.

3.4.3 Upper-Bound class design

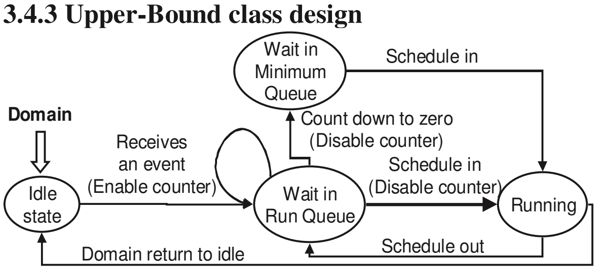


Fig. 8 State diagram of Upper-Bound design

The domains in Upper-Bound class are moved from run queue to Minimum queue when the domain has waited for a bound time (BT). In implementation, we also use a counter as BT counter. The counter is disabled in default, and is enabled when domain receives an event. The scheduler counts down the enabled counters periodically. When the domain is either scheduled in processor or promoted to Minimum queue, the counter is set back to disabled. This design can limit the maximum value of SI-Delay to provide acceptable I/O performance for Upper-Bound domains.

4. Experimental results and analysis

4.1 Experimental environment

The testing platform is composed of an Intel Core 2 Duo E6420 system with 2GB RAM and Xen-3.1. Domain0 runs Ubuntu 7.1 with 512MB RAM. DomainUs run Ubuntu dapper, and each has 2 VCPUs and 128MB RAM. All DomainUs share the read-only filesystem and the Linux 2.6.18 kernel. We also use an Intel Pentium4 system running native Linux as testing endpoint. Testing endpoint connects to testing platform through a 100Mbps local network. The endpoint has three missions; (1) to generate I/O requests for the testing platform (2) to be the control console of test platform (3) to collect the measurements from the testing platform.

4.2 Test case

To unify test scenarios, there are eight Domains, including D_1-D_8 in various test scenarios. And we design several types of domains, as follows:

I. BURN: Running two infinite loops to keep domain busy. This emulates a domain running CPU-Bound processes.
II. STREAM: Running a networking benchmark to transmitting TCP flow to testing endpoint. This emulates a domain running bandwidth-intensive processes.
III. IDLE: Do nothing. To keep the number of domains.
IV. PING: Stay idle, only response to ping packets. This emulates a domain running latency-sensitive processes.
V. PERIOD: Running a process which sending a UDP packet to testing endpoint every 100ms. This type is designed to test the timer accuracy.
VI. PING+BURN: A BURN domain responses to ping packets. This emulates a domain running both CPU-Bound and I/O-Bound processes.

4.3 Numeric results

4.3.1 Performance evaluation of Minimum class

We use the following settings to test Minimum class design. D_1-D_6 are composed of BURN domains and IDLE domains, D_7 is STREAM domain, and D_8 is PING domain. D_0, D_7 and D_8 are set to Minimum class to reduce SI-Delay on forwarding I/O. D_1-D_6 are set to Best-Effort class. We verify the performance of the scheduler by inspecting the ping latency of D_8 and the bandwidth of D_7 to estimate the response time and throughput.

As depicted in Table 3, the proposed scheduler can effectively eliminate SI-Delay to achieve high and stable I/O performance. As a result, the response time and throughput are independent of processor loading.

Table 3 Ping latency and throughput

Types of D_1-D_6	Avg. Ping Latency (ms)	MAX Ping Latency (ms)	VAR of Ping Latency	Bandwidth (Mb/s)
IDLE*6	9.484	11.200	3.640	93.420
BURN*1 IDLE*5	9.505	11.200	3.627	93.450
BURN*2 IDLE*4	10.186	11.200	3.305	93.460
BURN*3 IDLE*3	9.734	11.200	3.503	93.440
BURN*4 IDLE*2	9.394	11.200	3.328	93.430
BURN*5 IDLE*1	9.811	11.200	3.845	93.580
BURN*6	9.578	11.200	3.124	93.470

4.3.2 Performance evaluation of Fixed-Scope class

In this test, D_0 is set to Minimum class; D_1-D_7 are BURN domains and D_8 is Period domain of Fixed-Scope class. D_8 issues UDP packets periodically to the endpoint. We record the time interval between the receptions of UDP packets at endpoints. The interval time is used to measure the timer accuracy. The jitter of packets delay is neglectable. In default, the counter is set to 100ms.

As depicted in Table 4, due to SI-Delay, the original Credit Scheduler can not guarantee the constant time interval of successive packets. Thus, the average time interval is larger. However, through Fixed-Scope class design, the virtual timer error is only 10ms, and the average time interval matches 100ms. That is, such a design ensures the timer accuracy in VMM system.

Table 4 Time intervals between receiving UDP packets

	Avg.	MAX	MIN	VAR
Credit	155.55	300	10	6656.982
Fixed-Scope	100.000	110	90	35.176

4.3.3 Performance evaluation of Upper-Bound class

To test the effectiveness of Upper-bound class, D_1-D_6 are specified as BURN domains, D_7 is STREAM domain, and D_8 is PING+BURN domain. Among them, D_8 is set to Upper-Bound class, and D_0 is set to Minimum class.

As depicted in Table 5, the Upper-Bound class provides better I/O performance by restricting the maximum value of SI-Delay. However the CPU utilization is related to BT. This is because that the slice of the domain running on processor is fixed, and the promoting action gives the domain more processor time potentially. We can turn the BT to adjust it. This is a trade-off between I/O performance and fairness of sharing processor resource. As shown in experiment results, this design indeed improves I/O performance even in the case of high processor loading.

Table 5 Ping latency and CPU usage

	Credit	Upper-bound (BT = 30ms)	Upper-bound (BT = 50ms)
Avg. (ms)	196.160	19.280	30.649
MAX (ms)	646.000	40.900	60.700
MIN (ms)	0.129	0.113	0.209
VAR	18730.02	133.767	335.513
Total CPU usage %	26.9	32.4	26.6

5. Conclusion

In this paper, a real-time domain scheduler is proposed to support VMM. The proposed scheme efficiently distributes processor resource according to the defined domain classes, including Minimum, Fixed Scope, Upper-Bound and Best-Effort. The defined classes satisfy the demands from a variety of VM applications by offering the restriction of SI-Delay. To verify the proposed scheme, an open source of VMM, Xen, is employed. In the experiments, we show that real-time scheduler can indeed provide high and stable I/O performance for GOS. This enhancement will conduct new applications to consolidate heterogeneous multimedia servers into the VM system with quality of services.

6. References

[1] R. P. Goldberg, “Architecture of virtual machines”, Proc. of AFIPS National Computer Conf., July, 1973.

[2] A. Singh., “An Introduction to Virtualization”, http://www.kernelthread.com/publications/virtualization

[3] J. E. Smith, R. Nair, “The Architecture of Virtual Machines”, Computer, IEEE Computer Society, May, 2005.

[4] P. Barham et al., “Xen and the Art of Virtualization”, Proc. of the 19th ACM symposium on Operating systems principles, Oct., 2003.

[5] S. Govindan et al., “Xen and co.: communication-aware CPU scheduling for consolidated xen-based hosting platforms”, Proc. of the 3rd Int. Conf. on Virtual execution environments, June, 2007

[6] H. Umeno et al., “New Method for Dispatching Waiting Logical Processors in Virtual Machine System”, Proc. of the 29th Annual Int. Computer Software and Applications Conf., July, 2005.

[7] D. Ongaro et al., “Scheduling I/O in Virtual Machine Monitors”, Proc. of the 4th ACM SIGPLAN/SIGOPS international conference on Virtual execution environments, March, 2008.

[8] I. Pratt, “Xen Virtualization”, Linux World 2005 Virtualization BOF Presentation.

ACKNOWLEDGEMENT

The authors would like to thank Institute of Information Industry(III) and National Science Council in Taiwan R.O.C for partially financial supporting under the grant NSC95-2221-E-006-289-MY2.

An Architecture for an Operational Picture System for Crisis Management

Erland Jungert and Niklas Hallberg
FOI (Swedish Defence Research Agency)
Box 1165, S-581 11 Linköping, Sweden
{Jungert, nikha}@foi.se

Abstract: ***Crisis management is to a large extent a collaborative activity between different organizations. It is also a complex activity where conditions within the environment, upcoming conditions and incidents must be considered and acted upon, using available and appropriate resources. Efficient crisis management depends on a high degree of situation awareness. Hence, to provide and present adequate information are high priority requirements on crisis management systems. This paper presents a system for rich and user adopted operational pictures, which should be regarded as a component possible to integrate in crisis management systems. The main capability of the proposed system includes an extensive set of views where a view corresponds to a representation of a portion of the operational picture. The view concept makes it possible to dynamically adapt the operational picture to meet the needs of the users. To achieve a high level of adaptability and interoperability the system is based on a service-oriented architecture. The role concept is used to enhance the ability to adapt to the actors and the current situation. Each role has a set of views connected to it. The system is designed to deliver consistent, current and timely situation descriptions.***

1. INTRODUCTION

Crisis management requires the ability to collaboratively act upon upcoming situations, which could be difficult to predict. Thereby, different combinations of crisis management actors and their systems must rapidly be put together to act jointly in case of emergencies [12]. Therefore, crisis management systems must fulfill the requirements of flexibility, adaptability, scalability and interoperability. Systems that meet these requirements ought to be based on a suitable architecture that beneficially is service-oriented [13], [11] and [9]. Further, crisis management systems require suitable user interfaces adjustable to the users´ needs, their activities and the situation at hand, i.e., adaptation to the different roles performed by each actor [14]. This is of particular importance at operational level in local and regional communities where heterogeneous groups of actors are directly engaged to solve the crisis. At a general level these categories of actors include (1) the fire brigades, (2) the police departments, (3) the healthcare providers, (4) emergency call centers, (5) the municipal councils, and (6) the county administrative board. These categories have their dedicated duties and missions to carry out. Thereby, they are assigned to perform their roles within the crisis management process based on the current situation, for which they need valid support. In order to enhance the actors' capabilities to manage activities, in a variety of different situations, decision-support tools must be provided and suited to fit the special needs of each role(s) of the actors. Examples of such tools are support for obtaining situational awareness, for planning, and for coordination of teams to jointly carry out their activities. In order to create a solid ground for decision making commanders must consequently be supplied with information that enhances their understanding of the situation, i.e., to support their awareness of the situation and the on-going processes. Hence, awareness is a necessity for effectual use of crisis management resources to minimize the loss of material, human suffering and lives. Thereby, adequate crisis management requires continuous collection of data that reflect the situation, including the ongoing crisis and the actual status of available resources. The collected data must be analyzed, stored, processed, and visualized to fit the needs of the users. In order to support the users to obtain situational awareness, crisis management systems must thus provide operational pictures, e.g. [1]. Further, from the users' perspective, the reliability in the information is of utmost importance. A difficulty related to this is that the collected data contains different types and degrees of uncertainty that must be considered when performing the analysis and the use of the results for the decision-making.

The operational picture used as a basis for crisis management must be possible to share among actors to make it possible for them to operate together. However, it is not necessary for all involved actors to share the same operational information. On the contrary, it is only the information that is relevant to have in common that should be shared and, thereby, different actors should just share distinctive sets of information [12]. However, what really is relevant information must be determined by authorized decision makers. Hence, depending on the differences in activities and responsibilities assigned to the different actors, the operational picture system must be adjustable with respect to the information and functions needed for its operations. Information presented by the operational picture system includes images, maps, tables, drawings, and textual comments. This requires the operational pictures' ability to present a large number of information structures. Hence, the concept can be seen as an approach for presentation of images in a broader sense and for different purposes. An operational picture system can be seen as a more extensive type of decision support system with a central position in a command and control architecture.

In this work, a *view oriented operational picture system* (VOOPS) is proposed as a means to enhance heterogeneous user categories' usage of common operational pictures to increase their situational awareness. The main focus is an architecture of a system for handling operational information by means of a service-based structure that can be adapted to the roles of the different actors and their work situation. The operational information will be presented to them by means of the views in a suitable way. VOOPS thus supplies the users with sets of views, adapted to the particular user categories by means of available services directly associated with the views. Each view includes and presents a dedicated set of information. A view is a representation of a situation from the perspective of a related *set of concerns* [8]. Each instance of a view is produced by the use of one or several services. As in Service Oriented Architectures (SOA), services are seen as a mechanism by which consumers' needs and producers' capabilities are brought together [10]. There the roles of the service providers and the service consumers can be played by humans or machines. Services could be *elementary services* produced by a single unit or *composite services* conducted by other services [2]. Hence, a view is a service that commonly is produced by several other, often subordinated, services. Further, the various views are part of a higher order service called a *section*. Each such section is specialized with respect to the functionality it provides. However, the structure of the sections is module oriented. This makes VOOPS, as a whole, adjustable to different applications and user roles, allowing different sets and numbers of sections to be integrated depending on the actual needs, i.e. the user requirements based on their activities, workload, and missions and their roles. Hence, the approach taken can be seen as a system for handling operational pictures, which constitutes a central part of a C^2 system.

2. THE OPERATIONAL PICTURES CONCEPT

To provide actors with means for improved awareness, systems with the ability to create a description of the situation, i.e., an operational picture, are needed. In the proposed system, VOOPS, each user has his or her own share of situational information, called an *individual operation picture* (IOP). Further, modern crisis management systems must provide the ability for the users to share information with other users. The information shared between groups of users corresponds to a *common operational picture* (COP). This information could be shared between two or several users. The complete set of situational information is called the *entire operational picture* (EOP), which is the sum of all IOPs. The relations between these three types of operational pictures are illustrated in Figure 1.

3. THE VIEW CONCEPT

Views should be seen as a concept to customize operational information, within heterogeneous groups of users, to their different needs of knowledge. Hence, based on the views, users could construct their own IOP based on the EOP. A view instance is the information that a view represents at a given moment. Each view instance is related to a set of concerns, which define a portion of the data, e.g., found in databases, obtained from sensors, and/or given in intelligence reports. Consequently, the system includes a set of view instances where each instance includes application dependent information. To handle these view instances, mechanisms to support the users' interactions with the views are required, i.e., services. Thus, the users should be able to manipulate the views depending on their information needs, which as a consequence causes the creation of new and from a time perspective subsequent instances. Therefore, VOOPS provide functionality that can be used for creating a view by filtering information and/or combining view instances. Further, various functions must be provided to carry out information searches and selections either directly in a view, from other available data sources or from existing repositories by means of available services.

A view corresponds to a service with the capability to view a part of the entire operational picture and to manipulate it. Through a view it is possible to actively call a set of lower order services that can carry out actions similar to commands needed to manipulate the views. From a system perspective the approach can be seen as a hierarchical service structure where manipulation and presentation of the view instances is the lowest level. On top level reside the section level services.

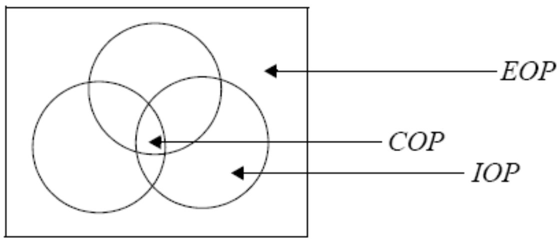


Figure 1. The relationships between the entire operational picture (EOP), the individual operational picture (IOP) and the common operational picture (COP).

4. SITUATIONAL DESCRIPTION VIEWS

Systems for generation of operational information, i.e. the system for generation of sets of view instances, must have capabilities for data analysis, data fusion, and for handling heterogeneous data, which may contain uncertainties. The systems must support presentation of coordinated information that often are completed with context information like maps, plans and organizational information. The information may come from external sources, which means that support for collection of such information must be available as well as methods for their analysis. Furthermore, information about allocated resources, specifically in the context of planning and plan verification is needed.

To support the operative control, the VOOPS architecture provides five different views in the operative control section (Figure 2). The five views are hierarchically organized in two levels with the *Current operational view* (COV) at the top level. The remaining four views are located on a secondary level and attached to the COV. The purpose of the COV is to present the current situation to users with specific roles, i.e. to present the information matching the current situation relevant to the users' assigned roles. The *Allocated resources view* (ARV) includes the resources allocated to the on-going missions. The *imported view* (IMV) shows view instances obtained from other actors. As these view instances may be shared they are part of the COP. The *A priori dataview* (APV) includes general database information like names of personnel and organizations that need to be engaged but also information concerning recommendations about what to do under certain conditions. The final view in the operative section is the *sensor data view* (SDV), which can be used to visualize sensor data. This view is of particular interest as it can be used to view sensor data directly. This is useful when data are difficult to analyze and interpret automatically, which can be due to uncertainties in the data.

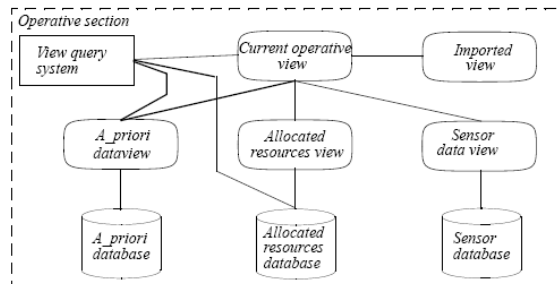


Figure 2. The operative control section and its six views.

The services associated with each view in VOOPS are specialized with respect to the view they are associated with, although the same services may be available to more than one view. From a user perspective it may not be enough to manipulate data from a single view. For this reason, it must be possible to request data from several views and to combine these data to create new kind of view instances. To allow this cross view information manipulation, a view query language (VQL) is provided that will be integrated in the operative section. The operative section has a central position in the VOOPS architecture for all user roles, and will be further discussed in section 7.

5. USER ROLES

Different actors have different responsibilities and tasks to carry out throughout the crisis management. In other words they are assigned different user roles that subsequently are called *roles* for short. Further, different tasks and responsibilities, and thereby different roles, require different information. For example, the fire department will need information about the construction of a building and the location of the water posts while the police does not. Thereby, crisis management systems must be sufficiently flexible and adaptable to the various roles that the actors are assigned momentarily. In this case, the realizations of the operational pictures must be adapted to the needs of the different roles. A role is defined generally with respect to its tasks and responsibilities, but it also depends on its organizational belonging. The use of roles and views in the definition of the system architecture makes the development of systems both user and usage oriented, i.e., the focus on the systems capabilities is to meet the different roles' momentary information needs. This approach also makes it possible to allow for a modularity that makes it possible to determine the set of views for each role including the services associated to each view. The services must be suited for the tasks to be performed by the users given their special roles. Roles of interest in local society crisis management include members from the following organizations:

- Medical care providers (MCP)
- Police department
  - Onsite command (OC)
  - Regional command (RC)
- Fire department
  - Onsite command (OC)
  - Local command (LC)
- Information coordinators
- Analysts
- Municipality board (MB)
- County administration board (CAB)

Each role is associated with the building blocks in VOOPS, i.e., the views include the services needed to collect the information necessary to carry out the assigned tasks. Through the adaptation of the system to the different roles it will be possible for the users to create their individual operational pictures. That is, the users' interact with the entire operational picture by means of the services associated with the views and dedicated to their roles. Further, the common operational pictures are maintained by the system momentarily. Thus, each view instance must indicate whether it is a member of the common operational picture or a member of the individual operation picture.

Table 1: Role/views attachments

Role	COV	CXV	IMV	HYV	ARV	PLV	APV	SAV	SDV
MCP	X		X		X	X	X	X	
PD, OC	X	X	X		X		X	X	
PD, RC	X	X	X		X	X	X	X	X
FD, OC	X	X	X		X		X	X	
FD, LC	X	X	X		X	X	X	X	X
SOS-Alarm	X	X	X		X	X	X	X	
Information co-ordinator	X	X	X				X	X	X
Analyst	X	X	X	X	X		X	X	X
MB	X	X	X				X	X	
CAB	X	X	X				X	X	

Further, it should be observed that, for instance, the Municipality Board and the County Administrative Board do not have any operative responsibilities; instead they coordinate the activities on a strategic level.

Each role must be assigned a set of views, which also may vary over time as the needs of the roles change. Currently nine different views have been identified. Of these nine views, five are part of the operative section as discussed in section 3. The four additional views are the *context view* (CXV), the *History view* (HYV), the *Planning view* (PLV) and the *Service availability view* (SAV). The nine views are, depending on the requirements of the roles, assigned the roles identified for the local community crisis management (Table 1).

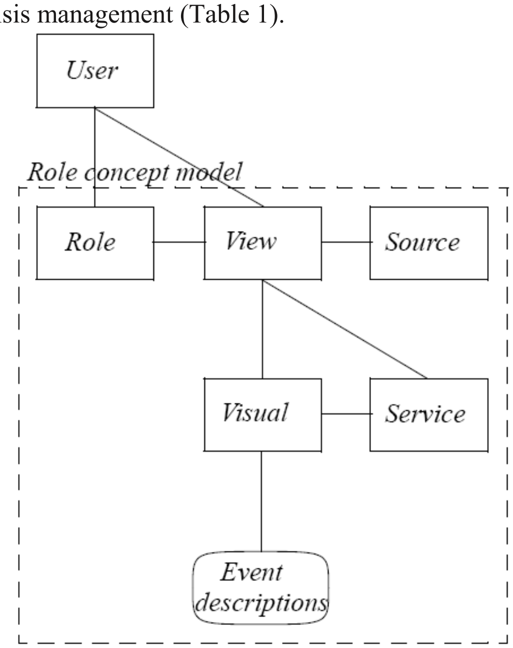


Figure 3. The mission support model and its agent relations.

6. THE ROLE CONCEPT MODEL

The VOOPS architecture is based on a structure, called the role concept model, which also can be seen as an ontology. The basic concepts of the model are: (1) views, (2) services, (3) roles, and (4) their relationships. Primarily the model is developed to provide for the mission support in the command and control process (Figure 3). The model illustrates how actors (users) relate to their role in the crisis management process; the relations between the views and the actors can also be seen. Furthermore, on the next lower level the relationships between the views and the services and the visuals are illustrated. A *visual* corresponds to a visualization of a view instance. Most relations in the model are of many-to-many type. For instance, an actor may use several views and a view may be used by several actors. Thus, a single view may be used by more than one role, and a certain role can be associated with a set of views including their visuals and services. Users may not have any direct link to the data sources. If this is the case, such information must be linked from a data source via a view by the users independently of the role. This is due to the need for verification of all kinds of data content, e.g., it may under certain weather conditions be difficult to determine whether an observed object is human or not.

The role concept model demonstrates not just the relations between roles, views and services but also their relations with the data sources, which is an important feature of the VOOPS architecture. Furthermore, for each role the generation of the individual operational picture is supported by means of the services through the set of related views, via the visuals and down to the event descriptions, where the corresponding operational information resides. This structure of the model also demonstrates how the situational awareness, required in command and control processes, is improved by means of the information in the lowest level of the model that contains the event descriptions of the on-going crisis. The event descriptions contain from the perspective of the actors collected descriptions of the events in the crisis. A similar concept model is also used in [6].

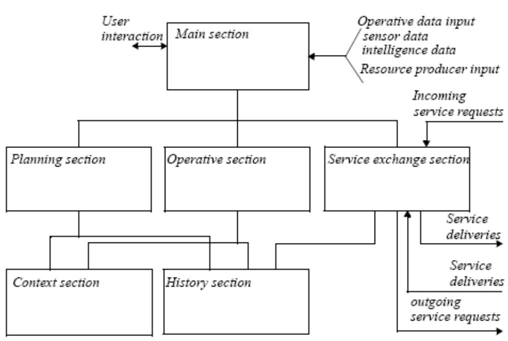


Figure 4. The section structure level of the system architecture of

7. THE VOOPS ARCHITECTURE

The architecture of the view-oriented operational picture system (VOOPS) is on the highest system level constituted by a set of sections (Figure 4). Each section is assigned a number of tasks supported by one or several views. Since this level of the architecture is application dependent not all of the sections are mandatory for all systems, i.e., dependent on what the system is supposed to support different sections have different levels of relevance. The section structure is basically hierarchical with three levels. Further, the service exchange section is available to external users as its intention is to allow requests for and distribution of view instances. The service exchange section permits the users to request services from other users. Each section has its own functionality, set of views and services.

The main section differs in various aspects from the rest of the sections. It does not include any views and is a control and interaction module that includes three major services. These are (1) *the operative section service*, (2) *the allocation planner service* and (3) *the service exchange service*. The operative section service supports the users in maintaining their individual operational picture. The support for this is carried out by means of the views available in the operative section. The allocation planner service, which is not needed for all roles, supports the allocation of the required resources necessary to perform the planning activities. The purpose of the last major service, the service exchange service, is to support requests for information from other actors via external services as well as to be supported by other actors through transmitted requests. The main section also includes the application database for received information, which may include view instances, sensor data, human observations, and intelligence reports. The main section can be seen as a command and control environment where the purpose is to control the general activities and to carry out the crisis management process.

The planning section is the main part of the allocation planning system, directly connected to the allocation planner in the main section. *The context section* includes the context information from which background information related to various view instances are selected, transferred and integrated with the current operational information to form a view instance or to be integrated with any view instance that requires background information. The context section includes information such as maps, satellite images, building blueprints, and locations of dangerous substances. *The history section* is a repository for the historic view instances of all views. Hence, when a view is updated a new view instances replaces the old one and the old view is transferred and stored in the history section. Whenever necessary, the historic view instances can be reinstalled for, e.g., analysis of incidents and review of taken actions. This is particularly relevant to analyze because of what has happened during a certain time frame; that is, when for instance several view instances for the same views are scrutinized

Commonly, during crisis situations large amounts of data are obtained, provided by different types of sensors and observers. These data are associated with uncertainties of various kinds. To handle these uncertainties, methods for sensor data fusion are required [7]. Further, information fusion methods are also needed, especially to support services in the operative section but also to perform risk analyses. The query system used in the main section works on streams of incoming sensor data and includes mechanisms for sensor data fusion. It is also designed to work in a sensor data independent way where the sensor data sources deliver heterogeneous data that must be possible to transform into a unified structure. The query language is described e.g. in [4], [5] and from a scenario driven perspective by Camara and Jungert [3].

8. CONCLUSIONS

This paper presents an operational picture system, and its architecture, that could be integrated as a component in a command and control system for crisis management. The aim of the operational picture system is to enhance the crisis management actors' situational awareness. However, this is not trivially developed. Information about the same issue should not be inconsistent. Hence, there is a need, to produce application adapted information to each type of users. To fulfill this requirement an approach for presentation of user adapted and consistent operational information has been proposed. The approach is based on a service oriented structure. Among these services, the ones that constitute views are the most fundamental as they are designed to increase the users' situational awareness with regards to the ongoing crisis.

The view instances correspond to an event description at a give time. Hence, a role's set of view instances corresponds to fragments of the entire operational picture (EOP) at a given period in time. Further, the proposed system also provides functionality for planning of activities and for allocation of required resources.

The view and service concepts are not just assets for presentation of operational information they are also a structure for handling and for manipulation of operational information to enhance the actors' situational awareness. The system can be extended to fit other user categories and needs which make the system a useful tool in other contexts as well.

ACKNOWLEDGEMENTS

This work has become possible due to grants from KBM (Swedish Emergency Management Agency).

REFERENCES

1. Artman, A. (1999) Situation awareness and co-operation within and between hierarchical units in dynamic decision making, *Journal of Ergonomics*, 42, 11, pp 1404-1417.
2. Benatallah, B., Sheng, Q. Z.., Dumas, M. (2003) The Self-Serv environment for Web services composition. *Internet Computing, IEEE*, 7, 1, pp 40-48.
3. Camara, K., Jungert, E. (2007), A visual query language for dynamic processes applied to a scenario driven environment, *Journal of Visual Languages and Computing*, Vol 18, No 3, 2007, pp 315-338.
4. Chang, S.-K., Jungert, E., X. Li, X. (2006), A Progressive Query Language and Interactive Reasoner for Information Fusion Support, *Journal of information Fusion*, Vol. 8, No 1, pp 70-83.
5. Chang, S.-K., Costagliola, G., Jungert, E., Orciuoli, F. (2004), Querying Distributed Multimedia Databases and Data Sources in Information Fusion Applications, *IEEE Trans. on Multimedia*, Vol. 6, No. 5, October 2004, pp 687-702.
6. Chang, S.-K., Jungert, E. (2007) A Self-Organizing Approach to Mission Initialization and Control in Emergency Management, *Proceedings of the International Conference on Distributed Multi-media Systems*, San Francisco, CA, September 6-8, 2007.
7. Hall, D. L., Llinas, J. (Eds.) (2001) Handbook of Multi-sensor Data Fusion, *CRC Press*, New York, 2001.
8. IEEE 1471 (2000) *IEEE Recommended Practice for Architectural Description of Software-Intensive Systems*, Std 1471-2000, New York.
9. Jungert, E., Hallberg, N., Hunstad, A. (2006) A Service-based Command and Control Systems Architecture for Crisis Management, *The International Journal of Emergency Management*, Vol. 3, No. 2, 2006, pp 131-148.
10. MacKenzie, C. M., Laskey, K., McCabe, F., Brown, P. F., Metz, R., Hamilton, B. A (eds.) (2006) Reference Model for Service Oriented Architecture 1.0 Committee Specification 1, 2 August 2006 http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=soa-rm (2008-01-24)
11. McGovern, J., Ambler, S. W., Stevens, M. E., Linn, J., Jo, E. K., Sharan, V. (2003) The Practical Guide to Enterprise Architecture, *Prentice Hall*, PTR, New Jersey.
12. Mendonca, D., Jefferson, T., Harrald, J. (2007) Collaborative adhocracies and mix-and-match technologies in emergency management, *Communications of the ACM*, 50, 3, pp 44-49.
13. Rechtin, E. (1999) System architecting of organizations: Why Eagles can't swim, *CRC Press*, Boca Raton, FL.
14. Wybo, J. L., Kowalski, K. M. (1998) Command centers and emergency management support. Safety Science, 30, pp 131-138.

Using Low-Fidelity Simulations to Support Design of Decision-Support Systems for Command and Control Applications

Jiri Trnka, Helena Granlund, and Rego Granlund
Division of Human-Centered Systems, Department of Computer and Information Science
Linköping University
Linköping, Sweden
{jirtr, helgr, reggr}@ida.liu.se

***Abstract*—Modern decision-support systems for command and control may lead to radical organizational and technological changes. This paper describes a methodology to support designers and developers of such system in order to identify this type of changes in the design process. The methodology is based on a shorter series of low-fidelity simulations. The analyses do not focus on individual actions of the decision-support system operators, but instead on impacts of the systems on situated and joint actions of the command and control teams in terms of communication and coordination. The simulation series presented in the paper shows how the use of a decision-support system combining real-time data from various sensors led to changes in the content and structure of the communication as well as the resource coordination. Findings from the discussed simulations suggest that more attention should be given to possible impacts of decision-support systems on the joint actions of the command and control teams, as well as how these teams' actions are situated within the larger command and control system**.

I. Introduction

Many organizations working with emergency management invest in various information and communication systems, hoping to increase performance and control capabilities in their everyday work, especially in critical situations. Systems that are particularly relevant in this context are decision-support systems (DSS) for command and control applications. Such DSS integrate various technologies, such as multimodal communications, geographical information systems, real-time sensors and advanced (non-conventional) databases. DSS for command and control are designed to support commanders in tasks such as real-time assessment of responded events, real-time access to essential information about the environment in which commanders are working, and, not in the least, information exchange between commanders. An important feature of DSS for command and control is that they often support multiple commanders working collaboratively.

Authors working in various fields, such as computer supported cooperative work [1], distributed cognition [2] and cognitive systems engineering [3] all emphasize the importance of scrutinizing the usefulness and the actual effect of new systems when applied in practice in addition to assessment of potential effects of new technologies from a theoretical understanding. It is thus essential that the design process of DSS for command and control takes into account not only how a command and control system should behave, but also what the actual command and control system does and how it performs when using new DSS [3, 42]. A DSS for command and control needs to be empirically evaluated in an appropriate context as an integral part of the design process [43]. In this sense, context represents emergency response related circumstances, facts and conditions, which influence the actions and behaviour of commanders responding to an emergency [3].

To study events and human experience in situations where geographically distributed commanders use various technological systems under heavy workload is however difficult. Real emergency response operations are thus rarely reviewed in sufficient detail to gain insight for the design purposes. Simulations are often the only way to confront and analyze these situations and systems [5]. The simulations that are relevant in this context are 'human-in-the-loop' simulations, which are scenario-based and executed in real-time. These simulations are interactive multi-person settings, where the participants face tasks conducted in actual time, and where the development of the tasks can be described as dynamic [6, 7].

The aim of this paper is to present how low-fidelity simulations can be used to analyze the impact of distributed multimedia DSS on command and control teams in terms of effects on coordination and communication in order to support the design process.

II. Theoretical Framework

This research focuses on design of novel DSS to support collaborative work and processes of command and control teams. A team is a number of people (team members) engaged in a set of goal oriented activities, which are carried out in a collaborative manner [8]. Team members often have explicit roles and tasks; they also have access to different information and often use special tools [9, 40]. The team members' actions are also interrelated and interdependent [8]. In the context of this research, we look at a specific type of teams: distributed teams in control situations, such as military, emergency management and transportation. This type of teams in these circumstances has been termed ‘command and control teams’ [10, 37].

An important characteristic of command and control teams is that they do not act in isolation but are a part of a larger system, a so called ‘command and control system’. A command and control system is in principle a distributed supervisory control system with the goal to utilize coordination of resources, which the system is in control of [11]. Command and control systems are characterized by highly complex and dynamic span of interactions, and medium to low coupling between their parts [12]. In order to perform effectively and achieve their goals, parts of the command and control system may need to operate on different time scales and maintain a certain amount of freedom of action to be adaptive [14, 15]. The actions of the command and control system are shaped by different constraints, such as circumstances, system’s capacity, and used technological systems. These constraints not only limit the range of possible actions but also provide opportunities for certain actions to take place [20]. The process of command and control can thus be seen – from the perspective of a command and control team – as the enduring management of internal and external constraints in order to achieve external goals (defined by the command and control system) as well as internal goals (defined by the command and control team) [13]. The team’s ability to manage these constraints and coordinate its work within these constraints, determines if the team is able to accomplish its tasks and with what outcome [13, 20].

If team members are to function as a team that strives towards its goals, they need to coordinate their actions in accordance with these – external and internal – goals [8, 17]. How the team’s actions are organized and coordinated underlines the functioning and performance of the team [10]. Effective internal and external coordination is carried out by means of communication [8, 47]. Different configurations of teams, for example, collocated vs. distributed and hierarchical vs. networked, have an impact on communication and can lead to different outcomes of teamwork [49, 19]. The nature of communication in such teams may differ in a number of ways, such as quantity and content of the exchanged data, compared to collocated teams [9, 19]. Likewise, different communication settings have an impact on the conditions of collaborative work and the coordination capabilities of the teams [8, 16].

The support that command and control DSS provide allows novel ways of distributing and executing tasks as well as organizing work [45]. For example, the mimetic representation of reality influence orders, the way they are communicated, and consequently the overall performance of the command and control system [4]. In other words, many DSS for command and control may lead to radical organizational and technological changes [42]. The design process of DSS for this type of applications therefore has to take into account that actions taken – in our case by command and control teams – are situated and collective [48]. This means that not only individual actions of commanders interacting with a DSS should be studied, but also joint actions of the command and control team, which are situated in the context of the command and control system.

III. Methodology to Support a Design process

The methodology used to support the design process of DSS is based on a series of simulations using a microworld. Microworlds are low-fidelity computer simulations, providing a computer-generated task environment that has complex, dynamic and opaque characteristics [21, 22, 23]. Many of the characteristics in the microworlds are similar to the characteristics of tasks that people normally encounter in ‘real life’ situations, allowing controlled studies of decision-making, both by individuals and teams of decision-makers [22, 26, 27, 28]. The simulation sessions are usually rather short, between 20-30 minutes per session; at the same time, larger sets of sessions are commonly run. As the simulation is fully computer-based, advanced monitoring tools can be used, and data on wide range of parameters can be gathered. Microworld simulations have been used in a number of studies on team decision-making and performance [21, 29, 24, 30, 31, 32]. Microworlds have also been used to study cultural differences in teamwork [33, 34], as well as to investigate effects of information systems on situational awareness [25, 35, 36]. For an elaborated discussion about the use of low-fidelity simulations see [44], and microworlds in research [22, 38].

A. *Simulation Outline*

The basis for the analyses to support the design process of DSS for command and control, presented in this paper, was a series of simulations. In the simulations, command and control teams facing a fire-fighting task were studied. The circumstances were one or more forest fires occurring in a certain geographical area with built-up areas. The task was to extinguish the fire and to protect the built-up areas. The teams were organized as distributed command and control teams with a two-level hierarchical commanding structure. Each team consisted of six persons, where three participants acted as members of a command post and three as ground-chiefs. Each ground-chief controlled three firefighting units. The team capacity was defined by the number of commanders and the organization of the teams. The teams had a limited number of resources to use (nine firefighting units, see Figure 1).

The command post had an access to a low-fidelity DSS, which had the essential functionality of modern command and control DSS for this type of applications. The DSS integrated data from several different sensors: weather, resources, and fire-outbreaks. These sensors were emulated by the simulation environment. The DSS provided the commanders with real-time updated data about (a) meteorological conditions (wind speed and direction), (b) position of the resources and their status, and (c) fire-outbreaks and status of fire-extinguishing efforts. The sensor data were shown on topographical and orthophoto maps in different scales between which the DSS operator could swap. A screenshot of the DSS interface is shown in Figure 2. The command post had access to three text-based communication terminals – of which one was a part of the DSS and the other two were stand-alone – for communication purposes between the command post and the ground-chiefs.

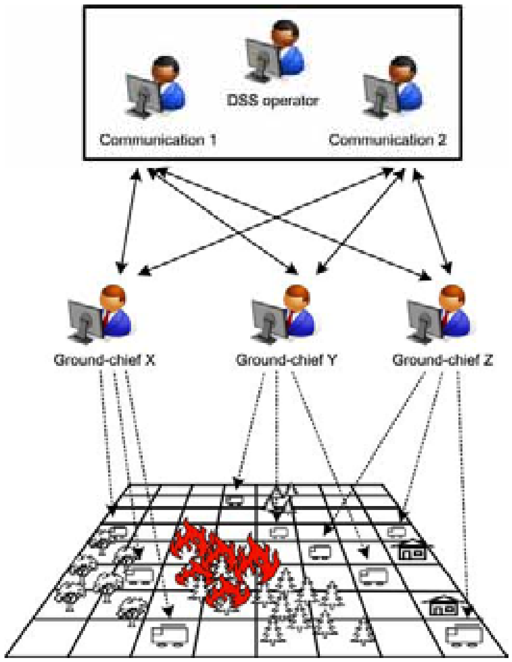


Figure 1. Organization of the command and control teams: (a) one command post with DSS operator and two communication officers, and (b) three ground-chiefs controling three firefighting units each, nine in total.

The ground-chiefs controlled their firefighting units in the simulation through a computer terminal, where other units as well as fires in close proximity to their units were shown on a single layer digital map. Eight teams were studied. Every team performed five trials, lasting for 20 minutes each. The teams were also given an introduction to the simulation for about 20 minutes. The participants were university students (Swedish males and females).

B. *Analytical Focus*

The analyses focused on the way the command and control teams coordinated their actions. The coordination was assessed by analyzing the teams' communication. The teams participating in the simulation were temporary teams, i.e., the teams did not pre-exist. The team members were not trained to use any type of command and control DSS. This is in contrast to real command and control teams, which normally have some previous shared experience, and are trained in using their systems. To minimize the impact of this fact we used only the fifth trial in the simulation in our analyses.

The analyses took into account the volume and distribution of the communication as well as the actual content of the exchanged messages. The exchanged messages were analyzed by categorizing their content according to eleven categories (see Table I). These eleven categories were developed from a modification of thirteen categories used by Svenmarck & Brehmer [39] in their study of distributed command and control. Most of the categories in the Table I are self-explanatory. However, the distinction between the two different types of commands needs clarification. 'Mission' order is an order with a high degree of freedom. 'Direct' order is an order with a high degree of precision, which leaves little room for own initiative. The communication distribution was assessed by calculating the socio-metric status [41]. The socio-metric status is a metrics to measure 'how busy' a team member (as a communication node) is in relation to the other team members. The metrics allow comparison of the communication distribution and load. The socio-metric status is given by the following equation (1):

$$\text{Socio-metric status (ST)} = (1/(g-1)) \sum (x_{ji} + x_{ij}) \quad (1)$$

where g is the total number of team members involved in communication, i and j are individual team members, and x_{ij} are the values representing the number of exchanged messages from team member i to team members j. The socio-metric status of each team member can be then compared with the criterion of key communication agent. The criterion is a sum of mean and standard deviation (2):

$$\text{Criterion (CR)} = \sqrt{(1/(g-1))} \sum (x_i - ((1/g-1) \sum (x_i)) \quad (2)$$

If the socio-metric status of a team member is higher than the criterion this team member has a key role for communication in the team (for detailed explanation, please see [41]).

IV. Analyses and Discussion

A. *Communication Characteristics*

The use of DSS influenced how the commanders engaged in communication. The classification of the communication content revealed that the command post sent to the ground-chiefs in average only 1-2 messages of the type 'question' (category (1) and (2), see Table I). Most of the communication was also initiated by the command post as the commanders did not have to wait for feedback from the ground-chiefs. This observation is also supported by the number of messages 'information about own activity', which corresponded to 5 messages in average (3.2 % of the exchanged messages). As the DSS provided new ways of obtaining feedback the commanders at the command post could see actions of the

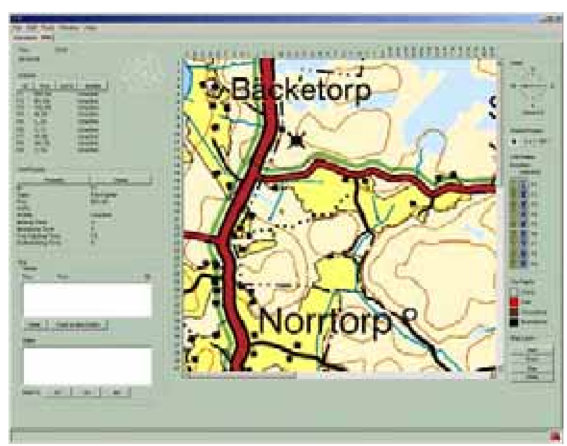


Figure 2. A screenshot of DSS used in the simulation.

TABLE I. COMMUNICATION CLASSIFICATION SCHEME

Eleven categories of communication content			
Question	*Information*	*Order*	*Other*
(1) About fire	(3) About fire	(6) Mission order	(8) Request for help
(2) About other persons activity	(4) About own activity	(7) Direct order	(9) Request for clarification
	(5) About other persons activity		(10) Acknowledge-ment
			(11) Miscellaneous

ground-chiefs as well as effects of the ground-chiefs' actions directly in real-time through the DSS. This resulted in changed communication practice. In traditional one-to-one communication in command and control operations the ratio between sent and received messages is around approximately 1:1 (see for example [18]). The data from the simulations show that this ratio was approximately 3:1 (sent:received) from the perspective of the command post. This is a significant change of the communication practice and further evaluation is necessary in order to identify the impacts of the changed practice. This includes areas, such as effects on awareness of the individual commanders in the team with respect to ongoing actions, development of the operations, intent, commander's workload, etc.

The use of DSS also influenced what was communicated between the command post and the ground-chiefs. The major category of the communication content was 'order' (category (6) and (7), see Table I) representing 57% of the exchanged messages. The ratio between the 'mission' and the 'direct' orders was approximately 1:2. A part of the messages under category 'direct' order was related to the use of DSS. These messages had a form of direct navigational support to the ground-chiefs; for example: COM-1 to ground-chief 1 (GC-1) "go south of stone-valley", COM-1 to GC-1 "to the right, towards (coordinates) AT23", COM-1 to GC-1 "good". This type of communication (navigational support) is a new type of cognitive work, which took place thanks to the use of the DSS. The navigational support thus represents a new task, which requires the commanders at the command post to work on additional – a very short – time scale.

B. *Coordination Strategies*

The DSS did not limit the command and control teams' choice of a coordination strategy. However the navigational support the DSS allowed was incorporated differently in the different coordination strategies the teams chose. All the command and control teams participating in the simulations organized the commanders at the command post in the same way with respect to communication: two communication officers and one DSS operator (see Figure 1). The DSS operator never engaged in communication with the ground-chiefs. In all eight teams the communication was concentrated on one of the communication officers. This communication officer was the team member with the highest socio-metric status; the officer was also the only team member who had the key communication role in the team's communication (see Table II). Though the teams organized their communication in similar way the communication volume was ranging from 34 to 276 exchanged messages (109 exchanged messages in average). By combining the communication distribution and volume together with the communication content analyses four types of coordination strategies were identified among the teams.

First, one communication officer (COM-1) had the role of a primary communication point for all three ground-chiefs; the other communication officer (COM-2) acted as a support function. COM-1 gave all orders (both 'direct' and 'mission' orders). COM-2 provided primarily information about the ongoing fire development and navigational support to the ground-chiefs. This approach was preferred by three teams (we further identify this strategy as *support*).

Second, COM-1 and COM-2 had the same role and were each assigned one or two ground-chiefs to coordinate. Both the COM-1 and COM-2 gave 'mission' and 'direct' orders. This approach was preferred by two teams.

Third, COM-1 and COM-2 had the same role and both coordinated all three ground-chiefs. The COM-1 and COM-2 divided the coordination of the ground-chiefs dynamically on the basis of geography and responsibility for particular fires/fire-sections. Both COM-1 and COM-2 gave 'mission' and 'direct' orders to the ground-chiefs. This approach was preferred by two teams.

Fourth, COM-1 and COM-2 has the same role and both coordinated all three ground-chiefs without any apparent organization of the communication. Both COM-1 and COM-2 gave 'mission' and 'direct' orders to the ground-chiefs. This approach was preferred by one team.

The way the navigational support was incorporated in the four coordination strategies was different. In the coordination

TABLE II. SOCIO-METRIC STATUS AND CRITERION OF KEY COMMUNICATION AGENT (T = TEAM NUMBER; CR = CRITERION; N = NUMBER OF TEAM MEMBERS WITH ST > CR; WHO(1) = TEAM MEMBER WITH ST > CR (COM STANDS FOR COMMUNICATION OFFICER); ST(1) = ST VALUE OF WHO(1); WHO(2) = TEAM MEMBER WITH THE SECOND HIGHEST VALUE OF ST (GT STANDS FOR GROUND-CHIEF); ST (2) = ST OF WHO(2))

Socio-metric status and criterion of key communication agent						
T	*CR*	*N*	*Who (1)*	*ST (1)*	*Who (2)*	*ST (2)*
1	10,1	1	COM	13,0	GC	5,8
2	18,7	1	COM	21,3	GC	17,0
3	9,5	1	COM	11,5	GC	7,5
4	5,8	1	COM	7,0	GC	3,8
5	12,8	1	COM	15,8	GC	8,3
6	19,6	1	COM	23,8	GC	13,0
7	40,5	1	COM	42,0	GC	37,3
8	13,0	1	COM	15,0	GC	10

strategy *support* the navigational support was separated from the rest of the communication compared to the other coordination strategies. The communication of the teams using the *support* strategy was characterized by the lowest number of exchanged messages (in average 49 messages) from all the other teams. Though, it is not apparent if this difference in the number of exchanged messages was caused by the coordination strategy in terms of assigning the navigational support to COM-2.

V. Concluding Remarks

Modern DSS for command and control may lead to radical organizational and technological changes. It is therefore important that the design process takes into account not only individual actions of the commanders operating the DSS, but also joint actions of the command and control team, which are situated in the context of the command and control system.

The simulation series and their analyses, presented in this paper, showed that the real-time feedback on the development in the environment, status of the resources and effects of own actions the DSS provided had impact on how the commanders engaged in the communication and what was communicated. The communication practice changed from the ratio between sent and received messages 1:1 to 3:1 from the perspective of the command post. The commanders at the command post also carried out new type of cognitive work, navigational support. The DSS did not limit the command and control teams' choice of coordination strategy. However the navigational support the DSS allowed was incorporated differently in the different coordination strategies the teams chose. The findings from the simulations highlight the need for attention to possible impacts of DSS on the collective and situated aspect of the command and control teams' actions.

The discussed methodology, which is based on low-fidelity simulations, provided a possibility to investigate the actual work and the interaction within the command and control teams when using the DSS. In this sense the low-fidelity simulations proved to be a feasible methodology to support the design process of DSS for command and control applications. The methodology also allows testing of various design assumptions and ideas, such as a variety of features of the designed DSS, but also different organization settings and team structures. The advantages of this methodology are the possibility to collect a wide range of qualitative and quantitative data, to carry out larger repeatable simulation series and to compare parallel designs.

Acknowledgment

This paper is based on research funded by the Swedish Emergency Management Agency (SEMA).

References

[1] K. Schmidt, and L. Bannon, "Taking CSCW seriously: Supporting articulation work," *Computer Supported Cooperative Work*, vol. 1, no. 1, 1992, pp. 7-40.

[2] E. Hutchins, *Cognition in the Wild*. MIT Press/Cambridge MASS, 1995.

[3] E. Hollnagel, and D.D. Woods, *Joint Cognitive Systems: Foundations of Cognitive Systems Engineering*, Taylor & Francis, 2005.

[4] B. Johansson, H. Artman, and Y. Waern, "Technology in crisis management systems – ideas and effects," *J. Research and Problem Solving in Organizational Communication*, vol. 2, no. 3, 2001, pp. 247-258.

[5] J. Brynielsson, *Gaming Perspective on Command and Control*, Stockholm: Royal Institute of Technology, 2006.

[6] B. Brehmer, "System design and the psychology of complex systems," in *Empirical Foundations of Information and Software Science*, J. Rasmussen and P. Zunde, Eds. New York: Plenum Publishing Cooperation, 1987.

[7] D. Crookall, and D. Saunders, "Towards an integration of communication and simulation," in *Communication and Simulation*, D. Crookall, Ed. Multilingual Matters, 1998, pp. 3-32.

[8] J. Orasanu, and E. Salas, "Team decision making in complex environments", in *Decision making in action: Models and methods*, G.A. Klein, J. Orasanu, R. Calderwood and C.E. Zsambok, Eds. Norwood, NJ: Ablex Publishing, 1993, pp. 327-345.

[9] H. Artman, "Team situation assessment and information distribution," *Ergonomics*, vol. 43, no. 8, 2000, pp. 1111-1128.

[10] P.E. Jones, and P.H.M.P. Roelofsma, "The potential for social contextual and group biases in team decision-making: biases, conditions and psychological mechanisms," *Ergonomics*, vol. 43, no. 8, 2000, pp. 1129-1152.

[11] L. Shattuck, and D.D. Woods, "Communication of intent in military command and control systems," in *The Human in Command: Exploring the Modern Military Experience*, C. McCann and R. Pigeau, Eds. New York: Kluwer Academic/Plenum Publishers, 2000, pp. 279-292.

[12] C. Perrow, *Normal Accidents: Living with High-Risk Technologies*, New York: Basic Books, 1984.

[13] P-A. Persson, *Towards a Grounded Theory for Support of Command and Control in Military Coalitions*, Linköping: Linköping University, 1997.

[14] B. Brehmer, "Modern information technology: timescales and distributed decision making," in *Distributed Decision Making: Cognitive Models for Cooperative Work*, J. Rasmussen, B. Brehmer, and J. Leplat, Eds. John Wiley & Sons, 1991, pp. 193-200.

[15] B. Brehmer, and P. Svenmarck, "Distributed decision making in dynamic environments: time scales and architectures of decision making," in *Contributions to Decision Making – I,* J.-P. Caverni, M. Bar-Hillel, F.H. Barron and H. Jungermann, Eds. Elsevier Science, 1995, pp. 155-174.

[16] B. Johansson, and E. Hollnagel, "Pre-requisites for large scale coordination," *Cognition, Technology & Work*, vol. 9, no. 1, 2007, pp. 5-13.

[17] B. Johansson, *Joint control in dynamic situations*, Linköping: Linköping University, 2005.

[18] J. Trnka, B. Johansson, and R. Granlund, "Information Support in Collaborative Command and Control Work – An Empirical Research Using a Role-Playing Exercise Approach, " *Eleventh International Command and Control Research and Technology Symposium*, Cambridge, 2006.

[19] J.E. Driskell, P.H. Radtke, and E. Salas, "Virtual teams: effects of technological mediation on team Performance," *Group Dynamics: Theory, research and Practice*, vol. 7, no. 4, 2003, pp. 297-323.

[20] R. Woltjer, *On how constraints shape action*, Linköping: Linköping University, 2005.

[21] P. Svenmarck, and B. Brehmer, "D3Fire, an experimental paradigm for the study of distributed decision making," in *Distributed Decision Making*, Roskilde, B. Brehmer, Ed. Risö National Laboratory, 1991.

[22] B. Brehmer, and D. Dörner, "Experiments with computer-simulated microworlds: escaping both the narrow straits of the laboratory and the deep blue sea of the field studies," *Computer in Human Behaviour*, vol. 9, no. 2/3, 1993, pp. 171-184.

[23] R. Granlund, *C3Fire – A Microworld Supporting Emergency Management Training,* Linköping: Linköping University, 1997.

[24] R. Granlund, *Monitoring Distributed Teamwork Training*, Linköping: Linköping University, 2002.

[25] R. Granlund, "Monitoring experiences from command and control research with the C3Fire microworld," *Cognition, Technology & Work*, vol. 5, no. 3, 2004, pp. 183-190.

[26] G. Rolo, and D. Diaz-Cabera, "Empirical Evidences about Process Control Systems Simulation," *First Workshop on Cognitive Research with Microworlds*, Granada, 2001, pp. 15-21.

[27] R. Granlund, B. Johansson, M. Persson, H. Artman, and P. Mattsson, "Exploration of methodological issues in micro-world research – experience from research in team decision making," *First Workshop on Cognitive Research with Microworlds*, Granada, 2001, pp. 73-82.

[28] C. Gonzalez, P. Vanyukov, and M.K. Martin, "The use of microworlds to study dynamic decision making," *Computers in Human Behaviour*, vol. 21, no. 2, 2005, pp. 273-286.

[29] D.E. Howie, and K.J. Vicente, "Measures of operator performance in complex, dynamic microworlds: advancing the state of the art," *Ergonomics*, vol. 41, no. 4, 1998, pp. 485-500.

[30] R. Granlund, "Monitoring experiences from command and control research with the C3Fire microworld," *Cognition, Technology & Work*, vol. 5, no. 3, 2003, pp. 183-190.

[31] R. Granlund, and B. Johansson, "Monitoring Distributed Collaboration in C3Fire Microworld," in *Scaled Worlds: Development, Validation and Applications*, S.G. Schiflett, L.R. Elliott, E. Salas and M.D. Coovert, Eds. Aldershot, Ashgate, 2003.

[32] M-E. Jobidon, R. Breton, R. Rousseau, and S. Tremblay, "Team response to workload transition: the role of team structure," *Proc. 50th Annual Meeting of the Human Factors and Ergonomics Society*, San Francisco, 2006.

[33] I. Lindgren, and K. Smith, "Using microworlds to understand cultural influences distributed collaborative decision making in C2 settings," *Proc. 11th Int. Command and Control Research and technology Symposium (ICCRTS)*, Cambridge, 2006.

[34] I. Lindgren, and K. Smith, "National patterns of teamwork during an emergency management simulation," *Proc. 50th Annual Meeting of the Human Factors and Ergonomics Society*, San Francisco, 2006.

[35] H. Artman, and R. Granlund, "Team situation awareness using graphical or textual databases in dynamic decision making," in *Ninth Conference of Cognitive Ergonomics,* T. Green, L. Bannon, C. Warren and J. Buckley, Eds. Limerick, 1998, pp. 149-155.

[36] B. Johansson, J. Trnka, and R. Granlund, "The effect of geographical information systems on a collaborative command and control task," in *Fourth International Conference on Information Systems for Crisis Response and Management (ISCRAM)*, B. Van den Walle, P. Burghardt and C. Nieuwenhuis, Eds. Delft, 2007, pp. 191-200.

[37] P.C. Rasker, W.M. Post, and J.M.C. Schraagen, "Effects of two types of intra-team feedback on developing a shared mental model in command & control teams," *Ergonomics*, vol. 43, no. 8, 2000, pp. 1167-1189.

[38] B. Brehmer, "Some reflections on microworld research," in *Scaled Worlds: Development, Validation and Applications*, S.G. Schifflet, L.R. Elliott, E. Salas and M.D. Coovert, Eds. Ashgate Cornwall, 2004.

[39] P. Svenmarck, and B. Brehmer, *D3FIRE - An Experimental Pradigm for the Study of Distributed Decision Making,* Upsala University, 1994.

[40] G.A. Klein, and M.L. Thordsen, *Cognitive processes of the team mind,* Yellow Springs: Klein Associates.

[41] R.J. Houghton, C. Barber, R. McMaster, N.A. Stanton, P. Salmon, R. Stewart, and G. Walker, "Command and control in emergency services operations: a social network analysis," *Ergonomics*, vol. 49, no. 12, 2006, pp. 1204-1225.

[42] M.L. Cummings, *Designing Decision Support Systems for Revolutionary Command and Control Domains*, University of Virginia, 2004.

[43] L. Adelman, "Experiments, quai-experiments, and case studies: a review of empirical methods for evaluating decision support systems," *IEE Transactions on Systems, Man, and Cybernetics*, vol. 21, no. 2, 1991, pp. 293-301.

[44] J.L. Weaver, C.A. Bowers, E. Salas, and J.A. Cannon-Bowers, "Networked simulations: new paradigms for team performance research," *Behaviour Research Methods, Instruments, & Computers*, vol. 27, no. 1, 2000, pp. 12-24.

[45] D.D. Woods, and S. Dekker, "Anticipating the effects of technological change: a new era of dynamics for human factors," *Theoretical Issues in Ergonomics Science*, vol. 1, no. 3, 2000, pp. 272-282.

[46] D.D. Woods, and E. Hollnagel, *Joint Cognitive Systems: Patterns in Cognitive Systems Engineering*, CRC Press, 2006.

[47] N. Wiener, *Cybernetics or Control and Communication in the Animal and the Machin*e, Cambridge: MIT Press, 1948.

[48] M. Arvola, *Shades of use: The dynamics of interaction design for sociable use*, Linköping: Linköping University, 2004.

[49] H. Artman, and M. Persson, "Old practices – new technology: Observations of how established practices meet new technology," in *Designing Cooperative Systems*, R. Dieng, A. Gibson, L. Kersenty, G. De Michelis, Eds. Amsterdam: Ios Press Ohmsha, 2000, pp. 35-49.

Clustering of Texts Facilitating Efficient Searching in Spatial Search Engines

R. Göbel and R.P. Lano
University of Applied Sciences Hof
Alfons-Goppel-Platz 1
95028 Hof, Germany

***Abstract-*[1]The University Hof develops a spatial search engine for security applications. In contrary to existing Internet search engines, support of spatial search criteria is a crucial requirement in this application context. A key challenge for such a search engine is the construction of index structures supporting efficient searching with combined spatial and textual search criteria. Existing database index structures cannot easily be extended for this purpose since they use a "divide and conquer strategy" splitting sets of entries into disjoint subsets according to a search condition. This approach is not directly applicable for text searching since two documents have usually many terms in common. This paper proposes a new approach by grouping documents according to terms which do not occur in a set of documents (forbidden or exclusion terms). In this case some groups of documents can be ignored at least for some searches. This paper analyses this approach by presenting results of first simulations. In addition it proposes a concept of a self organizing cluster by identifying appropriate exclusion terms and transferring documents between nodes to maximize the number of exclusions terms on the nodes of the cluster.**

I. Introduction

The Internet becomes increasingly important as a source of information for disaster management and disaster recovery. In this application context users often access recent news for getting context information during an evolving crisis. As an example an aid organization may collect information about the infrastructure to prepare its engagement in the case of an earthquake or a tsunami. Also a local organization may benefit from recent information on regional events, which may have an impact on rescue activities. On a political level advisors may monitor activities in a distant country by analyzing articles from news providers in this country.

Starting point for the acquisition of information from the Internet are usually search engines. Current search engines, however, fall short to provide sufficient support in this context. A key issue is the missing support for spatial search criteria. Without spatial search criteria it is difficult or even impossible to define sufficiently precise conditions for retrieving documents referring to a certain location.

The support of spatial criteria in search engines has been investigated in different research projects (see for example [9], [10], [15]) A key challenge for this type of search engines is the deduction of the geographical position from texts found in the Internet (e.g. [1], [2]). In the case of a spatial search engine for security applications this deduction is slightly simpler because only news need be considered here. This type of documents provides geographical positions by place names of well known locations. A concept for such a specialized spatial search engine has been proposed in [3] and implemented as a prototype.

The other key challenge for spatial search engines is the fast access to a large number of documents using search conditions combining space and time (efficient searching). This challenge needs to be also addressed in its full complexity for a spatial search engine in the security domain. In fact efficient searching is probably a more critical issue in this domain due its tight time constraints. So far relatively little is known about efficient searching for combined textual and spatial search criteria. Therefore one focus of our research is on hybrid index structures.

Index Structures

Key for fast searching is a so called "index structure". Index structures support quick navigation to search results avoiding the checking of every individual entry. A typical index structure is the B-Tree (see [4]) which is available in every Object/Relational Database Management System. Text Databases use an "Inverted Index" which is more appropriate in this context. An inverted index manages a set of terms where every term points to all documents containing it. This approach might be combined with other index methods (e.g. [11]) to improve efficiency (e.g. for substring searches).

Spatial searches are usually supported by the so called R-Tree (see [12]). Although R-Trees are probably the best index method for spatial searches at present, it is difficult to ensure fast response for very large databases in every case (see for example [13]). Therefore a wide number of extensions had been proposed (e.g. [5], [6], [14]) which do solve some but not all problems. For some special cases, however, this tree could be optimized to meet efficiency requirements (see for example [7]).

The spatial search engine needs to support search conditions combining textual and spatial search criteria. This means that also specific search indexes need to be available for this purpose. In general three different options are possible:

1) Independent indexes are built for the spatial criteria and the full text search. Separate searches are executed for every

1 This work has been funded by the Deutsche Forschungsgemeinschaft and the Oberfrankenstiftung.

criterion. The intersection of all intermediate result sets will provide the final result set. The approach for generating the intersection might use techniques like marking algorithm or bit lists. Note that this approach is slow if very large intermediate result sets need to be handled.

2) The index structures for spatial and full text searches are connected with each other. One approach is a single R-Tree containing all locations and every location is the starting point for an inverted index organizing documents related to this location (option 2.1). The other approach is an inverted index where every term is a starting point for an R-Tree organizing the geographical positions of documents with this term (option 2.2).

3) A search could be supported by a single index managing both positions and terms in a single tree.

Our existing prototype uses an index structure starting with an R-Tree containing all geographical locations. Every geographical location in this tree points to an inverted index managing all documents for this location (option 2.1). This option and the other option 2.2 were also described in [15]. This paper compares these two structures by applying them to some test data. These first tests seem to indicate that searches can be processed faster for option 2.2. This paper, however, does not provide any worst case analysis for time and space. In particular the space requirements are not considered at all.

A more formal analysis of option 2.1 and 2.2 is given by [16]. This analysis shows that option 2.1 is space but not time efficient. In contrary option 2.2 supports efficient searches but uses too much space to be applied in a larger application.

The key problem with option 2.1 are conditions which require the searching of a significant part of the initial R-Tree but the search term is only contained in few or even no inverted indexes reachable from the visited leaf nodes. This situation can be avoided by including information about available terms higher in the index structure. Therefore [16] proposes a new index structure extending an R-Tree by bit lists at every node. Every bit in this list provides the information whether the related term is contained in the sub tree below this node. This approach is currently investigated in an experimental implementation. The key challenge here is to reduce the length of the bit list to a reasonable size. A potential solution might be the exclusion of very rare terms and very frequent terms from the bit list since different efficient solutions exist for these types of terms.

New Strategy

In this paper we propose a different solution based on option 2.1. This solution is based on a cluster computer, where every node is an independent system managing a set of documents which is disjoint from the sets of documents on other nodes. In this case a search is automatically distributed to all nodes and processed in parallel by every individual node. With this approach every node of the cluster needs to search a smaller amount of data reducing the overall search time.

In contrary to a conventional cluster we investigate a solution which dynamically reorganizes the locations of documents in the cluster to reduce the nodes which need to be searched for a single query. Since every document refers to more than one location and contains a larger number of terms, it is not obvious how to group documents on nodes with respect to these criteria. The key idea here is to group nodes according to criteria which are **not** satisfied for a given group. In this case it is possible to ignore this group of documents for this criterion. In this paper we focus on the absence of certain terms in documents as the criterion for grouping them.

This paper starts with a simulation investigating the feasibility of this approach: Then the efficiency of the cluster approach is analyzed. Finally we introduce a concept for automatic relocation of documents in a cluster to increase the number of "forbidden terms" on nodes.

II. SIMULATION

To analyze the proposed strategy we performed a set of simulations. The simulations use as parameters the number of nodes of the cluster (N), the number of documents (D), the total number of terms in all documents (T), the number of terms per document ($T_D<T$), the number of "critical" search terms ($T_C<T$), the number of forbidden terms per node ($T_F<T_C$) and the number of search terms in a query (T_Q). The number of "critical" search terms refers to those terms which occur only in a limited number of documents. Search conditions with these terms may require that a large part of an R-Tree needs to be searched but only few or no found documents contain these terms. This is exactly the worst case scenario for our index structure starting with an R-Tree and followed by inverted indexes.

A key challenge of our approach is to assign as many forbidden terms as possible to the nodes of a cluster. In a simulation we investigate the dependency of the number of forbidden terms from these other parameters as well as the dependency between the number of forbidden terms and the number of nodes which need to be searched. For this purpose the simulation performs the following steps:

1) In an initial step, all documents are assigned random terms, where the number of terms for each document is equal to T_D. In a similar way the simulation randomly assigns a bit-list of T_F forbidden terms to each node.

2) Next, assign the documents to the nodes. Simply compare the terms a document contains with the forbidden list of each node, and if permitted, place the document in that node. If the document can not be placed in any node, we put it in the "Trash" node. Ideally, at the end, there should only be a very small number of documents in the Trash.

3) Finally, we are interested to find out, how many nodes need to be searched for a query with T_Q terms. Only those nodes need to be searched where none of the search term is a forbidden term (implicit logical "AND" between search terms).

The simulation uses a random number generator providing equally distributed data. This means for example that all terms occur with equal probability in a document or a forbidden term list. This approach is a worst case scenario since terms from real documents are by far not equally distributed. An unequal distribution, however, can be used to increase the number of forbidden by selecting rare terms here.

The first diagram directly analysis the relationship between the number of forbidden terms per node and the number of documents assigned to the trash. We have used 50 nodes, 100,000 terms, 1000 critical search terms and 1000 terms for every document. The simulation generates 10,000 documents and tries to find a node for every document. If no node could be found then the document ends up in the trash.

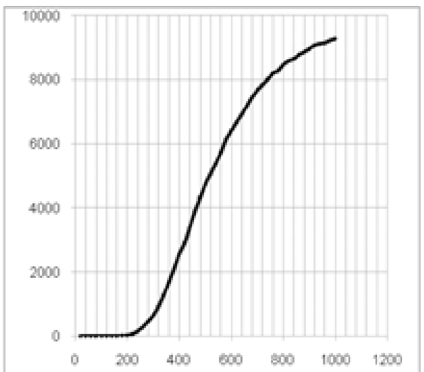


Fig. 1. Trash Size as a Function of the Number of Forbidden Terms

This diagram shows that the trash remains almost empty up to 200 forbidden terms per node. Above 200 forbidden terms the size of the trash grows quickly to unacceptable values. A trash size is probably acceptable if the trash does not contain more terms than the other nodes. With 50 nodes the trash should probably not contain more than 2% of all documents. In our simulation we reach this figure with around 250 forbidden terms.

We have verified the robustness of the simulation by varying the values for nodes and documents. With up to 1,000,000 documents we do not get significantly different results. Also increasing the number of nodes changes the simulation results only slightly. A larger number of nodes make it easier to find a node for a document. Nevertheless the size of the trash node decreases only very slowly with an increasing number of nodes.

Increasing the total number of terms beyond 100,000 has only a limited effect on the results of the simulation as well. Increasing the number of critical search term results in a small decrease for the number of documents in the trash, if the difference between the total number and this figure is not significantly reduced. Decreasing this difference significantly will, however, result in a much larger trash size. In Figure 2 we increase the number of critical search terms stepwise from 1000 to 10,000 and decrease the number of total terms from 100,000 to 10,000 at the same time. Here we assume 100 forbidden terms per node. The other parameters are the same as above.

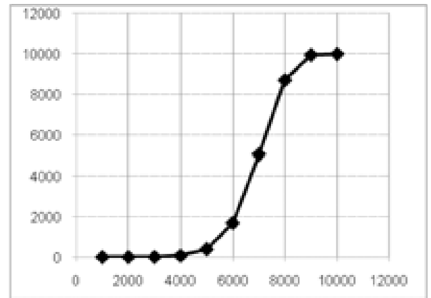


Fig. 2. Trash Size vs. Increasing T_C and Decreasing T

The results show that the worst case is given, if the number of critical search is identical to the total number of terms. It is probably interesting to analyze an acceptable number of forbidden terms for this case. The results of this simulation are given in Figure 3.

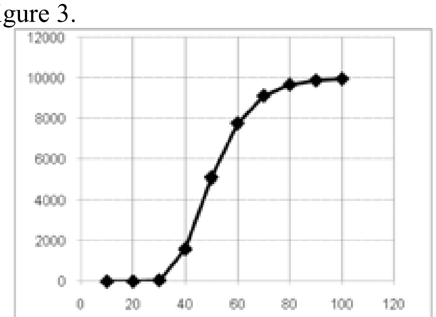


Fig. 3. Trash Size vs. Number of Forbidden Terms for $T = T_C$

In Figure 3 an acceptable number of forbidden terms is well below 40. This value is probably too low in most application contexts. As a consequence we need to ensure that the number of critical search terms is significantly smaller then the total number of terms. On the other hand experience from real search engines show, that the total number of terms for a larger number of documents is very high and often above 1,000,000 terms. Since very rare terms and very common terms need not to be considered in this context (other solutions exist in this case), it is a realistic assumption to identify a set of critical search terms which is much smaller than the total number of terms.

Figure 4 provides the results of a simulation analyzing the percentage of nodes which need to be visited. We relate this figure with the number of terms in a query increasing from 1 to 10. For the other parameters we choose 50 nodes, 100,000 terms, 1,000 critical search terms and 200 forbidden terms per node.

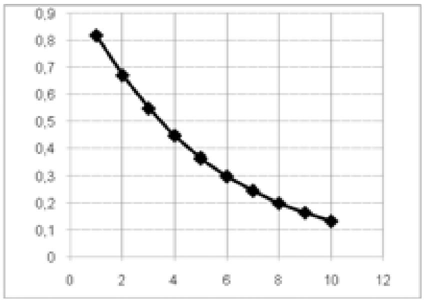


Fig. 4. Search Effort as Number of Nodes vs. Number of Query Terms

Figure 4 shows that single term queries still require that most nodes need to be searched (more than 80%). For the more common multi-word queries, however, a significant reduction can be expected. As an example only half of the nodes need to be searched for a three word query.

III. Analysis

In general the reduction of search time for a cluster computer strongly depends on the effectiveness of the index structure. Only a small reduction can be expected for an effective index structure. Assume for example that the search time grows logarithmic ($O(log(n))$) with the number of documents n then the benefit of distributing the data to k nodes is relatively small ($O(log(n/k))$). On the other hand the search time significantly decreases if we assume a linear time complexity for a search ($O(n/k)$).

In our case we use an R-Tree followed by inverted indexes for every location in a leaf node. In the best case this index would ensure a logarithmic search time complexity for a tight search condition identifying only a small number of locations, if the R-Tree is constructed according to the condition from [7]. In the worst case, however, a large part of the R-Tree would need to be searched as well as all inverted indexes connected to locations in a visited leaf node. This means that the search effort would grow at least linearly with the number of nodes in the R-Tree.

The number of locations in the R-Tree and the number of terms in the inverted indexes grow slower then linear with the number of documents. With Heaps' law [8] the total number of terms t is related to the summarized size n of all documents by the following formula:

$t = \alpha * n^{\beta}$

The constant β is usually significantly smaller than 1 (e.g. 0.4 – 0.6). The constant α strongly depends on the type of terms and documents.

If we assume documents of approximately equal size then we may assume that n represents the number of documents in this formula. We may use this formula as well for the number of locations derived from these documents, since these locations are derived from place names. It should be however noted that at least the constant α will be significantly smaller in this case.

This formula provides an immediate estimate for the search time of the R-Tree depending on the number n of documents using the O-notation and thus ignoring the constant α here.

$O(n^{\beta})$ search time complexity for R-Tree

The number of terms in an attached index structures is also given by this formula with other constants α' and β' depending on the subset of documents n' managed by this index. The search time in an inverted index usually grows not worse than logarithmic with the number of contained terms. This means that we may approximate the search time relative to the number of managed documents as follows:

$O(log(\alpha' * n'^{\beta'})) = O(log(\alpha') + \beta' * log(n'))$

If we ignore the constants and assume that n' grows linearly with the total number of documents then we get the following search time complexity for searching a single inverted index:

$O(log(n))$

This asymptotic search effort is needed for searching every inverted index linked to a found location. Since we search the full R-Tree we get the following worst case complexity for searching the combined index structure:

$O(n^{\beta} * log(n))$

The following table shows the growth of search time with the number of documents in the worst and the best case. We have used 0.5 as the value for $\beta = 0.5$ (square root).

n	log(n)	n^{β}*log(n)
100,000	5	1,581
1,000,000	6	6,000
10,000,000	7	22,136
100,000,000	8	80,000

From this table we can immediately deduce the consequences of distributing the data to 10 nodes of a cluster by comparing the values from the previous row with the current row. Only in the last column this approach leads to a significant decrease in search time. Still this decrease is by far not linear with the size of the cluster.

Our proposed approach will not decrease the search time further. Instead it avoids that all nodes of the cluster need to be searched for every single query. The saved computation time may be used for other tasks (e.g. processing of more queries in a given time interval). As a consequence this method will reduce the total computation time for queries and improve the benefit of adding additional nodes to a cluster.

IV. Overall Architecture

The above simulation results are encouraging, and therefore in the remainder we are considering an implementation strategy considering these results. The overall architecture of the system will follow that of a typical search engine and is comprised of the following components:

1) A Crawler: navigates the Internet by starting at a given set of URI's, retrieving all documents reachable directly or indirectly from these initial URI's. In our case the crawler includes a finder starting at seed URI's from a database and identifying for UIR's in all reachable documents. The documents itself are retrieved by a document builder.

2) A Data Archive: holds all information about documents needed to support searches. This data archive contains an analyzer extracting terms and phrases from documents. For our search engine the analyzer will also derive a set of locations from each document by identifying place names in the contained text. The information derived by the analyzer is stored in appropriate index structures (Metadata). The documents could be stored as well to provide a cache, if an original document has been changed or is (temporarily) not accessible.

3) A Search Interface: (typically a web interface) provides access to the Search Engine for end users.

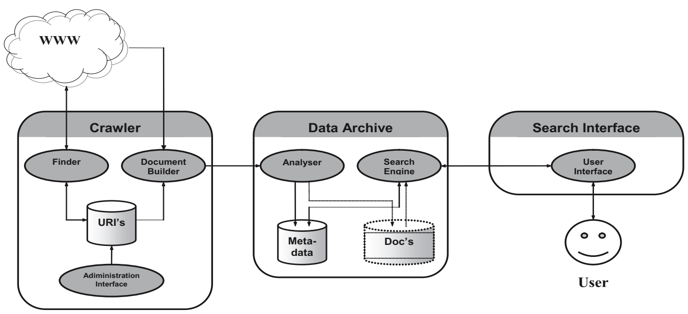


Fig. 7. Overall Architecture of System.

Every node of the cluster will operate its own crawler and its own data archive. The data archives will provide standardized interfaces facilitating access of remote user interfaces to search their data holdings. In addition the nodes of the cluster will inform directories about the set of terms in their local documents. These directories will provide inverted indexes containing terms pointing to all relevant nodes. The relationship between a term and the nodes is given by a bit list where every bit corresponds to a node of the cluster.

A search interface will query such an inverted index first to find all relevant nodes for a subsequent document search. For this purpose the interface will retrieve the corresponding bit lists of all terms from the query and process these bit lists with logical operators corresponding to the structure of the query.

A key functionality of the cluster is the dynamic relocation of documents to increase the number of forbidden terms on a node. The directory plays a central role here since it needs to maintain statistics about the frequency of terms in documents. In particular the inverted index of a directory will contain the global number of documents containing a term. In addition the local number of documents containing a term will be provided for every node. This information will be used by scheduled tasks on the individual nodes trying to indentify terms as candidates for relocation. Good candidates for this relocation are globally rare terms which local frequency is significantly below their global frequency. In this case the node would try to move the few local documents with this term to other nodes.

The relocation of terms is supported by a second inverted index on directory nodes. This inverted index contains all forbidden terms pointing to the corresponding nodes. Also here the relationship between terms and nodes is given by bit lists. A bit in this list is set to zero if it corresponds to a node where the term is forbidden.

The bit lists for all terms from a document are retrieved and combined with a logical "AND". Those bits which still contain a "1" refer to nodes which may take this document. If no node could be found for a document, then the relocation is interrupted and the term is stored for a limited time in a queue (“relocation avoidance queue”) to avoid another relocation attempt during this time period.

The inverted index is also used to identify a node for a new document discovered by a crawler. A crawler will not necessarily store a found document in its local data archive since the document may contain some forbidden terms from this node. If this is the case another node is selected for the document and a reference to this document is stored for the local crawler to prepare for future visits of the document site.

In some cases it may happen that a new document cannot be assigned to any node. In this case the “most forbidden term” (many nodes forbid this term) will be ignored and the system retries to find a node. If again no node could be found then the next “most forbidden terms” are stepwise ignored until a node could be found. The conflicting terms are removed from this node as forbidden terms and stored temporarily in the “relocation avoidance queue”. The required information about the number of nodes which do not contain a term needs to be stored and maintained in the inverted index as well for every term.

The efficiency of the proposed relocation approach strongly depends on the efficiency of the two inverted index in the directory. If we assume a limited number of nodes (< 10,000) then it is probably feasible to keep these inverted indexes in main memory. The access to the terms could be supported by a data structure like the Patricia tree [11]. A Patricia Tree ensures an access time growing linearly with the size of the search term but not growing at all with the number of stored terms. Joining bit lists using Boolean operators is one of the most efficient operations anyhow on a standard computer if the bit lists are short. The length of these bit lists are directly limited by the number of nodes in the cluster. This means that the total time for operations on these inverted indexes is small relative to all other operations like searching documents on nodes or moving documents from one node to another.

V. Conclusion

Clustering of texts using forbidden terms seems to be a realistic option for reducing the overall search effort in classical and spatial search engines. Our simulation provides evidence for this thesis if some constraints are considered. The key constraint is the identification of a subset of critical terms for which the proposed index structure fails in supporting efficient searching. This subset of critical terms needs to be significantly smaller than the total set of terms from all documents.

The required dynamic relocation of documents for increasing forbidden terms on nodes can be efficiently supported by inverted indexes combined with bit lists. This data structure supports the efficient identification of nodes as search targets as well as candidate nodes which may take a document.

One of our next steps will be the integration of the proposed method into the prototype of our search engine. We are currently preparing a cluster with 50 nodes for this purpose. This cluster will allow the investigation of the relation between parameters of this method and the overall search efficiency using a realistic set of documents in the envisaged application domain. As an example we will investigate different methods for the choice of critical search terms and heuristics selecting the “right” node for a document.

REFERENCES

[1] E. Amitay, N. Har’El, R. Sivan, A. Soffer: “Web-a-Where: Geotagging Web Content”, Proceedings SIGIR’04, ACM Press, pages 273 – 280, Sheffield, UK, 2005

[2] J. Ding, L. Gravano, N. Shivakumar: “Computing Geographical Scopes of Web Resources”. Proc. 26th VLDB Conference, Cairo, Egypt, 2000

[3] R. Göbel, L. Colaiacomo, A. de la Cruz, “A Spatial Search Engine Supporting Security Applications,” *Proceedings Fourth Conference Image Information Mining, ESA Workshop Proceedings*, 2006.

[4] R. Bayer, C. McCreight: “Organization and Maintenance of Large Ordered Indexes,” Acta Informatica 1(3), pp.173-189, 1972.

[5] N. Beckmann, H. P. Kriegel, R. Schneider, B. Seeger: "The R*-tree: An Efficient and Robust Access Method for Points and Rectangles", Proc. ACM-SIGMOD International Conference on Management of Data, pages 322 - 331,,Atlantic City (NY), May 1990

[6] S. Berchtold, D. A. Keim, H. P. Kriegel: "The X-Tree: An Index Structure for High-Dimensional Data", Proc. International Conference on Very Large Data Bases, pages 28 – 39, Mumbai (Bombay), India, 1996

[7] R. Göbel: “Towards Logarithmic Search Time Complexity for Two-Dimensional R-Trees”, Proceedings of Second International Conference on Systems, Computing Sciences and Software Engineering 2006, Springer, to appear in 2007

[8] H.S. Heaps: “Information Retrieval – Computational and Theoretical Aspects”, Academic Press, 1978

[9] Jan Torben Heuer, Sören Dupke: “Towards a Spatial Search Engine Using Geotags”, Proceedings GI-Days 2007 - Young Researchers Forum, Florian Probst and Carsten Keßler (Eds.), IfGIprints 30. ISBN: 978-3-936616-48-4, 2007

[10] C. B. Jones, A. I, Abdelmoty, D. Finch, G. Fu, S. Vaid: “The SPIRIT Spatial Search Engine: Architecture, Ontologies and Spatial Indexing”, Proceedings Third International Conference on Geographic Information Sciences, pages 125 – 139, Adelphi, MD, USA, 2004

[11] D. R. Morrison, “PATRICIA - Practical Algorithm To Retrieve Information Coded in Alphanumeric,” J. ACM 15(4), pp.514-534, 1968.

[12] A. Guttman, “R-Trees, A Dynamic Index Structure for Spatial Searching,” Proc. ACM SIGMOD Conference, Boston, pp.47-57, 1984.

[13] K.V. Ravi Kanth, A. Singh, “Optimal Dynamic Range Searching in Non-Replicating Index Structures,” Proc. International Conference on Database Theory, pages 257 - 276, Springer Verlag Berlin Heidelberg, 1999.

[14] T. Sellis, N. Roussopoulos, C. Faloutsos: "The R+-Tree: A Dy-namic Index for Multi-Dimensional Objects", Proc. International Conference on Very Large Data Bases, pages 507 – 518, Brighton, September 1987

[15] Y. Zhou, X. Xie, C. Wang, Y. Gong, W.Y. Ma, “Hybrid Index Structures for Location-based Web Search,” Proceedings of the 14th ACM international conference on information and knowledge management, Bremen 2005.

[16] R. Göbel, A. de la Cruz, “Computer Science Challenges for Retrieving Security Related Information from the Internet,” GMOSS Publication, 2007.

A System Concept for Intelligent Surveillance of a Critical Infrastructure

Erland Jungert, Christina Grönwall, Niklas Hallberg, Fredrik Lantz
FOI (Swedish Defence Research Agency)
Box 1165, S-581 11 Linköping, Sweden
{jungert, Stina, Nikha, flantz}@foi.se

Abstract: *The protection of critical infrastructures has become an increased necessity over the last few years. Essential parts of the protection of such objects are the usage of sophisticated (1) methods and techniques for data collection, (2) event analysis and (3) support for decision making. An approach to the development of this type of surveillance systems is a systems architecture to which sensors for data collection are integrated together with various decision support tools. In this work, the focus has been to develop a system for early determination of potential threats. A system architecture, which is modular, is based on three basic concepts, i.e. services, views and user roles. The input data are generally originating from sensor systems, which in order to conveniently support the users include resources for information fusion. Basically, in this work, a system architecture is described including an underlying approach for the determination and analysis of potential antagonistic threats.*

1. INTRODUCTION

Protection of critical infrastructures has recently become necessary due to a number of spectacular attacks during the last decade. For this reason, a system architecture for protection of primarily connection nodes in critical infrastructures, e.g. nodes of electrical power supply networks, is proposed. Such nodes can be substations, nuclear power plants or other similar establishments. The focus in this work has been of the first type. The approach is based on three fundamental notions or concepts, i.e. services, views and user roles. To form the basis for the eventual system structure an investigation of the substations has been carried out. This study was concerned with how to determine events that possibly can be serious and for this reason either will cause alarms or in cases where the consequences of the event is unpredictable just a type of warning.

In Garcia [3] physical protection of infrastructure nodes are discussed from a technical point of view. Trevedi et al. [12] discuss distributed interactive video arrays for event capture and enhanced situational awareness. Other works of interest, which also concern the protection of critical infrastructure, are e.g. [9] and [13]. Other aspects of concern in modern systems for infrastructure protection are that they must be based on flexible system architectures. Generally this leads to systems that are service based. Works that emphasizes on service oriented approaches are for example [4] and [8]. In [5] Hu et al. gives a survey on visual surveillance with a focus on object motion and behaviors.

The structure of this paper is as follows. Section 2 defines the basic problem addressed here. Events of interest for determination during protection of the critical infrastructure nodes are discussed in section 3. Methods for sensor data analysis are presented and discussed in section 4. In section 5 are methods for data analysis for intelligent surveillance discussed. The system architecture is presented in section 6 while some aspects on possible scenarios suitable for testing the system are discussed in section 7. The conclusions of the work appear in section 8.

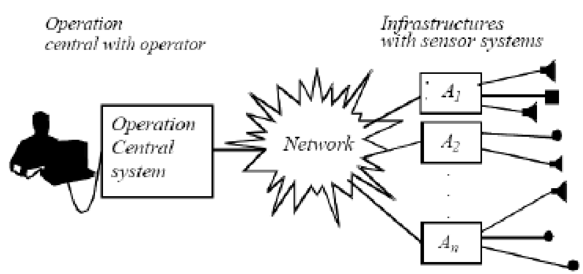


Figure 1. The surveillance system and its subparts.

2. PROBLEM AND REQUIREMENTS

The problem discussed in this work has been to determine early and reliable warnings of events and other activities that may lead to threats against the infrastructure nodes in focus. To solve this problem a system for surveillance of primarily people and vehicles that temporary reside around the objects to be protected is developed. Other objects of interest for registration are e.g. cameras, binoculars, weapons and tools. Thus a surveillance system must have capabilities for registering not only the objects of interest but also for determination of whether the activities performed correspond to potential threats to

the infrastructure nodes subject to protection. To achieve this, the following general systems requirements must be full-filled:
- Automatic collection and analysis of information from sensors and other intelligence sources.
- Handling and storage of large heterogeneous data volumes generated by the data sources.
- Compilation and maintenance of operational pictures to improve the users' situational awareness.
- Support for an efficient user dialogue.
- Means for adequate decision making.

As a consequence, three main system components were identified. A sensor system positioned at each node in the infrastructure, an information system at the operation central and to connect them a communication network. The latter has not been of any concern in this work. This structure is illustrated in Figure 1. The sensor system can handle a large number of different sensors and it has capacity to automatically carry out all necessary sensor data analysis and to transfer the resulting information via the network to the information system at the operation central. The operator working at the operation central reacts upon every alarm that is sent from any of the nodes by the system.

3. EVENT IDENTIFICATION

An event can be seen as a description of an activity carried out by an actor at a certain location during a period in time. The activity is generally homogeneous and is therefore said to be elementary. A set of elementary events that occur in a time sequences are here called compound events.

Identification of events which can lead to alarms and for which actions must be taken are necessary to determine early in the development process as this support identification of the sensors to be used. For this reason, a number of possible events have been identified for determination of useful sensors.

A further aspect to consider in relation to these possible events is to view them as elementary parts of more complex events. For example, a person can arrive at the area in his car and then take a walk towards the node, then walk back and drive away again. Obviously, this combined event can be split into arriving in the car, walk towards the object, walk back to the car and finally drive away. This method was used here to determine which sensors that may be used for the different events. However, due to limitations in space just a few of these event descriptions are presented here.

The area outside the nodes is split into three different subareas, i.e. 1) the area outside the fence (a 10 m wide zone), 2) connecting roads and the zone nearby them and 3) distant areas (10 - 100 m outside the fence). The fist example given here is concerned with the area close to the fence while the second example is concerned with the distant area.

Example 1:
Event I: A person is walking around the fence.
Risk: The person is preparing a break-in or sabotage.
Solution: A person's presence can be detected with a Doppler-radar or with a video camera. The person's movement can be followed by cameras of visual or IR types depending on whether day or night capacity is required. Magnet sensor can be used to detect movements of persons and vehicles. Geophones can be used for detection of walking/running people.

Signal and image processing methods are required for determination of people, their movements and tracks. By using image generating sensors the length of the person can also be determined as well as whether it is a human or an animal. With high resolution sensors even arm movements can be determined, which can be used for identification of a person. By using situation analysis the person's movement can also be analyzed with respect to the establishment and the time of the day.
Example of sensors: Doppler-radar, cameras for the visual or for the IR spectra, magnet sensors and geophones.

Example 2:
Event II: A person is observing the establishment for a long period.
Risk: The person is preparing a break-in or sabotage.
Solution: Sensors to be used and the signal or image analysis methods are similar to those given in event I. Laser-radar can be used to penetrate the parts of the terrain. Change detection with laser-radar is also possible. Furthermore the size of the person can be determined as well with the laser-radar as well as the position of the person in question. In the situation analysis sensor information can be combined with other stored information to determine whether the situation is normal or not.
Example of sensors: Doppler-radar, cameras for the visual or the IR spectra, magnet sensors, geophones and laser sensors (gated viewing or laser-radars).

As many elementary methods are of combined type the system must be able to split such combined events into elementary events. These parts of the system are always carried out by the system at the different infrastructure nodes

4. SENSORS FOR INTELLIGENT SURVEILLANCE

Sensors for intelligent surveillance must not be too expensive to buy or operate. Many systems that are available in a military context are too expensive for a civilian customer. The willingness to procure systems for security is governed by regulations and limited to the minimal level that passes these regulations. Still, there are many sensors that can be considered for a system, depending on the type of critical infrastructure. CCD-cameras are a natural component, but cannot always be relied on. Doppler radars, microphones, magnetometers, geophones, IR-cameras and lasers are other potential choices.

A robust configuration of sensors is a combination of Doppler radar, TV- and IR-cameras, and acoustic sensors. The primary role of the Doppler radar is to detect and localize moving targets at distances up to to a distance of 1-2 kilometers, which is sufficient for this application. A Doppler radar can also be used for classification of the targets using the distinct motion characteristics of different types of targets. The acoustic sensors are assembled in arrays, which makes detection, tracking and classification of vehicles up to 1-2 kilometers possible. However, the sensors are sensitive to environmental disturbances, e.g. wind. CCD- and IR-cameras are used for tracking and classification of the targets. Although lighting can be used in the vicinity of the infrastructure, IR-cameras must be used at longer distances. Furthermore, the cameras must be directed to potential targets to be useful at long distances.

5. DATA ANALYSIS FOR INTELLIGENT SURVEILLANCE

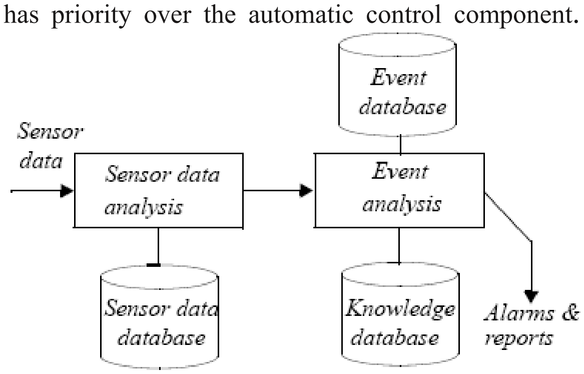
The data analysis is divided into four different processes. There is the sensor data analysis, the event analysis, the situation analysis and the impact analysis. The role of the processing will be described below. However, all levels must contend with uncertainty in their object representations, models, etc. An important capacity of the system is to be able to propagate, aggregate and visualize uncertainties in the system. The sensor data analysis will work in a probabilistic framework. These uncertainty descriptions are translated into a multi-hypothesis, qualitative framework for the operator in the event, situation and impact analysis.

5.1 Sensor data analysis

The sensor data analysis is an automatic process performed in two steps; one for each individual sensor and one for the joint set of sensors. The purpose is to generate descriptions, i.e. tracks, of the objects of interest, in this case humans and vehicles. The goal is to create a joint and maximally accurate description for all sensor systems. It concerns the traditional data fusion subjects of determining the objects position, movement, class and attributes. The algorithms involved are algorithms for tracking, association and classification.

The capacities required is to: 1) Detect and localize humans and vehicles in data from all sensors. 2) Track the movement of humans and vehicles in data from single sensors and in data from multiple sensors. 3) Determine relevant attributes for the objects, e.g. their color and size. 4) Associate objects from different sensors or times. Every object must be given a unique identifier by the system. 5) Fuse the information from different sensors. 6) Separate adults, children and animals into different classes.

A system for intelligent surveillance must also contain components for automatic sensor control, besides the direct control issued by users. Automatic control is used to determine the distribution of tasks in the joint sensor system. The tasks of the sensor system include search for new objects, tracking of known objects and classification of detected objects. Most importantly is determination of the time distribution allocated for sensors in searching for new objects and the tracking of known objects. The user has priority over the automatic control component.



.Figure 2. The sensor data analysis module followed by the event analysis module.

5.2 Event analysis

Event analysis is a process that is performed to provide an operator with a high-level description of the activity in the vicinity of the infrastructure, i.e. event descriptions. It is a processing step that follows the sensor data analysis and cannot always be performed fully automatically. An event is a description of the activity, the actor, the location and the time interval for the activity. The actor is identified by a class descriptor, a length estimate, characteristic visual features from the cameras, e.g. a histogram describing the objects composition of colors and the unique system identifier. Location concerns both a quantitative component, i.e. a coordinate, and a qualitative component, i.e. a particular part of the facility or its environment, e.g. "outside the fence", "at the gate". Activity also concerns a quantitative component, i.e. velocity and motion of extremities, and a qualitative component, i.e. the movement and behavior of the object in relation to the facility, e.g. "moving around the facility", "entering the facility". New events are instantiated when a new object enters the surveillance volume or when an object changes its location or activity qualitatively.

The ultimate purpose of the event analysis is to determine, which events that are important to notice. The event analysis has the capacity to determine: 1) the objects location in relation to the facility. 2) The movement of the objects in relation to the facility. 3) The movement and behavior over longer periods of time. 4) Which objects that are connected, e.g. which people that belong to a certain vehicle, which people that belong to an observed group. 5) Which events that occurs. 6) Which events, or composite events, that are cause for an alarm/warning. Other information than sensor data are handled in the event analysis process. Firstly, this is information regarding the geometry and geography of the infrastructure. Secondly, this regards information about normal activity at different times and at different locations at the infrastructure.

Event analysis requires algorithms for association. When considering longer periods of time, e.g. hours, this capacity is required for the event analysis. Algorithms for aggregation of tracks, events and for anomaly detection, are required [6], [11] and [10].

Event analysis is a fundamental pre-requisite for intelligent surveillance. It is particularly important when surveillance is performed in non-restricted areas where people and vehicles move freely. Many surveillance systems today are built using a simple strategy for warnings where the system warns the user when an object is within the sensor field-of-view. Such a strategy is insufficient when people can have legitimate reasons for being in the coverage of the sensors. In this case the sensors must determine the activity and class of the object before issuing a warning. To be able to do that, the object must be tracked for an extended period of time and multiple sensors are often required. It is essential that the system can react differently in parts of the sensors coverage, typically outside or inside a fence.

6. THE SYSTEM ARCHITECTURE

The system architecture proposed here, is based on the three concepts services, views and roles. These three concepts have a strong impact on both the structure of the system and the usage of the system.

6.1 The service concept

The concept of service is defined differently in different domains. The service concept used in this paper is based on the idea proposed in service oriented architecture (SOA) [8], which emphasis to create systems with more lose couplings between the different software components. In SOA a service is seen as a unit of work done by a service provider to achieve desired end-results for a service consumer. Thus services are independent of how they are implemented but they need to include explanations on how to make use of them, vital qualities, and what is produced, i.e., the effect of making use of the service. The service can thus be seen as an interface between the producers and the consumers. Further, an additional implication is that users not need to have knowledge about how and by whom services are produced. Services are searched for and considered based on how a specific realization meets the qualities of the services.

In the proposed system the consumers correspond to the operators at the operation central and the providers are the sensor system at the infrastructure nodes or some other component carrying out specified functions for information collection or manipulation. A service can in this context be seen as an function, which the surveillance systems performs as a result of a specific call, which can be made by a user or by other services, manually or automatically.

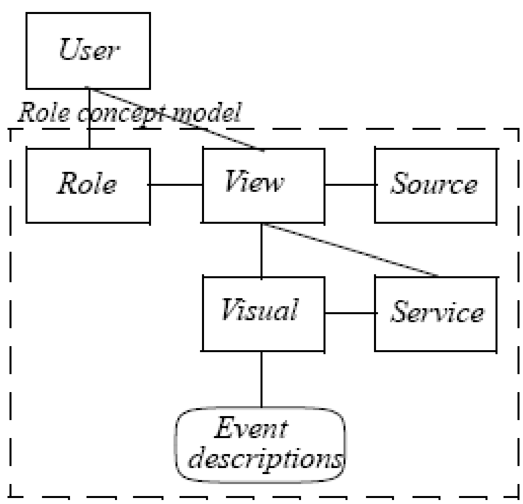


Figure 3. The role concept model.

6.2 The role concept model

The system architecture includes a hierarchical structure, called the role concept model, which also can be perceived as an ontology. The basic concepts of the role concept model are: (1) views, (2) services, (3) roles, and (4) their relationships. Primarily the model is developed to provide for mission support in the command and control process (Figure 3). The model illustrates how users relate to their role and views in the surveillance process. Furthermore, on the next lower level the relationships between the views, the services and the visuals are illustrated. A visual corresponds to a visualization of a view instance. Most relations in the model are of many to many type. For instance, a user may use several views and a view may be used by several users. Thus, a single view may be used by more than one role, i.e., a certain role can be associated with a set of views including their visuals and services. Users may not have any direct link to the data sources. However, in reality such information must be linked from a data source via a view by the users. This linking is due to the need for verification of all kinds of data content, e.g., to under certain weather conditions determine whether an observed object is human or not.

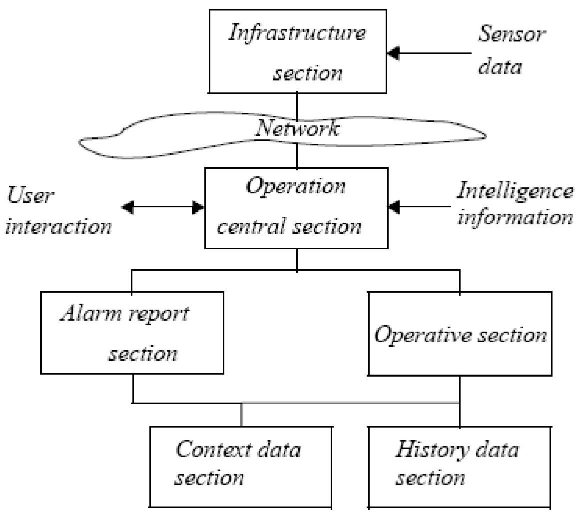


Figure 4. he architecture of the surveillance system.

The role concept model illustrates not only the relations between the roles, views and services but also their relations with the data sources. For each role, its individual operational picture is supported by means of the services through the set of related views, via the visuals and down to the event descriptions, where the operational information resides. This demonstrates how the situational awareness, required in command and control processes, can be improved by means of applications of the services through which the event descriptions of the on-going crisis are created. The event descriptions contain from this perspective data and information about the crisis.

Two main roles have been identified for the given application, i.e., the operator role and the analyst. The operator role is designed for the surveillance of the infrastructure nodes while the analyst's task is to analyze events that have occurred over certain and sometimes longer periods in time.

6.3 System architecture

The system architecture contains six different parts called sections where each section can be seen as a service and where each section generally contains at least a single view with its subordinate services (Figure 4). Of these six sections, five reside in the operation center system while an instance of the sixth section is present in each infrastructure node, i.e., the Infrastructure section, which includes the sensor systems and the capabilities for sensor management, sensor data and event analysis. The root of the sections in the operation central system, called the Operation Central section, includes the basic means for user interaction but also for reception of intelligence information and the information from the Infrastructure sections. This section does not contain any views but includes services for calling the Alarm report and the Operative sections.

The Alarm report section includes a view containing a list of observed events where some may be alarms while others may be of a lower grade of warnings to which the operator may decide to react or not. The Operative section is the most complicated and will be discussed further more deeply in section 6.4. On the lowest level of the architecture reside the Context data and History data sections, which include views supportive functionality to the views in the Alarm report and the Operative sections. These sections are not directly accessible from the Operation central section. Instead they are called by the alarm and operative sections. The purpose of the Context data section is to supply the users with necessary information about the environments, e.g., map information. The History data section includes the History data view composed by a repository for all created view instances. The information in this view is used by the analyst to determine whether any threats are emerging at the node.

6.4 The system sections

The operative section (Figure 5) is equipped with three views, i.e., the Operative view containing the current situational or operational descriptions of an event subject to an alarm; this corresponds to the actual operational picture described in relation to its context. Attached to this view are the Sensor data view and the Resource data view. The Sensor data view is essentially intended for direct visualization of sensor data, e.g., data from a video. The Resource data view is intended for presentation of information of concern during an event and may include information about police or FD and information about what to do under certain conditions [2].

Two decision support tools are integrated as well, i.e. a view query system, of dynamic type [1], and an information fusion system for situation and consequence analysis.

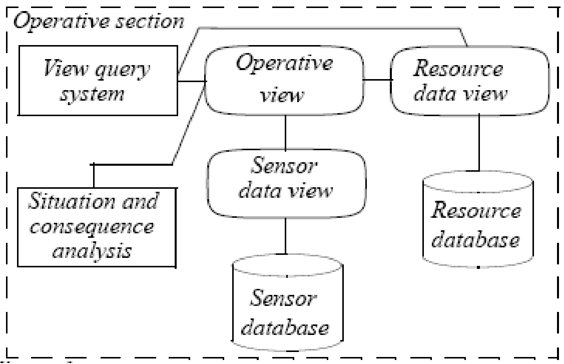


Figure 5. The operative section with its views and decision support tools.

Beside these three views, two decision support tools are integrated as well, i.e. a view query system and an information fusion system for situation and consequence analysis. The query language should be of dynamic type [1].

Figure 6 illustrates the functional links between the Operational central section and the Infrastructure sections. The purpose of these two types of associations is to provide support for the management of the sensors at the infrastructure nodes and for the acquisition of data from the sensor systems at the nodes, which correspond to event information although sensor data can be transferred as well.

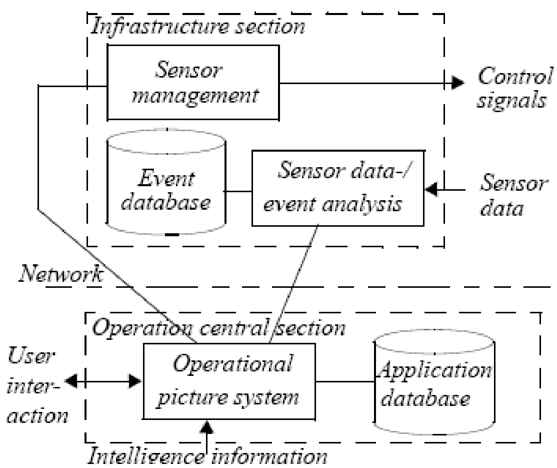


Figure 6. The infrastructure and the operation central sections and their connections.

6.5 View services

Included to each view is a set of services. Some of these services represent user commands in ordinary information systems. Since, it is more convenient for the users if the interaction with the system is homogeneous: Hence, commands like operations will also be handled as services. Furthermore, three classes of services are available: (1) automatic services activated directly by the system due to, e.g., an alarm, services available to (2) the operators and (3) to the analyst. Some of services may be found in more than one class. Generally, the users interact with just a single view at a time, i.e., the current view. To interact with some other view users must switch views by means of a service call.

As the number of services is fairly large just a few examples from the operative view are presented here:

Get attribute values from a particular object (carried out by means of the query tool).

Change current view (to a new arbitrary view).

Show earlier view instances (from History view).

6.6 Situation analysis

The purpose of the situation analysis is to support the users to in the assessment of the joint situation at all the nodes of the infrastructure. The analysis is based on events that have occurred at the facilities, together with relevant intelligence and the system knowledge of the relation between events and situations. A situation is in this case a higher level description of a composition of events, i.e., regarding surveillance and trespass towards facilities. It differs from the event analysis as it is performed only when the users queries for a situation and not continuously. The longer time perspective (days, weeks, months) is an additional difference. Analysis regarding changes in the pattern of events involving the entire infrastructure may lead to other conclusions than when only concerned with a certain facility.

It is difficult to determine the situation solely on the basis of collected sensor data. In many cases it will be the intelligence data and knowledge about the characteristics of the infrastructure that permit conclusions to be drawn. Sources of intelligence data can be police as well as employees or others that have observed something of interest to the infrastructure. The intelligence data must be assembled and integrated into the system by the users. Intelligence reports can potentially contain many different types of information and can be structured in different ways. The situation analysis is restricted to handle structured information of two types. First, intelligence reports may contain information about events in the same manner as determined by the event analysis, including estimates of certainties. Second, the intelligence may consider general information about the threat against a certain facility or the public that use the facility. The situation analysis will not be able to analyze other types of intelligence automatically. It is the system knowledge about the relation between actors, facilities and activities that determine how a situation should be interpreted. The system concept does not include algorithms for automatic learning. System knowledge must be elicited from experts.

The capacities needed for the situation analysis are similar to those needed for event analysis regarding detection of anomalies and aggregation of events. However, as the scope of the analysis is different data mining algorithms are needed. Furthermore, the situation analysis must be integrated with the query system.

6.7 Impact analysis

The impact analysis must support the operators in the assessment of the consequences of different situations. Depending on the level of ambition, this can be trivial or extremely difficult to accomplish. If the analysis focus on the facility, the consequences of the situation can be obvious to operators once a situation is recognized. The analysis must be different if the focus is on the impact for people using the facility, its employees or if a broader societal impact must be determined. Impact analysis services can for instance be used to determine: (1) How many people and how large an area that is influenced of a facility failure of some kind. (2) If there is something of particular value at the facility that can be stolen or destroyed. (3) The capacity and attitude of actors in the area of a facility. (4) How long time it may take to replace or repair damaged parts of a facility.

Impact analysis is performed on the users' requests. The relevant information resulting from the impact analysis will automatically be made available to the users once a query for such analysis is issued. This will be provided through documents that are connected to situations, facilities and actors. With this information available, the user must be able to determine the consequences of the situation. The impact analysis will not give suggestions of what and how to act in a certain situation. This is contained in the resource view in the operative section of the system

6. SCENARIOS

This section presents an example of a set of scenarios and the interpretations made by the intelligence surveillance system. The scenario describes antagonistic observations made at a facility. A car arrives at the facility at two different occasions. On the first occasion nobody leaves the car. The next day, the car returns and a person steps out of the car and walks towards the gate of the facility. The car is stolen. The system warns the operator that there is potential threat against the facility. In Table 1, a breakdown of the events that the system registers and a description of how the system reacts when the events occur is given. There are three ways that the system handles events. A0 - information about the event is stored locally but no message is sent to the operator. A1 - a message about the event is sent to the operator as information. A2 - an alarm is sent to the operator.

Table 1: Scenario example Part 1, Time: 23:46, Date: 23 November.

Step	Event	System action
A blue Volvo is approaching the facility, via the access road.	Z.vehicle approach facility **via the access road**.	A0
The Volvo stops in front of the facility gate and the engine is turned off. Nobody leaves the car.	Z.vehicle stop **at gate**.	A1
A few minutes later, the car engine is turn on again.	Z.vehicle start **at gate**.	A0
The car turns around and leaves the facility.	Z.vehicle depart facility **via the access road.**	A0

In the tables, object identifiers are printed with capital letters with a.class suffix to indicate class label, actions are underlined and locations are printed in bold. The time intervals of events, in the scenario, are excluded as well as the characteristics needed for recognition of objects.

There are no reasons for cars to be at the facility at night time. Therefore, the system notifies this event to the operator. When the car turns up again the following day, the car is recognized by the system as the same as the day before. This is also notified to the operator. At this time, the registration number is caught by a camera. The car turns out to be stolen when investigated by the analyst.

7. CONCLUSIONS

This paper presents a system and its architecture for intelligent surveillance of critical infrastructure nodes. The objective of the work has been to develop a system architecture for determination of antagonistic events, by means of the sensor systems covering the infrastructure nodes, registration and different types of analyses of events going on at these nodes.

Table 2: Scenario example Part 2, Time: 23:12, Date: 24 November.

The same blue Volvo is approaching the facility via the access road.	Y.vehicle approach facility via the access road.	A0
The car stops at the gate. This time, the registration number is visible to one of the cameras.	Y.vehicle stop at gate.	A1
A person leaves the car and approaches the gate.	X.person comes out of Y.vehicle. X.person approach gate.	A0
The person returns to the car and enters it.	X.person approach Y.vehicle. X.person enter Y.vehicle.	A0
A few minutes later, the car is started again.	Y.vehicle start at gate.	A0
The car is turned around and driven away from the facility.	Y.vehicle depart facility on approach road.	A0
The analyst queries the system about events at the facility because of the notified events. As the registration number is noticed it is checked with the police.		

The surveillance system is based on the principles of (1) services, (2) views and (3) user roles. Furthermore, the architecture is based on the role concept model, which in a logical way demonstrates the relationships between not only the three basic concepts but also with the data sources. The model can be seen as a foundation for a system for management of operational pictures where the views support both visualization of operational information and application of a fairly large number of services for maintenance of the operational picture.

ACKNOWLEDGEMENT

This work was made possible due to a contract with Swedish National Grid (Svenska Kraftnät).

REFERENCES

[1] Ahlberg C., Williamson C. and Shneiderman B. (1992), *Dynamic queries for information exploration: an implementation and evaluation*, in 'Proc. of Conf on Hum Fact in Comp Systems (CHI 92)', ACM Press, pp. 619-626.

[2] Chang, S.-K., Jungert, E., *A Self-Organizing Approach to Mission Initialization and Control in Emergency Management*, Proceedings of the Int Conf on Distributed Multi-media Systems, San Francisco, CA, Sept 6-8, 2007.

[3] M. L. Garcia, *The design and Evaluation of Physical Protection Systems*, Elsevier, Boston, 2001

[4] Hu J. and Grefen P. (2003), *Conceptual framework and architecture for service mediating workflow management*, Information & Software Techn., vol. 45, iss 13, pp. 929.

[5] Hu, W., Tan, T., Wang, L., Maybank, S., A Survey on Visual Surveillance of Object Motion and Behaviors, IEEE Transactions on Systems, Man and Cybernetics, Vol. 34, No., 3, pp. 334-352, Aug. 2004.

[6] Ivanov, Y., A., Bobick, A., F., Recognition of Visual Activities and Interactions by Stochastic Parsing, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 8, pp. 852-872, Aug. 2000.

[7] MacKenzie, C. M., Laskey, K., McCabe, F., Brown, P. F., Metz, R., and Hamilton, B. A (eds.) *Reference Model for Service Oriented Architecture 1.0* Committee Spec Aug 1-2, 2006 http:// www.oasis-open.org/committees/tc_home.php?wg_abbrev=soa-rm (2008-01-24)

[8] Rodosek G.D. (2003), A generic model for IT services and service management, IFIP/IEEE Eighth International Symposium on Integrated Netw Management, pp. 171-184.

[9] Scalingi, P.L., *DOE energy industry program on critical infrastructure protection*, Power Eng Socy Winter Meeting,. IEEE Vol 1, 28 Jan.-1 Feb. 2001 pp 104 - 107

[10] Shi, Y., Huang, Y., Minnen, D., Bobick, A., Essa, I., *Propagation Networks for Recognition of Partially Ordered Sequential Action*, Proceedings of the IEEE Comp Vis and Pattern Recogn (CVPR 2004), Washington, 2004.

[11] Stauffer, C., Grimson, W., *Learning patterns of activity using real-time tracking*, IEEE Transactions on PAMI, Vol. 22, No. 8, pp. 747-757, Aug. 2000.

[12] Trivedi, M.M. Gandhi, T.L. Huang, K.S., *Distributed interactive video arrays for event capture and enhanced situational awareness,* Intelligent Systems, Sept.-Oct. 2005, Vol 20, No 5, pp 58- 66.

[13] Wolthusen, S.D., *GIS-based command and control infrastructure for critical infrastructure protection*, Critical Infrastr Protect, 1st IEEE Intl Workshop, Nov 3-4. 2005.

Emergency Control in Shipyard Safety Systems*

Augusto Celentano, Fabio Furlan
Università Ca' Foscari Venezia
auce@dsi.unive.it, fabiofurlan@iol.it

Stefano Burcovich
Fincantieri SpA
stefano.burcovich@fincantieri.it

Abstract

Large shipyards for gigantic cruise ships require a special attention to safety due to the huge number of workers involved and to the complex structure of spaces. A critical issue of such environments is to keep the environment under control to avoid a disordered evacuation in case of an emergency. After introducing the basic issues related to safety in shipyards, we discuss the design of an information system for checking the shipyard safety status, supporting synthetic information visualization techniques about escape routes and dynamic evacuation plans. We discuss also the problems of communicating with workers in case of evacuation using visual cues to signal the correct escape routes.

1. Introduction

Large shipyards for gigantic cruise ships require a special attention to safety due to the huge number of workers involved, the complex structure of spaces, the large range of activities performed inside, and the spread in people ability to react to adverse situations. One of the most critical problems to be faced in case of an adverse event is the evacuation of workers. Any satisfactory solution to this problem is based on the possibility to drive the evacuation in an ordered way. Conversely, the most critical situation is to face a disordered evacuation, defined as a type of evacuation which adds damages and injuries to people beyond those due to the disaster.

At the Computer Science Department of Ca' Foscari University in Venice, Italy, we have developed a study about shipyard safety systems with Fincantieri SpA, the largest Italian shipbuilding company. The study aimed to design: (a) a comprehensive information system for monitoring the shipyard in normal and emergency situations; (b) a set of visual interfaces helping persons in charge of safety to get a synthesis of the escape routes situation; (c) algorithms for computing dynamic evacuation plans; (d) personal devices and visual cues to drive workers to safe places in case of emergency. From the perspective of the emergency responsible, the main goal of a safety system is to overcome the limits of current practices, based on static evacuation plans and large signboards, whose size and content is defined by a set of official regulations, but that do not adapt well to the shipyard dynamics. During the building of an environment wide and complex like a cruise ship plans must be changed dynamically with the evolution of the environment itself and the degree of completion.

In this paper we focus on the issues related to the design and management of information about the ship status during its building, and to the the visual representation of critical situations that could lead to emergency problems.

2. Regulations about shipyard safety

The prevention of emergency situation is ruled in Italy by Law 626 [1], which has introduced professional roles devoted to safety control and environmental requirements targeted to risk protection. Among the duties of the emergency responsibles, the most important are risk evaluation, the adoption of suitable measures to protect the yard and the people working in it, and the continuous monitoring of the work places. To achieve their goal, emergency responsibles and emergency workers must be supported by information management systems specifically targeted to the yard environment.

In case of emergency, Law 626 defines a number of concepts and procedures targeted to evacuate the workers by driving them through safe escape paths under any adverse condition. To this end, qualitative as well as quantitative parameters are associated to the escape paths and must be monitored to assure a continuous and up-to-date assistance to the workers in danger. The parameters describe features such as the level of risk of the places, the number of different escape routes, the width and capacity of the doors, passages and stairs (called *checkpoints*, their features are discussed in Section 5); they also constrain the presence of oc-

[0]Part of the content of this paper has been presented in a preliminary form at the *HCI for Emergency* workshop, in conjunction with *CHI 2008 Conference*, Florence, Italy. This work has been supported by Fincantieri SpA.

clusions due to work progress and define the size and shape of the signboards used to signal the escape routes. Such regulations apply during ship building in order to guarantee that suitable escape routes exist from any place, sized on the presence of workers. At the same time, they allow a safety responsible to check if some limit is trespassed, providing the means for restoring a safe configuration.

Law 626 applies to every type of building. As a shipyard presents specific problems due to the pervasive use of metal as building material, and to the high risk of fire caused by fuel operated tools, a set of international regulations also apply, called SOLAS (*Safety Of Life At Sea*) [2]. SOLAS regulations define primarily parameters related to fire resistance, which evolve as a deck environment, initially empty, is progressively populated with halls, cabins, rooms, replacing a large open space with small closed ambients. SOLAS rules therefore progressively overlap the rules defined by Law 626 as building progresses. This situation, if not properly driven, may cause conflicts rather then reinforce safety, also because the risk evaluation made at design time must be continuously compared with the actual work progress.

3. Safety constraints

Besides architectural issues deriving from the frequent changes in the spatial configuration, three classes of constraints are relevant to safety.

Technical constraints. Deploying communication and sensing infrastructures is difficult due to the dynamics of the environment. Large spaces being progressively split into small closed environments, cables and sensors cannot be placed freely; moreover, the large amount of metal used limits radio communication.

Human constraints. Workers in a shipyard are very heterogeneous, coming from many countries, with different base cultures and different skills. As many as 70 different languages have been counted, English being often unknown; emergency messages might be misunderstood, or not understood at all.

Organization constraints. Changes in the environment are recorded with variable delays, thus leading to inconsistency between the ship state known by the persons in charge of safety and the real state, possibly suggesting emergency actions that refer to an outdated ship configuration. The most critical concern is that workers often do not know the real danger level to which they can be exposed, possibly taking wrong decisions in case of emergency.

Figure 1 shows four views of a shipyard illustrating typical cases that can affect safety. The four pictures refer to normal, allowed situations, not to exceptional or incorrect situations. The top left picture shows one of the staircases external to the ship hull, giving access to the decks. The stairs are not connected to every deck, resulting in more articulated routes to escape on ground. The top right picture shows a deck hosting the passenger cabins before walls' installation: the environment is easy to walk, but lacks precise reference points. The bottom left picture shows the same deck after cabins' mounting: places and routes are clearly identifiable, but the narrow passage can be difficult to traverse in emergency due to partial occlusion caused by carts and building materials. Finally, the bottom right picture shows a staircase passage closed by a safety ribbon, hence not usable.

Figure 1. Views of a shipyard

4. Evacuation management

The main goals of a correct evacuation are to be fast and ordered, to give correct information to people in danger, and to keep control over the environment. They are achieved by a correct mix of prevention and intervention. Prevention mainly requires the analysis of the potential sources of emergency problems, through a network of sensors monitoring the state of the ship [3]. Intervention requires fast delivering of correct information, as well as the limitation of damage propagation with physical and organizational means.

Therefore, an evacuation management system operates in two distinct phases: observation time and emergency time.

Observation time. During the normal work information about the ship state is collected and processed, generating knowledge about the workers flows and the rooms occupation. This phase prepares the system to plan an evacuation in case of emergency.

Emergency time. The system delivers reliable information about how to leave the ship in safety, providing both on-

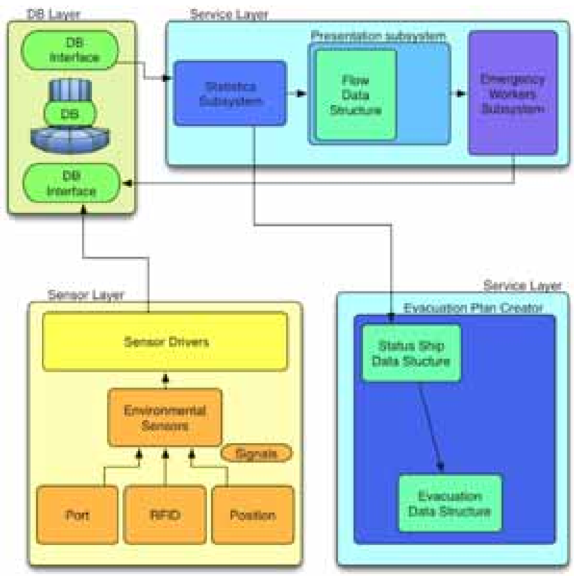


Figure 2. The emergency management system architecture

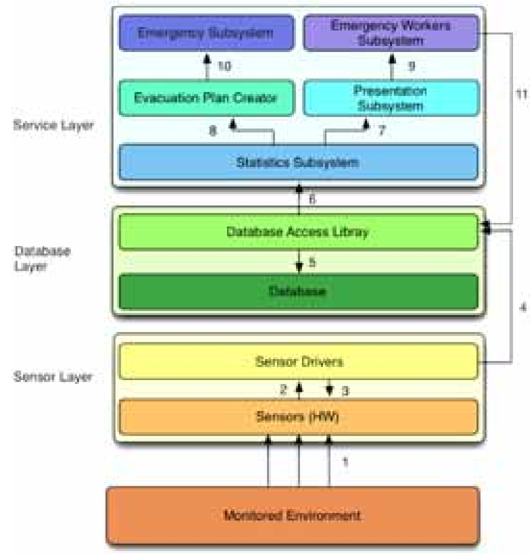


Figure 3. Information flow among the system components

board emergency workers and rescue teams with detailed information on the state of the ship, on the evacuation progress and on the number of saved people.

Both phases rely on the management and the delivery of critical information, hence depend on how the information is presented, since an emergency situation benefits from pre-attentive stimuli and unbiased signals universally interpreted.

5. The emergency management system

A ship is spatially organized in a three level hierarchy [4]: *decks*, which mark a vertical decomposition, *main vertical zones* (MVZ) corresponding to the longitudinal watertight compartments, and *rooms*, like halls and cabins, which are the areas in a MVZ which can be traversed in case of an emergency, if free from temporary occlusions. Locations in a ship are thus identified by three coordinates related to the deck, the longitudinal frame (counted onward and backward from the rudder), and the ship side (left, center, right).

Two other elements are of primary importance for monitoring the environment and guiding the workers in case of emergency: the *checkpoints* and the *well known points*. The simplified concept of communication passage between two environments used in buildings, e.g., a door, is unsuitable in a shipyard, due to the large number of types: doors, openings, stairs, steps, etc.. The concept of checkpoint (CKP) is introduced to define any passage in the environment through which the number of persons crossing at the same time is limited and can be measured. Well known points (WKP) are environments with special architectural and functional clues like theaters, halls, restaurants, lift lobbies, suites, shops, panoramic walks etc., that can easily identified by everyone onboard.

A database is built around such a ship structure. It has a fixed content, describing the ship structure, and a variable content defined by data collected from sensors monitoring the ship areas during building, which change according to the actual ship state.

5.1. The system architecture

The emergency management system architecture is shown in Figure 2. The *sensor layer* manages the hardware and software (drivers) infrastructure needed to store in the database the information on the ship status.

The *statistics subsystem* is part of the *service layer* and manages a data structure containing information about workers occupation and flow through the different ship rooms. Raw data are read from the database and processed to inizialize the statistics subsystem. The *presentation subsystem* uses information from the statistics subsystem and from sensors to create visual maps about the ship and workers situation.

The *evacuation plan creator* computes the best escape ways, storing the information in a data structure that will be used by the *emergency workers subsystem*.

In case of emergency, the emergency workers subsystem uses the data provided by the evacuation plan creator to signal the best escape ways through the environment devices

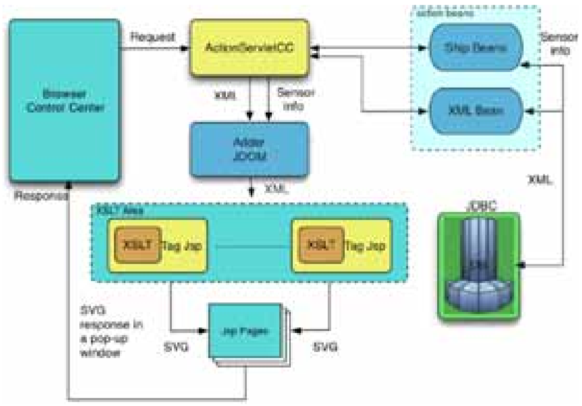


Figure 4. The control center

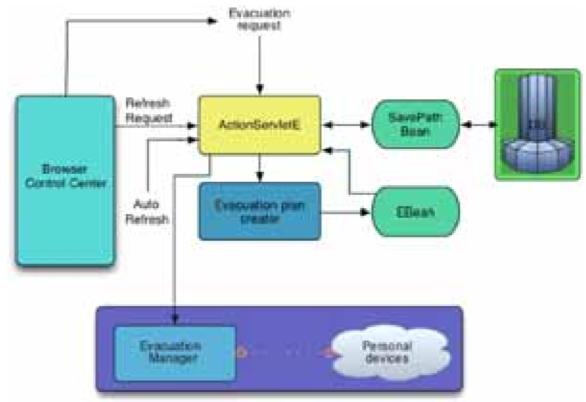


Figure 5. The evacuation plan creation

located in proper places along the escape routes, and possibly through the portable devices worn by the ship workers.

5.2. The information flow

Figure 3 illustrates the information flow between the sensor, database and service layers. The numbers in the following description match the numeric labels on the arcs of the figure. (1) The sensors monitor the workers' location inside the environment. (2) Signals coming from the sensors are processed and converted into usable information; (3) sensors can be configured according to the progression of the building. (4) Sensors send updated information to the database. (5) The database is constantly updated to reflect the current ship situation both in normal situation and in an emergency situation, concerning ship workers, onboard emergency workers and rescue teams involved. (6) The statistics subsystem creates and stores the data structures containing raw information related to people's flow, overcrowding and ship status. (7) The raw statistic data are combined and interpreted to provide meaningful and reliable information to the onboard emergency workers. (8) Collected statistic data are also used to create an evacuation plan, dynamically updated as the workers flow changes. (9) Visual information is presented on the personal devices of the onboard emergency workers. (10) If the ship must be evacuated, visual information about the safe escape routes is sent to the emergency subsystem that manages the emergency panels and the workers personal displays. (11) Feedback is sent to the emergency database in order to identify critical points such as persisting overcrowding, occlusions and changes in the escape routes due to the emergency evolution.

5.3. The emergency control center

Figure 4 shows the control center operations. The emergency responsible queries the system about the status of a specific ship area. A control servlet receives the request and activates two components that manage the occupation indices of the ship areas (*Ship bean* component), as described in Section 6, and the ship planimetry (*XML bean* component) expressed in a XML dialect and stored in the DB. An *Adder* component superimposes the two information by modifying the ship planimetry to add visual information about the current workers occupation and flow [5]. The result is translated to SVG graphics, visualized in a browser window in the control center.

Emergency plans are maintained starting from initial plans defined according to the official safety regulations. They are periodically updated as the ship building proceeds and the occupation of the shipyard areas evolves. Figure 5 shows a fragment of the system devoted to plan update. Information about the current ship status is extracted from the DB and passed to an emergency plan creator, an algorithm that builds a network of connections among nodes representing the locations that can be traversed in emergency. The arcs are labeled with cost values that are function of several variables: the most important are the properties of material as defined by the SOLAS regulations, biased by the knowledge about the current emergency dynamics and the people flow. Due to space limitations we do not elaborate on this algorithm, but it's worth to note that, since it must be fast and it must converge to a solution, euristics are adopted to speed up the computation, even if they do not provide optimal solutions.

6. Checking occupation and people flow

Two synthetic features are able to anticipate critical situations if kept under continuous control: *congestion index* and *flow index*, respectively measuring the ratio between people and environment size and capacity (showing overcrowding), and between people flowing through passages and passage capacity (showing critical situations due to es-

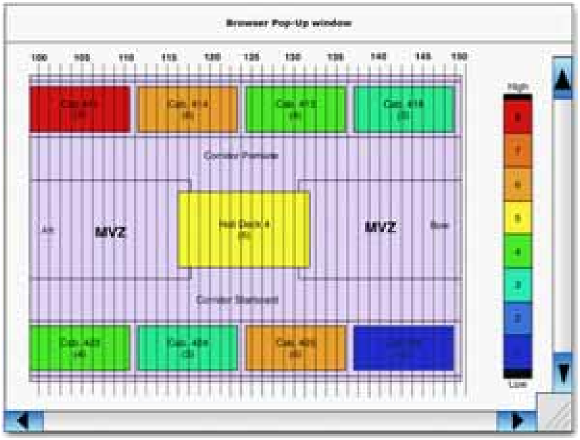


Figure 6. The interface for occupation control

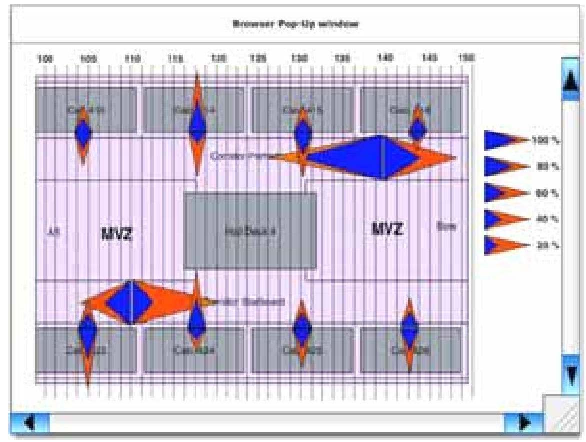


Figure 7. The interface for people flow control

caping persons congestion). In most cases the capacity of environments and passages is fixed by safety regulations, but the presence of temporary occlusions due to building material and the reaction of persons to an emergency situation might reduce the capacity and cause the limits to be surpassed.

To display the ship environment visual information processed by the sensor level is overlapped to the ship map. Visualization is thus linked to the ship structure, and may concern a bridge, a MVZ across all the bridges, a MVZ of a single bridge, or the whole environment of a bridge (least level of granularity).

Information about occupation is represented using colors; each color represents a different level of occupation, as defined by a legend on the right side of the display. Warm colors represent high workers concentration, cold colors represent low workers concentration. In Figure 6 the rooms of a MVZ are colored and every room displays the number of persons inside. In such a way both qualitative and quantitative information is displayed, highlighting the real danger level.

The representation of workers flow is based on the checkpoints. Every checkpoint has a theoretical capacity: the number of persons that can cross it in a unit period of time without danger. Work tools and materials laying in the environment can reduce the checkpoint capacity, so we must know which is the real capacity and how much it is reduced with respect to the theoretical capacity. Both must be displayed to identify situations where the current state of an environment or of a checkpoint could increase the risk level and lessen the evacuation speed. A similar concept is the comparison between the current persons flow and the supported flow as defined by the actual capacity, to identify situations where the dynamics of workers moving could create bottlenecks.

These two pairs of measures are represented using *flames*, two overlapped triangles of different color: blue for the theoretical capacity, red for the real capacity (Figure 7). In a similar way, green and red flames visualize the ratio between the actual flow and the real capacity. Flames are drawn on the ship map, aligned with the checkpoints. The triangle vertex represents the flow direction. The theoretical capacity is represented in the background since it is alway greater or equal to the real capacity, and the overlap between the triangles immediately shows their relationship.

7. Communicating with workers

Most of the information about the ship actual configuration collected during the prevention phase is used to compute and communicate to workers the escape routes in case of emergency. Such information, and the way of presenting it, depends on technical radio communication limits and on some non technical issues.

Workers usually do not know the whole ship but only the part in which they work. Suggesting them a path through unknown ship areas increases the risk; longer escape paths can be safer if they cross only areas to which workers are accustomed, where they can find known visual cues.

Different ship areas are exposed to different risk levels. Normative institutions have issued a classification of danger in different environments. A good escape route crosses areas with decreasing danger level.

Since people work in groups, the evacuation procedure is safer if during the escape the cohesion of the group is maintained. This principle is very important because the ability to help each other is increased by people being used to work together, by speaking the same language, by being used to understand each other, and by being able to integrate their partial knowledge of the environment into a more complete view of the situation.

To signal the escape routes three issues are important: (1) to take care of the changes in the environment due to the building progress, from a skeleton of wide spaces to a complex structure of small rooms, requiring to adapt the granularity and the range of the signals; (2) to differentiate the stimuli used to signal escape routes and wrong paths, using visual signs for positive stimuli and auditory signs for negative stimuli; (3) to avoid the use of text in favor of graphics and symbols independent from specific languages and cultures.

According to these issues, we have proposed two approaches: a weak approach and a strong approach, differentiated by the spatial granularity of signs and information delivery. A strong approach requires positioning a larger number of visual escape signs, and such positioning cannot be done in a highly dynamic and highly incomplete environment. A weak approach is based on fuzzy cues which refer to a few well known locations, easy to identify according to their function. From a general point of view, during the building of a ship the system should evolve from a weak to a strong approach. During a simulation it was evident that due to the little detail present in initial phases of ship building, environments are poorly distinguishable unless they are focal points in the ship (e.g., a hall, which can be easily recognized even if incomplete), or are connection nodes, such as stairs and elevators. Such environments are generally known by all workers, and being a few (with respect to cabins and corridors) are easily identifiable, therefore are marked as well known points.

8. Discussion

The evolution of an event like a fire is not impulsive or instantaneous. It starts in a localized area and may extend potentially to all the ship. But its speed allows people in charge of safety to adopt local strategies and to follow its evolution, starting with the evacuation of a limited number of people close to the fire center, and proceeding to a total evacuation only if the event cannot be controlled. A plausible strategy can be based on three elements: (1) An evacuation signal (visual and/or acoustic) must be forwarded only to people inside the area subject to an immediate risk, through the personal device and the signs in the focal locations. (2) A feedback signal must be received by the people in charge of the emergency management, by monitoring the position of people at risk, checking that they are moving in the right direction. (3) In case the feedback shows immobilized persons, personal devices should be used as beacons to guide the rescue squad.

The work presented here has focused mainly on the design of the information system able to provide timely information for knowing (and managing) the risks due to an emergency, and on the visual aspects of emergency information management. Future work will explore in details the dynamics of emergency and the communication between the control center and the workers. We can anticipate some themes that deserve attention.

Sensor reliability. Sensors measuring the workers presence and flow could be damaged by a fire, creating critical holes in the monitoring of the evacuation process. Damage could prevent data to be read, but could also cause the transmission of altered data. Hence, diagnostic functions must check if the sensors work correctly. Besides using redundancy and transmission channels devoted to check the sensor availability, the shipyard environment is equipped with many types of sensors, among which the fire and smoke detectors are the most reliable due to their primary importance in detecting emergencies. They can be used to circumscribe the areas whose sensing information requires further check in case of a fire.

Signboards, public displays and personal devices. While large electronic displays could be useful for delivering public information during normal operations, their use is not advisable in emergency because they violate two important requirements: (1) since the environment changes continuously, there is a high risk to damage devices devoted to emergency information, whose position must change with the building progress; (2) during an evacuation visual stimuli (and audio stimuli, which are not described here but have been considered in our work) are very simple, e.g., based on colors matching the escape routes, and personalized according the path the worker is following. They are better identified through personal devices and simple color marks, e.g., with laser-like light that can be seen also in presence of smoke.

References

[1] DECRETO LEGISLATIVO 19 Settembre 1994, n. 626. Attuazione delle direttive CEE riguardanti il miglioramento della sicurezza e della salute dei lavoratori sul luogo di lavoro, 1994 (in Italian). http://www.oftal.it/legge626.htm.

[2] International Maritime Organization. *SOLAS, Safety of Life at Sea*. London. UK, 2004.

[3] L.M.Ni, Y.Liu, Yiu Cho Lau, A.P.Patil. *LANDMARC: Indoor Location Sensing Using Active RFID*. Kluwer academic Publisher, 2004.

[4] Holland America Line. Hull.6079, *General Arrangement Plan*. Fincantieri, 2005.

[5] Ben Shneiderman. *Designing the user interface: strategies for effective human-computer interaction*. Addison-Wesley, 2004

A Table-based Lively Interface for Collaborative Music Performance

Takeshi Nakaie[1*], Takayuki Koyama[2], and Masahito Hirakawa[2]
[1]Interdisciplinary Graduate School of Science and Engineering
Shimane University, Japan
[2]Interdisciplinary Faculty of Science and Engineering
Shimane University, Japan
hirakawa@cis.shimane-u.ac.jp

Abstract

Computer music is one of active research fields in information technology in recent years, and systems having powerful facilities for creation of music have been developed. In those systems, however, the user feels difficulty in recognizing performance because of a lack of spatial sound presence especially in a collaborative music performance environment.

In this paper, we present a gesture-based computer music system in which a 16-channel sound table is prepared as its key component. The system captures, for each user, his/her manipulation of a stick-type device in a 3D space over the table and generates sound at the position that is associated with the manipulation. This helps users to have a feeling of their play. In addition, computer-generated graphic images are projected onto the tabletop so that users can interact with the system through both auditory and visual channels.

1. Introduction

It is fun and enjoyable for people to create music as well as listen to music. Towards development of a new musical expression, people have taken various technologies into music creation.

DAW (digital audio workstation) which is a system designed to record, edit and play back digital audio is a pioneering work in this movement. Trials of developing new digital instruments, or instruments which are augmented by digital technologies, for musical expression and artistic performance have also been proposed. Research examples include Hydraulophone [1], Mobile Clavier [2] and Overtone Violin [3], while commercial products are available such as Yamaha's Silent Guitars. However, in those trials, the user is expected to have a good knowledge of music editing or performance.

There have been demands for computer music systems which can help nonprofessionals to create music. Systems which have been proposed so far include Musical Sketch Pads [4], Sibelius [5], and SONASPHERE [6], where music is specified by manipulating graphical objects on the computer screen. While they allow the user to create music with ease and fun, another interesting approach from an advanced user interface point of view is tangible and gestural computer music. Many interesting trials have been investigated so far [7]-[14]. Our proposal in this paper belongs to it.

Meanwhile, most of the systems assume a small number of speakers, even though multiple users play collaboratively with the system. A problem for this is that users have difficulty in associating generated sound with their performance. This causes a lack of realism for collaborative music performance.

We propose in the paper an interactive computer music system providing a lively sound space. Actually the system is equipped with a table in which 16 speakers are placed in a 4 x 4 layout. Sound can be output at any position on the table by properly controlling the loudness for 16 speakers. We call this table Sound Table.

In addition, we provide a stick-type input device as a means of exhibiting user's musical ideas through its manipulation in a 3D space over Sound Table. The system then projects computer-generated graphical images onto Sound Table. Users can interact with the system through both visual and auditory channels.

Rest of the paper is organized as follows. Related works are explained in Section 2. Section 3 presents an overview of the system, and implementation issues of input and output processing are described in Section 4. Section 5 concerns interaction with the system in a music performance application. Finally, in Section 6, we conclude the paper.

* T. Nakaie is now with Dai Nippon Printing Co., Japan.

2. Related Works

There have been futuristic computer music creation tools which do not request the player to have a special skill or knowledge in operating them.

SONASPHERE [6] is such an example. Functional units such as sound samples, effects, and mixers are represented as small spherical objects floating within a 3D space on a computer screen. Music composition is made by connecting objects along a flow of signal. Interestingly, properties of an object are determined with reference to other objects in the space, and composition of music leads to attractive visual motion of objects.

While SONASPHERE is a laptop interface, some proposed computer music systems at which music composition is carried out by physical interaction on a tabletop surface [7]-[11] or through body motion [12]-[14]. We believe the tabletop approach is promising, and will focus on it hereafter.

Sound Rose [7] and the system in [8] are musical table systems with touch sensitive interfaces. The user creates music by touching the table with finger(s).

Block Jam [9], Audiopad [10] and reacTable [11] are tangible tabletop instruments for computer music, where manipulation of plastic blocks on a table corresponds to music composition. While each block in Block Jam has itself a visual display and gestural/clickable inputs, Audiopad and reacTable have a facility of depicting the change of the block position and orientation graphically on the table.

The system we propose in this paper uses a table as a means for interaction as is similar to the above trials. However, a music sequence is specified by the 3D motion of a stick-type device over the table, rather than a touch or an arrangement of music blocks on the table. Furthermore, our system is capable of outputting multiple sounds at different positions on the table. This helps each of the users to sense others' performance, which is essential for collaborative music creation.

Transition Soundings [15] and Orbophone [16] are specialized interfaces using multiple speakers for interactive music making. A large number of speakers are mounted in a wall-shaped board in Transition Soundings, while Orbophone houses multiple speakers in a dodecahedral enclosure. Both systems are deployed for sound art, and their goal is different from ours.

Another related approach of using multi-channel speakers appears in [17]. It is specially designed for CAVE-like multi-screen projection display, and is rather complex and expensive. We use conventional speakers and sound boards, and no specialized hardware is used at all.

3. Collaborative Multimodal System

3.1 System overview

The system we propose in the paper is designed to provide users with effective visual and auditory feedback in creating music collaboratively. Figure 1 shows a physical setup of the system. The system is organized by Sound Table as its central equipment, a pair of cameras, a video projector, and a PC (not shown).

Figure 1. System setup

Sound Table is a table in which 16 speakers are equipped in a 4x4 matrix, as shown in Fig. 2. It is of 90cm width and depth, and 73cm height. The distance between two adjacent speakers is 24cm.

Two 8-channel audio interfaces (M-AUDIO FireWire 410) are equipped to the PC (Apple Mac Pro), and connected to Sound Table through a 16-channel amplifier. The top of Sound Table is covered by a white cloth so that computer-generated graphical images can be projected onto it. The video projector which is mounted at the top of the system is provided for this purpose.

Users create music by swinging a stick-type device over Sound Table. Figure 3 shows the device whose base unit is the Nintendo Wii Remote. The Wii Remote has unique and powerful capabilities as a gesture controller, and many trials of using it have been investigated (e.g., [18] for music composition).

We customized the Wii Remote by newly attaching an infrared LED to it. Also the LED is covered by a semi-transparent cap to have the light diffused, as shown in the figure. In consequence, even though the device is turned below to the table, its position can be sensed by the cameras which are mounted at the top of the system.

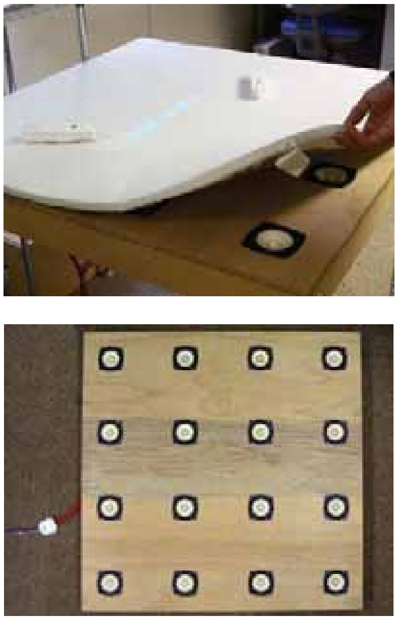


Figure 2. Sound Table

Figure 3. A stick-type input device

3.2 Processing flow

Figure 4 shows a processing flow of the system. As a tool for implementing software modules, MAX/MSP with Jitter is used.

The system receives inputs from a pair of cameras and an accelerometer provided in the stick-type input device (the Wii Remote). 3D position and motion tracking data are analyzed to recognize user's gesture. We define 5 primitive gestures as will be explained in 3.3. Here, the task of identifying gestures is separated from that of their semantic interpretation in a specific music application so that development of the application can be made easier. The result of interpreting a gesture is expressed in a message and sent to audio and visual output processing parts. Here we adopt OSC (Open Sound Control) protocol [19] for message communication among the components.

In receiving the message, the system executes polyphony control to sound multiple notes at one time. Sound is then generated and comes out through 16 speakers. Meanwhile, graphical objects are generated in OpenGL and projected onto the top of Sound Table.

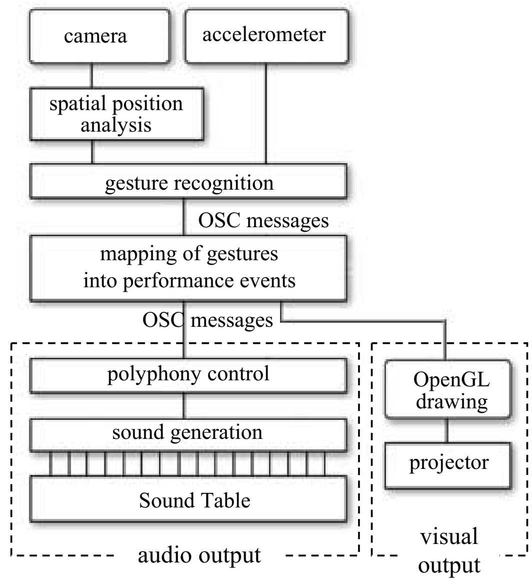


Figure 4. System processing flow

3.3 Primitive gestures

In our system, the user expresses him/herself in gestures in a 3D space, like a composer does on a stage. We define five gestures which are simple, natural, and powerful so as to be used in a variety of music applications. They are *tap*, *drag, attack*, *flick*, and *tilt*, as explained below (see Fig. 5).

Tap is a gesture of touching the head of the stick device down onto the tabletop. It specifies the 2D position on the table and corresponds to a mouse click operation. Internally, when a tap gesture is recognized, a message with its coordinate values is generated.

Drag is achieved by bringing the stick device with keeping its head on the table. It corresponds to a mouse drag operation. Messages with 2D coordinate values of the focus point are generated at certain intervals during the drag gesture.

Attack is a gesture of swinging the device downward quickly, as is seen in the drum play. 2D coordinate values of the attack point, height of the device at the lowered position, and velocity of the attack motion are considered as message parameters.

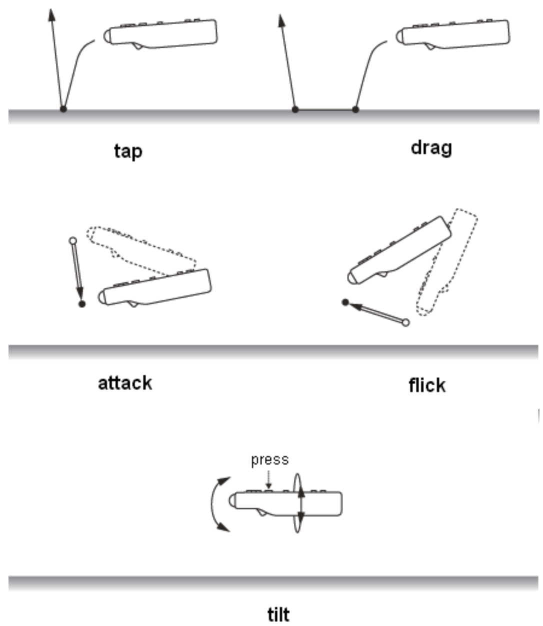


Figure 5. Five primitive gestures

Flick is a gesture of removing (a virtual object on the table) with a light quick blow. This gesture is expressed internally by a message with parameters of device position and motion velocity.

Tilt is a gesture of tilting the stick device. This gesture should be carried out while pushing a button of the device. Though it is possible to have acceleration along three axes by the use of an accelerometer in reality, we remove the yaw motion from consideration since we consider that it is not essential. Furthermore we can reduce gesture recognition error. Posture data of the stick device in 3D and its height from the tabletop are sent with a message.

In the next chapter, we will explain a mechanism of input processing (i.e., motion capturing) and output processing (i.e., sound control) in detail. An application of the gestures into music performance will be demonstrated in Section 5.

4. Input and Output Processing

4.1 Capturing of gestures

There are possible object tracking techniques which suit our system setting. Capacitive sensing with a mesh-shaped antenna, or SmartSkin [20], is a smart approach. However, the sensing area is limited to the surface of the sensing table. Use of a supersonic sensor is another possibility. It can catch the height information, but doesn't work properly in a multi-player collaboration environment. Magnetic tracking technology shows good performance in preciseness. However, we were afraid of the influence of magnetism caused by speakers. Also it is rather expensive.

Finally we decided to apply a vision-based technique. For each of the LEDs (i.e., stick devices), the system detects its position in a pair of images captured by two cameras. The system then calculates the distance between the LED and cameras using triangulation.

Meanwhile, the motion of a stick device is captured by an accelerometer provided in the Wii Remote. This means that it is needed for the system to have a connection between motion (sensor-based) data and position (vision-based) data. However, images which are taken from cameras are in grayscale and no additional features are available. There is no simple way to identify which stick device it is. The system can make a misinterpretation when multiple devices overlap each other.

In order to solve this problem, the system keeps comparing 3D motion features which are extracted from camera input images with those from the accelerometer, and decides one-to-one correspondences between LED marker images and stick devices (see Fig. 6). Even if a correspondence becomes wrong at one time, it can be corrected properly.

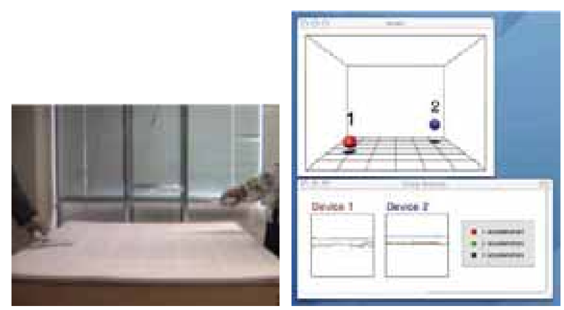


Figure 6. Motion matching of vision-based and sensor-based outputs

4.2 Positioning the sound

Sound Table is designed to be used by multiple users. Multiple different sound objects may appear at the same time, where each of them is associated with an individual user's action. It is essential that the system helps users to identify to whom the sound objects belong.

While 2-channel stereo and 5.1-channel surround sound frameworks for generation of a virtual sound space are available and common in these days, the best spot for listening is fixed. If the listener is out of the spot, a reality of the sound space cannot be maintained any more.

Sound Table adopts 16 speakers to overcome this problem. The system controls the loudness of each speaker so that sound can be positioned at any point irrelevant to the speaker position.

We extendedly apply the well-known sine-cosine pan law provided originally for regular stereo sound management to spatial sound management. The conventional sine-cosine pan law is explained as follows: For a pair of two adjacent speakers, the loudness of one speaker is determined by multiplying the input by the sine of a control value, and that of the other speaker is by multiplying the input by the cosine value. Energy, i.e., loudness of sound, is maintained as constant regardless of the position.

This law is applied to both horizontal and vertical directions of a speaker matrix. Suppose, for example, the compensated results in horizontal and vertical directions for an upper left speaker are l_h and l_v, respectively, as illustrated in Fig. 7. Loudness for the speaker is calculated by l_h*l_v.

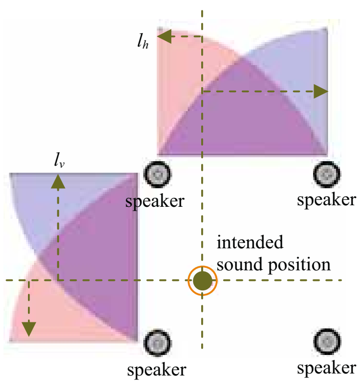


Figure 7. An extended sine-cosine pan law in a 2D sound field

Every sound object can be moved continuously from one place to another by controlling the loudness of each of the speakers. Experimental results showed that this sound positioning scheme works with reasonable performance [21].

Here it is noted that the reason why we adopt such multi-speaker table configuration instead of using a speaker prepared in the Wii Remote is a demand of keeping sound objects at their own position.

5. Playing with System

We have developed a music performance application of the system in order to demonstrate its strengths in collaborative computer music domain. It should be mentioned that the system is not a replacement of conventional music instruments, but is provided to help nonprofessionals enjoy performing music together as a new tool for expressing their ideas.

In order to allow users play with music, sound objects are visualized as colored circles on the table, as shown in Fig. 8. Five operations for manipulation of the sound objects are prepared as listed in the following, which correspond to the gestures in 3.3.

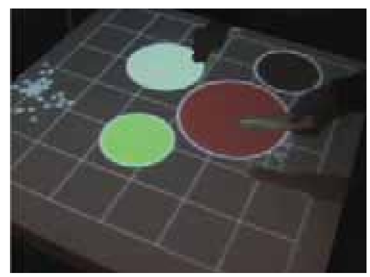


Figure 8. Playing with music

- *Drop* (implemented by tilt) is an operation to newly create a sound object. If the user tilts the stick device with keeping a button pressed, a sound object is generated and spilled over the table. Colored circles appearing on the table in Fig. 8 are such objects. Their size and loudness depend on the position of the stick device. Higher the spatial position, larger the circle size and louder the generated sound.
- *Play & stop* (implemented by attack) is an operation to switch between play and stop of sound. When the device is swung down over a sound object which is sounding, sound is terminated. At the same time, its color becomes black to see the change. If the gesture is applied again, the object restarts sounding.
- *Move* (implemented by drag) is an operation to move a sound object to any position on the table. Needless to say, the associated sound moves along the drag gesture.
- *Change* (implemented by tap) is an operation to change sound sources. A different color is assigned to each of the sound sources. At each time when a tap gesture is

placed on a sound object, its color and sound source change.

- *Remove* (implemented by flick) is an operation to delete a sound object. When the user flicks a certain sound object, it flows out of the table with graphical effects (see the lower image in Fig. 8).

6. Conclusion

This paper proposed a computer music system which allows multiple users to play music collaboratively by their hand gestures in a 3D space.

A musical table containing 16 speakers, called Sound Table, is provided for the system. Users can feel sounds coming up from any arbitrary points on the table irrespective of actual speaker positions.

We also discussed a gesture-based music performance scheme in which a stick-type device is considered a tool for interaction. A mechanism of vision- and sensor-based tracking of user's position and motion was presented. Computer-generated graphic images are projected onto the tabletop so that multiple users can interact with the system through both auditory and visual channels.

Moreover, we presented a framework for system development using the OSC protocol. Development of an application is separated from identification of gestures.

Future work will involve designing more sophisticated operations for music making and conducting experiments on the design of the system. In addition, we now plan to apply the system to other applications such as life review.

7. Acknowledgements

The authors would like to thank Minoru Itakura for his contribution in implementing the system.

8. References

[1] S. Mann, R. Janzen, M. Post, "Hydraulophone Design Considerations: Absement, Displacement, and Velocity-Sensitive Music Keyboard in which Each Key is a Water Jet," Proc., ACM Int'l Conf. on Multimedia, pp. 519-528, 2006.

[2] Y. Takegawa, M. Tsukamoto, T. Terada, S. Nishio, "Mobile Clavier: New Music Keyboard for Flexible Key Transpose," Proc., Int'l Conf. on New Interfaces for Musical Expression, pp.82-87, 2007.

[3] D. Overholt, "The Overtone Violin," Proc. Int'l Conf. on New Interfaces for Musical Expression, pp.34-37, 2005.

[4] http://www.creatingmusic.com/

[5] http://www.sibelius.com

[6] http://www.sonasphere.com/

[7] A. Crevoisier, C. Bornand, S. Matsumura, and C. Arakawa, "Sound Rose: Creating Music and Images with a Touch Table," Proc., Int'l Conf. on New Interfaces for Musical Expression, pp.212-215, 2006.

[8] P. Davidson and J. Y. Han, "Synthesis and Control on Large Scale Multi-Touch Sensing Displays," Proc., Int'l Conf. on New Interfaces for Musical Expression, pp.216-219, 2006.

[9] H. Newton-Dunn, H. Nakano, and J. Gibson, "Block Jam," Proc., ACM SIGGRAPH2002 Conf., p.67, 2002.

[10] J. Patten, B. Recht, and H. Ishii, "Interaction Techniques for Musical Performance with Tabletop Tangible Interfaces," Proc., ACM Conf. on Advances in Computer Entertainment Technology, 2006.

[11] S. Jorda, G. Geiger, M. Alonso, and M. Kaltenbrunner, "The reacTable: Exploring the Synergy between Live Music Performance and Tabletop Tangible Interfaces," Proc., ACM Conf. on Expressive Character of Interaction, pp.139-146, 2007.

[12] H. H. S. Ip, B. Kwong, and K. C. K. Law, "BodyMusic: A Novel Framework Design for Body-driven Music Composition," Proc., ACM Int'l Conf. on Advances in Computer Entertainment Technology, pp.342-345, 2005.

[13] J. Zigelbaum, A. Millner, B. Desai, and H. Ishii, "BodyBeats: Whole-Body, Musical Interfaces for Children," Proc., ACM CHI Conf., pp.1595-1600, 2006.

[14] S. Mann, "The Andantephone: A Musical Instrument that You Play by Simply Walking," Proc., ACM Multimedia Conf., pp.181-184, 2006.

[15] D. Birchfield, K. Phillips, A. Kidane, and D. Lorig, "Interactive Public Sound Art: A Case Study," Proc., Int'l Conf. on New Interfaces for Musical Expression, pp.43-48, 2006.

[16] D. Lock and G. Schiemer, "Orbophone: A New Interface for Radiating Sound and Image," Proc., Int'l Conf. on New Interfaces for Musical Expression, pp.89-92, 2006.

[17] T. Ogi, T. Kayahara, M. Kato, H. Asayama, and M. Hirose, "Immersive Sound Field Simulation in Multi-screen Projection Displays," Proc., Eurographics Workshop on Virtual Environments, pp.135-142, 2003.

[18] S. Toyoda, "Sensillum: An Improvisational Approach to Composition," Proc., Int'l Conf. on New Interfaces for Musical Expression, pp.254-255, 2006.

[19] M. Wright and A. Freed, "Open Sound Control: A New Protocol for Communicating with Sound Synthesizers," Proc., Int'l Computer Music Conf., pp.101–104, 1997.

[20] J. Rekimoto, "SmartSkin: An Infrastructure for Freehand Manipulation on Interactive Surfaces," Proc., ACM CHI2002, pp.113-120, 2002.

[21] T. Nakaie, T. Koyama, and M. Hirakawa, "Development of a Collaborative Multimodal System with a Shared Sound Display," Proc., Ubi-media Conf., 2008 (to appear).

Layered Context Modeling in Situated Information Processing

Augusto Celentano
Dipartimento di Informatica, Università Ca' Foscari Venezia
auce@dsi.unive.it

Abstract

In this paper I propose a model for situated information processing based on a set of layers, each related to a different processing phase and to a different context type. The model is illustrated with reference to a system for processing semantically rich documents, typical of business and juridical domains. Starting from the conceptual level of document definition, through operations of selection, specialization, instantiation and presentation document are adapted to the situation to which they apply and to the local context in which they are issued.

1 Introduction

The term *situated computing* parallels the term *context-aware computing* and is progressively replacing it in the literature on mobile and contextually informed system design [10]. Situated computing draws from situated cognition [8] "which holds that where are you and when are you there matters, and that the state you are in affects what you do" [12].

The word *situation* is used together with or instead of the word *context* to denote the environment affecting the way a user executes a task, interacts with a device, accesses information, and so on. The concepts subsumed by the two words present differences (sometimes subtle) in different authors. In most cases the situation is a component of the user context, referred to personal traits rather than to the surrounding environment, to the equipment, to the available resources, etc. The situation describes the user with respect to the environment, while the context traditionally describes the environment with respect to the user.

I would suggest a different view on situation and context. According to authoritative dictionaries, a *situation* is the status of a part of the world with respect to conditions and circumstances; it defines the relations between something and its surrounding [15]. In my perspective the situation extends the notion of context, which remains primarily attributed to physical, observable features of the user environment. The situation tries to capture aspects of the user environment at a higher level including primarily the user goals, plans, activity and history [13]. As such, situation is not related to the immediate and evolving status of the user, but reflects his/her attitudes and motivations with respect to a wider plan of actions related to a goal.

Keeping distinct the two concepts naturally leads to a layered model of context definition and analysis. The model I propose is focused on data intensive applications, where information goes through several processing stages including operations like selection, specialization, instantiation and formatting of its content. I shall take as a working case study the processing of documents characterized by rich semantic content, defined by formal regulations, having important social effects on the users. Such documents are contracts, agreements, acts, notary deeds, in general documents with a juridical relevance, whose contents change as a function of the situation they address. Each final document serves a well defined case within the context of the user and of the surrounding normative world.

The discussion could as well be focused on the domains of data intensive web sites, e-commerce applications, e-government portals, etc.. As juridical documents, such applications are built according to social and juridical rules, and must adapt to the user's goal, profile and environment. However, dealing with prose text rather than with records and tables might better express the requirements of complex and flexible adaptation at several levels, including issues related to the correct linguistic expression, to domain related jargon and to the writing style.

The paper is organized as follows: after briefly reviewing some literature on situation, context and adaptation in Section 2, Section 3 presents a reference framework for adaptable documents definition and processing. Section 4 introduces the concept of layered context awareness, and Section 5 defines four levels of situated information processing related to different context layers. Section 6 discusses issues related to situation and mobility, and draws the concluding remarks.

2 Situation, context and adaptation

Situated cognition, like context aware computing, builds on the adaptability of information content, applications and interaction through the formalization of how a situation modifies the perception a user has of the environment, hence the way information processing and interaction systems work.

The notion of context has been widely extended since the initial focus on location; the definition by Dey [9] is still taken as authoritative, but many authors try to make more explicitly the presence of components related to the user "per se" (see for example [14, 17]). Extending the analysis of the user situation to include the user plans and goals in executing tasks and accessing information requires a view of the context also in cognitive and linguistic domains [18]. A complete view of context is presented by Bradley and Dunlop [4] who explore a multidisciplinary approach addressing in a unified view different domains.

The notion of context as a description of the situation of a user executing an application has greatly influenced the adaptation of information and documents, evolving from ad hoc processing up to automated systems for generating and tailoring content and presentation to the user environment and needs. Adaptation occurs at several levels: media and layout [22], logical structure [3], inter-media relations [5], and cognitive impact [11]. Consequently, context modeling has been extended from the simplest model based on unstructured attribute/value pairs up to models based on logic and ontologies [7, 20, 19].

Context adaptation of structured information in the area of personal databases has been investigated by the Context-ADDICT (Context-Aware Data Design, Integration, Customization and Tailoring) project[1]. The project aims at supporting mobile users through the dynamic hooking and integration of new information sources, delivering only appropriate, context-based portion of data to their mobile device. Adaptation is based on an ontological representation of the application domain and of data source contents [1, 2, 21].

3 Semantic-based document processing

In scenarios bound to juridical and normative domains the way a document is drawn depends on precise rules that enforce a change or set a status in the world concerning the document issuer, the addressee and the social relations around them. In a broad sense, such documents are the results of the execution of a procedure which follows a plan defined by a set of regulations adapting to variants depending on the situation raised by the procedure actors.

The two terms *situation*, introduced in Section 1, and *case* denote two related but distinct concepts. From an information systems point of view, the concept of situation is closer to the idea of class than to the idea of instance. For example, a situation is the set of features and circumstances which distinguish a purchase of an apartment in a condominium within the class of real-estate transactions; the situation describes the rules, the clauses and the constraints that apply specifically to apartments and not, e.g., to houses or to industrial buildings.

A *case* is a specific occurrence of a situation, in which all the information about the subjects and objects involved is instantiated. Completing the example above, a case is the purchase of such and such apartment, with such and such sellers and buyers, according to such and such financial conditions.

The gap between the two concepts is large, because situations can be more or less detailed, thus representing a continuous range of information types and instances up to a point in which all information is instantiated. Somewhere in this range there is a border, which depends on the domain and is subject to some degree of discretionality.

The paper [6] introduces the concept of *situated document* to denote documents whose processing is defined by the user situation, which does not include only physical or technical context parameters, but also (and mainly) the role of the document, the effects of the document on the user and on his/her social relations, and the user goal, sharing many features of the scenario of this work. I shall take from the architecture for situated documents only the basic ideas described hereinafter, which characterize a broad class of semantic-based adaptable documents, and information containers in general.

Document collection. Documents are instances of classes organized in collections (called *suites* in [6]) grouping related documents. A collection is bound to an application domain targeted to a specified goal. It defines both the goal and the set of documents needed to pursue it. An example of a document collection is the set of documents needed to buy an apartment in a condominium. According to the Italian law, the collection includes a preliminary contract (usually a private agreement), a notary deed of purchase, and documents for the real-estate registry office, the tax offices, the municipality archives, etc.. Other documents, such as a mortgage deed, a declaration of marital status of the buyer, a procurement act, and so on, might be required according to the *situation* in which the apartment is sold, related to the seller, the buyer, the apartment status, and a set of condition set by the local administration which may be different in different locations.

Conceptual document. Each document of a collection is a *conceptual document*, a model for a class of documents whose overall structure, goal and meaning is defined, but where specific data, dependent on the actual case, are not

[1] http://poseidon.elet.polimi.it/ca/

defined. For example, a deed of purchase has an almost invariant structure in which actual data of deed subjects and objects can be embedded. It may have different *clauses*, each corresponding to a different status of the apartment (e.g., there is a mortgage on it), of the seller (e.g., is a minor), of the buyer (e.g., is represented by a proxy), or of the location (e.g., the building in which the apartment is of historical interest, and subject to a bond).

Concrete document. A conceptual document collection is instantiated into a set of concrete documents through a sequence of processing phases pertaining to an adaptation plan that develops in several layers.

Adaptation to the situation concerns primarily the choice of the documents of the suites to be instantiated; this phase is driven by the situation at its most general level. Then, for each selected document type, the adaptation concerns the selection of the pertinent variable clauses which complete the invariant content with details about the specific situation to be described. The outcome of these two phases is a set of document templates which can be used for a class of cases differing only in the instances of the subjects and objects involved, not in the situation they apply.

A semantic-based document model must include at least three elements: (1) a structure, almost invariant, which defines the document main role and the type of content, as required by the class of situations in which it is used; (2) a set of variable and possibly optional components, called *clauses*, which depend on a subset of the information describing the situation; and (3) a set of data instances describing the actual case.

Up to now I have not considered the document final presentation. I have introduced selection and adaptation of the document content, without considering how the content appears. Nevertheless, the presentation and its variants are one of the relevant outcomes of the part of the context that describes the environment, the devices, the user preferences, i.e., the elements that do not influence the document content and meaning, but help to convey it in the optimal way.

This element enters as the final phase of adaptation in the process described above, and is represented by a further component of the document model: a set of rules which define how to instantiate the physical form of the document, according to a set of parameters which describe the related part of user context.

4 Situation and layered context-awareness

The instantiation/adaptation process described in previous section relies on four types of information which enter in four stages of document processing: (1) the external situation and the user goal; (2) the information about facts specializing the situation to a narrow case class; (3) the details about the data instances of the specific case; (4) the environmental context defining the document physical aspect.

To characterize the four information types, we need to define what we mean with *context* in a more precise way.

As recalled in Section 2, there is still some vagueness about what the word *context* means. There is a general agreement about associating it to the environment surrounding the execution of an application, or the processing of an information. Anyway, what is context as opposed to application data is still matter of discussion.

The scenario described in this paper is a good example of this debate. Since early '80s the automatic generation of documents in some well defined domains has been a reality. The business domain and the juridical domain are the two areas which have gained benefit from such automation. Systems for drawing contracts and notary deeds have been designed and are widely used since then. At that time, the word context had no special meaning, and almost all the office document processing applications used all the data at the same level. A database holding all the relevant data was consulted to extract different types of information, such as: switches to select the proper template, coherent with the set of data instances; flags signaling the need to include or exclude clauses; actual data instances to fill out the blanks in document templates; rules to verify the coherence of the different document clauses [16]. Often, also information about the final presentation, such as the document format and the type of printing has to be known during all the process. A few better structured application programs used to differentiate at least the selection of the document template and the final format from the data needed to instantiate the document.

Since then, the (once) new concept of context, new information structuring methodologies and new models for knowledge representation have modified the scenario, and we are now able to distinguish, practically if not theoretically, the different types of data occurring in a complex information processing application, and to arrange them in layers, each layer having a specific goal with respect to the final result.

In processing a document we proceed through a series of phases, in each phase selecting, modifying, adapting or instantiating some components of a comprehensive document model including document types, structures, templates and sentences in natural language. A set of rules supplement the model defining, in some way, constraints and mutual relationships. A simple example, solved at different levels of smartness in old automatic processing systems, is the linguistic concordance among the grammatical properties of the text and the variable data to be embedded. According to the juridical style of language used in notary deeds, which is very precise, a sentence stating that a (set of) purchaser(s) buys (buy) an apartment by a (set of) seller(s) has four lin-

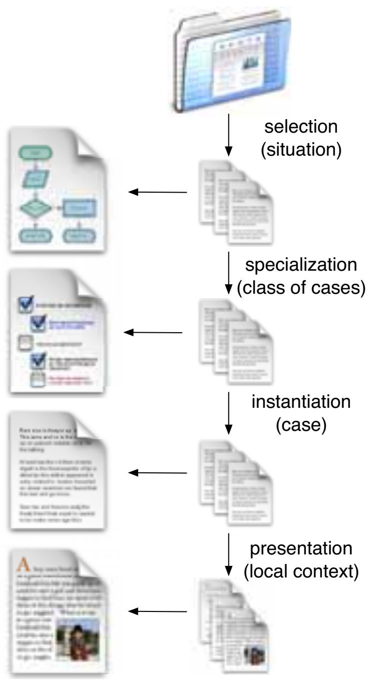


Figure 1. Layered document processing

guistic variants according to the cardinality of each party in the deed, and more complex situations can be raised by gender biased expressions.

The processing phases, and the related ruling descriptions, are not at the same semantic level. Each phase uses a set of data for a specific goal, and produces a specific type of transformation. Some data are used only in one level, other data are used, with different roles, in different levels.

I denote this situation with the term *layered context-awareness*: at each layer a context environment is defined, which is used to modify the way an information is generated, processed or used. The way the context is modeled and described in concrete may depend on the layer. While admitting that the choice of a suitable model is the crucial point of this approach, in this paper I shall overview the pros and cons of some models, leaving to future work the design of suitable and formally processable choice criteria.

5 Layered document processing

Document processing is done for each layer assuming three types of input: (1) a document structure (a type, a template), which is the source for the adaptation of that phase; (2) a context instance, suitable for that processing level; (3) a set of data representing the situation or case specific information for that processing layer. Figure 1 illustrates the four layers.

Layer 1: Selection. The uppermost layer of processing is related to the choice of the documents in a suite; the choice depends on the situation in its widest meaning. It depends on the documents' role for fulfilling the user's expectation. The goal is linked to the application domain, to the initial state of the user world according to the domain, and to the final state to be reached. It could therefore be described with a reference to the semantics of the part of the world involved.

It is worth to note that in this layer the distinction between *context* and *application data* might blur more than in the other layers. If we think of automatic generation of notary deeds, it might be hard to say that the information about what deeds need to be generated is not part of the application input data.

Nevertheless, at this stage we could well ignore the *information* about the subjects and objects involved (which is known as the case), but we need to know which is their situation. We are reasoning at an abstract level about the identities (the instances), but at a concrete level about the circumstances. Hence, it's plausible to assume that most of information at his level is about context and not about content.

Layer 2: Specialization A second layer of processing is the identification, for each document selected at layer 1, of the relevant clauses. The situation is described in more specific terms because it is (also) based on actual data about the subjects and objects of the document. For example, while the presence of a mortgage is an explicit information that must be provided as part of the context in which the transaction is done, the presence of a minor among the parties, identified by the birth date, which is application data, requires he or she to be represented by a proxy, or the presence of witnesses to validate the deed (thus becoming an element of the situation).

Layer 3: Instantiation. The third layer concerns directly the composition of the document content. The document template, selected in the suite and adapted in its skeleton to receive the appropriate data, is filled with the variable data that describe the subjects and objects of the transaction. We assume that in this layer the final presentation is not yet taken into consideration, therefore it is most concerned with processing application data: the contribution of context is therefore very marginal. Nevertheless, some dependences from parameters not immediately recognizable in the application data, still not related to the physical aspect of the document, could be found at this level.

The more immediate (and maybe abused) example is the adaptation of the document language. Should such a change be reflected in the document template used? As such, it should be raised to the previous layer. Or should it correctly be considered as part of the instantiation, since the *meaning* of the template in the current case does not change, but only its external appearance (even if it is not matter of simple formatting)? This case, in its simplicity, shows the type of ambiguity that even in a controlled architecture like the one proposed here affects the classification of context with respect to application data.

Another type of context, still more difficult to grasp, is the choice of the linguistic (or literary) style of the final document. As we are facing prose text and formally relevant documents, processing is not just a matter of filling a few placeholders in a template. Natural language processing techniques exist that may help to produce good text according to adaptable grammar and rethoric rules. In this case, however, we risk to assign to the word *context* a different and broader meaning, involving the cultural environment of the users. The concept of adaptation broadens too; besides trespassing the scope and the goals of this paper, approaching such an issue requires a different focus and a different skill.

Layer 4: Presentation. The fourth layer includes the final instantiation, i.e., the adaptation of the document to the local context of the user, of the environment, of the device, according to the requirements, the guidelines and the techniques so deeply discussed in the literature about context-awareness. Therefore, I do not enter into this level, which is anyway important in resource critical scenarios such as those based on mobile technology.

The processing functions needed in the different layers demand different models for defining the situation and the context. In the first layer, selection, the situation definition must express a possible complex knowledge about the application domain procedures and about the role of information. It must be rich enough to express relationships between classes of facts, which can be implemented as relationships between sets. Models based on ontologies, despite their complexity, could serve the purpose; for efficiency reasons, this level could be categorized as a collection of *sub-situations*, each stable enough to be selected with simple mechanisms. Models based on logic could also be used profitably, mainly for expressing direct dependences between facts related to the situation and single conceptual documents.

In the second layer, specialization, the context denoting the situation is derived from a subset of data about the transaction, plus a set of information about how to execute the transaction. This set will not be used to adapt the document text, rather it will be used to select the proper clauses. A suitable model to describe the interrelationships between these two classes of information must be able to manage rules and constraints, therefore a logic based context model is to be preferred.

At the third layer, instantiation, much depends on what is considered context adaptation with respect to document pre-selection, and to what extent a natural language based processing is considered necessary, as discussed above. At the most demanding side of the problem, however, I do not see context models so rich to account for such a complex processing, except that models and formalisms for knowledge representation could help to reason about context too.

The fourth layer, presentation, mimics the adaptation functions commonly assigned to context-aware systems, therefore a large fan of context models exists to support it.

6 Situated computing and mobility

In the current literature context awareness and situated computing are strongly related to mobility, and have received increasing attention from the growth of mobile information processing. Nevertheless, the concepts are disjoint. Mobility induces technology driven constraints, requires best use of resources both in terms of computing and communication power, and challenges the user cognitive effort. Mobility spontaneously leads application designers to consider context-awareness as a fundamental feature, since the information processed is subject to change during a session: information related to the location and to the environment, resources related to the communication and to the device, and the goal of the user may change as the situation of the user and the surrounding environment change.

Is situation-awareness a typical property of mobility? The answer is, of course, yes. Is situation-awareness meaningful only in presence of mobility? Despite the largest effort has been devoted to study situation-awareness in relation to a changing user environment, the answer is *no*. Situation expresses a concept broader than the environmental context defined in pervasive systems, and applies also when the environment does not change.

As mobile devices increase their capabilities, the difference with desktop devices will no longer be a matter of functions, but of opportunity. A wide and authoritative scientific literature foresees the use of personal mobile devices to carry out most of the activities today done with desktop computing devices: the management of medical records and medical tests in the healthcare domain, the electronic transactions in the business domain, on-site assistance in case of disaster in the emergency domain, and so on.

The information processed in such cases needs to be adapted to the user goals, before than being adapted to a device. What part of information about a patient is relevant for a doctor depends on what role the doctor is playing when at the patient home, in the hospital, in the operating room, in a mobile rescue unit.

The layered context architecture illustrated in Section 4 is only a first step that needs to be completed with inter-layer paths, which assure that information processed in a higher layer can be further processed in a lower layer. For example, going back to the juridical domain used throughout this paper, the selection of a conceptual document within a document suite must be compatible with the possibility of processing the document down to the concrete instance. A professional user accessing a real-estate transaction document suite through a mobile device could receive all the information about a deed to be stipulated, but could not be able to produce the final document, e.g., because of limited access to public registry offices from a mobile device, a process which is managed at document instantiation time.

Such a requirements impacts on the models used for defining the context, and on the processing architecture that integrates context and application data processing, and requires to analyze at a deeper extent the ideas here proposed as a first step to approach the problem.

References

[1] C. Bolchini and E. Quintarelli. Filtering mobile data by means of context: a methodology. In *Proc. OTM Workshops 2006 - Context-Aware Mobile Systems*, pages pp. 1986–1995, 2006.

[2] C. Bolchini, F. A. Schreiber, and L. Tanca. A methodology for very small database design. *Information Systems*, 32(1):61–82, March 2007.

[3] S. Boll, W. Klas, and J. Wandel. A Cross-Media Adaptation Strategy for Multimedia Presentations. In *ACM Multimedia*, pages 37–46, 1999.

[4] N. A. Bradley and M. D. Dunlop. Toward a multidisciplinary model of context to support context-aware computing. *Human-Computer Interaction*, 20(4):403–446, 2005.

[5] A. Celentano and O. Gaggi. Context-aware design of adaptable multimodal documents. *Multimedia Tools and Applications*, 29(1), 2006.

[6] A. Celentano and F. Pittarello. Situated multimodal documents. In *MIMIC '07, 1st Int. Workshop on Management and Interaction with Multimodal Information Content*, Regensburg, Germany, September 2007.

[7] H. Chen, F. Perich, T. Finin, and A. Joshi. SOUPA: Standard ontology for ubiquitous and pervasive applications. In *MobiQuitous 2004, 1st Annual International Conference on Mobile and Ubiquitous Systems, Networking and Services*, pages 258–267, Cambridge, MA, 2004.

[8] W. J. Clancey. *Situated cognition: on human knowledge and computer representations*. Cambridge University Press, 1997.

[9] A. Dey. Understanding and Using Context. *Personal Ubiquitous Computing*, 5(1):4–7, 2001.

[10] P. Dourish. A foundational framework for situated computing. Position paper for the CHI 2000 Workshop on Situated Computing: A Research Agenda.

[11] N. Elouazizi and Y. Bachvarova. On cognitive relevance in automatic multimodal systems. In *Proc. IEEE Conf. on Multimedia Software Engineering*, pages 418–426, Miami, FL, 2004.

[12] J. S. Gero. Situated computing: A new paradigm for design computing. In A. Choutgrajank, E. Charoenslip, K. Keatruangkamala, and W. Nakapan, editors, *CAADRIA03*, pages 579–587, Bangkok, 2003.

[13] K. P. Hewagamage and M. Hirakawa. Situated computing: A paradigm to enhance the mobile user's interaction. In S. K. Chang, editor, *Handbook of Software Engineering and Knowledge Engineering*. World Scientific, 2000.

[14] A. Jameson. Modeling both the context and the user. *Personal Technologies*, 5(1):29–33, 2001.

[15] Merriam-Webster. *Merriam-Webster's Collegiate Dictionary, 11th Edition*. Merriam-Webster, 2003.

[16] P. Paolini, F. Garzotto, S. Mainetti, F. Barbic, A. Celentano, and F. Liguori. Knowledge based document generation. In W. Lamersdoff, editor, *Office Knowledge: representation, management and utilization*. North-Holland Publ. Co., 1988.

[17] D. Petrelli, E. Not, M. Zancanaro, C. Strapparava, and O. Stock. Modelling and adapting to context. *Personal Ubiquitous Comput.*, 5(1):20–24, 2001.

[18] A. Schmidt. A layered model for user context management with controlled aging and imperfection handling. In *MRC 2005, Modeling and Retrieval of Context*, volume 3946 of *LNCS*, pages 86–100. Springer Verlag, 2006.

[19] T. Strang and C. Linnhoff-Popien. A context modeling survey. In *UbiComp, 1st International Workshop on Advanced Context Modelling, Reasoning and Management*, pages 34–41, September 2004.

[20] T. Strang, C. Linnhoff-Popien, and K. Frank. CoOL: A context ontology language to enable contextual interoperability. In *DAIS 2003, Distributed Applications and Interoperable Systems, 4th IFIP WG6.1 International Conference*, volume 2893 of *LNCS*, pages 236–247, Paris, France, 2003. Springer Verlag.

[21] L. Tanca. Context-based data tailoring for mobile users. In *Proc. BTW Workshops 2007*, page 282295, 2007.

[22] J. van Ossenbruggen, J. Geurts, F. Cornelissen, L. Rutledge, and L. Hardman. Towards second and third generation web-based multimedia. In *Proc. Tenth International World Wide Web Conference*, pages 479–488, Hong Kong, 2001.

Case Studies on Augmented Mirrors that reflect Context

Kaori Fujinami
Department of Computer, Information and Communication Sciences,
Tokyo University of Agriculture and Technology
fujinami@cc.tuat.ac.jp

Abstract

In this paper, we present two case studies that augment the notion of a traditional mirror with context reflective functionalities. Contextual information contributes to a user to take an appropriate action with less effort because it is inherently extracted from huge amount of or unstructured information. A mirror that reflects not only physical appearance but virtual one, context, should support a user's decision making and enhance the awareness of his/her health. The first case aims to investigate the interaction method with the mirror that presents context helpful for decision making for the next appointment. The persuasive nature, self-reflectiveness, of a mirror is focused in the second case, where one's history of daily physical activity, i.e. walking, and the sharing with others is the context.

1. Introduction

A mirror has been utilized by people since ancient days[12]. It reflects physical objects appearing in front of it. Often we become inquisitive about our physical appearance by looking at a mirror. Moreover, while using a mirror we can also see and comprehend what is happening in the backgrounds, e.g. someone is passing, a kettle is boiling, etc. This reflective nature is essential for a mirror. We usually stay in front of a mirror for a period of time. These characteristics suggest us the acceptability of presenting information on the surface of the mirror. Furthermore, we often think about immediate events or something important e.g. schedule of the day, weather forecasting at today's destination, weight increasing, etc., while performing something else there. We consider such information represents a user's context. Contextual information contributes to a user to take an appropriate action with less effort because it is inherently extracted from huge amount of or unstructured information. Therefore, a mirror that reflects not only physical appearance but virtual one, *context*, should support a user's decision making and enhance the awareness of his/her health.

In this paper, we present two case studies that augment the notion of a traditional mirror with context reflective functionalities. The first case aims to investigate the interaction method with the mirror that presents context. Information based on a person's schedule is presented with two levels of abstraction, that supports one's decision making. The target user is identified through the combination of low-level sensors for privacy reason. We then focus on the persuasive nature, i.e. self-reflection, of a mirror as a second case, where the history of daily number of steps is presented to make people aware of health status in a subtle manner. In both cases, an augmented mirror is realized by attaching a translucent mirror board on an ordinary computer monitor, where a bright color behind is seen through while an object in front of the board is reflected in the dark colored area.

The rest of the paper is organized as follow. The first and the second cases are investigated in section 2 and 3, respectively. Section 4 examines related work, and we concludes the paper in section 5.

2. Case1: Supporting Decision Making

2.1. Overview

In the first case, the mirror displays information relevant to a person in front of it on the periphery of his/her sights. We can change our behavior once we notice some remarkable situation. The mirror has been designed to start presenting information with abstract images so that a user could perceive the contents at a glance after the identification (Figure 1-(a)). Upon the request from a user, it turns to show textual messages for clear understanding. In Figure 1-(b), a scrolling message that describes the image appears on the bottom of the display. The detail can be found in [4].

2.1.1 User Identification

A *user* is implicitly identified in a privacy aware manner. This is an important aspect to augment a mirror that requires user identification since the place around mirror is highly privacy sensitive. We have determined to integrate three types of low level information, rather than utilizing a single and rich data source like a video camera. They are 1) the existence of a person in front of a mirror using an infra-red range finder, 2) the collocation of a personal (unshareable) artefact with a mirror, and 3) utilization of the personal artefact. Such an everyday artefact that can be utilized as a representative of a person includes a safety razor, a toothbrush, a comb, etc. We have decided to use a toothbrush for detecting a user because it is gender neutral and utilized by almost everyone.

2.1.2 Schedule-driven Context

We brush our teeth, make up, etc., that is normally done in situations where we are considering immediate events or something important over the day to come in the morning. We consider information related to an immediate event is useful in such a case. We also often change our decision if someone tells us about weather forecasting, traffic accident, etc. Considering these facts, we have selected three types of information to be displayed that can affect our behavior and supports decision making: 1) weather forecasting (state and temperature) at the destination where he/she will go, 2) state of a public transportation system that the person is going to use, and 3) information about the next appointment. These types of information can remind us of taking something required, e.g. umbrella and documents, and offers us the opportunity to take an alternative route to go to the destination, or rush ourselves. So, we consider that it is appropriate for information presented on a mirror.

2.1.3 Two Modes of Information Rendering

Although the above listed information is useful, it might disturb a user's primary task if all information appears on the surface of a mirror. Therefore, we have decided to apply two types of information rendering and interaction methods with the mirror. By default, the mirror does not display anything, so it looks like just a mirror. However, when a person appears in front of AwareMirror and a toothbrush is utilized, then AwareMirror enters into the *abstract mode*. In the abstract mode, it plays a role of an ambient display, which shows abstract information on the periphery of a user's line of sight (Figure 1-(a)). Abstract information can be represented using images and colors we are familiar with and therefore easy to understand at a glance. This is also a solution that takes into account of the characteristic of the place, *in front of a mirror*, since some people often take their glasses off while brushing their teeth and washing their faces. Therefore, we consider that it is useful for them to be just notified of some important events.

The other mode, *detailed mode*, however, shows information in more detail using text information (Figure 1-(b)). The mode transition is caused by the person explicitly. Hence, we consider that a feeling of disturbance does not come up in his/her mind. Both modes automatically return to the default screen on detecting a user's leaving.

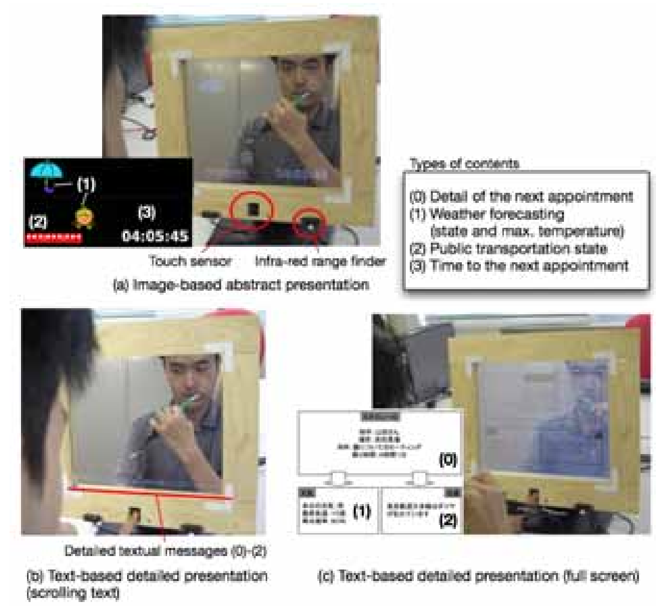


Figure 1. Using an Augmented Mirror

2.2 Evaluation on Interaction

The interaction method with the mirror is evaluated with a WoZ method[2] and in-situ experiments.

2.2.1 Methodology

Contrastive functionalities were implemented, and tested together with the original ones. The points of the comparison are as follows.

1. Automatic vs. Manual activation of the first (abstract) presentation

2. Existence vs. Little-existence of the reflective nature in detailed presentation

3. One phase (only textual message) vs. Two phase information provision

Note that the manual activation of the first scene was realized by touching the same area as the mode transition. The version with little reflective nature can be seen in Figure 1-(c). The message is drawn in black on white background, that reduces the reflection on the surface due to the characteristic of a magic mirror. Therefore, it looks like a traditional computer screen, rather than a mirror.

A WoZ method was applied to obtain comprehensive view while an in-situ experiment was conducted for intimate findings. Fourteen subjects (12 students from outside the authors' laboratories and 2 company employees) were recruited for the WoZ experiment. Each subject was directed to use the system, e.g. start brushing teeth, pressing the touch sensor, and the observer actually controlled the system.

The in-situ experiments were conducted with totally four persons, i.e. one family (parents and an adult daughter) and a company employee living alone. A working prototype was installed in the living room at each home for 17 and 22 days, respectively. The subjects were asked to change the properties for the different experimental conditions by themselves. In the following discussion, the results of the above two experiments are combined.

2.2.2 Results and Discussions

Activation of Abstract Information Mode The half of the subjects (7 people) preferred the automatic activation that does not require any explicit direction. The comments are that even simple touching was forgettable in the busy morning, and sometimes it is forgettable. This indicates that they preferred to confront the system as an ordinary mirror. Notably, all the four subjects from the in-situ experiments liked this version. On the other hand, the subjects who liked manual activation needed the rendered information only when they wanted to see it. This comes from their situations where they usually had no specific plan that requires decision making. For example, an accidental situation of a train to take can be shown while the normal case does not render any information. These facts suggest that preserving a mirror's original functionality is important for all the subjects.

Detailed Information Rendering The full screen version (Figure 1-(c)) was preferred by 68.8% of the subjects due to readability of presented information. It shows all the information in a screen at a glance although it has little reflective nature. In contrast, the scrolling text version (Figure 1-(b)) requires a user to follow the message to the end, and thus he/she needs to pay more attention. The reflective nature of the mirror would rather obstacle. The subjects pointed that they noticed the mirror was not functioning during the full screen. But they did not mind it since the time to check the detailed information was short because of the readability. This suggest that the full screen version contributes to reduce the interruption into the primary task consequently. On the contrary, a subject who had negative impression on the full screen version requested to see her face in the whole period of brushing teeth. This indicates the desire to keep the original function intact at any time, that is supportive of our initial design principle.

Two Phase Information Presentation Ten subjects (66.7%) wanted the two phase information presentation, i.e. the detailed information with textual message is presented after a graphical abstract information. The major reasons are that 1) they do not always need detailed information and 2) it is difficult to read the textual message from the beginning in case that they do not wear glasses. The subjects who participated to the in-situ experiment turned to the textual expression mode when they actually wanted to know the detail. Two subjects changed their behavior twice based on the information (one for low temperature and the other for an accident of public transportation). Such a decision making support during an ordinal task is what the mirror is aiming at.

Furthermore, a subject looked forward to seeing a change in the image when he stands in front of the mirror. We could say this is interesting and important aspect in a daily life computing paradigm. Such an aspect is basically ignored by a traditional "efficiency-based" computing paradigm. Also in this point, the two phase presentation proved useful.

3 Case2: Supporting Self-Reflection

3.1 Overview

In the second case, another mirror was developed to enhance the awareness of health as an instance of the *persuasive technology*[3]. The presented information is considered as a context of a user. It represents his/her situation regarding walking exercise. The same superimposing technique was adopted. However, in this version, only the graphical presentation was utilized. The presentation was designed to have an unpredictable nature as well as competition/collaboration aspects with others so that he/she could continue the exercise to have a fun with it. More detail can be found in [5].

3.1.1 Ambient Walking Persuasion

We have specified four strategies for walking encouragement: pleasurable interaction with the mirror through daily activity, day-by-day history of walking, an abstract form of a character's appearance and facilitating inter-personal encouragement.

We consider that an appropriate level of unpredictability is a source of pleasure, which attracts people to the display, and thus becomes a trigger to be aware of their health. We have utilized a flock of life-like characters to represent day-by-day history of walking of a week or more. This is often called *swarm intelligence*. A character corresponds to the achievement of a day, and moves autonomously in the screen following the *Boids* algorithm[11]. As a result of local decision based on the neighboring view, a wide variety of flock forms emerge. Furthermore, an explicit interaction by a user is realized by mouse clicking, where a click act as "a stone thrown in a pond". Each character *escapes* from the position. A user can make a change (destroy) on a flock and watch the process of re-organization.

The daily result is defined as a ratio of a day's steps to the day before. If a day's result is better than the nearest preceding day, a user might feel satisfied even though the number of steps itself is small. On the other hand, a bad result should make him/her recall the day's activity compared with the preceding day, and contribute to the improvement of the next day's activity. Also, the daily goal is set to a certain number of steps, i.e. 8000, to provide a feeling of accomplishment. In Figure 2-(c), 4 days' history (from 24th to 27th December) is shown, where the bright color on 24th indicates that the number of steps exceeded the goal of a day. We consider these representation can be a clue for people to review the activity of a day, by which they can effectively develop strategies to get better results.

Drawing a realistic character, e.g. bird, might allow people to have intimate feelings and become more aware of the states. However, subjects with low motivations tend to refuse the system when characters do not grow or die, rather than watch their unhappy appearance[9]. So, we have decided to provide a character with an abstract form of appearance. Finally, inter-personal encouragement has been realized in a way that flocks for a group of users are presented in the same display, and each member in different flocks try to avoid collision when they come to close. People may have feelings of cooperation or competition by the *group context*.

3.1.2 Mapping Daily Walking to a Character

The shape and the behavior of an individual character, *boid*, reflects the healthiness of a person's body. A triangle shape is utilized to clearly indicate the moving direction. The width of the triangle varies according to the ratio of a day to the day before, while the area is constant. Figure 2-(b) illustrates the parameters and three examples of boids with (i) lower, (ii) equal, and (iii) higher ratio. The more a person walks compared to the day before (i.e. the bigger the ratio), the sharper the shape becomes (small w).

The behavior of a boid is determined by the perception of a small neighborhood[11] and the speed of moving. Here, the view is defined by distance ($dist$) and range (deg), and becomes longer and wider when the ratio gets high. This comes from our general understandings that a person who exercises well and is thus active in mind is more able to sense external world. So, he/she can find more members of the same group to form a flock as well as obstacles to avoid. The velocity (v) is proportional to the ratio, that allows a boid with high ratio to move faster.

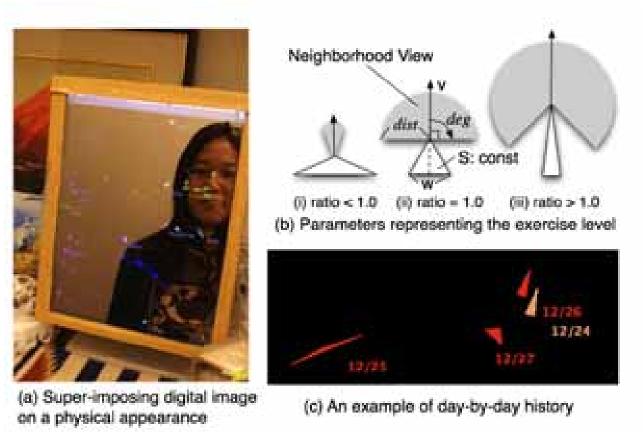


Figure 2. A scene of utilizing the display and the parameters controlling a boid's behavior

3.2 Preliminary Experiment

3.2.1 Methodology

In-house experiments have been conducted to investigate the acceptance of the concept. The system has been deployed at two homes and an office with totally 6 people (two males and four females). The ages range from 23 to 71 years old with 30's as the most frequent range of the age. The displays are connected via the Internet for the inter-personal encouragement. The mirrors were installed outside the lavatory to avoid damaging machines by dew condensation. The alternative placement has been decided based on the preferences of the subjects. They are 1) on the shoe cupboard in the entrance, 2) on a drawer in the living room, and 3) on the fridge in the laboratory.

In this experiment, we have utilized a pedometer (TANITA FB720[15]) that is commercially available. Everyday, a user manually registers the number of steps into the system. Three months have passed since the installation. Comparative versions that do not show information on a mirror surface have also been developed and tested: 1) without a translucent panel (black background) and 2) with the background of randomly selected picture (like a digital photo frame).

3.2.2 Results and Discussions

Although we have not yet confirmed significant improvement in the number of daily steps, 5 of 6 subjects have become aware of their daily steps. Every subject preferred the mirror version because of its natural look. The version without mirror is clear to see, but they told it was unnatural into daily environment. On the contrary, a digital photo frame is getting popular, however the combination of a picture with the boid characters looked still strange to them.

A subject spent long time with the mouse clicking interaction, where she enjoyed destroying the flocks and watching the re-organization process many times. During the process, she talked with herself in a mirror and checked her appearance against the movement of her flock. She thought about her health as well as other things that were too embarrassed to tell other people. We can say this is "deep interaction" with a mirror, that certainly made an impact on her activity implicitly or explicitly.

We have confirmed the effect of the pleasurable interaction, the day-by-day comparison, and the interpersonal encouragement. Some subjects went further when they took a dog for walk. They were pleased when they saw the blinking boid, i.e. the daily goal was achieved, and even a sharp boid that was not blinking. Also, emergent behaviors of boids like chasing and skirmish were pleased, and the subjects looked to forward to seeing new ones. One subject disliked the disordered flock with a flat and isolated boid, e.g. labeled with "12/25" in Fig. 2-(c), since she likes to be a person with prompt action and self-organized. Then, she intentionally walked more than or equal to the day before to maintain "beautiful" flock. A housewife subject liked the expression based on the ratio of two consecutive days. She was satisfied with the increasing number even though the number itself is small because she is aware of shortage of exercise and the increasing is the progress for her. The members in distributed sites talked about their achievement in each group, which generates the atmosphere or *social pressure* to be aware of their steps.

4. Related Work

Much work on an augmented mirror is found. The first criteria is the method of super-imposing. We have taken the approach where the reflective nature is realized by a physical mirror, i.e. translucent mirror. Interactive Mirror [8] and Miragraphy [6] take the same approach. The other method, the utilization of a video camera, allows a system to create modified image of a person who is in front of the mirror [1, 16] and show stored images that present the events happened there [10]. The virtual mirror approach provides flexibility in image processing, however it contains the feeling of privacy violation even though it processes the captured image in a real time manner. Furthermore, the delay to show images would bring discomfort to a person, that is critical in an appliance that is utilized in daily lives.

Regarding the application and the interaction method with an augmented mirror, information that supports decision making like ours has also proposed in [8], where gesture and touching is utilized for the manipulation of presented information. Also, Miragraphy [6] provides more active involvement. Information about clothes and their coordination is shown by the detection of an RFID tag, and the history of daily clothes captured by a video camera is presented. To the best of our knowledge, a user explicitly *utilizes* the functionalities provided by these systems, that could make a gap between a user and a system. Our interaction method described in section 2.1.3 would provide very small gap since the involvement is designed to be the extension of the activities in front of an ordinal mirror. The Persuasive Mirror [1] shows a predicted image of a person in front of a mirror to change the unhealthy habits. The image acts as a warning message to the person, that is more direct than our approach. Here, the persuasiveness of the image itself is the key to the success. In contrast, our approach presents day-by-day history of walking, that provides a user with room to interpret the meaning and the reason. We consider that is more effective to be aware of his/her health.

The work of Rodenstein [13] applies the opposite notion of a mirror and the biggest characteristic of a window, *transparency*, to overlap information in cyber space, i.e. short-term weather forecast, onto a view of outside. Information visualization with artistic impression like Informative Art [7] and InfoCanvas [14] suggests an "augmented paintings", which can naturally fit into daily living. Their roles are presenting information from the beginning, while AwareMirror has more strict constraints that it should not change its original usage as a mirror while offering additional values, *personalized information.*

5. Concluding Remarks

In this paper, two case studies on augmented mirrors that reflect context were presented. The first case aims to investigate the interaction method with an augmented mirror that presents context to help decision making to a person in front of it. The contents, sensing modality, and interaction method have been determined based on the characteristics of the place where a mirror is installed and a user's primary task with a mirror. The interaction method was evaluated by comparative functionalities. The result shows that the majority of the subjects preferred the two phase information provision, automatic activation. However, they preferred the detailed mode rendering with less reflective feature of a mirror. We consider that a designer of an augmented mirror needs to take care of the controllability of a system and the efficiency of understanding information, rather than adhering to keeping an artefact's original functionality intact. This may sound trivial in traditional interactive system. However, we believe the findings would contribute to update the design principle of a new type of a display that provides information through its original usage.

In the second case, we focused on the persuasive nature, *self-reflectiveness*, of a mirror. The context that the mirror presents is the history of daily number of steps. Furthermore, the histories of other persons are shared, that acts as a *group context*. This was intended to make him/her aware of the health status in an ambient manner. Through the in-situ experiment, we have confirmed the acceptance by the subject compared to non-mirror versions. Also, a promising effect of *deep interaction* with the augmented mirror has been found. We are planning to conduct comparative studies with various aspects for more thorough evaluation. This includes 1) other forms of an ambient display, 2) the level of unexpectedness in the behavior of a flock, and 3) the reality of the character in terms of understandability, intimacy and emotion against a negative state.

Acknowledgments

This work has been supported by The Ministry of Education, Culture, Sports, Science and Technology in Japan under a Grant-in-Aid for Division of Young Researchers, and by the foundation for the Fusion Of Science and Technology (FOST).

References

[1] A.Andres del Valle, et al.. The Persuasive Mirror: Computerized Persuasion for Healthy Living. *Human Computer Interaction International (HCI International)*, 2005.

[2] N. Bernsen, et al. *Designing Interactive Speech Systems — From First Ideas to User Testing.* Springer Verlag, 1998.

[3] B. J. Fogg. *Persuasive Technology: Using Computers to Change What we Think and Do.* Morgan Kaufmann Publishers, 2003.

[4] K. Fujinami, et al. AwareMirror: A Personalized Display using a Mirror. In *Proc. of Pervasive2005*, pages 315–332, 2005.

[5] K. Fujinami, et al. A Case Study on an Ambient Display as a Persuasive Medium for Exercise Awareness. In *Proc. of Persuasive2008*, pages 266–269, 2008.

[6] HITACH, Human Interaction Laboratory. Miragraphy. URL: <http://hhil.hitachi.co.jp/products/miragraphy.html> (in Japanese).

[7] L. E. Holmquist, et al. Informative art: Information visualization in everyday environments. In *Proc. of GRAPHITE '03*, pages 229–235, 2003.

[8] T. Lashina. Intelligent bathroom. In *Ambient Intelligence technologies for wellbeing at home, a workshop on European Symposium on Ambient Intelligence (EUSAI2004)*, 2004.

[9] J. J. Lin, et al. Fish'n'Steps: Encouraging Physical Activity with an Interactive Computer Game. In *Proc. of UbiComp2006*, pages 261–278, 2006.

[10] P. Markopoulos, et al. The photomirror appliance: affective awareness in the hallway. *Personal and Ubiquitous Computing*, 10(2-3):128–135, 2006.

[11] C. W. Reynolds. Flocks, Herds, and Schools: A Distributed Behavioral Model. In *Proc. of SIGGRAPH'87*, pages 25–34, 1987.

[12] Richard L. Gregory. *Mirrors in Mind*. W. H. Freeman and Company, 1997.

[13] R. Rodenstein. Employing the Periphery: The Window as Interface. In *Extended abstract of CHI'99*, pages 204–205, 1999.

[14] J. Stasko, et al. Personalized Peripheral Information Awareness through Informative Art. In *Proc. of Ubicomp2004*, pages 18–35, 2004.

[15] Tanita Co. 3-Axes Pedometer. URL: <http://www.tanita.com>.

[16] K. Ushida, et al. i-mirror: An Interaction / Information Environment Based on a Mirror Metaphor Aiming to Install into Our Life Space. In *Proc. of the 12th International Conference on Artificial Reality and Telexistence (ICAT2002)*, 2002.

Capturing and Sharing Sonic Experiences through Web-Based Acoustic Maps

Daniela Fogli[1] and Elisa Giaccardi[2]
[1]*Dipartimento di Elettronica per l'Automazione, Università di Brescia, Italy*
fogli@ing.unibs.it
[2]*Department of Computer Science, University of Colorado at Boulder, USA*
elisa.giaccardi@colorado.edu

Abstract— *Due to the widespread availability of web mapping services, Internet users can now contribute cartographic content and easily create, modify, and share geographic maps. Several web mapping services that are currently available enable users to geographically locate text, pictures, and movies on a map and then share them with family, friends, and other interested users. However, few services give emphasis to the importance of sounds in human experience. In addition to contributing to filling an existing gap in collaborative web mapping, the authors argue that using sounds represents a personal and powerful way to remember, re-encounter, and gain new understandings of everyday space—thus facilitating expression and interpretation of personal meanings through richer and more immersive map-based interactions. The paper describes a collaborative mapping system, TheSilence.org, based on Web 2.0 technologies, that enables users to contribute and share sounds. The importance of sounds in collaborative web mapping with respect to multimedia content and immersive map-based interaction is discussed through the results of a pilot study.*

Index Terms—*web mapping, collaborative systems, graphical user interfaces, map-based interaction, multimedia, sounds*

1. Introduction

Due to the increasing sophistication in web technologies and the widespread availability of web mapping services, web-based maps are becoming popular tools to trace and visualize location-based information. Internet users can now contribute cartographic content and easily create, modify, and share geographic maps. This opportunity offers people new ways to connect with each other and with their urban and natural environments, opening up the possibility for mapping and visualizing both individual and collective meanings ascribed to particular locations [2][6].

Different kinds of content can be contributed and shared through web-based maps. Several web mapping services are currently available that enable users to geographically locate text, pictures, and movies on a satellite or street map ("geotagging") and then share them with family, friends, and other interested users. However, few web mapping services enable users to contribute and share sounds; the importance of sounds in human experience with respect to immersive map-based interaction is usually not given much emphasis.

Sounds play an important role in understanding the nature of the acoustic environment, or "soundscape," in which we live [16], and sounds can significantly contribute to improve the quality of our lives [5]. This paper describes a collaborative mapping system based on Web 2.0 technologies for the creation and sharing of web-based acoustic maps. The system enables users to map and annotate the soundscapes of urban and natural environments, and to use sounds to capture and share their sonic experiences and express the personal meanings they associate with these sounds.

We conducted an experimental activity in Boulder, Colorado, in the summer of 2007. It involved 20 participants for a period of 6 weeks and allowed us to evaluate the usability and usefulness of such a mapping system. Much data was collected, and several findings emerged from the analysis of these data. In this paper, we discuss in particular those findings concerned with the importance of sounds in collaborative web mapping with respect to multimedia content and immersive map-based interaction.

2. Related Work

"Geospatial web," or "geoweb," is a term used to indicate a new infrastructural paradigm that enables users to navigate, access, and visualize geo-referenced data on the web as they would in the physical world [13]. This architecture offers users the possibility to easily create, modify, and share online maps. Popular web mapping services, such as Google Maps [10] and Google Earth [9], enable users to create personalized 2D and 3D maps and annotate them with texts, photographs, and videos. Maps enriched by users' commentaries, pictures, and movies are increasingly offered also by "mashups," such as WikiMapia [19] and CommunityWalk [3]. Even media-sharing services, such as Flickr [4] and YouTube [12], are now either providing collaborative mapping features or extending their uploading capabilities to enable users to add location metadata to their photos and videos and to use these "geotags" to organize and search multimedia content by place. However, despite the growing popularity of collaborative web maps, few services enable users to contribute and share sounds, giving little emphasis to the importance of sounds in human experience.

A project attempting to fill this gap is Wildsanctuary [18]. Wildsanctuary aims to add environmental sounds to Google Maps and Google Earth to provide users with an immersive experience of place through access to the world's soundscape. While offering access to a large archive of field sounds, the project does not provide collaborative mapping features for users to contribute audio files and engage directly in soundscape creation. On the map provided by Wildsanctuary, sounds are represented by icons that visually indicate the type of environment in which a sound was recorded. Listening to the sound requires the user to click on the icon, and then click on the play button that appears in the pop-up window—thus interrupting the flow of soundscape exploration. Despite the attempt to create a more immersive environment by "giving voice to the map" [18], the absence of collaborative mapping capabilities and interaction features in Wildsanctuary prevents users from exploring personal sonic experiences and limits map-based interaction with respect to multimedia content.

We argue that the main limitation of current web maps is that they visually display where information is located but they neither display how information is interpreted nor encourage any interpretation; this is true even when maps are enriched by textual, visual, or other multimedia content. Mapping and visualization of location-based information highly benefit from cartographic semantics supporting expression and sharing of personal perceptions and interpretations [6]. From this perspective, using sounds not only enriches maps, filling an existing gap in collaborative web mapping, it also represents a personal and powerful way to remember, re-encounter, and gain new understandings of everyday space. The use of sounds can encourage users' expressions and interpretations through sensorially richer and more immersive map-based interactions.

3. TheSilence.org

The application for collaborative web mapping presented in this paper is called TheSilence.org and is part of the long-term project Silence of the Lands [17]. Silence of the Lands is a cross-media infrastructure that uses sounds to promote the active role of local communities in the interpretation and management of their urban and natural environments [7].

Based on the belief that sounds are important and personal elements of our encounters with environmental settings [5][16], the goal of Silence of the Lands is to engage people in listening to each other's experiences, connecting with each other's perceptions, and unfolding new understandings of the places in which they live and that they share. This goal is undertaken by enabling people to capture their sonic experiences and then create and share soundscapes of the places where sounds were recorded. Users record sounds by using a Personal Digital Assistant (PDA), which we call a *sound camera*, outfitted with GPS mapping software. Recorded sounds are then uploaded to an online database through the web application for collaborative mapping, TheSilence.org, where they are associated with their owners and placed on a map. The result is an acoustic map that changes over time according to users' perceptions and interpretations of their environmental settings.

Silence of the Lands was initiated by Giaccardi in 2005 and currently involves an international collaboration among the Center for LifeLong Learning & Design of the University of Colorado at Boulder (USA), the University of Brescia (IT) and the Institute of Digital Art and Technology of the University of Plymouth (UK).

4. Creating and Sharing Acoustic Maps in TheSilence.org

TheSilence.org is an application for collaborative web mapping that supports users in the creation and sharing of acoustic maps.

4.1 Overview of Information Architecture

The information architecture of TheSilence.org is shown in Figure 1.

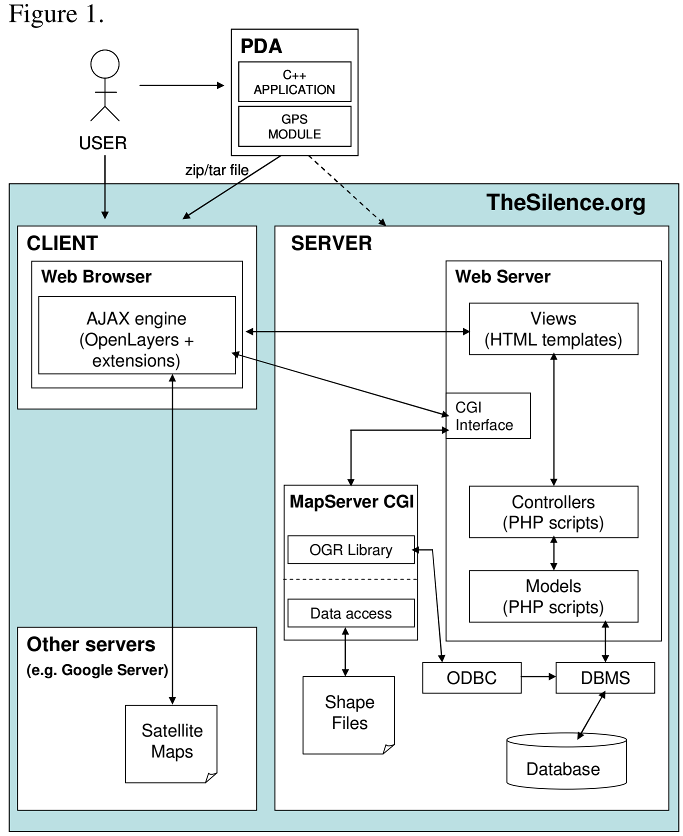


Figure 1: The information architecture of TheSilence.org.

In order to contribute his/her content on TheSilence.org, the user interacts at first with a PDA application that has been properly developed to collect ambient sounds from the natural environment and geo-reference them with Global Positioning System (GPS) data. This application produces a file archive (zip or tar) that contains the recorded sounds as mp3 files and the GPS data related with the sounds, which determine their location in time and space and permit the visualization of sounds on an interactive map. The archive can be easily uploaded through the web application in a MySQL database. During uploading, a validity check is performed on the data, and errors generated by the GPS device are recognized and managed before data registration on the database. A unique identifier is assigned to each registered sound; therefore, if two sounds are recorded in the same geographical position, they are stored in the database as two different records. The visualization of the sounds on the map will then depend on filtering options (see Subsection 4.3).

The database is successively accessed and updated through the web application TheSilence.org by means of two main classes of functionalities: (1) functionalities for managing and representing data, such as user profiles and logging information; and (2) functionalities for interacting with maps and sounds.

The first class of functionalities has been implemented in PHP

according to a Model-View-Controller architecture; it exploits the CakePHP framework in order to guarantee code reuse and maintainability. This part of the application runs on the web server and interacts directly with the database.
The second class of functionalities has been developed as a Rich Internet Application (RIA) and runs on the client. In order to limit page refresh, provide users with an easy access to sounds and associated data, and support efficient map navigation, we have chosen AJAX technology to implement the RIA. We have used and extended the AJAX-based library OpenLayers to visualize the map in the web browser and provide functionalities for map zooming and panning, sound playing, and pop-up activation and filling. OpenLayers acquires satellite images from existing services such as Google Maps and Yahoo! Maps, and then overlaps information layers to build the interactive map. In our architecture, information layers are created through MapServer, a Geographic Information System (GIS) installed and running on the web server. Using MapServer as a Web Feature Service (WFS), the system connects to the database and creates an information layer with geo-referenced sounds. Next, OpenLayers interfaces to MapServer to obtain the information layer and create sound objects (*markers* in the OpenLayers terminology) with which users can interact. Flash MP3 Player is used to play sounds. It is controlled through JavaScript APIs activated on marker events.

4.2 Edit Mode: Collaborating with Acoustic Web Mapping

The geographical position, time, and date associated with sounds are entered automatically in the application during the uploading phase; they are obtained by processing the data in the file archive generated by the sound camera. Then, users can visually associate their own sounds with a color and also annotate sounds with verbal tags and textual narratives. This provides sounds with user-generated metadata about how their owners have subjectively interpreted them and what personal meaning they associate with them.
We expect tagging to be an important feature for personalizing soundscape exploration, just as color rating has proven to be an important activity for the effectiveness of web mapping and visualization [6]. Color rating contributes to the mapping and visualization of what we have called “affective meaning” [6], expressing users’ likes and dislikes in an immediate and personal fashion and encouraging serendipitous map-based interactions. The color scheme adopted is based on the principles of the Abaque de Régnier, a method used today in areas as diverse as human resources, regional planning, and sustainable development to help people express themselves and build shared understanding of common problems [15]. The method uses a color-coded scale that includes dark green, light green, yellow, light red, and dark red. This scale ranges from the most favorable position (dark green) to the most unfavorable (dark red). Additionally, white and black indicate whether the respondent does not have any opinion (white) or refuses to answer (black). For the sake of our project, we use only dark green, light green, yellow, light red, and dark red. White is used exclusively in the Edit mode to indicate to a registered user what sounds he/she has not yet annotated.

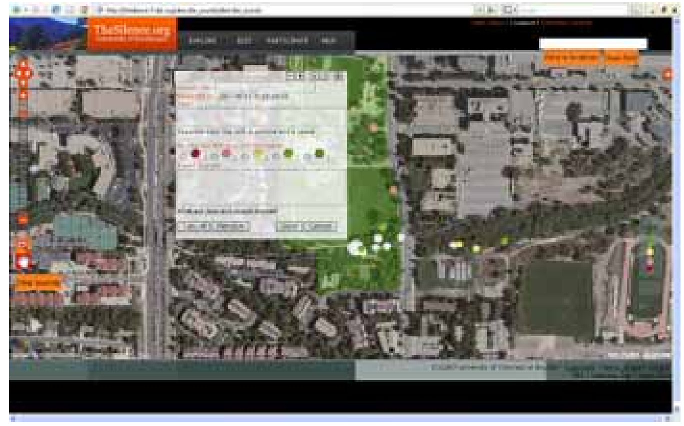


Figure 2: Sound annotation and color rating in the Edit mode.

Sounds are visualized on the map as dots. In the Edit mode, registered users can visualize on the map their own sounds and distinguish between sounds they have already annotated (colored dots) and sounds they have not yet annotated (white dots). Mousing over a dot enables users to “preview” the sound, whereas clicking on a dot selects the sound—which then can be either endlessly reproduced or paused—and automatically opens a pop-up window asking for color rating and sound annotation (title, tags, and sound journal; see Figure 2).
In the acoustic web mapping of TheSilence.org, sound editing aims to recreate the sensory experience of “listening to the land” and to maintain the unique connection to place that is responsible for the meaning one may attribute to a recorded sound. By listening to a recorded sound during the editing process, users can sonically recall their personal experiences. Furthermore, visualizing it directly on the map in relation to previously recorded and rated sounds facilitates the elicitation of users’ environmental knowledge and emotional response with respect to the specific place where the recording was performed.

4.3 Explore Mode: Immersive Map-Based Interaction

Through users’ personal recordings, ratings, and annotations, a shared representation of the acoustic environment is produced in the form of an immersive soundscape in which both registered and unregistered users can seamlessly browse from one sound to another.
From a visual standpoint, the map presents as a satellite image with brilliant green and blue layers for aesthetic appeal and easy detection of urban and natural areas. Street and hybrid views are also available, but the satellite view (with green and blue layers for open space and rivers where geographic information is available) is the system default. On this map, colored dots serve the function of both locating where a sound was recorded and visualizing the interpretation ascribed to that sound. Colored dots visually define both the space mapped by users (geo-location) and the subjective quality of the places users encountered (color rating) [6].
The dots assume different sizes according to the zoom level: the higher the zoom level, the smaller the dot (from a cloud of small dots to the view of a clickable individual dot). This supports the localization and signification of individual sounds, as

well as the immediate visualization of public trends. A single dot expresses a sound location and the user's personal interpretation with respect to the sonic experience captured by the recording, whereas color patterns allow immediate identification at a bird's-view of areas of positive agreement (majority of greens), negative agreement (majority of reds), uncertainty (majority of yellows), dissension (mixed colors), and anomalous positions (isolated red dots within a majority of greens and, vice versa, isolated green dots within a majority of reds).

From a sonic standpoint, consistent with map-based interaction in the Edit mode, the map enables immediate audio streaming. A 60-second snippet of the recorded sound can be played and put into a loop by simply mousing over the dot. Browsing on the map from one sound to another causes sounds to fade in and out, generating the impression of a living soundscape. A sound control panel (play, pause, volume) and relevant information (date/time, recorder, title, tags, description, and users' comments) are provided by clicking on a single dot and accessing the pop-up window associated with the selected sound (Figure 3).

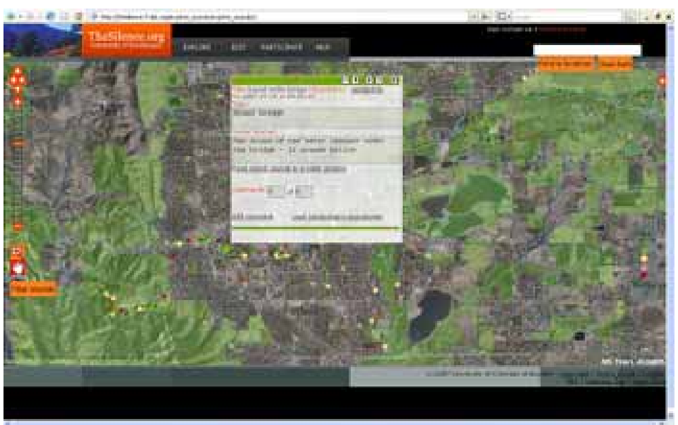


Figure 3: Accessing sound-related information in the Explore mode.

Within this multimedia environment, the Explore mode provides users with the capability to navigate the soundscape by filtering the cartographic output according to several criteria. Users can choose and combine several filtering options by selecting (a) the color of the sounds to be visualized, (b) the users who performed the recordings, and (c) the time and season of the recordings (or date/time in advanced searches), as well as by entering tags of interest. Filtering options become useful also when different sounds have been recorded in the same geographical position: they can be alternatively visualized on the map according to the values of the other attributes (e.g., recording time).

Consistent in both the Edit mode and the Explore mode, the aesthetically and emotionally engaging map-based interaction of TheSilence.org aims to recreate the sensory experience of "listening to the land." In the Explore mode, the unique connection to place responsible for the meaning one may attribute to a recorded sound is explored by allowing several permutations of the cartographic output. These permutations enable multiple readings of the same territory and can be performed (both visually and sonically) at different levels: from the local level of the individual (single sound and single color), to the regional level of social patterns (local clusters of sounds and colors), to the global level of the community (overall trends of sounds and colors with respect to a specific community). In turn, these readings can be conducted at the local level of a specific site, the regional level of a specific topological area, or the global level of a larger geographic area.

5. Pilot Study

In collaboration with the City of Boulder Open Space and Mountain Parks (OSMP) and Water Quality departments, we engaged the local community of Boulder, Colorado, in capturing and sharing sonic experiences for a period of 6 weeks [8]. The goal of the pilot—a joint public program called Community of Soundscapes—was to evaluate the usefulness, impact, and potential applications of acoustic web mapping technology. In particular, we were interested in investigating how an immersive multimedia interaction with geo-located sounds can encourage people to reflect on their perception and interpretation of the natural environment and facilitate the sharing of such perceptions and interpretations with others.

5.1 Methodology

A sample of 20 volunteers (4 males and 16 females) participated in the pilot study. Their ages ranged from 20 to 62 years. They all held higher education degrees and represented varied professional backgrounds, including writers, engineers, scientists, managers, designers, educators, therapists, musicians, and college students. Participants in the pilot were asked to capture their sonic experiences and upload sounds on TheSilence.org, where they could annotate and share them with other participants. They were asked to take at least three sound walks in the period from July 2007 to September 2007 (two mandatory and one of their choosing). A total of 1338 sounds were recorded, and 567 sounds were selected and made available on the web application. We triangulated different kinds of qualitative data collected from (a) three focus groups, (b) two questionnaires (a pre- and a post-questionnaire) and (c) unstructured interviews and direct observations conducted during participants' activities. Section 6 provides findings relevant to evaluate and discuss the importance of sounds in collaborative web mapping. Other and more general aspects of user interface and map-based interaction are briefly summarized in this section and have been more extensively discussed by us elsewhere [6].

5.2 Findings

5.2.1 Capturing and Sharing Sonic Experiences

One of the general themes that emerged from the pilot is whether users felt able to express and share their sonic experiences and, through them, their personal perspectives on the environment. The possibility of associating colors and textual narratives to sounds and immediately accessing them through the map highlighted a broad range of personal connections to place, providing a discursive context for choice and rating of the sounds mapped. In general, pilot results seem to suggest that being able to record and contribute sounds, locate them on the map, rate them, and annotate them by color and textual narratives allow people to capture and share perceptions and interpretations of the environment in which they live in a novel and engaging way.

5.2.2 Exploring the Soundscape

Another general theme that emerged from the pilot is whether users felt interested in exploring and understanding other people's experiences. Generally speaking, participants appreciated the possibility of enjoying sounds collected by other participants and sometimes could visualize in their minds the place where these sounds were recorded. Other participants emphasized the differences in perceptions and interpretations that emerged within the community and stressed the enrichment they gained from these differences. Colors in particular appeared useful to participants, guiding them in map-based interaction. Some comments provided a helpful account of how participants read the map and what aspects were paid more attention. In general, participants tended to explore the extremes (areas with a lot of red dots or dark green dots). In fact, at a bird's-eye view, color patterns allowed participants to identify which areas were more contaminated with traffic and which were quieter, and this identification guided decisions about places to visit.

Overall, the combination of colors and sounds seemed to be particularly effective. Through visual clues, participants were able to notice differences and in general find their own way through map exploration and interpretation. At the same time, geo-location proved useful to make users aware of the dependency between an individual sound, its color rating, and the environmental context of the recording. This aspect made users curious and engaged.

6. Sounds in Collaborative Web Mapping

6.1 Importance of Sounds

Pilot results suggest that sounds are an important and personal element of the everyday environment people experience. In the pre-questionnaire, we asked explicitly: "How do sounds affect your appreciation of the natural environment (or anything, for that matter)?" Some of the answers were significant:

"Without sound, the world would be less connected and intimate. We share in part through sounds. Then we understand and appreciate."

"I enjoy the sounds I encounter in the natural environment, sometimes more than the scenery. I think I feel this way because sounds tell us so much more about a place than we can pick up visually."

These answers seem to express the need for collaborative mapping systems where sounds (not only visual content) can be used to reflect on one's own experience of place, and connect one's personal experience to that of other individuals. Pilot results demonstrate that such activities require an easy and engaging map-based interaction with sounds, similar to the one designed for TheSilence.org. At the question "What features of the web application did you find most engaging?" answers similar to the following were the most common:

"The ease of hearing sounds. By moving the pointer, you can actually 'cruise' along a trail."

"Ability to rollover data points and hear the sound instantaneously … I like that you only had to move the mouse over the icon to hear the sound."

"Definitely the ease of listening. Just rolling over the dots to hear the sounds is quick and easy for exploring."

6.2 Taking Pictures versus "Taking" Sounds

Through the post-questionnaire question "How would you describe the difference between taking and sharing pictures versus taking and sharing sounds?" we tried to understand what kind of experience is afforded by capturing and sharing sonic experiences through our sound mapping technology, and how this differs from creating and sharing pictures. Some answers to this question well highlight the dynamic nature of sounds with respect to pictures and visual content in general. For example, one participant commented:

"A picture is fun to take and share, but sounds are more interactive to me and cause more of a deeply moving response than a still photo."

Another participant commented on the different sense of connection to a place that sounds can evoke compared to pictures:

"Sounds are more intimate ... Pictures feel more detached to me, like we're observers only. Sounds feel more like being immersed in the experience."

Participants generally agreed that sounds are a trigger more powerful than pictures in helping to remember, re-encounter, and gain new understandings of everyday space:

"Sounds have more room for the imagination."

"The sound evoked more memories than pictures."

These answers support the argument that web mapping is much more engaging and fun when different senses are involved, and that sounds play a role in conveying emotions and values often more important than visual content.

6.3 Fostering Novel Practices through Acoustic Web Mapping

The collaborative sound-mapping technology proposed encourages novel interaction experiences with web-based maps, fostering novel phonographic practices of sound recording and sharing that we call "acoustic web mapping." Pilot results suggest that sound (1) is fun, (2) is an important natural resource, and (3) can serve as a critical social indicator. The experimental acoustic web mapping technology used during the pilot makes phonographic practice mobile, location-based, and potentially accessible to anyone equipped with a GPS-enabled mobile device and Internet access ("just in my purse," as a participant indicated). This possibility democratizes sound recording and sharing, and opens the door to a broad range of recreational activities ("sound as fun"). As suggested by the pilot participants, these activities may include mapping one's neighborhood, recording a friend's child speaking, recording a vacation as a way to remember and share one's memories with friends and family, and creating artistic sound recordings.

Sound can also be considered "as a natural resource," as suggested by the naturalists and educators involved in the pilot as stakeholders. In this context, the same technology can be used as a tool for amusing and spontaneous forms of informal education. For example, collaborative sound mapping can support the interpretative experience of natural parks visitors, helping them connect emotionally to places, bringing them closer to wildlife that they sometimes cannot even see, and aiding them in understanding natural resources [1]. Additionally, it can also support longitudinal mapping and monitoring of fragile and critical wildlife areas, thus providing naturalists and land managers with a tool for assessment and decision-making and virtual nature-based tourism ("ecotourism") offerings for areas closed to the public.

Finally, as prospective stakeholders suggested on the basis of pilot results, sound can also serve "as a social indicator." The same technology can enable administrations and public organizations to visualize how communities perceive and interpret their environmental settings and how these public trends change over time. Collaborative sound mapping can represent here a participatory and democratic way to engage communities on issues such as land management, urban development, environmental justice, and noise pollution, with the goal of supporting public awareness and social dialogue over sustained periods of time. Studies have demonstrated that people do not act unless they connect with the problem emotionally [14]; collaborative sound mapping is an intimate way to connect emotionally to places and develop a sense of ownership and commitment toward the environment.

7. Conclusions and Future Work

This paper has focused on how to enable users to contribute and share sounds. The work fills an existing gap in collaborative web mapping and highlights the importance of sounds in human experience with respect to map-based interaction. The system described, TheSilence.org, brings together sounds and affective mapping and visualization of geo-located content [6] with the goal of generating an aesthetically and emotionally engaging map-based interaction that recreates the unique and sensory experience of "listening to the land." Through this approach, users can sonically recall or imagine one's personal experiences by listening to the recorded sounds. Visualizing and contextualizing the sound directly on the map facilitates the elicitation of users' environmental knowledge and emotional responses with respect to the specific place where the recording was performed.

The pilot study results presented and discussed in the paper suggest that sounds are an important element of the everyday environment, and that sounds can effectively help reflect one's experiences of place and connect to other people's experiences. These activities require an easy and immersive map-based interaction, similar to the one users experienced through the described system. Pilot results also highlight the dynamic nature of sounds with respect to pictures and visual content in general, suggesting that sounds are a trigger more powerful than pictures in helping people to remember, re-encounter, and gain new understandings of everyday space. Collaborative web mapping is more engaging and fun when different senses are involved, and sounds in particular play an important role in conveying emotions and values. Finally, pilot results suggest that what we have called "acoustic web mapping" is likely to foster novel interaction experiences and novel practices of sound recording and sharing.

Future work includes an exploration of the social networking aspects of acoustic web mapping, with the goal of creating a platform that can be used in different ways and elicit different forms of community engagement. Further evaluation studies will include a collaboration with the Off-Campus Student Services and the Restorative Justice Program at CU-Boulder to assist college students in understanding and experiencing the impact their own and their friends' recreational activities may have on the neighborhoods in which they live. The project, which is scheduled for 2009, will provide a second case study in the different domain of environmental justice.

Acknowledgments

The authors thank Ilaria Gelsomini, Francesca Pedrazzi, Guido Pollini, Gianluca Sabena, and Chris Speed for their contribution to the development of the collaborative web mapping application. This research was supported by: (1) the National Science Foundation grant IIS-0613638, (2) CU-Boulder's Outreach Committee grant 2006-2007, (3) the Università degli Studi di Brescia and (4) the Arts Council England.

References

[1] Caputo, P. Sound Interpretation, *Legacy*, 2007, 18(1): 10-11.
[2] Celentano, A., Mussio, P., Pittarello, F. The Map is the Net—Towards a Geography of Social Relationships, *Workshop on Map Based Interaction in Social Networks (MapsInNet07), INTERACT2007*, Rio de Janeiro, Brazil, September 2007.
[3] *CommunityWalk*, http://www.communitywalk.com/
[4] *Flickr*, http://flickr.com/map/
[5] Gallego-Juàrez, J. A., Acoustics for the 21st Century, *Proc. ICA 2007*, Madrid, Spain, October 2007 (CD-ROM).
[6] Giaccardi, E., Fogli, D. Affective Geographies: Towards Richer Cartographic Semantics for the Geospatial Web, *Proc. AVI 2008*, Naples, Italy, May 2008, 173-180.
[7] Giaccardi, E., Palen, L. The Social Production of Heritage through Cross-Media Interaction: Making Place for Place Making, *International Journal of Heritage Studies*, 2008, 14(3), in press.
[8] Giaccardi, E., Freeston, J., and Matlock, D. Community of Soundscapes: Results and Evaluation, *Internal Report,* University of Colorado at Boulder, September 2007.
[9] *Google Earth*, http://earth.google.com/.
[10] *Google Maps*, http://maps.google.com/.
[11] Krause, B., Giving Voice to the Map: The "There" in "Where," *Where 2.0 Conference*, May 29-30 2007, San Jose, CA.
[12] *YouTube*, http://www.youtube.com/
[13] Leclerc, Y. G., Reddy, M., Iverson, L., Eriksen, M. The GeoWeb — A New Paradigm for Finding Data on the Web, *Proc. ICC2001*, Beijing, China, August 2001.
[14] Pooley, J. A., O'Connor, M. Environmental Education and Attitudes: Emotions and Beliefs Are What Is Needed, *Environment and Behavior*, 2000, 32(5): 711-723
[15] Régnier, F. *L'Enterprise Annonce la Couleur: Gerer les Divergences Leviers d'Efficacite Creatrice*. Paris: Les Editions d'Organisation, 1984.
[16] Schafer, R. M. *The Tuning of the World: Toward a Theory of Soundscape Design*. Philadelphia: University of Pennsylvania Press, 1980.
[17] *Silence of the Lands*, http://www.thesilence.org/
[18] *Wildsanctuary*, http://earth.wildsanctuary.com/
[19] *WikiMapia*, http://www.wikimapia.org/

Towards Multi-provider LBS Visual Portals

Silvia Gordillo[1,2], Robert Laurini[3], Catalina Mostaccio[1], Françoise Raffort[3], Sylvie Servigne[3]
[1] LIFIA, Universidad Nacional de la Plata, Argentina.
Gordillo@lifia.unlp.edu.ar, Catty@lifia.unlp.edu.ar
[2] CICPBA, Argentina.
[3] LIRIS, INSA de Lyon, France.
Robert.Laurini@insa-lyon.fr, Francoise.Raffort@insa-lyon.fr, Sylvie.Servigne@insa-lyon.fr

Abstract

Location-Based Services (LBS) are more and more commonly accessed through PDA or smart phones. As portals deriving of services delivered by a unique provider are easy to define, the objective of this paper is to address the problem of generating portals allowing the accesses of services provided by several companies. In other words, the goal is to generate automatically a unique visual multi-provider portal in a context of portal interoperability. After the analysis of examples in order to define more clearly spatial concepts such as reference location and scopes of services, various organizations of portals are examined, namely text-based, icon-based, map-based and street-view. We conclude this paper by recommending map-based and street-view systems, together with XML extensions which must supplement LBS metadata in order to generate those visual portals.

Keywords: *Location-Based Services, GUI for PDA, portal interoperability, GIS, Visual portal*

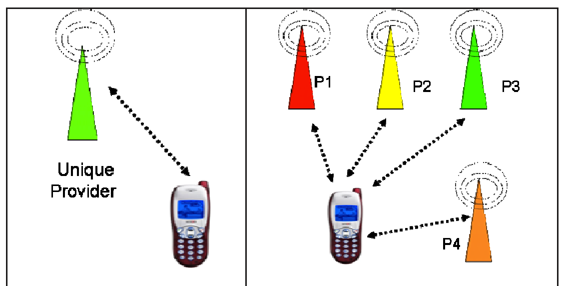
1 – Introduction

The huge majority of studies regarding Location-Based Services (LBS) concern a unique provider of services; however comparing to the evolution of mobile telephones, it is important to begin studies relative to multiple providers; in other words, even if norms and standards will emerge in the nearby future, interoperability of location-based services will be an important theme of research.

One of the problems we have to face is the problem of organizing portals. This problem can be stated as follows (Figure.1): "*Suppose I have a list of p providers, each of them providing s_p LB services, how to organize a portal such as all LB services can be presented in a friendly way*"?

Spiekermann [6] defines LBS "*as services that integrate a mobile device's location or position with other information so as to provide added value to the user*". And later, he distinguishes two categories, namely, person-oriented LBS (such as looking for nearest friend) and device-oriented LBS (such as car tracking). In Wikipedia [7], LBS are defined by a quite similar text as "*information services accessible with mobile devices through the mobile network and utilizing the ability to make use of the location of the mobile device*".

Let us remember that some LB services can be either push, which means that they are offered to the user prior his/her consent, or pull, which means that the service list is the result of a query linked to the profile of the user. In this study, we will overall assume that the user has given his profile, and several services coming from several providers respond to this profile. In addition perhaps emergency push services can also be present, for instance to announce floods or tornadoes when necessary. Moreover, even if SPAM services may exist, they will not be considered in this study.



(a). The present situation: a unique LBS provider	(b). The nearby future: a single PDA is connected to several LBS providers

Figure 1. *From unique to multiple LBS providers.*

Indeed location means several things for any LBS. For instance municipal services are linked to a city, and are no more valid in a neighboring city, whereas a shoe shop does not consider city boundaries since it will attract as many customer as possible. From those basic examples, two concepts are emerging:

- the **reference location** which corresponds to the specific location of the service's provider; for instance the shoe shop itself or the administrative buildings,
- and the **scope**, that is to say the area in which the service is offered, for the city, within the

city boundary, or for the shop its marketing zone.

This paper will address this problem and present some solutions. After having examined a few examples, several portal organizations will be studied, together with some graphic semiology remarks.

An initial remark is if there is no interoperability between portals, a PDA linked to several providers must assign a window per portal as exemplified Figure 2.

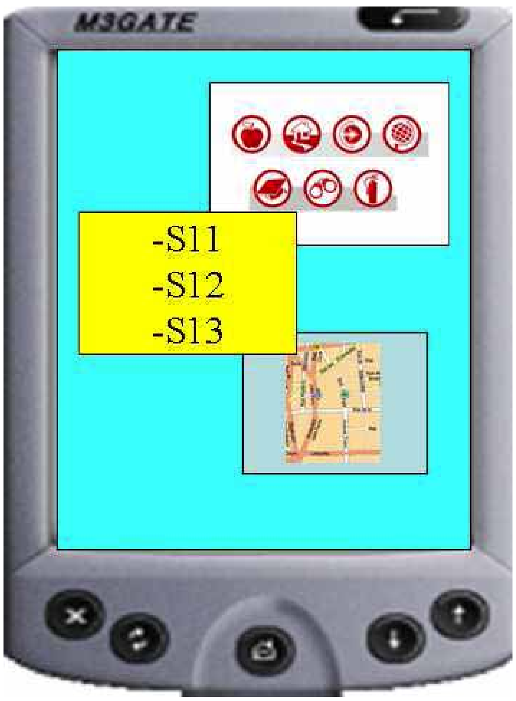


Figure 2. *PDA linked to several service providers assigns one window per provider if no portal interoperability is envisioned.*

2 – Analysis of some examples

Instead of presenting a formal theory about multi-provider LBS, in order to design portals, we will consider a few examples of applications in which services can be requested. In order to illustrate the concerns, we will explain each application in terms of their:

- **Reference location:** specific location of services' providers,
- **Scope:** locations where services must be physically offered,
- **Time:** time in which services are available, and
- **Visualization:** the way in which the user perceives the result.

2.1 – Municipal Administration

We consider a municipal administration which offers different kinds of services to citizens in a city. The city administration may be distributed in many buildings related to several addresses in which different services are provided. The scope of services provided by the Municipal Administration is limited to its jurisdiction; so the user will have services available only in this defined scope. These services in turn will be provided in the hours in which the city hall is opened. The mobile device will show the location of the user, the location of each building belonging to the city hall and the services provided for each of them. In this case we have:

- Reference location: addresses of municipal buildings,
- Scope: the whole city (crisp area),
- Time: opening hours,
- Visualization: one icon per department/per building.

2.2 – Real Estate Agent

This example considers a real estate agent offering services for selling and renting apartments, houses and commercial premises. The scope corresponds to the locations of each apartment, house or premises the agency has to rent or sell. Services are available in the hours in which the agency is opened. Regarding visualization, a map with one icon per apartment, house or premises will be displayed with different representations depending on the nature of the concerned object. For each one of these objects additional information about the particular characteristics or for example, schedules to visit them can be shown.

- Reference location: address of the office,
- Scope: several locations of apartments, etc.
- Time: opening hours,
- Visualization: one icon per object.

2.3 – Meteorology

Meteorological services are very common and offered by different providers in several places. In this case we are interested in the climate around the zone where the user is. Of course, the user may ask for information like temperature, humidity, rains, etc. in any time and even in another place out of his neighborhood; in this case the scope could be considered as the whole Earth. The visualization will show a set of animated icons (like the TV forecast) in a continuous space. In order to bind this space we propose to limit it to 1 km^2 around the user.

- Reference location: meteorological station, or headquarters,
- Scope: the neighborhood of the user,
- Time: any time, present, past or future,
- Visualization: animated icons.

2.4 – Emergency information

In this case we suppose there is a provider who offers services in emergency situations such as natural disasters, catastrophes, etc. These kinds of services are generally "push"; the user receives the information when he/she is entering in the dangerous zone.

The scope will probably be a zone with fuzzy boundaries established according to the type of event; this fact will be also reflected in the visualization. Different labels of security limits could be defined, for example with different colors.

- Reference location: the headquarter of the agency,
- Scope: the epicenter of the zone,
- Time: when necessary,
- Visualization: emergency zone.

2.5 – Public transportation

Services in the context of public transportation are very common and useful not only for local users but also for tourists. Information about bus lines, bus stops, and places where buy tickets are offered by the provider. The scope of services is the locations representing bus lines, stops etc. represented as a set of scattered points and interconnected lines. The time when services are able depends on the schedule of the transport company. The visualization will be a map showing the classical transport network.

- Reference location: the headquarters of the transport company,
- Scope: the various bus lines, bus stops, points of selling tickets,
- Time: when the services are available,
- Visualization: bus stops, line directions (scale).

A variant can be a sightseeing tour in which the scope is the bus circuit that is a closed polyline generally linking the various landmarks of a city.

Surely, some other examples can be studied and analyzed with those criteria, but from those examples, we can state the following:

- The reference location which appears perhaps primarily as the more important aspects, sometimes is not so easy to define (weather forecast) or not important for the user (headquarters of a public transportation company).
- Those reference locations can often be modeled as 2D points.
- The scope can be geometrically defined by several ways, as set of points (real estate), or a network (public transportation, a crisp area (municipality), a fuzzy area (marketing zone, see Example Figure 3) or even a very huge continuous space (meteorology) which must be limited to the user's neighborhood for practical reasons.

Figure 3. *Example of a marketing area of a shop represented as a fuzzy set; the curves represent the various membership degrees at various percentiles, the first one 10 %, the next one 20 %, etc.*

Another aspect is linked to the PDA's screen displaying a spatial window around the user. Depending on the scale this spatial window will represent a territory. Only services the scope of which intersects the spatial window will be considered and laid out. For instance, suppose you are looking for a flat to rent, and there is one at the vicinity (displayed in the spatial window); for this apartment you can easily get information; but when you want to visualize the realtor's address (reference location) the spatial window will be moved accordingly, in order to find the adequate scale showing three locations, user, flat and real estate agency.

After having analyzed those examples, let present some considerations for the designing of a multi-provider portal.

3 – From non-integrated to integrated portals

Designing portals and especially visual portals is a complex task, because we need to decide what kind of metaphor we want to use and how to continue the metaphor without problems of consistency. We will first analyze non integrated portals, and then detail some integrated portals.
In [2], several kinds of portals were examined; among them we can distinguish

- text-only portals, in which services are given by names,
- icon-based portals, in which services are illustrated by icons or thumbnail images,
- map-based portals, in which all services are located in a single map.

However, the previous portals were designed for a single provider. To study the multi-provider case and to illustrate the concepts, let us take an example of three providers offering each of them several services.

- Provider 1 (generalist):
  - Yellow pages,
  - Nearest restaurant,
  - Nearest ATM.
- Provider 2 (local information):
  - Museum opening hours (HP),
  - Shoe shop sales,
  - Chinese restaurant,
  - Italian Restaurant,
  - Church website
- Provider 3 (municipal information):
  - Local event calendar (Sports, Culture, etc.),
  - Located municipal services,
  - Weather forecast.

3.1 – Without/before integration

Without integration, the PDA's screen is divided into sections, each of them being assigned to a single provider. Figure 4a gives the principles of organization, whereas Figure 3b gives an example illustrating various types of portals which must be integrated into a single portal.

(a). The PDA's screen is split in sections, each of them assigned to a unique provider	(b). Example with different types of portals, text-only, icon-based, map-based

Figure 4. *Example of multi-provider LBS portals without integration.*

3.2 – Integrated portal

As previously explained, the objective is to design a unique portal integrating the various component portals into a single consistent portal. However if among the list of services, there is one service which is proposed by several providers, it will be presented several times letting the user the choice (maybe with different prices). The key-idea to design such an integrated portal will be based on the Shneidermann's mantra [5] "*Overview, zoom-and-filter, details-on-demands*", in which for our case:

- "**overview**" will correspond to the unique portal allowing the accesses of all services; here, according to this mantra, we need to propose a global organization aiming at connecting to each of all services, whatsoever the provider must be;
- "**zoom-and-filter**" refers both to geographic and semantic zooming and filtering; in other words, the list of provided services must be reduced;
- "**details-on-demand**" refers to additional details when a subset of services is selected; maybe this subset could a reduced to a single service.

Let us now examine various possibilities of portals, textual, iconic, cartographic and street-view.

a/ Text-based

With this presentation all service names are displayed, for instance either by alphabetic order, or ranked according to some preferences as given in the user's profile.

(a) Services ranked by alphabetic order.	(b) Services ranked by user's profile order.

Figure 5. *Examples of text-only multi-provider portals.*

If the same service is proposed by several providers, the name of the provider can be concatenated to the name of the service.

The main advantage of this layout is its completeness and its easiness to be generated. However, the main drawbacks are:

- The length of the list can overpass the size of the screen, essentially if we want to get a readable list,
- There is no indication of the scope of the information,
- Spatial characteristics are not used,
- All scaling aspects are lost, for instance no in- or out-zooms are possible.

Finally, this solution is an easy one to generate, but the main characteristics of LBS are not used.

Moreover, Shneiderman's mantra is not followed because it looks semantically complex to zoom and filter this list.

b/ Icon-based

This solution implies that each service is assigned a very understandable icon. In case when the same service is offered two times, the same icon will be presented twice. To distinguish them, perhaps the logo of the provider can be added.

Figure 6. *Example of services in which reference locations are mapped with icons, shapes corresponding to services and colors to providers.*

Since icons are often difficult to understand, this solution is not recommended. In this paper, we will not continue to study this solution.

c/ Map-based

In this case, a base map is used in which all LB services can be laid out in both senses of reference locations and scopes.

When reference locations exist, we can easily position the service on the map. Whereas mapping a scope implies the description of the zone above the base map.

Here, in- and out-zooming are very easy to perform and the main characteristics of LBS are used.

Figure 6 illustrates an example in which each service is assigned a shape or an icon whereas colors can distinguish providers.

In the next section, other issues concerning this solution will be developed.

d/ Street-view

Figure 7. *Using street-view as a way of presenting services for a pedestrian.*

Another interesting approach is to use street-views as exemplified in Figure 7. This style of presentation is very interesting for pedestrians, especially for those who have difficulties when reading maps. However, only reference locations which exist along the street can be mentioned. A special algorithm must be created taking the perspective of the street-view into account in order to position exactly the services.

One of the main difficulties resides in the exact location of a service. Imagine that the reference location of a small service (for instance an apartment to rent) is located at the fifth storey of a building; do we have to position the service at the fifth storey or at street level? We do recommend in this case positioning all services at street level.

Let us mention that this kind of display is very interesting in picture-aided navigational systems designed especially for pedestrians as detailed by (Laurini et al, 2008) [3]. Inn this case, the picture can be supplemented with icons for LB locations.

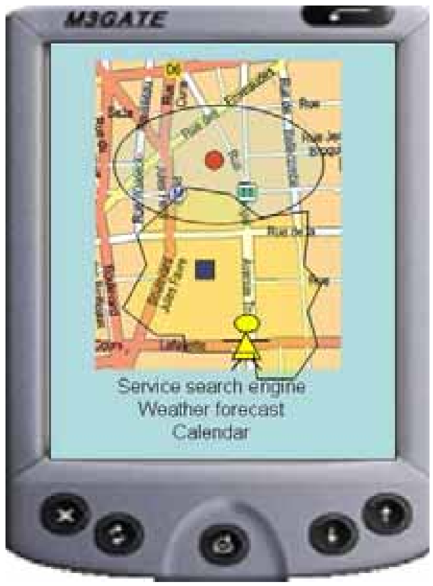


Figure 8. *Example of visual representation with reference locations and scopes.*

4 – Elements of graphic semiology

As reference locations are punctual, scopes can have different mathematical representations. A possibility could be to fill the scope area with graded colors with transparency such as the base map can be always readable. Such an example is given Figure 8.

An interesting solution could be to use the idea of Burigat and Chittaro (2007) [8] to put a sort of signal along each icon to show the relevancy of the displayed LB services.

5 – Conclusions

The goal of this paper was a preliminary study concerning the design of portals in case of multiple companies providing each of them several services. Two concepts were emphasized, reference location and scope. In this paper, we advocate that a unique visual portal can be an adequate solution allowing accesses to all services, whatever may be their respective providers. The idea of using either a map-based or a street-view solution instead of a text-only solution is very promising.

As reference locations are easy to depict and scopes can having more complicated mathematical shapes, adequate representations must be designed without altering the understanding of the base map; a solution based on graded and transparent colors filling the scope area seems an interesting answer. But for that some complementary cognitive studies must be carried out.

In order to realize the correct positioning of reference locations and scopes, information regarding geometric shapes must be provided in the service metadata. An extension of XML perhaps based on GML [4] can be an excellent candidate to describe the issues.

To conclude this paper, let us say that we need experimentation and also cognitive studies to propose a complete definitive system.

References

[1] Burigat S., Chittaro L., (2008) "*Interactive visual analysis of geographic data on mobile devices based on dynamic queries*", Journal of Visual Languages and Computing 19(1), pp. 99-122.

[2] Laurini R., Servigne S. (2007) "*Visual Access to City's websites: A Challenge for PDA's GUI*". Published by "*Journal of Visual Languages and Computing*", 18 (3), pp.339-355, June 2007.

[3] Laurini R., Gordillo S., Raffort F., Servigne S., Rossi G., Wang N., Fortier A. (2008) "*A Primer of Picture-Aided Navigation in Mobile Systems*". In proceedings of the SDH 2008 "International Symposium on Spatial Data Handling", Montpellier, June 23-25, 2008.

[4] Ogc: http://www.opengeospatial.org/standards/gml

[5] Shneiderman, B (1996) "*The Eyes Have It: A Task by Data Type Taxonomy for Information Visualization*". In Proceedings of the IEEE Symposium on Visual Languages, Washington, IEEE Computer Society Press, 1996, pp. 336-343.

[6] Spiekermann, S., (2004) "*General Aspects of Location Based Services*", in: "*Location Based Services*" edited by Agnes Voisard and Jochen Schiller, Morgan Kaufmann, San Francisco, pp. 9-26.

[7] Wikipedia http://en.wikipedia.org/wiki/LBS

Metaphorical Cartography for Knowledge Creation and Sharing

Augusto Celentano[1], Piero Mussio[2], Fabio Pittarello[1]
[1]Dipartimento di Informatica, Università Ca' Foscari Venezia
{auce,pitt}@dsi.unive.it
[2] Dipartimento di Informatica e Comunicazione, Università di Milano
mussio@dico.unimi.it

Abstract

In this position paper we discuss the role of geographical metaphors as tools for representing, organizing and accessing heterogeneous shared information, typical of the Web 2.0. Metaphors represent concepts and relations of a knowledge domain with symbols taken from another domain, usually simpler and more immediate to understand. The paper discusses issues related to the use of metaphorical cartography for creating and sharing knowledge.

1. Introduction

The most visible difference between the old world wide web and the so called Web 2.0 is the progressive fading of separation between information producers and consumers, that mimics the lack of distinction between writers and readers in the historical Ted Nelson's vision of hypertexts [6]. As in classical hypertexts, the problem of finding the right information or the right document is central in Web 2.0; users have different attitudes towards information organization and classification, stemming from individual needs and experience, hardly fitting rigid and predefined conceptual schemas.

Indeed, Web 2.0 is only the most recent and evident scenario in which heterogeneous shared information plays a central role. The issues we shall discuss in this paper apply as well to collaborative systems, cooperative design environments, social communities, etc. In such scenarios a problem arises for an author: how to organize the knowledge being produced, and how to describe it to other humans so that the knowledge can be shared with them and they can be engaged in the process of knowledge production. One possible solution is to identify a metaphor [2]; to this end, two points should be considered.

First, humans grow and live in a common real space. They all experience reasoning and representing space to survive [4]. All users experience movement in real territorial space, moving around in their daily life; they experience the difficulties of climbing a mountain or the fatigue of covering a long distance. Quoting Kuhn, "space is fundamental to perception and cognition because it provides a common ground for our senses as well as for our actions ... space is not just any domain of experience, but the most important one, and as such uniquely qualified as a metaphor source" [5].

Using a map, a description of a territory to move in, is also a common experience: this experience is however mediated, and depends on the culture, technical skill and ability of both the author of the map and the user. The history of cartography shows how the progress in science and discoveries has produced maps that are not only more complete and accurate in the Earth representation, but also more apt to address specific goals, due to users agreement on symbols and conventions. On the other side, most of the difficulties in interpreting ancient maps come from the gap between the historical and cultural environment of the map designer with respect to the archeologist discovering them Figure 1 shows a Babylonian world map of 600 b.C. on the left, and the archeologist reconstruction on the right. The interpretation rised from archeological findings [1].

Second, technology makes new tools available: the emergence of geoweb, offering users new possibilities of

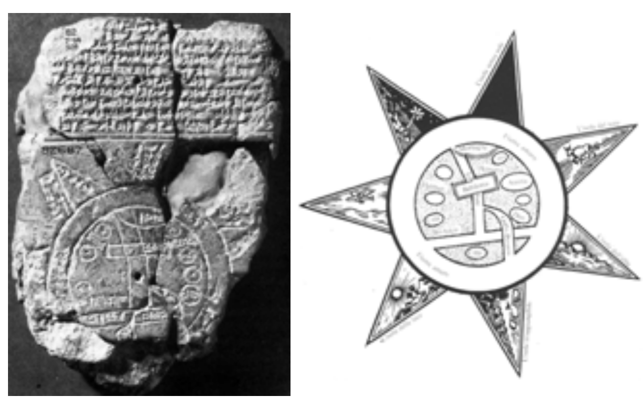


Figure 1. A Babylonian world map (600 b.C.) and the archeologist reconstruction.

interaction and knowledge creation and stimulating the rise of new user capabilities. Web services such as Google Maps enable people to add content to a map and to share it.

In the Web 2.0 world maps have become one of the favorite means to represent information placement. As geographic maps permit the understanding of a land, helping people to understand the spatial relations in it, to move and to meet, hence to communicate, abstract and metaphorical maps permit understanding of a common semantic domain, communication and sharing of knowledge.

2. The raise of a new metaphor

In the past, the metaphor of knowledge as a territory and the use of maps as a new medium for representing knowledge organization became widely popular [2]. Knowledge can be represented as a territory populated by information items, identified by metaphorical landmarks that people recognize based on common sense, human experience and specific cultural biases. Maps become the boundary objects for people to discuss on knowledge evolution.

In the digital world, maps evolve from mute displays of a geographic reality to digital interactive and pro-active tools, whose content develops in time and whose physical appearance can be determined on the fly at use time. Users themselves may become (co-)authors of the map, directly contributing to its evolution. Maps become the medium to update and enrich knowledge.

The basic idea is to mimic the old technique of mentally associating knowledge to spatial and geographical entities to remember and retrieve its chunks [8]. In analogy, knowledge is organized as a virtual geographical space but now stored in a system outside the mind. Digital tools allow users: a) to organize their growing knowledge as a virtual space, b) to communicate this organization to other humans as maps representing the virtual space from different points of view, and c) to contribute new content to the space, hence to the map.

A human can thereafter guide other humans in the exploration of stored knowledge organizing trips through the space, described by tracks on the map, commented by texts (narrations) illustrating the different fragments of knowledge being visited and their properties.

If we focus our attention to bi-dimensional maps that can be viewed using a standard web browser, we can count a huge number of proposals, from the well-known Google Maps to other interesting alternatives [11, 15]. The reasons for such a success stem from the need of a shared classification of knowledge, from the evolution of the web and of the devices toward mobility and ubiquity, and from the user participation.

Knowledge classification. In a situation where people use different personal schemes for classifying an ever-growing amount of information a geographically inspired representation can be seen as a stable and sharable representation of classification methods, being strictly associated to the physical concreteness of the human being and of the territory where he/she lives.

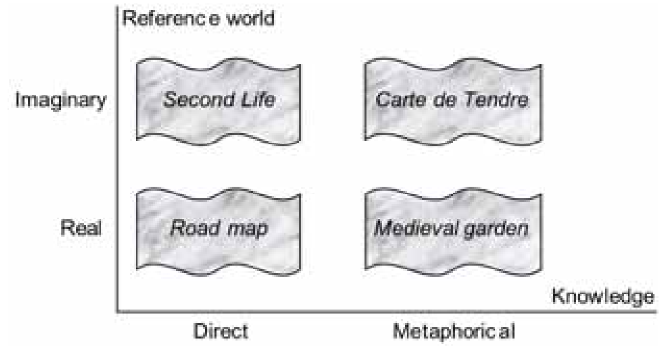


Figure 2. The map classification space

Mobility and ubiquity. In the Web 2.0 age the role of mobility with personal devices is greatly increased; people appreciate the possibility of having appropriate information associated to the location they are currently visiting.

Role of the users. As stated in the introduction, the difference between the information producer and consumer is fading. New hardware and software devices allows users to produce information, to associate it painless to the geographical location and to share it with other users.

3. Between reality and metaphor

We go a step further, and distinguish between a symbolic, thematic annotation of a real land, and a metaphorical representation of knowledge mapped through the symbols of geography; the two representations are end-points of a continuous range where classification is fuzzy, and combination of realistic and metaphoric elements can co-exist.

Figure 2 displays such end-points on the horizontal axis. The label *direct* denotes the case where the map components denote the properties of real objects having a specific location on the territory. The label *metaphorical* denotes the case where the map components are mapped to objects and concepts that are not located on a land, but may benefit from a representation based on similarity with land features. The vertical axis is related to the type of world making the map basis; *real* and *imaginary worlds* can be used as territories on which the map is drawn.

While the plane defined by the two axes can host a continuous range of map types, Figure 2 displays four cases deriving from the combination of the two end limits.

The *Map of Second Life* [12], the artificial world where people may meet in a virtual land using a virtual counterpart, is a good representative of imaginary worlds; most of

Figure 3. A metaphorical medieval garden

Figure 4. The Map of Tenderness (English translation)

the map contents are not metaphorical, because they represent objects of the virtual land, such as a (virtual) building.

The Road *map* is the classical example of direct representation of a real world, whose goals is to assist travelers by showing them a symbolic scaled representation of what they are seeing (or will see) when moving in the real environment.

The map of a *medieval garden*, the so-called *hortus conclusus* (Figure 3), is anchored to the real world, but the elements pictured have a metaphorical meaning: the well, the fountain or the tree placed in the garden center represents the tree of life, i.e., the source of knowledge; the four paths that divide the garden enclosure in four quadrants represent the four angles of the universe. The garden is therefore a metaphor of the universe and of the spiritual life, even if represented through concrete land objects.

The *Map of Tenderness* shown in Figure 4 (*Carte du Pays de Tendre* in the original French version) is the first and best known of a series of metaphorical maps representing an abstract domain [7, 13]. It appears in a novel written by M.me de Scudéry in middle XVII century, and represents the "stations of love" along a path of increasing degrees of affection. The path starts at the bottom of the map in the "*New friendship*" town and leads to tenderness through a smooth river and straight paths crossing small villages, each representing a good feeling. The path is also populated by dangerous feelings to be avoided, such as the forbidden "*Rock of Pride*", leading to adverse moods like *indifference* and *enmity*, represented by a lake and a sea at the two sides of the map. The metaphor is impressive on the emotional side; we shall discuss it in some detail in the next section.

Metaphorical maps, located in the upper right quadrant of Figure 2, are in our view the most interesting ones, since they match the human ability to perceive space and environment with the relations linking entities and concepts in the represented domain. Knowing the composing elements of the map (the base vocabulary of the metaphor) and their layout we are able to interpret the knowledge subsumed by the geographical metaphor.

The correspondence between a representation and the represented knowledge, however, can be satisfied at different degrees; consequently the interpretation of the metaphor can be more or less easy and safe, as art and literature (mainly the classical poetry) put into evidence. To cite a simple case, Dante's *Comedy* is full of metaphors and allegories that have engaged critics for centuries, in many cases without coming to a common interpretation.

4. Metaphor structure, definition and interpretation

A metaphor relates two different words, in the case of our interest a domain of knowledge and a geographical one, with the aim of clarifying some aspect of the first through the features of the second.

The entities of a knowledge domain are classified in *concepts* and *relations*. Concepts can be organized in classes that represent their common aspects.

The author of the metaphor establishes a map from the concepts, classes and relations of the knowledge domain into elements and relations of a geographical space. The association is established because the author identifies some features of the geographical element as explicative of some feature of the knowledge element, he wants to highlight. To represent his view, the author represents the metaphor representing the metaphorical geographical space as a map, drawn following a set of abstraction rules. However, the name attributed to a map entity is the name of the knowledge domain entity associated to the geographical entity being represented. A reader of the map interprets it by first trying to associate the map element to geographical entity and then to derive the author view on the original knowledge domain entity. We discuss an ex-

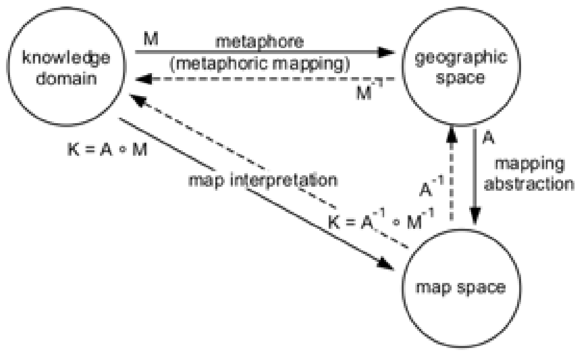


Fig.5 metaphorical map construction (solid lines) and interpretation (dashed lines)

ample to clarify these processes.

4.1 The metaphor of the Map of Tenderness

According to M.me de Scudéry's view, in the knowledge domain of love and affection, examples of concept classes, concepts and relations are:

Concept classes and instances

cause of tenderness (*disposition*, *esteem*, *gratitude*)
origin of tenderness (*new friendship*)
good feeling (*kindness*, *submission*, *attentiveness, ...*)
bad feeling (*negligence*, *inequality*, *oblivion*, ...)
obstacle (*pride*)
tenderness (*tender by disposition*, *tender by respect*, *tender by gratitude*)
failure (*enmity*, *danger*, *indifference*)

Relations

cause of tenderness flows from *orgin of tenderness* to *tenderness*
cause of tenderness leads to *tenderness* through *good feeling*
cause of tenderness leads to *failure* through *bad feeling*
good feeling connects origin of *tenderness* to *tenderness*
bad feeling connects origin of *tenderness* to *failure*

The author associates concept in this domain with geographical entities to express her view on each concept or relation. In her vision, the *tenderness causes* flow across increasing feelings, leading from a new friendship to three types of tender, three different affection states. *Tenderness causes* have a behavior similar to a flow of water, a river, while the different affection states are represented by solid, large cities where the evolution of feeling may stop for a while or forever . Some feelings are smaller cities, which are not meant as destinations but as intermediate steps leading to tenderness. Indeed, the disposition river flows smoothly and directly, because natural disposition needs no other feeling to reach tenderness. Gratitude and esteem, on the contrary, are grown by intermediate feelings that increase the tenderness, but also reveal risks to deviate from the right path with bad feelings, leading to indifference and enmity.

Tenderness is the spiritual destination of good feelings, but since the human nature is not only spiritual, it must be controlled, otherwise an excess can lead, through the dangerous sea of senses, to the unknown lands of the passion.

The metaphorical correspondence between concepts and relations in the spiritual domain has a counterpart in the map, as represented in Figure 5. The solid lines represent the process of metaphor definition, the dashed lines the metaphor interpretation.

4.2. Metaphorical map interpretation

Interpretation of a metaphorical map requires an inverse reasoning: an observer looks at the map and interprets the symbol, first relating them to their meaning in a geographical metaphorical space. By interpretation, the observer infer that in the metaphorical space M.me de Scudery conceived a river named Disposition, leading to the town of Tender by Disposition, and two main paths of villages, labeled with names of feelings, leading to Tender by Esteem and Tender by Gratitude. Deviations from the main paths cross villages labeled with names of bad feelings leading to the waters of Indifference and Enmity. In a second step of reasoning, the reader recognizes the original meaning of this scenario, representing the stages leading from a new friendship to tenderness, to indifference or to enmity, according to the feeling expressed.

4.3. A different metaphor built on the same geographical space

A different metaphor on the same land has been suggested by Jean-Luc Michel. In his web site [9] he proposes, rewriting the content of the *Map of Tenderness*, a map of *Research Tenderness* (*Carte du Tendre de la Recherche*), in which the concepts of a sentimental domain are replaced by concepts related to the scientific research: the *Sea of Speculation*, the *Lake of Empirism*, the villages *Demonstration*, *Imagination*, etc., marking the itineraries leading from *New Research* to *Scientific verification* (*Crible de Scientificité*), as depicted in Figure 6.

4.4 Topology and metrics

The quality of representation/interpretation is also influenced by the level at which the metaphor is designed, i.e., by the use of metric and topological properties together with the basic symbology.

At the basic level a map can be read as a collection of symbols, including neither topological nor metric con-

Figure 6. A metaphor about scientific research on the geographical shape of Map of Tenderness

straints. The metaphorical territories are meaningful only for the symbols they contain. This level is not useful to represent knowledge at a usable level, and is sometimes drawn under artistic and emotional perspectives.

At a second level, many metaphorical territories drawn by artists in the past or visible in the Web use topological constraints, e.g., adjacency, with limited presence of quantitative elements often used to compare rather than to measure. Typical examples are the metaphoric political maps frequently used in the 19th century. Such maps embed qualitative knowledge about the diplomacy and the political relations in Europe and in the World in the shape of topological relations among the map components.

The full concept of map includes metric properties as meaningful components of the representation. Spatially quantified relations in metaphors can convey meanings, contributing therefore to design the richest map type. However, translating quantity on a metaphorical plane is more difficult than translating quality. Metrics are frequently and easily associated to thematic maps, because they can relate metric properties with the values of variables. For example, a thematic map might relate the size of a region with the number of children born in the last year, or with the percentage of land covered by woods.

Metaphorical maps consider knowledge in a broader sense; they may use metric properties to give a feeling of the strength or of the difficulty of an abstract concept. For example, the length of a road may be used to give a feeling of the difficulty of achieving a goal. Therefore an important issue is how to let the reader know which are the meaningful relationships for a given map, in order to avoid misunderstandings. The introduction of a legend, explicitly declaring the mapping, is mandatory in many cases.

5. Building and browsing metaphorical interactive maps

Most of the available tools may treat different classes of objects and are focused on specific segments of the scheme of Figure 2: for example a G.I.S. is usually focused on the (*direct*, *real*) case, because it permits to place different classes of vector objects in the real place; on the other side, authors of maps belonging to the (*metaphorical*, *imaginary*) case may use generic tools for artistic vector graphics, such as Adobe Illustrator™.

A relevant exception to this situation can be found in the Google Maps. While most of the maps built with this tool belong to the (*direct*, *real*) case, because they represent real objects represented on a georeferenced map, there are also some notable exceptions. For example, the site *Slurl.com* [12] uses Google Maps to present a preview of interesting locations available in the *Second Life* world, belonging to the (*direct*, *imaginary*) case. Another interesting example is the interactive version of the *MiddleEarth* [10], related to the famous *The Lord of the Rings* novel. The same site permits to visualize *Second Life* locations that have a strong metaphorical meaning and therefore belong to the (*metaphorical*, *imaginary*) case such as the *Palace of Memory* (Figure 7).

Besides, the Google Maps tool may be used also for displaying maps containing objects that have been built with a metaphorical aim and therefore belong to the (*metaphorical*, *real*) tuple. For example, the maze has been often used in the Christian tradition as a metaphorical representation of the pilgrimage to Jerusalem. Most mazes are represented in the pavement of famous cathedrals, such as Chartres, but there also gardens representing mazes that could have been used by religious communities during centuries. A collection of outdoor mazes from different parts of the worlds is available on Google Maps [14].

These examples show that Google Maps, originally conceived as a tool for displaying *real* geographical ob-

Figure 7. The Palace of Memory in Second Life

jects without any metaphorical aim, can be profitably used also for creating and managing metaphorical maps. Its components represent a distilled subset of objects, widgets and interaction modalities that are widely used in the domain of WIMP interfaces.

The adoption of similar components by other Web 2.0 map applications such as Yahoo Maps! and Microsoft Live Search is a confirmation of their suitability by ordinary users, who are not trained in geographical information systems. Therefore, these components should be carefully analyzed as an interesting starting point for the proposals aimed at building interactive metaphorical cartographies in the context of the current WIMP paradigm.

6. Conclusion

Maps are representations of the real geographic space and – through metaphoric mappings – of other domains of human knowledge. They are useful tools to travel in the real world as well as to metaphorically navigate other domains: the travel from *new friendship* to *tenderness* in the map of Figure 4 is a clear example. The rise of new technologies, such as Google Maps and other geographic mapping systems, opens the way to the creation of interactive maps, which become new media for knowledge representation, sharing, navigation and incremental creation.

A relevant issue reserved to future work is the automatic generation of maps on the basis of a well-defined set of relationships between concepts and geographic elements, defined at design time. The mapping might be domain dependent, shifting therefore the expressivity of metaphorical maps from the single item to the class of metaphorical objects related to a specific knowledge area. The final result might include also the generation of the map legend, in order to smooth ambiguity problems that may be associated to the reading of metaphorical maps.

References

[1] C. Bruni. *I pionieri delle carte geografiche*. Sperry Computer System, 1983 (in Italian).

[2] E. Churchill. Ps and Qs. What's In A Name? Idioms, Metaphors, and Design, *Interactions*, 15, 1, 2008.

[3] P. Codognet. An historical account of indexical images: from ancient art to theWeb. *Proc. IEEE Symposium on Visual Languages*, 1999.

[4] A. Glenberg. Comprehension While Missing the Point: More on Minimalism and Models, *Psycoloquy*, 1993.

[5] W. Kuhn. Handling Data Spatially: Spatializing User Interfaces. In *Proc. 7th Int. Symposium on Spatial Data Handling*, SDH'96, Advances in GIS Research II, IGU, 1996.

[6] T. H. Nelson. *Literary Machines*. T. H. Nelson, Swarthmore, Penn., 1981.

[7] G. F. Orenstein. Journey Through Mlle. De Scudery's Carte De Tendre: A 17th Centruy Salon Woman's Dream, *Femspec* 3.2, 2003.
http://www.femspec.org/samples/ salon.html.

[8] F. A. Yates. *The Art of Memory*. University of Chicago Press, Chicago, 1966.

Web sites

[9] Carte du Tendre de la Recerche. http://www.cetec-info.org/jlmichel/Page.de.Jean-Luc. MICHEL.html

[10] Google Middle Earth.
http://www.elfenomeno.com/GoogleMiddleEarth/

[11] Live Search Maps. http://maps.live.com/

[12] Location-based linking in Second Life.
http://slurl.com/

[13] Love's Topography: la Carte de Tendre.
http://strangemaps.wordpress.com/2008/02/18/245-love's-topography-la-carte-de-tendre

[14] Maze Collection.
http://maps.google.com/maps?q=http://bbs.keyhole.com/ubb/download.php?Number=927611&t=k&om=1

[15] Yahoo Maps! http://maps.yahoo.com/

3D City Models as part of an Open Geospatial Data Infrastructure

Jürgen Bogdahn, Volker Coors

HFT Stuttgart - University of Applied Sciences, Schellingstr. 24, 70174 Stuttgart, Germany
{juergen.bogdahn; volker.coors}@hft-stuttgart.de

Abstract

Three dimensional city models get more and more into focus of professionals and non-professionals. Applications like Google Earth, with a strong level of awareness, provide 3D geospatial data to a huge user community. However, the Google approach, all data for free but stored in Google databases, does not always meet the expectations of private and public owners of geodata, especially 3D city models in this case. They are willing to share their data, but data management should be under their responsibility and control. In order t solve this issue the already existing open geospatial infrastructure has to be extended to cover 3D city models as well. The paper presents actual developments towards this 3D open geospatial infrastructure and shows the use of the developed framework within two projects. In the Virtual Environmental Planning system (VEPs) project a Web 3D Service (W3DS) is used to provide models over the internet that are integrated into an online application. The second project, MoNa3D, investigates the use of distributed 3D city models to support way-finding within a mobile navigation system.

Keywords: geospatial data infrastructure, 3D urban models, mobile navigation

1. Introduction

The interest in creating 3D city models has increased rapidly over the last few years and these models are already entering the field of navigation systems and applications on portable devices. Especially larger cities like Berlin (Döllner et al., 2006), Hamburg or Stuttgart started to build a 3D city model to support city marketing, urban planning, emergency management and other fields. In principle, these models can be integration into virtual globes like 'Grifinor' (Kolar, 2006), GoogleEarth, WorldWind. One kind of navigation in the future might be visual navigation support using a distributed geospatial data infrastructure. As an example, on a trip from Stuttgart to Boston, a personal mobile navigation device will connect to the server of Stuttgart municipality to receive a 3D city model. Based on this model, it will guide the user to the airport. In Boston, the same device will connect to another server and receive a 3D city model of Boston to guide the user to the hotel. In addition, tourist information from another data source could be integrated into the 3D city model.

A candidate for an interface to integrate 3D city models into an open geospatial data infrastructure is the Web 3D Service (W3DS). The W3DS defines parameters to query 3D models in scene graph formats such as KML, VRML, or its successor X3D. These scene graphs can be used by the client to visualize the 3D world and to enable the user to navigate in it.

In this paper we propose a framework to manage 3D urban data and provide a W3DS interface to access the data. The framework provides modules to integrate existing 2D GIS data into the model. This is essential if the 3D model is used in an application environment which should enable the user to get additional information about objects, for example.

One application that uses the presented framework is an online participation tool in the field of urban planning. This tool is developed as part of the Virtual Environmental Planning System (VEPS) project and enables citizens to comment on ongoing urban planning, supported by a 3D city model. Another application, that is currently under development in the MoNa3D project deals with mobile navigation support based on 3D city models (Coors & Zipf 2007). The requirements in the scope of these projects showed us that the W3DS is capable of providing 3D scenes but we identified missing functionalities that are needed to generate the specified functionalities of the applications. These missing functionalities will be described and possible solutions will be presented in this paper.

The following sections of the paper are structured in the following way: In section 2 we will describe how the Web 3D Service (W3DS) implementation utilizes the framework components and how it benefits from the modular architecture in order to fulfil its task. In section 3 we will describe the structure and use of the CityModel Administration Tool (CAT3D) framework to manage 3D urban data. In section 4 and 5 we will present two scenarios in which we integrated 3D city models into applications using the W3DS interface. Section 6 will show some shortcomings of the W3DS in our scenarios and section 7 will conclude the paper.

2. Web 3D Service

In order to achieve interoperability between distributed 3D city model databases it is essential to define a standard interface to access these data sources. One candidate of such an interface specification is the OGC Web 3D Service discussion paper. The Web 3D Service fills the gap between Web Feature Service (WFS) and Web Terrain Service (WTS). On the one hand the WFS delivers geo data objects in form of GML. This data must generally be converted on client side for visualization purposes (compare de Vries & Zlatanova, 2004). On the other hand the WTS delivers image data that can be visualized immediately in a browser, but it is not possible to navigate in a 3D world. 'The W3DS delivers a 3D scene graph instead of pre-rendered images like the Web Terrain Service (WTS). These scene graphs are rendered by the client application, e.g. by a 3D plug-in in case of a web-browser, and enable the user to navigate a 3-dimensional world' (Quadt/Kolbe, 2005).

Normally a W3DS could operate on top of a WFS and a WMS in order to provide 3D scenes. In that case the W3DS just needs to query the WFS for the requested features and to convert the WFS response to the appropriate 3D format. For X3D output this can simply be achieved by a XSLT processor as GML and X3D are both XML based formats. The WMS can be used to deliver terrain textures for the 3D model, the specific requests just have to be integrated into the scene graph in an intelligent way during the XSLT processing

The W3DS that has been developed on top of the CAT3D framework for the two scenarios described in this paper utilizes the components presented in section 3 to fulfil its specific task of providing 3D scene graphs. This approach is useful when not every data source offers a WFS interface or when merging data on GML syntax level is not sufficient. It would be a lot of overhead to parse the GML information again, then merge the data and create the output for the W3DS. Therefore our W3DS implementation uses the CAT3D data connector components to access data sources directly and to hold the data in the mapping layer for further processing. Each layer defined in the W3DS can be associated with a data source. These sources can also be distributed in a Local Area Network (LAN) or on the internet.

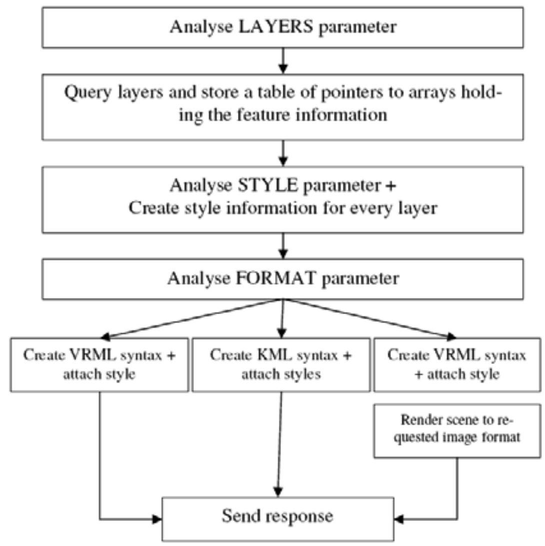


Figure 1: W3DS GetScene request example and the process of creating the response within the CAT3D framework (simplified).

After merging and manipulating data in the mapping layer to suit the model a format creator module can create the output format that was requested by the client. Therefore the framework supports the workflow of the W3DS very well. As you can see in fig.1 the service needs to query data from different thematic layers, apply a style, merge the data and produce the required output format. These tasks can be associated with the components groups described in section 3 (data connectors, data mapping, utilities and format creators).

3. CAT3D – An Open Framework for 3D Data Management

The W3DS implementation in our projects has been developed on top of the 3D CityModel Administration Toolkit (CAT3D) framework. This modular framework can store, manage and manipulate geo data and especially 3D city models. The core of the CAT3D framework is a geospatial database for efficiently storing and querying 3D urban models.

The framework consists of four parts: data connectors, data format creators, data mapping and utilities (fig. 2). The data connector modules can access different data sources like spatial databases, files or other web services. By implementing the logic of the source format they assure the consistency of the data and they manage the data transfer from the source format to the internal data representation. During the transfer process the data logic can be checked; for example, a geometry object can never exist without a relation to a feature. The data that is read from the data sources is converted into the data mapping layer. The data mapping classes describe a general form of features and geometries or in other words a general object model. This model is based on the OGC Simple Features definition and adopted some basic definitions but was also extended to describe 3D geometries (e.g. facets). The object model is used inside the framework to describe geospatial data, its form and structure, and all modules that are added to the framework must support this mapping model in order to be able to exchange data with other components. Before merging the two sets of data from distributed sources, it is also possible to transform parts of the data to fit into the model by using utility functions that work on the internal data format. The format creator modules in the framework are used to create special external data formats like VRML, SHP, KML and others.

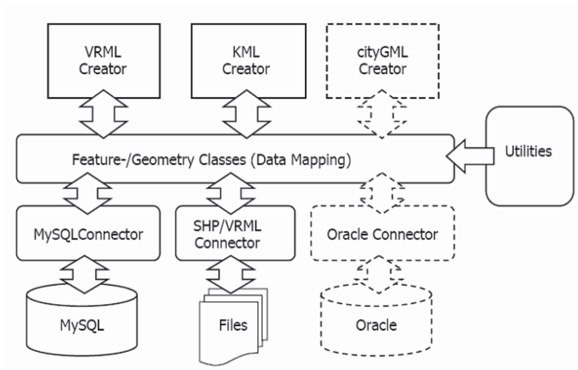


Figure 2: CAT3D Service Framework architecture

In contrast to the data connectors, which translate the source data schema to mapping classes, these modules translate the internal format to the external one. The special capability of the format creators is to convert the data of the feature classes to the external format, which sometimes is more restricted and cannot handle all the information provided in the source format, or only in a different way.

4. Online Public Participation using 3D City Models

Within the VEPs project the W3DS is used as in the context of an online participation tool for citizens. The VEPS Project aims at using internet-based 3D city and landscape models to improve ePlanning, consultation and communication of citizens' views on planning issues (Chen & Knapp, 2006). It integrates 3D city models, 2D map functionality and a discussion forum in one web-based client application. Online discussion and communication is managed by an open source online forum (www.phpbb.com). The contributions to the discussion forum can be linked to a specific 3D view or an individual 3D object in the planning area. The 3D window is not only used for visualization, it is even more a part of the user interface (Figure 3). It is possible to start a discussion about an object in the forum by simply clicking on it in the 3D scene. The position from which the user started this specific discussion is determined and needs to be stored together with the comment/discussion. When the discussion is opened again this position should be restored in the 3D view so that the user can see the visual aspect of the discussed issue and better understand the planning intention. Unfortunately the discussion forum in its original version does not support the storage of a spatial position together with the discussions and comments. Therefore the php-code and the database schema had to be extended in order to support the link of a geo-position together with a comment and to fulfil the requirements of the participation tool. These extensions allowed the use of the rich functionality of the discussion forum within the participation tool, but these changes were made only for this specific tool. In order to provide a more standardized way of storing comments with a spatial relation, the Comment Markup Language (CoML) was developed, a GML application schema, that enables the use of a WFS to manage and query the geo-comments created by the tool (Schill et al, 2007). The 3D scene as part of the user interface can also be used to query attributes and additional information about objects using the GetFeatureInfos request of a modified W3DS (see section 6). It is also possible to query all ongoing discussions about a specific object as the object ID is stored together with the comments as well. In this way the 3D model is not only visualization, but also a very intuitive tool to access additional data and semantics.

Another shortcoming of the W3DS interface that we have identified for our application is the query of terrain textures. The W3DS interface allows querying a model by defining a bounding box (bbox), therefore the area of the result scene is always different. A texture with fixed size can be regarded as inappropriate to serve as a texture for a dynamically queried terrain model. We would suggest a GetTerrainTexture function to be integrated into the W3DS interface to be able to load terrain textures without reloading the entire scene (see section 6). In order to be able to integrate the system into existing spatial data infrastructures (SDI) the tool incorporates web-applications on client side, which means the communication with the server remains in the responsibility of the specific application. The synchronization of the "views" is done on client side using Javascript functionality (fig.4). In that way one component can be replaced by another implementation quite easily and different distributed sources can be integrated. For example, if the municipality already offers a WMS, this service can be used as the back-end of the 2D view. In this specific scenario one can see that it is not only important to have a flexible framework to build the 3D model, the whole system needs to be able to utilize existing SDI components to avoid duplicate systems and redundant data.

Figure 3: Web3D Service integrated into an online participation platform.

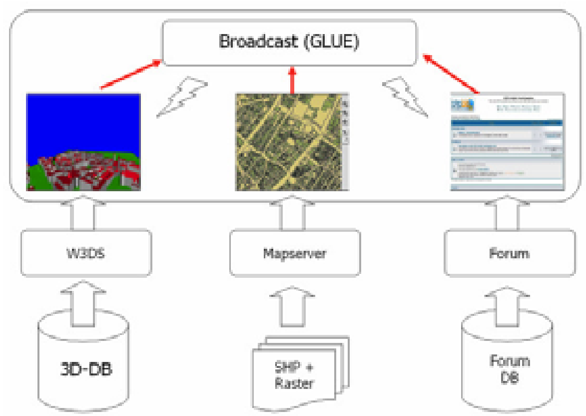


Figure 4: Distributed data sources used by VEPs participation tool.

5. 3D City Models supporting Mobile Navigation

There is a wide variety of techniques to present directions and support on mobile devices ranging from spoken instructions to 3D visualization. In order to produce a coherent and cohesive presentation it is necessary to adapt a presentation to the available technical and cognitive resources of the presentation environment. Coors et al. (2005) evaluated several means of route instructions to a mobile user. 3D visualization seems to be well suited where time and technical resources are not an issue, and where the available position information is imprecise: a 3D model of the area allows the user to search her environment visually for specific features, and then to align herself accordingly, thereby compensating for the imprecision. 3D urban models as integrated part of a navigation system will support pedestrian navigation as well as the "last mile" car navigation from a parking place to the final destination. Especially using public transport and finding the best exit at a larger tube station for instance is a typical three-dimensional problem that is hard to solve with traditional 2D navigation support. The two main companies for navigation data, TeleAtlas and Navteq currently offer 3D landmarks of almost all big European cities. For some cities even large 3D urban models are available (van Essen 2008). However, only a few research prototypes actually support mobile navigation with real 3D urban models on mobile devices such as smart phones (Coors & Schilling 2003, Nurminen 2006).

The aim of the project "Mobile Navigation with 3D City Models" (MoNa3D) is to develop and evaluate such a mobile navigation system. The main goals of the project are to provide a cognitive semantic route description by using landmarks in 3D, which allow context dependent personalized navigation support and to develop an approach to create suitable building textures using image processing methods and synthetic textures (Ebert et al., 2003) along with a corresponding compression system for an efficient storage, transfer, and rendering of 3D urban models with an open geospatial data infrastructure.

In this context the current support of data formats of the W3DS Interface is not sufficient. Neither VRML nor KML are really suited for mobile devices. Instead, an efficient 3D compression scheme is necessary for data transmission as well as fast loading of the transmitted data into the rendering engine. The Java Mobile 3D Graphics API (JSR 184) supports the more compressed M3G format for 3D graphics on mobile devices. While general purpose compression tools such as Win-zip or Gzip achieve average compression results on 3D data, specialized algorithms such as Delphi compression (Coors and Rossignac 2004) that are not based on textual representation but take into account the 3D model structure of the content propose a deflation of up to 95% (Eppinger and Coors, 2005).

A 3D city model usually consists of a terrain model and individual building models in different levels of details. The terrain model is stored as a regular grid, an irregular point set or TIN. From each of these different representations a triangle mesh can be created at run time on the server side. Building models are usually stored as polygonal boundary representations. Each polygonal boundary representation can always be transformed into a triangle mesh without any loss of information. A triangle mesh is usually represented by its geometry information, the vertex location of the mesh, and by its connectivity information, which describes the incidence relation between each triangle of the mesh's surface and its boundary vertices. Additional attributes such as color and texture information may be part of the description. However, in the proposed approach of synthetic textures, only a few parameters are stored per polygon or triangle patch. Based on these attributes, synthetic textures are generated during rendering on the mobile device (Coors 2008).

Currently, the components are integrated into a prototype navigation system based on a service based infrastructure including the discussed Web3D Service that grants access to the 3D urban models and a routing service based on the OpenLS interface. The client is implemented in Java ME using the M3G API in addition for 3D rendering on smartphones.

6. Shortcomings of the W3DS in the presented scenarios

Using the W3DS interface specification in the context of the presented applications some shortcomings showed up. The W3DS only provides a graphical representation of the geo data. In fact geo data, in most of the cases, consists of geometry and additional attribute data. This kind of additional information is not supported in scene graph structures, as these structures are developed to describe the visual aspect of a 3D model. However, in order to use 3D city models beyond visualization, the application needs access to the attribute data as well. We would suggest a "GetFeatureInfos" request like in the WMS specification. In contrast to the WMS GetFeatureInfo request, which includes the majority of the original map request to define which feature information is requested, the W3DS should include unique IDs into the scene graph for each feature. These IDs may be utilized in the GetFeatureInfos request to determine for which feature the information is needed. The IDs can be created by using the database primary keys or any other information that is appropriate to build a unique identifier (compare also Coors & Jung, 1998). The specific format of the ID should remain in the responsibility of the service provider. The general problem of the feature concept in a scene graph might be to define what a feature is. Some service providers might define that one feature is modelled as a FaceSet-Node (in case of VRML output). Another service provider would model a feature as a Group-Node that has several FaceSet child nodes. An application using both services might run into problems when it has to find the feature ID. A more general solution would be a GetFeatureInfos request with x, y, z-parameters. In this case the service would have to find the feature by a spatial operation to return the set of parameters. The challenge of attribute integration should definitely be discussed in more detail among providers and users of the W3DS interface as the additional attribute information is extremely important when the W3DS should support interactive 3D applications (see section 4).

Another problem in the public participation scenario was the requirement to switch terrain textures. Normally texture tiles (e.g. orthophotos) are used to texture terrain models. These tiles need to be merged dynamically to fit the bounding box of the scene. One solution would be to specify the texture in the style parameter, "produce" the appropriate texture when the scene is built and to create a file that can be referenced by the scene graph. In our scenario we wanted to be able to switch between different terrain textures (orthophoto, street map, etc.) without reloading the entire scene. Therefore we would suggest a GetTerrainTexture function that enables the client to query a texture by defining a bbox and a type. This function is not supposed to work as a WMS; as it does not merge different data layers to produce a map-texture. The function just offers an interface to query terrain textures of different types and dynamic size, even though it could forward the request on server side to a WMS in order to receive appropriate maps that can be used as terrain textures. The use of 3D styled layer descriptors as presented in (Neubauer&Zipf, 2007) might be a much more elegant solution to define layer styles and needs to be investigated for further development of our W3DS implementation, though it might not replace the GetTerrainTexture function.

For the mobile navigation scenario efficient and fast data transmission is going to be very important and therefore the W3DS implementation needs to support additional compressed and optimized data formats as output. This requirement is primarily not a problem of the W3DS specification but needs to be supported by the underlying framework or data management. Nevertheless it has to be investigated in the future if additional parameters for the GetScene request are required to query optimized models.

7. Conclusions and Outlook

In this paper the Web3D Service was explored as a connector for 3D city models to an open geospatial infrastructure. The W3DS was developed as part of the open source framework CAT3D and tested within two application scenarios, one in web-based public participation, and the other in mobile navigation support. Some shortcomings of the current specification of the Web3D Service have been identified. Extensions of the Web3D Service have been proposed to overcome these shortcomings. The Web 3D Service interface can be used to provide 3D scene graphs which are used by special viewers to visualize the 3D scene. For interactive 3D scenes providing additional information about objects the W3DS interface should be extended by a GetFeatureInfos request similar to the one in the WMS specification. A requirement that is specific to the participation tool is the creation of an Industry Foundation Classes (IFC) connector for the CAT3D framework, which would allow reading CAD models. This is a quite strong requirement at the moment because planning data must be integrated into the 3D model for the participation tool and therefore a connector which bridges the gap between CAD and GIS would be extremely useful. Another development step for the future would be the support of SLD3D as it is a more flexible way to define styles for 3D data layers.

Acknowledgements

This work has been supported by the Germany Ministry of Education and Research (BMBF) within the project MoNa3D and by the EU INTERREG IIIB within the VEPs project. Special thank to all project partners and co-workers for their valuable support and input.

References

Chen, Y. and Knapp, S. (2006): VEPS – Virtual Environmental Planning System: First Steps towards a web-based 3D-planning and participation tool. CORP 2006 & Geomultimedia06, Proceedings / Tagungsband; Editor / Herausgeber: Manfred SCHRENK, ISBN: 3-9502139-0-2

Coors, V. and Jung, V (1998): Using VRML as an interface to the 3D data warehouse. Symposium on the Virtual Reality Modelling Language(VRML98), Monterey, CA, USA, Feb. 16-19, 1998.

Coors, V., Elting, C., Kray, C., and Laakso, K. (2005): Presenting Route Instructions on Mobile Devices - From Textual Directions to 3D Visualization, to appear in in Kraak, M., MacEachren, A., Dyke, J.: Geovisualization, Taylor & Francis

Coors, V., and Rossignac, J. (2004): Delphi: geometry-based connectivity prediction in triangle mesh compression In The Visual Computer, International Journal of Computer Graphics, Vol. 20, Number 8-9, Springer Verlag, Berlin Heidelberg, November 2004, pp 507 – 520.

Coors, V., and Schilling, A. (2004): 3D Maps on Mobile Devices, In: Branki, C.; Unland, R.; Wanner, G. (Hrsg.): Multikonferenz Wirtschaftsinformatik (MKWI) 2004, Universität Duisburg-Essen, 9.-11. März 2004, Band 3: Mobile Business Systems, Mobile and Collaborative Business, Techniques and Applications for Mobile Commerce (TAMoCO). Essen

Coors, V. and Zipf, A. (2007): MoNa 3D --- Mobile Navigation using 3D City Models, 4th International Symposium on LBS and Telecartography 2007. Hongkong.

Coors, V. (2008): Synthetic Textures for 3D Urban Models in Pedestrian Navigation, TAMoCo 2008, Glasgow

de Vries, M. and Zlatanova S. (2004): Interoperability on the Web: The Case of 3D Geo-Data. P. Isaias, P. Kommers and M. McPherson (eds.), Proceedings of the IADIS International Conference e-Society 2004, Avila, Spain, July 16-19, pp. 667-674

Döllner, J.; Baumann, K.; Buchholz, H. (2006): Virtual 3D City Models as Foundation of Complex Urban Information Spaces. Proceedings / Tagungsband; Editor / Herausgeber: Manfred SCHRENK ISBN: 3-9502139-0-2 (from 2007: ISBN 978-3-9502139-0-4)

Ebert, D., F.K. Musgrave, D. Peachey, K. Perlin, S. Worley (2003): Texturing & Modelling. A procedural Approach. 3rd Edition, Elsevier Science

Eppinger, F. and V. Coors: Compressing 3-dimensional urban models In: Proceedings 4th Workshop on Dynamic and Multi-dimensional GIS (DMGIS 05), September 2005, International Archives of Photogrammetry and Remote Sensing, Vol XXXVI, Part 2/W29

Haist, J. and Coors, V. (2005): The W3DS-Interface of CityServer3D. Gröger/Kolbe (Eds.), Proc of the 1st Intern. Workshop on Next Generation 3D City Models, Bonn 2005

JSR184 (2005): Java Specifiaction Request 184: Mobile 3D Graphics API, Version 1.1, June 2005; and Java Specifiaction Request 297: Mobile 3D Graphics API, Version 2.0, Early Draft, August 2007

Kolar, J. (2006): On the road to 3D geographics systems : important aspects of global model-mapping technology. Innovations in 3d geo information systems. New York : Springer, 2006. p. 207-223 Conference: International Workshop on 3D Geoinformation, No. 1, Kuala Lumpur, Malaysia, August 7, 2006 - August 8, 2006.

Neubauer, S; Zipf, A.: "Suggestions for extending the OGC Styled Layer Descriptor (SLD) Specification into 3D – Towards visualization rules for 3D city models"; In: Urban and Regional Data Management; Coors, Rumor, Fendel, Zlatanova (Eds.), Taylor&Francis, ISBN 0415440599, 2007

Nurminen, A. m-LOMA - a mobile 3D city map Proceedings of the eleventh international conference on 3D web technology, 2006.

Quadt, U. and Kolbe, T.H. (2005): Web 3D Service. OGC Discussion Paper, Reference Number: 05-019

Rossignac, J. (1999): "EdgeBreaker: Connectivity compression for triangle meshes", IEEE Transactions on Visualization and Computer Graphics, 5(1), 47-61.

Schill, Ch.; Bogdahn, J.; Koch, B.; Coors, V: Public Participation Comment Markup Language and WFS1.1, In: Urban and Regional Data Management; Coors, Rumor, Fendel, Zlatanova(Eds.) , Taylor&Francis, ISBN 0415440599, 2007

van Essen, R.: „Maps Get Real, Digital Maps evolving from mathematical line graphs to virtual reality models." In: Oosterom, P. Et al. (Eds): Advances in 3D Geoinformation Systems, Lecture Notes in Geoinformation and Cartography, Springer, 2008.

Dynamic Access of XML Document for Multi-User

Jeang-Kuo Chen*, Kuan-Chang Lu

Department of Information Management, Chaoyang University of Technology

jkchen@cyut.edu.tw, s9514637@cyut.edu.tw

Abstract

XML is a simple and flexible text format. Quick developments of XML makes the data management need an effective mechanism for data exchange and electronic commerce to support the XML document access by multiple users concurrently. The access method with of single-user for XML document is unsuitable in a multi-user environment. Therefore, we propose an access method with concurrency control of multi-user for a dynamic XML document. With the techniques of concurrency control, our method allows multiple users to concurrently and correctly access the same data without any error occurrence.

Keyword：XML, Concurrency Control, Lock, Dynamic Access

1. Introduction

The XML (eXtensible Markup Language) [14] supports various commercial transactions such as electronic commerce, data exchange, etc [1, 8, 12]. XML can carry a description of its own format, which is a special characteristic of XML. With a DTD or an XML Schema, an XML document can be verified whether it is a valid document or not. As the quantity of XML documents increases gradually, it is necessary to manage XML document by the database management system to facilitate management and access of XML documents [6, 11,13]. A concurrency control mechanism is required to maintain multi-user accessing the same data at the same time. If XML database is accessed by many transactions without any concurrency control, some unpredictable problems such as lost update, dirty read, or incorrect summary [4] may occur. An example is illustrated as follows. An XML document is used to record the data of a bookstore. When two transactions access the same magazine data at the same time without any concurrency control, some unexpected mistakes may happen. Suppose the stock of a magazine in the bookstore is 10. Transaction A (T_A) will subtract 3 form the stock because 3 magazines are sold. Transaction B (T_B) will add 5 to the stock because 5 magazines are stocked. T_A includes three steps (1)reading the stock of the magazine, (2) subtracting 3 form the stock, and (3)updating the new stock into database. T_B also includes three steps too (1)reading the stock of the magazine, (2)adding 5 to the stock, and (3)updating the new stock into database. If T_A and T_B execute sequentially (T_A then T_B, or T_B then T_A), the result is correct as shown in Table 1. However, if T_A and T_B execute alternately, as shown in Table 2, the lost update problem occurs because T_A loses the new stock value at the order 6 of Table 2. Form the above example, we can understand the importance of concurrency control in a database with multi-user access.

This paper proposes a set of concurrency control algorithms for search and insertion of elements in an XML document. With the locking technique, the share lock (s-lock) [3] is used to lock an element by one transaction before reading the element. The s-lock is sharable with other s-locks which means an element can be s-locked by more then one transactions that can read the element concurrently. The exclusive lock (x-lock) [3] is used to lock an element by one transaction before changing the contents of this element. The x-lock is exclusive with other s-locks or x-locks that means an element can be x-locked only by one transaction at one time. The s-locks are compatible but x-locks. An x-lock is incompatible with any other s-locks and x-locks. A transaction must lock an element before accessing the element and release (unlock) the lock as soon as possible after handling the element.

2. Relative technologies

2.1 XML

XML is a kind of mark language derived form SGML [15]. W3C proposed two types of XML[14]. The first type is called well-formed while the second type is called valid. A correct XML document must be well-formed. A well-formed XML document is valid if It is verified by a DTD (Document Type Definition) or an XML schema. An example of a DTD and related document are shown in Fig.1.

2.2Concurrency control

Concurrency control offers an efficient and reliable mechanism for database under the environment of multi user. This mechanism guarantees that even transactions work execute alternately, the result is still correct. There are three main concurrency control methods, locking, time stamp, and optimistic [4].

The lock method allows many transactions access data item at the same time. Lock is one variable that describes the state of this item. If one data item has been locked by T_A, only T_A can access it, else this item can lock by any other transactions, this is binary lock protocol. The drawback of binary lock scheme is only one transaction can access database item. Therefore, lock type evolve to share lock and exclusive lock. When a data item has share lock, meat that data item can read by others transaction. If one data item locks by exclusive lock, that data item can’t access by other transaction.

Another protocol of lock named two phase locking protocol. This protocol includes two phase, growing and shrinking phase. When transaction gets lock on data item, this transaction is during the growing phase. After all of target of transaction already get locks, transaction works. When transaction is finished, transaction should unlock all data items, this transaction enters shrinking phase. In shrinking phase transaction can't lock any data item, only can release lock. This rule guarantees that deadlock never happen. The two phase locking protocol has a drawback. That is if one lock request for data item by T_A first, that data item will release when T_A is finished, that data item can not access by other transaction. The efficiency will reduce. One lock protocol named lock-coupling protocol can solve this problem. When one transaction needs to lock two data item sequentially, the first item should be release before second item is locked.

Time stamp method utilizes the administrative system of the database to produce a unique identifier time stamp for every transaction. When one transaction accesses any data items, should compare time stamp of transaction and data item first [4]. The optimistic technique allows any transaction access data item. This method check when transaction finishes. The transaction accesses data in local copy. When prove result is correct upgrade the local copy into the database [4].

The paper [7] propose three kinds of core protocols on a list type data structure. the noses in this list type includes content and pointer points to other nodes. Those protocols are NODE2PL, NO2PL and OO2PL, they works on different area in index.

The concept of paper [10] uses list locking protocol and * - factoring method to execute concurrency control. Method of paper [9] XLP is concurrency control method designing for XPath. An intact Xpath query is “Axis: : Node-Test [predicate]”. XPath query starts at root node. Transaction need access relevant nodes for search destination nodes, those nodes need concurrency control to deal with. XLP regards XML as a tree structure, while execute one Xpath query, the nodes will be P-locked if pass by, when nodes didn’t satisfy node-test or predicate, p-lock will be release. If nodes are satisfied request of query, those P-locks will be upgrade to other lock, for each operation like insertion lock, deletion lock, read lock, and write lock.

3. Search and insert operations

At a broad view of the method that introduce at preceding paragraph. The shortcomings of XLP are two. First is lacked insertion or modification operation, can only search data in XML document. Second shortcoming is XPATH designed to professional user. The general user is difficult to use XPATH as query tool. Methods of paper[7], use too much lock types on pointers and nodes. When destination element is middle of the child nodes, the quantity of lock will be huge, that will spend time and resource to deal with. Because above-mentioned reasons, we proposes a suitable method for search and insertion concurrency control algorithm. Search method is utilizing those keywords to find out the proper element in document. Insert algorithm can find out a proper parent element quickly, and insert new element as a sub element at that element.

3.1 Index tree structure

Use m-way tree as the index structure of XML to accelerate data access on XML document. The reason is structure of XML similar to m-way tree, it don’t limit child node quantity or key value of node such as B-tree [2] or R-tree [5]. Node structure in this index tree is shown in Fig.2. Each node includes the pointer named tag_name, content, and child. Those pointers are pointing to the XML element of tag name, position of XML element content, if content don’t have real data, this field will be null and Child1, child2, …,childn pointing to each child nodes of this node.

3.2 Lock mode

The lock types are share-lock and exclusive-lock. The relations of these two locks please consult Fig. 3. S-lock expresses share-lock and x-lock expresses exclusive-lock. O shows the two locks are compatible, X means incompatible. Share-lock(s-lock) uses at transaction need to read the data of node. When node being s-lock by one transaction, other transaction permit a s-lock only. Exclusive-lock(x-lock) uses at transaction need to modify the data of node. When a node has an x-lock or s-locks, other x-locks will refused, until the x-lock at this node is x-unlock.

3.3 Search operation

The search operation ascertains that tag name, content or attribute of node included keywords or not. Use the breadth-first search while visit the nodes. List some explanation of nouns that will be used in the course of searching operation as follows. In our definition, R is an index root node, Q is a queue for store index node of XML document waiting to check. The r is one queue for store index node of XML document included all keywords. The KEY is a set of keywords (K1, K2, K3, …). The execution step of search operation shown in Fig.4.

```
Algorithm SEARCH(R, KEY）
Input char KEY[1..n];
      //keywords K1 ,K2 ,… ,Kn//
      node R; //root node//
Output：queue r;
        //Queue stores node satisfied keywords//
```

```
Begin
  node N;
  // current node do keywords match//
  queue Q; //queue for store index node of XML document
         waiting to check//
  queue r; //quere for store index node of XML document
         included all keywords //
  s-lock(R);
  Add R to Q;
  while Q is not empty, do
    N← get the first element form Q;
    if N's tag_name, content, or attributes include
    KEY, then
      add N to r;
    end if;
    for each child i of N, do
      s-lock(i);//Lock this node, other transaction can not
            modify it.//
      add i to Q;
    end for;
    s-unlock(N);//Match operation is finished, allow other
            transaction access it.//
  end-while;
  return r;
End SEARCH.
```

3.4 Insert operation

When inserting a new node N at M, transaction should take x-lock on M, then insert N to M as child node. When insert operation is finished, x-unlock M and s-unlock all node in PATH sequentially, the transaction finished. From root node to M node, we regard this route as big node of a logic node, without meddles with other transactions. The execution step of insert operation shown in Fig.5.

```
Algorithm INSERT （R,,N,PATH）
Input：m-way tree R;node N; //a new node //
      queue PATH[1..n]; //queue stores the nodes in path
      form root to M, excludes M//
Begin
  for each node i, form top to bottom in PATH, do
      s-lock(i);
      //lock nodes in path to ensure those nodes not
      modified by other transaction.//
  end for;
  x-lock(M);//lock M to ensure M do not access by other
        transaction. //
  add node N to M as M's child node; //insert node//
  x-unlock (M);//Allow M access by other transaction.//
   for each node i, form top to bottom in PATH, do
      s-unlock(i); //Allow nodes in path access by other
transaction.//
  end for;
End INSERT.
```

4. Example

This case describes the execution of search and insertion operations. The Fig.6 is the logic structure of XML file of Fig.1(b), each circle represents a data of one XML element included tag name, attribute, content, and each pointer points to one child node. An XML element represent with the tag name in following explanation, tag name (attribute) represent an element with attribute. XML document permit of elements have same tag name. In Fig.6 each nodes will distinguish by number, expression is tag name(serial number). The serial number is accorded with from top to bottom, the order from left to right and begun at 1 sequentially.

The user inputs the key word "system", the root node “Book list” will be add to Q first, then s-lock(Q), then compare the key word included in the tag name or attribute of “Book list” or not, because “Book list” don't include the key word, then s-lock child node of “Book list” “Book_type(2)”, “Book _ type(3)” then add into Q. Obtain the node “Book_type(2)” from Q named N, s-lock (N). The N's tag_name, content, or attributes exclude KEY, then s-unlock(N), s-lock child nodes of N “Book(4)”, “Book(5)” and add into Q. Take out “Book(4)”from Q named N, check N find attribute included key word, so add N into r, s-unlock N. Take out the next node “Book(5)” from Q named N, Repeat above-mentioned steps, until Q is empty. The last result is Book (4), Book (5), these two nodes are destination.

5. Conclusion

This paper proposes a search and insertion method of concurrency control technology on XML document. The index tree uses m-way tree. Use two kinds of lock whose are share lock and exclusive lock. While search operation is executing, use birth first search travels the node and compare tag name, attribute, content and keywords. When insert a new node, locking nodes from root until destination node sequentially, after insertion, the lock on nodes will release. This paper proposes a tool for search and insertion. Use simple lock types for prevention of system judgment relation between each locks for increase system efficiency, and guarantee correct result at the same time.

References

[1]A. Bonifati, S. Ceri, and S. araboschi, "Pushing Reactive Services to XML Repositories Using Active Rules", ***The 10th International Conference on World Wide Web***, pp. 633-641, 2001.

[2]R. Bayer and M. Schkolnick, Concurrency of Operations on B-trees, ***Acta Inforrnatica*** pp. 1-21, 1977.

[3]J.K.CHEN,Y.F.HUANG, .H.CHIN,"A Study of Concurrent Operations on R-Trees" ***Information Sciences, Volume 98, Number 1***, pp. 263-300, 1997.

[4]R. Elmasri and S. B. Navathc, "***Fundamentals of Database Systems, 4th Education***." Addison Wesley, 2003.

[5]A. Guttman, "R-trees: A dynamic index structure for spatial searching" ***Proc.ACM SIGMOD Annual Meeting***, pp. 47-57, 1984.

[6] G. Governatori, B. Stantic , and A. Sattar1, "Handling of Current Time in Native XML Databases", ***17th Australasian Database Conference Vol. 49***, pp. 175-182, 2006.

[7]S. Helmer, C. Kanne, and G. Moerkotte, "Evaluating Lock Based Protocols for Cooperation on XML Documents", ***ACM SIGMOD Record Vol. 33, Issue 1***, pp.58-63, 2004.

[8]W. Hümmer, A. Bauer, and G.r Harde, "XCube – XML for Data Warehouses", ***the 6th ACM International Workshop on Data Warehousing and OLAP***, pp. 33-40, 2003.

[9]K. F. Jea and S. H. Chen, "A High Concurrency XPath-Based Locking Protocol for XML Databases", ***Information and Software Technology Vol. 48***, pp. 708–716, 2006.

[10]E. Lee, "Multi-Granularity Locks for XML Repetitive", ***IEEE Fourth Annual ASIC International Conference on Computer and Information Science***, pp. 222-227, 2005.

[11]E. J. Lu, R.H. Tsai, and S.H. Chou," An Empirical Study of XML/EDI", ***Journal of Systems and Software Volume: 58, Issue: 3***, pp. 271-279, 2001.

[12]T. Milo, S. Abiteboul, B. Amann, O. Benjelloun, and F. D. Ngoc, "Exchanging Intensional XML data", ***ACM Transactions on Database Systems Vol. 30, Issue 1, pp***. 1-40, 2005.

[13]S. Natu and J. Mendonca "Digital Asset Management Using A Native XML Database Implementation" ***Proceedings of the 4th Conference on Information Technology Curriculum CITC4 '03***, pp. 237-241, 2003.

[14]W3C,ExtensibleMarkupLanguage(XML)1.1, http://www.w3.org/TR/2006/REC-xml11-20060816/

[15]W3C, Standard Generalized Markup Language, ***http://www.w3.org/MarkUp/SGML/***

Order	transaction	operation	stock
1	T_A	Read x	10
2	T_A	x=x-3	10
3	T_A	Write x	7
4	T_A	Commit	7
5	T_B	Read x	7
6	T_B	x=x+5	7
7	T_B	Write x	12
8	T_B	Commit	12

Table 1. Correct result.

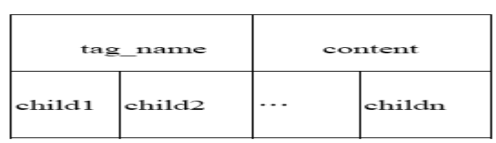


Fig. 2 Node structure.

	s-lock	x-lock
s-lock	O	X
x-lock	X	X

Fig. 3 Relation of two locks.

Order	transaction	operation	stock
1	T_A	Read x	10
2	T_B	Read x	10
3	T_A	x=x-3	10
4	T_B	x=x+5	10
5	T_A	Write x	7
6	T_B	Write x	15
7	T_A	Commit	15
8	T_B	Commit	15

Table 2. Incorrect result.

```
<?xml version="1.0" encoding="Big5"?>
<!DOCTYPE booklist[
<!ELEMENT booklist (book_type+)>
<!ELEMENT book_type (book*)>
<!ATTLIST book_type type CDATA #REQUIRED>
<!ELEMENT book (title,author+,publisher+,price+,year+)>
<!ATTLIST book ISBN CDATA #REQUIRED>
<!ELEMENT author (#PCDATA)>
<!ELEMENT publisher (#PCDATA)>
<!ELEMENT price (#PCDATA)>
<!ELEMENT year (#PCDATA)>
]>
```

Fig. 1(a) A DTD.

```
<?xml version="1.0" encoding="Big5"?>
<booklist>
    <book_type type="computer science">
        <book ISBN="115225722">
            <title> Database System </title>
            <author>Luce White</author>
            <publisher>Mcgraw Hill</publisher>
            <price>800</price>
            <year>2006</year>
        </book>
        <book ISBN="552489547">
            <title> information system management
            <author>Bill Wu</author>
            <publisher>Flag</publisher>
            <price>600</price>
            <year>2007</year>
        </book>
    </book_type>
    <book_type type="novel">
        <book ISBN="514569874">
            <title>Run</title>
            <author>Edwin King</author>
            <publisher>C.C.</publisher>
            <price>320</price>
            <year>2006</year>
        </book>
        <book ISBN="458474257">
            <title>Right now</title>
            <author>Jean Mear</author>
            <publisher>D.K.</publisher>
            <price>180</price>
            <year>2005</year>
        </book>
    </book_type>
</booklist>
```

Fig. 1(b) An XML document

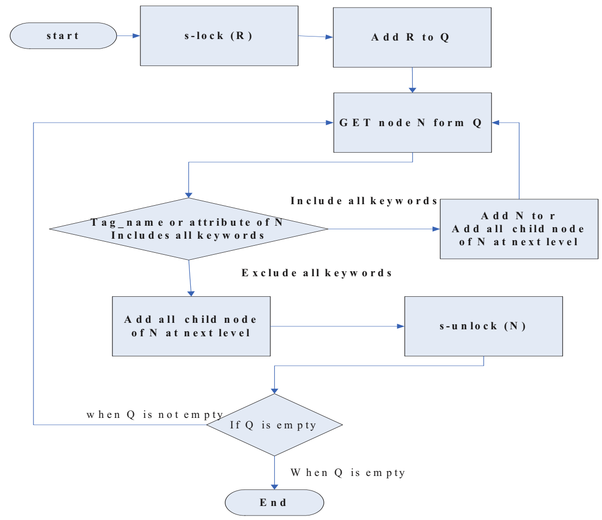


Fig. 4 Flow of search operation.

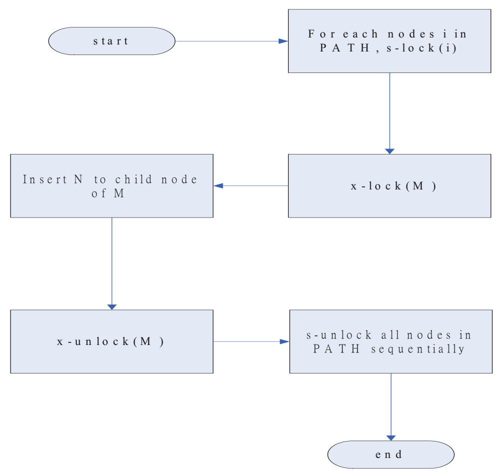


Fig. 5 Flow of insert operation.

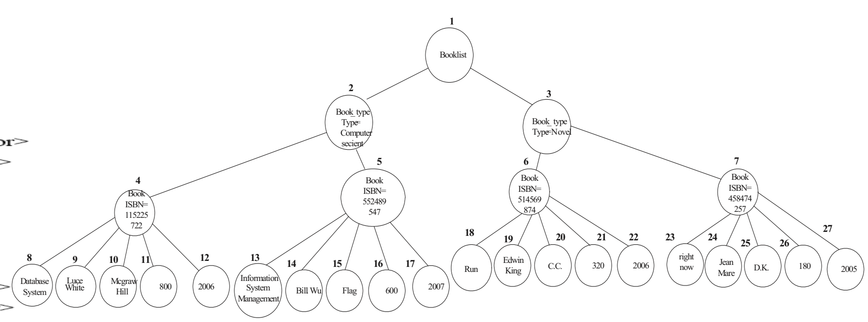


Fig 6. A logical tree of XML document.

A Multimedia Data Streams Model for Content-Based Information Retrieval

Shenoda Guirguis[1], Rohit Kulkarni[2] and Shi-Kuo Chang[1]
[1]Department of Computer Science
University of Pittsburgh
Pittsburgh, PA 15260 USA
chang@cs.pitt.edu
shenoda@cs.pitt.edu
and
[2]Department of Information Science
University of Pittsburgh
Pittsburgh, PA 15260 USA
rok26@pitt.edu

Abstract: *Multimedia applications nowadays are becoming the standard, for they utilize the enormous human brain computational power. In the past, the Relational Database Model was generalized to the Multimedia Databases Model. Orthogonally, the Relational Database model was recently generalized to the Data Streams Model, as the technology advanced and Data became bulky and unbounded in size, utilizing sensor networks. In this paper, we take one more step of generalization, by providing a Multimedia Data Streams model. We give the formalization in details of the multimedia data streams problem. The objective is to furnish a formal framework to efficiently design a Multimedia Data Streams (MMDS) schema that achieves an efficient performance in regard with content based retrieval. We also extend the Functional Dependency Theory and the Normalization Framework to handle Multimedia Data Streams. Finally, we discuss algorithmic methods of generating Continuous Multimedia Queries, along with concrete examples for further illustration.*

1. Introduction

Multimedia Databases have been used in many applications for over two decades now. The internet boom has increased and called for this trend, introducing many new interesting applications related to query processing and content based retrieval. In [4], a Normalization Framework for Multimedia Databases was provided, as a generalization to the Traditional Relational Database model. The data dependency theory was also extended to include dependencies involving different types of multimedia data. Recently, a new research area is being proposed called Human Computation [5], which utilizes the ultimate powerful computational units all humans possesses; i.e. human brain. Starting with CAPTCHA, Luis et. al., proposed many Multimedia Games (applications that utilizes Multimedia databases) to utilize Human Computational power in solving some difficult problems [5]. On the other hand, the data stream management concept was first introduced in [9], and later a full-blown Stream Management Systems were proposed, such as STREAM [3], Borealis [1], and Aurora [2]. And couple of commercial Data Stream Management Systems (DSMSs) are now available [10], [11]. Data streams emerged as the new data management model that can handle the huge massive updates of data collected from sensor networks or monitoring applications. In the push-model for the data streams, data updates arrives along the time with high frequency, while a certain set of queries (continuous queries) reside in the data stream server to process incoming data [2]. In other words, users register their continuous queries (CQs), and data will be pushed into these queries as they arrive. This model is to be opposed to the pull-model of traditional databases. Data streams have many critical and important applications ranges from network monitoring applications, to military applications, passing through environmental monitoring. Data streams was even more promoted by the back then new technology of sensor networks; small devices that can easily be spread over a large area to collect some information. From what precedes, one senses that soon enough, as technology advances, Multimedia data will become massive and unbounded; i.e. a stream of multimedia objects, or what we call: Multimedia Data Streams. It is hence required to generalize the Multimedia Dependency Theory and Multimedia Normalization to be applicable for Multimedia Data Streams (MMDS) and to provide a model for MMDS, based upon which, we can draw some performance guarantees such as minimum output rate per query, maximum supported queries per site, and maximum supported streams per site. The Dependency

Theory should also be extended to included dependencies among Multimedia Data Streams.

In this paper we provide a Multimedia Data Streams model. We give the formalization in details of the multimedia data streams problem. The objective is to furnish a formal framework to efficiently design a Multimedia Data Streams (MMDS) schema that achieves an efficient performance in regard with content based retrieval. We also extend the Functional Dependency Theory and the Normalization Framework to handle Multimedia Data Streams. Finally, we discuss algorithmic methods of generating Continuous Multimedia Queries, along with concrete examples for further illustration. Multimedia data streams have potential critical applications.

The paper is organized as follows. Three motivating examples are presented in Section 2. Section 3 provides the mathematical model of the multimedia data streams along with performance bounds. Dependency Theory is extended in section 4. A concluding remark and related work are given in sections 5. Due to lack of space the algorithms for Multimedia Data Streams Continuous Querying and a set of illustrative examples are not included in this paper, but can be found in the full technical report by the authors.

2. Motivating Examples

Consider the following three motivating examples from a health application, another from a security application, and the last from an aircraft database system.

Motivating example 1: Security System: Given a security system, having video cams fixed in hidden places in some secured building or store. Typically a guard or two keep rotating between the captured videos and they interact when something suspicious or a threat shows. However, this suffer two major problems: first, inevitable human error, and second, if a WANTED person shows up, it is less likely that the security guard can pay attention, and hence interfere before-crises, as opposed to after-crises. However, if we have a Multimedia Data Stream management system, that receives the stream of videos captures, or stream of frames, and has a relation (i.e. a table) with images of all WANTED criminals, and another relation with images of all white weapons. Then one can register couple of useful Continuous Queries registered to automate (and hence make it robust) the task of the guards, and help them doing there job and avoid human errors. Such CQs include:
1) Tell me whenever you see an object similar to any white weapon.
2) Tell me whenever you see a WANTED person.
3) Tell me whenever an object (or more) in 2 frames within the last 30 seconds moves in a very "violent" manner.
4) If any of the above, start recording the video on a "Threat" clip for future reference.

Thus, if the guard missed anything, or had to leave his place for whatever reason, the system can help him. Moreover, new things that he were not able to do, such as identifying WANTED criminals is now achievable.

Motivating example 2: Health System: One of the major data streams applications is the health monitoring systems, where the patients heart rate, or body temperature is continuously monitored, and a CQ to report when these values goes beyond the normal values are being continuously executed. Imagine the case for some diseases when an X-Ray photo need to be inspected to follow up progress, or identify anomalies, this is a multimedia stream, that can be fed to a multimedia data stream management system with the appropriate queries. This can be also applied to may be some health experiments done on animals, the animal reaction could be captured as photos or X-Rays, and these could be fed to the MMDSMS for anomalies discovery.

Motivating example 3: Aircraft: Assume we have a multimedia database with multimedia objects being still images of aircraft types. And then we have a multimedia stream of video frames taken by some airport's security cameras. The images are taken inside the terminal, and outside as well. So, the input multimedia stream might contain frames that contain aircrafts. It might be useful to recognize the aircraft type and the location, to make sure that every thing is going accordingly, and there are no violations of the airport rules. This could be modeled as a multimedia continuous query. The advantage of applying a multimedia data stream approach is the ability to have up to date rules, and allowing exceptions in emergency cases, yet, the precision and the 24 hours monitoring advantages are also gained as opposed to rile on human monitoring.

3. Multimedia Data streams model

In this section we will give some definitions then we will provide the mathematical model of the Multimedia Data Streams problem. At the end, we will try to give a concrete example.

3.1) Terminology and Definitions:

After we provide a list of the terminology used, we first define what we mean by Multimedia Data Stream then we

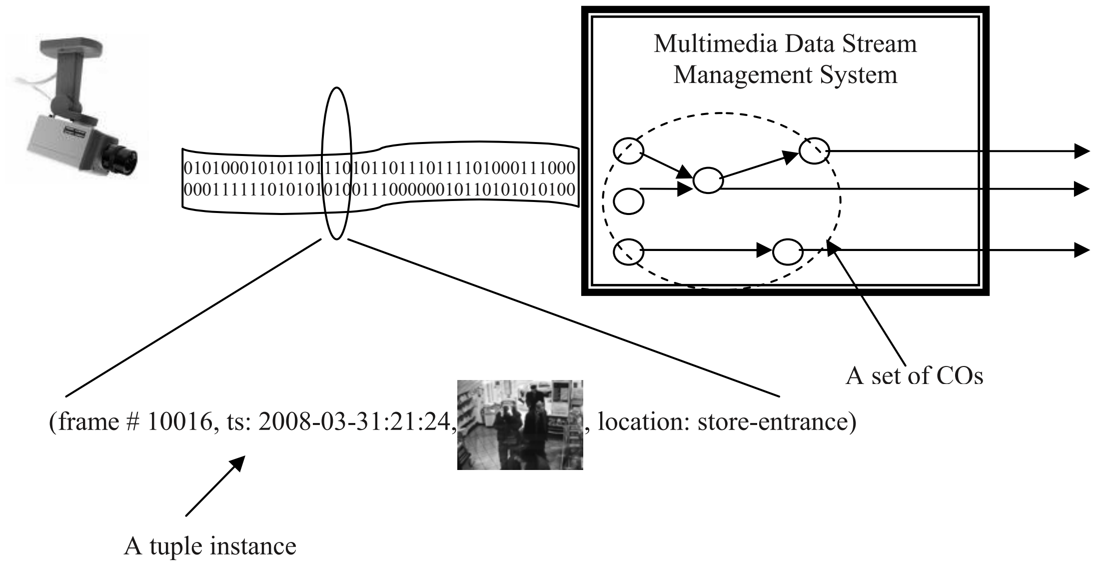


Figure 1: A Multimedia Stream illustration

define some mathematical terms. The definitions 1 through 6 serve the first part, while the rest of the definitions serve the second part.

Terminology, Mathematical terms, variables, and notations:

Let S be a schema, which is a set of streams and multimedia streams S:{S_i}

- Each stream S_i consists of a set of attributes, each is of a certain predefined type T_i, where the type could be any traditional data type (such as number, string, boolean, timestamp, etc) or any multimedia data type (*Micon*), such as: *icon (image)*, *earcon (audio)*, *vicon (video)*, etc. Formally:

$$S_i = \{X_i \mid 0 \leq i \leq n \text{ and } X_i \text{ is of Type } T_i\}$$

- A multimedia continuous query (m-CQ) is a computational structure or graph - that could be seen as a network of operators - that consists of a set of nodes representing the operators, and a set of edges indicating the data flow from one operator to antoher. Each m-CQ should have at least one input multimedia stream, and exactly one output stream.

$$\text{m-CQ: } (S_1, S_2, \ldots S_k) \rightarrow S_{new}$$

In other words, m-CQ = G(V,E); that is each m-CQ is basically a DAG (Directed Graph), where V (the set of vertices) denotes the operators, and E (the set of edges) define the flow of data. The set V must contain at least one multimedia operator (conversion (σ-operator) or fusion operator (ϕ-operator)).

- Let C_i = Cost (m-CQ$_i$) denotes the total cost in terms of CPU time consumed to process a single tuple through all operators of m-CQ$_i$.

- Let S_i = Selectivity (m-CQ$_i$) denotes the selectivity of m-CQ$_i$. Where selectivity is defined as the probability of producing one output tuple for processing one input tuple. This is usually less than one, since each tuple is typically matched against some predicates in each operator. However, some special operators (such as Join operators) have selectivity greater than one.

- Let r_i denotes the average response time of tuples produced by m-CQ$_i$ where the response time is defined as the time span between when the input tuple was first available for processing in the input buffers, and when the corresponding tuple was available for dissemination in the output buffers.

- Let R_i denotes the output rate of producing output tuples of m-CQ$_i$

Now we define precisely what we mean by a Multimedia Data Stream. First, we review what a data stream is, and what a multimedia icon (Micon) is.

Definition 1. Data Stream: A Data Stream is a huge sequence of tuples according to a certain schema that keeps arriving to a Data Stream Management System.
Definition 2. Micon: A Micon is a multimedia icon that could be: text (*ticon*), still image (*icon*), audio (*earcon*), video (*vicon*), or a set of the *Micons*.

Definition 3. Multimedia Data Streams: A Multimedia data stream (MMDS) is a data stream (as defined above) that contains at least one *Micon* as one of its attributes, according to a certain schema.

Please note that the data stream schema refers to the specification of the attributes' (fields') names, and their types, that constitutes any tuple that belongs to that stream, as well as the relationships and/or interactions among different streams. In a sense, a stream in the data streams model is comparable to a relation in the relational database model. However, it should be noted that a relation could be easily modeled as a stream. This is further clarified in the Generalization or Special Case discussion below. An example of a multimedia stream is the video stream captured by some security camera. It would follow the schema below, which is further illustrated in the Figure 1.

Video (frame-number, time-stamp, one-frame-of-video-data, location-id).

Generalization or Special Case? The question at hand now is: does multimedia data stream model generalize the data stream model, or the other way around? In fact, the multimedia data stream is a special case of any data stream. However, the MMDSMS allows both multimedia data streams, and non-multimedia (regular) data streams. Hence, the model of the streams that are allowed to be registered in a MMDSMS should be of the general type. Now, we will furnish some definitions related to a Multimedia Data Stream Management System (MMDSMS), then we will provide some mathematical definitions that will be used in the model we are proposing.

Definition 4. Continuous Query: A Continuous Query (CQ) is a query registered by a user at the MMDSMS that is to be executed - theoretically - forever. If the CQ includes one or more multimedia operators; transformation or fusion operators –as defined later - then the CQ is called a multimedia CQ, or *m-CQ* for short.

Since a data stream in naturally unbounded in size, some query operators can not be applied except for a certain portion (window) of the data, such as aggregate and join operators. Such operators are called windowed operators, and are defined as follows.

Tuples could be of one out of three types: add, delete and the update tuples. Each tuple has both a unique identifier and a timestamp that is used to order the tuples.

Definition 5. Windowed Operators: A windowed operator is to be executed over a certain portion of the data stream (window) as opposed to the whole stream.

The window is used to bound an unbounded computation, such as a join of 2 streams or an average (or any aggregate) function of a stream. The window is specified using a range (size) and a slide (step). A window could be time-based, or tuple-based and could be sliding (if the step is less than the size), or tumbling (if the step equals the size).

Tuple-based window is a window specified in terms of number of tuples. While a time based window is specified in terms of time units. Note that it is easier to deal with tuple-based windows, since the size of each window in terms of number of tuples is fixed, while in time-based windows, each window instance size vary based on the arrival process. Now we are ready to define a MMDSMS.

Definition 6. A Multimedia Data Stream Management System: A MMDSMS is a virtual center (could be physically distributed) that receives the tuples of all streams registered within it (both multimedia streams and regular streams). Users use the MMDSMS's user interface to register a set of CQs and the *m-CQ*s. The MMDSMS responsibility is to continuously process the tuples under real-time constraints with respect to the registered CQs, and to disseminate the results back to the users.

The architecture of the MMDSMS is illustrated in Figure 2. We are now ready to give some mathematical definitions:

Definition 7. Multimedia Tool Box of Type 1: A multimedia tool box of type 1 (*mtb1*) is some black box (ready made) tool box for processing the multimedia data of "any" type, with the property that the performance is optimized with respect to Quality of Service (QoS); which is denoted by the processing time here.

Definition 8. Multimedia Tool Box of Type 2: A multimedia tool box of type 2 (*mtb2*) is some black box (ready made) tool box for processing the multimedia data of "any" type, with the property that the performance is optimized with respect to Quality of Data (*QoD*); which is denoted by the quality of the multimedia objects that belongs to the output stream.

Now, we are ready to formulate our problem.

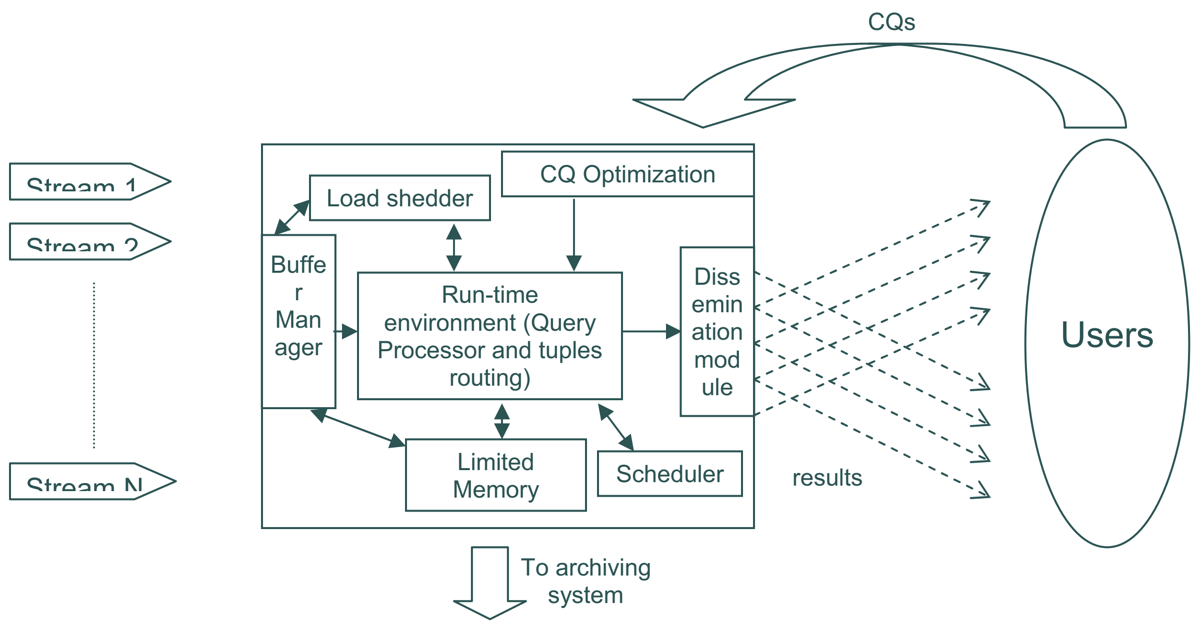


Figure 2: MMDSMS Architecture

3.2) Problem Formalization:

Given a set of streams, with stream S_i a certain multimedia stream of a certain schema $\{X_i\}$, with X^m_i of a *Micon* type, that is T^m_i= *Micon*, and given a Computational Network which constitutes the optimized query execution plan for a set of *CQ*s and *m-CQs*, this Computational Network is composed of both data processing operators (selection, projection, and join) and the special multimedia operators (σ-operator & ϕ-operator), it is required to compute a lower bound on the quality of service provided, measured by the average output rate R_i of each query.

3.3) Performance Bounds:

Typically, there is a tradeoff between QoS and QoD metrics. Hence, we assume an adaptive processing algorithm that given a certain QoS threshold (t_{qos}) and a certain QoD threshold (t_{qod}), the algorithm then starts utilizing mtb1, and when the QoD gets below t_{qod}, the processing switches automatically to *mtb2*. Then as QoS gets below t_{qos}, the algorithm switches back to *mbt1*, where *mtb1* and *mtb2* are as defined above.

Thus, we need to calculate the following bounds

1- Assuming a certain total cost of a given computation network (*C*), what are the bounds of the output rate, as a QoS.
2- Assuming a desired minimum QoS (output rate) for a given computation network, what is the bound on the number of data streams (maximum possible number of streams to be supported).
3- Similarly, and complementary, given a desired minimum QoS, and a given set of registered multimedia data streams, what is the bound on the complexity of the computation network (i.e. the cost *C*).

3.3.a) Output Rate Bound:

Given a computational network of cost *C*, what would be the output rate? Assuming that the filters of the query have selectivity *S*, then the computational network (the optimized continuous queries plan) generates an output tuple with probability *S* every *C* time units. guaranteed output rate is simply

$$R = \frac{S}{C} \qquad \ldots(1)$$

Where *S* is calculated as the multiplication of the individual operators' selectivity along the path of a single tuple, that is:

$$S = \prod_{i \text{ is an operator}} s_i \qquad \ldots(2)$$

Where s_i is the selectivity of operator *i*. And C is the expected cost. Assume that our QoS/QoD adaptive algorithm uses *mtb1* at q of the time, while it uses *mtb2* (1-q) of the time. And the cost of *mtb1* and *mtb2* at operator i are c_{i1} and c_{i2} respectively. Then the expected cost C is calculated as the cost of the first operator plus the cost of the second operator if it made it through the first operator and so on. Thus, the total cost could be written as follows:

$$C = \sum_i \left[\left(q c_{i1} + (1-q) c_{i2} \right) \left(\prod_{j=1}^{i-1} s_j \right) \right] \quad \ldots(3)$$

3.3.b) Number of Supported Streams Bound:

Assuming a desired minimum QoS (output rate) to be R_{min}, for a given data stream management system, what is the bound on the number of data streams (maximum possible number of streams to be supported)? Adding new data stream means accepting new queries to the system, which means an increase in the cost of the computational network. Assuming that the cost of the computation network is a function of the number of streams n, that is $C = f(n)$, then n is governed by the rule:

$$R = \frac{S}{C} \geq R_{\min}$$

$$C \leq \frac{S}{R_{\min}}$$

$$\text{that is}: n \leq f^{-1}\left(\frac{S}{R_{\min}}\right) \quad \ldots(4)$$

However, the relation between C and n (i.e. f) is not a simple one, that it might not be computationally feasible to calculate f^{1}. So, to work around this problem, one can calculate an average query cost, and thus, one can base the bound based on this average query cost. Say the average query cost is C_{avg} then

$$R = \frac{S}{n.C_{avg}} \geq R_{\min}$$

$$n \leq \frac{S}{C_{avg}.R_{\min}} \quad \ldots(5)$$

3.3.c) Computation Network Complexity Bound:

Similarly, and complementary, given a desired minimum QoS, and a given set of registered multimedia data streams, what is the bound on the complexity of the computation network (i.e. the cost C)? Trivially this bound is governed by equation (4) above.

3.3.d) A Concrete Example:

Consider the following scheme: Imagine a stream of video frames captured from a certain security monitoring camera:

Video (frame-number, time-stamp, one-frame-of-video-data, location_id)

The frame-number attribute here becomes the unique identifier of the video frame (a tuple in the stream).

A derived relation can be computed from the above relation, for example,

Aircraft (frame-number, one-frame-of-video-data, type-of-aircraft)

There is a multimedia dependency from video-data to type-of-aircraft. The dependency relations tell us how these streams are related.

A possible $m\text{-}CQ_1$ could be:

```
Notify me every 30 seconds of all video frames within the past 30 seconds that contains a weapon that you receive at Video stream, from the entrance or the exit locations.
```

Clearly this *m-CQ* includes the following operations (or manipulations) over the Video multimedia streams:

1) Tumbling Window: a window of size 30 sec, and step 30 sec should be first defined.

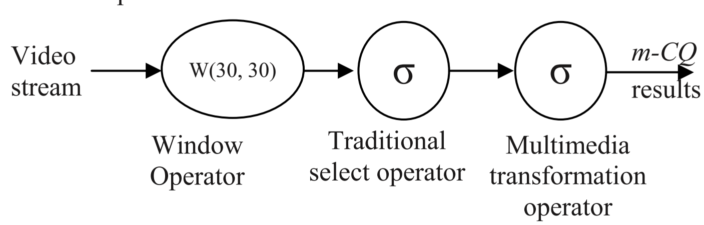


Figure 3: Illustration of m-CQ_1

2) Over the previous window, there is a filter that filters out the frames not coming from the entrance or exit locations. Or in data base language, a Select operator with predicate set such that the location_id

field equals to that of the entrance, or that of the exit sites.

3) For all frames that go through the above 2 operators, some mtb1 or mtb2 should be utilized to detect weapon objects. This is considered a transformation operator (σ-operator).

The corresponding Query Graph G($m\text{-}CQ_1$) (or the Computational structure in other words) is illustrated in Figure 3.

Now, assuming the cost of this computational network/structure is C_1 then, the MMDSMS can provide the following guarantees:

1) If this is the only stream registered, and this is the only m-CQ registered, then the QoS (output rate) is guaranteed to be:

$$R_1 = \frac{S_1}{C_1}$$

Where S_1 is the selectivity of the *m-CQ*, and C_1 is the cost.

2) Assuming a desired minimum QoS of R_1 then the maximum possible number of streams that could be supported is limited by the relation that:

$$\text{overhead} = \frac{S_1}{R_1} - C_1$$

Where the overhead represent the additional cost the additional streams processing might induce.

3) Similarly, assuming this stream to be the only stream, then the maximum number of queries is bounded with the same exact relation above (in bound 2). The difference is in the way this overhead is calculated. In the bound of number of streams, the overhead should reflect the cost of adding support for new streams, while here the overhead is basically the cost of the new operators needed to process the additional m-CQs.

Another possible $m\text{-}CQ_2$ over the MMDS schema given above could be:

```
Notify me every minute of all video
frames within the past 10 seconds
that contains an aircraft of type
"XYZ".
```

Clearly this *m-CQ* includes similar window operator. However, filtering here requires an additional join operator. Here the multimedia dependency plays a role. Since there is a dependency between the video data, and the aircraft type, we can simply use the derived stream, instead of the original stream, and then we can join based on the *frame_number* attribute to get the rest of the data. The query is illustrated in the figure 4.

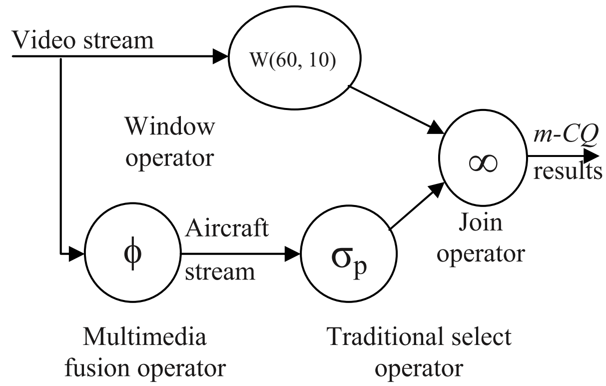


Figure 4: Illustration of m-CQ2

4. Extended Dependency Theory

In this section, we extend the Multimedia Dependency Definitions from [4] to generalize to the multimedia streams case. As in [4], in order to evaluate the similarity between multimedia objects of two tuples, we need to use tuple-distance functions. The tuple-distance function summarizes the different distance functions applied to the elements of the 2 tuples under comparison. So, basically, a distance function is applied on corresponding attributes of the two tuples, then a function that takes these distances as input, will produce the final distance between the two tuples is called the tuple-distance function. What about tuples timestamps? Clearly, the notion of sliding windows on data streams can be utilized as follows for distance functions: if the two tuples belong to the same window then they might be considered for similarity, o.w., they may not. This way, we bound the calculations needed. And since the window specifications are set by the user, this way of utilizing the windows for calculating tuples distances does reflects user best interest. We are ready now to give some definitions.

Definition 4.1. MS-Similarity: Let ϖ be a tuple distance function on a relation R, and t be a maximum distance threshold, and x and y be two tuples in R, we then say that x is type MS-similar (Multimedia Stream Similar) to y with respect to ϖ, denoted x $\cong^{\varpi(t)}$ y iff x and y belongs to the same window, and $\varpi(x, y) \le t$

Definition 4.2. (*type-MS functional dependency*): Let R be a relation with attribute set U, and X, Y $\subseteq$ U.

$X_{g1(t')} \rightarrow Y_{g2(t'')}$ is a *type-MS functional dependency* (MSFD) relation *iff* for any two tuples t1 and t2 in R, if $t_1[X] \cong_{g1(t')} t1[X]$, then $t1[Y] \cong_{g2(t'')} t2[Y]$, where t' and t'' are similarity thresholds, and $t_i[X]$ is the projection of the tuple t_i over the set of attributes X, and similarly is $t_i[Y]$.

In English, this definition typically says: there is a *type-MS functional dependency* (MSFD) from set of attributes X (under MS-Similarity g1(t')) and the set of attributes Y (under MS-Similarity g2(t'')) if and only if, any two tuples that are MS-similar under g1(t') this implies that these very two tuples are also MS-Similar under g2(t''). It reads that the set of attributes Y are *type-MS functionally dependent* on attributes X.

Using these definitions, the set of inference rules presented in [4] still holds for multimedia data streams case. And similarly the type-M Multi-valued dependency (MMD) can be generalized to the *type-MS Multi-Valued dependency* (MSMD) and the also the type-M join dependency (MJD) to the *type-MS Join dependency* (MSJD).These set of functional dependency can be utilized to normalize the schema at design time.

5. Discussion

In this paper we presented a novel model to model multimedia data streams as a potential tool for many useful applications such as security and health systems as described in the motivating examples earlier.

To the best knowledge of the authors of this paper, this is the first work that tries to model multimedia data streams. However, there is a lot of research done on both multimedia databases, and on data streams separately and orthogonally. For multimedia databases, the work [4] is the inspiring work of this paper, as in [4] the dependency theory was generalized for multimedia databases, and a normalization framework was proposed. Other work in content-based retrieval and algorithms has also been proposed in the multimedia database literature. The literature of data streams is focused on performance fine tuning; minimizing response time, as a QoS metric, or improving freshness, as a QoD metric [7]. Also, scaling with burst arrivals of data, and peak loads gained some attention (Load Shedding) [6]. Also, some work was proposed for Mining Data Streams [12].

References:

[1] D. J. Abadi, Y. Ahmad, M. Balazinska, U. C etintemel, M. Cherniack, J.-H. Hwang, W. Lindner, A. Maskey, A. Rasin, E. Ryvkina,N. Tatbul, Y.Xing, and S. B. Zdonik. The design of the borealis stream processing engine. In *CIDR*, pages 277–289, 2005.

[2] Daniel J. Abadi, Don Carney, Ugur C¸ etintemel, Mitch Cherniack, Christian Convey, Sangdon Lee, Michael Stonebraker, Nesime Tatbul, and Stan Zdonik. Aurora: a new model and architecture for data stream management. *The VLDB Journal*, 12(2):120–139, 2003.

[3] Arvind Arasu, Brian Babcock, Shivnath Babu, Mayur Datar, Keith Ito, Itaru Nishizawa, Justin Rosenstein, and Jennifer Widom. Stream: The stanford stream data manager (demonstration description). In *SIGMOD '03: Proceedings of the 2003 ACM SIGMOD international conference on Management of data*, New York, NY, USA, 2003.

[4] S.K. Chang, V. Deufemia, and G. Polese. A Normalization Framework for Multimedia Databases. *IEEE Transactions on Knowledge and Data Engineering*, v. 19, issue 12, Dec 2007.

[5] http://www.cs.cmu.edu/~biglou/research.html, 2008

[6] N. Tatbul and S. Zdonik. Window-aware load shedding for aggregation queries over data streams. In *Proc. of VLDB Conference*, 2006.

[7] Mohamed A. Sharaf, Panos K. Chrysanthis, Alexandros Labrinidis, and Kirk Pruhs. Algorithms and Metrics for Processing Multiple Heterogeneous Continuous Queries. *ACM Transactions in Database Systems (TODS)*, March 2008.

[8] S.K. Chang, Gennaro Costagliola, Erland Jungert and Karin Camara. Intelligent Querying Techniques for Sensor Data Fusion. Chapter from the book *Intelligent Techniques for Warehousing and Mining Data Streams*.

[9] Jianjun Chen, David J. DeWitt, Feng Tian, and Yuan Wang. Niagaracq: a scalable continuous query system for internet databases. In *SIGMOD '00: Proceedings of the 2000 ACMSIGMOD international conference on Management of data*, pages 379–390.ACM, 2000.

[10] http://www.coral8.com/, 2004.

[11] http://www.streambase.com, 2006.

[12] M. Gaber, A. Zaslavsky, and S. Krishnaswamy. Mining data streams: a review. *SIGMOD Record*, 34(2), 2005.

Area-Based Data Mining for Querying on Spatial Data

Jeang-Kuo Chen*, Wen-Ting Chien
Department of Information Management, Chaoyang University of Technology
jkchen@cyut.edu.tw, s9414633@cyut.edu.tw

ABSTRACT

Data mining can be applied to many fields for extracting useful information from a lot of original data. As the R-tree is the most popular index for spatial data access, we try to apply data mining to R-tree to extract some characteristics of spatial object. Three mining items are proposed to mine area-based data of spatial objects. They are *the minimum rectangle of covering overlap regions* (*MRCOR*), *the overlap region with the maximum area value* (*ORMAV*), and *the region overlapped by the most objects* (*ROMO*). Every node in R-tree has the three mining items. The values of these items are pre-calculated as follows. First, the values in leaf nodes are computed. Then the values in non-leaf nodes are processed with an upside-down level-by-level way. The values in each parent node are computed by collecting the related values in the child nodes of this parent node. Finally, the root gathers together the total information of each mining item. When querying area-based information of spatial objects, some available mining items at the upper level of R-tree can omit the overhead of searching and computing at the lower level of R-tree.
Keyword: data mining, spatial data, R-tree, overlap

1. Introduction

Data mining means to extract meaningful potential data from a large number of old data [13]. The application of data mining is quite wide such as electronic commerce, financial application, customer relationship management, cross selling analysis, decision support, web page retrieval, geographic information system, image retrieval, etc [6, 14, 15, 18, 21]. Data mining can be applied to query for object data to expand the benefit of searching result. Some representative multi-dimensional data indices such as Gird file [16], K-D Tree [3], K-D-B Tree [17], and R-Tree [12], can be classified into (1) point data index and (2) spatial data index [10]. The former is built for point data such as Grid file [16], K-D Tree [3], and K-D-B Tree [17]. A point substitutes an object in the data space which is partitioned into several subspaces to index the points in these subspaces. The latter is built for spatial data such as R-tree [12], R^+-tree [18], R^*-tree [2], and X-tree [4].

Some researches of data mining application in index have been proposed [1, 3, 5, 6, 8, 9, 11, 19] such as STING (STatistical INformation Grid) [19] for point data, but rare related researches for spatial data [6]. We propose three data mining items for R-tree to supply special area-based data such as area value, overlap region of spatial object, etc. Except the traditional data such as object location, object range, the access of spatial objects also contains some statistical data such as area value, overlap region. There are different computing ways for the demand of mining items at different range. For example, the area of available farms in a town is different from that in the whole country, but the result of the latter must be the accumulation of many results of the formers. Therefore, the data mining items of each node in R-tree can be computed in advance. When statistical information is needed, it is not necessary to collect the data mining items of all nodes in R-tree because those data is available at the nodes of upper level. The time of searching and computing at the nodes of lower level can be omitted.

2. Related work

Proposed by Guttamn [12], the R-tree is a multi-dimensional, dynamic searching index for spatial object. An R-tree consists of leaf nodes and non-leaf nodes. Each node is composed of several entries and an entry represents the information of a specific object or a node. The format of an entry in a leaf node is (*I, tuple-identifier*). The *tuple-identifier* is a pointer to address a spatial object while the *I* denotes a minimum bounding rectangle (MBR) that covers a spatial object tightly. The format of an entry in a non-leaf node is (*I, child-pointer*). The *child-pointer* is a pointer to point a child node of a parent node containing the *child-pointer* while the *I* denotes an MBR that covers the child node tightly. Let *M* be the maximum number of entries in a node and *m* be the minimum number of entries in a node. The R-tree satisfies the following properties. (1) The root node has at least two children unless it is a leaf node. (2) Each node contains entries between *m* and *M* unless it is the root. (3) All leaf nodes appear on the same level.
An example with *m*=2 and *M*=3 is shown in Figure 1. Figure 1(a) shows the distribution of spatial objects while Figure 1(b) shows the corresponding R-tree. The root node 'a' contains two child nodes 'b' and 'c', the node 'b' contains

three child nodes ‘d’, ‘e’, and ‘f’, and the node ‘c’ contains two child nodes ‘g’ and ‘h’. The leaf nodes ‘d’, ‘e’, ‘f’, ‘g’, and ‘h’ have objects (***A***, ***B***), (***C***, ***D***), (***E***, ***F***, ***G***), (***H***, ***I***), and (***J***, ***K***), respectively.

The most obvious difference between spatial object and point object is that the former has variation of object area while the area of the latter is always zero. Therefore, the object-area related mining is essential for spatial object application. To mine area related information for spatial object, Chen and Chien [6] proposed five mining items, *object area* (a), *mean of object area* (m_a), *standard deviation of object area* (s_a), *minimum area* (min_a), and *maximum area* (max_a).

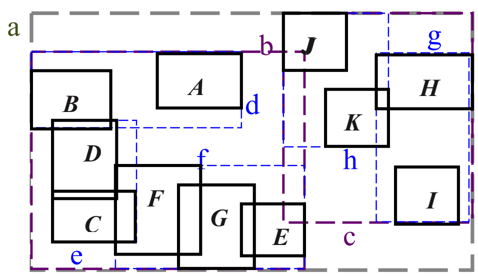


Figure 1(a). Distribution of spatial objects.

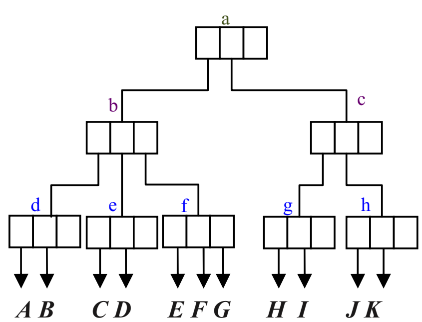


Figure 1(b). The corresponding R-tree.

Figure 1. Spatial objects and R-tree.

3. Data mining Items

In addition to the mining items in [6], we hope to supply more useful and advanced information for some special applications. With mining items, each node in R-tree can suppose rich information. The values of mining items are pre-calculated as follows. First, the related values of mining items in leaf nodes are computed. Then the related values in non-leaf nodes are processed repetitively with an upward direction level-by-level. Finally, the root node gathers together the total information of each mining item in non-root nodes. The object-area related mining is an important theme for spatial object application. Besides, the overlap is a special characteristic of spatial object. We propose three mining items for R-tree to supply special area-based data. The three mining items are *the minimum rectangle of covering overlap regions* (*MRCOR*), *the overlap region with the maximum area value* (*ORMAV*), and *the region overlapped by the most objects* (*ROMO*). The descriptions of these data mining items are listed as follows.

(1) *MRCOR*: the minimum rectangle of covering overlap regions in a specific range

MRCOR is composed of two coordinates (x_1, y_1) and (x_2, y_2) that represent the lower left and the upper right coordinates of a bounding rectangle, respectively. With *MRCOR*, we can find all the overlap regions in a specific range. When *MRCOR* is recorded in a leaf node, the formula is $MRCOR = \bigcup(\forall_{i \neq j}(O_i \cap O_j))$, where O_i and O_j are two different objects which overlap. When *MRCOR* is recorded in a non-leaf node, the formula is $MRCOR = \bigcup_i (MRCOR_i)$, where $MRCOR_i$ is a smaller *MRCOR* of objects in a certain node i which is a child of a node p that includes *MRCOR*. Thus, *MRCOR* gathers $MRCOR_i$ of each child node i of node p within a bounding rectangle. Each $MRCOR_i$ again is the minimum rectangle which covers $MRCOR_j$ of each child node j of node i. According to this recursion way, each related data about the minimum rectangle of covering overlap regions in each node can be determined. Figure 2 shows an example of *MRCOR*. In Figure 2(a), there are five spatial objects O_1, O_2, O_3, O_4,

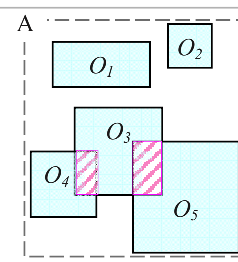


Figure 2(a). Distribution of spatial objects.

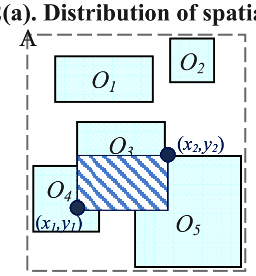


Figure 2(b). *MRCOR* of region A.

Figure 2. An example of *MRCOR*.

and O_5 with the two overlap regions (O_3, O_4) and (O_3, O_5) presented by the rectangles with skew lines. Figure 2(b) shows the *MRCOR* of region A with reverse skew lines and *MRCOR* = ((x_1, y_1), (x_2, y_2)).

(2) *ORMAV*: the overlap region with the maximum area value in a specific range

With *ORMAV*, we can find the largest region from all overlap regions in a specific range. If the candidate region has more than one, select the one overlapped by the most objects. Therefore, *ORMAV* includes two parameters ($Rect_O$, n_O). The $Rect_O$ is composed of two coordinates (x_1, y_1) and (x_2, y_2) to represent the lower left and the upper right coordinates of the selected region, respectively. The n_O is the number of objects overlapped with $Rect_O$. When *ORMAV* is recorded in a leaf node, the formula is $ORMAV = \max_{i \neq j}(\forall(O_i \cap O_j))$, where O_i and O_j are two different objects. When *ORMAV* is recorded in a non-leaf node, the formula is $ORMAV = \max_i(ORMAV_i)$, where $ORMAV_i$ is the overlap region with the maximum area value in a certain node *i* which is a child of a node *p* that includes *ORMAV*. Thus, *ORMAV* is finding the maximum $ORMAV_i$ of each child node *i* of node *p*. Each $ORMAV_i$ again is determined by finding the maximum $ORMAV_j$ of each child node *j* of node *i*. According to this recursion way, each related data about the overlap region with the maximum area value in each node can be determined. Figure 3 shows an example of *ORMAV*. In Figure 3(a), there are five spatial objects O_1, O_2, O_3, O_4, and O_5 with the five overlap regions (O_1, O_3), (O_1, O_4), (O_3, O_4), (O_1, O_3, O_4), and (O_2, O_5) presented by the rectangles with skew lines. These five overlap regions have the same area value. The second parameter, n_O, is compared to choose the region overlapped by the most objects. Therefore, the selected overlap region is (O_1, O_3, O_4). Figure 3(b) shows the *ORMAV* of region A with reverse skew lines and *ORMAV* = (((x_1, y_1), (x_2, y_2)), 3).

(3) *ROMO*: the region overlapped by the most objects in a specific range

Using *ROMO*, we can find the region overlapped by the most objects in a specific range. If the candidate region has more than one, select the one with the maximum area value. Therefore, *ROMO* includes two parameters (n_R, $Rect_R$). The n_R is the number of objects overlapped with $Rect_R$. The $Rect_R$ is composed of two coordinates (x_1, y_1) and (x_2, y_2) to represent the lower left and the upper right coordinates of the selected region, respectively. When *ROMO* is recorded in a leaf node, the formula is $ROMO = \max_{i \neq j}(count(O_i \cap O_j))$, where O_i and O_j are two different objects. When *ROMO* is recorded in a non-leaf node, the formula is $ROMO = \max_i(ROMO_i)$, where $ROMO_i$ is the region overlapped by the most objects in a certain node *i* which is a child of a node *p* that includes *ROMO*. Thus, *ROMO* is finding the maximum $ROMO_i$ of each child node *i* of node *p*. Each $ROMO_i$ again is determined by finding the maximum $ROMO_j$ of each child node *j* of node *i*. According to this recursion way, each related data about the region overlapped by the most objects in each node can be determined. Figure 4 shows an example

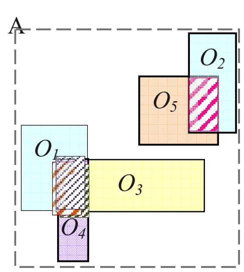


Figure 3(a). Distribution of spatial objects.

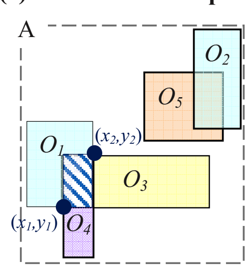


Figure 3(b). *ORMAV* of region A.

Figure 3. An example of *ORMAV*.

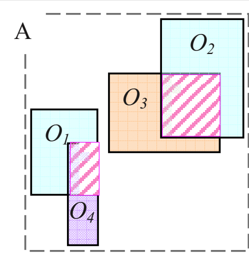


Figure 4(a). Distribution of spatial objects.

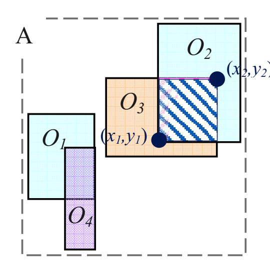


Figure 4(b). *ROMO* of region A.

Figure 4. An example of *ROMO.*

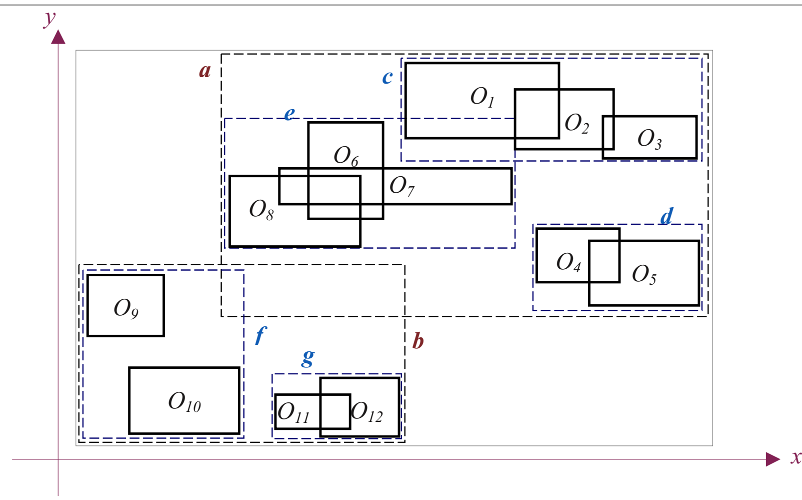


Figure 5(a). Distribution of soil heavy-metal pollution areas.

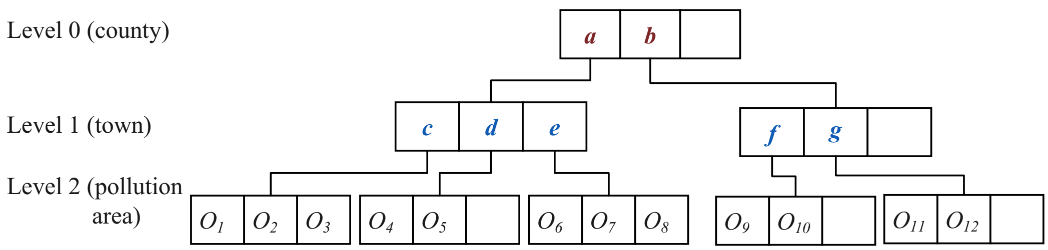


Figure 5(b). The corresponding 3-level R-tree.

Figure 5. An example of data mining for soil heavy-metal pollution area.

of *ROMO*. In Figure 4(a), there are four spatial objects O_1, O_2, O_3, and O_4 with the two overlap regions (O_1, O_4) and (O_2, O_3) presented by the rectangles with skew lines. These two overlap regions have the same overlapped objects. The second parameter, $Rect_R$, is compared to choose the overlap region with the maximum area value. Therefore, the selected region is (O_2, O_3). Figure 4(b) shows the *ROMO* of region A with reverse skew lines and *ROMO* = (2, ((x_1, y_1), (x_2, y_2))).

4. Data mining arrangement

We modify the data structure of R-tree to include data mining items first. The data structure of R-tree is modified as (E[$1, 2, \ldots, M$], *MRCOR*, *ORMAV*, *ROMO*), where the $E[i]$ represents the entry of child node and M is the maximum entries of each node. When the child node is a non-leaf node, every $E[i]$ includes a couple of (I_i, *child-pointer*$_i$). When the child node is a leaf node, every $E[i]$ includes a couple of (I_i, *tuple-identifier*$_i$). The values of these items are pre-calculated as follows. First, the values in leaf nodes are computed. Then the values in non-leaf nodes are processed with an upside-down level-by-level away. The values in each parent node are computed by collecting the related values in the child nodes of this parent node. Finally, the root gathers together the total information of each mining item.

We give an example to explain the R-tree associated with the data mining items bringing additional benefit. Figure 5 shows that people mine the soil heavy-metal pollution area for certain two counties in Taiwan. Figure 5(a) shows the distribution of soil heavy-metal pollution areas. Figure 5(b) shows the corresponding 3-level spatial object index of R-tree and each node contains child nodes between two and three. The root node contains two child nodes '*a*' and '*b*' whose represent the soil heavy-metal pollution areas of *a* county and *b* county, respectively. The node '*a*' contains three child nodes '***c***', '***d***', and '***e***' whose represent the soil heavy-metal pollution areas of *c* town, *d* town, and *e* town, respectively. The node '***b***' contains two child nodes '***f***' and '***g***' whose represent the soil heavy-metal pollution areas of *f* town and *g* town, respectively. The leaf nodes '***c***', '***d***', '***e***', '***f***', and '***g***' have objects (O_1, O_2, O_3), (O_4, O_5), (O_6, O_7, O_8), (O_9, O_{10}), and (O_{11}, O_{12}) whose represent each soil heavy-metal pollution area, respectively. The related values in leaf nodes are abstracted from objects. Then these values are gathered repetitively for non-leaf nodes with an

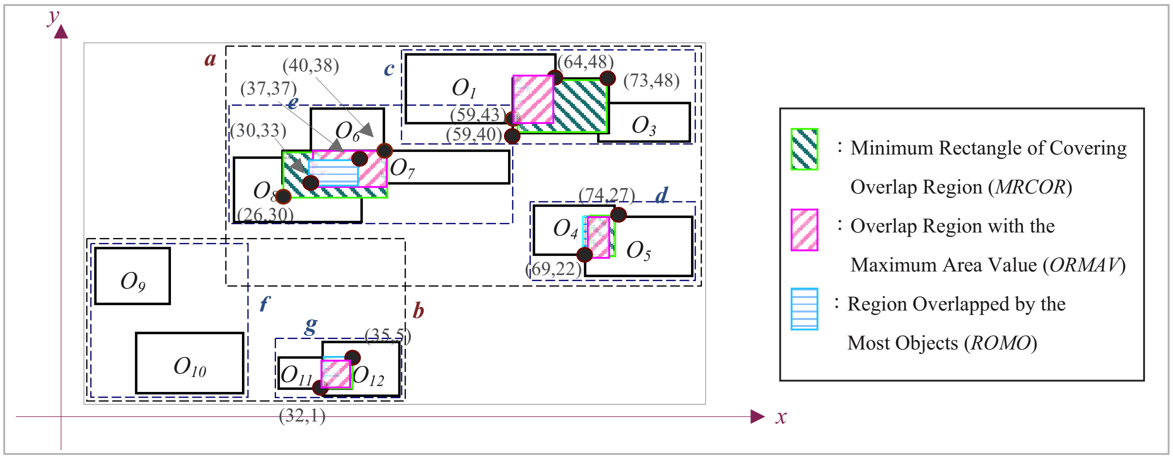


Figure 6. The multiple pollution areas.

Table 1. The statistical information of each spatial object at level 2.

Level 2	O_1	O_2	O_3	O_4	O_5	O_6	O_7	O_8	O_9	O_{10}	O_{11}	O_{12}
MRCOR	*null*	*null*	*null*	*null*	*null*	*null*	*null*	*null*	*null*	*null*	*null*	*null*
ORMAV	*null*	*null*	*null*	*null*	*null*	*null*	*null*	*null*	*null*	*null*	*null*	*null*
ROMO	*null*	*null*	*null*	*null*	*null*	*null*	*null*	*null*	*null*	*null*	*null*	*null*

Table 2. The statistical information of each spatial object at level 1.

Level 1	*c*	*d*	*e*	*f*	*g*
MRCOR	((59, 40), (73, 48))	((69, 22), (74, 27))	((26, 30), (40, 38))	*null*	((32, 1), (35, 5))
ORMAV	(((59, 43), (64, 48)), 2)	(((69, 22), (74, 27)), 2)	(((33, 33), (40, 38)), 2)	*null*	(((32, 1), (35, 5)), 2)
ROMO	(2, ((59, 43), (64, 48)))	(2, ((69, 22), (74, 27)))	(3, ((30, 33), (37, 37)))	*null*	(2, ((32, 1), (35, 5)))

Table 3. The statistical information of each spatial object at level 0.

Level 0	*a*	*b*
MRCOR	((26, 22), (74, 48))	((32, 1), (35, 5))
ORMAV	(((33, 33), (40, 38)), 2)	(((32, 1), (35, 5)), 2)
ROMO	(3, ((30, 33), (37, 37)))	(2, ((32, 1), (35, 5)))

upside-down level-by-level way. At last the root node gathers together the summary statistical information of each mining item. Therefore, each node has statistical and spatial information as Table1, Table 2, and Table 3 where the "*null*" represents the pollution area with single soil heavy-metal pollution. Figure 6 show the multiple pollution areas associated with mining items *MRCOR*, *ORMAV*, and *ROMO*.

Appling data mining to R-tree make each node has plentiful information which is according to spatial position of node. For example, the node '*c*' records related soil heavy-metal pollution area information of *c* town. The multiple pollution areas in *c* town are distributed over the spatial coordinate ((59, 40), (73, 48)). Also, the largest and the most serious multiple pollution areas are distributed over the spatial coordinate ((59, 40), (73, 48)). Following this rule, the multiple pollution areas in *d* town are distributed over the spatial coordinate ((69, 22), (74, 27)). This region is the largest and the most serious multiple pollution areas the same. The multiple pollution areas in *e* town are distributed over the spatial coordinate ((26, 30), (40, 38)). The largest multiple pollution area is distributed over the spatial coordinate ((33, 33), (40, 38)) and the most serious multiple pollution area is distributed over the spatial coordinate ((30, 33), (37, 37)). After obtaining the related information of *c* town, *d* town, and *e* town, we can determine the related information of *a* county. The multiple pollution areas in *a* county are distributed over the spatial coordinate ((26, 22), (74, 48)). The largest multiple pollution area is distributed over the spatial coordinate ((33, 33), (40, 38)) and the most serious multiple pollution area is distributed over the spatial coordinate ((30, 33), (37, 37)).

Each node of R-tree has statistical and spatial information. We can obtain information quickly through corresponding specific node. For example, people want to know the pollution area related information in *a* county. It only needs to access the corresponding node '***a***' to obtain statistical information quickly. If people want to know the detail, only needs to drill down the pollution area of *c* town, *d* town, or *e* town to obtain information. It doesn't need to compute all related pollution area information with upward direction level-by-level repeatedly and these queries can be answered quickly. It reduces operation time, and gathers statistical and spatial information quickly.

5. Conclusion

We propose three area-based mining items *MRCOR*, *ORMAV*, and *ROMO* for advanced applications. With the three mining items, each node in R-tree can suppose rich information. For example, we can get the minimum rectangle of covering overlap regions in a specific range by *MRCOR*. We can get the overlap region with the maximum area value in a specific range by *ORMAV*. We can get the region overlapped by the most objects in a specific range by *ROMO*. The related values for the mining items in leaf nodes are abstracted from objects. Then these values are gathered repetitively for non-leaf nodes with an upward direction level-by-level. The root node finally gathers together the total information from all its child nodes. Each node in R-tree has statistical and spatial information. The nodes at higher level have aggregate information from the nodes at lower level. That reduces searching and computing time because we can obtain necessary information from the nodes at higher levels.

6. Reference

[1] R. Agrawal, J. Gehrke, D. Gunopulos, and P. Raghavan, "Automatic Subspace Clustering of High Dimensional Data for Data Mining Applications", Int. Conf. Management of Data, pp.94-105, 1998.

[2] N. Beckmann, H. P. Kriegel, R. Schneider, and B. Seeger, "The R^*-tree: An Efficient and Robust Access Method for Points and Rectangles", In Proc. of ACM SIGMOD Int. Conf. Management of Data, pp. 322-331, 1990.

[3] J. L. Bentley, "Multidimensional Binary Search Trees Used for Associative Searching", Communications of the ACM, Vol. 18, pp.509-517, 1975.

[4] S. Berchtold, D. A. Keim, and H. P. Kriegel, "The X-tree: An Index Structure for High-Dimensional Data", In Proc. 22th Int. Conf. on VLDB, pp. 28-39, 1996.

[5] J. W. Chang and D. S. Jin, "A New Cell-Based Clustering Method for Large, High-Dimensional Data in Data Mining Applications", In Proc. of the 2002 ACM symposium on Applied computing, pp.503-507, 2002.

[6] J. K. Chen and W. T. Chien, "The Research for Data Mining Applied to Spatial Data Index", National Computer Symposium, pp.504-514, 2007.

[7] R. Cooley, B. Mobasher, and J. Srivastava, "Web Mining: Information and Pattern Discovery on the World Wide Web", IEEE Int. Conf. of Tools with Artificial Intelligence, pp.558-567, 1997.

[8] V. Estivill-Castro and I. Lee, "AMOEBA: Hierarchical Clustering Based on Spatial Proximity Using Delaunay Diagram", In Proc. 9th Int. Spatial Data Handling, pp.10-12, 2000.

[9] C. Faloutsos and K. Lin, "Fastmap: A Fast Algorithm for Indexing, Data Mining and Visualization of Traditional and Multimedia Datasets", ACM SIGMODE Conf., pp. 163-174, 1995.

[10] V. Gaede and O. Gunther, "Multidimensional Access Methods", ACM Computing Surveys, pp. 170-231, 1997.

[11] S. Guha, R. Rastogi, and K. Shim, "ROCK: A Robust Clustering Algorithm for Categorical Attribute", In Proc. 1999 Int. Conf. Data Engineering, pp.512-521, 1999.

[12] A. Guttman, "R-Tree: A Dynamic Index Structure for Spatial Searching", In Proc. ACM SIGMODE, pp.47-57, 1984.

[13] J. Han and M. Kamber, "Data Mining: Concepts and Techniques", Morgan Kaufmann Publisher, 2001.

[14] C. Kleissner, "Data Mining for the Enterprise", In Proc. 30th Int. Conf. on Vol. 7, pp.295-304, 1998.

[15] R. Kosla and H. Blockeel, "Web Mining Research: A Survey", ACM SIGKDD Explorations, Vol. 2, pp.1-15, 2000.

[16] J. Nievergelt, H. Hinterberger, and K.C. Sevcik, "The Grid File: An Adaptable, Symmetric Multikey File Structure", ACM Trans. Database System Vol. 9, No. 1, pp.38-71, 1984.

[17] J. T. Robinson, "The K-D-B-tree: A Search Structure for Large Multidimensional Dynamic Indexes", In Proc. of the ACM SIGMODE Int. Conf. on Management of Data, pp.10-18, 1981.

[18] T. Sellis, N. Roussopoulos, and C. Faloutsos, "The R^+-Tree: A Dynamic Index for Multi-Dimensional Objects", In Proc. 13th VLDB conf., pp. 507-518, 1987.

[19] W. Wang, J. Yang, and R. Muntz, "STING: A Statistical Information Grid Approach to Spatial Data Mining", In Proc. 23th Int. Conf. on VLBD, pp.186-195, 1997.

[20] J. Xing, C. Yingkun, X. Kunqing, M. Xiujun, S. Yuxiang, and C. Cuo, "A Novel Method to Integrate Spatial Data Mining and Geographic Information System", IEEE Int. on Vol. 2, pp. 795-798, 2005.

[21] D. Zhang and L. Zhou, "Discovering Golden Nuggets: Data Mining in Financial Application", IEEE Transactions on Vol. 34, pp.513-522, 2004.

Proceedings
International Workshop on
Visual Languages Computing
(VLC 2008)

Co-Editors

Giuseppe Polese, *University of Salerno, Italy*
Giuliana Vitiello, *University of Salerno, Italy*

Forestry Visualization and Analysis Using Lidar Data

Minho Park
Stephen F. Austin State University
Department of Computer Science
Nacogdoches, Texas
parkm@sfasu.edu

Ian Brooks
Stephen F. Austin State University
Department of Computer Science
Nacogdoches, Texas
z_brooksir@titan.sfasu.edu

Abstract

We present an open source-based visualization and analysis system environment for a forestry-lidar dataset. In a nutshell, the forestry-lidar dataset is a collection of spatial points representing forest features on the ground. There are numerous different methodologies and techniques for visualizing, managing, and manipulating lidar data sets. But most system environments fall short in terms of usability, functionality and flexibility. Our goal is to study efficient methods for managing and manipulating lidar dataset by focusing on effective ways to visualize and analyze forest structures. To support the system environment, the open source 3D Blender software is employed as a comprehensive visualization tool.

1. Introduction

A lidar dataset, usually referred to as "lidar point cloud," is composed of numerous points with the x-, y- and z-coordinate of each point recorded. Data derived from lidar points can be used to measure forest attributes at individual tree level (total height and crown height, etc.) or stand level (volume, basal area and biomass, etc.) [1-4]. There are numerous 3D/2D-based visualization systems [5-7] for the lidar dataset.

Our study follows the same line but enhances the capability and flexibility of extracting and quantifying forest features from lidar points. Most systems attempt to provide a fully functional environment for visualization but fall short in terms of usability, functionality and flexibility. For example, many research groups employ commercial software as a visualization tool, however most systems do not provide additional capabilities, such as analysis and statistics, and most computer algorithms for visualization are hidden from the user.

Our goal is to study efficient methods for managing and manipulating lidar datasets by focusing on effective ways to visualize and analyze forest structure. We develop an open source-based visualization and analysis environment for forestry-lidar data. Our system is based on volumetric pixel (voxel) structure, which divides lidar point cloud into cubes in 3D space. Included in the system is the ability to adjust the height of a cube automatically, so that the forest structure of canopy, stem and understory can be separated. Application programs for statistics and characteristics of forestry data is created as well. With these tools, forest attributes can be quantified and fused with other remote sensing data for better classification of forest type.

In addition, based on lidar point cloud, we develop 3D geometry algorithms for tree detections and surface model structures to construct a virtual forest environment allowing users to estimate real forest structures as well as assisting forest managers in decision making.

2. Related Work

Lidar research has been around since the early 1980s, and there have been a number of researchers that have investigated the possible uses. Research in forestry has included discovering individual tree location, forest structure, canopy surface evaluation, as well as virtual forest. These forestry features will allow future automated systems to make estimations on proper forestry management as well as discover tree growth trends.

2.1. Individual Tree Location

Individual tree location is among the most important features of a lidar system. The ability to discover the location and height of individual trees allows researchers to determine the number of trees in an area as well as the forest density. This could allow an area of land to be surveyed numerously over a large period of time, and that would allow for area specific growth rates. Other possibilities for growth rates could include the effects on increased/decreased pollution on forest growth.

The Tiff system by Qi Chen [8] has been able to discover tree location, height, crown area, trunk height, and leaf area. The system that we have developed is able to discover tree location as well, but it does not use a completely original tree detection algorithm because it is similar to the one used by Tiede, Hochleiter, and Blaschke [9].

2.2. Forest Structure

Forest structure is also an important feature of any lidar system. As mentioned before, it can allow for discovery of growth rates in specific areas. The forestry structure evaluation can view the average height, density, and distance between the trees. Information could be used to determine the species of tree that inhabit the area. Researchers have used grid-based [10, 11] or query-based [12] approaches to analyze/visualize forestry structures. The system that we have developed breaks the lidar data into cubic area based on their location on a grid. Once the data is placed into cubes, it is able to be viewed by layers (a horizontal collection of cubes), cylinder (a vertical collection of cubes), or by individual cubes. The layers can be broken into up by a user specified distance or by natural breaks in the data.

2.3. Three Dimensional Representation

3D visualizations for the forest as well as the canopy have become more common with the increased computational power of the PC. The system from [12] was able to represent the forest and canopy, but we believe the researchers Suarez, Ontiveros, Smith, and Snape [13] did better job for the canopy. They were able to break down the limitation of the lidar data as well as produce a highly detailed display of the surface.

The system that we have developed has used a similar method for producing a 3D model for the canopy surface. It also produces the ground surface model as well as virtual forest. The virtual forest consists of tree models in their estimated locations and ground surface.

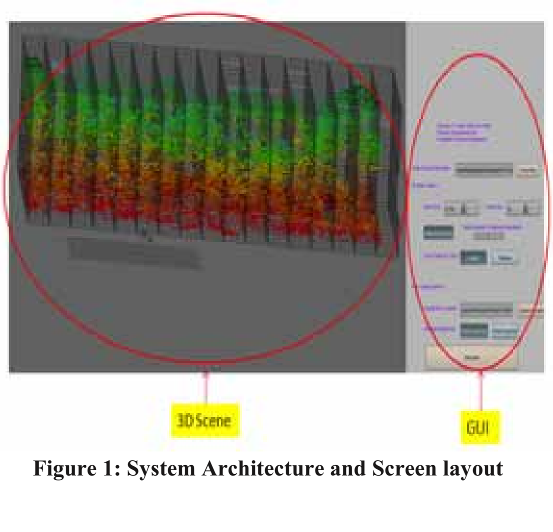


Figure 1: System Architecture and Screen layout

3. Overview

In this section, we explain the system framework and our approach as well as an overall process for visualization and analysis using the forest lidar data.

3.1. System Framework

Figure 1 depicts the architectural structure and main screen layout for the visualization and analysis system. The open source 3D Blender software is employed as a comprehensive visualization and analysis tool.

The system is composed of 3D Scene and a graphical user interface. Within the 3D Scene, lidar data (i.e., lidar point cloud = visualization of lidar data using simple 3D geometric objects), statistics, surface models, and a virtual forest are dynamically visualized through user interactions. To support user interactions, we develop a graphical user interface in Python and OpenGL. This interface has six functions: *Visualization, Creation of Natural Boundary, Surface Model Generation, Tree Detection, Creation of Virtual Forest, and Statistics.*

- *Visualization* - Visual display of the lidar point cloud in Blender 3D Scene. Each point is represented using a tetrahedron, which is color according to its height.
- *Creation of Natural Boundary* - The user is able to select natural boundaries to determine the number of horizontal layers. They are given a choice to from a range of 2 - 7. Natural boundaries are discovered by viewing all of the z values, and looking for "breaks" in the data.
- *Surface Model Generation* - An object is created to represent the canopy of the forest. The data is divided into cylinders which are the height of the point cloud field. The highest point in each cylinder is selected, and each point is linked together with its directly connecting neighbors. Faces of the canopy object are created on the inside of the links.
- *Tree Detection* - Trees are discovered by breaking the data into cylinders, and comparing the highest point in each cylinder to its directly connecting neighbors. If a point is taller then all of its neighbors then it is determined to be a tree.
- *Creation of Virtual Forest* - The virtual forest is created by discovering the location of the trees, and the generation of the ground surface. Tree models are position according to their coordinates, and the models are scaled according to the tree's height.
- *Statistics* - Statics are created by breaking the data into cubic area, and discovering the points which fall within then cube boundaries. The statics such as density and average distance between the points are generated by comparing the points inside of the cube, and comparing the cubes with one another.

3.2. Overall Process

A simplified overall process for visualization and analysis in a forestry domain is shown in Figure 2. The following steps describe the general methodology for forestry visualization and analysis using lidar datasets:

1. Generate lidar point cloud in 3D space
2. Analyze lidar point cloud using Natural Boundary algorithms
3. Generate statistics data in text files based on step 2
4. Detect trees and tree tops and create surface models based on lidar point cloud using tree detection and surface model algorithms
5. Using detected tree tops and data filtering algorithms, generate a virtual forest environment

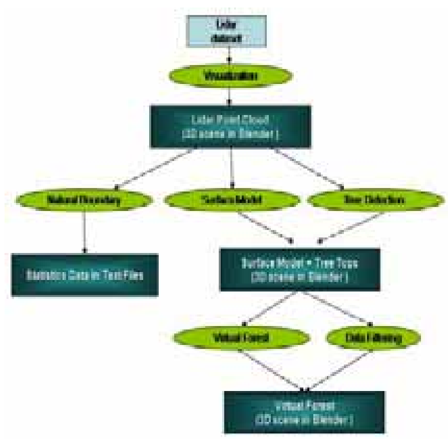


Figure 2: Overall Process

4. Methodology

4.1. Filtering the Point Cloud

The raw data comes in a comma spaced text file, and each data point has the following attributes: id, x position, y position, and z position. These values represent where the point was found on the Earth's surface. This file is read in using python scripts, and each data point is placed in a list along with its attributes.

Once the data has been placed into a list, the data can be easy searched through to discover important values. The max and min x, y, and z values are all discovered, and they are important for future rendering.

4.2. Rendering the Point Cloud

With the points in a list, they can easily be rendered. They are displayed in the Blender software using a python script that accesses the Blender API. The object that was chosen to represent the point is a tetrahedron. A tetrahedron is a three sided pyramid, and it was selected because it is the most basic 3D object. The tetrahedron has four points and four faces. In the center of the bottom face is the location of the point.

The color of the tetrahedron is determined based on the height of the point, and the range is determined with the z max and z min (zmax – zmin). There are seven different color zones which start from lowest to highest. Those color zones are red, orange, yellow, lime green, green, and forest green. The color zones were selected in order of the visual light spectrum with out crossing into the blue color zones. Figure 3 displays the point cloud seen from a top view.

4.3. Creating the Grid

The grid is created by the first determining the width and length of the cube. Those values are equivalent, and they are both user selected from the range of 5 – 20 units. The vertical measurement that makes up the layer size is determined by the user as well. They user has an option to use natural breaks, or they are able to use the same value from cube size. Figure 4 displays the grid on a small data set from Figure 3.

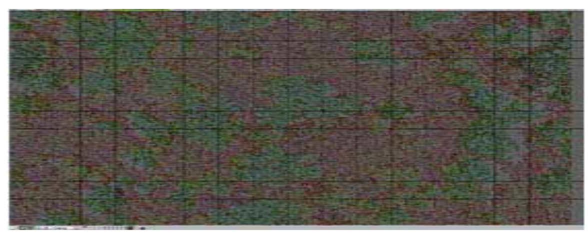


Figure 3: Top view a data set. Each color point displays a different point in the point cloud. The greener color points are taller than the red and orange color points.

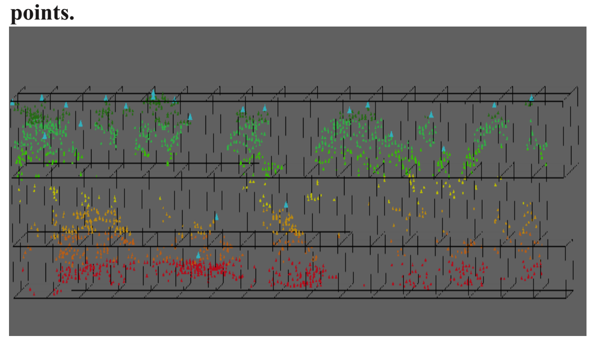


Figure 4: Display a different data set from Figure 3, but from a side view. This view allows the user to view the cubes and the points in them.

4.4. Creating and Filling the Cubes

Once the values are determined and the grid is created, the intersection of the grid lines creates the cube boundaries. The cubes are numbered to identify them, and then it is determined which points fall in the boundary lines of the cube. A list is created from the points in the cube, and statistics are determined. The important statistics are density, average distance between points, and other average values.

4.5. Layers and Natural Breaks

The layers are created by all of the cubes that fall in the same vertical plane. These layers are later displayed on their own tab in Blender for future evaluation. The vertical measurement which determines the number of layers is either the user selected value or from a natural break. If the user selects natural break, they are able to select from a range of 2 – 7. Natural breaks in the determined by viewing all of the data point's z values, and determining where a discontinuous break in the graph line appears.

Natural breaks are important because they can help determine the forest structure. A person with forestry knowledge could use the information to determine the age or species of a forest. Figure 5 displays a graph with 2 natural-points displays as an example.

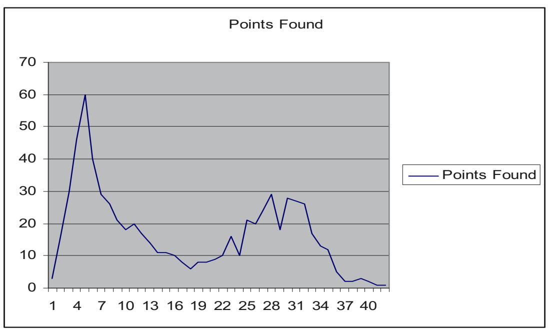


Figure 5: displays a graph with natural breaks at height 17 and 29. There are other natural breaks in the data, but if the user were to select two then these would be found.

4.6. Tree Detection

Tree detection begins with reorganizing the data into cylinders instead of cubes. The cylinders are the height of the total z range, and their width and length are determined by the user from a range of 1 – 10 units. The smaller the value the higher the detail, but the downside is greater amount of processing timer. Once the cylinders boundaries are determined, the points which fall inside of those cylinders are placed into a list. All of the point's heights are compared with one another, and the highest point is selected for each cylinder. These points are then compared to the highest values of their directly connected neighbors. If a point is taller then all of its neighbors then it is assumed to be a tree. These points are added to a list of tree points, and are used for future rendering.

There is an obvious problem of possibly selecting two points from the same tree. This can be partially alleviated by selecting a cylinder width and length that provides enough detail to find the tree, but not to over count them. From our testing, I found unit size of 3 to work well, but this is not extensive testing. In the future more testing on accuracy will have to be conducted along side with field surveys.

Once the discovered tree list is created, the tree points are rendered using a light blue tetrahedron that is larger then the rest of the points. This helps the user clearly identify the tree location in Blender. Figure 6 displays the blue tetrahedrons.

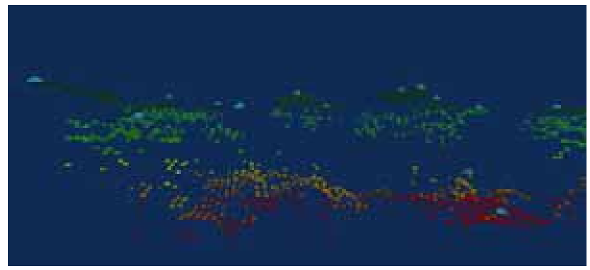


Figure 6: Displays the tree locations which can be identified by the blue tetrahedron.

4.7. Canopy Surface Rendering

The canopy surface is rendered using the same cylinders from the tree discovery algorithm. A python script takes the highest points from each cylinder and creates an edge (line) between it and all of its connecting neighbors. The faces are created inside anywhere three edges are connected. Once all of the faces are created, they are unioned to create a single object. This object represents the canopy surface. Figure 7 displays the surface object of the same data set from Figure 3.

4.8. Ground Surface Rendering

The ground surface is rendered much in the same way as the canopy. Instead of selected the highest point in each cylinder, the lowest point is selected. This is not enough to create the ground surface object because in areas where a tree appears they often shadow out the lower points. To overcome this problem, we must estimate where the ground will be. This is done by taking the lowest value in the cylinder and comparing against the average z value of all of the points. If the low value is greater then the average height, the point is defiantly not a low value. The non true low value is replaced by the z minimum which was found earlier. Simply replacing the value is not the best estimation, but it is the simplest way to estimate what the surface will look like. Figure 8 displays the ground surface under the tree models.

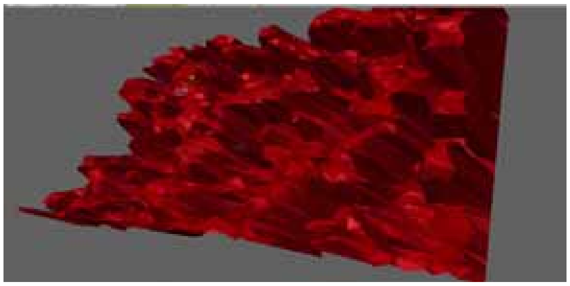


Figure 7: Displays the canopy object from a side angle. It allows the user to see the high and low spots from the cloud field.

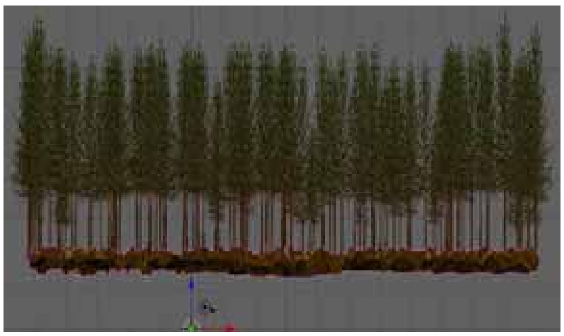


Figure 8: Displays the virtual forest from a side view. It includes the ground surface and tree models.

4.9. Tree Model Generation

A tree model generator by Sergey Prokhorchuk [14] is used in Blender to create tree models. The tree model selected uses the minimum values of variables to create a tree model that is not complex. Other models were created but they were large in memory size which made them difficult to replicate many times. Figure 9 also displays the virtual forest, but with a single tree selected.

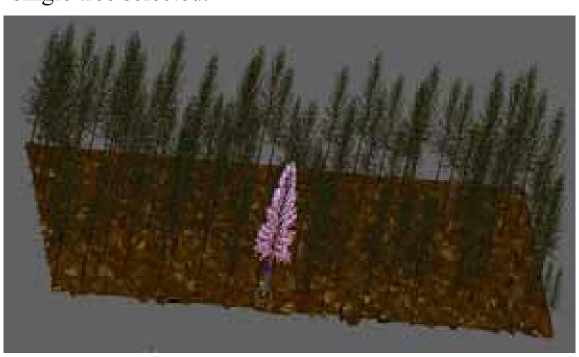


Figure 9: From the same data set as Figure 8, but viewed at a different angle. The pinkish colored tree is a selected tree model.

4.10. Virtual Forest

The virtual representation of the forest is generated by placing ground surface object and tree objects together. Once the ground object is placed, trees are placed according to their x and y positions on the grid. The tree's size is scaled according to the height to create trees that are realistic looking. The virtual forest is an estimation of what we believe the forest to look like with a single species of tree. Figures 8 and 9 display the virtual forest.

5. Conclusion and Future Work

We have presented a general process for visualization and analysis of lidar data in a forestry domain and have discussed and explained our implementation approaches under the Blender environment. Our goal is to study efficient methods for creating visualization and analysis system environments as well as a virtual forest environment from plain lidar data sets by focusing on effective ways to visualize, manage, and analyze individual tree attributes and structures.

Recently, we have been interested in including a modelling and simulation environment for future work. We refer to this environment as the *integrative multimodeling environment* [15], which allows the user to switch between information, geometry, and dynamic models while never leaving the immersive scene. This will require more advanced visualization and involve a complicated multimodel environment.

Acknowledgements

We would like to thank ORSP at Stephen F. Austin State University for support of this research. We thank Dr. I-Kuai Hung of Forestry Department at Stephen F. Austin University for providing lidar datasets and giving fruitful suggestions.

References

[1] Harding, D.J., M.A. Lefsky, G.G. Parker and J.B. Blair, "Laser altimeter canopy height profiles methods and validation for closed-canopy, broadleaf forests", *Remote Sensing of Environment,* Vol. 76, pp 283-297, 2001.

[2] Persson, Å., J. Holmgren and U. Söderman, "Detecting and measuring individual trees using an airborne laser scanner", *Photogrammetric Engineering and Remote Sensing* Vol. 68, pp 925-932, 2002.

[3] Schardt, M., M. Ziegler, A. Wimmer, R. Wack ank J. Hyyppä, "Assessment of forest parameters by means of laser scanning", *International Archive of Photogrammetry and Remote Sensing* XXXIV(part 3A), 2002, pp 302-309.

[4] Hyyppä, J. and M. Inkinen, "Detecting and estimating attributes for single trees using laser scanner",

Photogrammetric Journal of Finland vol. 18, pp 43-53, 2002.
[5] Arizona State University Active Tectonics Research Group, http://lidar.asu.edu/LViz.html
[6] Remote Sensing Application Center, http://www.fs.fed.us/eng/rsac/fusion/
[7] Martin Isenburg and Jonathan Shewchuk, http://www.cs.unc.edu/~isenburg/googleearth/
[8] Chen, Q., "Airborne lidar data processing and information extraction", *Photogrammetric Engineering and Remote Sensing,* vol. 73, no. 2, pp 109-112, 2007.
[9] Tiede, D, Hochleitner, G, and Blaschke, T., "A Full GIS-Based Workflow for Tree Identification and Tree Crown Delineation Using Laser Scanning", in *Proc. CMRT05,* 2005, pp 9-14.
[10] Means, J.E., S.A. Acker, B.J. Fitt, M. Renslow, L. Emerson and C.J. Hendrix, "Predicting forest stand characteristics with airborne scanning lidar", *Photogrammetric Engineering and Remote Sensing* vol. 66, pp 1361-1371, 2000.
[11] Næsset, E., "Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data", *Remote Sensing of Environment,* vol. 80, no. 1, pp 88-99, 2002
[12] Kao, D, Kramer, M, Luo, A, Dungan, J, and Pang, Alex, "Visualizing Distributions from Multi-Return Lidar Data to Understand Forest Structure", *the Cartographic Journal*, vol. 42, no. 1, pp 35-47, June 2005.
[13] Suarez, J, Ontiveros, C, Smith, S, and Snape, S, "Use of Airborne LiDAR and Aerial Photography In the Estimation of Individual Tree Heights In Foresty", *Computers and Geosciences*, vol. *31*, pp 253-262.
[14] Prokhorchuk, S., http://www.geocities.com/bgen3/
[15] Park, M. and Fishwick, P., "Integrating Dynamic and Geometry Model Components through Ontology-based Inference," *Simulation*, vol. 81, no. 12, pp. 795-813, December 2005.

Human-Computer Interaction in Emergency Management

Davide De Chiara[1], Vincenzo Del Fatto[2], Luca Paolino[1], Monica Sebillo[1], G. Tortora[1], G. Vitiello[1]
[1] *Dipartimento di Matematica e Informatica, Università di Salerno*
84084, Fisciano (SA), Italy
{vdelfatt, lpaolino, msebillo, tortora, gvitiello}@unisa.it
[2] *LIRIS – INSA de Lyon*
69621 – Villeurbane Cedex - France
+33 472438172
vdelfatt@unisa.it

Abstract

In case of emergency, slow or obstructed communication lines may make exchange of information difficult or even impossible. Such a problem is particularly important when we manage maps, which are usually complex to deal with. They require a long time for transmission and can generally be viewed only after complete acquisition. In order to mitigate this issue, a promising approach is to transmit a reduced map first and, in subsequent steps, its progressive refinements until reaching an acceptable detail. However, choosing the first and the subsequent reduction levels is a complex problem and users are often let to decide on a case-by-case basis.

In this paper, we present an empirical experiment we performed to determine which reduction levels we should apply in case of emergency by analyzing the perception of map changes with respect to the original map. On the basis of such results, we also derived a set of rules that allow us to determine appropriate reduction levels on the basis of the number of vertices and the number of polygons.

Keywords. Emergency Management, Geodata Transmission, empirical Experiment.

1. Introduction

Fires, inundations and outbreaks are events which are potentially very destructive both in terms of human lives and property damages. In order to reduce the impact of these events, decision makers must be able to manage emergencies in a very fast way and to quickly organize the intervention of people who work on field. In this sense, a very important support has been given in the last years by geographical information systems (GIS) which, thanks to their ability to analyze spatial data by means of specific models, allow officials to produce the right data at the right time, displayed in a logical fashion, so as to respond and take appropriate actions. Using a GIS, decision makers can pinpoint hazards and begin to evaluate the consequences of potential emergencies or disasters. When hazards are viewed with other map data (streets, pipelines, buildings, residential areas, power lines, storage facilities, etc.), emergency management officials can begin to formulate disaster mitigation, preparedness, response, and possible recovery needs [3]. However, the correct managing of such events not only depends on the analysis of the spatial data and on the production of a correct solution, but it is also the combination of many other factors. Among them, special focus deserves the speed by which decisions are broadcast towards the regions struck by an event. Transmitting a very large and detailed map may cause fatal delays, if, e.g., slow or obstructed communication lines cause transmission failures and require to repeat the operation. Therefore, producing a complete solution might not be appropriate.

One method for solving this issue is to transmit the map in a progressive way, namely a reduced/light map is first sent, then a successive packet of points is sent to achieve a new detailed version of the map, and so on, passing through different versions of the map until the required detail is reached [7]. The first and the subsequent details may be calculated by means of the Douglas-Peucker (RDP) reduction algorithm [2, 5]. However, choosing which reduction levels should be used to calculate each intermediate map is a complex problem and users are often let to decide on a case-by-case basis, and very little human guidance or automated help is generally available for that [7]. As an instance, transmitting an unrecognizable map first could be useless as well as selecting a level which does not substantially improve the previous one. In

order to avoid such problems, we performed an experiment aimed to identify which reduction levels better highlight substantial changes of map perception. To this aim, we identified five levels of perception, namely 1) No simplification detected, 2) Some minor simplifications detected, 3) Some evident simplifications detected, 4) Some substantial simplifications detected and 5) The map can no longer be recognized. Then we evaluated at which simplification rates, on the average, participants reached each change perception level. Following the experiment, we derived a set of rules that allow us to determine appropriate reduction levels, on the basis of the number of vertices and the number of polygons.
In the remainder of this paper, we describe the experiment setting and present, in order, independent and dependent variables, participants, tasks and results. A discussion on the achieved results concludes the paper.

2. Independent Variables

In our experiment, the independent variables are the number of vertices (NV) and the number of polygons (NP), which represent the inner and the outer complexity of a map, respectively.

Number of Vertices. The number of vertices is the first independent variable we took into account in our experiment. For the sake of clarity, we call vertex a point which lies on either the start or the end of a map segment.

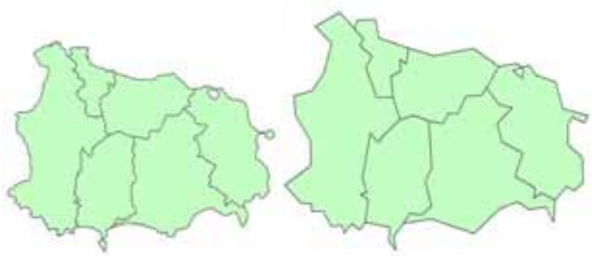


Figure 1. An example map represented through two different sets of vertices.

NV	Range
1	between 7000 and 8000
2	between 46000 and 47000
3	between 100000 and 120000

Figure 2. Categorizing the number of vertices

It is intuitive that the higher the number of vertices used to represent a geographic object, the more is it detailed and hence the better is it perceived by humans. Usually, this value may be represented through a continuous variable ranging from 0 to infinitive. However, for our aims we decided to categorize it by setting three levels indicating three different size orders (see Fig. 1 and Fig. 2).

Number of Polygons. The number of polygons is the second variable we managed. It seems plausible that this factor affects the user's perception because the higher the number of objects visualized, the higher the number of details perceived in the map. As for the number of vertices, we decided to categorize the NP variable with the following three different levels (see Fig. 3 and Fig. 4).

Figure 3. An example map showing the same county drawn with a different number of polygons.

NP	***Range***
1	less or equal to 5 polygons
2	between 6 and 10 polygons
3	between 11 and 30 polygons

Figure 4. Categorizing the number of polygons

3. Dependent variable

Usually during a map simplification process, user's perception of changes may vary from very little perception to a point when the simplification becomes evident. Generally, we may identify the following five levels of perception (LP) of changes:

1. *(LP1)No simplification detected*, in this case the participant does not recognize any difference between the original map and the simplified map.
2. *(LP2)Some minor simplifications detected which do not affect the appearance of the map.* In this case, some differences are perceived but they are not relevant. As an example, simple linearization of raw lines or smaller contour lines due to less overlapped points.
3. *(LP3)Some evident simplifications detected* which do not alter the meaning of the map. For instance, when rough lines get fairly straight.
4. *(LP4)Some substantial simplifications detected* which affect the information the map conveys. The map is

still recognizable but some changes may hide relevant details. As examples, some peninsulas may disappear; boundaries may deeply alter their position, and so on.

5. *(LP5)The map can no longer be recognized*.

The dependent variables are the Simplification Rates (SR1, SR2, SR3 and SR4), which are defined as the RDP simplification rates that we observe when participant starts to see transitions between two successive levels of perception of changes (1-to-2, 2-to-3, 3-to-4, 4-to-5) with respect to the original map.

4. Participants

The participants we used in this research were students following the degree programme of Computer Science at the faculty of Science, University of Salerno (Italy). The complete list of individuals involved 72 participants divided into 9 groups, 6 participants per group. In order to make such groups as independent as possible we decided to randomly assign participants to each group.

The rationale behind the choice of 9 groups is that 9 is exactly the number of different combinations of values for NP and NV. Thus, the 6 participants composing each group were associated with a real map having a combination of the two factors.

5. Apparatus

As for the apparatus, we exploited three software applications, namely, SYSTAT™ [7], MapShaper [1] and ESRI ArcView™ [3]. They were run on a Windows XP© Professional platform mounted on a computer based on a Pentium Centrino™ processor with 1G RAM, a 72 G HD at 7200 rpm and a 15. 4'' multi resolution display.

For our investigation, we exploited SYSTAT™ ver.12, an advanced software application for statistical analysis.

We used MapShaper for producing the reduced maps. It is a free online editor for Polygon and Polyline Shapefiles. It has a Flash interface that runs in an ordinary web browser. Mapshaper supports the Douglas-Peucker (RDP) line simplification algorithm.

Finally, we used the ESRI ArcView 3.2 to produce the source maps we used in the experiment and to express them also in terms of the combination of the factor levels for, respectively, the number of polygons (NP) and the number of vertices (NV).

6. Tasks

Each participant performed the following steps. The original map and a simplified version of it were successively shown for ten second intervals. Each group worked on a different map. Then, according to the scale concerning the perception of changes, we asked each participant for a value between 1 and 5.

This step was repeated, increasing the simplification rate, until either participants reached the maximum simplification rate or the maximum of the scale.

For each participant we reported the value of SR, namely the simplification rate registered when s/he detected each level LP.

			Number of vertices								
			7K-8K			46K-47K			100K-120K		
			Aver.	D.S.	IC	Aver.	D.S.	IC.	Aver.	D.S.	IC.
Number of polygons	<5	1-to-2	76.83	13.11	10.48	84.83	11.57	9.25	79.67	12.13	9.70
		2-to-3	87.17	6.94	5.55	94.33	4.97	3.97	94.33	4.97	3.97
		3-to-4	94.5	6.41	5.13	100	n.d.	n.d.	100	n.d.	n.d.
		4-to-5	Ov100	n.d.	n.d.	Ov100	n.d.	n.d.	Ov100	n.d.	n.d.
	6-10	1-to-2	71.5	8.78	7.03	78.67	14.49	11.59	87.17	7.91	6.33
		2-to-3	89.33	8.69	6.95	92.50	7.12	5.70	94.17	2.86	2.29
		3-to-4	97.83	5.30	4.25	96.67	5.50	4.40	100	n.d.	n.d.
		4-to-5	Ov100	n.d.	n.d.	Ov100	n.d.	n.d.	Ov100	n.d.	n.d.
	11-30	1-to-2	78.83	14.92	11.93	80.83	7.91	6.33	83.83	11.60	9.28
		2-to-3	94.67	8.48	6.78	93.17	4.12	3.30	97.67	3.61	2.89
		3-to-4	97.67	3.61	2.89	98.83	2.86	2.29	100	n.d.	n.d.
		4-to-5	Ov100	n.d.	n.d.	Ov100	n.d.	n.d.	Ov100	n.d.	n.d.

Table 1. Perception of changes summarize

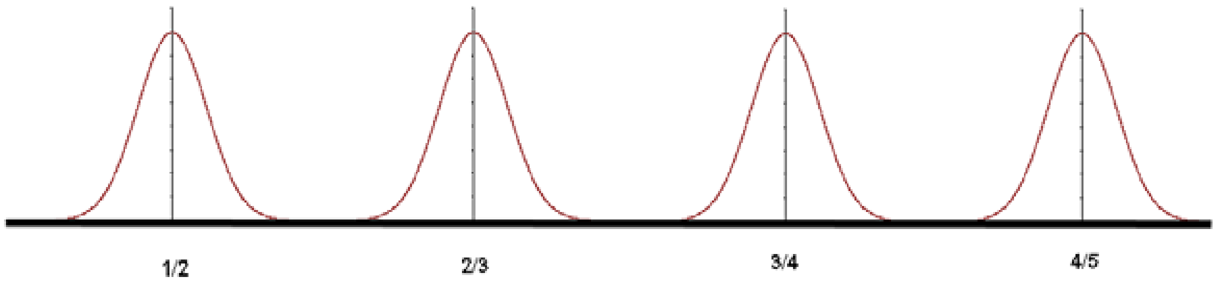


Figure 5. Gaussian curves representing the transitions between each successive LP

7. The Subjects

Table 2 shows the perception values of a participant gained through a record card. 7k-8k and 6-10 represent the characteristics of the map which was presented to the participant, values 6, 12, …, 100 are the simplification rates and 1, 2, 3, 4 and 5 are the levels the participant indicated for each simplification.

		NV							
		7k-8k							
NP	**6-10**	6	12	18	25	32	38	44	50
		1	1	1	1	1	1	1	2
		56	62	68	75	81	87	93	99
		2	2	2	2	2	3	4	5

Table 2. A record card storing a user's perception Results

In order to answer the rationale of this paper we collected the *SR* values that participants perceived when we showed them the simplified versions of the maps. Then, we identified which *SRs* correspond to the transition between any successive category levels $SR(LP_i/LP_{i+1})$. As an example, the SR levels detected for one of the participants' card of Table 2 were 44 for 1-to-2, 81 for 2-to-3, 87 for 3-to-4, and 93 for 4-to-5.

Table 1 shows basic statistics (Average, Standard Deviation and 95% Interval Confidence) summarizing results of the experiment for each category of *NV* and *NP*. Within the table, *Ov100* means that the transition has not been perceived before the last simplification, while *n.d.* means that such a value cannot be calculated. As we expected the $SR(LP_i/LP_{i+1})$ averages may be approximated by some Gaussian curves as shown in Figure 5. For each average, we also measured the relative 95% interval of confidence. As a matter of fact, it indicates the interval where the $SR(LP_i/LP_{i+1})$ average should fall at a probability of 95% [6].Actually, claiming that the average of $SR(LP_i/LP_{i+1})$ is A_i with interval of confidence at 95% (IC_i) means that:

$Pr(A_i - IC_i < SR(LP_i/LP_{i+1}) < A_i + IC_i) = 0.95$,

or also that the probability that a map transmitted with a reduction level of A_i - IC_i is associated with LPj, with $j < i$, is greater than 0.95. In a formal way, we could claim that:

$Pr(LP_{Ai\text{-}ICi} < i) = 0.95$

where $LP_{Ai\text{-}ICi}$ is the level of perception change that the participant perceive when see a map with *SR* equals to Ai - ICi.

We can then use the $SR(LP_i/LP_{i+1})$, so determined, to define a smart sequence of map reductions that can be applied on a progressive map transmission as described in the Introduction.

Given a pair *NP*, *NV* indicating the number of points and the number of polygons describing a map, we can define the sequence of map reduction levels, *MR4(NP, NV), MR3(NP, NV), MR2(NP, NV), MR1(NP, NV)* where:

- MR4(NP, NV) is the reduction level where the meaning of the map will be recognized by at least 95% of the participants, namely it is the highest reduction for which $Pr(LP_{A5\text{-}IC5} < 5) = 0.95$
- MR3(NP, NV) is the level where at least the 95% of the participants do not perceive substantial simplifications on the map, namely it is the highest reduction for which $Pr(LP_{A4\text{-}IC4} < 4) = 0.95$
- MR2(NP, NV) is the level where the map has the property that at least the 95% of the participants do not perceive evident simplifications on the map, namely it is the highest reduction for which $Pr(LP_{A3\text{-}IC3} < 3) = 0.95$
- MR1(NP, NV) is the level where the transmitted map has the property that at least the 95% of participants do not perceive minor simplifications on the map, namely it is the highest reduction for which $Pr(LP_{A2\text{-}IC2} < 2) = 0.95$

As an example, let us suppose we have to quickly transmit a map representing areas struck by an earthquake. It contains spatial information regarding the evacuation plan and evacuation ways. Properties characterizing the map are 47000 vertices, and approximately 30 polygons. We want to identify which reduction levels better fit these requirements. Table 3 shows the experiment details in correspondence of such properties.

Level Transition	Aver.	D.S.	Conf.
1-to-2	80.83	7.91	6.33
2-to-3	93.17	4.12	3.30
3-to-4	98.83	2.86	2.29
4-to-5	Ov100	n.d.	n.d.

Table 3. A detail of the data summarize table showing simplification levels with regards 47K points and 30 vertices.

This kind of map is always recognizable because no participants have associated reduced map with Level 5. In this case, the point 4-to-5 does not exist, then it is not necessary to define the MR4 value. As for MR3, corresponding to the point 3-to-4, the average is 98.83 with a confidence value of 2.29 at 95%. It means that we can reduce the map to 96.54 % (98.83 - 2.29) and obtain a new version with no substantial simplification problems but with some evident changes. Then, we can send a packet of vertices which can be added to the vertices previously transmitted corresponding to a map reduced to the MR3 level. In correspondence with 2-to-3, the average is 93.17 with a confidence value of 4.12. It implies that a map reduced at 89.05% (93.17-4.12) has some minor simplification problems which are not evident. Thus, as previously done, we can send a packet which integrates the previous one until MR2 is reached. Finally, we can transmit a new packet whose points allow to integrate the last reduced map to a new version having the property that participants should perceive no simplifications at 95%. As a matter of fact, on the average, the 1-to-2 value is approximately around 80.83% with a confidence value of 6.33. It means that we can generate a map reduced at 80.83 – 6.33 = 74,50% which cannot be distinguished from the original.
In conclusion, in the given example, MR's are as follows: MR4= undefined, MR3=98.83, MR2=89.05 and MR1 =74.5.

8. CONCLUSION

This paper has focused the attention on a particular problem of the emergency management, namely the necessity to rapidly send maps in case of events. We highlight that progressive transmission may be useful when events happen because it allows to deliver manageable and immediately usable maps also in case of slow connections. We put also in evidence in this paper that intermediate reduced maps should be accurately chosen to avoid transmission of unrecognizable maps or intermediate maps which give no additional information with respect to the previous ones. To correctly select these levels we performed an empirical study which identifies them by analyzing the perception of changes of reduced maps with respect to the originals. Results show that a reduction of transmitted data may be applied without losing significant details so minimizing transmission data and speeding up the emergency management.

9. REFERENCES

1. Bloch, M., Harrower, M.. MapShaper.org: A Map Generalization Web Service. *In: Proc. of Autocarto 2006*, June 26-28, 2006 Vancouver, Canada.
2. Douglas, D., and Peucker, T. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature, *The Canadian Cartographer 10(2)* (1973), 112-122.
3. ESRI ArcView 3.2. www.esri.com
4. ESRI, GIS for emergency management, http://www.esri.com/library/whitepapers/pdfs/emermgmt.pdf
5. Hershberger, J. and Snoeyink, J. Speeding up the Douglas-Peucker line simplification algorithm. *In: Proc. 5th Intl. Symp. on Spatial Data Handling, Vol 1* (1992), 134–143.
6. Sheldon, M.R. *Introduction to probability and Statistics for Engineers and Scientists, 2nd Ed*, Academic Press, CA, USA, 2003.
7. SYSTAT ver.12. www.systat.com
8. Zhao, H., Li, X., Jiang, L., A modified Douglas-Peucker simplification algorithm. In: *Proc. of Geoscience and Remote Sensing Symposium (IGARSS '01),* (2001), Sydney, NSW, Australia
9. M. Zhou, M. Bertolotto, Implementation of Progressive Vector Transmission Using a New Data Structure and Modified RDP Algorithm. In: Proceedings GISRUK 2004, Norwich, UK, (2004), 69-72.

Augmented Tour of Archaeological Site by means of Adaptive Virtual Guide

Andrea F. Abate, Giovanni Acampora, Stefano Ricciardi

Università degli Studi di Salerno, Dipartimento di Matematica e Informatica
84084 Fisciano (Salerno), Italy
{abate, gacampora, sricciardi}@unisa.it

ABSTRACT - *This study presents a novel framework exploiting augmented reality to visualize a synthetic 3D virtual guide inside an exhibit or a public gallery to assist visitors wearing a Head Mounted Display during their visit and providing them with both visual and informative enhancements not available in a standard tour. The human-avatar interaction is approached through a model based on timed automata to address the conversational issues and to simplify framework's design and checking. The paper includes an application of the proposed methodology to the "avatar assisted tour" of a roman villa, located within the archaeological site of Pompei, as a case study.*

1. INTRODUCTION

The main focus of this study is the proposal of a novel framework to assist the visitors of cultural heritage sites by means of a virtual guide able to provide various level of information about the exhibits exposed. The proposed architecture exploits augmented reality and stereoscopic vision to enable the user wearing an Head Mounted Display (HMD) to see a realistic 3D avatar visualized into the actual exhibit space. The virtual assistant is able to harmonize its goal (to guide and assist the user throughout the tour) to the user's needs (for instance modifying the path according to user's preferences) thanks to an interaction model based on timed automata, addressing the temporal aspects of human-avatar communication, and enabling the system designer to verify the formal correctness of the model itself. The idea of virtual assistants is not new, as shown in [1] where a classification of these entities in three different categories (avatars, assistants and actors) is provided and applied to the definition of an asynchronous hierarchical architecture for virtual assistant management and interaction, tailored for virtual museum applications. In the same applicative field Stock et al.[2] presented an avatar based model to enhance cultural heritage fruition, proposing a suite of interactive and user-adaptive technologies (animated agents), while Badler [3] proposed an approach to virtual humans based on a parallel-finite-state machine controller implementing a "sense-control-act" architecture providing reactive behaviors to drive virtual humans through complex tasks. On a higher level of abstraction Peters [4] presented a psychological evolutionary model for virtual conversation based on intelligent agents exploited to achieve a perception of user's behaviour based on the concept of "state-of-mind", while others authors are exploring the role of semantic memory as a premise to avatars featuring more human-like artificial minds. The topic of avatars visualized in real time through immersive displays and mixed reality techniques has also been investigated by authors like Lee et al. [5] and Ogi et al. [6], exploiting live video as a source for silhouette mapped avatars (video-avatars) to achieve tele-presence and interaction [7] in collaborative virtual environments. Even on the man-machine interaction and interfacing topic there are several contributions related to the proposed work. For instance, the timed automata theory has been applied to man-machine interaction in Loer and Harrison [8] which describes how work representations might be generated using a process based on model checking. Sequences or traces are explored that either satisfy properties or demonstrate the failure of a property. In Norstrom et al. [9] timed automata based event driven system are designed by means of meta-models. The proposed approach to human-avatar interaction differs from all the aforementioned studies as it combines a completely synthetic virtual assistant (a real 3D character opposed to a live captured "video" replica of a real user, acting "autonomously" and not according to a puppeteer) to an augmented reality environment (allowing user to see the virtual assistant together with virtual objects placed into the real environment) and to an interaction model based on timed automata, featuring avatar-to-user synchronization and dynamic adaptation to user's preference.

2. FRAMEWORK ARCHITECTURE

As described in the introduction, the proposed framework is intended to assist the user during a tour to a set of exhibits (stations) located within a cultural heritage site by means of a virtual guide (see Fig. 1). In a typical applicative scenario, the system firstly recognizes the user typology (children/adults, experts/novice, etc.) via a previously assigned RFID tag, then it selects the appropriate virtual guide appearance and the related behavioural model, affecting the way in which the tour is accomplished (i.e. how many stations, where they are located and in what order). The interaction paradigm is based on a bidirectional asymmetric communication between the visitor and the avatar, as the avatar can provide info about the tour through speech synthesis while the visitor's position and head orientation are exploited to inform the avatar where, along the visiting path, the visitor is located and where he is looking at. The avatar and any virtual object related to the guided tour are visualized onto the actual environment by means of a visual engine which is responsible for the 3D registration of the virtual models according to the current user's point of view and which also performs the stereo rendering of the registered models in real time. The avatar's behaviour is the crucial component for the whole system, as it directly affects the way in which the virtual character interacts with the visitor. The proposed approach to synthetic behaviour design is based on a combination of finite not deterministic automata and timed automata [10], as in a given time each interacting object can be in a specific state and can change such state due to either specific temporal events or according to changes in the state of other objects. One of the main advantages offered by automata based systems is their verification capability through, for instance, the reachability analysis.

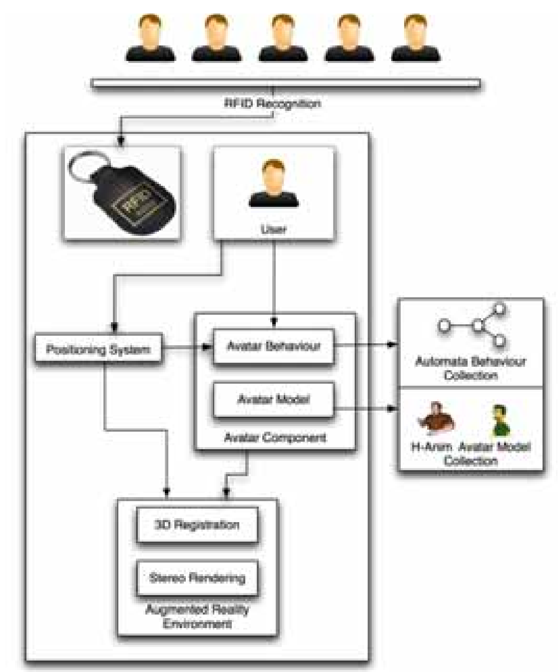


Figure 1. A schematic view of the proposed Framework

This feature, indeed, allows to check since from the design stage if a given state or a defined set of states will be reached during system simulation or at runtime. Three main components are required to simulate and validate the aforementioned interaction paradigm: the Visitor component, the Avatar Component and the Positioning System. They represent the events and the information sources involved in the interaction and relevant to the design of the avatar's behaviour. The communication between each component is achieved by means of shared variables and via a parallel composition operator able to allow interleaving of actions as well as hand-shake synchronization.

2.1 Framework Components

While in the design stage, the *Visitor Component* models the visitor's behaviour. Indeed, user status can be modeled through a simple non-deterministic finite automaton because a generic visitor can be considered as a discrete entity assuming three states: *waiting*, *visiting* and *moving*. When the user is in *waiting* status, the virtual character checks user's position exploiting the data captured by the positioning system. Once the user's position is acknowledged as a valid location, the assistant starts the appropriate animation and the user automaton transits in *visiting* mode until the presentation ends or the visitor changes its position. In both cases the visitor automaton transits in *moving* state. Formally, let A_{user} be the non-*deterministic* finite automaton describing the user:

$$A_{user} = \langle Q_{user}, \Sigma_{user}, \delta_{user}, q_0, F_{user} \rangle$$

where Q_{user} is the set of user states, Σ_{user} is a symbols alphabet representing the actions involving the user, δ_{user} is a transition function mapping the user's status changes, q_0 is the user's initial interaction state and F_{user} represents the ending states for human-avatar interaction. Formally:

- $Q_{user} = \{start, moving, visiting\} \cup dummyStates$;
- $\Sigma_{user} = AssistantSignals \cup PositioningSystemSignals$;
- $q_0 = start$;
- $F_{user} = \{start\}$, $\delta_{user}: Q_{user} \times \Sigma_{user} \rightarrow Q_{user}$

Q_{user} contains the *waiting*, *moving* and *visiting* states (together with a collection of dummy states used to perform not significant transition) as defined at the beginning of section; Σ_{user} contains the pair of 'symbols' set; the first collection is used to change the visitor status according to event triggered by the virtual assistant, whereas, at same way, the second collection is exploited to manage the event generated by the Positioning System automata. $q_0 = waiting$ means that the user starts its interaction only when the assistant is in ready mode, i.e., an interaction process cannot start during avatar animation. $F_{user} = \{waiting\}$ means that each human-avatar interaction session terminates in the *waiting* state in order to allow the user to start the next tour in a transparent way. The meaning of δ_{user} transition function is depending upon application. Section 4 will show some samples of interaction transitions defined by δ_{user}.

The *Avatar Component* is the collection of system components defining the rule set exploited by the virtual character in order to enable an advanced conversation. Indeed, this module is capable of analyzing inputs coming from both user and the positioning system in order to change the avatar's status and, consequently, its representation. In other words, the Avatar Component realizes an event-driven behaviour where, typically, events are triggered according to user's position or to temporal events. The design of the Avatar Component is based on timed automata. Differently from standard automata, timed automata use timers to perform transitions among states. More precisely, a timed automaton is a standard finite-state automaton extended with a finite collection of real-valued clocks. The transitions of a timed automaton are labeled with a guard (a condition on clocks), an action, and a clock reset (a subset of clocks to be reset). Intuitively, a timed automaton starts execution with all clocks set to zero. Clocks increase uniformly as time passes while the automaton is within a node. A transition can be taken if the clocks fulfill the guard. By taking the transition, all clocks in the clock reset will be set to zero, while the remaining ones keep their values. As a result, every transition occurs instantaneously. Formally, the state of an automaton is represented by a pair including a control node and a *clock assignment*, i.e. the current setting of the clocks. Transitions in the semantic interpretation are either labeled with an action (if it is an instantaneous switch from the current node to another) or a positive real number i.e. a time delay (if the automaton stays within a node letting time pass). In other terms, the timed automata are finite-state machines constrained by timing requirements so that they accept (or generate) *timed words*. Though this component is strongly application oriented, a set of design guidelines are provided in the following lines. Recalling the aforementioned timed automata definition, a set of actions Act_{Avatar}, a set of clocks C_{avatar}, a set of locations N_{avatar} and a collection of transition edges E_{avatar} have to be introduced to build the required automaton. N_{avatar} is a collection of different states in which avatar can be (analogous to Q_{user} set defined in previous subsection). Our approach defines the N_{avatar} set as the union of different set of states:

$$N_{avatar} = \bigcup_{i=1}^{k} animation_i \cup \bigcup_{i=1}^{l} communication_i \cup \{start\}$$

While in the $start$ dummy state, the avatar waits for a command from other subsystems (for instance it waits for user position's changes). The set $\bigcup_{i=1}^{k} animation_i$ is the collection of states affecting the avatar's appearance. Analogously, $\bigcup_{i=1}^{l} communication_i$, is the collection of states associated to virtual assistant in order to communicate with the user. To address the command reception issue, the virtual assistant has to synchronize its behaviour with the user's activity; this is possible by means of a pair of synchronization signal collections aggregated into the Act_{avatar} set:

$$Act_{avatar} = VisitorSignals \cup PositioningSystemSignals$$

where *VisitorSignals* is a collection of synchronization signals interacting with visitor automaton, whereas *PositioningSystemSignals* represent the signals allowing the interaction with positioning systems. The visitor/avatar interaction is also addressed by means of a clock exploited only when the avatar behaviour automaton is in a state belonging to the set $\bigcup_{i=1}^{l} communication_i$.This clock, named *waitingClock*, represents a timer used by virtual assistant to wait for the visitor to arrive in a valid station. Once *waitingClock* expires, the assistant automaton will transit in an animation state in order to communicate the right site position to visitor. Summarizing the previous description, a high level definition of E_{avatar} is given. Avatar's actions depends either upon q different kind of user's command or upon positioning data and, at same time, on the expiration of w different kinds of timer. Then:

$$E_{avatar} = \bigcup_{i=1}^{q} command_i \cup \bigcup_{i=1}^{w} expiration_i$$

where each $command_i$ is an edge connecting two vertex s_1 and s_2 with $s_1 \in (\bigcup_{i=1}^{l} communication_i)$ and $s_2 \in (\bigcup_{i=1}^{k} animation_i)$ representing active interaction animations, whereas, $expiration_i$ is an edge connecting two vertex s_1 and s_3 with $s_1 \in (\bigcup_{i=1}^{l} communication_i)$ and $s_3 \in (\bigcup_{i=1}^{k} animation_i)$ which represents interrupted animations.

The *Positioning System* simulates the motion tracking hardware and the processing of positional/rotational data streams aimed to detect the current user's location and head orientation. Like for the Visitor Component, these systems are simulated by means of non deterministic automata in order to check the overall systems behaviour. Further details about positioning and tracker systems will be provided in section 4.

2.2 Visual Engine

The purpose of the *Visual Engine* (VE) is to manage the avatar and the virtual (eventually animated) objects which augment the real environment surrounding the visitor, operating at their higher level of abstraction by means of the H-Anim ISO humanoid representation standard to model the avatar and through an XML description to represent other virtual objects . The output of this engine is a complete animated scene in VRML format, ready to be rendered by AR Interface. A detailed description is provided in the following subsections 3.2.1 to 3.2.4.

2.2.1 XML based representation for virtual objects and operating environment

From the VE viewpoint, virtual objects and operating environment are represented with two level of abstraction. At the lower level for each virtual object and for the whole operating environment a corresponding VRML file is provided. A typical VRML file associated to operating environment includes a 3D point cloud or polygonal mesh describing the real environment with the desired level of detail, while a VRML file associated to each virtual object available includes vertices, polygons, mapping coordinates, texture references, material description and animation data if any. Each 3D coordinate referenced in any VRML file, including avatar's low level description, is expressed with respect to the same reference system, assuring that all visual components share the same 3D space. At the higher level a single XML file represents virtual objects and operating environment. This description contains references to each VRML file and various kind of additional info by means of specific tags and attributes. This XML description provides a convenient way to manage the whole scene, allowing the VE to easily retrieve and manipulate every object or even part of them by extracting the XML fragment associated to the objects required.

2.2.2 *Animation Engine*

The Animation Engine manages any animation performed by avatar as well as its navigation through the operating environment. Moreover, it manages virtual objects animation and synchronization. The approach to real time avatar animation exploited in the Animation Engine is based on the decomposition of a set of complex actions in simpler and shorter motions, previously built without limitations in term of software tools or technology involved (any animation creation/editing software or mocap systems) and stored in an opportune repository called Avatar Motion Database. Indeed, H-Anim avatar's description represents the virtual character in term of a hierarchically arranged set of bones, each one affecting the surrounding polygonal surface, so that any action to be performed corresponds to an animation which in turn is a timed sequence of skeleton transformations (bones rotations around their local reference systems or overall body traslation relative to its center of mass). As the variety of action required for the avatar and their complexity may lead to a huge number of transformations (animation frames) we represent each complex action (for instance "point object switch_1") in term of concatenation of "atomic" animations. To join atomic animations, a well known technique called "motion blending" is exploited, resulting in almost seamless motion. Each action available for a particular application is included in the set:

$$HAnimSet = \{pointObject, goTo, lookAt, \dots\}$$

and is selected by the Integration Engine to be performed at a given moment. At this point the action name is resolved by the Animation Engine in a sequence of atomic animations depending on which class of motion (navigation, gesture, posture, etc) it belong to and also according to application oriented rules. For instance avatar navigation may require a walk cycle (a repeated atomic walk animation) between a starting and a destination point, whose variable length depends on the 3D position of avatar (*Humanoid node*) and the 3D position of destination (real or virtual object) and whose speed depends on the applicative context. The walking surface available for the avatar is defined as a sub-region of the whole environment in which some areas are flagged as obstacles. This representation is exploited to allow the avatar to move from a generic location to another exploiting a path-finding algorithm for autonomous virtual character based on a technique referred as "steering behaviour" to avoid collisions with obstacles including the real objects and the visitor as well. Avatar gestures like "point with finger" require to access the end effectors within the H-Anim model to dynamically modify the body joints involved via inverse kinematics. As said before, the Animation Engine is also responsible for the animation of virtual objects, which is completely defined at design time and embedded in the VRML object description (Object Database), but can be relocated, activated and stopped by means of a set of specific functions described below. Indeed, each virtual object is part of the set:

$$VirtualObject = \{object_1, object_2, \dots\}$$

and can be visualized at virtual world coordinates provided by the function:

$$coordinates: VirtualObject \rightarrow \mathbb{R}^3$$

and expressed with respect to a referring object given by:

$$referringObject: VirtualObject \\ \rightarrow VirtualObject \cup \{environment\}$$

Its animation will start at a time defined by function:

$$startTime: HAnimSet \times VirtualObject \rightarrow \mathbb{R}^+ \cup \{-1\}$$

and will terminate at time:

$$endTime: HAnimSet \times VirtualObject \rightarrow \mathbb{R}^+ \cup \{-1\}$$

where the -1 allows to delay the start or the end of animation over multiple actions.

2.2.3 *Augmented Reality Interface*

In the proposed approach the real world is augmented through a fully interactive synthetic character aimed to assist user on selected situations. This avatar can also interact with simpler context dependant 3D objects (animated or not) which can provide further visual help. According to a well established methodology the AR interface is based on a real time stereoscopic rendering engine able to sample user's motion data such as head's position and orientation, then processing these data to transform the virtual content as seen from user's point of view and coherently to a 3D model of surrounding environment, a crucial task referred as 3D registration. Indeed, any AR architecture requires a precise registration of real and virtual objects, in other terms the objects in the real and virtual world must be properly aligned with respect to each other, or the illusion that the two worlds coexist will be compromised. To achieve this not trivial goal two main requirements have to be satisfied: the position and orientation of user's head have to be precisely tracked at a high sample rate; the physical world or at least the objects relevant to the application has to be precisely measured in the same 3D space in which the user operates. Details about the technology adopted to capture user's head position/rotation to properly register the virtual content with respect to the real world are provided in section III. At runtime, the AR Interface pass the VRML descriptions outputted by the Visualization

Engine to the Rendering Engine. At this point two rendering cameras (one for each eye) are built, matching the exact position/orientation of user's eyes, transforming each vertex of each virtual object (including the avatar) to be displayed onto the real scene accordingly. Two renderings (for the left and right eye) are then calculated and coherently displayed through the HMD (see Fig 2). The rendering engine has been tailored to optical see-through HMD, but it can be easily adapted to video see-through displays as well. A partial or total culling of the avatar may be performed whereas it is partially or totally behind a real object, based on a previously built 3D model of the real environment (see section 3. for more details on this topic).

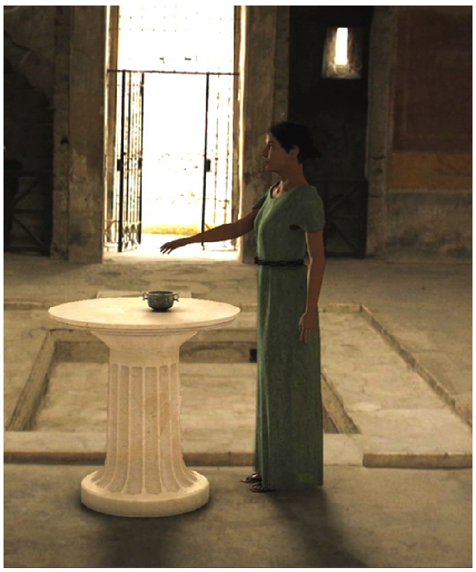


Figure 2. A third person view of the virtual guide next to a reconstructed table, as seen trough AR.

2.2.4 Speech Engine

As previously mentioned, one of the aims of the avatar as the visitor's virtual guide is to provide useful context adaptive information about the exhibits visited during the tour. This activity is fulfilled by means of a speech capability based on a speech synthesis engine coupled to a specific speech database. The latter is organized in layers which reflect the user typologies addressed by the framework and the avatar's appearance. In other terms for each description associated to a given station along the tour there are different speeches corresponding to various deepening levels and to different avatars (tourist, child, expert, female, male).

2.2.5 Integration Engine

The *Integration Engine* integrates the behavioural model and the visual engine in order to build a final runtime application. It allows to merge the information inferred by the automata network with the avatar's and virtual objects geometry and with the information coming from the operating environment (application-oriented data). This integration is accomplished defining a *behaviourAnimation* function associating each animation state of the *avatar behavioural automaton* with the corresponding avatar actions and object animations, as described below:

$$behaviourAnimation: \bigcup_{i=1}^{k} animation_i \cup \{start\} \rightarrow HAnimSet \times \mathcal{P}(VirtualObject)$$

This function maps a generic avatar's behavioural state, $animation_i$, in an ordered pair where the first component represents the action the avatar have to undertake and the second component represents the set of virtual objects active during this action (if any). In other words, this pair corresponds to an XML model provided as an input to the Visual Engine in order to setup the appropriate VRML scene to be rendered by the AR-interface.It has to be remarked that the *behaviourAnimation* function does not deals with operating environment data in a direct way. Indeed, operating environment data are embedded into each state of behavioural timed automata, i.e., in the set $\bigcup_{i=1}^{k} animation_i \cup \{start\}$.

3. THE FRAMEWORK AT WORK

The whole framework described above has been applied to a real applicative context represented by a roman villa located in the archaeological site of Pompei, and known as the Menandro's Home. The environment involved in the experiments covers a fraction of the overall villa's surface which is composed of many rooms surrounding a typical roman open ceiling garden. According to the archaeologists, eight locations have been selected for their particular significance. Considering the specific application (indoor environment) and the level of precision needed by augmented reality to operate, the user's tracking requirements has been addressed by means of a 6DOF (six degrees-of-freedom, three for x-y-z position and three for head's yaw, pitch, roll) ultrasonic based wireless motion tracking system produced by Intersense Corp. featuring a scalable coverage volume, high sample rate and a positional/rotational measurement accuracy adequate for this application. The ultrasonic sensors, arranged in strips, have been suspended to the ceiling of the selected rooms, excluding the garden and other open ceiling spots in which visitors could not be accurately tracked by this technology. As the avatar should be realistically occluded by real object while navigating the villa, a 3D model of the accessible environments and of the main visible objects is created. To this aim, firstly, a 3D scanning of the operative volume, including every object available, is conducted via a time of flight laser range scanner from 3rdTech. Inc. Then, the point cloud is simplified and used as a 3D reference template to model by hand a new low resolution version of the environment, featuring only the elements visually relevant to the avatar rendering and to its navigation. The HMD adopted is a Cybervision Visette Pro SXGA whose two video input are linked via a special

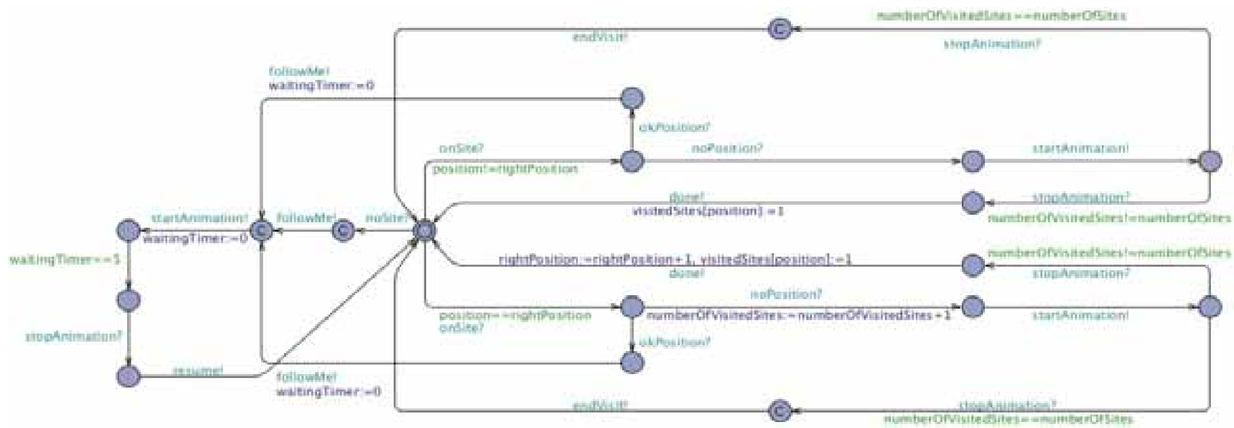


Figure 3. An example of timed automata modeling the avatar's behaviour.

adapter to a backpack portable PC based on Intel Duo core 2 2,6 GHz processor delivering the processing power required to support the various engines and equipped with a Nvidia 8600M GT graphics board exploited for the real-time rendering of the virtual contents. Additional lithium batteries are included in the backpack to supply power to the tracker and to the HMD. The equipment is exposed in an dedicated stand near the entrance of the archaeological site. According to the proposed approach, the different framework components for this application have been designed by means of specific automata and then validated through appropriate model checking queries. In Fig. 3 a transition graphs representing the behavioral automaton is shown.

4. CONCLUDING REMARKS

In this paper a complete framework for avatar-human interaction experienced through augmented reality has been presented. The proposed approach exploiting timed automata to allow synchronized communication and interaction between the user and the virtual character, proved to be suited to the case study described and arguably extensible to a wider applicative range as well. Further advantages of this methodology include a straightforward way to design and check the interaction framework, while the low computing load related to the automata based engines make them particularly suited to the real time response required by an augmented reality application. As the typical interaction paradigm between humans is based on known rules involving the reciprocal attention of both subjects (by means of eye movements, gestures, postures, etc.), future works could enhance the level of empathy in avatar-user communication implementing avatar's face expressivity according to user's face visemes and user's voice recognition could also be exploited to allow a full verbal interaction. A further enhancement to the framework could address the interaction among multiple avatars and between them and multiple user.

5. REFERENCES

[1] K. Mase, Aspects of Interface Agents: Avatar, Assistant and Actor," in Proceedings IJCAI'97 Workshop on Animated Interface Agents, 33-37.

[2] O. Stock, M. Zancanaro, P. Busetta, C. Callaway, A. Kruger, M. Kruppa, T. Kuflik, E. Not, C. Rocchi, Adaptive, intelligent presentation of information for the museum visitor in PEACH, User Model User-Adapted Interaction , Springer Science+Business Media B.V., 17, 257–304.

[3] I.N. Badler, Real-Time Virtual Humans, in Proceedings of the Fifth Pacific Conference on Computer Graphics and Applications, Seoul, South Korea, October 1997, 4–13.

[4] C. Peters, Foundations for an agent theory of mind model for conversation initiation in virtual environments, in Proceeding AISB 2005 Symposium on Virtual Social Agents. Hatfield, England.

[5] S. Lee, I. Kim, S. C. Ahn, H. Ko, M. Lim, H. Kim, Real time 3D avatar for interactive mixed reality, in Proceedings of the 2004 ACM SIGGRAPH VRCAI '04, Singapore, pp. 75-80.

[6] T. Ogi, T. Yamada, K. Tamagawa, M. Kano, M. Hirose, Immersive tele-communication using stereo video avatar, in Proceedings of IEEE VR, pp . 45–51.

[7] S. Lee, I. Kim, S.C. Ahn, M. Lim, H. Kim, Toward immersive telecommunication: 3D video avatar with physical interaction, in Proceedings ICAT '05 Christchurch, New Zealand, pp. 56-61

[8] K. Loer, M. Harrison, Model-based formal analysis of temporal aspects in human-computer interaction, in Proceedings of the HCI2003 Workshop on the Temporal Aspects of Tasks, Bath, United Kingdom.

[9] C. Norstrom, A. Wall, W. Yi, Timed Automata as Task Models for Event-Driven Systems, in Proceedings of Sixth International Conference on Real-Time Computing Systems and Applications.

[10] R. Alur, D.L. Dill, A theory of timed automata. Journal of Theoretical Computer Science, 126(2):183–235.

A visual language for paper-based computer-aided form sketching

Farrugia P.J.
Concurrent Engineering Research Unit, Department of Industrial and Manufacturing Engineering, University of Malta, Malta

E-mail: pjfarr@eng.um.edu.mt

Borg J.C.
Concurrent Engineering Research Unit, Department of Industrial and Manufacturing Engineering, University of Malta, Malta

E-mail: jjborg@eng.um.edu.mt

Camilleri K.P.
Department of Systems & Control Engineering, University of Malta, Malta

E-mail: kpcami@eng.um.edu.mt

Abstract

Due its efficacy in rapidly externalizing form concepts, paper-based sketching is still extensively used by practising designers. It is a common practice that form concepts are sketched on paper prior to being translated into three-dimensional (3D) virtual models for further development. However, the user-interface of commercial Computer-Aided Design (CAD) systems does not support the automatic translation from sketches to 3D models. This research is concerned with the development of a visual language to link form sketching on paper with CAD. Due to the idiosyncrasy in form sketching, this language is prescribed to the designers to communicate their ideas to the computer. To this end, the prescribed sketching language, PSL, is based upon modelling principles commonly found in CAD systems. PSL is composed of user- and pre- defined graphical elements and a set of rules for drawing them on paper.

1. Introduction

Sketching is one of the oldest methods used by artists and designers to visually support the development of ideas [1]. Traditional sketching is still very popular amongst designers, as it provides an efficient means to instantly capture ideas, not only inside, but also outside their office [2].

It is a common practice that form concepts externalized by sketching are translated into 3D virtual models for further development. However, although improvements aimed at integrating sketching with 3D modelling technology have been recently registered in commercial CAD systems, such as in *Autodesk AliasStudio* [3], they do not support the *automatic* conversion of *paper* sketches into 3D models. A paper-based sketch is only used as a guide for the subsequent 3D modelling. The manual transfer of sketches into CAD systems is time-consuming since ‘the designer is forced to invest a lot of time and thinking to concretize the raw sketch’ [4]. In view of these issues, this research is concerned with the development of means to link form sketching on paper with CAD.

The rest of this paper is organized as follows. Section 2 discusses related work on visual languages developed for supporting sketching. By using examples of form representations with *PSL*, the type of symbols and language rules are explained in Section 3. The subsequent section treats the formalism of *PSL*. Section 5 discloses the results of an evaluation carried out to assess *PSL* and the implemented parser. A discussion on *PSL* and its significance follows in Section 6. Finally, Section 7 draws conclusions from this work, with focus placed on the contribution disclosed in this paper.

2. Related work

Visual languages for supporting computer-aided sketching were developed for various applications. For instance, Caetano *et al.*[5] developed a visual language to create graphical user-interfaces by freehand sketching, based on symbols representing typical user-interface elements (e.g. text boxes). Do [6] developed a visual language for representing architectural plans. Hammond and Davis [7] developed a visual language for sketching Unified Modelling Language (UML) diagrams. Collectively, these systems highlight the importance of exploiting sketching for a natural human-computer interaction. Yet, they also suggest that a visual language aimed at bridging the gap between paper-based sketching and CAD, in a product form design context, is still missing.

3. *PSL* symbols and rules

PSL exploits *re-usable symbols* which map two-dimensional (2D) form descriptions on paper into 3D geometries. *PSL* has four sets of symbols, Set_{PL}, Set_{3DOP}, Set_{3DPR}, and Set_{FF}, which represent respectively planes, 3D operators (e.g. *extrude*), 3D primitives (e.g. a *cylinder*) and form features (e.g. a *threaded hole*). Consider the *PSL* form representation of a perfumery bottle shown in Figure 1.

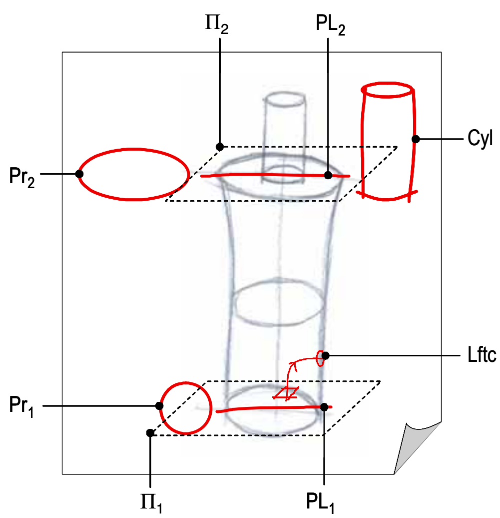
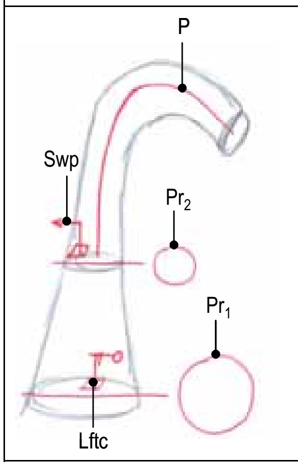


Figure 1. A *PSL* form representation

Note that the *PSL* annotation is drawn in a coloured pen, on the original pencil sketch. As shown in Figure 1, the two planes Π_1 and Π_2 are represented respectively by the plane lines PL_1 and PL_2. Essentially a plane line represents the side of the plane on which a user-defined profile, a 3D primitive or a form feature resides. In this example, Pr_1 resides on Π_1, whereas Pr_2 and the base of the cylinder primitive (represented by $Cyl \in Set_{3DPR}$) lie on Π_2. The loft symbol ($Lftc \in Set_{3DOP}$) operates on Pr_1 and Pr_2. Loft is a 3D operation commonly found in CAD systems by which complex geometries can be created by blending a series of cross-sections.

Table 1 illustrates further how the symbols in *PSL* accomplish the 2D-to-3D mapping. The rotational geometry of the vase is represented by a cross-section and two revolve symbols. In the second example, the loft symbol operates on Pr_1 and Pr_2, whereas the sweep symbol (*Swp*) operates Pr_2, along the path *P*. It must be mentioned that at present, *PSL* is restricted to forms having a 2D axis laying in the plane of the paper.

Table 1. Further examples of using *PSL*

PSL representation	3D model generated
	
	

As in any other language, the user of *PSL* has to observe a number of rules, in this case, to correctly represent a form. For instance, the axis of revolution passing through the centrepoint of each revolve symbol must not intersect the profile (see example 1 in Table 1). Another rule specifies that in order to assign a profile *Pr* to a plane line, *PL*, two spatial relationships, namely, (*projection-of Pr PL*) and (*closest-to Pr PL.ep1* | *PL.ep2*) must be observed. These relationships mean respectively that *Pr* must be drawn in the projection of *PL* and closest to either endpoint of *PL* (denoted by *PL.ep1* and *PL.ep2*).

Following the previous examples, essentially, with the *PSL* symbols and rules, the designer represents the construction of the form geometry in a CAD system.

4. *PSL* formalism

Several general approaches to the specification of visual languages may be grouped together as attributed multiset grammar based approaches [8]. These approaches are characterized by production rules which rewrite sets or multisets of symbols having geometric and semantic attributes associated with them. Besides specifying what elements compose the representation of a form geometry, the *PSL* rules specify constraints on the attributes of the *PSL* elements as in an attributed grammar. To this end, the underlying principle of such a grammar has been adapted to formalise and subsequently to parse *PSL*.

To explain how this is accomplished, the *PSL* representation (denoted by *PSL_rep*) of the second example in Table 1, is taken as a case study. As shown in the parse tree in Figure 2, *PSL_rep* is decomposed into two types of geometries which are represented by a 3D operation; in this case we have a loft geometry (*Lftc_op_geom*) and a sweep geometry (*Swp_op_geom*). Note that the order of plane lines (i.e. PL_1, PL_2) and profiles (i.e. Pr_1, Pr_2) in the parse tree is deduced by constraints on the geometric attributes of these elements. In addition, which 3D operator symbol acts on a particular profile, depends on the symbol's spatial position with respect to the plane line associated with the profile. In view of this, the production rules in the grammar consist of:

1. the *PSL* elements required to represent the construction of the particular form geometry (e.g. *Lftc* PL_n Pr_n PL_k Pr_k);
2. the constraints on the geometric attributes of the *PSL elements*, such as (*closest-to Pr PL.ep1* | *PL.ep2*) mentioned previously;
3. the attributes of a *form geometry*. One such attribute is the starting and ending point of a form geometry. For instance *Lftc_op_geom* starts from Pr_1 and ends at Pr_2 ;
4. the meaning of the form geometry derived as a result of the attributes in (3). This meaning is conveyed to a CAD system to eventually construct the 3D model. For example, the meaning of *Lftc_op_geom* is given by:

$$Lftc\_op\_geom.meaning = Lftc\_cmd(\left\langle Pr_1.\{p1\_x, p1\_y \ldots pn\_x, pn\_y \right\rangle , (\left\langle Pr_2.\{p1\_x, p1\_y \ldots pn\_x, pn\_y \right\rangle \}$$

where:

Lftc_cmd represents the loft command in the CAD system operating on the profiles Pr_1 and Pr_2, having the set of salient vertices *p1_x, p1_y… pn_x, pn_y*.

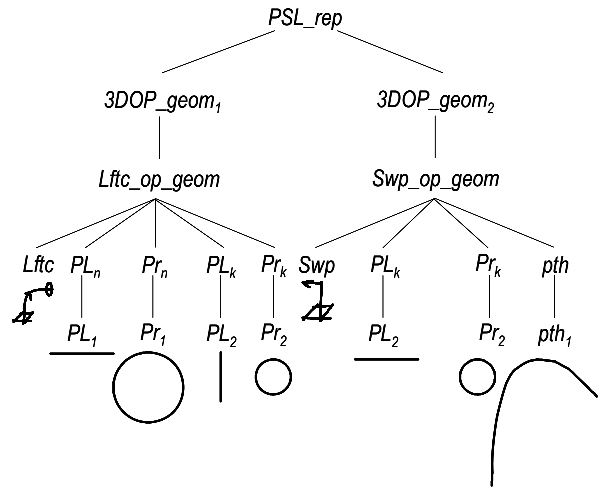
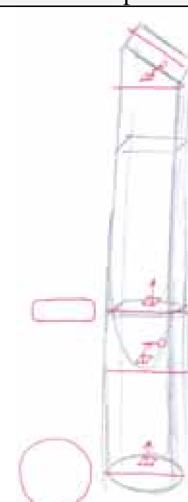


Figure 2. Parse tree of *PSL_rep*

5. Evaluation

The evaluation goal was to assess firstly the modelling capability of *PSL* for the objects it was applied to, and secondly the parser efficiency. Eleven existing objects were employed in this experiment, most of them requiring more than one 3D operation to represent the form (see example in Table 2).

Table 2. Example of an object used to test *PSL*

Existing object	*PSL* rep.	3D model
		

It must be emphasized that the evaluation was qualitative, aimed at assessing the semantic potential of *PSL* and related problems that may emerge. The paper-based sketch of each test object was scanned and parsed using the aforementioned attributed grammar.

It resulted that with the four classes of *PSL* symbols it was possible to *approximate* the form geometry of the eleven objects. Moreover, in all cases, the parser successfully derived the syntactic structure of each *PSL* representation and subsequently the 3D model was correctly generated.

At the same time, the experiment demonstrated that it was not possible to represent certain features with *PSL*, such as the hollow section in the object illustrated in Table 2.

6. Discussion

This paper argued that a *PSL* is required to bridge the gap between sketching and CAD. Computer-aided sketching systems which resorted to symbols as means to communicate form to the computer were developed [2]. However, they replace paper with a digital medium. By contrast, the work presented in this paper exploits 3D modelling symbols for *paper*-based computer-aided sketching.

We argue that *PSL* has implications on the development of new user-interfaces of CAD systems. In addition, based on the *PSL* parsing, a CAD system can reason out, directly from the original sketch, the most plausible configuration of an assembly (e.g. a flange and a shaft). Given that *PSL* is based on 3D modelling principles, we also argue that it promotes 3D modelling education; since freehand sketches are marked up with *PSL*, designers who are skillful in sketching, but who do not have CAD experience, can appreciate how their concepts are generated in CAD.

Future work is however required to extend the practical use of *PSL*, particularly to represent a wider range of geometries beyond those having a planar axis and to test further the efficiency of the parser.

7. Conclusions

This paper contributed a new visual language, *PSL*, whose novel characteristic lies in the exploitation of symbols which operate on user-defined profiles to map form concepts sketched on paper into 3D CAD models. Also this paper demonstrated how an attributed grammar approach can be applied to specify a visual language for product form design.

8. Acknowledgments

The authors wish to thank Alexandra Bartolo for her input in image processing. The financial support provided by the University of Malta, Malta, through research grants 'Emotional Support Tool for Early Product Design' (EST 73-751) and 'Innovative 'Early Stage' Design Product Prototyping' (IED 73-529) is also greatly appreciated.

9. References

[1] Muller, W., *Order and Meaning in Design*. 2001: Lemma Publishers.

[2] Farrugia, P.J., J.C. Borg, X.T. Yan, K.P. Camilleri, and G. Graham, *A Sketching Alphabet for Paper-based Collaborative Design.* Journal of Design Research (JDR), special issue on Fostering innovation during Early Informal Design Phases, 2007. **6**(1,2): p. 260-288.

[3] AliasStudio, *AutoDesk Alias Studio Tools*. 2006. retrieved from: www.autodesk.com/aliasstudio. Date accessed: 4th June 2008.

[4] Diehl, H., F. Muller, and U. Lindemann. *From raw 3D-Sketches to exact CAD product models - Concept for an assistant-system*. in *Workshop on Sketch-Based Interfaces and Modeling 2004, Eurographics Symposium*. 2004. Grenoble, pp.137-141.

[5] Caetano, A., N. Goulart, M. Fonseca, and J. Jorge. *JavaSketchIt: Issues in Sketching the Look of User Interfaces*. in *AAAI Spring Symposium on Sketch Understanding*. 2002. Stanford, CA, pp.9-14.

[6] Yi-Luen Do, E., *Investigation of Freehand Drawing as an Interface to Knowledge Based Design Tools*. 1998, Ph.D. Thesis, Georgia Institute of Technology.

[7] Hammond, T. and R. Davis. *Tahuti: A Geometrical Sketch Recognition System for UML Class Diagrams*. in *AAAI Spring Symposium on Sketch Understanding*. 2002. Stanford, CA: AAAI Press, pp.59- 66.

[8] Marriott, K., B. Meyer, and K. Wittenburg, *A Survey of Visual Language Specification and Recognition*, in *Visual Language Theory*, K. Marriott and B. Meyer, Editors. 1998. p. 5-85.

A Geometric-based Sketch Recognition Approach for Handwritten Mandarin Phonetic Symbols I

Paul Taele, Tracy Hammond
Sketch Recognition Laboratory
Department of Computer Science
Texas A&M University
Mail Stop 3112, College Station, TX 77839
{ptaele, hammond}@cs.tamu.edu

Abstract

Inputting written Chinese, unlike written English, is a non-trivial operation using a standard keyboard. To accommodate this operation, numerous existing phonetic systems using the Roman alphabet were adopted as a means of input while still making use of a Western keyboard. With the growing prevalence of computing devices capable of pen-based input, naturally sketching written Chinese using a phonetic system becomes possible, and is also generally faster and simpler than sketching entire Chinese characters. One method for sketching Chinese characters for computing devices capable of pen-based input involves using an existing non-alphabetic phonetic system called the Mandarin Phonetic Symbols I (MPS1). The benefits of inputting Chinese characters by its corresponding MPS1 symbols – unlike letters from its alphabetic-based counterpart – is that it retains the phonemic components of the corresponding Chinese characters. The work in the paper describes our geometric-based MPS1 recognition system, a system designed particularly for novice users of MPS1 symbols that gives reasonable vision-based recognition rates and provides useful feedback for symbols drawn with incorrect sketching technique such as stroke order.

1. Introduction

Unlike written English, where words can be represented by the twenty-six letters of the Roman alphabet, written Chinese consists of tens of thousands of distinct characters in common usage [9]. To accommodate the non-trivial operation of inputting Chinese characters into a computer, numerous existing Roman alphabetic phonetic systems were adopted as a means of input while still making use of the Western keyboard. For the growing number of computing devices (e.g., Tablet PCs, mobile computing devices) which allow input from a pen-based device, it would be more desirable to input symbols from a phonetic system, since it would be more natural than inputting symbols of a phonetic system from a software keyboard, and generally faster and easier than sketching entire Chinese characters.

One phonetic system which can benefit from pen-based input is the Mandarin Phonetic Symbols I (MPS1), a system developed and commonly used in Taiwan (Figure 1). The advantage of MPS1 is that it is a phonemic system which allows Chinese words – particularly Chinese Mandarin, the largest language in the Chinese language family – to be accurately represented by its symbols. This differs from other phonetic systems which use Romanization and which do not necessarily reflect the original sounds of the associated Chinese characters, such as the popular Hanyu pinyin, the legacy Wades-Gile, and the now obsolete Mandarin Phonetic Symbols II (MPS2) introduced simultaneously with MPS1. Given the prevalence of pen-based computer input, using MPS1 as a means of inputting Chinese characters has a growing relevance of providing an ideal and phonemic means of inputting Chinese characters.

Figure 1. The thirty-seven MPS1 symbols. The adjacent English letters are their Romanized equivalents (source: Wikipedia)

In order to accommodate the use of MPS1 as a means of input, sketch recognition techniques must be employed in order for a system to be capable of processing them. Techniques employed by those existing systems in general, such as written Chinese, typically fall into two categories. The first category uses gesture recognition, which requires characters to be drawn in a particular stroke order and direction before a character can be recognized. The problem with this approach is that it can cause users to be frustrated when a character that is identifiable by a human is misclassified by the system due either to incorrect stroke order. The second category uses vision-based recognition, which disregards aspects such as stroke order of a character in favor of its shape and structure in relation to previously trained data. The problem with this approach though is that it encourages and rewards incorrect stroke orderings. Since stroke ordering is key for learning written Chinese characters [6], and since symbols in MPS1 are derived from written Chinese, this latter approach is not ideal at a pedagogical level for allowing novice users to improve their sketching technique of MPS1 symbols.

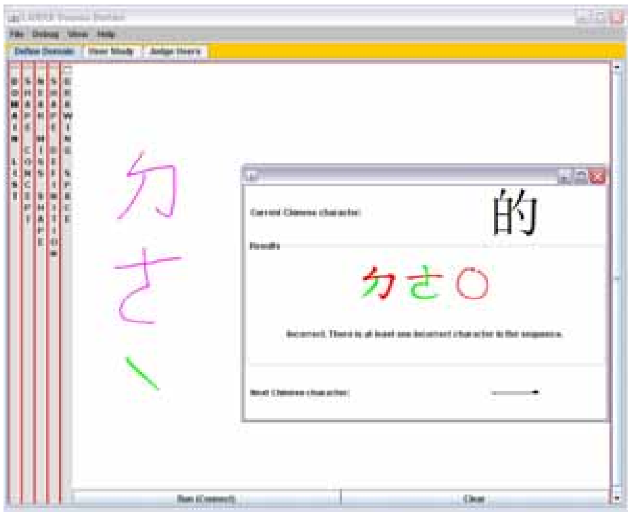


Figure 2. A sample application of sketching MPS1 symbols using our recognition system for determining both visual and technical correctness.

The work in this paper describes our geometric-based sketch recognition system for the phonetic symbols in MPS1, which can recognize the shape of the symbol regardless of how it was drawn and also identify the correctness of the stroke ordering (Figure 2). This is because our system retains relevant information from user sketches in identifying the level of sketch technique correctness in their MPS1 symbols. The benefits of our system includes giving reasonable recognition rates for classifying the visual-correctness of user-sketched MPS1 symbols, retaining useful sketch properties such as stroke order, and providing useful feedback to novice users for improving their sketch technique on those same symbols.

2. Previous work

Two different approaches have been explored in the domain of sketch recognition: vision-based and geometric-based. Their relation to our system is elaborated below.

2.1. Vision-Based Sketch Recognition Work

Various research groups such as IBM and Microsoft [8] have labs dedicated to research in vision-based sketch recognition specifically for handwritten recognition of written Chinese. Publicly available systems usually come in the form of input method editors, or IMEs, which are programs that enable users to input symbols in East Asian languages with either a keyboard or pen-based device [5]. The sketching functionality found in popular IMEs, such as the Microsoft IME for the Windows XP operating system, use vision-based techniques to recognize handwritten characters, but currently do not support direct sketching of phonemic systems such as MPS1 for inputting Chinese characters.

Three other systems do allow for written Chinese to be input directly using phonemic systems. The first system by Lai [4] uses a novel approach in allowing users to select MPS1 symbols manually with a pen-based device in a rapid, hierarchical fashion, Cheng's system [2] allows stylus-input of both Romanized- and phonemic-based phonetic systems, and the system by Bao [1] uses Hidden Markov Models which allows for specific Chinese characters possessing specific phonemic sounds as an alternative to MPS1 symbols for inputting written Chinese. The downsides of each of those systems are that they assume fluency of the systems and don't take advantage of natural input (i.e., Lai), discard useful data such as stroke order, direction, and explicit proportions of sketched input (i.e., Cheng, Bao), and make use of non-standard phonetic systems (i.e., Bao). In other words, these systems were designed for users already familiar with written Chinese and MPS1 symbols.

2.2. Geometric-Based Sketch Recognition Work

Various papers have been written in regards to the general area of geometric-based sketch recognition. One aspect of this area includes primitive shape recognition. These recognizers involve determining if a pen stroke is a certain type of stroke primitive such as an ellipse, line, point, or polyline, while each stroke primitive is assigned an independent classifier to aid these recognizers. For polyline classifiers in particular, strokes are segmented into their respective line components. One such segmenter by Sezgin [3] uses pen velocity and curvature data to find the corners of pen strokes for later segmentation. The classified strokes can then be referenced and later used for another aspect of geometric-based sketch recognition called domain shape recognizers.

For domain shape recognizers, knowledge-based shape recognizers are used to understand the geometric properties of shapes for a variety of domains. One system is LADDER [7], a sketching language for describing how sketched shapes can be drawn, displayed, as well as edited. Shapes are described using a set of geometric constraints, and sketches which best fulfill a particular shape description are described as such.

Since vision-based sketch recognition approaches have numerous problems in handling the type of functionality we desired for our system, we instead utilized a geometric-based sketch recognition approach to compensate for those deficiencies. By shifting the domain of MPS1 sketch recognition to a geometric-based approach for an audience of novice users, we succeed in creating an MPS1 sketch recognition system with reasonable accuracy while also providing valuable feedback for improving sketch technique.

3. Implementation

The system we desired should not only gave feedback on the visually correctness of MPS1 symbols (i.e., vision-based approaches), but should also provide additional feedback for its technical correctness (i.e., geometric-based approaches). Our system takes advantage of the strengths of both approaches by utilizing and extending existing geometric-based tools in a two-stage process. In the first stage, our system pre-processes user-sketched MPS1 symbols for visual correctness, regardless of technical correctness. When the user sketches a visually-correct MPS1 symbol, our system offers feedback by highlighting the relevant sketch for its successful recognition. Once the MPS1 symbols are recognized, post-processing is done on the sketched symbol by analyzing temporal data of the primitive shapes on the recognized symbol. The criterion for technical correctness in our current system is stroke ordering, and the system provides different feedback depending on the sketched technique of the MPS1 symbol made by the user.

3.1. Resources

The symbols in the MPS1 domain consist primarily of polylines, so we required a robust primitive shape recognizer to classify sketched MPS1 symbols into computer-recognized form. Therefore, the Sezgin recognizer fit our needs to handle polyline classification. In addition, we desired a domain shape recognizer to handle the task of combining these polylines correctly to their recognized forms. The LADDER sketching language allowed us to do so by approximately describing MPS1 symbol geometrically for later recognition. The use of these two tools in conjunction aided us in the creation of our system.

3.2. Pre-processing: Building Shape Descriptions

The first recognition phase of our system involves visual correctness, which is done during the pre-processing of MPS1 symbols before analyzing it for technical correctness. We observed that the more complex symbols in the MPS1 domain were usually composed of existing MPS1 symbols. We exploit this trait by partially building the more complex symbols out of the simpler symbols. The first benefit of establishing all the visual connections for building more complex symbols in this manner was the elimination of duplicating shape descriptions from simpler MPS1 symbols into the more complex MPS1 symbols. Furthermore, since our system highlights recognized symbols which are visually correct, we can provide that same kind of feedback to the user in displaying partial correctness of the more complex symbols during sketching, as opposed to having the user sketch the entire complex symbols and not knowing which part of the sketch was responsible for the incorrectness in the misrecognition. Figure 3 shows a subset of the symbol visual relationship tree we used to aid in our construction of MPS1 shape descriptions.

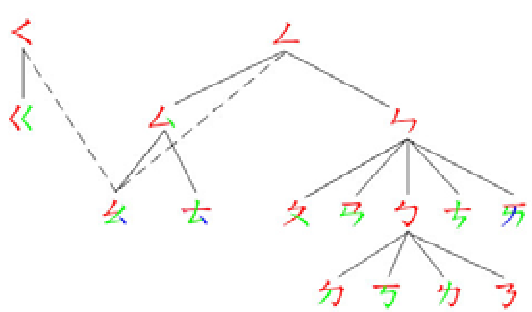
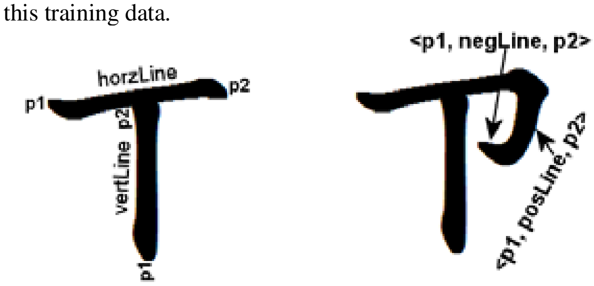


Figure 3. A subtree of the visual relationships between the various MPS1 symbols. Bold lines indicate direct relationships, and dotted lines indicate possible relationships ().

Once we developed a tree showing the visual relationships between the different MPS1 symbols, we concentrated on building shape descriptions for each of the individual symbols using the LADDER sketching language. Three important attributes in LADDER shape descriptions are: components, constraints, and aliases. Components establish the primitive shapes that make up a symbol, constraints define all pre-conditions that need to be fulfilled for proper classification, and aliases are user-created synonyms of components which allow for components to go by multiple descriptions. In order to facilitate the process of constructing shape descriptions, we conducted a user study involving natural sketching of MPS1 symbols by expert users. These expert users consisted of five university students of Chinese and Taiwanese descent, all of whom were fluent in written Chinese and MPS1 symbols. Our final shape descriptions for the set of MPS1 symbols were based on this training data.



(a) (b)

Figure 4. Visual shape and our assigned labels for two MPS1 symbols: (a) ㄒ [X] and (b) ㄗ [Z].

(a) Partial Shape Description for ㄒ(X)		
components:		
Line	horzLine	
…		
constraints:		
…		
aliases:		
Line	rightPoint	horzLine
…		

(b) Partial Shape Description for ㄗ(Z)		
components:		
Line	posLine	
Line	negLine	
constraints:		
not negSlope	posLine	
negSlope	negLine	
leftOf	posLine.p1	posLine.p2
leftOf	negLine.p1	negLine.p2
coincident	x.rightPoint	posLine.p2
coincident	posLine.p2	negLine.p1
aliases:		
…		

Table 1. Partial LADDER shape descriptions for (a) ㄒ [X] and (b) ㄗ [Z].

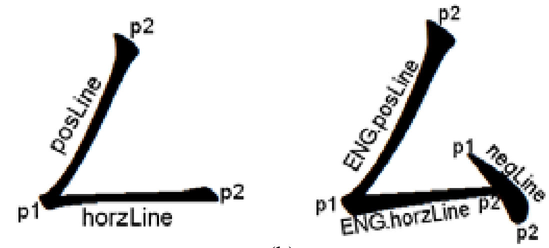
To demonstrate how we used LADDER to build these shape descriptions, we list the shape descriptions for a more complex MPS1 symbol which builds on a simpler MPS1 symbol. The symbols and the labels for the primitive line and point shapes are given in Figure 4 above. We also give a partial list of the corresponding shape descriptions to those symbols in Table 1. Due to space constraints, we only give the most relevant attributes to describe our process of building MPS1-specific shape descriptions. In defining shape descriptions suitable for sketched MPS1 symbols, we first list all single lines within that symbol and label them descriptively enough to indicate their orientation (e.g., "horzLine" denotes a horizontal line, and "posLine" denotes a line with a positive slope). If the MPS1 symbol is composed of other simpler MPS1 symbols, the simpler symbols become listed as components as well. It should be noted that two of the thirty-seven symbols – ㄛ(O) and ㄜ(E) – also contain arcs. To handle those two special cases, we approximated user-sketched arcs of those symbols as two coincident lines. We discovered that this was a better solution than incorporating arc primitives to our domain, because natural sketching of MPS1 symbols by users would prematurely trigger numerous negative instances of arcs that were non-existent when working solely with line primitives.

Secondly, we list the minimum constraints necessary to both enable correct recognition of symbols from our training data and also avoid shape collisions (i.e., cases of incorrect shape recognition due to one shape being too similar to another). We arrange the list of constraints into three groups so that shape descriptions can be rapidly constructed and easily referenced. The first group of constraints defines the orientation of line components, such as whether it is horizontal, vertical, positively-, or negatively-sloped. Since we gave descriptive names to these components, defining this group of constraints is highly streamlined. The second group explicitly defines spatial relationships between endpoints of lines. Since we used the Sezgin recognizer, only the endpoints of recognized line primitives are kept. LADDER denotes these endpoints as p1 and p2, so our method of explicitly defining spatial relationships of the endpoints makes it easier to construct our third group of constraints. This third group usually contains the largest list of constraints that describes how all the components in the symbol interact with each other. LADDER contains a variety of descriptive constraints, and we utilize a small subset of them to define those interactive relationships. The more important constraints in our domain include: *concident, sameX, sameY, near, closer, and contains.*

Lastly, we define aliases for existing line, point, and specific MPS1 shape components. Defining aliases are typically necessary to reference components of one MPS1 symbol into another symbol. The reason why the LADDER sketching language does not permit direct access to these components without aliases is to minimize any combinatorial problems which occur in globally searching through all possible geometric components in sketches to existing candidate shape descriptions. Aliases also give us an opportunity to give better or more pertinent labels to components of one symbol used in another symbol. This can be seen in Table 1, where the component *horzLine.p2* in the shape description of ㄒ is more descriptive as the alias *rightPoint* in the shape description of ㄗ. Aliases provide another important role in the post-processing phase of our system, which we shall demonstrate in the next section.

3.3. Post-processing: Temporal Data and Aliasing

The first recognition phase of our system involves technical correctness in terms of stroke order, since proper stroke order plays an important role in learning Chinese characters and MPS1 symbols in particular. Post-processing is named as such since we process stroke order technique correctness after visual recognition through LADDER's shape descriptions has already occurred. The use of aliases introduced in the previous section play a key role in determining the correctness of stroke order technique in user sketches. This is done by initially creating aliases for each line primitive within the MPS1 symbols. Aliases for each primitive line are then enumerated as *stroke#*, where # reflects the order of that stroke for that given line. We give an example of this aliasing process visually in Figure 5 and by their shape descriptions in Table 2 for both a simple MPS1 symbol and a complex MPS1 symbol partially composed by that simpler symbol.



(a) (b)

Figure 5. Visual shape and our assigned labels for two MPS1 symbols: (a) ㄥ [ENG], (b) ㄙ [S].

(a) Partial Alias List for ㄥ[ENG]		
aliases:		
…		
Line	stroke1	posLine
Line	stroke2	horzLine

(b) Partial Alias List for ㄥ[S]		
aliases:		
…		
Line	stroke1	eng.stroke
Line	stroke2	eng.stroke2
Line	stroke3	negLine

Table 2. Partial alias lists for (a) ㄥ [ENG] and (b) ㄙ [S].

After the sketched symbols are processed and classified in LADDER, we utilize these aliases in post-processing for stroke order correctness. This is first done by collecting all primitives recognized as lines for the classified MPS1 symbol. Since LADDER already segmented the line primitives in pre-processing, we collect each of those line primitives by their *stroke#* aliases in sequential order based on the temporal information stored in the raw sketch data. The collected lines are stored in a list, and we loop through each of the lines by their *stroke#* aliases to determine stroke order correctness. If a user sketched a MPS1 symbol in the correct stroke, the *stroke#* aliases would also be enumerated sequentially correct. For example, if a MPS1 symbol contains three line primitives, then a properly sketched symbol by stroke order should form a list of line primitives from its temporal information as [stroke1, stroke2, stroke3]. If the list shows that the strokes are not ordered in sequential order, then this means the user did not sketch the MPS1 symbol by its correct stroke order. This information thus becomes available for possible feedback.

3.4. The Application

In order to showcase our recognition system, we created a learning tool for novice MPS1 users to improve their knowledge and technique of these symbols (Figure 1). The application involves testing users on both the correct sequence of MPS1 symbols associated with Chinese characters and also the correct sketching technique of those symbols. Figure 6 shows the ten Chinese characters used in our application for testing our user of their MPS1 knowledge and writing skills. These characters were chosen because they are the ten most common found in the Chinese Mandarin language.

的 一 是 不 了
人 我 在 有 他

Figure 6. The ten Chinese characters used in our learning tool for testing users on their knowledge of these characters' associated MPS1 pronunciation.

In addition to the thirty-seven symbols in the MPS1 set, we also created shape descriptions for the five different tones in the Chinese Mandarin language. Since Chinese Mandarin is a tonal language, it is necessary to also test users on the tones of Chinese characters.

Figure 7. One of three possible display windows in our learning tool for critiquing the user's sketched symbols.

For our learning tool, the user first runs our system to display a window of a randomly selected Chinese character. The user then sketches that character's associated sequence of MPS1 symbols which represents its pronunciation. After the user finishes sketching what he or she thinks is the sequence, our learning tool processes the sketches for both visual and stroke order correctness. The user's performance for the sequence of MPS1 symbols results in the displaying of windows featuring one of three outcomes. The first outcome is that the sequence of MPS1 symbols has one or more incorrect symbols in the sequence, and the window indicates this and displays the correct sequence to the user. The second outcome is that the sequence is correct, but one or more of the symbols was not sketched in the correct stroke order. The window for this outcome indicates this case and also highlights all symbols in the sequence with the incorrect stroke containing the correct color-coded stroke order (Figure 7). The third and last outcome is that the user sketched the correct sequence and with the correct stroke order.

After the user completes sketching MPS1 symbols for all ten of our tested Chinese characters, our learning tool displays a results window of the user's performance. This results window can be found below in Figure 8.

We assigned a point-system to our particular learning tool in judging the performance of a user's knowledge in MPS1. For each tested Chinese character, a user receives ten points for a correct sequence of MPS1 symbols with correct stroke order, five points for a correct sequence with at least one symbol containing incorrect stroke order, and no points for an incorrect sequence. Our learning tool is easily scalable to include even more characters for use in a Chinese language curriculum.

Figure 8. The results window of our learning program for critiquing the user's performance for all ten tested Chinese characters.

4. Results

It is vital to gauge the performance of both our recognition system and our learning program in order to determine if they would be feasible in an actual learning curriculum. We first evaluated the performance of our recognition system on a set of MPS1 symbols sketched by five university students distinct from the users whose sketches were the basis of our LADDER shape descriptions. The new group of users was also fluent in written Chinese and MPS1 symbols, and the user study criterion was similar except we asked these new users to sketch MPS1 symbols as if they were using their sketches to teach novice users how to draw them. This change was requested because we wanted our system to test on model MPS1 symbols for use in a learning curriculum. With this new test data, our recognition system was able to achieve an accuracy rate of over 94%.

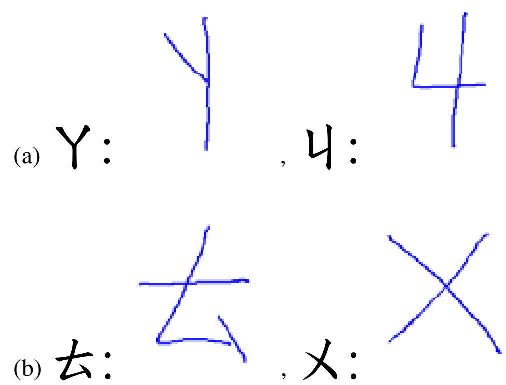


Figure 9. User-sketched MPS1 symbols by novice users which our system classified as incorrect: (a) Incorrect visual structure from the pre-processing phase, and (b) incorrect stroke order structure from the post-processing phase.

We then evaluated the performance of our learning tool which made use of the recognition system evaluated above. Our evaluation consisted of determining whether our system would be able to recognize the correctness of sketched MPS1 symbols by their visual structure, and also the correctness by their stroke order structure. To gauge the learning tool's performance, we conducted another user study involving a separate group of five users of American descent with no experience in written Chinese and MPS1 symbols. The task involved having them use our program for evaluating our learning tool designed for novice users. For each Chinese character given by our learning tool, we briefly showed these users a sheet of the correct MPS1 symbols corresponding to that character before their sketch. The result of this user study showed that our system was able to correctly differentiate visually-incorrect as well as stroke order-incorrect symbols (Figure 9).

5. Discussion

Based on the results given in the previous section, we were able to develop a recognition system with reasonable accuracy and a learning tool capable of providing valuable feedback for novice users. In regards to our recognition system's accuracy rate of 94%, we found this to be quite reasonable for the set of MPS1 symbols. We observed that our accuracy rates suffered primarily for the ㄛ and ㄜ symbols, since we approximated the arcs in those symbols as lines. Since the Sezgin recognizer used in our system performs best in a domain with polylines, we restricted our shape descriptions to lines in general. The only exception was for one of the tones used in our learning tool, which made use of the ellipse shape primitive unrelated to an arc shape primitive. Arcs are difficult primitives for our system to classify with the Sezgin recognizer, partly because there are many variations to draw arcs. On the other hand, we discovered that while users were not able to draw arcs approximated as lines in our recognition system correctly on the first try, they were able to adjust their sketches to cleaner arcs after a few tries to complete our learning program. We can rectify this problem of our recognition system handling arcs by utilizing a separate primitive shape classifier better designed for classifying arcs.

While the accuracy rate of our system is not as high as the 95% and better accuracy rates for various vision-based system in the larger domain of Chinese characters, our approach can perfectly differentiate between positive and negative examples in terms of stroke order. In other words, techniques employed in existing vision-based systems would correctly classify the intended MPS symbol sketched at a higher rate than our system, but they would do so even if those symbols contain incorrect stroke order. Such vision-based systems were not designed for novice users in mind, and thus would not be as suitable for teaching novice users on correct sketching technique of MPS1 symbols.

One other point in regards to our recognition system is that it currently handles only stroke order correctness in the post-processing phase. Several other characteristics important in Chinese characters and MPS1 symbols include stroke direction and line proportions. While we did not incorporate the testing of correctness for those characteristics, our recognition system can be easily extended to handle them using our aliasing technique. For stroke direction and line proportions, we can apply a similar scheme for determining proper stroke direction using the raw sketch data's temporal and spatial information, respectively.

6. Future Work

The previous section discussed several areas which can be improved and further extended. In addition to those areas, one future work we are currently investigating is employing our techniques to language curriculums which use Chinese characters, such as those found in Chinese and Japanese language programs. Numerous textbooks employed by these language programs still use paper-based approaches to test students on the difficult domain of written Chinese characters. We hope these programs can adopt our system to provide useful feedback on written technique core in language studies.

7. Conclusion

Inputting Chinese characters into a computer is far from a trivial task. One ideal method for doing so – while incorporating natural sketching and retaining unambiguous phonetic structure – involves the use of the Mandarin Phonetic Symbols I system. While existing vision-based systems are capable of handling sketching with MPS1 symbols, such systems were not designed for novice MPS1 users whom could benefit from valuable feedback on their sketching technique while learning symbols in that domain. Our geometric-based recognition system can provide such valuable feedback at comparable vision-based accuracy rates to existing vision-based system, and can also be extended for use in language curriculum-related applications.

8. Acknowledgements

This research is supported by the NSF IIS Creative IT Grant # 0757557 Pilot: Let Your Notes Come Alive: The SkRUI Classroom Sketchbook.

This research is also supported by the NSF IIS HCC Grant #0744150 Developing Perception-based Geometric Primitive-shape and Constraint Recognizers to Empower Instructors to Build Sketch Systems in the Classroom.

The authors thank the members of the Sketch Recognition Lab for their valuable assistance in this research.

9. References

[1] Bao, L., Toyama, J., and Shimbo, M. "Mandarin phonetic recognition using mixture hidden Markov models with time duration function." In *Fourth International Conference on Signal Processing Proceedings,* 1998.

[2] Cheng, E., Amar, R., and Matic, N. "Chinese Phonetic Handwritten Text Input for Mobile." In *UIST 2005: ACM Symposium on User Interface Software and Technology*.

[3] Hammond, T. and Davis, R. "LADDER, a sketching language for user interface developers." *Computers & Graphics*, vol. 29, no. 4, pp. 518-532, 2005.

[4] Lai J.T.C., Li Y., and Anderson, R. "Donuts: Chinese Input with Pressure-Sensitive Marking Menu." In *UIST 2005: ACM Symposium on User Interface Software and Technology.*

[5] Lunde, K. "CJKV Information Processing." Sebastopol: O'Reily & Associates, Inc., 1999.

[6] National Taiwan Normal University. "Practical Audio-Visual Chinese 1 Textbook, Vol. 1." Taipei: Cheng Chung Book Company, Ltd., 2004.

[7] Sezgin, T.M., Stahovich, T., and Davis, R. "Sketch based interfaces: early processing for sketch understanding." In *PUI '01: Proceedings of the 2001 workshop on Percerptive user interfaces*, (New York, NY, USA), pp. 1-8, ACM Press, 2001.

[8] Tay, Y.H., Khalid, M., and Yusof, R. "Online Chinese handwritten character recognition: a brief review." http://citeseer.ist.psu.edu/596162.html

[9] Yang, Y.Y. "Adaptive recognition of Chinese characters: imitation of psychological process in machine recognition." In *IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans*, Vol. 28, No. 3, 1998.

Sketch-Based Retrieval of Drawings using Topological Proximity

Pedro Sousa   Manuel J. Fonseca
Department of Computer Science and Engineering
INESC-ID/IST/Technical University of Lisbon
R. Alves Redol, 9, 1000-029 Lisboa, Portugal
pedrocrs@gmail.com, mjf@inesc-id.pt

Abstract

Currently, there are large collections of drawings from which users can select the desired ones to insert in their documents. However, to locate a particular drawing among thousands is not easy. In our prior work we proposed an approach to index and retrieve vector drawings by content, using topological and geometric information automatically extracted from figures. In this paper, we present a new approach to enrich the topological information by integrating spatial proximity in the topology graph, through the use of weights in adjacency links. Additionally, we developed a web search engine for clip art drawings, where we included the new technique. Experimental evaluation reveal that the use of topological proximity results in better retrieval results than topology alone.

Figure 1. Prototype for drawing retrieval.

1. Introduction

Nowadays, there are a lot of vector drawings available for integration into documents, either on the Internet or on clip art collections sold in optical media. This large number of drawings makes traditional searching mechanisms, based on browsing and navigation in directories or complex mazes of categories, inadequate. Furthermore, solutions using keywords or tagging are also impracticable since they have to be generated manually. A more adequate solution must take into account information automatically extracted from the content of drawings, instead of information manually generated by people. Although there are several solutions for Content-Based Image Retrieval (CBIR), they cannot be applied to vector drawings, because these are described in a structural manner, requiring different approaches from those used for raster images.

In our prior work [4], we proposed an automatic visual classification scheme based on topology and geometry, to retrieve vector drawings. Our solution takes advantage of users' visual memory and explores their ability to sketch as a query mechanism. We used a graph-based technique to describe the spatial arrangement of drawing components, coding topological relationships of inclusion and adjacency through the specification of links between nodes of the graph. Additionally, we used a multidimensional indexing method that efficiently supports large sets of drawings, in combination with new schemes that allow us to hierarchically describe drawings and subparts of drawings by level of detail. This way we are able to perform searches using coarse approximations or parts of the desired drawing.

In this paper, we propose and evaluate a new mechanism to describe the spatial arrangement of elements in a drawing, which takes into account their proximity. To validate this we developed a prototype for the retrieval of clip art drawings, in SVG format (see Figure 1). The prototype allows the search of drawings using sketches, keywords and query by example. Experimental evaluation with users showed that the inclusion of information about proximity in the topology graph increases the precision of our system.

The rest of the paper is organized as follows: Section 2 provides a summary of related work in content-based retrieval of drawings. In section 3 we present an overview of our system architecture. Section 4 describes how we code proximity into our topology graph. In Section 5, we describe the prototype and the experimental evaluation, comparing the solutions with and without proximity. Finally, in section 6 we conclude and enumerate some future work.

2. Related Work

In the past years there has been a great focus in querying Multimedia databases by content. However, most such work has focused on image databases disregarding the retrieval of vector drawings. Due to their structure these require different approaches from image-based methods, which resort to color and texture as main features to describe content. In the next paragraphs we describe some approaches for content-based retrieval of drawings.

One of the first works dedicated to the retrieval of drawings was Gross's Electronic Cocktail Napkin [7]. This system addressed a visual retrieval scheme based on diagrams, to indexing databases of architectural drawings. Users draw sketches of buildings, which are compared with annotations (diagrams), stored in a database and manually produced by users. Even though this system works well for small sets of drawings, the lack of automatic indexation and classification makes it difficult to scale the approach to real collections of drawings.

In the work of Beretti and Del Bimbo [1] shapes from a drawing are decomposed into tokens that correspond to protrusions of the curve. To compute the similarity between shapes, authors verify if the two shapes share tokens with similar curvature and orientation, within a given threshold. However, the efficiency of the similarity computation depends on the number of tokens in each shape and does not take into account the token order.

Leung and Chen proposed a sketch retrieval method [10] for general free-form hand-drawings stored as multiple strokes. They use shape information from each stroke exploiting the geometric relationship between multiple strokes for matching. Later on, authors improved their system by considering spatial relationships between strokes [11]. Authors use a graph based description, similar to ours, but describing only inclusion, while we also describe adjacency. Their technique has two drawbacks, complexity, since they use a restricted number of basic shapes (circle, line and polygon) and scalability.

Another approach for matching hand-drawn sketches is the line-based representation of sketches proposed by Namboodiri and Jain [13]. In order to skirt around the problem of identifying basic shapes from a sketch, a drawing is represented as a set of straight lines, which is very dependent of the way users draw sketches.

Liang et al. [12] developed a solution for drawing retrieval based on our prior solution [4]. Authors included some differences, such as the use of eight topological relationships and relevance feedback. Additionally, they segment sketches using vertices, drawing speed and curvature. By using eight topological relationships, the description and comparison will be more restrictive, producing less results.

Pu and Ramani, developed two methods to describe drawings as a whole [9]. One uses the 2.5D spherical harmonics to convert a 2D drawing into a 3D representation, which is independent to rotations. The other method, the 2D shape histogram, creates a signature with the shape distribution, by computing values for points in the surface of the shape. This method is independent of transformations, insensible to noise, simple and fast. After experimental evaluation, authors decided to combine both methods to get a better descriptor and to increase the system accuracy.

Recently Hou and Ramani [8] presented an approach for contour shape matching of engineering drawings, inspired by the divide and conquer paradigm. They divide the original shape into two levels of representation, a higher level with structure and a lower level with geometry. During matching, they first use the structure level and then the geometry level, to find similar shapes.

From the content-based retrieval systems described above we can observe two things: most published works rely mainly on the geometric description of drawings (mainly contours), discarding the spatial arrangement of drawing items. Second, those who use topology to describe the content of drawings do not explore the proximity between drawing elements, to get more precise results.

3. Overview of the System

The new algorithm developed to code proximity between items in a drawing was integrated in our general framework for sketch-based retrieval of drawings, developed previously [4]. To give context to the reader and to explain some of the topics needed to describe our new proximity mechanism, we shortly present an overview of the overall framework, describing its main components.

Our framework allows the classification, indexing and retrieval of complex vector drawings, such as CAD drawings or clip art drawings. To that end, it uses spatial relationships, geometric information and indexing mechanisms, as illustrated in the architecture on Figure 2.

3.1. Classification

In the context of vector drawings, features such as color and texture, used mainly in the domain of digital images, are not very expressive. Instead, features related to the shape of objects (geometry) and to their spatial arrangement (topology) are more descriptive of drawing contents. So, in our framework we focus on topology and geometry as main features.

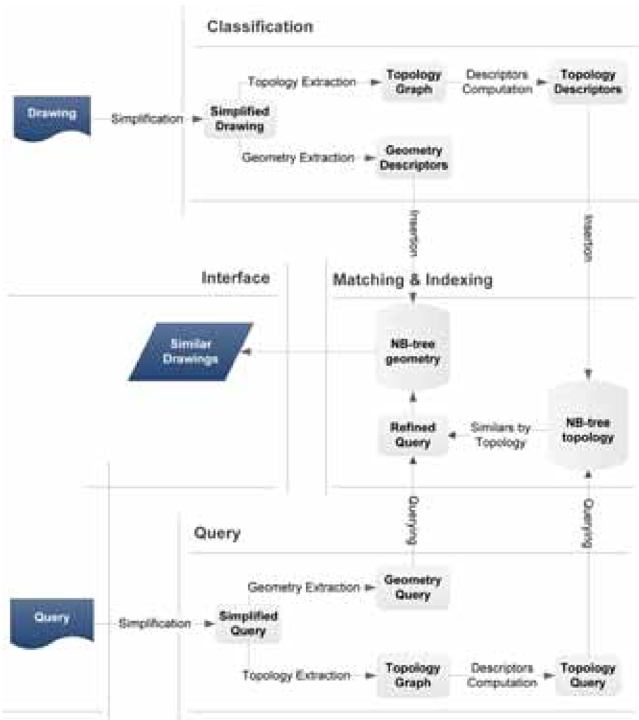


Figure 2. Framework architecture.

Our classification process starts by applying a simplification step, to eliminate most useless elements. The majority of drawings contain many details, which are not necessary for a visual query and increase the cost of searching. We try to remove visual details (i.e. small-scale features) while retaining the perceptually dominant elements and shapes in a drawing. This way we reduce the number of entities to analyze in subsequent steps of the classification process, speeding up queries.

After simplification we identify visual elements, namely polygons and lines, and extract geometric and topological information from drawings. We use two relationships, **Inclusion** and **Adjacency**, which are a simplified subset of the topological relationships defined by Egenhofer [3]. Relationships thus extracted are compiled in a **Topology Graph**, where "parent" edges mean Inclusion and "sibling" connections mean Adjacency, as illustrated in Figure 3. While these relationships are weakly discriminating, they do not change with rotation and translation.

Since graph matching is a NP-complete problem, we are not directly using topology graphs for searching similar drawings. We use the corresponding graph spectra instead. For each topology graph to be indexed in a database we compute descriptors based on its spectrum [2]. In this way, we reduce the problem of isomorphism between topology graphs to computing distances between descriptors. To support partial drawing matches, we also compute descriptors for sub-graphs of the main graph. Moreover, we use a new way to describe drawings hierarchically, by dividing them in different levels of detail and then computing descriptors at each level [4]. This combination of sub-graph descriptors and levels of detail, provides a powerful way to describe and search both for drawings or sub-parts of drawings.

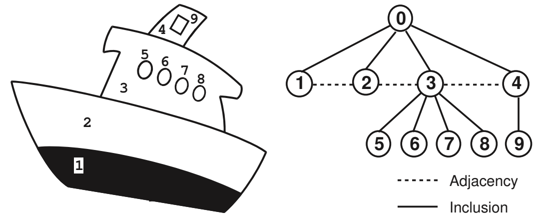


Figure 3. Drawing and Topology graph.

To acquire geometric information about drawings we use a general shape recognition library called CALI [6]. This enables us to use either drawing data or sketches as input. We obtain a complete description of geometry in a drawing, by applying this method to each geometric entity of the figure. The geometry and topology descriptors thus computed are inserted into two different indexing structures, one for topological information and another for geometric information, respectively.

3.2 Query and Matching

Our system includes a Calligraphic Interface to support the specification of hand-sketched queries, to supplement and overcoming limitations of conventional textual methods. The query component performs the same steps as the classification process, namely simplification, topological and geometric feature extraction, topology graph creation and descriptor computation. This symmetrical approach is unique to our method. In an elegant fashion two types of information (vector drawings + sketches) are processed by the same pipeline.

To improve the searching performance while using large databases of drawings, we included a multidimensional indexing structure in our framework. This indexing structure, the NB-Tree [5], is a simple, yet efficient indexing structure, which uses dimension reduction. It maps multidimensional points to a 1D line by computing their Euclidean Norm. In a second step points are sorted using a B+-Tree on which all subsequent operations are performed.

Computing the similarity between a hand-sketched query and all drawings in a database can entail prohibitive costs especially when we consider large sets of drawings. To speed up searching, we divide our matching scheme in a two-step procedure. First, we select a set of drawings topologically similar to the query, then we use geometric information to further refine the set of candidates.

4. Topological Proximity

In our previous solution we converted spatial relationships (inclusion and adjacency), between visual elements in a drawing, into a topology graph as illustrated in Figure 3. This graph has a well defined structure, being very similar to ”a rooted tree with side connections”. It has always a root node, representing the whole drawing. Sons from the root represent the dominant blocks (polygons) from the drawing, i.e. blocks that are not contained in any other block. The next level of the graph describes polygons contained by the blocks identified before. This process is applied recursively until we get the complete hierarchy of blocks. As a conclusion, we can say that each graph level adds more drawing details. So, by going down in the depth of the graph, we are ”zooming in” in drawing details.

To skirt the problem of graph isomorphism, we use the graph spectra to convert graphs into feature vectors. This way, we reduce the problem of isomorphism between topology graphs to the more simple computation of distances between descriptors.

To generate the graph spectrum we first create the adjacency matrix of the graph, second we calculate its eigenvalues and finally we sort the absolute values to obtain the topology descriptor (see Figure 4). The resulting descriptor is a multidimensional vector, whose size depends on graph (and corresponding drawing) complexity. Very complex drawings will yield descriptors with higher dimensions, while simple drawings will result in descriptors with lower size.

We assume that our topology graphs are undirected graphs, yielding symmetric adjacency matrices and assuring that eigenvalues are always real. Furthermore, by computing the absolute value and sorting it decreasingly, we exploit the fact that the largest eigenvalues are more informative about the graph structure. Additionally, the largest eigenvalues are stable under minor perturbation of the graph structure [2], making the topological descriptors also stable.

Although, isomorphic graphs have the same spectrum, two graphs with the same spectrum need not be isomorphic. More than one graph can have the same spectrum, which gives rise to collisions similar to these in hashing schemes.

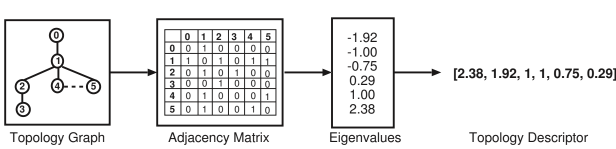


Figure 4. Block diagram for topology descriptor computation.

However, from experiences performed with 100,000 randomly generated graphs versus a set of 10 candidate similar graphs, we have observed that collisions with descriptors of very different graphs still allow us to retrieve the most likely graphs reliably.

While this solution produced good results in the past, we notice that in some cases results could be improved if we take into account the distance between the visual elements in a drawing. To that end we devised a new mechanism to include proximity into our topology graph. Our goal is to be able to differentiate between a drawing with two polygons which are close together and a drawing with two polygons that are far apart, as illustrated in Figure 5.

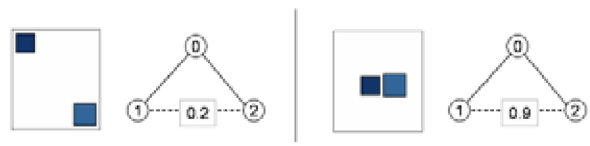


Figure 5. Using the adjacency weight to differentiate between far and near objects.

To code proximity in the topology graph, we associate weights to the adjacency links of the graph. While in our previous solution we only have an adjacency link when two primitives are connected, now we compute the (normalized) distance between two elements and use this value as the weight of the link. This change in the weights of the topology graph does not affect the stability and robustness of eigenvalues, as ascertained by Sarkar and Boyer [14].

5. Experimental Evaluation

We developed a search engine prototype for vector drawings, using our sketch-based retrieval framework and the new mechanism to describe topological proximity. The database of the system was filled with a set of SVG clip art drawings and experimental evaluation with users was carried out to compare the accuracy of the new algorithm against the previous one.

5.1 Indagare - The Drawing Search Engine

Our drawing search engine prototype, called Indagare (see Figure 1), supports the retrieval of SVG clip art drawings, using sketches, an existing SVG drawing or keywords as queries. This prototype integrates all the functionalities provided by the framework, namely, simplification mechanisms, an indexing structure to optimize the search, geometric description of visual elements and the new developed algorithm to take advantage of proximity.

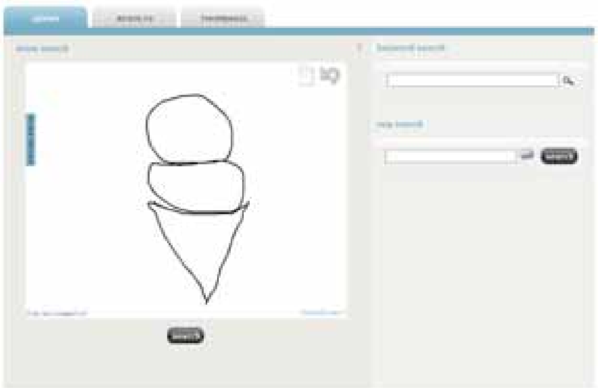


Figure 6. Sketch to search for an ice-cream.

Figure 7. Returned results for the query sketched in Figure 6.

Figure 6 shows the sketch of a query, while Figure 7 presents the results returned by the implied query. If the user wants, he can submit an existing drawing in SVG format or search by keywords (input fields on top right of Figure 6). Moreover, users can also select one of the results and use it to perform Query-By-Example.

5.2 Experimental Results

To evaluate our new approach of coding proximity into the topology graph, we carried out an experiment with ten users. Six of them were male and four were female, with ages between 18 and 58 years old. None of them had previous experience with tablet devices or any other pen-based system.

Our data set of clip art drawings was composed of 20 categories of five drawings each, selected from the OpenClipart library, yielding a total of 100 vector drawings.

Tests were conducted in two steps. First, we collected the queries by asking each user to draw three sketches, using a digitizing tablet: a balloon, a car and a house. Afterwards, and only at this time, we show all the 100 drawings in the database and requested them to identify the drawings that they considered most similar to each of the sketches they drew, without taking into account their semantic value.

The second step was carried without users' intervention. From the similar drawings selected by the participants, we identified the five more voted, and considered those as the "correct" results for each sketch. Then, we submitted the three sketched queries from each participant to the system and collected the returned results. We configured the system to retrieve 30 results for each query. With these results we computed precision and recall values.

In this experimental test, we evaluated four different system configurations. Besides testing the use of proximity we also evaluated the order in which we perform the matching steps. Typically, our framework performs first a comparison by topology and then compares the geometry of those topologically similar. Here in these tests, we also tested the other possibility, first a comparison by geometry and then by topology. The goal was to check which feature produces best results as a first filter, geometry or topology.

In summary, we tested the following configurations: i) topology plus geometry; ii) topology with proximity plus geometry; iii) geometry plus topology; and iv) geometry plus topology with proximity. To evaluate the quality of the retrieved results, we calculated precision & recall levels for each configuration, using the 11-Point Interpolated Average Precision method. Precision is the fraction of retrieved drawings that were relevant, while recall is the fraction of relevant drawings that were retrieved. The mean precision for each recall value, of the four configurations is presented in Table 1.

Table 1. Precision of the four configurations.

Recall	Conf. i)	Conf. ii)	Conf. iii)	Conf. iv)
0	0,123	0,130	0,096	0,099
0.1	0,123	0,130	0,096	0,099
0.2	0,123	0,130	0,096	0,099
0.3	0,109	0,116	0,082	0,084
0.4	0,109	0,116	0,082	0,084
0.5	0,039	0,057	0,043	0,045
0.6	0,039	0,057	0,043	0,045
0.7	0,000	0,004	0,005	0,005
0.8	0,000	0,004	0,005	0,005
0.9	0,000	0,000	0,000	0,000
1	0,000	0,000	0,000	0,000

The first thing that we can observe from Table 1 is that filtering firstly by topology yields better results than by geometry. Second, by introducing the proximity notion in the topology graph we can improve precision in both configurations (Topology filtering and Geometry filtering). However, with geometry filtering we only achieve a 0,1% increase, while in the topology filtering the improvement reaches one percent.

The small improvement in the Geometry filtering configuration was foreseeable, because the adjacency weights only play a relevant role in the topology refinement. Therefore, if the geometry filtering retrieves poor results, there is not much that the adjacency weights can do.

6. Conclusions and Future Work

In this paper, we propose a new way to describe the spatial arrangements of visual elements in a drawing. We included the notion of proximity and coded it in the topology graph through the use of adjacency weights. This new algorithm was integrated in our generic framework for sketch-based retrieval of drawings, which recast the general drawing matching problem as an instance of graph matching using vector descriptors. Topology graphs, which describe adjacency and containment relations, are transformed into descriptor vectors, using spectral information from graphs.

The use of proximity to describe the spatial arrangement gets our matching algorithm closer to the human perception, and therefore improving the retrieval effectiveness of our system. This improvement was validate through experimental evaluation with users.

Despite the complete spatial characterization of drawings provided by the use of topological relationships and proximity, the improvement achieved was small (only 1%). These results confirm informal conclusions achieved previously. Clip art drawings, contrarily to technical drawings, are more geometric than topological. Moreover, during tests with users we observed that users typically draw a very small number of shapes, and consequently do not specify topology, but only geometry.

So, improvements in the topological algorithm will produce a small impact in the final results. Additionally, our experimental tests showed that the geometric filtering needs to be improved. To overcome this we are currently developing a new algorithm to compare the geometry between drawings. Informal tests with a preliminary version revealed significant improvements in the precision values, which make us believe that we will be able to achieve better retrieval results in a very near future.

Acknowledgments

This work was funded in part by the Portuguese Foundation for Science and Technology, project A-CSCW, PTDC/EIA/67589/2006.

References

[1] S. Berretti and A. Del Bimbo. Retrieval by shape similarity with perceptual distance and effective indexing. *Multimedia, IEEE Transactions*, 2(4):225–239, 2000.

[2] D. Cvetković, P. Rowlinson, and S. Simić. *Eigenspaces of Graphs*. Cambridge University Press, United Kingdom, 1997.

[3] M. J. Egenhofer and K. K. Al-Taha. Reasoning about gradual changes of topological relationships. *LNCS*, 639:196–219, 1992.

[4] M. J. Fonseca, A. Ferreira, and J. A. Jorge. Content-Based Retrieval of Technical Drawings. *International Journal of Computer Applications in Technology (IJCAT)*, 23(2–4):86–100, 2005.

[5] M. J. Fonseca and J. A. Jorge. Indexing High-Dimensional Data for Content-Based Retrieval in Large Databases. In *Proceedings of the 8th International Conference on Database Systems for Advanced Applications (DASFAA'03)*, pages 267–274, Kyoto, Japan, Mar. 2003. IEEE Computer Society Press.

[6] M. J. Fonseca, C. Pimentel, and J. A. Jorge. CALI: An On-line Scribble Recognizer for Calligraphic Interfaces. In *Proceedings of the 2002 AAAI Spring Symposium - Sketch Understanding*, pages 51–58, Palo Alto, USA, Mar. 2002.

[7] M. Gross and E. Do. Demonstrating the Electronic Cocktail Napkin: a paper-like interface for early design. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI'96)*, pages 5–6, 1996.

[8] S. Hou and K. Ramani. Structure-oriented contour representation and matching for engineering shapes. *Computer Aided Design*, 40(1):94–108, 2008.

[9] P. Jiantao and K. Ramani. On visual similarity based 2d drawing retrieval. *Computer Aided Design*, 38(3):249–259, 2006.

[10] W. H. Leung and T. Chen. User-Independent Retrieval of Free-Form Hand-Drawn Sketches. In *Proceedings of the IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP'02)*, volume 2, pages 2029–2032, Orlando, Florida, USA, May 2002. IEEE Press.

[11] W. H. Leung and T. Chen. Hierarchical Matching for Retrieval of Hand-Drawn Sketches. In *Proceedings of the IEEE International Conference on Multimedia and Exposition. (ICME'03)*, volume 2, pages 29–32, Baltimore, USA, July 2003. IEEE Press.

[12] S. Liang, Z.-X. Sun, B. Li, and G.-H. Feng. Effective sketch retrieval based on its contents. In *Proceedings of Machine Learning and Cybernetics*, volume 9, pages 5266–5272, 2005.

[13] A. M. Namboodiri and A. K. Jain. Retrieval of on-line hand-drawn sketches. In *Proceedings of the 17th International Conference on Pattern Recognition, ICPR'04, Vol. 2*, pages 642–645, Washington, DC, USA, 2004. IEEE Computer Society.

[14] S. Sarkar and K. Boyer. Quantitative Measures of Change Based on Feature Organization: Eigenvalues and Eigenvectors. Technical report, Image Analysis Research Lab, University of South Florida, 1996.

STRAT: A Sketched-Truss Recognition and Analysis Tool

Joshua M. Peschel
Hydrology Research Laboratory
Zachry Department of Civil Engineering
Texas A&M University
College Station, Texas 77843-3136
peschel@tamu.edu

Tracy A. Hammond
Sketch Recognition Laboratory
Department of Computer Science
Texas A&M University
College Station, Texas 77843-3112
hammond@cs.tamu.edu

Abstract

The statically-determinate, pin-connected truss is a basic structural element used by engineers to create larger and more complex systems. Truss analysis and design are topics that virtually all students who study engineering mechanics are required to master, many of whom may experience difficulty with initial understanding. The mathematics used to analyze truss systems typically requires lengthy hand calculations or the assistance of proprietary computer-aided design (CAD) programs. To expedite work in this domain, we propose: STRAT (Sketched-Truss Recognition and Analysis Tool), a freehand sketch recognition system for solving truss problems. The STRAT system allows users to rapidly determine all of the unknown forces in a truss, using only a hand-drawn sketch of the truss itself. The focus of this article covers the design methodology and implementation of the STRAT system. Results from a preliminary user study are also presented.

Introduction

Structural engineers rely upon simplified mathematical descriptions of physical systems for analysis and design. Whether it is for an underground tunnel, a high-rise office building, or the wing of an airplane, engineers utilize these simplified models to work with larger and more complex systems. One basic structural element used by engineers is the statically-determinate, pin-connected truss (Figure 1). Formally, a pin-connected truss can be defined as a rigid structure that consists of slender members interconnected by joints. The practical use of the truss is to create a support system for distributed forces over relatively large distances. Examples of distributed forces carried by trusses may include the weight of a roof or the pressure underneath an airplane wing. The internal configuration of a truss is that of well-defined triangular subunits, which provide strength and stability. The employment of truss systems in design necessitates an understanding of engineering mechanics, as well as the additional use of computer-aided design (CAD) programs that allow engineers to evaluate multiple solution scenarios.

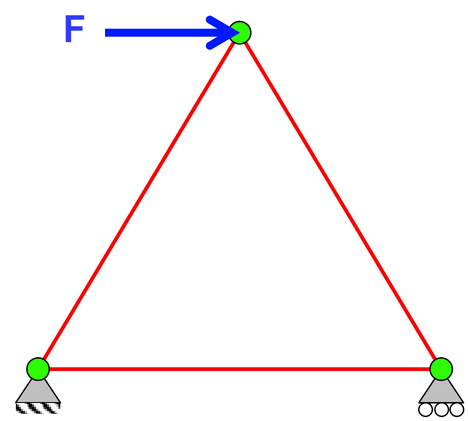


Figure 1. Example of a statically-determinate truss.

Some common CAD-based programs used for truss analysis include: Visual Analysis [1], Truss 2000 [2], and RISA-2D [3]. Each of these CAD programs, however, utilizes a standard drag-and-drop graphical user interface. To create a truss representation in these types of CAD programs, the user typically selects pre-defined truss elements, such as a joint or member, from a tool palate and then places each item on a pre-defined grid in the desired configuration. Member and joint connections are often made on-the-fly, along with force placements; properties such as member length and force magnitude are usually associated with the truss elements once the system has been manually assembled. After the assembly process is complete, a mathematical analysis for all of the unknown forces can then be performed.

In this article we propose: STRAT (Sketched-Truss Recognition and Analysis Tool), a freehand sketch recognition system for working with truss problems. Sketch recognition may be formally defined as the automated identification of shapes and objects that are drawn by freehand sketching alone. Sketch recognition is uniquely different from other drawing applications, including CAD, because it does not rely on predefined objects to create text or drawings – only freehand strokes are used as inputs. The recognition process usually consists of multiple levels, ranging from low-level recognition that may include simple lines and polygons, up through higher-level recognition usually involving more complex elements such as engineering schematics.

Research in the area of sketch-based truss analysis can be traced back to Ivan Sutherland. In his original 1963 work, Sutherland used the drawing and computational capabilities of Sketchpad to determine optimal geometric solutions for truss member lengths [4]. But this work lay mostly dormant for decades until later research efforts began to address freehand sketch recognition. In the general context of engineering mechanics, Stahovich developed an application to use freehand sketching for interacting with certain related classes of mechanical design problems [5]. LaViola and Zeleznik created a sketch-based interactive solver capable of understanding mathematical symbols and basic mechanics problems [6]. However, in both research efforts, neither system specifically addresses recognizing and solving truss problems. Hutchinson et al. used an existing freehand sketch recognition interface for finite element model (FEM) assembly [8]. Truss problems were addressed in this work but the FEM approach is an advanced technique involving additional layers of computational complexity not needed for standard truss analysis.

The primary rationale for the proposed STRAT application is that virtually all students of engineering mechanics spend a significant amount of time learning to solve truss systems 'by hand'. Therefore, freehand sketching of a truss system is a common skill learned by all of these students, well before CAD software packages are introduced to them. Learners could use the STRAT application to quickly verify answers to hand-worked problems; likewise, they could use the application to evaluate incremental solution steps. The mathematics used to analyze trusses often requires lengthy hand calculations and, working with pre-defined tool palate objects may involve a steep learning curve – much more so than freehand sketching alone. Because truss elements, e.g. members, joints, pins, etc. have reasonably well-standardized hand drawn forms in both academic and professional practice, the expected variability of freehand sketching would be quite low; thus, a simple yet reliable sketch recognition system could be developed to expedite and improve work in this domain.

The focus of this research article covers the application methodology of the proposed STRAT system, which includes primarily, recognition and analysis capabilities. Results and discussion from a preliminary user study are presented, along with the conclusions and future directions for this work.

Application Methodology

A framework for both recognizing and solving a robust set of truss problems, whether CAD or programmatically-based, necessarily involves several layers of information processing. When only freehand sketching is used as an input, additional layers must be considered. The STRAT application consists of two main information processing layers, with three sub-layers in each primary layer (Figure 2). The following sections outline the specific processing layer details.

Sketched-Truss Recognition

The initial information processing layer for STRAT is the sketched-truss recognition layer, representing the human-computer interface. Within this processing layer, user data are transformed from basic pen strokes into a recognized truss schematic. There are three sub-layers to the sketched-truss recognition layer: (1) stroke parsing, (2) low-level recognition, and (3) recognized associations.

Stroke Parsing. When a user draws, they will either form simple or complex strokes. Simple strokes in STRAT are defined as either single lines or circles – they are singular strokes that can be recognized initially as geometric primitives. Complex strokes, as defined in STRAT, are all other non-simple polyline segments. An example of a complex stroke would be a triangle or an arrow (Figure 3a). When a stroke in STRAT is recognized as complex, a corner-finding algorithm is employed to locate parsing points [8]. The parsing is done to address the stroke and drawing order variability that often occurs among users. Once the parsing sequence is complete, for a given complex stroke, each of its component sub-strokes will have been recognized as either a single line or circle. When all complex strokes have been simplified, low-level recognition can take place.

Low-Level Recognition. The low-level recognizer used in STRAT is a geometric rules-based algorithm. Simple

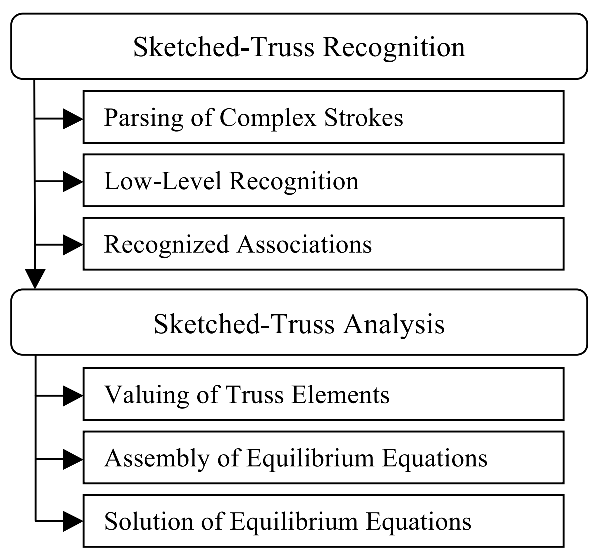


Figure 2. STRAT information processing layers.

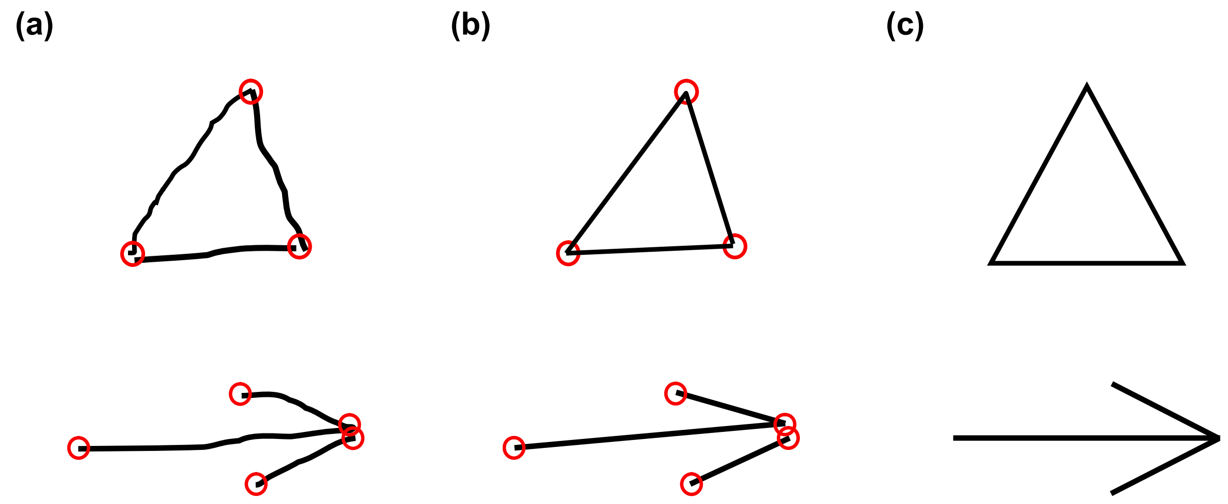


Figure 3. Sub-layer illustrations from the sketched-truss recognition layer: (a) parsing of complex pen strokes, (b) low-level processing of polylines to form geometric primitives – shown here are a triangle and an arrow being formed, and (c) recognized associations of low-level recognized primitives – shown here is a triangle being classified as a reaction support and an arrow is interpreted as a force.

strokes (lines or circles) are now put into context to form other geometric primitives, such as triangles or arrows (Figure 3b). This approach to recognition is taken for two reasons. First, by simplifying each stroke prior to recognition, the variability that may exist in comparison to correctly indentifying more complex gestures directly (such as an arrow) is likely reduced. Second, with a composite sequence of recognized strokes, the user is given the flexibility to remove one or more possibly unwanted segments without deleting the entire original segment. After all of the existing strokes are recognized, they move to the recognized association layer for assembly into a truss schematic.

Recognized Association. Geometric primitives that have been recognized possess context within a truss system. To give the appropriate context to recognized primitives, we again employ a rules-based algorithm where object types are assigned. For example, circles represent joints, arrows represent forces, etc. More complexity arises when reaction supports (small triangles) exist because larger triangles are also formed by the structural members (Figure 3c). To mitigate confusion between these two object types, our algorithm compares the relative sizes and spatial orientations using common-sense reasoning (i.e. we would not expect to find a reaction support in the center of the truss). A final step taken for recognized association is an examination of determinacy. To be statically-determinate, a two-dimensional truss system must contain zero degrees-of-freedom (DOF). The actual DOF can be numerically determined using the following expression:

$$DOF = 3(n-1) - 2j \qquad (1)$$

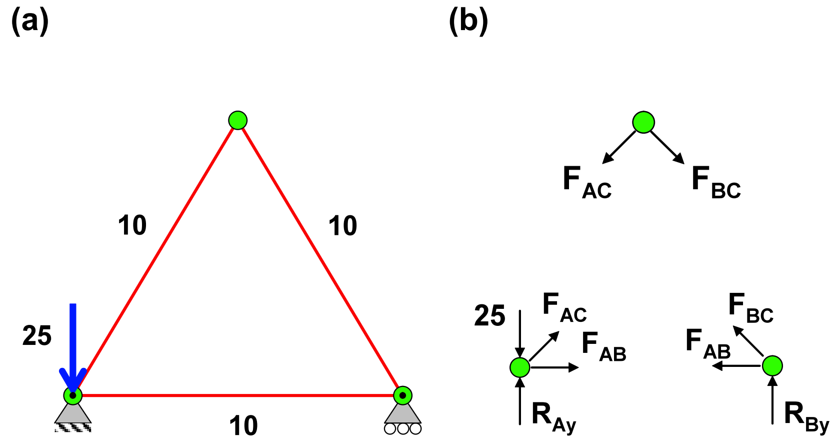
where n is the number of structural members and j is the number of joints. A non-zero result for the calculation would indicate that the truss system was statically-indeterminate. At present, the STRAT application can only evaluate statically-determinate truss systems. Once a truss schematic is correctly assembled, it moves to the next information processing layer.

Sketched-Truss Analysis

The second information processing layer for STRAT is the sketched-truss analysis layer. Here, recognized objects from the truss schematic are given numerical meaning and the unknown values desired are solved for. There are three sub-layers to the sketched-truss analysis layer: (1) element valuing, (2) assembly of equilibrium equations, and (3) solution of equilibrium equations.

Element Valuing. When truss systems are analyzed, engineers are interested in the forces that are carried by the structural members in response to applied external forces. These internal forces result in members being in either tension or compression, and they change in a predictable manner with changing external forces, as well as changing geometric configurations. To capture these configurations, numeric values must be assigned to the recognized truss elements. This is accomplished by assigning lengths to members and magnitudes to external force arrows. In the STRAT application, users may enter these values directly onto the truss schematic (Figure 4a). When this step is complete, equation assembly can begin.

Equation Assembly. All truss systems can be solved using an assemblage of static equilibrium equations. To



(c)

$$\begin{pmatrix} 1 & 0 & 0 & 0.5 & 1 & 0 \\ 0 & 1 & 0 & 0.86 & 0 & 0 \\ 0 & 0 & 0 & 0 & -0.5 & -1 \\ 0 & 0 & 1 & 0 & 0.86 & 0 \\ 0 & 0 & 0 & -0.5 & 0.5 & 0 \\ 0 & 0 & 0 & -0.86 & -0.86 & 0 \end{pmatrix} \begin{bmatrix} R_{Ax} \\ R_{Ay} \\ R_{By} \\ F_{AC} \\ F_{BC} \\ F_{AB} \end{bmatrix} = \begin{pmatrix} 0 \\ 25 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Figure 4. Sub-layer illustrations from the sketched-truss analysis layer: (a) valuing of recognized truss elements, (b) assembly of static equilibrium equations into a matrix format, and (c) solving of assembled static equilibrium equations.

determine the unknown forces in the members, a system of linear equations is developed. In the STRAT application, the 'method of joints' is used to create this system of equations. The method of joints technique simply examines each truss joint, and creates two separate static equilibrium equations that are the summation of x- and y-direction forces, respectively (Figure 4b). The standard positive x- and y-direction orientation for forces is assumed. Forces on angles are decomposed into x- and y-components using their relationship to the assumed positive x-axis; decomposed forces are recombined after the solution process. Unknown reaction forces are always assumed to act into the joints and unknown member forces are always assumed to act away from the joints. Positive or negative numeric values obtained from solving the system of equations determines the final force orientation. Once all of the static equilibrium equations are determined for each joint, they are assembled into an array for solving.

Equation Solving. When the system of linear equations is assembled for solving, it takes the form Ax = b, where A is the numeric array of truss element coefficient values, x is the vector of force identities, and b is the vector populated with external force values, if any are present at the individual joint (Figure 4c). To solve this expression, we simply employ an algorithm that inverts the array A and multiplies it by b (i.e. $x = A^{-1}b$). Determined values are then associated with all unknown forces. Sign also plays an important role in the solution process. A positive numeric value for an unknown reaction force indicates that its orientation is into the joint; a negative value would indicate that the reaction force acts away from the joint. Likewise, a positive member force implies that the member is in tension, while a negative value means that the truss member is in compression. All of the unknown forces solved for are written to a table of values displayed next to the truss schematic.

User Study Results

To evaluate the recognition and analysis capabilities of the STRAT application, a preliminary user study was performed. Ten users consisting of structural engineering faculty, staff, and students were asked to complete two separate sketch recognition application tasks on a Tablet PC (Figure 5). The user study initially began with a warm-up for each participant that lasted approximately two minutes. This was to familiarize each user study participant with the Tablet PC; brief explanations of the sketch recognition tasks were given and users were allowed to experiment with drawing basic shapes before the study began. The first part of the user study was to draw ten geometric primitives typically found in truss systems (e.g. lines, triangles, circles, etc.) (Figure 5a). In this task, users were instructed to reproduce the geometric primitive they were shown on the screen. Each user submitted three samples for each geometric primitive; therefore, a total of 300 samples were taken for evaluation. The second part of the user study involved drawing complete truss systems (Figure 5b). Two separate trusses were shown on-screen and each user was asked to submit two samples for each truss, resulting in a total of 40 samples being taken for evaluation. For the second part of the study, users were instructed to reproduce the shown truss in the manner that they would normally draw it. This was done to capture the variability among users who may draw certain aspects of the truss system with subtle differences, such as reaction supports

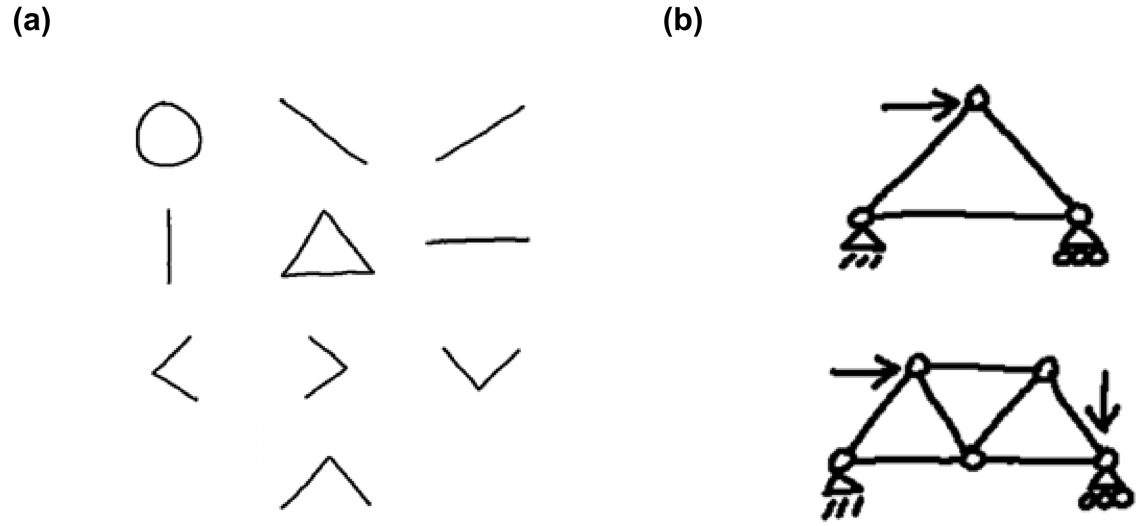


Figure 5. The two separate tasks assigned in the user study: (a) sketch recognition task I and (b) sketch recognition task II.

and force arrows. The following section discusses the results obtained from the preliminary user study.

Discussion

The first part of the preliminary user study resulted in a 100-percent recognition rate for all of the geometric primitive samples taken. This meant that all of the sample gestures were parsed correctly, then recognized, and reassembled into the correct recognized associations. This outcome was somewhat expected because stroke inputs were generally simple and clean, without much complex stroke parsing involved. This part of the user study was done to evaluate the low-level recognition capabilities for the STRAT application. Results from the second part of the user study were considered on an all-or-none basis, meaning that either every element of the truss system was correctly identified, or the truss system as a whole was not recognized. An overall accuracy of 70-percent was obtained from the samples in the second part of the user study. This lower level of accuracy can likely be attributed to one of the following three causes, following the examination of the original input data (Figure 6). First, some users when instructed to draw the truss system how they would normally do so provided reaction supports in the form of circles (Figure 6a). The consequence to this action, along with the overlapping location of the joints resulted in not all reaction supports being recognized for these truss sketches. Therefore, the truss system was not fully recognized. The second issue observed from the sample data was excessive overtracing and scribbling (Figure 6b). In the context of drawing a truss, scribbling to darken in the joint locations is not an uncommon occurrence. However, in the context of most sketch recognition applications, scribbling over a gesture means that the gesture should be deleted. Thus, when users performed this action, the recognized joints would be erased. The final occurrence observed that is believed to have contributed to lower accuracy was extended strokes (Figure 6c). An example of this would be when a user sketched the collective bottom chords of the truss in one stroke. The stroke itself was recognized as a simple line and therefore not parsed into additional members. Although, other strokes were connected to it at interior locations and even if the user indicated that joints were at these locations, the long stroke was not segmented.

Evaluation of the second information processing layer, the sketched-truss analysis, was not included in the original collection of user study data. We have tested the solver by manually adding data to the recognized associations and performance has been at 100-percent. There are obvious consequences to not fully correctly recognizing the truss system; chief among these is an inability to continue to the solver stage or the returning of an incorrect result. Therefore, we have focused primarily to date on the sketched-truss recognition portion of the STRAT application.

Future Work

There are several directions for the continuance of this work. Most pressing among these directions is improving the overall low-level recognition (and recognized association) accuracy of the STRAT application. We see this going in one of two possible ways. First, we add and evaluate additional rules to the low-level and recognized association layers. The only issue in doing so would be overcoming existing sketch recognition dogma, such as using scribbles to delete strokes. The alternative is to employ an additional object-based interface, such as

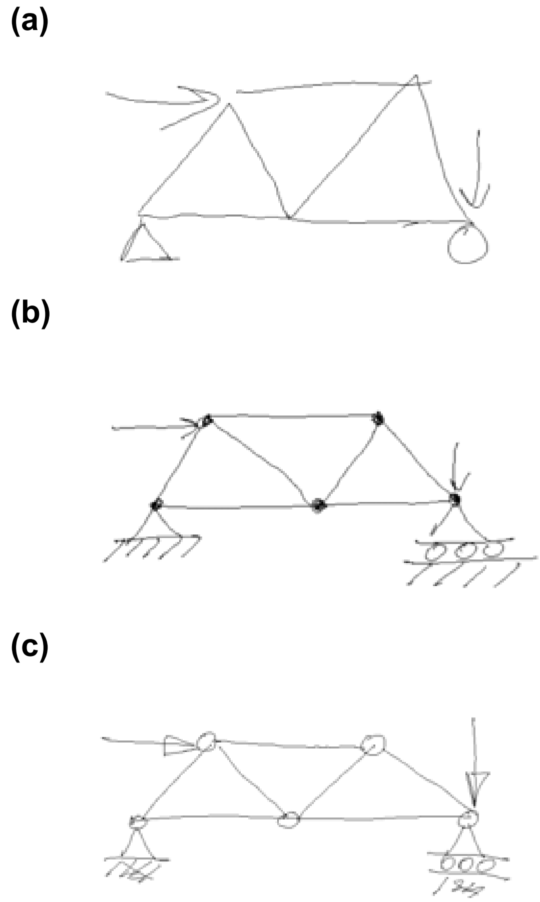


Figure 6. Examples of truss sketching variability.

LADDER, for the low-level recognition and recognized associations [9]. LADDER does not provide the same degree of stroke segmentation as STRAT, but that could be considered a trade off in favor of improved accuracy. Also, the solver component has worked well manually but an additional user study that takes into account element valuing also needs to be performed, as well as an overall usability study. Additional areas of advancement for the STRAT application include extending the recognition and analysis to statically-indeterminate trusses. Statically-determinate trusses only rely upon equilibrium equations for their solution, while statically-indeterminate truss systems require additional constitutive equations, such as those based on material properties.

Conclusions

This paper presented the application methodology and implementation of STRAT, a sketched-truss recognition and analysis tool. A preliminary user study was conducted to evaluate the recognition component of the application. Basic geometric primitives were all perfectly recognized when individually input; complete truss systems were recognized at an all-or-none accuracy of 70-percent. Work on the STRAT application is being continued and techniques to improve the overall accuracy are being investigated.

Acknowledgments

Joshua Peschel would like to recognize the Zachry Department of Civil Engineering for its financial support through the Tommy E. Knight '61 Endowed Fellowship. This work was funded by NSF IIS grant 0744150: Developing Perception-based Geometric Primitive-shape and Constraint Recognizers to Empower Instructors to Build Sketch Systems in the Classroom.

References

[1] Visual Analysis 6.0 Users's Guide, Available at: http://www.iesweb.com/support/, Last accessed: May 1, 2008.

[2] Truss 2000: Analysis and Design Spatial Trusses User's Guide, Available at: http://www.newsoft-eng.it/Truss_eng.htm, Last accessed: May 1, 2008.

[3] Risa-2D Version 8.0 User's Guide, Available at: http://www.risatech.com/risa-2d.asp, Last accessed: May 1, 2008.

[4] I.E. Sutherland, "Sketchpad: a man-machine graphical communication system," In *25 Years of DAC: Papers on Twenty-Five Years of Electronic Design Automation*, pp. 507–524, 1988.

[5] T.F. Stahovich, "SketchIT: a sketch interpretation tool for conceptual mechanical design," Massachusetts Institute of Technology Artificial Intelligence Laboratory Technical Report No. 1573, March 2002.

[6] J.J. LaViola and R.C. Zeleznik, "MathPad^2: a system for the creation and exploration of mathematical sketches," In *ACM SIGGRAPH '07 Courses*, August 2007.

[7] T.C. Hutchinson, R. Kuester, and M.E. Phair, "Sketching finite-element models within a unified two-dimensional framework," *J .of Computing in Civil Engr*, vol. 21, no, 3, May 2007.

[8] J.M. Peschel and T.A. Hammond, "A higher-order kinematics approach to digital stroke segmentation," unpublished.

[9] T. Hammond and R. Davis, "LADDER: A Sketching Language for User Interface Developers*," Computers & Graphics*, vol. 29, no. 4, pp. 518–532, 2005.

On the Semantic Labeling of Hand-drawn Sketches

Gennaro Costagliola[1] and Alberto Greco[2]
[1]*Dipartimento di Matematica ed Informatica – Università di Salerno*
[2]*Dipartimento di Scienze Antropologiche – Università di Genova*
E-mail: gcostagliola@unisa.it, greco@unige.it

Abstract

We propose a preliminary study for the application of Computational Linguistics techniques such as the Semantic Role Labeling of natural languages (SRL) to the case of sketch understanding. The idea of our approach originates from a psychological experiment where two groups of students were asked to produce and then recognize hand-drawn sketches, respectively. The paper presents a procedure to apply Semantic Role Labeling to the case of hand-drawn sketches and provides an example of application. This study seeks to enrich the body of research related to sketch understanding and reasoning.

1. Introduction

In any application based on hand drawn sketches it is very important to evaluate how much concepts and tasks are appropriately described in a pictorial form.

Knowing the semantic roles played by the components that appear in a hand-drawn sketch is of major importance for understanding its underlying meaning and for further reasoning. The inherent picture ambiguity and the way its constituents are spatially related can lead to very different interpretations of the same sketch.

Research on sketch understanding has taken many directions in the past decade and mostly it has tended to concentrate on the interpretations of the strokes forming the sketches [4]. In this paper we present our main ideas on how computational linguistics techniques can be used to address the problem of semantic sketch understanding. This is achieved by manually tagging constituents of sketch sentences with corresponding constituents of natural language sentences. We believe that the proposed approach may be a good starting point to be used when considering verbal descriptions in multimodal interactive interfaces, especially when meanings of speech and drawing have to be merged [1].

In particular we are interested in the use of semantic role labeling [7], one of the most promising approaches lately developed for natural language processing with interesting applications in the field of information extraction, question-answering, natural language database querying, spoken dialogue systems, machine translation, story merging, and others. While some of these applications are common for hand-drawn sketches, such as for query-by-sketch, [9], the others can provide new interests for the use of sketches.

In the following we recall the main concepts of Semantic Role Labelling as used in Computational Linguistics, describe the psychological experiment that gave start to this research, give a procedure for semantically labeling a hand drawn sketch from a verbal description of it and give some definitions of *consistent* hand-drawn sketches with an example showing the use of semantics hand-drawn sketch labeling.

2. Semantic Role Labeling

Let us recall the main characteristics of SRL. "Semantic Role Labeling involves the determination of domain-independent semantic relations among the entities and the events they participate in. Given a sentence, one formulation of the task consists of detecting basic event structures such as "who" did "what" to "whom", "when" and "where". Recently, the compilation and manual annotation with semantic roles of medium-large corpora – the PropBank [12, 13], NomBank, and FrameNet [5, 6] initiatives – has enabled the development of statistical approaches specifically for the task of semantic role labeling. SRL, especially focused on the labeling of verbal arguments and adjuncts, has become a well-defined task with a substantial body of work and comparative evaluation (e.g., see [7], CoNLL Shared Task in 2004 and 2005, Senseval-3)", [15].

As an example of SRL let us consider the following sentences referring to the frames of Trading and Judgement, respectively:

[$_{\text{temporal}}$ At the end of the day] , [$_{\text{things being traded}}$ 251.2 million shares] **were traded** . (TRADING)

[$_{Judge}$ She] **blames** [$_{Evaluee}$ the Government] [$_{Reason}$ for failing to do enough to help]. (JUDGEMENT)

The identification of such event frames will have an important impact in many Natural Language Processing applications such as Information Extraction [14], Question Answering [11], Machine Translation [2], as well as Story Merging [10]. "Although the use of SRL systems in real-world applications has so far been limited, the outlook is promising over the next several years for a spread of this type of analysis to a range of applications requiring some level of semantic interpretation." [15]. In our opinion these ones will naturally include sketch understanding.

There are many algorithms for semantic shallow parsing to assign semantic roles. This paper will specifically refer to the online UIUC parser [17]. The parser not only provides semantic role labeling of the sentence parts but also outputs a syntactic parse tree according to the Charniak parser, [3], allowing us to recover as much information as possible on each part of the sentence.

3. Experimental setup

The main goal in designing the experiment was to empirically identify correspondences between semantic features and pictorial elements and to explore a method for analyzing efficacy: operationally defined as the difficulty in matching pictures with original sentences, [8].

The experiment was divided into two stages and utilized two separate groups of participants – one for each stage of the experiment. In the first stage, the participants acted as "subjects" whilst in the second stage the participants acted as "judges".

In the first stage, participants were asked to represent graphically a series of sentences describing a number of different types of situations. Sentences could describe either mathematical or non-mathematical situations. There were spatial situations (relations such as "above", "nearby", etc.), set situations like in arithmetic situations (where a set is divided into parts or subsets), and time situations (where events occur at different stages in time). Examples of sentences are:

- The house where Alan, Burt, and Chris live.
- Burt had 15 books. He bought 8 more, now he has 23 books.
- In Alan's garden there are 50 trees. Burt has more trees than Alan.

In the second stage, other participants acted as judges; their task was to match sentences (in random order) with the pictures drawn by the subjects in stage 1. The aim of this procedure was to evaluate the appropriateness of the pictures to their intended purpose, i.e. communicating information to other people. Participants in stage 2 were not taught explicit procedures, in order to encourage them to develop their own implicit procedures as they encountered more and more complex and abstract situations.

From the experiment, it has been observed that in many cases it is possible to associate parts of the drawings with parts of the originating verbal descriptions. In particular, there is evidence that the expressiveness of the drawings depends on how "well" the parts of the drawing represent the corresponding parts of the verbal descriptions and how the spatial relations (implicit ones such as nearby, above, etc. and explicit ones such as arrow, link, etc.) between the sketched parts recall the properties described in the verbal description such as possession, location, etc.. Furthermore, the appropriateness of the pictures to their intended purpose appears to be enhanced when consistent use of visual and spatial relation descriptions are used. These considerations allow us to state that it is appropriate to label parts and relations in a hand-drawn sketch with semantic roles in order to enhance sketch understanding and reasoning.

4. Semantic Sketch Role Labeling

In order to define the semantics of sketches we need to start creating a corpus of annotated sketches similarly to what it has been done with PropBank and FrameNet [12, 13, 6]. To do so we propose a way of semantically labeling sketches as shown by the procedure in Figure 1. Here each action is represented by a circled number indicating the temporal sequence of execution. The input sentence is partitioned and its parts are semantically annotated through the SRL technique (action 1 in Figure 1). Parts of the hand-drawn sketch are then labeled with the resulting sub sentences (action 2). Finally, the semantically labeled sketch is derived by merging results from the previous actions (action 3).

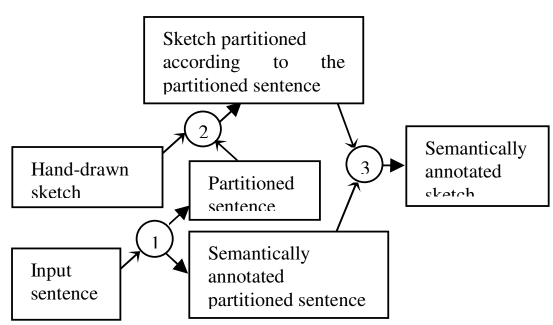


Figure 1. The Sketch Semantic Role Labeling procedure

Note that the procedure guarantees that the corresponding parts of the sketch and of the sentence are properly aligned.

The parts of the hand-drawn sketch to be labeled must be chosen accordingly. In fact, a hand-drawn sketch is a visual sentence and as such it is composed of a set of (spatially) related visual symbols or constituents [16]. An important task in the user sketch-sentence association is then the identification of the visual symbols and of the (spatial) relations to be labeled. It must be noted that, as in [16], relations in a picture can be either *implicit* such as the spatial relations nearby, below, above, etc. or *explicit* such as arrows, links, etc., i.e., relations with a visual representation.

As an example let us consider one of the sentences of the experiment in Section 3:

In Alan's garden there are 50 trees.
Burt has more trees than Alan

Figure 2 shows the hand-drawn sketch produced by a participant of stage 1 graphically depicting the sentence. In this figure, Alan and Burt are depicted as letters A and B, respectively, the garden as a waved line with flowers underneath, the two sets of trees as two tree shaped figures with the multiplication sign and a number to indicate quantity.

Table 1 provides a potential manual labeling of semantic roles for the picture in Figure 2 resulting from the procedure proposed in Figure 1.

In particular, each row contains a part of the hand-drawn sketch (column 2) drawn by the participant, the substring of the sentence with which the sketch has been manually annotated (column 1) and the annotations on the sub-sentence resulting from the Charniak parser and the Prop-Bank corpus (columns 3 and 4, respectively). While the association between the substrings and the sketch parts have been done manually, the Charniak parser and Prop-Bank annotations are done automatically by the UIUC parser, [17], based on the SRL techniques recalled in Section 2.

5. An SSRL Application

In this application we are interested in checking the semantic consistency of a hand-drawn sketch with respect to a pre-built corpus based on information coded as in Table 1. In the following, by "checking the consistency" we mean checking if the drawing style is compatible with one of the drawing styles coded in the corpus.

We distinguish three types of drawing style. Type 1 indicates the same use of spatial relations to describe

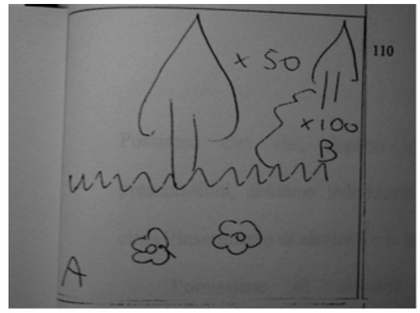
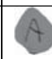
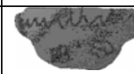
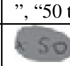
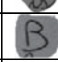
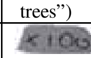


Figure 2. Participant 2 hand-drawn sketch

Table 1. Sketch Semantic Role Labeling for Fig. 2

Sentence	**Participant sketch**	**Charniak parser annotations**	**PROP-BANK annot.**
In			location [AM-LOC]
Alan		NNP (Proper noun, singular)	
's	nearby("Alan", "garden")	POS (Possessive ending)	
garden		NN (Noun, singular)	
there	nearby("garden", "50 trees")		
are			V: be
50		CD (cardinal number)	patient [A1]
trees.		NNS (Noun, plural)	
Burt			owner [A0]
has	nearby("Burt", "more trees")		V: have
more		JJR (Adjective, comparative)	possession [A1]
trees		NNS (Noun, plural)	

semantic connections between constituents of a sketch; type 2 indicates that semantic constituents (visual symbols) have a similar appearance; type 3 indicates that both type 1 and 2 occur. An example is shown in Figure 3: sketches (A) and (B) use the same style only with respect to type 2, while sketches (A) and (C) use the same style only with respect to type 1.

In the sequel we will only consider type 1 drawing style; in other words, two hand-drawn sketches are similar if they have a consistent use of spatial relations.

John and Mary live in a house:

(A) (B) (C)

Figure 3. The semantic concept “live” is expressed through the spatial relations “contain” in (A) and (C), and “nearby” in (B)

As an example let us consider the following sentence: **Fred owns a house** and its semantic and syntactic labeling as produced by [17], and shown in Figure 4.

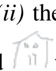
Fred	owner [A0]	
owns	V: own	
a	possession [A1]	
house		

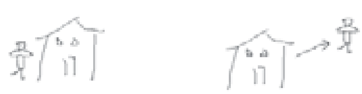
Charniak's Parse Tree

```
(S1 (S (NP (NNP Fred))
       (VP (VBZ owns)
           (NP (DT a)
               (NN house)))))
```

Figure 4. Semantic role labeling and Charniak annotation

Two possible hand-drawn sketches corresponding to the given sentence are shown in Figure 5 below.

In both cases *(i)* and *(ii)* the graphical sketch  is assigned with “Fred” and  with “house”. In case *(i)* the verb “owns” is associated to the relational fact nearby(“Fred”, “house”) or, more specifically, nearby(owner, possession). In case *(ii)*, “owns” will be associated to the relational fact arrow(owner, possession).



(i) *(ii)*

Figure 5. Hand-drawn sketches for the sentence **Fred owns a house.**

By properly matching the semantic role labeling of the sentence “Fred owns a house” and considering its visual associations against the labeling and associations of Table 1 it becomes possible to state that the hand-drawn sketch of Figure 5(i) has the same type 1 drawing style of Figure 2, while the hand-drawn sketch of Figure 5(ii) has a different drawing style.

6. Conclusions

In this paper we have presented a way for annotating hand-drawn sketches semantically. We derived its validity from the results of a psychological experiment and based its implementation on the computational linguistics technique of Semantic Role Labeling. We finally gave an example of application showing how to check for consistency between hand-drawn sketches.

7. References

[1] A. Adler and R. Davis, Speech and Sketching: An Empirical Study of Multimodal Interaction, EUROGRAPHICS Workshop on Sketch-Based Interfaces and Modeling, M. van de Panne, E. Saund (Editors) (2007)

[2] H. C. Boas. Bilingual FrameNet dictionaries for machine translation. In Proceedings of LREC 2002, 1364--1371, Las Palmas, Canary Islands. (2002).

[3] E. Charniak, Statistical parsing with a context-free grammar and word statistics, Procs of the 14th National Conference on Artificial Intelligence AAAI Press/MIT Press, Menlo Park (1997).

[4] R. Davis, Magic Paper: Sketch-Understanding Research, Computer (Vol. 40, No. 9) pp. 34-41, Sept 2007.

[5] C. J. Fillmore, C. Wooters, and C. F. Baker. Building a large lexical databank which provides deep semantics. InProceedings of the Pacific Asian Conference on Language, Information and Computation (2001).

[6] FrameNet - http://framenet.icsi.berkeley.edu

[7] Gildea, D., Jurafsky, D.: Automatic labeling of semantic roles. Computational Linguistics 28(3) pp. 245-288 (2002).

[8] Greco A. A method for a cognitive analysis of graphical presentations. Tech. Rep. 0801, Lab. of Psychol. and Cog. Sci. Univ. of Genoa, (2008).

[9] T. Hisamori, G. Ohashi, Query-by-sketch interactive image retrieval using rough sets, ISIC. IEEE International Conference on Systems, Man and Cybernetics, pp.1223-1229, Oct. 2007.

[10] F. McNeill, H. Halpin, E. Klein, and A. Bundy Merging Stories with Shallow Semantics. Knowledge Representation and Reasoning for Language Processing Workshop at the EACL conference, Genoa, Italy, (2006).

[11] S. Narayanan, S. Harabagiu. Question Answering based on Semantic Structures, International Conference on Computational Linguistics (COLING 2004), Geneva, Switzerland, August 22-29, (2004).

[12] M. Palmer, D. Gildea, and P. Kingsbury. The Proposition Bank: An annotated corpus of semantic roles. In Computational Linguistics, 31(1), (2005).

[13] PropBank - http://www.cis.upenn.edu/~mpalmer/project_pages/ACE.ht

[14] M. Surdeanu, S. Harabagiu, J. Williams, and P. Aarseth. Using predicate-argument structures for information extraction. In Procs of the ACL, Sapporo, Japan, (2003).

[15] Special Issue on Semantic Role Labeling http://www.lsi.upc.edu/~carreras/srlcl.html

[16] G. Costagliola, A. De Lucia, S. Orefice, G. Tortora, A Framework of Syntactic Models for the Implementation of Visual Languages. Procs. 13th IEEE Symposium on Visual Languages, Capri, Italia, 58-65, 1997.

[17] UIUC parser - http://l2r.cs.uiuc.edu/~cogcomp/srl-demo.php

Spider Diagrams Augmented with Constants: A Complete System

Gem Stapleton
Visual Modelling Group
University of Brighton, UK
g.e.stapleton@brighton.ac.uk

Abstract

The use of visual languages in computing is varied, ranging from system modelling to the display of data analyzed in computation processes. A prominent example of a visual notation is the Unified Modelling Language (UML), designed for use by software engineers. Constraint diagrams were proposed as an alternative to the UML's Object Constraint Language. Spider diagrams form a fragment of constraint diagrams, and their applications are more broad than just placing constraints on software models, including information visualization and displaying the results of database queries. This paper focuses on spider diagrams augmented with constants that represent specific individuals. We present a sound reasoning system for spider diagrams with constants and establish completeness. Furthermore, the technique used to prove completeness can be adapted to give rise to a decision procedure.

1 Introduction

It is widely recognized that diagrams play an important role in various areas particularly in many aspects of computing, including visualizing information and reasoning about that information. Diagrams are often useful for conveying complex information in accessible and intuitive ways. This is one reason behind the widening perception of the importance of diagrams in computing systems and more widely.

Software engineers form one group of users that need formal languages to specify and design complex systems. Ideally, their software specifications should be accessible to all stakeholders involved in the modelling process, including customers, managers and programmers. There is extensive use of diagrams to model software, with the Unified Modelling Language (UML) being an industry standard, mainly visual, notation. The only non-diagrammatic part of the UML is the Object Constraint Language (OCL) which is designed to place formal constraints on software models. It, therefore, seems sensible to offer formal diagrammatic notations for the purpose of precise, yet accessible, software specification.

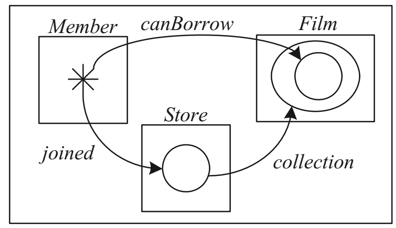


Figure 1. A constraint diagram.

Constraint diagrams were introduced in [10] as a way to visualize object-oriented invariants in the context of the UML. They have been used to develop high-level models independently of UML [8, 12]. Building on Euler and Venn diagrams, constraint diagrams contain *spiders* to indicate existential and universal quantification and use arrows to make statements about binary relations. For example, the constraint diagram in figure 1 expresses that people can borrow only books that are in the collections of libraries that they have joined. A formalization of constraint diagrams can be found in [4] and a generalized version of them is formalized in [16].

The language of spider diagrams [9] forms a fragment of the constraint diagram language. The only spiders present in spider diagrams in [9] represent the existence of elements (called *existential spiders*) and arrows are not permitted. The spider diagram d_1 in figure 2 expresses, by the inclusion of the curve $Lions$ inside $Cats$, that all lions are cats and, in addition, it expresses that there is a cat, which may or may not be a lion by the use of the existential spider (i.e. the tree). The spider diagram d_2 expresses is something which is not a dog. It has been shown that the spider diagram language is equivalent in expressive power to monadic first order logic with equality [18].

The diagrams d_4, d_5 and d_6 include constant spiders; these are labelled trees whose nodes are visualized as squares. The diagram d_4 tells us that tom is a cat; d_5 tells

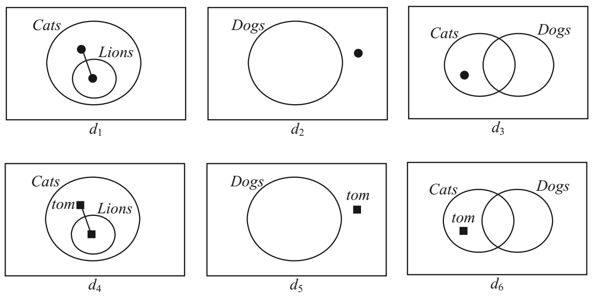


Figure 2. Spider diagrams.

us that tom is not a dog. From d_4 and d_5 we can deduce d_6 which tells us that tom is a cat but not a dog. By contrast, from d_1 and d_2 we cannot deduce d_3 (something is a cat but not a dog). There are many notations related to spider and constraint diagrams; see [1, 6, 15, 19] for examples.

There are a number of examples of spider diagrams being used in practice, such as assisting with the task of identifying component failures in safety critical hardware designs [2] and in other domains such as [14]. They have also been used (but not explicitly) for displaying the results of database queries [20], representing non-hierarchical computer file systems [3], in a visual semantic web editing environment [13, 21] and other areas [7, 11]. Virtually all of these application areas, constants are used to represent specific objects, thus highlighting the importance of augmenting spider diagrams with constants.

The contribution made in this paper is to provide a set of sound and complete reasoning rules for spider diagrams augmented with constants. Section 2 overviews the syntax and semantics of spider diagrams with constants. A set of reasoning rules is presented in section 3. Soundness and completeness is established in section 4; the proof strategies are only sketched due to space limits. The technique used to prove completeness can be trivially adapted to provide a decision procedure for spider diagrams with constants.

2 Syntax and Semantics

We give an informal overview of the syntax and semantics of spider diagrams with constants; a formalization can be found in [17]. The spider diagram d_1 in figure 2 contains two labelled, closed curves called *contours*. The minimal regions that can be described as inside certain (possibly no) contours and outside the remaining contours are called *zones*; d_1 contains three zones whereas d_2 contains just two zones. *Missing zones* are zones which could be present in a diagram, given the contour label set, but are not present; for example, in figure 2, d_1 has exactly one missing zone: that which is inside $Lions$ but outside $Cats$. The diagram d_2 does not have any missing zones.

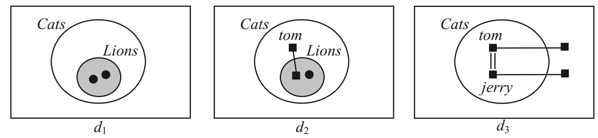


Figure 3. Shading and ties.

Spiders are placed in *regions*; a region is a set of zones. The region in which a spider is placed is called its *habitat*. Visually, a spider is represented by a tree whose nodes are either all round or all square; these nodes are called *feet*. Spiders with round feet are *existential* whereas those with square feet are *constant*. Constant spiders are labelled. The contours and spiders are all contained by a boundary rectangle which illustrates 'where the diagram stops'. Diagrams enclosed by such a rectangle are called *unitary diagrams*. In a unitary diagram, no two labels appear twice. Moreover, the labels used for constant spiders are never used for contour labels. This applies globally, in that any constant spider label never labels a contour and vice versa.

Unitary diagrams also contain further syntax: *shading* and *ties*. Shading is placed in zones, as demonstrated in figure 3. Ties can be placed between any pair of constant spider feet that are placed in the same zone. We treat the relation of two feet being joined by a tie as transitive: given any constant spider feet f_1, f_2 and f_3, if f_1 is joined to f_2 by a tie and f_2 is joined to f_3 by a tie then f_1 is joined to f_3. Moreover, we also assume that each foot is joined by a tie to itself (i.e. the relation is reflexive). This simplifies the formalization of the semantics; see [17]. Visually, ties are a pair of parallel line segments, like an equals sign, also demonstrated in d_3, figure 3. Note that, rather than drawing all ties between feet, we draw essentially a spanning forest of the graph that represents the 'is joined to be a tie' relation. Given two constant spiders, their *web* is the set of zones in which their feet are joined by a tie; in d_3, figure 3, the web of tom and $jerry$ is the single zone inside $Cats$.

We note that ties could also be used to connect existential spider feet. Indeed, they could also be used to connect an existential foot to a constant foot. However, for any diagram that incorporated such ties there exists a semantically equivalent diagram that does not contain such ties. This is not the case for ties between constant spider feet. It is straightforward to extend the work in this paper to the case where these additional ties are permitted.

In addition to the above, we take the symbol $\perp$ to be a unitary diagram. Unitary diagrams form the building blocks of *compound diagrams*: if d_1 and d_2 are spider diagrams then so are $(d_1 \vee d_2)$ and $(d_1 \wedge d_2)$.

For the semantics, regions represent sets, as illustrated in the introduction. Spiders represent elements in the sets represented by their habitats and distinct spiders represent distinct elements unless they are joined by a tie. Constant spiders represent specific individuals, just like contours represent specific sets; these individuals and sets are determined, in part, by their labels. Slightly more formally, an *interpretation* consists of a universal set, U, and a mapping of contour labels to subsets of U and constant spider labels to elements of U. The mapping of contour labels extends to the interpretation of zones and regions; see [17]. Two constant spiders represent the same individual if and only if they both represent an individual which is in the set denoted by some zone in their web. Shading places an upper bound on set cardinality: in the set represented by a shaded zone, all of the elements must be represented by spiders.

In figure 3, d_1 asserts that all lions are cats, there are at least two lions (by the use of two existential spiders) and there are no more than two lions (by the use of shading); in other words, all lions are cats and there are exactly two lions. The diagram d_2 expresses that all lions are cats, tom is a lion or a cat, something else is a lion and there are at most two lions. The diagram d_3 uses a tie to indicate that tom and $jerry$ are the same individual whenever they are both cats whereas if at least one of them is not a cat, they do not denote the same individual (due to the absence of a tie between their feet in the zone outside $Cats$).

The diagram $\perp$ is interpreted as false. The semantics extend to compound diagrams in the obvious way. Informally, we say that an interpretation is a *model* for a diagram when it 'agrees with the meaning of the diagram' as described above. Diagrams which have models are said to be *satisfiable*. The following theorem tells us that unitary diagrams are not capable of expressing contradictory information.

Theorem 2.1 *Let* d $(\neq\perp)$ *be a unitary spider diagram with constants. Then* d *is satisfiable.*

Given diagrams d_1 and d_2, we say that d_1 **semantically entails** d_2, denoted $d_1 \vDash d_2$, if all of the models for d_1 are also models for d_2.

3 Reasoning Rules

We will now develop a set of sound and complete reasoning rules for spider diagrams with constants. All of the reasoning rules given for spider diagrams without constants in [9] can be modified (sometimes non-trivially) and extended to spider diagrams with constants. In addition, new rules are also required.

3.1 Unitary to unitary reasoning rules

We introduce a collection of reasoning rules that apply to, and result in, a unitary diagram. For example, we can delete shading from a unitary diagram. All of the rules in this section are extended from those for spider diagrams without constants. The first rule allows the introduction of a shaded zone to a unitary diagram, d.

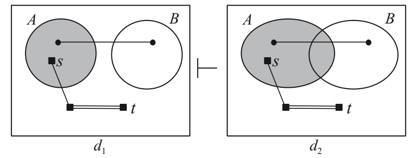


Figure 4. Introducing a shaded zone.

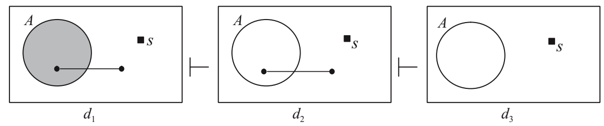


Figure 5. Erasing shading and spiders.

Rule 1 ***Introduction of a shaded zone*** Let d_1 be a unitary diagram that has a missing zone. If d_2 is a copy of d_1 except that d_2 contains a new, shaded zone then d_1 may be replaced by d_2 and vice versa.

In figure 4, rule 1 (introduction of a shaded zone) is applied to d_1 to give d_2. Applying the introduction of a shaded zone rule results in a semantically equivalent diagram. The next two rules are not information preserving.

Rule 2 ***Erasure of shading*** Let d_1 be a unitary diagram with a shaded region r. Let d_2 be a copy of d_1 except that r is completely non-shaded in d_2. Then d_1 may be replaced by d_2.

In figure 5, rule 2 (erasure of shading) is applied to d_1 to give d_2. Applying this rule 'forgets' the upper bound on the cardinality of the set represented by the region from which shading is erased.

Rule 3 ***Erasure of an existential spider.*** Let d_1 be a unitary diagram containing an existential spider e with a completely non-shaded habitat. Let d_2 be a copy of d_1 except that d_2 does not contain e. Then d_1 may be replaced by d_2.

In figure 5, rule 3 (erasure of an existential spider) is applied to d_2 to give d_3.

3.2 Unitary to compound reasoning rules

We now specify five further rules, each of which is reversible, that allow a unitary diagram to be replaced by a compound diagram, including a rule that allows us to introduce a contour. In the spider diagram without constants

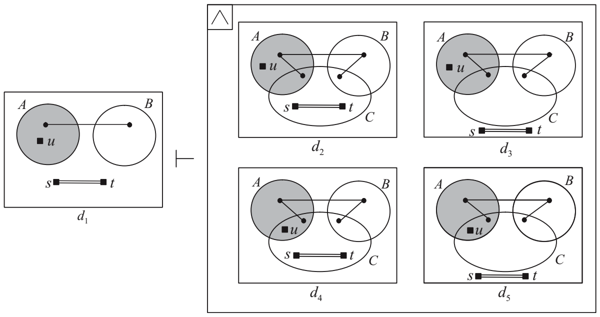


Figure 6. Diagram C-extensions.

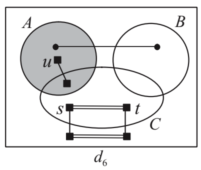


Figure 7. Incorrectly introducing a contour.

system, the introduction of a contour rule applies to, and results in, a unitary diagram. In figure 6, we can introduce the contour with label C to d_1. When we do so, each zone must split into two new zones, thus ensuring that information is preserved. The existential spiders' feet bifurcate, one new foot is placed in each new zone of the habitat. More care must be taken with the constant spiders, however. Consider, for example, the constant spiders s and t. The individual represented by both s and t must be either in $C - (A \cup B)$ or in $U - (A \cup B \cup C)$. The constant spider u represents an individual that is either in $A - (B \cup C)$ or $(A \cap C) - B$. This gives rise to four possibilities, shown in d_2, d_3, d_4 and d_5. We call these four diagrams the *C-extensions* of d_1. The diagram d_1 is semantically equivalent to $d_2 \vee d_3 \vee d_4 \vee d_5$. We could replace d_1 with the disjunction of just two unitary diagrams, each with u having two feet: one foot in the zone just in A, the other in the zone inside A and C.

It is not the case that the single unitary diagram d_6, in figure 7 is semantically equivalent to d_1. The constant spiders s and t must represent the same individual in d_1 but this is not the case in d_6, since the semantics of ties are zone based. To define the introduction of a contour rule, we first define the component parts of the resulting disjunction. We call these component parts *L-extensions*, where L is the introduced contour label.

Definition 3.1 *Let d_1 be a unitary diagram such that each constant spider in d_1 has exactly one foot. Let L be a contour label that is not in d_1. Let d_2 be a unitary diagram such that each constant spider in d_2 has exactly one foot. If the following conditions hold then d_2 is an **L-extension** of d_1.*

1. *The labels in d_2 are those in d_1, together with L.*
2. *The constant spider labels match.*
3. *The zones in d_2 are as follows:*
   (a) *each zone in d_1 is split into two zones in d_2, one inside and the other outside L;*
   (b) *shading is preserved in that if a zone is shaded in d_1 then the two zones it splits into in d_2 are both shaded in d_2 and no others are shaded in d_2.*
4. *The existential spiders match and their habitats are preserved under 'zone splitting'.*
5. *The habitat of each constant spider, c, in d_2 is either the zone inside L or that outside L (but not both) arising from its habitat it d_1.*
6. *Spider webs are preserved. Since constant spiders are single footed, this means that spiders joined by a tie in d_1 have the same habitats as each other in d_2.*

The set of L-extensions of d_1 is denoted $\mathcal{EXT}(L, d_1)$.

Rule 4 ***Introduction of a contour label*** Let d_1 $(\neq \perp)$ be a unitary diagram such that each constant spider has exactly one foot. Let L be a label not in d_1. Then d_1 may be replaced by $\bigvee_{d_2 \in \mathcal{EXT}(L,d_1)} d_2$ and vice versa.

Rule 5 ***Splitting spiders*** Let d be a unitary diagram with a spider s touching precisely every zone of two disjoint regions r_1 and r_2. Let d_1 and d_2 be unitary diagrams that are copies of d except that neither contains s, but instead each contains an extra spider, s_1 and s_2 respectively, whose habitats are regions r_1 in d_1 and r_2 in d_2. If s is a constant spider, then

1. s_1 and s_2 have the same label as s and
2. any ties joined to any given foot of s in d are joined to the appropriate foot of s_1 in d_1 or s_2 in d_2.

Then d can be replaced by the diagram $d_1 \vee d_2$ and vice versa.

Figure 8 illustrates an application of the splitting spiders rule. The spider s in d splits into two spiders, one in d_1, the other in d_2. Intuitively, the individual represented by s is either in the set $U - (A \cup B)$ or the set $A \cup B$.

Rule 6 ***Excluded middle*** Let d be a unitary diagram with a completely non-shaded region r. Let d_1 and d_2 be unitary diagrams that are copies of d except that d_1 contains an extra existential spider whose habitat is r and, in d_2, r is shaded. Then d can be replaced by the diagram $d_1 \vee d_2$ and vice versa.

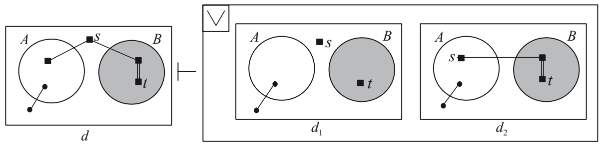


Figure 8. Splitting spiders.

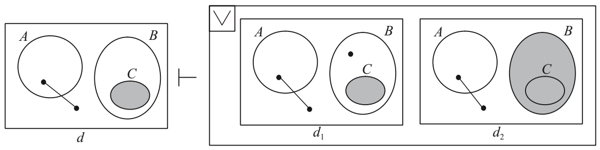


Figure 9. Excluded middle.

The diagram d in figure 9 can be replaced by $d_1 \vee d_2$ by applying the excluded middle rule.

Given a unitary diagram, d, that has only non-empty models (in which case d contains at least one spider), we can deduce that the individual represented by a constant spider, t, belongs to one of the sets denoted by the zones in d. Moreover, this individual must either be the same as, or different from, the elements already represented in d. As an example, consider d in figure 10 which has only non-empty models. Thus, in any model for d the constant spider t represents some individual. Then t is in A, B or $U - (A \cup B)$. If t is in A then it must equal s, since the region inside A is entirely shaded, shown in d_1. If t is in the set B then it may be either equal to, or different from, the element represented by the existential spider in B in the diagram d; these cases are represented by d_2 and d_3 respectively. Finally, if t is not in A or B then t must be in $U - (A \cup B)$, represented by d_4. The diagrams d_1, d_2, d_3 and d_4 are called t*-extensions* of d. For simplicity, we only add a constant spider to a diagram when all present spiders have exactly one foot; such a diagram is called an α**-diagram**.

Definition 3.2 *Let d_1 be a unitary α-diagram containing at least one spider and let t be some constant spider (label) that is not in d_1. Let d_2 be a unitary α-diagram. If the only*

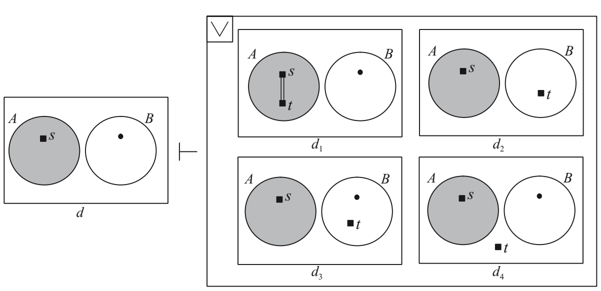


Figure 10. Diagram t-extensions.

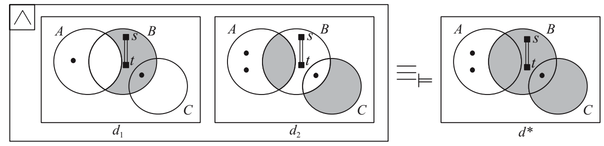


Figure 11. Combining diagrams.

*differences between d_1 and d_2 are the following then d_2 is a **t-extension** of d_1:*

1. *d_2 contains t with an arbitrary single zone habitat,*
2. *if the habitat of t is shaded then it is joined to another constant spider by a tie or there is one fewer existential spider in that zone (but not both),*
3. *if the habitat, z, of t is not shaded then either*
   - *(a) the number of existential spiders inhabiting z is reduced by one, or*
   - *(b) the number of existential spiders inhabiting z is the same, or*
   - *(c) t is joined by a tie to some constant spider also inhabiting z and the number of existential spiders is the same.*

The set of all t-extensions of d_1 is denoted $\mathcal{EXT}(t, d_1)$.

Rule 7 ***Introduction of a constant spider*** Let d_1 be a unitary α-diagram that contains at least one spider and let t be a constant spider not in d_1. Then d_1 can be replaced by the diagram $\bigvee_{d_2 \in \mathcal{EXT}(t,d_1)} d_2$ and vice versa.

Introducing the constant spider t to d in figure 10, by applying rule 7 results in $d_1 \vee d_2 \vee d_3 \vee d_4$.

The final rule in this section, called *combining*, replaces two unitary α-diagrams, with the same zone sets and constant spider label sets, taken in conjunction by a single unitary diagram. In figure 11, we illustrate the combining rule. We combine $d_1 \wedge d_2$ to give d^*. Any shading in either d_1 or d_2 occurs in d^*. Moreover, the number of spiders in any zone in d^* is the same as the maximum number that occur in that zone in d_1 or d_2. The diagram $d_1 \wedge d_2$ is semantically equivalent to d^*.

We now give a further example in a build-up to the definition of combining. In this example we derive results by working at the semantic level, but we will define the combining rule at the syntactic level. In figure 12, d_1 and d_2 contain contradictory information. We observe the following.

1. The zone inside A but outside B and C is shaded in d_1 and contains more spiders in d_2. Moreover, z_1 represents the empty set in any model for d_1. In any model for d_2, z_1 does not represent the empty set.

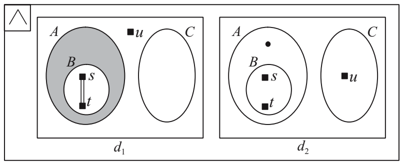


Figure 12. An unsatisfiable diagram.

2. The constant spider u has different habitats in the two diagrams. In any model for d_1, u represents an individual that is not in the set $A \cup C$. In any model for d_2, u represents an individual in the set C.

3. The constant spiders s and t are joined by a tie in d_1 but not in d_2. In any model for d_1, s and t represent the same individual, but in any model for d_2 they represent distinct individuals.

From any one of these three observations we can deduce that $d_1 \wedge d_2$ is unsatisfiable.

Definition 3.3 *Let d_1 and d_2 be unitary α-diagrams such that one of the following three conditions holds.*

1. *The zones are the same and the constant spider labels are the same.*

2. *The zones are the same and d_1 or d_2 is entirely shaded.*

3. *$d_1 = \perp$ or $d_2 = \perp$.*

Then d_1 and d_2 are called ***similar****.*

Definition 3.4 *Similar unitary α-diagrams d_1 and d_2 are said to be in* ***contradiction*** *if and only if one of the following four conditions holds.*

(i) Either $d_1 = \perp$ or $d_2 = \perp$.

(ii) There is a zone that is shaded in one diagram and contains more spiders in the other.

(iii) There is a constant spider with different habitats in d_1 and d_2.

(iv) There are two constant spiders that are joined by a tie in one diagram but not the other.

Lemma 3.1 *Let d_1 and d_2 be similar unitary α-diagrams.Then d_1 and d_2 are in contradiction if and only if $d_1 \wedge d_2$ is unsatisfiable.*

For completeness, it is sufficient to stipulate that the combining rule only applies to diagrams that have the same sets of constant spiders or that are contradictory.

Definition 3.5 *Let d_1 and d_2 be similar unitary α-diagrams Then their* ***combination****, denoted $d^* = d_1 * d_2$, is a unitary α-diagram defined as follows.*

1. *If d_1 and d_2 are in contradiction then $d_1 * d_2 = \perp$.*

2. *Otherwise $d^* = d_1 * d_2$ is a unitary α-diagram such that the following hold.*

   (a) The set of zones in the combined diagram is the same as the set of zones in the original diagrams.

   (b) The shaded zones in the combined diagram are shaded in one (or both) of the original diagrams.

   (c) The number of existential spiders in any zone in the combined diagram is the maximum number of existential spiders inhabiting that zone in the original diagrams.

   (d) The constant spiders in the combined diagram are the same as those in the original diagrams.

   (e) The habitats of the constant spiders in the combined diagram are the same as those in the original diagrams.

   (f) The webs of the constant spiders in the combined diagram are the same as those in the original diagrams.

Rule 8 ***Combining*** *Let d_1 and d_2 be similar unitary α-diagrams. Then $d_1 \wedge d_2$ may be replaced by $d_1 * d_2$, and vice versa.*

3.3 Logic reasoning rules

We now introduce a collection of rules, all of which have (obvious) analogies in logic. For space reasons, we give few details; throughout, D_1, D_2 and D_3 are arbitrary diagrams.

1. **Connect a diagram** D_1 can be replaced by $D_1 \vee D_2$.

2. **Inconsistency** $\perp$ can be replaced by D_1.

3. $\vee$**–Idempotency** D_1 may be replaced by $D_1 \vee D_1$ and vice versa.

4. $\wedge$**–Idempotency** D_1 may be replaced by $D_1 \wedge D_1$ and vice versa

We also assume that we have associativity, commutativity and distributivity.

3.4 Obtainability

To conclude this section on reasoning rules we define obtainability.

Definition 3.6 *Let D_1 and D_2 be two diagrams. Then D_2 is* ***obtainable*** *from D_1, denoted $D_1 \vdash D_2$, if and only if there is a sequence of diagrams $\langle D^1, D^2, ..., D^m \rangle$ such that $D^1 = D_1$, $D^m = D_2$ and D^k can be transformed into D^{k+1} by an application of a reasoning rule, for each k (where $1 \leq k < m$). If $D_1 \vdash D_2$ and $D_2 \vdash D_2$, we write $D_1 \equiv_\vdash D_2$.*

4 Soundness, Completeness and Decidability

To prove that the system is sound, the strategy is to start by showing that each reasoning rules is sound. The soundness theorem then follows by a simple induction argument. Due to space reasons, we omit the proofs.

Theorem 4.1 ***Soundness*** *Let D_1 and D_2 be spider diagrams. If $D_1 \vdash D_2$ then $D_1 \vDash D_2$.*

The completeness proof strategy for spider diagrams without constants given in [9] extends to the more general case here. The extended strategy, outlined in figure 13, is as follows. Suppose that $D_1 \vDash D_2$. The aim is to transform D_1 and D_2 into disjunctions of unitary α-diagrams using reversible rules, where, roughly speaking, each unitary part has some specified contour label set and constant spider label set.

Firstly, we split the constant spiders in D_1 and D_2 until, in each unitary part, all the constant spiders have precisely one foot, giving D_1^S and D_2^S respectively. This allows us to add contours to the unitary parts in both D_1^S and D_2^S using the reversible rule 4 (introduction of a contour label), until each (non-false) unitary part has the same contour label set, L. This gives D_1^L and D_2^L respectively. For the next step, zones are introduced to each unitary part until all (non-false) unitary parts have the same zone set, Z. This is done using the reversible rule 1 (introduction of a shaded zone) and yields D_1^Z and D_2^Z respectively. Now we obtain α-diagrams using the reversible rule 5 (splitting spiders), yielding D_1^α and D_2^α respectively. The formalization of the diagrams D_i^L, D_i^Z and D_i^α generalize those given in [9] for spider diagrams without constants.

We wish to introduce constant spiders to each side until each unitary part has the same constant spider label set. However, we can only introduce constant spiders when our diagrams contain at least one spider (ensuring non-empty models). Thus the next step we take is to apply the excluded middle rule to both sides until all the (non-false) unitary parts are either entirely shaded or contain a spider. The reversible rule 7 (introduction of a constant spider) is then applied, introducing constant spiders to all unitary parts that contain a spider, until all such unitary parts have some specified constant spider label set, C. This gives D_1^C and D_2^C respectively. We now apply rule 8 (combining) and some

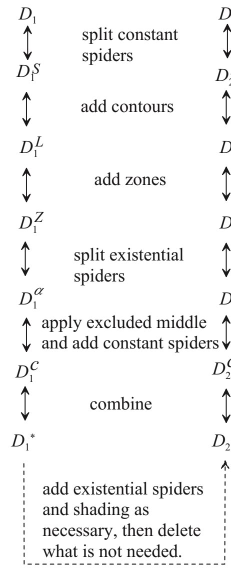


Figure 13. Proof strategy.

logic rules to remove all the conjuncts, giving two disjunctions of unitary α-diagrams, D_1^* and D_2^*. All of the unitary parts of D_1^* and D_2^* are either

1. $\perp$ or
2. have zone set Z and are entirely shaded and contain no spiders or
3. have zone set Z and constant spider label set C.

We note that $D_1 \equiv_\vdash D_1^*$ and $D_2 \equiv_\vdash D_2^*$, since all the rules applied so far are reversible. The diagram D_1^* (D_2^*) is a type of normal form that reflects the semantics of D_1 (D_2) in a clear manner.

We now apply the excluded middle rule to D_1^* until there are sufficiently many existential spiders and enough shading to ensure that each unitary part on the left hand side syntactically entails a unitary part of D_2^*, giving D_1'. Hence $D_1 \vdash D_1^* \vdash D_2^* \vdash D_2$ and, therefore, $D_1 \vdash D_2$.

The major differences between the completeness proof strategy here and that for spider diagrams without constants are the addition of the first step (splitting the constant spiders), with knock on changes to details of the other steps; and the insertion of an extra stage between splitting existential spiders and combining diagrams. Showing that $D_1^* \vdash D_2^*$ is also more challenging.

Theorem 4.2 ***Completeness*** *Let D_1 and D_2 be spider diagrams with constants. Then $D_1 \vDash D_2$ implies $D_1 \vdash D_2$.*

The proof of completeness provides an algorithmic method for constructing a proof that $D_1 \vdash D_2$ whenever

$D_1 \vDash D_2$. It is simple to adapt this algorithm to determine, given any D_1 and D_2, whether $D_1 \vdash D_2$.

Theorem 4.3 ***Decidability*** *Let D_1 and D_2 be constraint diagrams. There exists an algorithm that determines whether $D_1 \vdash D_2$.*

5 Conclusion

In this paper we have developed a sound and complete reasoning system for spider diagrams augmented with constants. The rules were largely extensions of those for spider diagrams without constants. However, some extensions were not straightforward, such as for the rule that allows the introduction of a contour and that which allows diagrams to be combined. In the latter case, this is due (in part) to the increased number of ways that two unitary diagrams can be inconsistent.

The proof of completeness is constructive, in that it provides an algorithm to convert a premise diagram into a conclusion diagram using the inference rules presented. Whilst we omitted details of the completeness proof due to space reasons, the proof strategy was a generalization of that for spider diagrams without constants. We note, though, that the actual details of the completeness proof are more complex when constants are involved.

In the future, we plan to integrate constant spiders into constraint diagrams. The work in this paper lays the foundations for developing a reasoning system for constraint diagrams augmented with constants. We also plan to investigate how to efficiently automate the search for readable proofs when using spider diagrams augmented with constants, following [5].

Acknowledgement This research is supported by EPSRC grant EP/E011160 for the *Visualization with Euler Diagrams* project.

References

[1] L. Choudhury and M. K. Chakraborty. On extending Venn diagrams by augmenting names of individuals. In *Proceedings of 3rd International Conference on the Theory and Application of Diagrams*, volume 2980 of *LNAI*, pages 142–146. Springer-Verlag, March 2004.

[2] R. Clark. Failure mode modular de-composition using spider diagrams. In *Proceedings of Euler Diagrams 2004*, volume 134 of *Electronic Notes in Theoretical Computer Science*, pages 19–31, 2005.

[3] R. DeChiara, U. Erra, and V. Scarano. VennFS: A Venn diagram file manager. In *Proceedings of Information Visualisation*, pages 120–126. IEEE Computer Society, 2003.

[4] A. Fish, J. Flower, and J. Howse. The semantics of augmented constraint diagrams. *Journal of Visual Languages and Computing*, 16:541–573, 2005.

[5] J. Flower, J. Masthoff, and G. Stapleton. Generating readable proofs: A heuristic approach to theorem proving with spider diagrams. In *Proceedings of 3rd International Conference on the Theory and Application of Diagrams*, volume 2980 of *LNAI*, pages 166–181, Cambridge, UK, 2004. Springer.

[6] E. Hammer. *Logic and Visual Information*. CSLI Publications, 1995.

[7] P. Hayes, T. Eskridge, R. Saavedra, T. Reichherzer, M. Mehrotra, and D. Bobrovnikoff. Collaborative knowledge capture in ontologies. In *Proceedings of the 3rd International Conference on Knowledge Capture*, pages 99–106, 2005.

[8] J. Howse and S. Schuman. Precise visual modelling. *Journal of Software and Systems Modeling*, 4:310–325, 2005.

[9] J. Howse, G. Stapleton, and J. Taylor. Spider diagrams. *LMS Journal of Computation and Mathematics*, 8:145–194, 2005.

[10] S. Kent. Constraint diagrams: Visualizing invariants in object oriented modelling. In *Proceedings of OOPSLA97*, pages 327–341. ACM Press, October 1997.

[11] H. Kestler, A. Muller, T. Gress, and M. Buchholz. Generalized Venn diagrams: A new method for visualizing complex genetic set relations. *Journal of Bioinformatics*, 21(8):1592–1595, 2005.

[12] S.-K. Kim and D. Carrington. Visualization of formal specifications. In *6th Aisa Pacific Software Engineering Conference*, pages 102–109, Los Alamitos, CA, USA, 1999. IEEE Computer Society Press.

[13] J. Lovdahl. *Towards a Visual Editing Environment for the Languages of the Semantic Web*. PhD thesis, Linkoping University, 2002.

[14] L. Niebrój. Defining health/illness: Societal and/or clinical medicine? *Journal of Physiology and Pharmacology*, 57(4):251–262, 2006.

[15] S.-J. Shin. *The Logical Status of Diagrams*. Cambridge University Press, 1994.

[16] G. Stapleton and A. Delaney. Evaluating and generalizing constraint diagrams. *Journal of Visual Languages and Computing*, available online, 2008.

[17] G. Stapleton, J. Taylor, J. Howse, and S. Thompson. The expressiveness of spider diagrams augmented with constants. *Journal of Visual Languages and Computing*, available online, 2008.

[18] G. Stapleton, S. Thompson, J. Howse, and J. Taylor. The expressiveness of spider diagrams. *Journal of Logic and Computation*, 14(6):857–880, December 2004.

[19] N. Swoboda and G. Allwein. Using DAG transformations to verify Euler/Venn homogeneous and Euler/Venn FOL heterogeneous rules of inference. *Journal on Software and System Modeling*, 3(2):136–149, 2004.

[20] J. Thièvre, M. Viaud, and A. Verroust-Blondet. Using euler diagrams in traditional library environments. In *Euler Diagrams 2004*, volume 134 of *ENTCS*, pages 189–202. ENTCS, 2005.

[21] Y. Zhao and J. Lövdahl. A reuse based method of developing the ontology for e-procurement. In *Proceedings of the Nordic Confernce on Web Services*, pages 101–112, 2003.

A Visual Language for Data Integration

Fausta Cosenza, Vincenzo Deufemia, Massimiliano Giordano, Giuseppe Polese, and Genoveffa Tortora

Dipartimento di Matematica e Informatica, Università di Salerno
Via Ponte don Melillo, 84084 Fisciano (SA), Italy
fausta.cosenza@gmail.com, {deufemia, mgiordano, gpolese, tortora}@unisa.it

Abstract

Many current approaches for constructing data warehouses rely on graph based visual languages to simplify the ETL programming task. In this paper we propose a new visual language for ETL production, which enhances graph based languages with the spreadsheet metaphor. The latter naturally lends itself to the data manipulation and data transformation task, giving immediate feedbacks on the effects of the planned data transformation processes. We also provide a grammar for the proposed language.

1. Introduction

Data warehouses heavily rely on huge amount of data collected from operational data sources and loaded in multidimensional databases through the so called ETL (Extraction,Transformation and Loading) process. Usually, ETL tools employ a data centric workflow to extract data from the sources, clean them from logical or syntactical inconsistencies, transform them into the format of the data warehouse, and eventually load them into the data warehouse.

ETL and Data Cleaning tools are estimated to cost at least one third of the effort and expenses in the budget of a data warehouse [2]. Moreover, Inmon estimates that the ETL process costs 55% of the total costs of data warehouse runtime [5]. Therefore, in order to cut costs, the design, development, and deployment of ETL processes, which are often performed in an ad-hoc, in-house fashion, needs uniform modeling, design, development methodology foundations, and automated tools that can boost productivity and reuse. To this end, several tailored approaches have been introduced to tackle the ETL task [3, 9]. In particular, some of them employ UML as a conceptual modeling language [8, 14], whereas others introduce a generic graphical notation for the same task [16]. A graph-based model for the definition of the ETL scenarios is presented in [17]. In particular, an ETL scenario is modeled through a special graph called Architecture Graph [15, 17], where nodes are used to represent activities, data stores, and their respective attributes. Activities have input and output schemata, whereas provider relationships relate inputs and outputs between data providers and data consumers.

A further extension of these approaches complements existing graph-based modeling of ETL activities by adding graph constructs to capture the semantics of insertions, deletions, and updates. The proposed model also supports negation, aggregation, self-joins, and a systematic way of transforming the so called “Architecture Graph” to allow zooming in and out at multiple levels of abstraction.

Graph-based models are considerably powerful, but inspite of their apparent simplicity they tend to produce huge and complex graph rapresentations of ETL processes. For their inherent nature graphs are powerful to express flow of data, and projection, but at the same time they provide a weak metaphor for supporting selection, trasformation, and mapping.

In this paper we propose a visual language introducing a novel approach to ETL process representation. Basically, we introduce the concept of “Visual Spreadsheet” (as proposed in some old commercial product like “Spreadsheet 2000” [10]) in the ETL desing process. Basically, the designer builds a Spreadsheet by selecting attributes from data sources, or creating new ones. On the spreadsheet s/he can define all kinds of transformations (as usual in spreadsheets), and apply filtering and projections operators through visual metaphors. We also present a visual grammar for modeling the proposed visual language from which we have automatically generated a compiler.

The paper is organized as follows. In Section 2, we discuss related work. Section 3 introduces the proposed language, and a formal grammar specification of it. In section 4 we present some experimental results. Finally, in section 5 we provide conclusions and insights for future work.

Figure 1. General job structure.

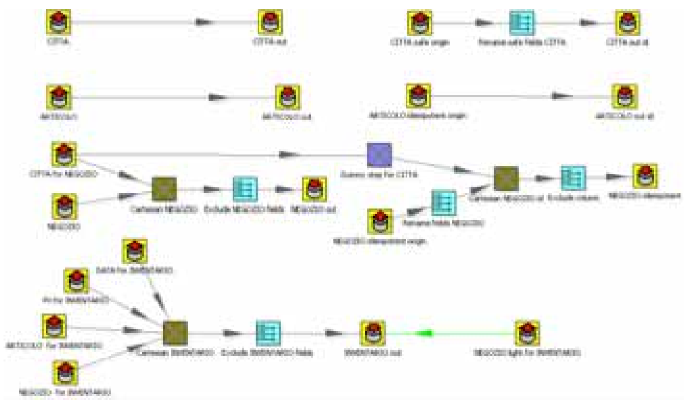


Figure 2. Data integration design with the tool Kettle

2. Visual Languages for ETL

In this section we present a survey on the use of visual languages in ETL programming, explaining how existing works can be improved.

2.1 Commercial Products

The following commercial products supporting the ETL process have been analyzed: Talend, Datastage and Kettle.

These tools share some common characteristics in the graphical user interface. In particular, they all provide a visual environment composed of: a Repository, that is, a toolbox gathering all the items that can be used to describe the business model or to sketch a job design; a Graphical workspace: a flowchar editor to manipulate both business models and job designs. Within the graphical workspace jobs all have the classical acyclic graph structure, starting from one or more sources, and ending in one or more destinations. Figure 1 shows the activities characterizing a job. Moreover, part of the graphical workspace contains a set of icons useful to design the data workflow.

In Figure 2 we show an example of data integration design performed with the Kettle tool.

2.2 Scientific Community

The efforts of the research on the data integration process have been devoted to:

- The optimization of the Data Integration process. In particular, there is a research concerning logic optimizations of the ETL process [11, 12], which models such a problem as a state space search problem, providing a heuristic based algorithm aiming to minimize the execution cost of the ETL workflow. Moreover, the concept of architectural graph is used for modeling the data integration dataflow [13]. Such a representation includes data stores, activities, and their components, as graph nodes connected through arcs representing one of four types of relations: instance-of, part-of, regulator, and provider. Then, special metrics are applied to such a structure, Dependence, and Responsability, in order to measure the degree of dependence between activities. However, such a model is difficult to represent and to comprehend, because of the huge quantity of activities involved in the data integration process.
- The individuation of templates for the data integration process [17]. In particular, research aims to detect templates representing frequently used ETL activities, including their semantics and interconnections.
- Extension of the semantics of ETL activities [18]. The semantics of data integration activities has been extended to include: negations, aggregations, and self joins; semantic complement of queries; transformation of the graph through zoom-in/out for multiple abstraction levels.
- On-line data refreshment approach. It refers to the concept of Active Datawarehouse [6], which represents the trend to update the data warehouse as soon as possible. Thus, ETL activities are implemented on queue networks, emplyoing Code Theory to predict performances and tuning of the operation.
- Engineering studies for evaluating ETL. An approach has been provided to evaluate commercial ETL tools [4]. The approach aims to: identify selection criteria that are essential for tool evaluation; develop scenarios to examine each criteria; develop quantitative measures to evaluate several aspects on the use of the ETL.
- Conceptual mapping of the ETL design based on cubes. It has been derived an approach introducing the concept of CommonCube [7], in order to represent the cubes into the final datawarehouse. The CommonCube highlights data transformation rather than loading. The approach defines the ETL mapping to capture the semantics of several cardinality relationships between source and destination attributes, and it represents a start point for the design of ETL. However, such approach provides quite a complex formal definition, so that a generic ETL designer is not sufficiently skilled to master it.

- Definition of automated processes for the maintenance of ETL design. It has been defined an approach for automatically generating an incremental ETL process starting from an entire ETL process [19]. The approach is based on the maintenance of materialized views, but it still suffers from the problem of the model complexity.

From the analysis of the above mentioned approaches we have detected the following problems:

1. Overall Design difficult to comprehend;
2. Impossibility to monitor operations and to perform a test on them;
3. Unintuitive operation management;
4. Lack of feedbacks on the results of executed operations;
5. Necessity to complete the design and to run it before understanding if the achieved result matches the one expected;
6. Lack of a visual language based formalization of generic functionalities enabling the construction of templates that can be used for future tasks;
7. Lack of a methodology related to the maintenance of the ETL process, both from the point of view of possible modifications of the involved activities, and of modifications on the data sources feeding the datawarehouse;

It turns out that is necessary to derive tools for optimizing the data extraction, transformation, and loading process, given the huge quantity of data to be manipulated. Such tools should provide visual representations easy to understand and to use for the Data Integration designer.

3. The proposed visual language

The proposed visual language has a hierarchical structure consisting of an "activity level" language whose symbols can be annotated with sentences of a "job level" language. The activity level language consists of a simple iconic language based on a directed graph, whereas the second language is based on the spreadsheet methaphor. In particular, the language at the first level is used to model the activities of a set of ETL jobs, whereas the second level language exploits the well-known spreadsheet metaphor to effectively assist the designer in structuring data integration activities. The first level resembles the well-known activity diagram, hence we focus on the description of the second level language.

3.1 Spreadsheet based visual programming

A typical spreadsheet based visual language is the one provided within the commercial tool Spreadsheet 2000, built for the Mac OS platform. Although such a tool was not more expressive than a traditional spreadsheet, it considerably simplified several complex operations through the use of graphical interaction metaphors, consequently reducing errors.

In our second level visual language we exploit the characteristics of spreadsheets in the construction of ETL procedures, in order to simplify their construction and to overcome many of the above mentioned problems arisen with traditional graph based approaches.

It is well known that spreadsheets provide a concise data representation, and it provides immediate and continuous feedbacks on the requested transformation operations, hence enabling a "trials and errors" approach, which can potentially boost user productivity. Another important advantage of spreadsheets is that most of their operations can be easily translated in a declarative language, such as SQL. At the same time, spreadsheet still support imperative procedures, which are particularly useful in the data transformation and cleaning phases.

3.2 Language elements

In what follows we describe the symbols of the proposed language, which are graphically shown in Figure 3:

- The datasheet: it is a graphical element integrating the functionalities of a dataset and those of a spreadsheet. As shown in Figure 4, the datasheet has a typical table layout, and it offers the usual operations of a spreadsheet.

- The New Column symbol: it is used to add new columns to a datasheet.

- The filter: it is a graphical element that allows to apply a filter on a given datasheet; the application of a filter is conditioned on a constraint that can be specified by means of the operator "condition";

- The join symbol: it enables join operations between datasheets. This operator is typically used in conjunction with the "condition" operator, but as opposed to the filter operator it is associated to two datasheets.

- The condition symbol: it allows to specify a condition in the context of other operations.

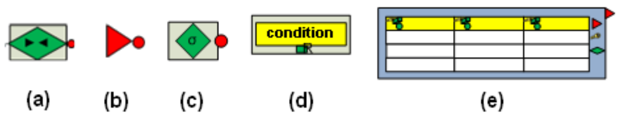


Figure 3. The spreadsheet visual language symbols: (a) join; (b) new column symbol; (c) filter operator; (d) condition; (e) datasheet

3.2.1 DataSheet

The ETL process can be modeled through several datasheets. A Datasheet can be constructed by dragging an empty datasheets on the working area and adding attributes to it by means of drag&drop operations from the data sources, or through the definition of new attributes.

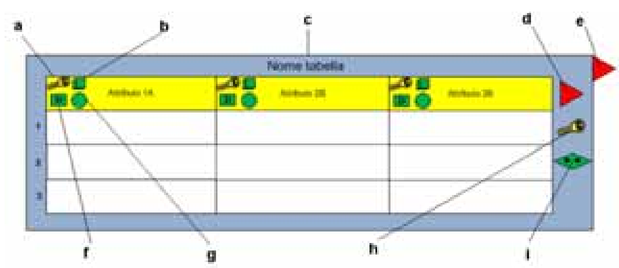


Figure 4. The datasheet symbol: (a) key property icon; (b) nullable icon; (c) datasheet name; (d) new column; (e) link to new table; (f) distinct selection; (g) filter definition; (h) surrogate key; (i) join operation

Through the datasheet it is possible to realize many operations:

- Selection: by dragging an attribute or a table of interest on an icon representing the empty datasheet;
- Transformation: it is possible to define a new attribute, or to alter an existing one, by applying a predefined function or a user provided one (or even by indicating a surrogate key), according to the typical syntactic style of a spreadsheet.

3.2.2 Filtering

The filtering operator enables us to filter the elements of a datasheet based on one or more conditions. The filter operator is associated with a "Reject" attribute that can be used to visualize all those elements of a source datasheet that do not match the filter condition.

A sample screenshot showing the application of a simple filter operator is shown in Figure 5(a).

3.2.3 Join Operations

The join operator enables the typical join used in relational databases (CROSS JOIN, INNER JOIN, OUTER JOIN) among several datasheets. It is used in conjunction with the "condition" operator, yielding a new resulting datasheet. Figures 5(b), 5(c), and 5(d) show the three types of join supported.

3.3. Grammar Specification

In the following we introduce the grammar for modeling the proposed visual language. To this end, we have used the eXtended Positional Grammar (XPG) formalism [1].

Let $DIS = (N, T \bigcup POS, S, P)$ be the XPG grammar for the proposed language. The set N of non terminal symbols is $\{Op, Join, Insert\}$, all having two attaching areas as syntactic attributes, called *1* and *2* as shown in Fig. 6. Op is the starting vsymbol type, i.e., $S = Op$.

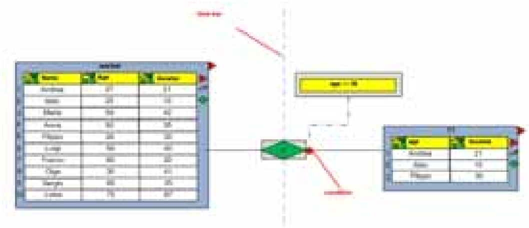
The set of terminals is given by T = {JOIN, INSERT, L, T, F, DL, COND, PH }. They are visualized in fig. 7. The set of relations POS contains the relation $\mathbf{LINK}_{i,j}$ which is defined as follows: a vsymbol x is in relation $\mathbf{LINK}_{i,j}$ with a vsymbol y iff attaching point (or region) i of x is connected to attaching point (or region) j of y, and will be denoted as **i_j** to simplify the notation.

Next, we provide the set of productions for describing the language.

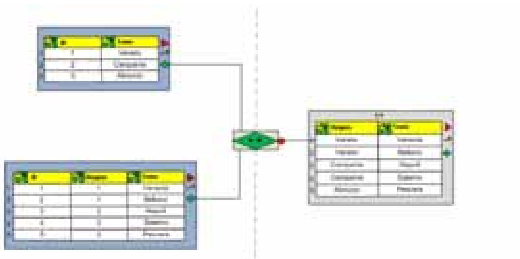
(1) $Op \rightarrow T \quad \Delta \quad (Op_1 = T_1, Op_2 = T_2, Op_3 = T_3)$

(2) $Op \rightarrow Op\prime \quad [3\_1] \quad L \quad [2\_1]$
$Join \quad [1\_2] \quad L\prime \quad [1\_3] \quad T \quad [2\_1]^{-2} \quad L\prime\prime \quad [2\_1] \quad T\prime$
$\Delta \quad (Op_1 = Op\prime_1, Op_2 = Op\prime_2,$
$Op_3 = Op\prime_3 - L\prime_1, Op_4 = Op\prime_4)$
$\Gamma \quad \{(P; |T_1 > 0| \quad or \quad |T_2 > 0| \quad or \quad |T_4 > 0|;$
$P_1 = T_1, P_2 = T_2, P_3 = T_3 - L\prime_1, P_4 = T_4),$
$(P\prime; |T\prime_2 > 0| \quad or \quad |T\prime_3 > 0| \quad or \quad |T\prime_4 > 0|;$
$P\prime_1 = T\prime_1 - L\prime\prime_2,$
$P\prime_2 = T\prime_2, P\prime_3 = T\prime_3, P\prime_4 = T\prime_4)\}$

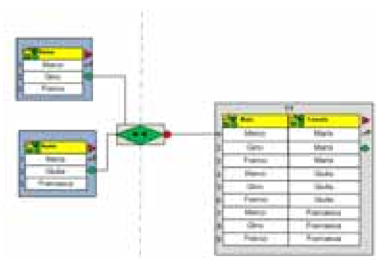
(3) $Op \rightarrow Op\prime[4\_1]L[2\_1]F[2\_1]$
$L\prime \quad [2\_1] \quad T \quad [2\_1]^{-2}L\prime\prime \quad [2\_1] \quad T\prime \quad [2\_1]^{-4}$
$DL \quad [2\_1] \quad COND$



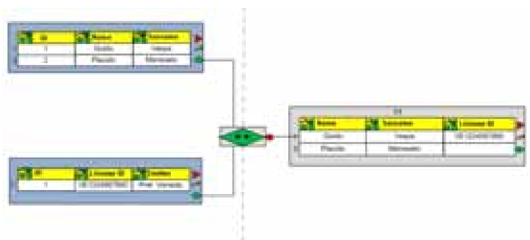
(a) Application of the filter operator on a datasheet



(b) Application of the inner join on two datasheets



(c) Application of the cross join on two datasheet



(d) Application of the outer join operator on two datasheets

Figure 5. Various Join and Filtering Operations

Figure 6. A typical visual representation of nonterminals for the proposed grammar

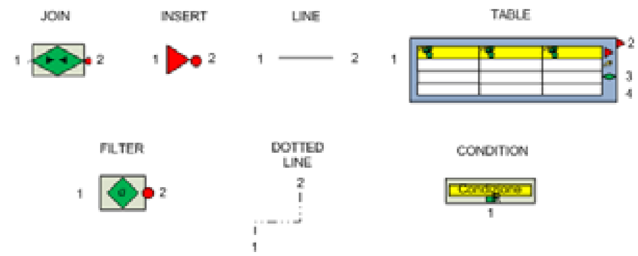


Figure 7. A typical visual representation of terminals for the proposed grammar

$$\begin{aligned}
&\Delta \quad (Op_1 = Op'_1, Op_2 = Op'_2,\\
&Op_3 = Op'_3, Op_4 = Op'_4 - L'_1)\\
&\Gamma \quad \{(P; |T_2 > 0| \quad or \quad |T_3 > 0| \quad or \quad |T_4 > 0|;\\
&P_1 = T1 - L'2, P_2 = T_2, P_3 = T_3, P_4 = T_4),\\
&(P'; |T'_2 > 0| \quad or \quad |T'_3 > 0| \quad or \quad |T'_4 > 0|;\\
&P'_1 = T'_1 - L''_2, P'_2 = T'_2, P'_3 = T'_3, P'_4 = T'_4)\}
\end{aligned}$$

$$\begin{aligned}
(4)\quad &Op \to Op' \quad [2_1] \quad L \quad [2_1]\\
&Insert \quad [2_1] \quad L' \quad [2_1] \quad T\\
&\Delta \quad (Op_1 = Op'_1,\\
&Op_2 = Op'_2 - L'1, Op_3 = Op'_3, Op_4 = Op'_4)\\
&\Gamma \quad \{(P; |T_2 > 0| \quad or \quad |T_3 > 0| \quad or \quad |T_4 > 0|;\\
&P_1 = T_1 - L'2, P_2 = T_2, P_3 = T_3, P_4 = T_4)\}
\end{aligned}$$

$$\begin{aligned}
(5)\quad &Op \to Op' < any > P\\
&\Delta \quad (Op_1 = Op'_1 + P_1; Op_2 = Op'_2 + P_2;\\
&Op_3 = Op'_3 + P_3; Op_3 = Op'_4 + P_4;)
\end{aligned}$$

$$\begin{aligned}
(6)\quad &Join \to JOIN \quad [2_1] \quad DL \quad [2_1] \quad COND\\
&\Delta \quad (Join_1 = JOIN_1; Join_2 = JOIN_2)
\end{aligned}$$

$$\begin{aligned}
(7)\quad &Join \to JOIN\\
&\Delta \quad (Join_1 = JOIN_1; Join_2 = JOIN_2)
\end{aligned}$$

$$\begin{aligned}
(8)\quad &Insert \to INSERT \quad [2_1] \quad DL \quad [2_1] \quad COND\\
&\Delta \quad (Insert_1 = INSERT_1; Insert_2 = INSERT_2)
\end{aligned}$$

$$\begin{aligned}
(9)\quad &Insert \to INSERT\\
&\Delta \quad (Insert_1 = INSERT_1; Insert_2 = INSERT_2)
\end{aligned}$$

4. Preliminary Evaluation

We have built a prototype tool embedding the proposed visual language and an XPG based compiler for it. The tool has been experimentally used for designing the ETL procedures of an industrial decision support systems in the context of a worldwide clothing selling company. We first employed three experienced programmers (1 senior and 2 junior), who were able to develop the requested ETL procedures in three man/months by using Java/JDBC/Oracle.

Successively, we employed five computer science students developing their project for a graduate Decision Support System course, asking them to develop the same ETL procedures by using our visual language based tool, assigning one student for each of the five data marts. Although they were less skilled than the people in the other group they were able to develop the same procedures in about 15 man/day, providing us with a first evidence of the effectiveness in using the visual approach with respect to traditional programming languages in ETL design.

5. Conclusions and Future Works

We have presented a visual language for ETL design in data warehouse development, experimentally proving its effectiveness in the context of a relevant industrial project.

In the future we aim to extend the first level visual language by providing new high level operations to better master the complexity of big projects. We plan to introduce several views to allow the designer to choose different levels of granularity during the design process. Moreover, we plan to create generic templates to enable an effective reuse of previous data integration projects, aiming to further reduce ETL development costs. Finally, we plan to investigate the possibility of using the visual sentences used in the data integration process as metadata to support future drill-through operation on the data warehouse.

References

[1] G. Costagliola, V. Deufemia, and G. Polese. A framework for modeling visual notations with applications to software engineering. *ACM TOSEM*, 13(4):431–487, Oct. 2004.

[2] J. T. C.Shilakes. Enterprise information portals.

[3] D. S. H. Galhardas, D. Florescu and E. Simon. Ajax: An extensible data cleaning tool. In *In Proc. of ACM SIGMOD*, page 590, USA, 2000.

[4] S. Henry, S. Hoon, M. Hwang, D. Lee, and M. DeVore. Engineering trade study: extract, transform, load tools for data migration. In *Systems and Information Engineering Design Symposium, 2005 IEEE*, pages 1–8, 2005.

[5] B. Inmon. The data warehouse budget, jan. 1997.

[6] A. Karakasidis, P. Vassiliadis, and E. Pitoura. Etl queues for active data warehousing. In *IQIS '05: Proceedings of the 2nd international workshop on Information quality in information systems*, pages 28–39, New York, NY, USA, 2005. ACM.

[7] Z. Li, J. Sun, H. Yu, and J. Zhang. Commoncube-based conceptual modeling of etl processes. *Control and Automation, 2005. ICCA '05. International Conference on*, 1:131–136 Vol. 1, 26-29 June 2005.

[8] S. Luján-Mora, P. Vassiliadis, and J. Trujillo. Data mapping diagrams for data warehouse design with uml. In P. Atzeni, W. W. Chu, H. Lu, S. Zhou, and T. W. Ling, editors, *ER*, volume 3288 of *Lecture Notes in Computer Science*, pages 191–204. Springer, 2004.

[9] V. Raman and J. M. Hellerstein. Potter's wheel: An interactive data cleaning system. In *The VLDB Journal*, pages 381–390, 2001.

[10] W. S. Building a visual programming language. *MacTech*, 13(4).

[11] A. Simitsis, P. Vassiliadis, and T. Sellis. Optimizing etl processes in data warehouses. In *ICDE '05: Proceedings of the 21st International Conference on Data Engineering*, pages 564–575, Washington, DC, USA, 2005. IEEE Computer Society.

[12] A. Simitsis, P. Vassiliadis, and T. Sellis. State-space optimization of etl workflows. *IEEE Trans. on Knowl. and Data Eng.*, 17(10):1404–1419, 2005.

[13] A. Simitsis, P. Vassiliadis, M. Terrovitis, and S. Skiadopoulos. Graph-based modeling of etl activities with multi-level transformations and updates. In A. M. Tjoa and J. Trujillo, editors, *DaWaK*, volume 3589 of *Lecture Notes in Computer Science*, pages 43–52. Springer, 2005.

[14] J. Trujillo and S. Luján-Mora. A uml based approach for modeling etl processes in data warehouses. In I.-Y. Song, S. W. Liddle, T. W. Ling, and P. Scheuermann, editors, *ER*, volume 2813 of *Lecture Notes in Computer Science*, pages 307–320. Springer, 2003.

[15] P. Vassiliadis, A. Simitsis, P. Georgantas, and M. Terrovitis. A framework for the design of etl scenarios. In J. Eder and M. Missikoff, editors, *CAiSE*, volume 2681 of *Lecture Notes in Computer Science*, pages 520–535. Springer, 2003.

[16] P. Vassiliadis, A. Simitsis, and S. Skiadopoulos. Conceptual modeling for etl processes. In *DOLAP '02: Proceedings of the 5th ACM international workshop on Data Warehousing and OLAP*, pages 14–21, New York, NY, USA, 2002. ACM.

[17] P. Vassiliadis, A. Simitsis, and S. Skiadopoulos. Modeling etl activities as graphs. In L. V. S. Lakshmanan, editor, *DMDW*, volume 58 of *CEUR Workshop Proceedings*, pages 52–61. CEUR-WS.org, 2002.

[18] P. Vassiliadis, A. Simitsis, and S. Skiadopoulos. On the logical modeling of etl processes. In *CAiSE '02: Proceedings of the 14th International Conference on Advanced Information Systems Engineering*, pages 782–786, London, UK, 2002. Springer-Verlag.

[19] X. Zhang, W. Sun, W. Wang, Y. Feng, and B. Shi. Generating incremental etl processes automatically. In *IMSCCS '06: Proceedings of the First International Multi-Symposiums on Computer and Computational Sciences - Volume 2 (IMSCCS'06)*, pages 516–521, Washington, DC, USA, 2006. IEEE Computer Society.

InfoHub: An Alternative Approach to Process Visualization

Daniel Gozick {daniel.gozick@gmail.com}
David Squire {squiredd@gmail.com}
Minho Park {parkm@sfasu.edu}
Stephen F. Austin State University
Department of Computer Science
Nacogdoches, Texas

Abstract

We present InfoHub, a cross-platform system based on process information that combines several established visualization techniques as well as some creative new techniques. Though process visualization is not a new topic in Computer Science, our series of tools brings the user new perspectives on traditional data to increase cognition of the current system. Our goal is to create a new metaphor for how people perceive and use the ubiquitous device known as a computer. This metaphor uses the computer hardware to depict system activities so that the user can become more familiar with how the computer functions. Working toward this goal, we utilized Processing, a Java based programming library, for graphics and XML for the data structure. While using this system, the user is then better informed, knowing more aspects of their hardware to better operate their computer as a whole.

1. Introduction

Information visualization has come a long way in recent years due to the overwhelming capability of computers that are readily available to perform the complex tasks needed to display the interpreted information. Information visualization is a valuable tool used to create interactive diagrams, animations, or images with the use of computer-supported systems to amplify cognition of the visual representation from the given data set [1]. There are numerous techniques to visualize different types of data sets such as bar charts, line graphs, scatter plots, star plots, time series, parallel coordinates, treemaps, and icicle trees [2-4]. Typically, these data sets are represented and/or categorized within one to n dimensions which will determine the type of visualization that will be used to best represent it.

Process visualization and monitoring, using analog and digital equipment, have been around for years. Today, there are numerous software based tools and libraries available which try to mimic the elder analog instruments in visualizing system information. Several of these tools and libraries are produced by Global Majic Software, Quinn-Curis Inc., LabVIEW System from National Instruments [5]. Additional software based applications and toolkits include the Grid Viewer [6], InfoVis Toolkit [4], and Prefuse Toolkit [7]. However, operating systems such as Windows and numerous Linux distributions contain a system monitor that enable users to view the current process in table form or as a two-dimensional line graph. Also, in the Unix environment, using the *top* command will provide continual reports to the user about the state of the system, including a list of the top CPU using processes [8].

The desktop metaphor translates stored computer data into meaningful graphics that are more familiar to users. This works well because it relates digital signals and information visual representations of real world objects. Our project attempts to describe a metaphor using real world objects that can actually be found inside the computer.

In this paper, we describe why knowing how computers work, even in the simplest form, will help educate students, consequently increasing performance and efficiency. Once given the process visualization, users will benefit from the improved features applied to several information visualization techniques that will be illustrated in this paper. This demonstrates how users are an integral part of the visualization process. How people perceive and interact with a visualization tool can strongly influence their understanding of the data as well as the system's usefulness [9].

To visualize process monitoring, we created a multi-viewing system, visualizing multivariate data from a list of sources which may change over time. This framework demonstrates the information collected from the system information collector and displays the activity, density, and hierarchies in various forms using InfoHub. The techniques used to visualize, ranging from common representations of 2D and 3D to creative representations, will aid in the assimilation and understanding of the data. This software implements some of the basic aspects of focus + context [10]. The computer is modeled in a generic way and the user has the ability to select a module to gain further understanding. While looking at the system as a whole, some details become apparent as the CPU calculates and memory is allocated. At a lower level, details related to other modules are still visible. This lends itself to better understanding of the inter-working between modules.

When implementing the process visualization, we used Processing, Java, and XML. Processing and Java were used to create the visualizations and graphical user interface. Also, XML was used to structure the data for more efficient interpretation by our program. This made the implementation distributed and modular in nature. This means that modules can be included or improved upon easily. Overall, XML provided a simple way of describing lists of multivariate data.

2. Related Work

Process monitoring has been around for years since the early days of industrialization. Analog equipment was first introduced, followed by the use of digital equipment. Finally, computers were introduced along with software, and they are used increasingly to monitor and visualize system processes [5].

2.1. Operating System Process Monitoring

Operating Systems such as Windows and Linux have system monitors which display the current processes and other information as shown in Figure 1. These monitoring systems allow the user to view the current processes with a graphical user interface.

Figure 1: Windows Task Manager. Visualizes the CPU usage and history in the amount used related to time sectioned by the number of CPU's or cores, memory usage and history.

2.2. Operating System Process Visualization

Operating Systems, such as Windows and Linux, also have system monitors which display a simple line graph of the CPU and memory history, as well as a bar chart displaying total CPU and memory usage as shown in Figure 1. This visualization gives users a better representation of how active or inactive their computer is.

2.3. Similar Techniques

For the process and memory visualization we added a hot and cold theme to help the user better identify the filtered information. Some of the visual techniques used were a modified three dimensional star plot and a three dimensional bar chart. A star plot is a graphical data analysis technique for examining the relative behavior of all variables in a multivariate data set [11].

2.4. Discussion

The desktop metaphor does well to depict the hierarchy of files on a computer, but it hides details about what is being done by the machine at a given time. This leads a user to think that this detail is beyond understanding. This is hardly the case as a user may benefit from knowing more about tasks being performed by a machine. By interacting with InfoHub on a daily basis a user can familiarize herself with how a computer typically operates. This will make errors and faults more noticeable to the user. Knowing that a process is running in an infinite loop and will never be available to perform requested actions would allow a user to close such an application and move on to other tasks. In terms of memory, the user would benefit from knowing if a process is continually allocating memory without permission. These examples only illustrate a few problems that a user may face. Our system is also useful for educational purposes. Users who see an animated approach to computer interaction may learn more about what is going on, or better yet, may be motivated to ask more questions. Therefore, this system is also useful in academia where topics of computer architecture are explored. The system has such a simple nature that it is useful for everyday computing and for further exploration into computing as a collection of inter-working devices.

3. Overview

The central goal of InfoHub is to provide a system environment that allows users to access system data that is generally hard to reach. This data represents information critical to the operating system that is brought to attention in graphical form. Thus, it provides a simple understanding of significant information extracted from the visualizations. Many aspects of a computer's work can be modularized, which can provide very useful information for the user. InfoHub also provides user interactions, making navigation easier. To visualize the data, a multi-viewing system was needed to best represent the given data. Multiple views can provide utility in terms of minimizing some of the cognitive overhead engendered by a single, complex view of data [12]. Each view can focus on several attributes that the user will be able to select from a drop-down menu, thus enabling the user to better understand a specific data set.

The motherboard model is the initial view of the system. As time passes, activity becomes apparent and is visually depicted by spheres of color which traverse the modules on the board. The zooming feature allows the user to view more details of the modules and messages being sent between them. This module acts as a navigation hub. From this point forward, other modules can be selected for a more detailed analysis. Likewise, the modules will allow for navigation back to the motherboard model.

The CPU module will demonstrate how a process is utilizing the processor's resources. A star spoked temperature wheel visualization technique is used to depict the process information as shown in Figure 2. Each process becomes a new spoke with a colored node at the end connecting to the center cube, which represents the CPU. The color representation demonstrates process information selected by the user, (i.e. CPU usage, processing time, or memory usage). We have used shades of blue and red to represent a temperature-like differentiation, where blue depicts low utilization, and red depicts higher utilization. The nodes are sorted based on user selection, creating the star spoked wheel which starts with high utilization and progresses in a clockwise fashion to the lowest utilization. From this view, when a process is selected it is illuminated. This creates a dialog box with additional detail for that process. The view is three dimensional and can be rotated by the user when the rotating mode is enabled.

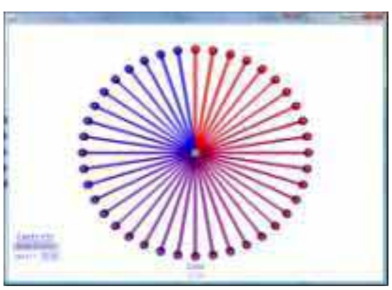


Figure 2: CPU visualization. Based on the Star Spoked Temperature Wheel. Shows the list of process usage based on the users selection.

The memory module depicts the memory usage for each process by utilizing a three-dimensional bar graph as shown in Figure 3. To avoid some processes hiding others, each process is sorted such that all of the large processes in memory are together on one side. Also, the hot and cold temperature theme is continued for better understanding. Rotating the display allows users to view all of the active processes. Each process is transparent so that almost all memory processes are visible simultaneously. Color is used to differentiate between processes.

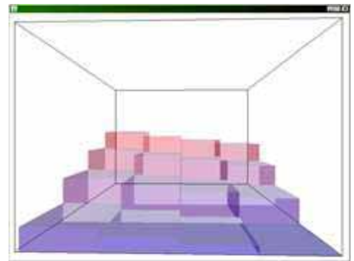


Figure 3: Memory Visualization. Three dimensional bar chart. Shows memory usage per process.

Several alternate approaches have also been implemented, which highlight selected processes. The Animated Memory System provides a view of process request allocation and de-allocation from the CPU. System memory is shown as a conceptual block on the right of the screen. Each process is given an area proportional to its size in memory. As the system makes requests to allocate or de-allocate memory, these requests are displayed by colored balls. These requests will appear on the left side of the screen and head towards the right seeking their location in memory as they travel. The color of the ball is determined by the type of request using the hot and cold theme. The system memory block also has the theme to depict the amount of activity from a process.

Another system overview displays the selected attributes in a "Koosh ball" fashion, which is known as the Fuzzball visualization. Figure 4 displays the system overview with the hot and cold theme. The selection of a specific process will highlight the areas that pertain to it.

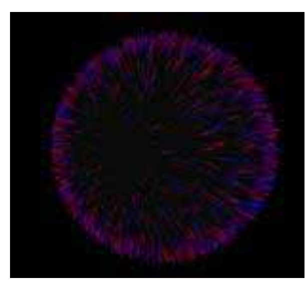


Figure 4: Fuzzball visualization. Displays features selected by the user.

4. Approach

The use of Java in this project allows it to be fully cross-platform capable. This means that the utility offered by the system can be reproduced for any architecture that has the Java Virtual Machine running on it. For this system, we used Processing [13] to display the graphical visualizations discussed above. Processing is a Java based programming library that allows for the quick creation of graphical interfaces for use within the project. These two decisions made a world of difference in our attempts to make a cross-platform and user friendly system. Linux was chosen in order to facilitate gathering the data used for visualization. This is important due to the open access to such information in a UNIX based operating system.

Animated Memory System shows a simplified overall methodology for process visualization. A data file is transformed into data which can then be visualized. Specifically, in Linux the proc directory contains data files that describe each process and the computer in general. These files are in a readable form and are updated over time. XML is used to simplify storage and maintain modularity caused by differences in operating systems. To create the XML from each operating system, a program is created to grab the information when called. Of course, this process needs to be specialized for each operating system because of accessibility issues. The XML file is then interpreted by another program that interprets and visualizes the data. The XML is modularized based on processes and time. This allows for any operating system with access to process information and system data to be used to generate an appropriate XML that will become a visualization of their data. Another benefit of this structure is that new modules can be added to the visualization program without necessarily modifying the XML generator, and conversely, recording new information into the XML will not require an immediate change to the visualization program. This flexibility will help future development of our project.

5. Conclusion and Future Work

We have presented an alternative visualization for process monitoring. And we have discussed and explained our implementation approaches under Processing, Java, and XML environments.

Our goal is to allow users to evaluate system functions, such as CPU, memory, file structures, and networking, using several visualization techniques that best represent the given data.

In the near future, we will include all of these system functions in our system environments as well as develop various visualization methods. Eventually, we will implement general visualization environments that allow users to choose how to visualize their system. In addition, various user interaction techniques, such as panning, zooming, filtering, searching, rotating, and three dimensional perspectives, will be included. However, this will require more modularity and functionality for cross-platform support. Additionally, we plan to support other operating systems. This will require changes to the program that generate the XML files. Our process should be generalized to other datasets in order to visualize different types of information.

Altogether, the generalization of our process will become a framework that will be useful for visualization of structured sets of data. Useful new thoughts could be generated by analyzing these datasets.

6. References

[1] Card, S., Mackinlay, J., Shneiderman, B., *Readings in Information Visualization Using Vision to Think,* Morgan Kaufmann Publishers, 1999.

[2] Spencer, R., *Information Visualization: Design for Interaction*, 2nd Edition, Pearson Education Limited, 2007

[3] Keim, D. "Information Visualization and Visual Data Mining," *IEEE Transactions on Visualization and Computer Graphics.* Vol. 7, pp 1-8, 2002.

[4] Fekete, J.D., "The InfoVis Toolkit," *IEEE Symposium on Information Visualization 2004.* pp 167-174, 2004.

[5] Matkovi c, K., Hauser H., Sainitzer, R., and Groller, M. "Process Visualization with Levels of Detail," *IEEE Transactions on Visualization and Computer Graphics.* Vol. 8, pp 1-8, 2002.

[6] Osborne, L., and Thalayattampalli, S., "Gridinfo: A Graphical Grid Monitoring Tool In Java," *Journal of Computing Sciences in Colleges, ACM*, pp 104-113, 2005.

[7] Heer, J., Card, S., and Landay, J. A. prefuse: a toolkit for interaction information visualization. Proc. CHI 2005, ACM Press, pp 421-430, 2005.

[8] Unix Top, http://www.unixtop.org/

[9] Troy, M., and Moller, T., "Human Factors In Visualization Research," IEEE Transactions on Visualization and Computer Graphics. Vol. 10, pp 72-84, 2004.

[10] Lamping, J., Rao, R., Pirolli, P., "A Focus+Context Technique Based on Hyperbolic Geometry for Visualizing Large Hierarchies." CHI 1995, pp 401-408, 1995.

[11] "Star Plot," *National Institute of Standards and Technology Dataplot Reference Manual,* pp 2-232 – 2-233, 1997.

[12] Wang Baldonado, M.Q., Woodruff, A., Kuchinsky, A., "Guidelines for Using Multiple Views in Information Visualization," AVI: Proceedings of the working conference on Advanced visual interfaces. pp 110- 119, 2000.

[13] Processing, http://processing.org/

Abstract Euler Diagram Isomorphism

Gem Stapleton
Visual Modelling Group
University of Brighton, UK
g.e.stapleton@brighton.ac.uk

Andrew Fish
Visual Modelling Group
University of Brighton, UK
andrew.fish@brighton.ac.uk

Peter Rodgers
Computing Laboratory
University of Kent, UK
p.j.rodgers@kent.ac.uk

Abstract

Euler diagrams are widely used for information visualization and form the basis of a variety of formal languages that are used to express constraints in computing. Tools to automatically generate and layout these diagrams from an abstract description have to overcome the fact that this problem is computationally difficult. We develop a theory of isomorphism of diagram descriptions and identify invariants of these descriptions under isomorphism. We can apply this theory to improve the efficiency of the generation of all abstract descriptions (up to isomorphism). We can also consider the production and use of libraries of diagrams with nice visual properties: by providing a normal form for the abstract descriptions we can improve efficiency of searches for isomorphic diagrams within such libraries and, moreover, utilize invariants for further efficiency savings. We produce an implementation of the theory and give an indication of the efficiency improvements.

1 Introduction

In this paper we develop the theory of isomorphism for Euler diagrams. Isomorphism of Euler diagrams is relevant in many applications where they are used. Examples of their application areas include any areas where representing relationships between collections of objects is helpful, such as [4, 9, 11, 13, 16, 22, 23, 24].

In computing, many modelling notations use closed curves as part of their syntax and can, therefore, be viewed as extending Euler diagrams. As an example, constraint diagrams [12] were designed for formal software specification as a potential replacement for the Object Constraint Language; see [10, 14] for examples of software modelling using constraint diagrams. A constraint diagram can be seen in figure 1.

The study of Euler diagram isomorphism is related to the study of other topics, such as graph isomorphism, where the study of optimizations for isomorphism checking are well advanced [15, 18, 19]. The abstract descriptions of Euler diagrams are equivalent to the abstract descriptions of hypergraphs [17], although the visualization method is different. This equivalence means that the techniques presented in this paper are applicable to hypergraphs.

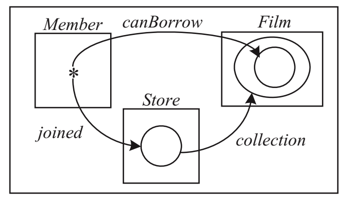


Figure 1. A constraint diagram.

After providing motivation (section 2) and background material (section 3) we develop the theory of diagram isomorphism, in section 4, and identify invariants of Euler diagrams in section 5. The problem of Euler diagram isomorphism checking is equivalent to hypergraph isomorphism checking. Applications of this theory include improving efficiency when generating abstract diagrams to populate a library, or searching for a diagram within such a library. We produce an implementation which performs isomorphism checking, making use of the invariants as optimisations. The amount of efficiency gain by using our methods is given by running the task of generating all abstract diagrams and comparing the effects of applying different invariants, detailed in section 6.

2 Motivation

In general, the generation problem is to find an Euler diagram that represents some specified collection of sets and possess certain properties, sometimes called wellformedness conditions. The input to an Euler diagram generation algorithm is an abstract description of the diagram to be generated. Various methods for generating Euler diagrams have been developed, each concentrating on a particular class of Euler diagrams, for example [2, 3, 6, 13, 25]. The existing generation algorithms use some method for embedding curves in the plane. Diagram generation using these algorithms is a time consuming process and, moreover, can result in layouts which are not aesthetically pleasing and, therefore, potentially hard to read.

A previously unadopted generation approach is to have a library of nicely drawn diagrams from which an appropriate diagram is selected. In this context, it is helpful to have a notion of isomorphism amongst abstract descriptions in order to facilitate the extraction of an appropriate Euler diagram from the library. When using a library of example Euler diagrams, each token in the library would be marked with is abstract description. For example, suppose we wish to generate a diagram that expresses $A \cap B = \emptyset$ and $C \subseteq B$, like d_1 in figure 2. The diagram d_1 is 'isomorphic' to d_2, which expresses $C \cap A = \emptyset$ and $B \subseteq A$. If the library contained d_3 then we could select d_3 and insert labels in the appropriate way to yield either d_1 or d_2. To extract d_3 from the library would require us to establish that its abstract description is isomorphic to that for d_1.

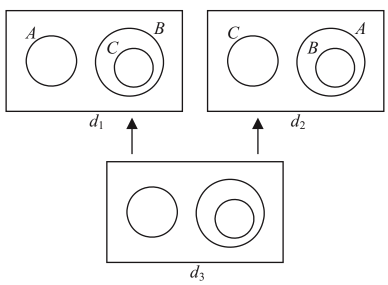


Figure 2. Using a library for generation.

We will need to be able to extract appropriate diagrams from the library in an efficient manner. Merely checking for isomorphism is time consuming and, if the library is large, will be infeasible in some cases. Thus, we require methods to partition the library of examples to reduce the number of checks that must be performed. When searching through a library, we can use invariants to subdivide the search space, reducing the number of brute-force checks to determine whether any given abstract description is isomorphic to another.

We can first partition the set of library diagrams by the number of labels they contain. Thus, we can break down the set of library examples into sets each of which contains all diagrams with a given number of labels. It is likely that this will be sufficient to provide an efficient search through the space when there are few labels. However, as the number of labels increases, the number of diagrams in the associated set can be very large. Thus, we could further subdivide these sets using more sophisticated isomorphism invariants. Obviously the order in which the subdivisions are performed can have an impact on the efficiency of the search through the library. In this paper, we identify a collection of invariants and in the future we will establish how best to use them for set-subdivision.

A further motivation for the work in this paper relates to listing all abstract descriptions, up to isomorphism, for various purposes. These include classifying types of diagrams that might be drawn nicely and being able to count various types of diagrams such as those which are 'atomic' (defined later). Once such a collection has been built up, we can use it to generate a library of drawn diagrams using previously developed generation techniques. Just as when searching through the library, will need to be able to efficiently check for isomorphism when producing a collection of all abstract descriptions.

3 Euler Diagrams: Syntax and Semantics

We briefly overview the syntax and semantics of Euler diagrams; the formalizations are adapted from [20, 21]. The diagram d_1 in figure 2 contains three *contours*; these are the closed curves labelled A, B and C. Contours represent sets and their spatial relationship is used to make statements about containment and disjointness of sets. So, in d_1 the contours assert that the sets A and B are disjoint because the contours do not overlap in any way; similarly $A \cap C = \emptyset$. The placement of C inside B expresses $C \subseteq B$. A *zone* is a maximal sets of points in the plane that can be described as being inside certain contours and outside the rest of the contours. The diagram d_1 contains four zones, one of which can be described as inside B but outside both A and C. A **drawn Euler diagram** is a finite collection of labelled closed curves, in which each label occurs at most once.

To formalize diagram descriptions, the input to a generator, all that is necessary is knowledge about the labels and the zones that are to be present. A zone description is taken to be a set of labels, in, following the style of [6]; the set in contains the labels of the curves that contain a particular zone. For example, the zone inside B but outside A and C in d_1 is described by $\{B\}$; given the labels A, B and C, we can deduce that the zone described by $\{B\}$ is outside A and C.

We can also talk about zone descriptions that are subsets of the label set in a diagram, d, but for which no zone occurs in d. For example $\{A, B\}$ describes a zone which is not present in d_1 (no zone is inside both A and B). Such zones are said to be *missing* from the diagram; d_1, therefore, has four missing zones. Further, we assume the existence of some fixed label set, $\mathcal{L}$, from which we choose the labels used in our Euler diagrams.

An **abstract Euler diagram** is an ordered pair, (L, Z) where $L \subseteq \mathcal{L}$ and $Z \subseteq \mathbb{P}L$. It is these abstract Euler diagrams that are the input to generation algorithms and, thus, from which a drawn Euler diagram is generated. We frequently blur the distinction between abstract diagrams and their drawn counterparts and simply refer to Euler diagrams; the context will make it clear whether we mean 'drawn' or 'abstract'. Similarly, a set of labels will be called a zone.

At the semantic level, an *interpretation* is a universal set, U, together with an assignment of a subset of U to each contour (strictly, to contour labels) which is extended to interpret zones. Formally, an **interpretation** is a pair, (U, Ψ) where U is any set and $\Psi\colon \mathcal{L} \rightarrow \mathbb{P}U$ is a function. In a diagram $d = (L, Z)$, a zone $in \subseteq L$ represents the set $\bigcap_{l \in in} \Psi(l) \cap \bigcap_{l \in L - in} (U - \Psi(l))$. An interpretation is a **model** for d if all of the zones which are missing from d represent the empty set.

4 Isomorphic Diagrams

Given a fixed label set, $L \subseteq \mathcal{L}$, the table below identifies how many abstract diagrams there are with that label set; the set of diagrams with label set L is denoted $D(L)$.

$\lvert L\rvert$	0	1	2	3	4	5
$\lvert D(L)\rvert$	2	4	16	256	65536	4294967296

Note that given a label set L and a subset, M of L, such that $|M| = |L| - 1$, we have $|D(L)|$=$|D(M)|^2$. Of course, many of these abstract Euler diagrams are isomorphic. For example, d_1 and d_2 in figure 2 have abstract descriptions $L = \{A, B, C\}$ and zone sets $\{\{A\}, \{B\}, \{B, C\}\}$ and $\{\{C\}, \{A\}, \{A, B\}\}$ respectively which are the same up to renaming the labels. Recall that a permutation of a set S is a bijection from S to itself.

Definition 4.1 *Let $d_1 = (L_1, Z_1)$ and $d_2 = (L_2, Z_2)$ be Euler diagrams. Then d_1 and d_2 are **isomorphic** if there exists a permutation, $\sigma\colon \mathcal{L} \rightarrow \mathcal{L}$, such that the image of σ when the domain is restricted to L_1 equals L_2 and which induces a bijection, $\Sigma\colon Z_1 \rightarrow Z_2$ defined by $\Sigma(z_1) = \{\sigma(l) : l \in z_1\}$; such a permutation σ is called an **isomorphism** from d_1 to d_2.*

There is potentially a large number of non-isomorphic diagrams that have the same number of labels; the table below gives the numbers for up to five labels, where $NID(L)$ denotes the set of non-isomorphic diagrams.

$\lvert L\rvert$	0	1	2	3	4	5
$\lvert NID(L)\rvert$	2	4	12	80	3984	37333248

The table below shows how many diagrams there are with $|L|$ labels and $|Z|$ zones up to isomorphism and indicates how many brute-force isomorphism tests we might need to perform if we first compare number of labels and number of zones.

$\lvert L\rvert$	$\lvert Z\rvert$								
	0	1	2	3	4	5	6	7	8
0	1	1							
1	1	2	1						
2	1	3	4	3	1				
3	1	4	9	16	20	16	9	4	1
4	1	5	17	52	136	284	477	655	730
5	1	6	28	134	625	2674	10195	34230	100577

The numbers in the above two tables are obtained from [8], after noting the analogy with switching theory. Even with only five labels, $|L| = 5$, the number of non-isomorphic diagrams with, say, 8 zones, is rather large at 100577. When $|Z|$ increases to 16, there are 5182326 non-isomorphic diagrams with five labels. Thus it is important to have further methods of partitioning the set of non-isomorphic diagrams.

4.1 Linking Isomorphism and Semantics

The notion of isomorphism is well-defined with respect to semantics: although isomorphic diagrams are not, in general, semantically equivalent but they are *expressively equivalent*, defined below.

Definition 4.2 *Let (U, Ψ_1) and (U, Ψ_2) be two interpretations. We say (U, Ψ_1) and (U, Ψ_2) are **permutation-equivalent** if there exists a permutation of $\mathcal{L}$, say $\sigma\colon \mathcal{L} \rightarrow \mathcal{L}$, such that for all $l \in \mathcal{L}$, $\Psi_1(l) = \Psi_2(\sigma(l))$.*

To illustrate, the interpretation with $U = \{1, 2, 3\}$, $\Psi(A) = \{1\}$, $\Psi(B) = \{2, 3\}$ and $\Psi(C) = \{3\}$ is a model for d_1 but not d_2 in figure 2. However, the interpretation $U = \{1, 2, 3\}$, $\Psi(C) = \{1\}$, $\Psi(A) = \{2, 3\}$ and $\Psi(B) = \{3\}$ is a model for d_2 but not d_1. These two models are permutation equivalent and d_1 and d_2 are isomorphic.

Definition 4.3 *Let $d_1 = (L_1, Z_1)$ and $d_2 = (L_2, Z_2)$ be Euler diagrams. Then d_1 and d_2 are **expressively equivalent** if there exists a fixed permutation $\sigma\colon \mathcal{L} \rightarrow \mathcal{L}$ such that the models for d_2 are precisely those which are permutation equivalent to the models for d_1 under σ.*

Theorem 4.1 *Two diagrams are isomorphic if and only if they are expressively equivalent.*

5 Isomorphism Invariants

In the worst case, determining whether two abstract Euler diagrams are isomorphic takes exponential time in proportion to the number of labels in the diagram. Of course, there are some obvious checks that one can perform to reduce the computations involved by using invariants. The invariants we define below each capture a certain type of structure present in Euler diagrams that is preserved under isomorphism. The two most obvious invariants are the number of labels present and the number of zones present. However, these invariants do not capture the complexity of the relationships between the closed curves (i.e. how the zone set relates to the label set). Thus, we define a range of invariants that go some what towards capturing features of the zone set and its relationship with the label set. For example, a diagram that contains exactly two zones inside one contour will not be isomorphic to any diagram that has no contour which contains exactly two zones. Our refined invariants take this kind of difference between non-isomorphic diagrams into account.

5.1 Zone Invariants

Trivially, we can use the number of zones as an isomorphism invariant.

Lemma 5.1 (Invariants) *Let $d_1 = (L_1, Z_1)$ and $d_2 = (L_2, Z_2)$ be isomorphic Euler diagrams. Then $|Z_1| = |Z_2|$.*

Rather than simply counting the zones, we can count the number of zones of each size, where the size of a zone is the number of labels it contains. To illustrate, isomorphic diagrams in figure 2 each have 1 zone inside 0 contours, 2 zones inside 2 contours, 1 zone inside 2 contours, and no zones inside all three contours.

Definition 5.1 *Let $d = (L, Z)$ be an Euler diagram. Then the* ***zone-partition sequence*** *associated with d, denoted $ZPS(d)$, is defined to be a sequence of natural numbers, $ZPS(d) = (s_0, s_1, ..., s_{|L|})$, where s_i is the number of zones in d with cardinality i (informally, s_i is the number of zones in d that are 'inside' i contours).*

Theorem 5.1 (Invariant) *Given two isomorphic diagrams d_1 and d_2, $ZPS(d_1) = ZPS(d_2)$.*

Zone-partition sequences provide a method of subdividing the space of abstract diagrams into smaller classes, beyond the subdivision provided by simply counting the labels and the zones. Indeed, when we use zone-partition sequences to subdivide our library of examples, we no longer need to consider the number of labels and number of zones, since these are derivable from $ZPS(d)$; the number of labels in d is the length of $ZPS(d)$ and the number of zones is the sum of the entries in $ZPS(d)$.

Hardly surprisingly, there are diagrams which are not isomorphic but that have the same zone-partition sequence. To illustrate, the non-isomorphic diagrams in figure 3 both have zone-partition sequence $(1, 2, 2, 0, 0)$.

5.2 Label Invariants

As with the number of zones, we can use the number of labels as an isomorphism invariant.

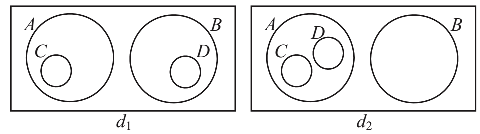


Figure 3. Zone partition sequences.

Lemma 5.2 (Invariants) *Let $d_1 = (L_1, Z_1)$ and $d_2 = (L_2, Z_2)$ be isomorphic Euler diagrams. Then $|L_1| = |L_2|$.*

We can further refine our subdivision of the set of atomic diagrams by considering how many zones are contained by each label. First we have the following lemma, whose proof proceeds by a simple induction on $|L|$.

Lemma 5.3 *For any Euler diagram $d = (L, Z)$ and label $l \in L$, there are at most $2^{|L|-1}$ zones, z, in d that contain the label l, that is $l \in z$.*

Definition 5.2 *Let $d = (L, Z)$ be an Euler diagram. Then the* ***label-partition sequence*** *associated with d, denoted $LPS(d)$, is defined to be a sequence of natural numbers, $LPS(d) = (t_0, t_1, ..., t_{2^{|L|-1}})$, where t_i is the number of labels in d which contain i zones (informally, t_i is the number of contours in d that contain i zones).*

The diagrams in figure 3, which have equal zone-partition sequence, have different label-partition sequences: $LPS(d_1) = (0, 2, 2, 0, 0, 0, 0, 0, 0)$ and $LPS(d_2) = (0, 3, 0, 1, 0, 0, 0, 0, 0)$.

Theorem 5.2 (Invariant) *Given two isomorphic diagrams d_1 and d_2, $LPS(d_1) = LPS(d_2)$.*

5.3 Using Labels and Zones

A further subdivision of the set of abstract diagrams can be achieved by considering in more detail the relationship between the labels used and the zones in which they occur. First, recall that a lexicographical ordering of sequences can be achieved by ordering $(l_1, l_2, ..., l_n)$ before $(m_1, m_2, ..., m_p)$ whenever there is a j such that $l_i = m_i$ for all $i < j$ and $l_j < m_j$.

Definition 5.3 *Let $d = (L, Z)$ be an Euler diagram. For each label, $l \in L$, we define a* ***label-zone sequence****, denoted $LZS(l, d)$ to be $(l_1, l_2, ..., l_{2^{(|L|-1)}})$ where l_i is the number of zones in d that contain i labels and which include l, that is $l_i = |\{z \in Z : l \in z \wedge |z| = i\}|$. Further, we define the* ***label-zone sequence*** *for d, denoted $LZS(d)$, to be the lexicographically ordered sequence of label-zone sequences for the labels in d.*

For example, in figure 3, the contour A in d_1 contains three zones; of these, one is inside just A and two are inside A and another contour. This gives $LZS(A, d_1) = (1, 1, 0, 0)$. Similarly, B has label-zone sequence $LZS(B, d_1) = (1, 1, 0, 0)$. For C and D we have

$$LZS(C, d_1) = LZS(D, d_1) = (0, 1, 0, 0).$$

Thus,

$$LZS(d_1) = ((0, 1, 0, 0), (0, 1, 0, 0), (1, 1, 0, 0), (1, 1, 0, 0)).$$

The label-zone sequence for d_2 is

$$LZS(d_2) = ((0, 1, 0, 0), (0, 1, 0, 0), (1, 0, 0, 0), (1, 2, 0, 0)).$$

This more refined invariant captures more accurately the relationship between labels and the zones they contain.

Lemma 5.4 *Given diagrams d_1 and d_2 such that $LZS(d_1) = LZS(d_2)$, we have the following*

1. $LPS(d_1) = LPS(d_2)$ *and*
2. $ZPS(d_1) = ZPS(d_2)$.

Theorem 5.3 (Invariant) *Given isomorphic diagrams d_1 and d_2, $LZS(d_1) = LZS(d_2)$.*

5.4 Refinements to Blocks

Refinements to the above invariants can be made based on the observation that one only requires so-called *atomic* diagrams from which one can generate all diagrams in a relatively simple way. An atomic diagram is one in which the curves form a connected component of the plane; non-atomic diagrams are called **nested** [7].

This notion of nesting at the drawn diagram level has an analogy at the abstract level, defined by considering the *dual graph* (in [5, 6] this graph is called the *superdual*). Given an abstract diagram, $d = (L, Z)$, the **dual graph** of d, denoted $dual(d)$, has Z as its vertex set and there is any edge between vertex v_1 and vertex v_2 if the symmetric difference of v_1 and v_2 contains exactly one label. Figure 4 shows a diagram with its dual graph.

Under certain well-formedness conditions, the absence cut-vertices in the dual graph corresponds to atomic diagrams [7]. We can use the dual graph to further partition the space of non-isomorphic diagrams. Recall, a block of a graph is a maximal connected subgraph that contains no cut vertex.

Definition 5.4 *A **block** of an abstract Euler diagram, $d_1 = (L, Z_1)$ is an abstract Euler diagram, $d_2 = (L, Z_2)$ such that the dual of d_2 is a block of the dual graph of d_1.*

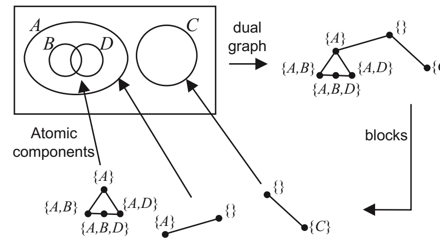


Figure 4. Finding diagram blocks.

An illustration can be seen in figure 4. The blocks of d can be found by computing the blocks of $dual(d)$ for which there are known algorithms, see [1] for example.

Definition 5.5 *Let d be an Euler diagram. The **blockwise zone-partition sequence** for d, denoted $BZPS(d)$, is the lexicographically ordered sequence of zone-partition sequences for the blocks of d.*

To illustrate, the nested diagram d_1 in figure 4 has blockwise zone-partition sequence

$$BZPS(d_1) = ((0, 1, 2, 1, 0), (1, 1, 0, 0, 0), (1, 1, 0, 0, 0)).$$

Theorem 5.4 *Given two isomorphic diagrams, d_1 and d_2, $BZPS(d_1) = BZPS(d_2)$.*

The definitions of the other sequences given above, such as the label-partition sequence, extend to blockwise sequences in a manner similar to that exemplified by the zone-partition sequence. Moreover, these further blockwise sequences are also invariant under isomorphism. We conjecture that these refined blockwise invariants will be particularly helpful when abstract descriptions have large numbers of labels.

5.5 Refinements to Complements

We note that the more zones there are in a diagram, the more checks we need to perform when seeking to establish whether two diagrams are isomorphic. When more than half of the zones are present, we can reduce these number of checks by comparing diagram *complements*.

Definition 5.6 *The **complement** of an Euler diagram, $d_1 = (L, Z_1)$, is an Euler diagram, $d = (L, Z_2)$, where $Z_2 = \mathbb{P}L - Z_1$.*

In figure 5, d_1 has complement $(L = \{A, B, C\}, Z = \{\{B\}, \{B, C\}\})$. Clearly, given a diagram, d, the complement of its complement is d.

Theorem 5.5 *Given isomorphic diagrams d_1 and d_2, their complements are isomorphic.*

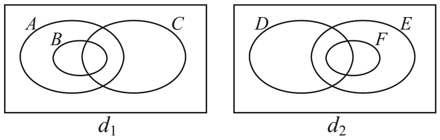


Figure 5. Diagram complements.

To illustrate the computational savings argument, suppose we wish to determine whether the diagrams d_1 and d_2 in figure 5 are isomorphic and that we have already established that the various sequences associated with them, which are invariant under isomorphism, are equal. Then we would need to construct bijections, $\sigma\colon \{A, B, C\} \to \{D, E, F\}$, and determine whether they are isomorphisms by checking that each zone in d_1 maps to a zone in d_2 (that is, check whether σ extends bijectively to the zones). Alternatively, we can check whether σ extends bijectively to the complements, reducing the number of checks (the complements of both diagrams contain just two zones whereas each original diagram contains 6 zones).

6 Implementation

In this section, an implementation of abstract Euler diagram isomorphism testing in a software system is described. The system implements the optimizations discussed previously for isomorphism testing. However, it also includes additional improvements for listing all the unique abstract diagrams for a particular number of sets.

When listing all abstract diagrams up to isomorphism, the implementation only iterates through abstract diagrams that do not have the empty zone, $\emptyset$, present. To get the full list of abstract diagrams, the ones in the generated list simply have to have the empty zone added to them. This reduces the number of abstract diagrams built by a half. For example, with two labels, we will generate the abstract diagram $(\{A, B\}, \{A\})$ from which $(\{A, B\}, \{\emptyset, \{A\}\})$ can subsequently be generated.

We now outline our method for producing all abstract diagrams with some fixed number of labels. Firstly, the Venn diagram with the required number of labels, $|L|$, is found, and each of its zones is assigned an integer, from one to the number of zones. For example, when we require two labels, and take $L = \{A, B\}$, the zones in the Venn diagram are (excluding $\emptyset$) $\{A\}, \{B\}, \{A, B\}$; each zone is then assigned a number, $\{A\} \mapsto 1, \{B\} \mapsto 2$ and $\{A, B\} \mapsto 3$. Using this, we can give a binary word that describes the zone set; for example, a diagram that contains the zones $\{A\}$ and $\{A, B\}$ would be described by 101 since, for example, $\{A, B\}$ was assigned to 3 and its presence is indicate by a 1 in the third place of 101.

We then iterate from zero to the number of diagrams to be listed (minus one), $|D(L)| - 1$ (we subtract one since we have already generated the Venn diagram), and produce an abstract diagram by treating them number as a binary. In our two-contour example, the number of diagrams to be listed is six, since we do not include the zone $\emptyset$; there are 12 non-isomorphic diagrams with two zones. It is easy to see that there are six binary words containing three bits over the alphabet $\{0, 1\}$.

As stated above, the presence of a zone is given by a 1 in the binary description. To further illustrate, the list of all zones for 3 contours is $(\{A\}, \{B\}, \{C\}, \{A, B\}, \{A, C\}, \{B, C\}, \{A, B, C\})$, seven zones (and 128 diagrams to be generated, as the empty zone is not present), with $\{A\}$ numbered one and $\{A, B, C\}$ numbered seven. Combination number 69, with binary representation 0100111, is the abstract diagram with zone-list $(\{B\}, \{A, C\}, \{B, C\}, \{A, B, C\})$.

The method above is used to produce all abstract diagrams with some fixed number of labels: each binary word is converted into a subset of the Venn zone set, giving all subsets of the Venn zone set. During the process of listing all such diagrams, any diagram which does not have A in its first zone are not generated. This is because such diagrams are isomorphic to another diagram with A in its first zone, and so the following processes do not need to be applied.

Each abstract diagram is converted into a normal form. This normal form maps the sets to labels so that they are in a particular order, with A the highest value label, B the next, and so on. The labels are, firstly, compared by checking each zone size (by zone size, we simply mean the cardinality of the zone). A label occurring in a smaller size zone than another label gets a higher value. Where there is no difference in the zone size that they appear in, then a label occurring in a zone with a higher ordered label where the other label does not gets a higher value. Otherwise, if an order between labels cannot be derived then they are ordered arbitrarily. The labels in zones are lexically ordered, then the zones in abstract diagrams are first ordered by size, then lexical ordering of same size zones is performed.

For example, given the diagram with zone set $\{\{B\}$, $\{D\}$, $\{A, C\}$, $\{B, C\}, \{B, C, D\}\}$, B and D are higher than A and C because they are in zones of size 1, B is ordered higher than D because it appears in zones of size 2. Moreover, C is higher than A because it appears next to B. Giving a mapping $\sigma(B) = A$, $\sigma(D) = B$, $\sigma(C) = C$ and $\sigma(A) = D$, the diagram has the normalized zone set $\{\{A\}, \{B\}, \{A, C\}, \{C, D\}, \{A, B, C\}\}$.

This means that many isomorphic diagrams will have the same label ordering. So, when listing all unique diagrams, if the normal form already occurs in the set of known diagrams, there is no need to do the further isomorphism tests. Of, course, this normalization repeats some of the steps in the isomorphism optimization given below, and the isomorphism tests for two diagrams that reduce to the same normal form would tend to be quick in any case, as the brute force part of the algorithm would not have many alternative mappings to test. However, it is useful to have a consistent way of displaying abstract descriptions. The set of known unique abstract diagrams is stored in a hash set. This speeds up the test to determine whether a normalized abstract diagram is already in the set.

If equality by normalization cannot be derived then the current normalized abstract diagram is compared against the set of known unique abstract diagrams by using an isomorphism test. This test performs the optimizations described previously in sections 5.1 to 5.3: the invariants are compared first and, if equal, a brute force algorithm has to be executed to perform isomorphism tests. First, a partition of the label set is discovered. Each set in the partition contains the labels that can map to the same choices of labels in the other abstract diagram. Hence, members of a partition are those labels with indistinguishable values from the normalization process. Given that the sets can be placed in order, further optimizations can be carried out, such as ensuring each pair of sets in order from both diagrams are of equal size.

This table gives the number times that pairs of diagrams have been identified as non-isomorphic by simply checking the invariant when generating a list of all unique abstract diagrams with a given number labels. For example, for diagrams with 4 contours, the zone-partition sequence alone identifies 3446662 diagram pairs as non-isomorphic. Both the normalization and the absence of the zone $\emptyset$ (that contained by no contours) optimization have been used in computing table entries. As a result, the number of times a brute-force tests for isomorphism has been applied is greatly reduced for the more effective invariants. For example, using the label-zone sequence to identify non-isomorphic diagrams, the number of brute-force isomorphism tests for four labels was reduced to just 1886; this is a very small fraction of the number of isomorphism tests that were saved by using the invariant, as detailed in the table (3468679).

Invariant	$\|L\|$			
	1	2	3	4
$\|L\|$	1	9	231	138645
$\|Z\|$	1	13	687	3018907
LPS	1	14	784	3413693
ZPS	1	15	814	3446662
LZS	1	15	822	3468679

7 Conclusion

We have defined a notion of isomorphism between abstract Euler diagrams and presented a collection of invariants that can be used to reduce the time taken to identify whether two abstract diagrams are isomorphic. The resulting efficiency savings are likely to be useful when using a library of examples to display Euler diagrams, for instance. Our implementation includes some optimizations and we have provided data comparisons to demonstrate the effects of utilizing different invariants. Further optimizations are possible and, since the problem of determining isomorphism between Euler diagrams is identical to the hypergraph isomorphism problem, optimizations can be drawn from graph isomorphism and should be very feasible. For example, we could convert to a constraint satisfaction task [15, 19] or adapt the graph isomorphism methods of Nauty [18].

Another useful application of our work is in the use of the invariants such as the zone-partition sequence to help build an initial library of examples, enabling the generation of a wide spread of diagrams; the idea being that drawn diagrams with nice visual properties are stored for each abstract diagram in the library. The intention would be that such that a library based generation system would have methods for combining diagrams, and the abstract diagram generation and comparison techniques could help guide the design of the rules for diagram combination.

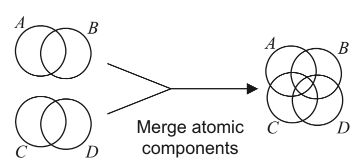


Figure 6. Combining diagrams.

It may be that diagrams in a library can be combined in sophisticated ways, such as taking two atomic components and merging them together to create a new diagram; figure 6 shows the merger of two atomic diagrams into a single atomic diagram (as opposed to a nested diagram). The theory required to facilitate this type of constructing is currently being developed and would reduce the number of diagrams required to be stored in the library.

Further detailed experiments to determine the best way in which to use the invariants to divide the search space are also required. The manner in which the invariants are used will impact the search through a library in order to extract a diagram which corresponds to an input abstract diagram description, for example.

Acknowledgements This research is supported by EPSRC grants EP/E011160/01 and EP/E010393/01 for the *Visualization with Euler Diagrams* project. Thanks also to Ebru Keyman for useful discussions on aspects of this paper.

References

[1] G. Chartrand and O. Ollermann. *Applied and algorithmic graph theory*. McGraw-Hill, 1993.

[2] S. Chow and F. Ruskey. Drawing area-proportional Venn and Euler diagrams. In *Proceedings of Graph Drawing 2003, Perugia, Italy*, volume 2912 of *LNCS*, pages 466–477. Springer-Verlag, September 2003.

[3] S. Chow and F. Ruskey. Towards a general solution to drawing area-proportional Euler diagrams. In *Proceedings of Euler Diagrams*, volume 134 of *ENTCS*, pages 3–18, 2005.

[4] R. DeChiara, U. Erra, and V. Scarano. A system for virtual directories using Euler diagrams. In *Proceedings of Euler Diagrams 04*, volume 134 of *Electronic Notes in Theoretical Computer Science*, pages 33–53, 2005.

[5] J. Flower, A. Fish, and J. Howse. Euler diagram generation. *Journal of Visual Languages and Computing*, available online, 2008.

[6] J. Flower and J. Howse. Generating Euler diagrams. In *Proceedings of 2nd International Conference on the Theory and Application of Diagrams*, pages 61–75, Georgia, USA, April 2002. Springer.

[7] J. Flower, J. Howse, and J. Taylor. Nesting in Euler diagrams: syntax, semantics and construction. *Software and Systems Modelling*, 3:55–67, March 2004.

[8] A. Harrison. *Introduction to Switching and Automata Theory*. McGraw-Hill, 1965.

[9] P. Hayes, T. Eskridge, R. Saavedra, T. Reichherzer, M. Mehrotra, and D. Bobrovnikoff. Collaborative knowledge capture in ontologies. In *Proceedings of the 3rd International Conference on Knowledge Capture*, pages 99–106, 2005.

[10] J. Howse and S. Schuman. Precise visual modelling. *Journal of Software and Systems Modeling*, 4:310–325, 2005.

[11] J. Howse, G. Stapleton, and J. Taylor. Spider diagrams. *LMS Journal of Computation and Mathematics*, 8:145–194, 2005.

[12] S. Kent. Constraint diagrams: Visualizing invariants in object oriented modelling. In *Proceedings of OOPSLA97*, pages 327–341. ACM Press, October 1997.

[13] H. Kestler, A. Muller, T. Gress, and M. Buchholz. Generalized Venn diagrams: A new method for visualizing complex genetic set relations. *Journal of Bioinformatics*, 21(8):1592–1595, 2005.

[14] S.-K. Kim and D. Carrington. Visualization of formal specifications. In *6th Aisa Pacific Software Engineering Conference*, pages 102–109, Los Alamitos, CA, USA, 1999. IEEE Computer Society Press.

[15] J. Larrosa and G. Valiente. Constraint satisfaction algorithms for graph pattern matching. *Mathematical Structures in Computer Science*, 12:403–422, 2002.

[16] J. Lovdahl. *Towards a Visual Editing Environment for the Languages of the Semantic Web*. PhD thesis, Linkoping University, 2002.

[17] E. Luks. Hypergraph isomorphism and structural equivalence of boolean functions. In *Proceedings of 31st Annual ACM Symposium on Theory of Computing*, 1999.

[18] B. McKay. Practical graph isomorphism. *Congressus Numerantium*, 30:45–87, 1981.

[19] S. Sorlin and C. Solnon. A global constraint for graph isomorphism problems. In *6th International Conference on Integration of AI and OR Techniques in Constraint Programming for Combinatorial Optimisation Problems*, pages 287–301. Springer, 2004.

[20] G. Stapleton, J. Masthoff, J. Flower, A. Fish, and J. Southern. Automated theorem proving in Euler diagrams systems. *Journal of Automated Reasoning*, 39:431–470, 2007.

[21] G. Stapleton, S. Thompson, J. Howse, and J. Taylor. The expressiveness of spider diagrams. *Journal of Logic and Computation*, 14(6):857–880, December 2004.

[22] N. Swoboda and G. Allwein. Using DAG transformations to verify Euler/Venn homogeneous and Euler/Venn FOL heterogeneous rules of inference. *Journal on Software and System Modeling*, 3(2):136–149, 2004.

[23] N. Swoboda and G. Allwein. Heterogeneous reasoning with Euler/Venn diagrams containing named constants and FOL. In *Proceedings of Euler Diagrams 2004*, volume 134 of *ENTCS*. Elsevier Science, 2005.

[24] J. Thièvre, M. Viaud, and A. Verroust-Blondet. Using euler diagrams in traditional library environments. In *Euler Diagrams 2004*, volume 134 of *ENTCS*, pages 189–202. ENTCS, 2005.

[25] A. Verroust and M.-L. Viaud. Ensuring the drawability of Euler diagrams for up to eight sets. In *Proceedings of 3rd International Conference on the Theory and Application of Diagrams*, volume 2980 of *LNAI*, pages 128–141, Cambridge, UK, 2004. Springer.

Visually Specifying Multi-Agent Simulation Environments with DIVAs

R. Steiner, R. Zalila-Wenkstern, J. Burnell

MAVs Laboratory, Department of Computer Science, University of Texas at Dallas
2601 North Floyd Rd, Richardson, TX 75080, USA
email: {rsteiner, rmili}@utdallas.edu,
jimburnell@student.utdallas.edu

Abstract

*Time and time again, computer scientists design beautiful systems that solve complex problems. However, a common shortcoming is the inability to extend that solution to individuals outside of computer science. In this paper, we present a tool that allows users outside of the computer science community to take advantage of the power of multi-agent simulation systems. Multi-agent simulation systems (MASS) are desirable since they can exhibit qualities such as emergence, self-organization, learning and adaptive behavior thus enabling the modeling and understanding of more real world complex application domains. Our domain independent tool, DIVAs (***D**ynamic **I**nformation **VI**sualization of **A**gent **s**ystems*) allows non-computer scientists to* visually *model, execute and evolve a multi-agent simulation without programming a single line of code. This visual programming paradigm allows for greater dispersal of MASS beyond the computer science field.*

1. Introduction

Multi-Agent Simulation technology has long belonged to the artificial intelligence, robotics, and agent communities. Multi-agent simulation systems (MASS) are desirable since they can exhibit qualities such as emergence, self-organization, learning and adaptive behavior thus enabling the modeling and understanding of more real world complex application domains [1]. However, the only way to define and run a simulation has been to program in a language such as Smalltalk, C++, and Java. This restriction constrains MASS technology to the computer science industry even though it has application across all disciplines (e.g., economics, social sciences, finance, etc.). Our goal is to bring this technology to those disciplines that can benefit greatly from the strengths of MASS.

With that goal in mind, our domain independent tool, DIVAs (***D**ynamic **I**nformation **VI**sualization of **A**gent **s**ystems*) allows non-computer scientists to *visually* model, execute and evolve a multi-agent simulation without programming a single line of code. This visual modeling paradigm allows for greater dispersal of MASS beyond the computer science field as well as making the modelers of these simulations more productive in the sense that simulation models can be created quicker and modifications to the behaviors can be made at runtime without stopping the current execution or creating a new model.

In Section 2, we discuss the motivation for our tool in more detail. In Section 3, we discuss what it takes to define the environment. In Section 4, we present our results and demonstrate the tool's usage to visually specify the environment. Finally, in Sections 5 and 6, we discuss related works, summarize our work and discuss future direction.

2. Problem

To best understand the problem we are solving, it is beneficial to discuss MASS in more detail.

In our work, we consider a specific category of Multi-Agent System, the Agent-Environment System (AESs) [2–4] on which to base our MASS tooling (see Figure 1).

We define an Agent-Environment System as a system composed of a) a set of interacting social-agents, b) a distinct open environment in which these agents are situated, and c) a mechanism for social-agent/environment interactions. A social-agent is a software entity which 1)is driven by a set of tendencies in the form of individual objectives; 2) can communicate with other agents; 3)possesses resources of its own; 4) acts in an environment; 5) has a partial representation of this environment; 6) reacts to external events; 7) possesses skills and can offer services. A cell manager-agent is a software entity which 1)is driven by a set of tendencies in the form of individual objectives; 2) can com-

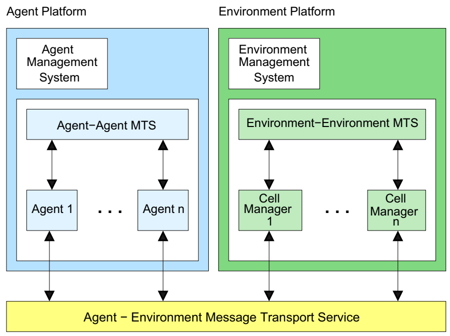


Figure 1. A General AES Architecture

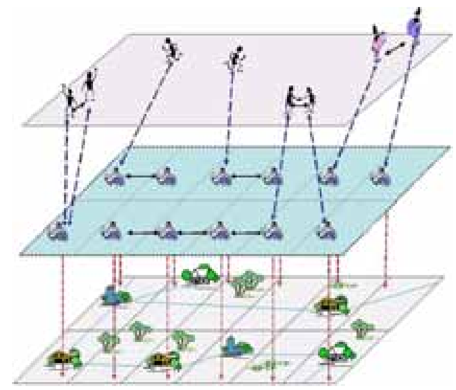


Figure 2. AES layers

municate with other environment entities; 3) possesses resources of its own; 4) has a partial representation of the agent population; 6) reacts to external events; 7) possesses skills and can offer services.

One fundamental idea of AES is that the environment is open (i.e., inaccessible, non-deterministic, dynamic and continuous) [5], and totally decoupled from the agents. The environment is considered to be a "living" entity that can be partitioned into easily manageable cells. Cells represent a division of continuous space. They are not necessarily contiguous and, hence, can form a network. This partitioning becomes critical when simulations become quite large (e.g., supporting thousands of agents). Cell Managers are design entities which are responsible for supporting the distributed and inaccessible feature of open environments [6–8]. Responsibilities include 1) managing its portion of the environment (i.e., cell); 2)informing other managers of any change that may affect their cell; and 3) informing its embedded agents of the latest state of the cell.

Additionally, because we are interested in visualizing location information at various *levels* of abstraction, the environment is defined as a hierarchy. The information contained in cells and location graphs is refined as the level of granularity decreases. Therefore, in order to manage the environment as a whole, each cell manager has to manage its connections with the cells it is linked to at the same level (i.e., horizontally), as well as at the upper and lower levels (i.e., vertically).

Using these definitions, one example of an AES is a social, geographically-based simulation system that consists of the population of social entities (i.e., agents) and their interactions and, the cell manager entities and their interactions as well as the interactions between the agents and their cell manager entity (see Figure 2).

For the sake of clarity, we will illustrate our implementation of an AES, DIVAs, using the example of a social simulation system where social entities represent "people" and the environment represents the world. The underlying structure of the environment is a graph where nodes represent locations and edges represent pathways between locations. People move in the environment using the graph. We assume that the simulation system is intended to show how the population and the environment react to external events. We also assume that the user requires a flexible architecture where agents and environment are decoupled so that if the agent requirements change, it will not affect the environment and vice versa.

Therefore, given the complexity of designing and implementing an AES, it is necessary to provide the user with a tool that removes the complexity of MASS thereby allowing the user to focus on defining the behaviors and properties that compose the model. In the next Section, we will focus on the steps required to define the environment for DIVAs.

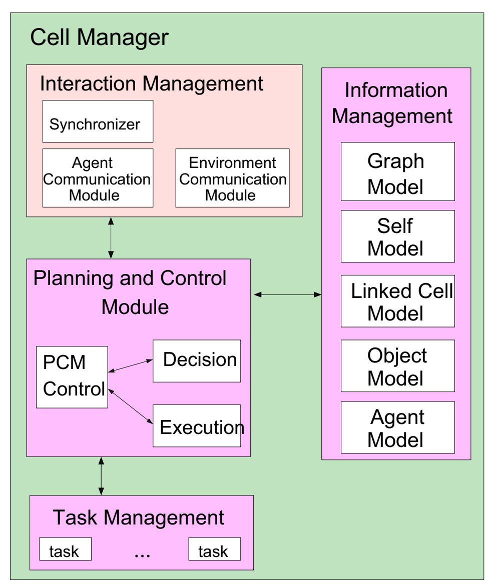


Figure 3. Cell Manager Architecture

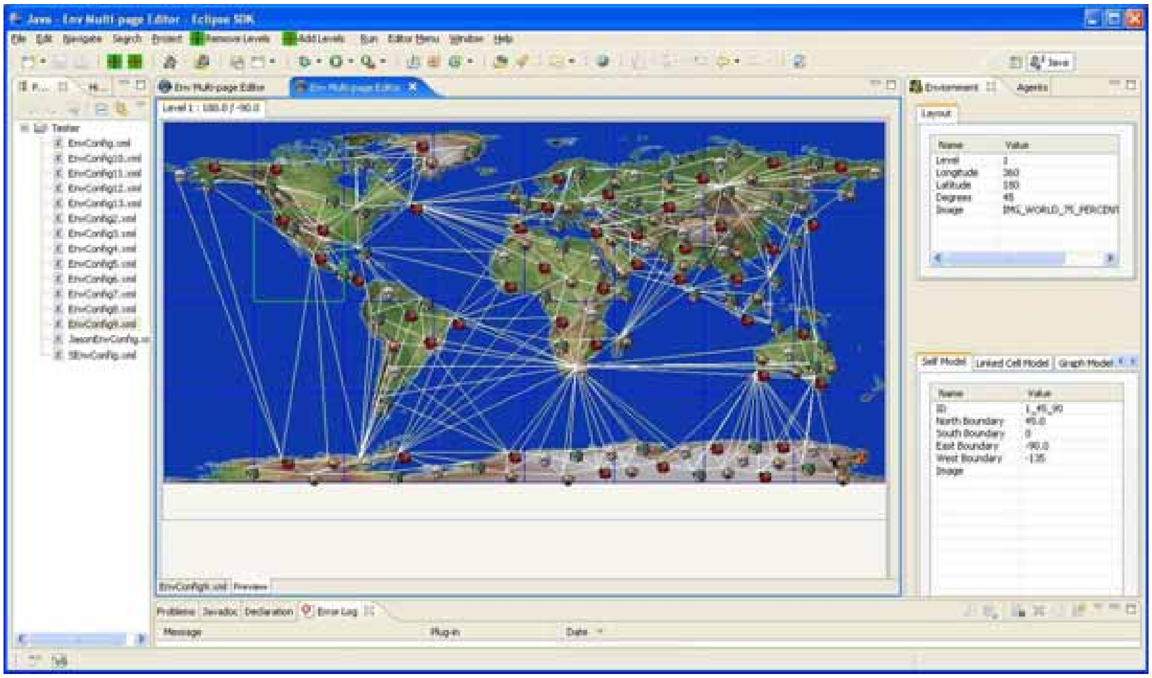


Figure 4. DIVAs Tool, Environment Specification View

3. Environment Specification

Given that, for DIVAs, the environment is a first-class entity which has responsibilities and tasks, and given that the environment is decoupled from the agents, it is reasonable to conclude that specification of the environment is more complicated than in traditional MASS.

Consider the cell manger's architecture in Figure 3. While the inherit framework of the AES provides the capability to communicate (i.e., *Interaction Management*), a rudimentary brain (i.e., *Planning and Control Module*) as well as simple services (i.e., *Task Management*), the definition of the information that enables the decisions made by the cell manager during the course of simulation execution is left to the simulation modelers.

Therefore, a simulation modeler must program classes that are responsible for *Information Management*. Such data includes: information about the map's nodes and edges (i.e., *Graph Model*); the cell manager's identity which includes the cell's relative position and communication handle (i.e., *Self Model*); the listing of other cell managers that share paths with the cell manager (i.e., *Linked Cell Model*); the listing of all passive objects within the cell manager's boundaries (i.e., *Object Model*); and, the listing of all agents currently within the cell manager's boundaries (i.e., *Agent Model*).

These classes would be specific to a certain simulation instance. If parameters change, the classes must also change. This process limits the productivity in creating different simulation models. For example, if the graph model changed, a programmer must change the code and recompile. Hence, it follows that simulation modelers must be programmers which limits users of MASS to be within the computer science field.

It is our premise that, if the programming complexity is removed, productivity is increased and MASS can be utilized by a broader community that does not have to include computer scientists. In the next Section, we discuss how we address the inherit complexity in modeling environments for AES.

4. Visually Specifying the Environment

DIVAs Tooling has been designed and implemented so that anyone can define, execute and dynamically modify a simulation model. In order to make the tool accessible to non-scientists, we took care to remove all programming requirements from the user. Defining a simulation model is as simple as filling in the blanks of tables which represent different pieces of the cell manager and agent architectures. We built DIVAs as a plugin to the open-source eclipse platform ecosystem thereby increasing user familiarity by utilizing a development environment that many other end-user tools are built upon. Additionally, since we have based it on eclipse, we can produce a Rich Client Platform (RCP) application which gives it a look and feel which is much simpler than standard eclipse. An example of the tool is shown in Figure 4.

To define a simulation model, the user simply completes tabular information similar to the table shown in Figure 5 which represents the *Self Model* of a cell manager in the environment. The relationship between the actual implementation and the visual abstraction can be seen clearly in Figure 6.

In the following paragraphs, we will discuss these concepts in more detail and, in this brief case study, we demonstrate the ease in which a user can specify the environment for a geographical-based simulation.

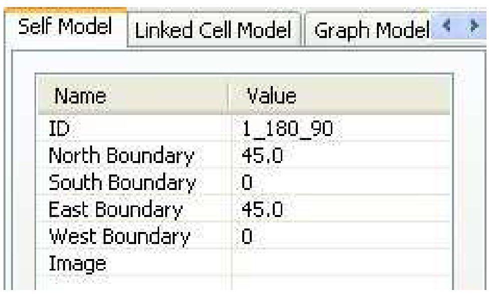


Figure 5. Specifying the Environment's Self Model

A user begins defining a simulation model by describing the overall map layout (i.e., Level 1). The dimensions of the map, in longitude and latitude, and the size, in degrees, of each cell must be defined. The user must also specify the image that will represent the level. So, by simply completing this data, the tool will produce a specification model in XML format that will be executed at runtime (see Figure 7). This process is repeated for each finer-grained level of abstraction that the user wishes to visualize.

For the environment, the items that must be specified correspond to the environment's architecture as discussed previously in Figure 3. These items map directly to the cell manager's Information Management modules as shown in Figures 3 and 8: Graph Model, Self Model, Linked Cell Model, Object Model and Agent Model.

Next, the user continues by defining each cell manager. As shown in Figure 5, the user must specify the image that will represent the area of the map that has been selected. All other information in the Self Model is populated by the tool. In this manner, all of the models in the cell manager's Information Management module are specified. This process continues for the Graph Model, where nodes and edges are defined and positioned, the Object Model where passive objects (e.g., hammer, nails) are defined and positioned on the map, and the Agent Model where agents are named and positioned on the map. The Linked Cell Model is inferred from the Graph Model and completed by the tool. The tool

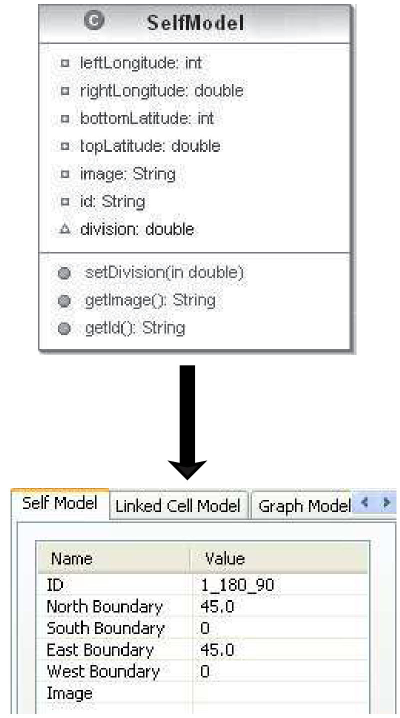


Figure 6. Correlation between implementation and visual modeling

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE environment-config>
<environment-config>
    <level id="1" longitude="360" latitude="180" division="45"
          image="IMG_WORLD_&%_PERCENT">
</environment-config>
```

Figure 7. Simulation Model snippet for the Environment Level Configuration

will produce the XML based simulation model and the simulator will execute it.

During execution, a user can chose to modify the model visually. For example, a node can be deleted from the graph and the reaction of the agents and other cell managers can be observed. This is extremely beneficial in order to observe adaptive behavior.

5. Related Works

A number of agent-based simulation systems have been described in the literature including MASON [9], Swarm [10], Repast [11], Ethos [12], Ascape [13], AgentSheets [14], XRaptor [15], CORMAS [16], SimAgent [17], and SeSAm [18]. Most of these systems are often restricted to

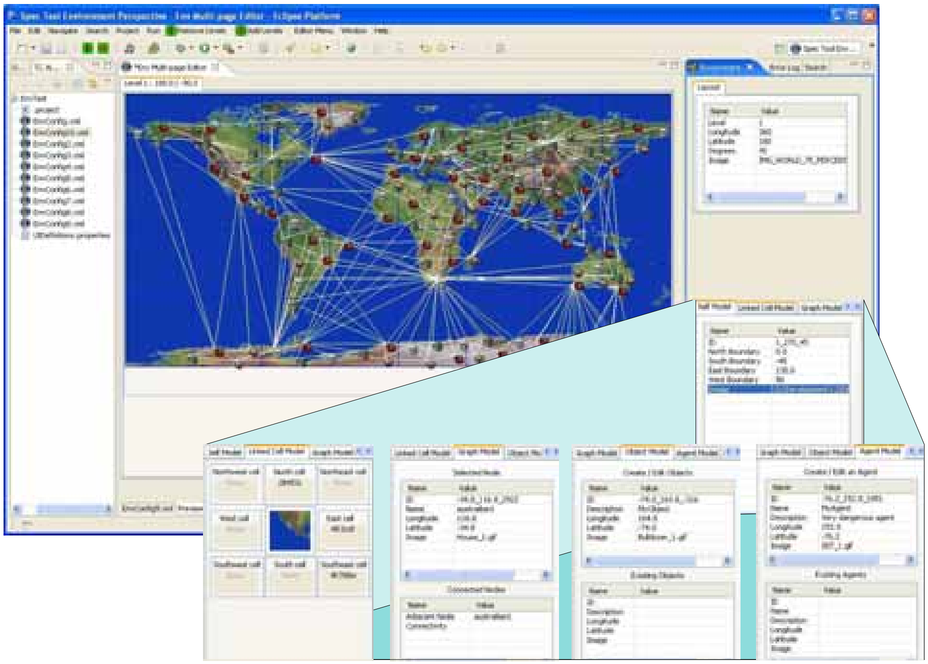


Figure 8. Modeling the DIVAs environment architecture

specific applications, and are difficult to adapt to other domains. In addition, these systems do not clearly separate between agents and environment responsibilities. Researchers agree that this is a substantial weakness in MAS and have recently organized workshops dedicated to the exploration of environments as first class entities [19–21].

We restrict further analysis to those MASS that best compare to our work such as NetLogo [22], MASON [9] and JACK [23, 24]. All of these requires programming to define a simulation instance. NetLogo has its own language, Mason has its own java extension API, and JACK is a mix of its own language and java.

None of the three MASS have models that can be changed during simulation execution. With NetLogo, deleting a node requires removing the node's co-ordinates from the code and then restarting the simulation. In MASON and JACK, deleting a node requires a re-compilation and re-execution of the simulation model.

To the best of our knowledge, none of these MASS provide a tool to visually specify the simulation model for the environment.

6. Conclusion

In this paper, we have described an AES framework and tool which abstract the complexities of MASS modeling. Our tool, DIVAs allows non-computer scientists to *visually* model, execute and evolve a multi-agent simulation without programming a single line of code. This visual modeling paradigm allows for greater dispersal of MASS beyond the computer science field across many different disciplines. Additionally, by simplifying the definition of simulation models, modelers will become more productive (e.g., defining and executing different models quickly).

Future work includes completion of the agent specification module and providing a additional module to allow the user to modify environment behavior at runtime.

References

[1] P. Davidsson, "Multi agent based simulation: Beyond social simulation." in *MABS*, 2000, pp. 97–107.

[2] R. L. Steiner, G. Leask, and R. Z. Mili, "An architecture for MAS simulation environments," in *Proceedings of 2nd Workshop on Environments for Multi-Agent Systems*, Utrecht, Netherlands, July 2005, also in Lecture Notes in Computer Science 3830. Springer Verlag (2006) 50-67.

[3] R. Steiner, G. Leask, and R. Z. Mili, "An architecture for MAS simulation environments," in *Proceedings of 2nd Workshop on Environments for Multi-Agent Systems, Lecture Notes in Computer Science*. Springer-Verlag, 2006, vol. 3830, pp. 50–67.

[4] R. Z. Mili, E. Oladimeji, and R. Steiner, "Architecture of the DIVAs simulation system," in *Proceed-*

ings of Agent-Directed Simulation Symposium ADS06, Huntsville, Alabama, April 2006.

[5] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*. Prentice Hall, 1995.

[6] R. Mili, G. Leask, U. Shakya, R. Steiner, and E. Oladimeji, "Architectural design of the DIVAs environment," in *Proceedings of 1st Workshop on Environments for Multiagent Systems*, New York, USA, July 2004, pp. 57–66.

[7] R. Z. Mili, R. Steiner, and E. Oladimeji, "DIVAs: Illustrating an abstract architecture for agent-environment simulation systems," *Multiagent and Grid Systems, Special Issue on Agent-oriented Software Development Methodologies*, vol. 2, no. 4, pp. 505–525, 2006.

[8] R. Steiner, "Engineering open environments for multi-agent simulation systems," Ph.D. dissertation, Dept. of Computer Science, University of Texas at Dallas, 2006.

[9] S. Luke, C. Cioffi-Revilla, L. Panait, and K. Sullivan, "Mason: A new multi-agent simulation toolkit." in *Proc. of the SwarmFest Workshop*, Michigan, USA, 2004.

[10] M. Nelson, R. Burkhart, C. Langton, and M. Askenazi, "The swarm multi-agent simulation system," in *Proc. Workshop on The Object Engine - Object-Oriented Programming Systems, Languages, and Applications (OOPSLA) '95*, 1994.

[11] N. Collier, "Repast: An extensible framework for agent simulation," Social Science Research Computing, University of Chicago, IL, USA," Technical report, 2002.

[12] J. Simao and L. M. Pereira, "Ethos: A mas framework for modelling human social behavior and culture," in *Proceedings of 4th Workshop on Agent-Based Simulation*, Montpellier, France, 2003.

[13] M. T. Parker, "What is ascape and why should you care." *Journal of Artificial Societies and Social Simulation*, vol. 4(1), 2001.

[14] A. Repenning, A. Ioannidou, and J. Zola, "Agentsheets: End-user programmable simulations." *Journal of Artificial Societies and Social Simulation*, vol. 3, 2000.

[15] P. Mossinger, D. Polani, R. Spalt, and T. Uthmann, "Xraptor - a synthetic multi-agent environment for evaluation of adaptive control mechanisms," F. Breitenecker and I. Husinsky, Eds. Elsevier: EUROSIM '95, 1995.

[16] F. Bousquet, I. Bakam, H. Proton, and C. L. Page, "Cormas: Common-pool resources and multi-agent systems," in *Lecture Notes in Artificial Intelligence v. 1416*. Springer Verlag, 1998, pp. 826–838.

[17] A. Sloman and R. Poli, "SIM AGENT: A toolkit for exploring agent designs," *Proc. Intelligent Agents Vol II (ATAL-95). Lecture Notes in Computer Science*, vol. 1037, pp. 392–407, 1996.

[18] F. Klügl, R. Herrler, and C. Oechslein, "From simulated to real environments: How to use sesam for software development," in *MATES 2003, LNAI 2831*, M. Schillo, Ed. Springer-Verlag, 1994, pp. 13–29.

[19] D. Weyns, H. V. D. Parunak, and F. Michel, Eds., *Environments for Multi-Agent Systems, First International Workshop, E4MAS 2004, New York, NY, USA, July 19, 2004, Revised Selected Papers*, ser. LNAI, vol. 3374. Springer, 2005.

[20] D. Weyns, H. V. D. Parunak, F. Michel, T. Holvoet, and J. Ferber, "Environments for multiagent systems: State-of-the-art and research challenges," in *Environments for Multi-Agent Systems, First International Workshop, E4MAS 2004, New York, NY, USA, July 19, 2004, Revised Selected Papers*, ser. LNAI, D. Weyns, H. V. D. Parunak, and F. Michel, Eds., vol. 3374. Springer, 2005, pp. 1–47.

[21] D. Weyns, H. V. D. Parunak, and F. Michel, in *Proceedings of 2nd Workshop on Environments for Multi-agent Systems*, Utrecht, Netherlands, July 2005.

[22] S. Railsback, S. Lytinen, and S. Jackson, "Agent-based simulation platforms: Review and development recommendations." in *Simulation*, 2006.

[23] A. H. Howden, Ralph Rnnquist and A. Lucas, "Jack summary of an agent infrastructure," in *Proc. of 5th International Conference on Autonomous Agents*, 2001.

[24] W. Howden, R. Ronnquist, A. Hodgson, and A. Lucas, "JACK intelligent agents," Available over the internet. Accessed October 2006, http://www.agent-software.com/shared/home.

Proceedings
International Workshop on Distant Education Technology (DET 2008)

Co-Editors

H. S. Chen, *National Taiwan University, Taiwan*
Tim Arndt, *Cleveland State University, USA*

Communities of Practice in Second Life: the case of Eclipse Italian Community

P. Maresca *, G. Scarfogliero*, L. Sorrentino*

**Dipartimento di Informatica e Sistemistica, Università di Napoli "Federico II"*
Via Claudio 21, Napoli
paolo.maresca@unina.it

Abstract

The Eclipse Italian Community recently had its second birthday. It has intent to develop synergies and projects around Open Source and so around a Community of Practice that aims to spread, improve and use Eclipse Platform. Many projects are ongoing actually, but one of most interesting challenges is represented by the communication tools that a Community of Practice can and must use to grow up. At the beginning of this year the Eclipse Italian Community decided to invest its own energies in testing the community activities in Second Life: has been founded the Eclipse-IT-SL project. A new workgroup has been created to study the problematic and the activities linked to this communication paradigm. This paper shows the state of the art of the new project of Eclipse Italian Community in Second Life. The test building is temporary hosted in IBM Eternal City island.

1. Introduction

In last 10 years the technological progress has radically changed the way of living in modern society: several are the changes it brought about many fields, but surely one of most important is related to development and continuous empowerment of global communication networks.
The Internet arrival, the broadband's diffusion, the mobile connectivity, have a real revolution in way of communicate and interact of people. Each one, in fact, by using his own computer, can communicate with wide world people and share an immense amount of information, whenever he wants. Moreover, people can add his own knowledge by publishing more contents. E-mails, discussion forums, newsgroups, mailing lists, are just some of most famous tools for modern communication.
The use of these technologies contributed to the development and the growth of many communities of practice, because only by using these tools it's possible to manage the interaction and the knowledge sharing between people belonging to different geographical areas, allowing them to stay in their workspaces or their houses, without moving to reach someone other. This possibility to interact and communicate, by using advanced tools for the diffusion of all multimedia data, allows a rapid expansion of Communities of Practice that, some times, can reach national extension or international extension if coupled with the use of English language.
The Eclipse Italian Community is a perfect example of this kind of Community of Practice: it was born in 2006, from the willingness of Italian Universities and some Companies involved in Eclipse-based software development, that, to improve their productivity, decided to share their knowledge and skills on eclipse usage and development. Thanks to Internet technologies and to the founder members perseverance, the Eclipse Italian Community has reached a National extension.
The complex web platform, hosted in University of Naples Federico II, allows to community users to manage their own avatars, where everyone can also write information about personal knowledge. It also allows users to interact with other members thought the use of many discussion forums, of private messaging or of emails, of development groups, of chats or of common repositories.
The Eclipse Italian Community has a web development team always looking for new interaction forms, able to improve the communication qualities between members, allowing people to exchange their knowledge in a more informal way, realizing a more pleasant and less alienating communication. So Eclipse Italian Community couldn't sit by one of biggest phenomenon of internet-based interaction technology of last two years: Second Life
Second-Life is the biggest virtual world ever conceived; a product distributed in 2003 by Linden Labs, it joins recreational aspects of a massive-multi-player videogame with social and economics ones. In this big virtual world, each user can create a personal avatar, a kind of virtual alias for himself, with which he can explore lands and interact with other people. Composed of islands, Second Life allows, using teleport, to move from a place to another and eventually to create own house or company office after buying a piece of land. SL in fact recognizes copyrights on created objects, so a user can sell his

creation receiving payment in a virtual coin, called Linden Dollar (L$) that can be converted in real US Dollars. The great peculiarity of this virtual world is that it can be expanded by the same users, buying new islands and building on them.
The potentialities of SL in the didactical area and in the diffusion of knowledge are evident: using a virtual world, in which each person can identify himself in a virtual avatar, joins the entertainment and the possibility of extend his own knowledge. In the same time it can involve new generations, very skilled in using complex gaming systems, allowing people to interact in a more informal way, realizing a more pleasant and less alienating communication.
So the web development team of Eclipse Italian Community decided to follow the example of many American Universities and several Organizations, that saw in Second Life an investment, and decided to establish an headquarter of our organization in Second Life where members can organize meetings and show the community developed products.
This development project has been named Eclipse-IT-SL.

In next section will be illustrated strategic targets of Eclipse Italian Community virtual headquarter building.
In last section will be shown the results and future development of Eclipse-IT-SL and Conclusions.

2. Eclipse Italian Community in Second Life

Building a headquarter for Eclipse Italian Community in Second Life, implies in primis the choice of a building that represents the organization and, in the meantime, that is functional to all community activities. Bearing in mind that SL is a videogame, is simple imagine that it allows realizing every kind of abstract building.

We looked for a kind of building, that, immersed in the territory, was of big visual impact and that represented, as well, the soul and aims of Eclipse Italian Community. The building must be erected on "IBM Eternal City" island, in which are already present very beautiful buildings. So we have chosen a lighthouse.

Allegorically a lighthouse can be intended as a point of reference that with his huge structure is erected sovereign and stationary in the surrounding chaos, place of certainties and order. The lighthouse (Fig. 1), in fact, is built by men pro men, and, in our case, hides inside a very comfortable place, where each one can find answers to his own questions and offer his know-how to the community. Supreme result of human hard-working, it well symbolizes the spirit of sharing knowledge and of the common growth of Eclipse Italian Community.

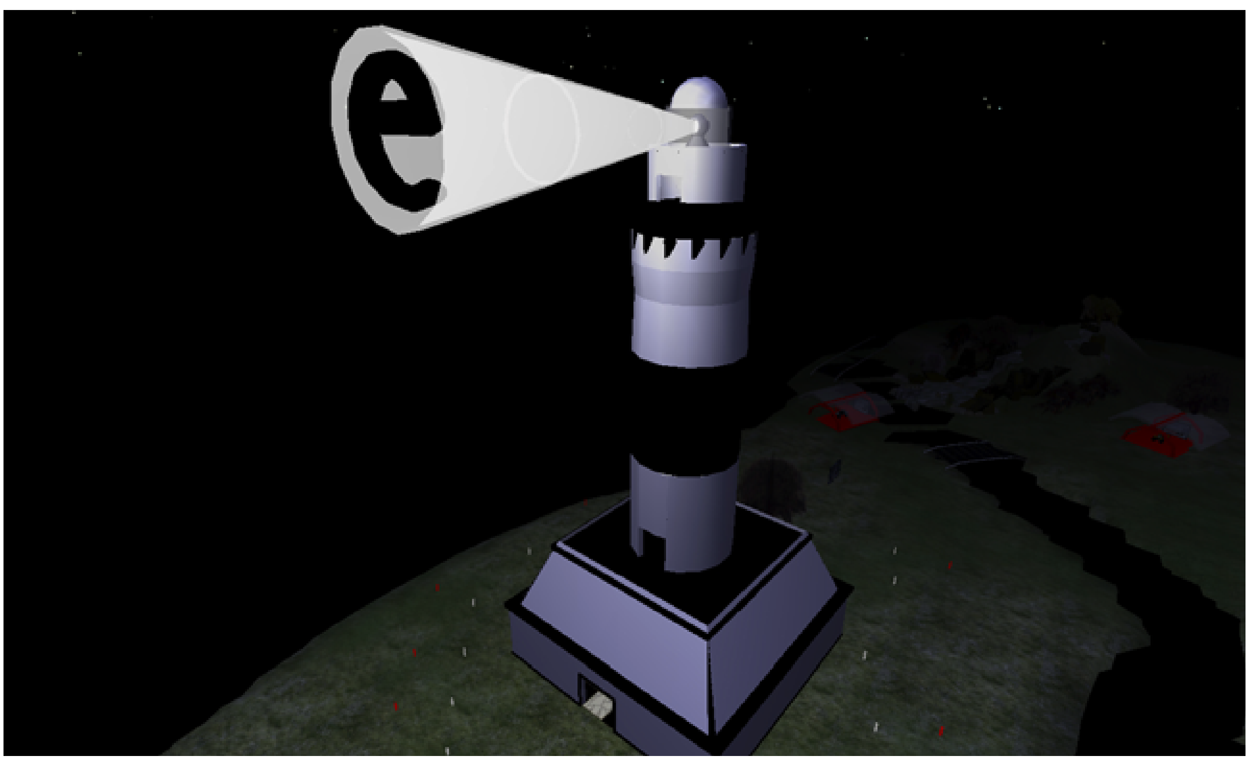


Figure 1. The Lighthouse of Eclipse Italian Community.

Figure 2. The Lighthouse base.

Figure 3. The Lighthouse tower.

The structure is developed in height ad is composed by two parts:

- The lighthouse base is a two-floor building (Fig. 2).
- The lighthouse tower has a cylindrical shape and terminates, on the top, with a circular room and the lantern.

The simple and essential shapes give geometrical rigour to the whole structure, that appears to an observer solid and robust; lack of windows and of any kind of opening along the tower, accentuates this sensation, and adds, however, a further meaning: the solid construction wraps knowledge among its walls, offering it protection from outside chaos.

The large entrances in the lighthouse base and in the circular room at the summit of the tower indicate, in the meantime, the structure (and the community) opening to the external world: the Man, knowledge keeper, is invited to enter and to take part to the virtual life of the Community.

The lighthouse base is composed by two rooms placed on two different floors:

- the hall (Fig. 4) is a wide square room, destined to host the acceptance and the info point of the community.

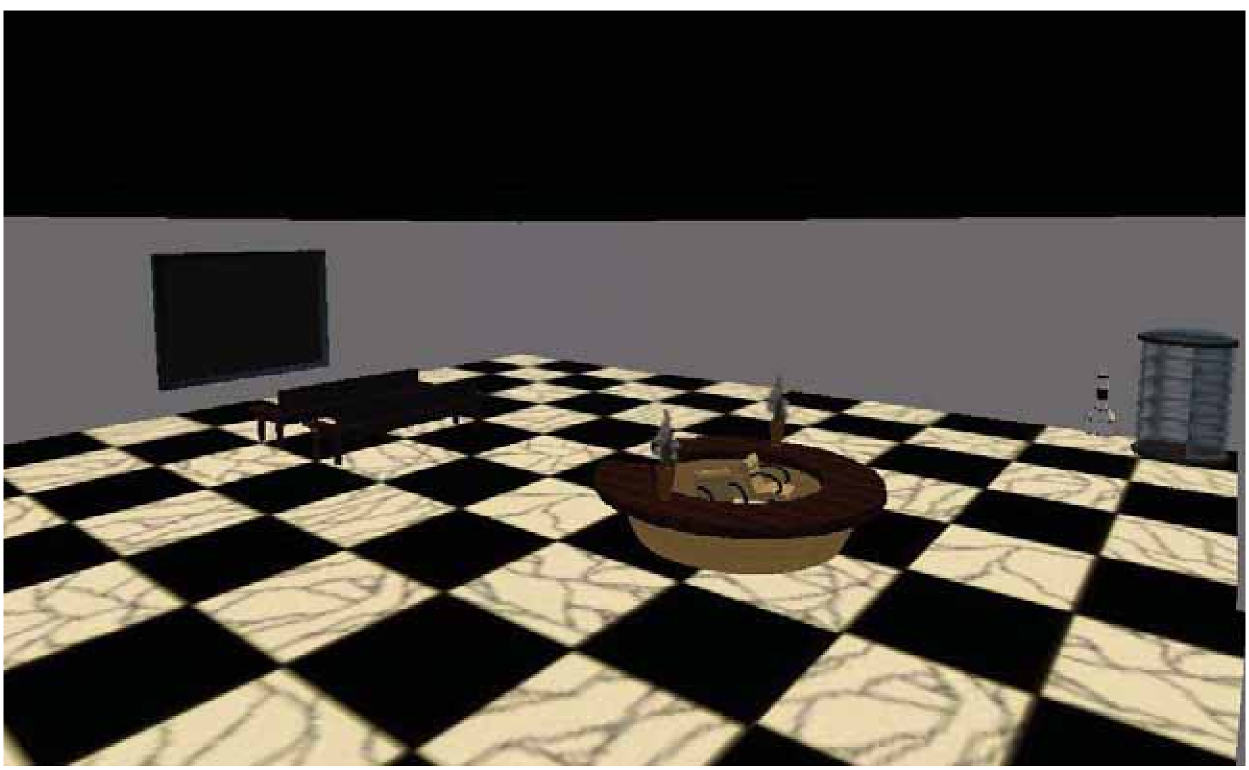


Figure 4. The hall of Eclipse Italian Community Lighthouse.

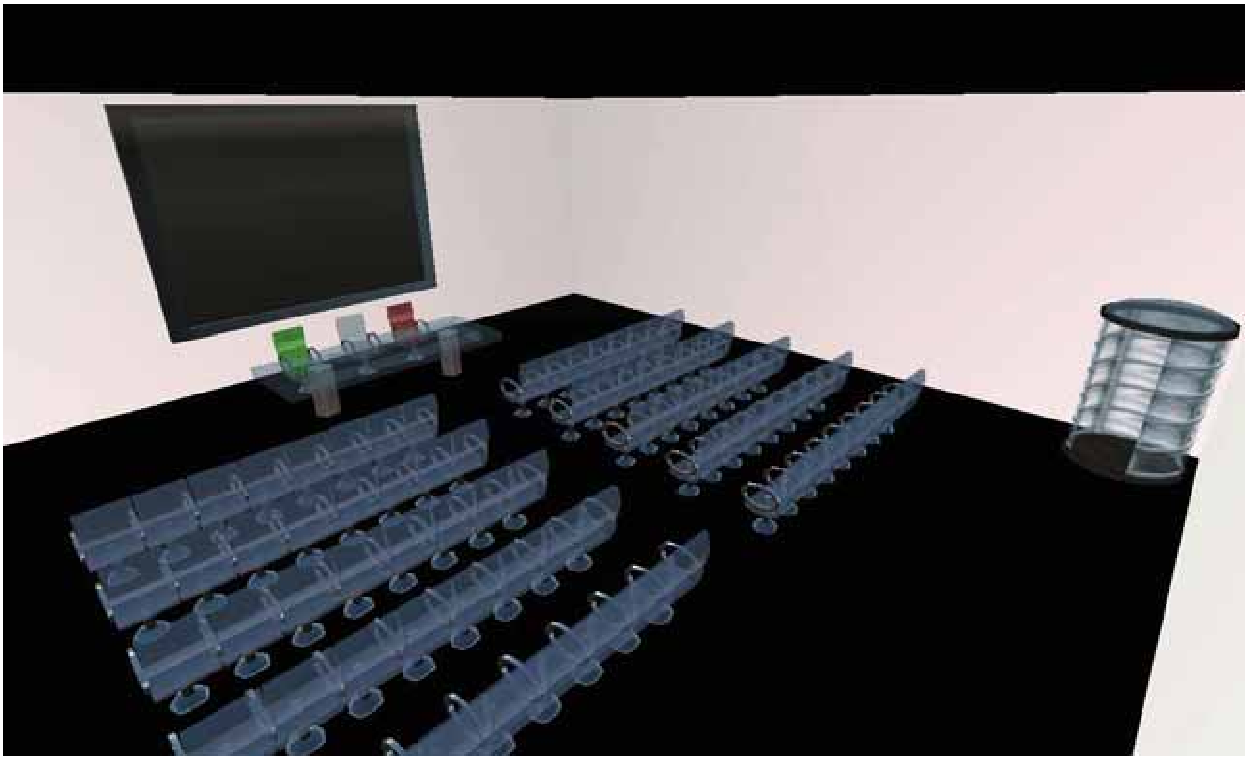


Figure 5. The conference room oe Eclipse Italian Community lighthouse

- The other room is a bit smaller than the hall, and hides inside a smaller area that will be used as conference room (Fig. 5).

In the conference room Eclipse users can have meetings and attend seminars or courses.

The tower is crossed, for almost all its height by a smaller glass cylinder, coaxial to the bigger one. It's used to have access to all the tree floors of the tower, flying inside of it. From the last floor (Fig. 6), using an elevator, you can reach the circular room, under the lantern.

The circular room (Fig. 7) will be used by the avatars of Community leaders to have private meetings.

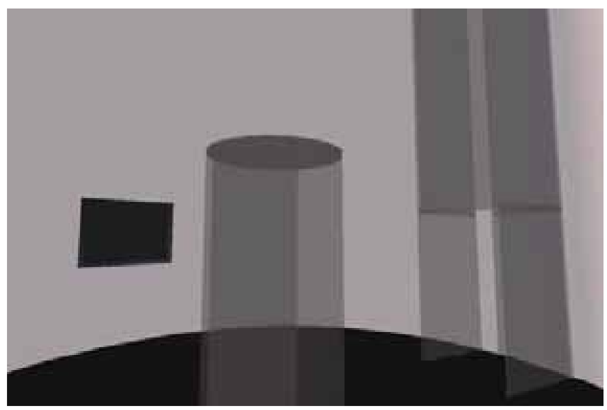


Figure 6. The last floor of Lighthouse.

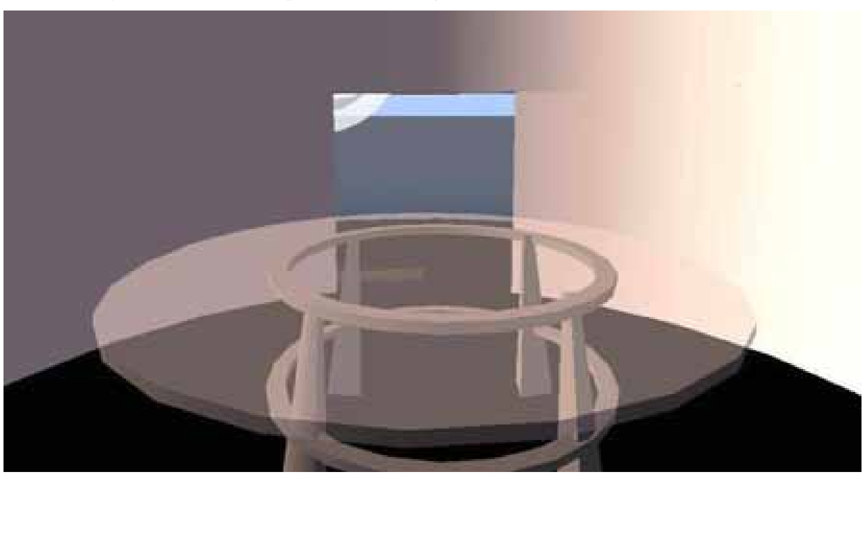


Figure 7.
The circular room of the Lighthouse.

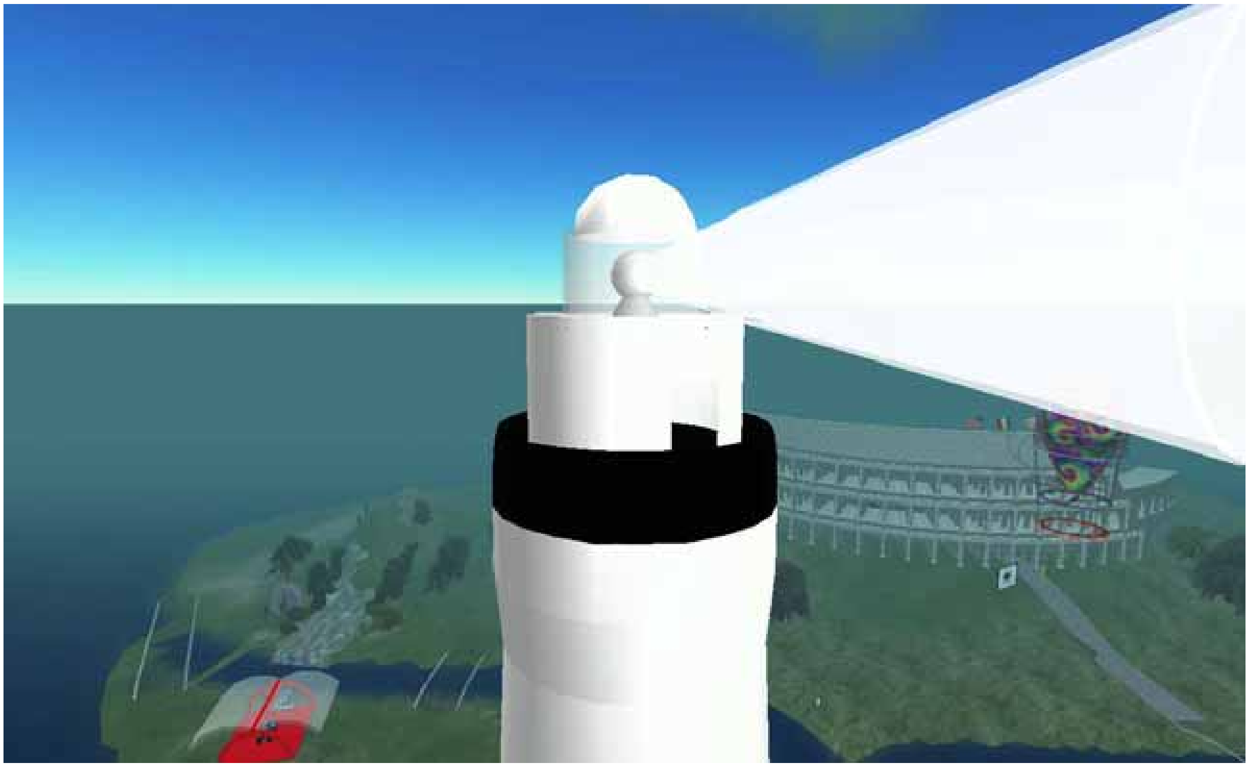


Figure 8. The lighthouse lantern.

The above area, delimited by transparent walls, contains the lighthouse lantern (Fig. 8): the light of knowledge, which intensity increases under the impulse of the community growth, drives peregrine looking for answers.

The cylindrical shape idea for the lighthouse tower was born form a focusing reflection on the global message that the construction must communicate to observers. The circumference is the geometrical shape from which derives the cylinder. It is a shape in which you can't distinguish the beginning from the end; it is the symbol of the principle from which everything has beginning and to which all came back, characterised by hostile contrast between expansive tendency and the harmony it communicates. So, the knowledge, following the same rules, expand itself to external world, in a controlled and harmonic way, keeping from this process new "learning", using it to feed and to renew itself, so to start a continuous improvement cycle.

3. Result analysis and Conclusions

The Eclipse-IT-SL project involves a workgroup that finished the external project of buildings and that is already building internal furniture to receive the avatars of people interested in this initiative and to have 3D-Community activities.

The lighthouse is provisionally placed in IBM Eternal City island.
Future developments consist of reach a communication paradigm for Eclipse Italian Community also through Second Life that allows the community to welcome more support requests. This will allow the community to get internationalized and most known with its projects beyond real space frontiers.

Bibliography

[1] Eclipse Italian Community, http://eclipse.dis.unina.it/
[2] IBM Academic Initiative, www.ibm.com/university/academicinitiative.
[3] Second Life, www.secondlife.com
[4] Second Life Wiki, http://wiki.secondlife.com/
[5] Second Life Italia, http://www.secondlifeitalia.com/

Acknowledgements

The Eclipse Italian Community thanks IBM Italy to have hosted the test project Eclipse-IT-SL on IBM Eternal City Island

Eclipse and Web Services

L. Stanganelli *, P. Maresca*

*Dipartimento di Informatica e Sistemistica, Università di Napoli "Federico II"
Via Claudio 21, Napoli
paolo.maresca@unina.it

Abstract

Web Services are the most promising SOC-based technology. They use the Internet as the communication framework and open Internet-based standards, including Simple Object Access Protocol (SOAP) for transmitting data, the Web Services Description Language (WSDL) for defining services, a lot of visual tools for representing web services (ex. Protège), and the Business process Execution Language for Web Services (BPEL4WS) for orchestrating services. The Eclipse platform is the ideal platform to sustain a WS approach given that SOC group together many intricately concepts, protocols and technologies that originate in a wide range of disciplines.

Many Eclipse projects are ongoing actually in the Eclipse platform, but one of most interesting challenges is represented by the logical composition of web services. This paper shows the state of the art of this projects, sustained from Eclipse Italian community, and shows a tool developed into eclipse environment that aims to compose web services in a logical manner. The same approach will be used in e-learning composing, in a logical manner, educational artefact, in order to satisfy a request.

1. Introduction

The service composition plane encompasses roles and functionality for aggregating multiple services into a single composite service. Resulting composite services can be used as basic services in further service compositions or offered as complete applications and solutions to service clients. Service aggregators accomplish this task and thus become service providers by publishing the service descriptions of the composite service they create. The aggregators also enforce policies on aggregate service invocation.

Currently, developers widely use the terms "orchestration" and "choreography" to describe business interaction protocols that coordinate and control collaborating services.

Orchestration [1] describes how services interact at the message level, including the business logic and execution order of interactions under control of a single end point.

It is an executable business process that can result in a long-lived, transactional, multistep process model. With orchestration, one of the business parties involved in the process always controls the business-process interactions.

Orchestration is achieved via BPEL4WS and other XML based process standard definition languages. Choreography is typically associated with the public (globally visible) message exchanges, rules of interaction, and agreements that occur between multiple business-process end points rather than a specific business process executed by a single party. Service choreography is achieved via the Web Services Choreography Description Language (WS-CDL), which specifies the common observable behavior of all participants. This sharp distinction between orchestration and choreography is rather artificial, and the consensus is that they should coalesce in a single language and environment. Some of the most notable research challenges for service composition in the near future include the following:

(i) Composability analysis for replaceability, compatibility, and process conformance, (ii) dynamic and adaptive processes and (iii) QoS-aware service composition, (iv) business-driven automated composition.

As far as the composability it is worth the pain of outline like this Service conformance ensures a composite service's integrity by matching its operations with those of its constituent component services. It imposes semantic constraints on the component services and guarantees that constraints on data that component services exchange are satisfied. Service conformance comprises both behavioral conformance as well as semantic conformance. The former guarantees that composite operations do not lead to spurious results, while the latter ensures that they preserve their meaning when composed and can be formally validated. Obviously the progresses of the composability are directly connected to the ability to compose services in logical way and from the possibility to have one extensible platform. The Eclipse platform is the ideal platform to sustain a WS approach given that SOC group together many intricately concepts, protocols and technologies that originate in a wide range of disciplines.

In fact Eclipse sustain well SOA (Service Oriented Architecture) that is the architecture used to design WS.

The future objective is that one to arrive to the point in which the Web Services are accessible automatically from the computers. With SOA an architectural model is identified it supports the use of services that satisfy the demands for the customers, rendering possible the uses of single applications like components of processes.
The following paper is organized in this way: chapter 2 is devoted to the semantic web discussion, chapter 3 speaks about logic composition and web services and shows the eclipse architecture and the environment that sustain the WS composition. Finally chapter 4 state the conclusion and the future development of WS project.

2. Semantic Web

In the last years migration from traditional Web towards the semantic Web is being assisted; the first one meant like a collection of documents and where the search of these happens by means of not optimizing search engines, the second instead represents the evolution of Internet from simple system of communication and document recovery, to "an intelligent" system in which the information will be comprised from specific software in a position to cooperating and assisting the customer in complicates task [2,3].
In order to realize the semantic Web the key elements are the knowledge representation, the agents and the ontology. The knowledge representation, comes faced through an more levels architecture constituted of: data, metadata, ontology (that is the semantic representation of data and metadata through specific languages). These levels are placed side by side from a technological structure whose main elements are constituted from the XML, RDF, RDFS and OWL like it is evident from figure 1 [4].

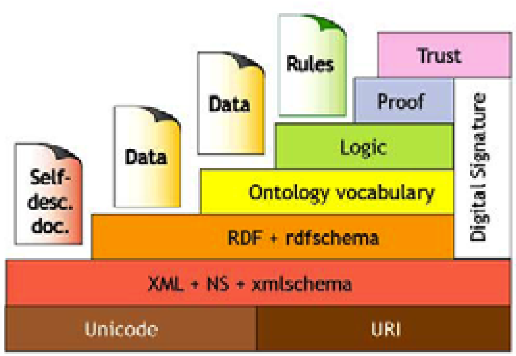


Figure 1. Semantic layer cake

The ontologies describe the semantic of the concepts and not the structure of the data, they expresses in formal way the relations and ties between the concepts. Ontologies are comprehensible from persons and machine are the base of the architecture and the applications of the Semantic Web.
OWL is constructed on RDF and RDFS and uses an XML based syntax. The "real" architecture of Web Service, that is constituted from a set of xml standard, embraces: SOAP(Simple Objet Access Protocol), WSDL (Web Service Definition Language) and UDDI (Universal Description Discovery & Integration). The interaction between different software within WS happens in this way: who wants to approach to a determined service will have to only obtain relative document WSDL, by means of the access to public XML registries (making itself part of the WS specification), to extrapolate the information for the understanding of the access modalities, and to this point you will be in degree, through SOAP, to communicate with it. Naturally in order to make this we need an hardware/software independent and extensible architecture. Eclipse is an optimal candidate in how much beyond to being easy extensible and fines platform concur to make to interact in various simple way software above all when there are many concepts and various technologies to put together.

3. Logic Composition methodology for WS their architecture and implementation

Composition of WS has a principal interest: to realize web services by using composition. More generally speaking we can have two WS composition type: (i) physical composition, (ii) logical composition. For the sake of the brevity we don't mention physical composition here and the reader can find some useful references in the bibliography. Here we want to speak about composition of web services that is composed of 4 steps as showed [5]:

1. *Service Representation*: Representing the available services and their capabilities.
2. *Requirements Specification*: Specifying the desired functionality of a new service.
3. *Composition*: Constructing a composition of available services that provides the desired functionality.
4. *Composite Service Representation*: Representing the new composite service and its capabilities so that it can be programmatically deployed, discovered and invoked.

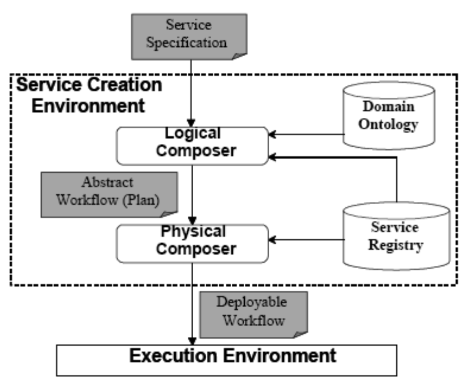


Figure 2. WS composition step

The composition step are summarized in figure 2 generally speaking composition type can be defined as:

1. Logical Composition: This phase provides functional composition of service types to create new functionality that is currently not available.
2. Physical Composition: This phase enables the selection of component service instances based on non-functional (e.g. quality of service) requirements, that would then be bound together for deploying the newly created composite service.
The literature until today has been interested to the physical composition of the WS by means of the use of WSDL [6], BPEL [7,8,9,10,11], but here the attention is wanted to be placed on the logical composition that use the ontology and the matching algorithm that it allows to select the services advertisement that more are approached the requested service. The matching algorithm for realizing the composition is that one of Paolucci [12,13,14]. In order to catch up our objective we have executed the following steps:

- to define the dominion by means of language owl;
- to describe the services using OWL-S [15];
- to import the into Eclipse Environment;
- to use Jena Libraries, under Eclipse platform, in order to query the dominion;
- to carry out the composition of the services.

The dominion is constituted from a set of resources and has been described by OWL language [16], in order to make that *protégé* tool has been used. After to have described the dominion both the advertised and requested services has been described by using *protégé* tool. Then by means of the jena libraries we have imported and queried our ontology in Eclipse environment (see fig. 3). The same Eclipse environment [17,18] hosts the java matching algorithm used to searches the services that more are approached the requested service and then they come composed so as to realize the demanded service. It's worth while to outline that Eclipse environment manage also the interaction with a DB2 data base in order to store and retrieve both the requested and advertised services.
In order to realize the composition of web services [19,20] beginning from a initial specification an approach based on the hypothesis that such requested services refer to well defined and described dominions by means of ontologies [21,22,23] has been used.
The environment is completely organized into an eclipse platform. Particularly the algorithm has been implemented in Java languages using also Jena libraries in order to import rdf/owl files generated from other environments (Protégé) as shown in figure 4.

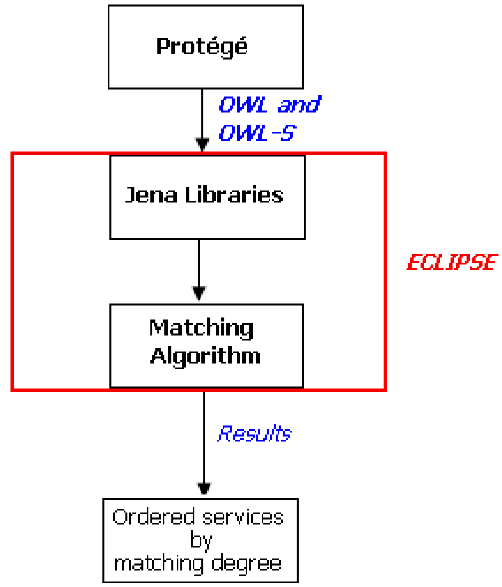


Figure 3. Logical Matching Algorithm Architecture inside eclipse environment

Particularly fig. 4 shows a snapshot of the environment execution menu constituted from the choice search for (I) advertisement (II) new advertisement (III) new services. Figure 5 show the matching results obtained composing the request with the advertisement.

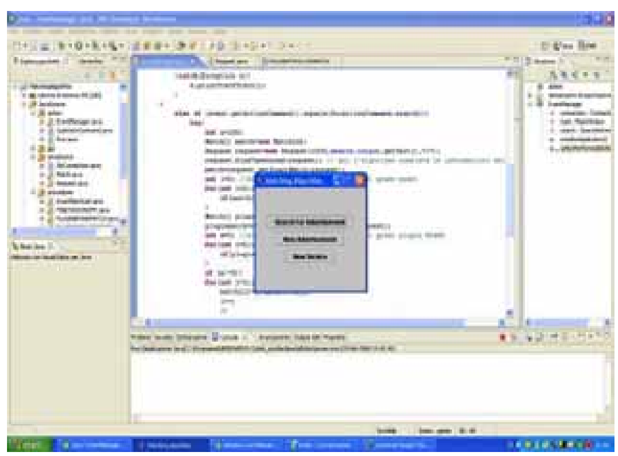


Figure 4. WS Eclipse environment: the menu search

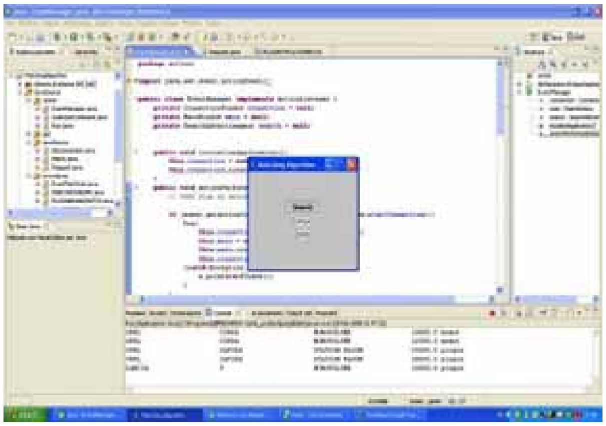


Figure 5. WS Eclipse environment: the ordered services by matching degree

The WS-Eclipse tool manage the repository using DB2 express C database that is a pure XML one.

4. Result analysis and Conclusions

The WS-Eclipse project involves a workgroup that is working around also with Italian industries. Up to now the environment is under experimentation and the first release has been done. For the next future we are planning to realize more importing and exporting feature from the eclipse environment and interface with physical composition and tools such as BPEL and BPEL4WS. An integrated environment that will cover all the methodology illustrated in the chapter 3 also will be a future goal for the project. Particularly we are working in order to extend the environment im a manner that it will be used also to compose didactic document or concept with the aims of satisfy user requests.

Bibliography

[1] M.P. Papazoglou,P.Traverso,S.Dustdar, F.Leymann, *Service-Oriented Computing: State of the Art and Research Challenges,* IEEE Computer november 2007, pp.38-45.

[2] R.Akkiraju, R.Goodwin, P.Doshi, S.Roeder, *A Method for Semantically Enhancing the Service Discovery Capabilities of UDDI,* IBM T. J. Watson Research Center.

[3] F.Lécué,A.Léger, *Semantic Web Service Composition Through a Matchmaking of Domain,* Proceedings of the European Conference on Web Services (ECOWS), 2006.

[4] S. Narayanan and S. McIlraith, *Simulation, verification and automated composition of web services,* Eleventh International World Wide Web Conference, pages 7–10, May 2002.

[5] V. Agarwal, K.Dasgupta, N.Karnik, A.Kumar, A.Kundu,S. Mittal, and B.Srivastava, *A Service Creation Environment based on End to End Composition of Web Services*, In Proceedings of the 14th International World Wide Conference, May 2005.

[6] E. Christensen, F. Curbera et al., *WSDL:Web Services Definition Language 1.1,* http://www.w3.org/TR/wsdl.

[7] A. Zaremski and J. Wing, *Specification matching of software components,* In ACM Transactions on Software Engineering and Methodology, 1997.

[8] A. M. Zaremski and J. M. Wing, *Specification Matching of Software Components,* ACM Transactions on Software Engineering and Methodology, 1997.

[9] B. Benatallah, M.-S. Hacid et al., *On automating web services discovery*, The VLDB Journal, 14(1):84–96, 2005.

[10] S. Agarwal and A. Ankolekar, *Automatic matchmaking of web services*, In WWW, 2006.

[11] T. Syeda-Mahmood, G. Shah, R. Akkiraju, A. Ivan, and R. Goodwin*, Searching Service Repositories by Combining Semantic and Ontological Matching*, In IEEE International Conference on Web Services (ICWS), 2005.

[12] M. Paolucci, T. Kawamura, T. R.Payne, Katia Sycara, *Semantic matching of Web Services Capabilities,* in First Intl.Semantic Web Conference,LNCS 2342:333-347,2002.

[13] M. Chung, H. Namgoong, K.-Il Kim, S.Jung,H.Cho, *Improved matching algorithm for services described by OWL-S, ISBN 89-5519-129-4, 2006.*

[14] M. Paolucci, K. P. Sycara et al., *Delivering semantic web services*, In WWW (Alternate Paper Tracks), 2003.

[15] M. Klusch, B. Fries et al., *Owls-mx: Hybrid owl-s service matchmaking*, In AAAI Fall Symposium Series, 2005.

[16] R.Guo, J. Le, X. Xia, *Capability Matching of Web Services Based on OWL-S,* Proceedings of the 16th International Workshop on Database and Expert System Applications (DEXA 2005).

[17] E. Gamma, K.Beck, *Contributing to Eclipse Principles, Patterns and Plug-in,* Addison-Wesley 2004.

[18] J. D'Anjou, S.Fairbrother, D. Kehn, J. Kellerman, P. McCarthy, *The Java Developer's Guide to Eclipse,* Addison-Wesley 2005.

[19] J.Lee,Yanggon Kim,Yoonho Kim,B.-H.Moon, *Business Process Integration with Web Services,* PSNPD/SAWN 2005.

[20] A.Kumar, A.Neogi, *Raising Programming Abstraction from Objects to Services*, IBM T. J. Watson Research Center, RI07003 in 2007.

[21] T. Syeda-Mahmood, G. Shah, R. Akkiraju, A. Ivan, and R. Goodwin, *Searching Service Repositories by Combining Semantic and Ontological Matching*, IEEE International Conference on Web Services (ICWS), 2005.

[22] J. Rao, P. Kngas et al., *Application of linear logic to web service composition*, In ICWS 2003, June.

[23] S. Narayanan and S. McIlraith, *Simulation, verification and automated composition of web services,* Eleventh International World Wide Web Conference, pages 7–10, May 2002.

Identity, Negotiation and Trust in a Distributed Knowledge Management Network

Giacomo Franco[1], Paolo Maresca[2], Giancarlo Nota[3]

[1] IBM Italia
[2] Dipartimento di Informatica e Sistemistica Università di Napoli "Federico II", Italy
[3]Dipartimento di Matematica e Informatica, Università degli Studi di Salerno, Italy

Abstract

The problem of knowledge management in a distributed network is receiving increasing attention from both scientific and industrial organizations. Research efforts in this field are motivated by the awareness that knowledge is more and more perceived as a primary economic resource and that, in the context of organization of organizations, the augmented management complexity of its whole life cycle requires new reference models.

In this paper, we build on recent research work to propose a distributed knowledge management framework that can be used in several application domains. First we discuss the role of identity, negotiation and trust in a distributed knowledge network, then we introduce a framework for the knowledge management that extends previous works towards the knowledge management in a network of organizations. Finally, we discuss a case study as a particular instance of the presented framework to handle the problem of risk management in enterprise alliance partnerships.

1. Introduction

The interoperability of organizations that engage in knowledge exchange to pursue innovation is one of the challenges of our age [17]. In fact, networks of organizations (NoOs) have been growing rapidly in number and scope in recent years, with the majority of business organization now belonging to at least one of such networks. Supply chain and virtual B2B are typical examples of network of organizations whose members work in close and continuous collaboration on projects, processes, products, services, to pursue a common strategy. Several issues push to interorganizational networking: faster time to market, needs to guarantee availability of resources and material, etc. Issues can be solved only if ability to use knowledge distributed inside the network of organizations increases and it allows to better concentrate on the growing of core competences. A network of organizations undoubtedly can offer risk and cost mitigation benefits when innovation programs are launched. An important contribution to the problem of knowledge exchange in a distributed system of systems is due to studies on communities of practice, social network, etc. and the interest to share knowledge in such virtual networks is continuously growing.

Despite strong motivations encourage research works in the direction of better virtual environments, conceptual barriers and limitations still exists.

A big issue is technology; it is mature enough to support communication between people and organizations, when most interactions involve exchange at data and information levels.

However, when people and organizations turn their attention to collaboration activities within the context of virtual networks [10], [18], [2], technology must evolve to embrace social needs easing interaction finalized to knowledge sharing (KS). But technology is only the tip of the iceberg in the sea of issues rising from paradigmatic differences in our understanding of what knowledge is.

"Paradigms" govern how knowledge flows between people and organizations and they impact on interorganizational knowledge exchange [6]. The debate about knowledge exchange focuses on two fundamental paradigms: *knowledge as object* and *knowledge as process*. The researchers and practitioners following the "Knowledge = Object" ("K=O") paradigm usually exploit the concepts from Information Theory in their understanding of Knowledge. The researchers and practitioners referring to the "Knowledge = Process" ("K=P") paradigm tend to take their concepts from philosophy, psychology or sociology.

Because of their different origins, the two tracks use different languages in their dialogues and thus confusion might arise. Presently, the "K=O" paradigm deriving from mechanistic world view of Cartesian Philosophy, has a widespread diffusion; it suggests us that a well done technological infrastructure might be perceived as a sufficient tool to enable exchange of knowledge. Instead, the "K=O" paradigm uses schemes that cannot embrace wealth of human behaviours during knowledge sharing. Characteristics of domain of knowledge are important issues, as demonstrated from open source software projects experiences. In fact, Open Source activities resemble the cumulative process of science [24]. Their KS success depends from the domain of knowledge that is the software development. To reach such results in KS both paradigms are useful, but they have to be integrated, as shown by results of studies on FLOSS communities [11]:
methodologically, we need more grounded, ethnographic-oriented research for our understanding of the socio-

technical practices of deployment, development and implementation of FLOSS in different contexts.

In FLOSS communities developers operate on a shared technical artefact (the code) that is tangible and they share the criteria for evaluating good designs. This context makes easier to build theories that work, because they construct simultaneously technical artefacts and their identity. The paradigms of KS have to be extended when they work in a further dimension, the interorganizational one, because they normally identify processes that are in place inside one organization. Processes related to KS between different organizations look very different [5] being mainly based on relationship mechanisms realized inside and between at least two levels of *receiving/giving* entities: organizations and employees.

This paper aims to investigate on models and technologies that can integrate paradigms of knowledge to provide support to distributed knowledge management (KM) in a context where people, groups or organizations exchange knowledge cooperating to reach a common goal. We draw on the research literature extending basic concepts of the Renaissance shop metaphora [12] of Eclipse Italian Community Governance that realizes a modern renaissance craftsman studio, a best of breed model for multiple corporations to create innovation networks, enabling cooperation on the development of product-ready open source software, and using the KM framework proposed in [14]. We focus on the aspect of social influence as a potential big barrier of KS between different organizations in terms of lacks of identity, negotiation and trust. Social influences involve people when they identify themselves with the single organization participating to the network or when their identity is associated to the emergent network of organization.

The structure of the paper is the following. In the next section we discuss the role of Identity, Negotiation and Trust in a distributed knowledge network. Then, in order to characterize the social influences emerging when a network of organizations is built, we propose a framework that integrates recent research results with the concepts of Identity, Negotiation and Trust. To show the framework applicability, in the fourth section a case study concerning the distributed risk management is discussed as a particular instance of the framework presented in section 3. A discussion about benefits and limitations of the proposed framework closes the paper.

2. Identity, Negotiation and Trust in a Distributed Knowledge Network

Peter Drucker observed in "The Fifth Discipline" that *knowledge is information that changes something or somebody, either by becoming grounds for action, or by making an individual (or an organization) capable of different or more effective action*. The concept can be extended to a network of people and organization as well.

Using the metaphor of complex systems, the whole of network of organizations can be perceived as a virtual meta-organization, with new properties emerging from interactions between organizations participating to the network.

To analyze from the KM view how this meta-organization emerges in interorganizational environment, and knowing that the results (good or bad) of emergence depend from organizations and individuals behaviours, we believe that strongest issues in interorganizational KS resides in identity and trust and how organizations use negotiation in building them. Furthermore, we believe that identity and trust are related in interorganizational environment and, obviously, they have to be addressed in some different way for organizations rather than for individuals. For example, to pursue what some authors [3] suggest to cultivate a community and to understand the main hurdles hindering the exchange of knowledge among members, we need to recognize emerging identity for groups of members and define what barriers are in place for them. Within NoOs an identity has to be defined for each organization belonging to network at meta-organization level. To ensure that members (organizations) do not misuse the shared knowledge (e.g., by taking advantage of confidential information), and to persuading them that the NoOs is a source of reliable information, trust has to be developed at meta-organization level. The paradox is that to control what emerges at meta-organization level we need to operate at organization level relying to what is done at individual level. To study these phenomena it is necessary to explore together three levels and to try to understand what happens.

Current literature gives definition of identity and trust and it emphasizes the impact they produce on community building. How could they be applied in the context of NoOs' KS? Different research oriented to issues of social, organizational and corporate identity have been recently explored [7]. On one hand, studies on social identity generally examine issues of cognitive process and structure and organizational identity, but research tends to address the patterning of shared meanings. On the other hand, studies of corporate identity tend to focus on products that communicate a specific image. Authors call for greater cross-fertilization of the identity concepts proposed by literature and discuss requirements for the integration of micro and macro-level analyses.

Indeed, identity is a subjective concept which changes over time and in different situations [13]. It is individualized through the narration of our own story; participating in new activities and social groupings encourages people to become different and this participation has an impact on their identities. Professions also play an increasingly important role as a source of identity. The sentence "defining a community" identifies the formation of an identity which provides the language through which we understand ourselves and the knowledge context in which he lies. Identity is tied up with one's commitments, values and goals. This nature of identity is fluid and it changes with organization's changing values and agendas. Wenger [23] states that building an individual identity consists of negotiating the meanings of our experience of membership in social communities.

When the action of building an individual identity is studied inside NoOs it appears that the negotiation is more complex than inside a community, due to three level paradox (individual, organization, meta-organization). People build their identity not only participating in activities of their own organization, but also (at the same time) participating in activities of meta-organization. When the strategy, the commitments, the values and goals are the same, the person behavior can be placed within the scope of a single organization.

If one of the above mentioned items changes, in the sense that people do not share anymore the same values, goals, etc., something changes. A person will need, above all, to negotiate his identity with himself and with others members of his organization as well, and later with other members of its meta organization. He will define in this way the meaning of his experience and what are the behaviors to hold in accordance with the actual situation. For example, if NoOs is a joint project between several organizations, and a leadership organization is formally defined, it will play the role of coordinator for the meta-organization and it will negotiate its identity from individual level to the meta-organization level. If a conflict of interests arises, potentially each participating organization will have to identify the correct behavior to cope with this situation, changing its behavior and, therefore, its identity. Whenever an equilibrium point cannot be reached a schizophrenic behavior of organization might appear.

Identity and Trust are strongly interconnected, especially inside NoOs. Some authors [9] [1] [19] assume that existence of body of common knowledge/practice, sense of shared identity and some common or overlapping values is likely to produce trust-based relationship, creating social conditions that are conducive to knowledge sharing.

Trust has been defined from many researchers in several disciplines. One intuitive definition of trust is "the confidence that a person has in his or her favorable expectations of what other people will do based on previous interactions" [8]. An extension of this definition has been exploited by community *Trust is the expectation that arises within a community of regular, honest, and cooperative behavior, based on commonly shared norms from the members of community*.

A frequently used definition of trust that summarizes both sociological and economical aspects is: *the willingness of a party to be vulnerable to the actions of another party based on expectations that the other will perform a particular action important to the trustor*[14]. This definition outlines a risk of an opportunistic behavior of partners, and it represents an appropriate reference to enable a suitable mechanism to reduce such a risk. In NoOs scenario several trust levels could appear resulting from trust concepts identified by other authors [15]:

- **Intraindividual Trust**: trust in one's own abilities and self-confidence basic trust (in others)
- **Interindividual Trust**: expectation based on cognitive and affective evaluation of partners
- **Intraorganization Trust**: trust in organization's own abilities and self-confidence basic trust (in other organizations)
- **Interorganization Trust**: expectation based on cognitive and affective evaluation of partners to follow established, or based on common ethical ground, rules of partnering
- **System trust**: trust in depersonalized systems/word that has to work and it is independent from people/organization (e.g. technology, regulations, economic systems, legal systems)
- **Knowledge object Trust**: A knowledge object could be certified by a third party.

All former levels of trust could be grouped in only two types of trust, *hard* and *soft* trust to simplify their usage in the context of this paper. *Hard Trust* has roots in things like contractual obligations or other legal and regulatory mechanisms, for example establishing coordinating roles such as sponsors, champions, facilitators, practice leaders, and infomediaries or making the access easy through the development of the technological infrastructure for KM. *Soft Trust* relies on cultural and social structures, reputation and interpersonal relations, such as relations inside community which bring benefits to members allowing KS, shared unwritten rules that members are expected to respect and a carefully selection of partners. Inside a NoOs trust is essential for all actions that involve interconnections between organizations, in particular for KS. Fears to receive a bad behavior from partners, to lose some competitive advantage sharing knowledge without any return or worse, gifting someone who could become a competitor and to make present and future common projects to fail, they are really a strong barrier to KS.

Other aspects, interconnected with identity and trust, have to be considered: *Negotiation* and its value inside NoOs. The value of *intermediation* inside community of practices (CoPs) has been identified in literature [3] and functions covering this activity seem to be fundamental inside CoPs. What about Negotiation in NoOs? If an intermediate facilitates the knowledge sharing inside a CoP and his role appears essential to make knowledge sharing to happen, it is possible to state that *without any negotiation between organizations inside NoOs it is not possible to perform any knowledge sharing activity*. The negotiation is performed in NoOs continuously:

1) it is performed during the identity building of each organization inside the network of employees (e.g. brand image and interests versus needs of network of organization);

2) resulting organization identity from previous step is negotiated with the organization identities coming from other organization in NoOs;
3) negotiation is performed to build the final identity of each organization and the meta-organization identity;
4) negotiation will be a need even to "certificate" the object (content) of knowledge (k-object) that is shared; the process of certification could be complex and deep as preferred, establishing associations of good reputation of members to the k-object, until a knowledge quality assurance process put in place.

The meta-organization could be identified by a leader that will assume the role of knowledge negotiator, a kind of supervisor performing activities that are normally done by a community leader. It identifies rules and control that they are respected from organizations, it works to realize NoOs development, motivating people to share knowledge at organizational and individual level; it facilitates knowledge transactions regulating competition taking into account macrostructural properties distinguishing formal work groups and networks of practice. Each organization participates to negotiation assuming the role of leader of its own community of employees, as it repeats the activities of meta-organization knowledge negotiator at organization level, working for and with individuals.

The last aspect, interconnected with others, to take into account is *Communication*. In every-day life much of distrust created is due to inappropriate communication on issues, feelings, intentions and opinions. Information that is relevant in NoOs should be given promptly and frequently together with revelation of negative aspects. In fact, production of trust rests on common knowledge base, which increases the predictability of partner behaviour through shared meanings. The object of communication is information, knowledge and/or emotion [21]. The role of communication inside NoOs is to balance, through communication skills, the lack of natural socialization appearing when partners have an asymmetric relation and work separately and in different contexts or cultures.

Communication validates the technology useful to growth and performing of social aspects (identity, trust and negotiation), cooperation and collaboration in NoOs. Communication is important because it rests on distributed knowledge inside the NoOs. Our claim is that the variant of the top-level conceptual framework for KM due to M. Stankosky and A. Murray (Fig.1) proposed in section 3 helps to understand:

1. how knowledge sharing can amplify human cognition at individual as well as at group or organization level,
2. how it could be pursued the community objective, reducing efforts required to share or exchange information and to keep remote people in touch so that the collective motivation is better oriented toward common goals .

Our understanding is that a Distributed Knowledge Management Framework can be exploited as a governance structure that helps to define how to manage social influence increasing awareness and encouraging collaboration and mutual accountability.

3. A Distributed Knowledge Management Framework

So far, much research work has been devoted to Knowledge Management foundation and models able to represent and manage knowledge within the context of a single institution. However, the experience already gained from social networks encourages the introduction of new models able to represent KM concepts in a distributed setting.

Although the top-level conceptual framework for KM due to M. Stankosky and A. Murray [34], reported in fig. 1.a), includes aspects regarding the Knowledge Transfer (these are further divided into the categories of social structures, sharing & dissemination, communication infrastructure and transfer protocol not shown in figure), we believe that an extension of this framework that includes the dimension of "social influences" is necessary to propose a distributed knowledge management framework (DKMF) that considers the ideas of identity, negotiation and trust, as discussed in the previous section, in order to better characterize the problem of KM in a network of organizations. Note that the life cycle concepts shown in fig 1.a) have been rearranged in order to better render the relationships between concepts belonging to different dimensions. As a matter of fact, suppose that the leader of an organization (dimension: four pillars of KM), decides to adopt knowledge assurance procedures (dimension: KM Life Cycle) in terms of confidentiality, non-repudiation, availability, integrity and access control that together contribute to the concept of trust (dimension: Social Influences). He might then start a decision process finalized to expose the organization to the influences of a social network either to consume or to produce knowledge; this requires identity, trust and negotiation that, in turn, relie on communication (dimension: social influences) to become part of a knowledge network. Indeed, the quality of knowledge assurance reinforces the trust of other organization that are then encouraged to consume knowledge generated, coded and ready to be transferred

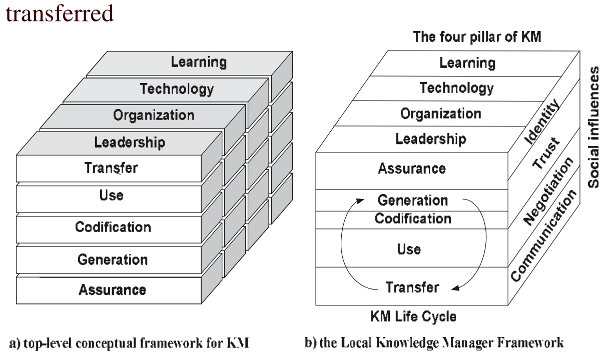


Figure 1. The conceptual framework for KM

Having characterized the dimension of social influence relating it to the four pillars of KM and to the KM life cycle, we are now ready to propose the DKMF shown in fig. 2 as a structure built starting from three fundamental parts: the Local Knowledge Managers (LKM), the Virtual Community Supervisor (VCS) and the Virtual Knowledge Repository.

LKM are the keepers of local knowledge i.e., the one maintained by an enterprise, that are available to share their knowledge with others community participants in reference to a given application domain. The role of the node VCS is of paramount importance in the start up phase of a new Distributed Knowledge Network; apart from the typical functions assigned to LKMs, it is capable to design a new knowledge network infrastructure, to assume the leadership for the government of a virtual community, to state the identity of the knowledge network, to define the four KM pillars at the meta level of DMKF.

Looking at KM models proposed in the literature we are conscious that their practical use requires the definition of some additional steps [4]. The DKMF is not an exception. The main phases for the realization of a knowledge network based on the proposed framework are

1. **Planning**: to define the four pillars (leadership, organization, technology, and learning) at the metalevel of virtual community. This includes the social structure, the communication infrastructure and the identification of transfer protocols.
2. **Tailoring**: to choose the application domain and the models for the knowledge representation; produce the first artefacts to share and disseminate.
3. **Relationships**: to state identities of participants, the negotiations and trust relationships; to define the rules that allow the LKM to consume/produce knowledge from/to the virtual repository.
4. **Use**: a LKM shares its knowledge with other community participants and receives from the virtual repository available knowledge according the transmission rules and modalities (e.g. on demand, planned transfer, asynchronous communication, e.g. announcements, notifications, etc.)

The virtual knowledge repository is the virtual place where shared knowledge can be stored and allows social aspects to work meaningfully. Indeed, people in NoOs could own a piece of tacit or explicit knowledge and are encouraged to consume/produce knowledge from/to a virtual repository when it is perceived as a centralized knowledge base even if it is really distributed in many virtual and physical places. The virtual knowledge repository for a NoOs is, from one hand, the personalized infrastructure of knowledge communication and storing (e.g. chat, Wiki, blogs, k-bases, reputation building software, social networking software, security access software, etc) that allows to consume/produce tacit and explicit knowledge. On the other hand it is a particular instantiation of the knowledge owned by human being and machines, both belonging the NoOs. The knowledge repository has to be intended as a common language that co-ordinate social activity, realized trough the infrastructure and the contents (this is a basic theme discussed in [24]: the principal function of language is to co-ordinate social activity, from which meaning itself arises).

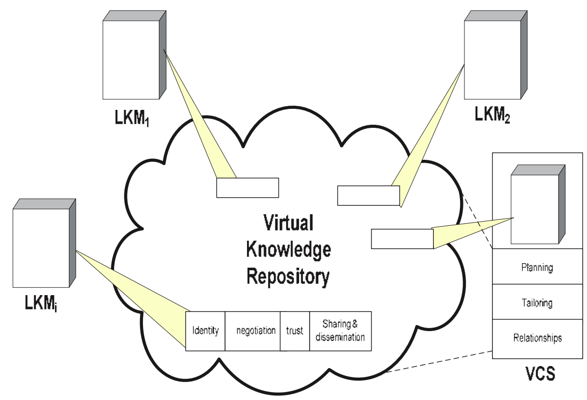


Figure 2. The Distributed Knowledge Management Framework

4. Risk Management in Alliance Partnership: A Case Study

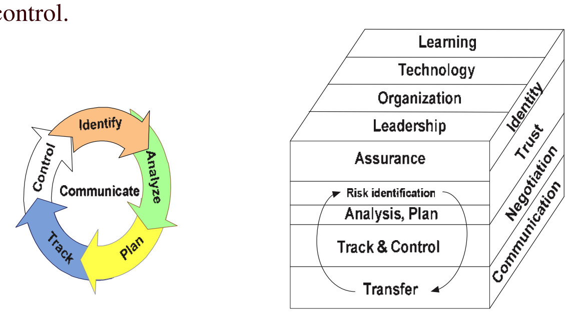
In the last few years the trend towards outsourcing partnership and alliance partnership is constantly increasing. This is due to several motivations such as, reduced cost, involvement of specialized competencies, better proposal capacity of goods and services, challenging project, innovation need, ecc.

Although these new forms of collaboration bring benefits in terms of time to market, reduced cost and quality of goods and services, the *risk exposure*, i.e. the loss deriving from an unwanted event weighed with probability of its occurrence, is difficult to handle in a typically distributed context [16].

In [14] current limitations to risk management are outlined especially in the context of a system of systems where interoperable risk management is much more difficult and poses new research questions concerning, for example, what information to exchange, which processes are relevant and what is the scope and nature of interaction between entities that participate in interoperable risk management.

We argue that the DKMF can be used as a reference to implement a distributed interoperable risk management system. Since it exists a large body of knowledge already developed to cope with risk exposure in enterprise risk management, it is possible to start from one of the proposed models as a building block in a distributed risk management scenario. For example, if we instance the LKM to the SEI paradigm we obtain the framework shown in fig. 3.a) where the main activities are:

Identify: to define the risks list
Analize: to prioritize risks and convert data into decision-making information
Plan: to decide what should be done about a risk or a set of related risks
Track: to acquire risk status and record it
Control: to decide for the best reaction when the risks manifest itself,
and the **Communication** activity is cross-activity in the sense that data or information handled by a certain activity can be communicated to whom it may concern with the purpose to maintain risk and risk loss under control.



a) the SEI risk management paradigm b) mapping the SEI paradigm to a LKM

Figure 3. Tailoring of a LKM node

Now, let as suppose that in an alliance partnership each participating enterprise adopts the SEI paradigm to manage risks. Even if internal risk can be profitably handled in the context of a single enterprise, many aspect concerning the risk exposure in a distributed environment are difficult to perceive. This is due to the shifty nature of risks in a cooperating alliance partnership; for example, if an unwanted event arises in a given enterprise having a potentially negative impact on other participants, the risk exposure could remain latent until a loss becomes manifest. Indeed, risk might result unidentified or hidden as a result of the enterprise opportunistic behaviour. Another aspect regards the risk migration; for example, moving workforce temporarily from one enterprise to another could undermine the capability of the first enterprise to match its own deadlines.

To gain insight about distributed risk management problem, we can apply the main phases of the DKMF as follows:

Planning: to state the four pillars at the appropriate level. In the example of alliance partnership, this consists in choosing the enterprise that assumes the leadership and assigning to it the role of VCS; then the social structure made of an organization of organizations is established, and the communication infrastructure is put forward together with the deployment of e-learning application oriented to the risk management in the alliance enterprise.

Tailoring: to map the SEI Risk Management Paradigm over LKM node as shown in fig. 3.b) and to produce, share and disseminate risk and risk exposure documentation concerning the application domain, e.g. a risk ontology in software engineering projects.

Relationships: to represent into the distributed risk management system involved people and enterprises that join the alliance (identification). The shared procedure for the exchanging of risk data is stated together with the contractual clauses that regulate the mutual interaction, and the criteria necessary to check if the clauses are satisfied (negotiation). The rules that allow the LKM to consume/produce knowledge from/to the virtual repository are defined and knowledge about risk events of each enterprise that might migrate and/or propagate to other participants increasing the global risk exposure is identified; finally, the appropriate handling procedures for risk migration/propagation is established (trust).

Use: to start the risk life cycle at each LKM, to start monitoring procedures for risk migration and propagation, to communicate the happening of risk event both at the LKM and DKMF levels, the planned handing procedures for local and global risk mitigation are eventually applied.

Once the framework is tailored to the problem of distributed risk management, the capability to reason about the concept of interoperable risk management increases. Essentially, the nature of agreements between different organizations, the modality of information sharing and the influences of operations performed by communicating entities can be investigated from the risk management perspective. The outlined method concerns either, the usual aspects of risk management in a single enterprise or the risk management within a distributed network of organizations and is defined and proposed as a means to control the collective behaviour of organizations.

5. Conclusion

The increasing importance of knowledge management in the context of an organization of organizations is due basically to economic and social factors. On the one hand, the awareness that knowledge is a primary asset in the scenario of growing competition drives enterprises to make alliances and to share their knowledge in order to create value through innovation. On the other hand, many organizations share and disseminate knowledge to pursue objectives of scientific and technologic divulgation, social inclusion and, in general, to improve standard of life.

Whatever is the motivation, the discipline of knowledge management exhibits a great support potential that can be exploited to reach either enterprise success or increased cultural level in social networks.

The framework introduced in this paper considers a variant of the top level conceptual framework for knowledge management due to A. Stankosky and A. Murray. Our proposal is oriented to the modelling of knowledge management in a distributed network of organizations and investigates the concept of identity, negotiation and thrust that allow, through communication between the participants, to share and disseminate knowledge in a virtual community.

Although the proposed framework is general enough to represent both existing implementations (notably the knowledge network implemented by the eclipse community) and future implementations (e.g., a network for distributed risk monitoring) in a variety of contexts, limitations exist. As the representation and management of knowledge is a complex field that aims to provide support to several different disciplines and relies itself on several disciplines to implement a Knowledge Management Program, we cannot expect that exists a catch all framework that is well suited for any application. In spite of these limitations, it is possible to recognize some common principles and theoretical constructs that can be profitably employed as a reference guide to model and implement Knowledge Management Systems in a distributed scenario. The DKMF presented in the paper takes into account the discussed limitations through the tailoring phase by means of which the peculiarities of a given application domain can be considered to qualify the framework.

References

[1]. Andrews K., Delahaye B., *Influences on Knowledge processes in organizational learning: The psychosocial filter*, Journal of management Studies, 37 (6), 797-810, 2000.

[2]. Baim S. A., Understanding Knowledge Bases and Building membership in Virtual Communities, Encyclopedia of virtual communities and technologies, Idea Group Reference, 2006.

[3]. BOLISANI E., SCARSO E., DI BIAGI M., *Economic Issues of Professional Communities*, Encyclopedia of Communities of Practice in Information and knowledge Management, Idea Group Reference, 2006

[4]. CALABRESE F. A., *The Early Pathways: Theory to Practice – a Continuum- in Creating the Discipline of Knowledge Management,* Michel Stankowski ed. Elsevier, 2005

[5]. CHEN S., DUAN Y., EDWARDS J, Inter-*Organisational Knowledge Transfer Process Model*, Encyclopedia of Communities of Practice in Information and knowledge Management, Idea Group Reference, 2006.

[6]. COOKE-DAVIEST., ET ALT., *We're non in Kansas Anymore, Toto*, PM Journal, PMI, June 2007

[7]. CORNELISSEN J. P., ., HASLAM A. S, BALMER J. M. T., *Social Identity, Organizational Identity and Corporate Identity: Towards an Integrated Understanding of Processes, Patternings and Products*, British Journal of Management, Vol. 18, S1–S16, 2007.

[8]. GEFEN, E-commerce: The role of familiarity and Trust, Omega 28,6, 725-737, 2000.

[9]. HISLOP D., *The paradox of Communities of Practice*, Knowedge Networks of Practice, Idea Group, 2004.

[10]. HUGHIRIAN P., *International Knowledge Transfer as a challenge for Communities of Practice*, Encyclopedia of Communities of Practice in Information and knowledge Management, Idea Group Reference, 2006.

[11]. LIN, Y. *"The future of sociology of FLOSS."* FirstMonday, 2005.*http://firstmonday.org/issues/special10_10/lin/index.html*.

[12]. MARESCA P., FRANCO G., *The Eclipse Italian Community e-governance the renaissance shop methaphora,* Eclipse.IT First International Conference, Napoli 2007.

[13]. MEDLIN C., QUESTER P., *Inter-firm trust: two theoretical dimensions versus a global measure*, 18th IMP-conference in Perth, Australia, 2002

[14]. MEYERS B. C., *Risk Management Considerations for Interoperable Acquisition*, CMU/SEI-2006-TN-032.

[15]. MEZGÁR I., *Building and Management of Trust in Information Systems*. Encyclopedia of Information Science and Technology (I) , 298-306, 2005

[16]. OLSON E. G. *Strategically managing risk in the information age: a holistic approach*, Journal of Business Strategy, vol. 26, n° 6, pp. 45,54, 2005

[17]. PANTELI N., *Trust and Conflicts within Knowledge Intensive Environments,* The Third European Conference on Organizational Knowledge Learning and Capabilities, Athens Greece, 5-12 2002.

[18]. PRIESTLEY L. J., *Knowledge Transfer within Interorganizational Networks*, Encyclopedia of Communities of Practice in Information and knowledge Management, Idea Group Reference, 2006.

[19]. ROBERTS J., *From Know How to Show How?*, Technology Analysis and Strategic Management, 12 (4), 429-443, 2000

[20]. STANKOSKY M. A., *Creating the Discipline of Knowledge Management*, Elsevier, 2005

[21]. SYDOW J., *Understanding the Constitution of Interorganizational Trust in Trust Within and Between Organizations*, Lane Christel and Bachman Richard (eds), Oxford University Press , 1998

[22]. TUOMI I. *What did we learn from Open Source?* FirstMonday. *http://www.firstmonday.org/issues/special10_10/tuomi/index.html*

[23]. WENGER E., William Snyder, Communities of Practice: The organizational Frontier, HBR Jan.-Feb. 2000

[24]. WINOGRAD T.,, FLORES F., *Understanding Computers and Cognition*, Addison-Wesley,1986

E-EXAMS CONTAINING "TIME-DEPENDENT TEST ITEMS WITH IMMEDIATE CHECK"

Paolo Maresca, Sergio Scippacercola

Dipartimento di Informatica e Sistemistica
Università degli studi di Napoli "Federico II" – Napoli (Italy)

ABSTRACT

A critical aspect of computer-assisted exams which is usually complained about, is the lack of interaction between the examiner and the person being examined. As an alternative, we here propose a model of e-exams organized in diversified test items some of them consisting in close-ended questions while others are "open-ended and time-dependent questions with immediate check". The latter type of questions allow to counterbalance the lack of examiner/learner interaction as the learner is given the opportunity of being informed about the test result immediately and thus may correct his/her response. Marks get as better as shorter the time lapse is between the question and the correct answer provided. In order to compare and assess the student's open-ended answer as matched with the key answer prepared by the examiner, a measurement of student answer similarity is proposed to be applied and adapted to the keywords subspace. The model described is applicable to the computer science domain but may be also extended to any other field of distance education and learning.

Keywords: e-learning, e-exam, computer assisted assessment, multi-choice test, decision support system.

1. INTRODUCTION

Computer-assisted education may be classified into: distance learning if geography is taken into account, e-learning if the use of technology is considered and on-line learning (or web-learning) if the teacher/student interactivity is the aspect we are mostly interested in [7]. E-learning, in particular, may be included in educational processes and frameworks where on-line teaching is matched to traditional teaching methodologies (classes, labs, workshops etc.). Unlike what happens in traditional educational approaches, the lack of the teacher in e-learning methodologies is one of the most critical aspect for the learner. This lack, during the learning phase is replaced by mentoring and monitoring activities aimed at reducing the risk of learners dropping out and thus controlling the learning process. During or after the teaching/learning phase competencies may be assessed through the administration of computer-based tests (CBT), also said e-exams [11]. The e-exam is a type of examination where questions and answers are recorded in a processor and may be immediately assessed or later. The types of testing activities provided (true/false, multiple-choice, etc.) highlight a particular underlying attitude of the student. On the other hand, open-ended questions, among other, allow the student to demonstrate knowledge acquired without being assisted, guided or disturbed by "ready-made" answers [12]. Furthermore, open-ended questions vis-a-vis close-ended ones solicit more detailed and richer responses provided in the form of well-constructed statements rather than triggering a short *reaction* conditioned by the testing modalities [10].

The objective of this paper is to propose a model of e-exam which also contains test items simulating the typical teacher/student interaction in an oral exam. The second section of the paper describes some typical aspects of the traditional evaluating and assessing approach and suggests which of these aspects is more suitable for being transferred into a computer-assisted examination in consideration of the shortest times required for test execution (*Time-saving principle*). In the third section the subsystems supporting the e-exam model system are briefly analyzed while in the 4rth section the measurement of answer similarity is proposed as a tool for comparing and assessing the key answer prepared by the teacher with the one provided by the student. Then in the fifth section the time-dependent/immediate check requirement is proposed as the most innovative feature of our e-exam model (Par. 5). In the sixth section the assessment procedure is presented functionally to the time-saving requirement illustrated before. The paper ends with experimental results, conclusions and prospects for the completion and dissemination of the model proposed.

2. FEATURES OF AN E-EXAM AND THE TIME-SAVING PRINCIPLE

Each oral examining approach offers advantages and disadvantages. Four distinct testing "techniques" or approaches characterizing traditional examining approaches may be distinguished and have been here reported in the form of charts (Fig. 1) containing nodes (questions) and edges (cross-relation between a question and the following) oriented in accordance with the testing approach adopted by the teacher during the exam. So we have identified four types of traditional testing procedures:

1) top-down;

2) bottom-up;

3) networked;

4) randomized.

Each node represents a question on a specific domain. In the first approach (top-down procedure) (Fig. 1, on the left) the teacher starts examining the student by asking him general questions progressively getting deeper and deeper. The second examining approach (bottom-up procedure) instead consists in detailed questions progressively getting more and more general to get to a final general synthesis of the topic proposed to the student. In the third approach (networked procedure) the examiner surfs through questions and topics in a logic relationship. In this case the teacher starts examining the student with any topic (any node of the network) (Fig. 1, on the right) and proceeds with topics logically related to the preceding questions.

The fourth approach consists in a set of questions randomly asked without any logic causative link between them. So in conclusion if the questions are logically cross-related the approach may be defined networked. Differently, if the questions are proposed randomly, the examining approach will be defined randomized. In practice there is no "pure" application of the procedures described above but mixed approaches are usually adopted by examiners.

On analyzing the four traditional approaches described above, and in view of adopting some of their features in an e-exam model, the first consideration stemming is that one single approach is to be chosen when designing an e-exam model (top-down, bottom-up, networked, randomized). The model proposed in the following paragraphs is based on a particular networked approach where the *n* questions relating a specific knowledge domain are networked randomly with both close- and open-ended questions. This architecture allows the check of various levels of knowledge as acquired by the learner.

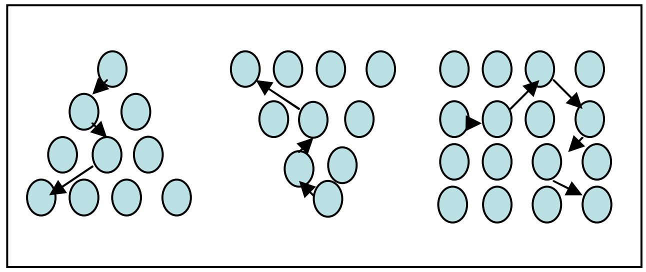


Fig. 1 – Charts representing different examining procedures with edges standing for hypothetical examining paths. Each node is a question. The top-down procedure is on the left, the bottom-up approach is in the centre while the networked one is to the right.

The second consideration is that the assessment process in a traditional exam is carried out by taking into consideration:

A) the percentage of correct answers vis-à-vis the total number of questions;

B) the percentage of correct answers after the teacher soliciting and help;

C) the total time elapsed between the question and the answer;

D) other elements depending on the exclusive teacher/student interaction and difficult to be classified such as the type of register adopted by the student, his accuracy, self-confidence and command on surfing through the specific domain, etc.

Items A and B above have already been adopted in the assessment process of existing e-exams. Traditional evaluation approaches are not only summative but take into account the items classified above as B and D. The *B* and *D* item depend on a specific relationship between the teacher and the student and cannot be easily automated.

As a matter of fact, each question requires a correct answer to be provided in a variable time depending on the level of difficulty of the topic presented. In addition to the answer appropriacy, time is a fundamental parameter to be taken into account in the assessment process. The time-saving principle basically prevails: "*P* score is assigned as inversely proportional to *t* time taken to provide the correct answer".

Also, we suggest proposing the student tests of different typology as this allows the teacher to detect underlying aspects of the student's know-how and so provide a clearer view of competencies achieved.

The e-exam model we propose here consists of a set of questions some of them open-ended, others multiple-choice. As concerns open-ended items three typologies have been identified:

• mono-key questions where the answer contains one single key word;

• multi-key questions if the answer contains more key words

• cross-related multi-key questions if the answer contains more cross-related key words (generally in a sequenced order).

Accordingly, some requirements of an e-exam model have been roughly identified in the following as essential if the e-exam is to replace a traditional type of exam successfully:

1) a set of questions of various types all requiring a dichotomic answer (true-false, fill-in-the-blank , match-type, multiple choice);

2) a set of open-ended questions with immediate check and possibility of correcting one's answer within a time limit.

For reckoning the final score of these tests the following is required:

• number of correct answers and the time taken to provide them must be easily accounted for (A, B and C percentage);
• to be provided with a facility which may counterbalance the lack of interaction in case of a wrong or partial answer.

The e-exam model proposed consists in a structure organized in a set of test items, some of them open-ended, the others multiple-choice.

3. THE ARCHITECTURE AND REQUIREMENTS OF AN E-EXAM MANAGING SYSTEM

The architecture of the system adopted for the arrangement of questions and the administration of this type of e-exam consists of two sub-systems [12]. The first subsystem is adopted by the teacher to build up the question database, the keys database, the organization of questions, the scores to be assigned, key words and their crossrelations, the time limit required for providing an answer per each question and to complete the test.

The first subsystem to be applied, named PARTES, is aimed at creating two databases. PARTES is basically employed by the teacher for the following purposes:

- changes the input text from a sequence of characters into a sequence of sentences and words.
- distinguishes per each sentence full words from empty ones (articles, linkers, etc.).
- measures the frequency of a full word in a sentence.
- displays the teacher each sentence with the relative distribution of words.
- proposes the teacher whether to discard or take into consideration the sentence as appropriate answer to a question. If this is the case, it accepts the drafting of the question by identifying the *full words* which may be considered correct in the student's answer.
- on the teacher's suggestion, checks and integrates the synonymous present in the *Synonym Database* and contributes to weighing single words contained in the "right" answer.

Each key prepared by the teacher is the result of a decomposition to obtain the most relevant words extracted from the noise.

The second subsystem named PARSTU is aimed at administering questions, acquiring the learner's answers and assessing them. Questions whether close- or open-ended are proposed in a random sequence and each answer is matched to the teacher's key by measuring its similarity in a 0-infinite range (0 in the case of perfect match between the learner's answer and the teacher's key and infinite in the case of fully wrong answers). The answer similarity measurement adopted compares relevant words, pre-recorded by the first subsystem, with the sentence keyed in by the learner. If there is a perfect match between the pre-recorded answer and the one provided by the student the aggregate score is assigned (e.g. 1), if the answer is partially correct or is incomplete half score is assigned (0.5), if it is fully wrong 0 score is assigned. In order to allow an effective process of answer acquisition the System must be fitted with the text editing, analysis and control facility.

The PARSTU subsystem (Fig. 2) is aimed at capturing and assessing the student's answers to questions proposed and is structured into two steps. In the first phase the subsystem:

1) randomly picks up from the Question Database the questions and helps the student writing the answer by means of a form. The subsystem lets the student write the answer as he likes;

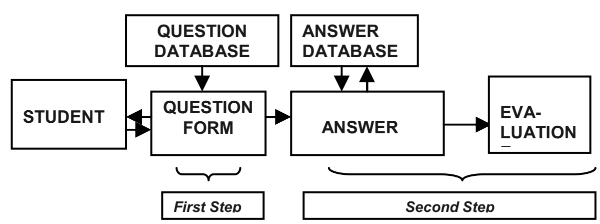


Fig. 2 – The PARSTU subsystem

2) modifies the answer text from a sequence of characters to a sequence of words;
3) measures the frequency of each word in a sentence.
In the second phase, by means of the similarity measurement described below, the subsystem compares relevant

words contained in the teacher's key with words of the student's answer.

4. STUDENT'S ANSWER SIMILARITY MEASUREMENT

In the case under test, starting from the strategies proposed by *LSA* [8], *Data* e *Text Mining* [4][1] and in particular the *Bag-of-Words* [2] we define the student's answer and the teacher's key as taken from the textbook. Each answer to a question is represented by the p key-words (k_i) identified in the textbook after eliminating each empty word and is represented by the vector $\boldsymbol{t}$ of p elements as follows:

$$\boldsymbol{t} = (k_1, k_2, \ldots, k_j, \ldots, k_p)$$

where k_j is a general relevant word identified in the teacher's key to the question. To take into account the frequency of words and the different relevance within the answer to a specific question, the teacher may introduce an adequate system of w weights per each answer and per each key word:

$$\boldsymbol{w} = (\varphi_1 w_1, \varphi_2 w_2, \ldots, \varphi_j w_j, \ldots, \varphi_p w_p)$$

where φ_j indicates the frequency of j-th word in the context of the correct answer and w_j indicates its weighting. The words of the student's answer to the i-th question are gathered in the vector $\boldsymbol{r}$ and the relative frequencies in the vector $\boldsymbol{f}$:

$$\boldsymbol{r} = (r_1, r_2, \ldots, r_i, \ldots, r_n) \qquad \boldsymbol{f} = (f_1, f_2, \ldots, f_i, \ldots, f_n)$$

where r_i represents the i-th word and f_i the relative frequency associated. It is assumed that the n words supplied by the student may be more or less the p words contained in the right answer. The $\boldsymbol{t}$ and $\boldsymbol{r}$ words are gathered in a matrix $\boldsymbol{A}(q,2)$ $(q \leq n+p)$ containing in its first column all the $\boldsymbol{t}$ and $\boldsymbol{r}$ words picked just once while in the second and third column, respectively, the weighed frequencies ($\varphi_j w_j$,) associated to the words of the right answer and the frequencies (f_j) associated to the word of the student's answer are indicated. That generates the matrix represented in Table 1.

As it is a frequency matrix, the similarity between the set of words represented by the vector $\boldsymbol{t}$ and that represented by the vector $\boldsymbol{r}$ will be measured. In this paper the vectors above are considered as two statistical variables: key-words in the teacher's key and words supplied by the student are to be intended statistical units for which the weighed frequencies are observed. We define the occurrence of a word (Tab. 1) as *left* if the word is used only in the key answer; as *double* if it is used both in the teacher's key and in the student's answer, and *right* occurrence if the word is used by the student but not in the key answer. A table is therefore generated (a matrix) as the one represented in Tab. 1, where the first column reports, just once, all the words employed both in the teacher's key and in the student's answer, associating their frequency of occurrence to each. Just for example three types of different occurrences are reported (*left*, *double* and *right*) for three words. In Tab. 1 each row reports the frequency of occurrence and the weights of words in the teacher's key (column 2) and the frequencies of occurrence of words in the student's answer (column 3). Answer similarity metrics may be applied but the χ^2 metric stands out from the others. In particular the weight of each root square deviation may be assumed as measurement of the relevance of the specific modality vis-à-vis the others [9]. Therefore we suggest [12][13] to use the χ^2 metric [3] [9] and let $d^2_{t,r}$ be the similitarity measure between the $\boldsymbol{t}$ and $\boldsymbol{r}$ vectors :

$$d^2_{\boldsymbol{t,r}} = \sum_{i=1,q} \left[\frac{a_{i1}}{a_{.1}} - \frac{a_{i2}}{a_{.2}} \right]^2 a_{..} \Big/ a_{i.}$$

where a_{i1} is the relative frequency of a word belonging to the set of teacher's keys and a_{i2} the frequency of a word in the student's sentence; $\boldsymbol{a}_{.1}$, $\boldsymbol{a}_{.2}$ and $\boldsymbol{a}_{i.}$ are, respectively, marginal total amounts of frequencies per column and per i-th row; $\boldsymbol{a}_{..}$ is the general total amount. The measurement of the distance between $\boldsymbol{t}$ *and* $\boldsymbol{r}$ provides a measurement of how similar is the student's answer to the teacher's key.

The answer similarity measurement proposed may be considered a measurement of the student's mistake rate in providing the answer. If the answer fully matches the teacher's key $d^2_{\boldsymbol{t,r}} = 0$. If differently the answer is missing, the measurement of answer similarity gets an infinite value. Depending on the question, thresholds are set within which the score is assigned. To determine the minimum threshold level the measurement value is calculated per each question in case of the occurrence of one or more relevant words with full weight. The distance thus obtained represents for that question the minimum threshold level to be achieved in order to obtain the minimum score fixed by the teacher.

Table 1 – Example of a matrix of frequencies of word occurrences in an answer (*left*, *double*, *right occurrence*).

Occurrence	k_i	φ_i	w_i	f_i
Left	UNICODE	1000	1	0
Double	ASCII	1000	1	1
Right	HOLLERITH	0	0	1

5. TIME-DEPENDENT TEST ITEMS WITH IMMEDIATE CHECK

The innovative aspect of this model is the proposal for introducing a partial simulation of the student/teacher interaction as concerns open-ended questions in the current e-exam System [13]. The student/teacher interaction is replaced by the display of the scale of test results and the possibility for the learner to key in other answers. The immediate display of the test results is implicitly a solicitation for the student to find another answer. The student while writing the answer can see the variable scale pointer moving up and down (Fig. 3 on the right). Per each correct word the student keys in, the pointer displays the relative result and the score goes up or down.

If the student feels like, he may try to answer the question better by keying in other words or by completing the answer provided. In case the student has not achieved the score desired he may insist in doing better or may give it up and pass onto the next question. The system however records the t_i time elapsing between the reading of the question *i-th* and the writing of the answer.

Fig. 3 - Form for "capturing" answers where the measurement scale is highlighted to allow the student to upgrade his response if he wants.

In Fig. 3 a type of test item is illustrated as having an immediate check ("list the most common fixed-length codes"). The student has entered the ASCII code as first. As the reference to other codes too is required the score obtained is not the highest, that is 1, but only 0,33. If the student wishes to better his score he may go on referring to other codes too and in this case the pointer will immediately move to 1 if the answer is correct. If the student does not intend to upgrade his answer he will proceed with other questions but will be not allowed to go back later. If he cancels his answer the pointer moves to zero.

6. ASSESSMENT OF AN E-EXAM

In this research, while adhering to the Classic Item Analysis [6], we will indicate by X_i the score assigned to the close-ended question with dichotomic result, by Y_i the score assigned to the test item with immediate check and by P the total score.

For the n close-ended answers included in the tests an X_i continuous variable associated to the answer provided to the i-mo question may be added which becomes zero value per wrong answer and 1 per right answer.

Per each test item with immediate check, in the preliminary drafting of the tests, the teacher records two threshold values in the subsystem (*S1* and *S2*) for the answer similarity index. This operation is fundamental as the answer similarity index varies from zero to the infinite and is indispensable to establish reference values for the acceptance or not of the answer:

- *S1* is associated to the acceptance threshold of a correct answer (variable index from zero to S1);
- *S2* indicates an incomplete answer (variable index from *S1* to *S2*).

If the answer similarity index reaches a value between 0 and *S1* the aggregate score assigned to the *i*-th test item is for ex $Y_i = 1{,}00$; if the index is between *S1* and *S2* the score is ($Y_i = 0{,}50$). If it exceeds the *S2* threshold the score is zero ($Y_i = 0{,}00$). For each test item with immediate check the teacher also sets the mean time allowed (T_i). The student answering within the mean time limit set (t_i) is awarded with a higher score if $t_i \leq T_i$ and penalized if $t_i > T_i$. Accordingly the rectified score Y_i^r will be:

$$Y_i^r = Y_i \frac{T_i}{t_i}$$

Lastly, by indicating with n the close-ended answers and m the test item with immediate check, P defines the aggregate score or aggregate assessment:

$$P = \sum_{i=1,n} X_i + \sum_{j=1,m} Y_i^r \ .$$

7. EXPERIMENTAL RESULTS AND CONCLUSIONS

Average times taken by the students for each answer have been calibrated by monitoring a group of volunteers. After calibrating answer times, the test has been administered to a sample of sixty-five students. Scores to be assigned ranged from zero to fourteen. Each test session lasted no longer than twenty minutes. The average time for completing the test was twelve minutes. Most students (65,8%) obtained a score between 6 and 10, 17,1% of the students obtained a high score (between 11 and 14). The method adopted therefore appeared quite discriminatory.

After completing the test, the students have been administered a questionnaire aimed at assessing their confidence in the new assessment method and the test user-friendliness, including any remarks and suggestions for improvement, etc. Among the remarks received from the students tested, the following have appeared worth mentioning: "the indicator is very good clear and synthetic"; "it would be better to be allowed to go back to the preceding questions to complete them"; "the e-exam is a better test than the oral exam". The student percentage approving the new exam has resulted very high (Tab. 2) and the introduction of the indicator to counterbalance the lack of teacher-student interaction has been positively judged. In conclusion, if we consider that the percentage of negative judgements on the e-exam has been low (6.5%) , we may certainly conclude that the e-exam assessment method is to be considered reliable.

The e-exam new assessment methodology takes into account both the introduction of test items with immediate check and the principle of saving the time elapsing between the question and the answer. The test items with immediate check may find application also in the test assessment process of other fields apart from the educational one. The computer-assisted exam may be applied both to a traditional learning course and a computer based one. This new approach may be applied also to training courses, competitions, text databases, classification of key words on the web, classification of complaints, determination of products having similar selling strategies, identification of stocks having similar trend, research on visual databases, etc. In conclusion, the methodology proposed, though currently restricted to the computer science domain and still missing significant validation tests, appears to be promising in terms of its functionality in measuring the student's mistake rate per each type of test administered also as concerns other disciplines.

Table 2 – Results of the questionnaire monitoring student's satisfaction with the new exam.

Question	**positive judgements**	**no reaction**	**negative judgements**
Do you feel confident in the e-exam?	73.7%	10.5%	15.8%
Do you appreciate the new test method displaying the measurement scale?	63.2%	21.0%	15.8%
Are questions clearly structured so as to trigger an immediate answer from you?	84.2%	10.6%	5.2%
Final judgement	73.4%	20.1%	6.5%

BIBLIOGRAPHY

[1]Aureli Cutillo E., Bolasco S. (2004), *Applicazioni di analisi statistica dei dati testuali*, Casa editrice Università LA SAPIENZA.

[2] Balbi S., Misuraca M., Pesi e Metriche nell'analisi dei dati testuali (2005), *Quaderni di Statistica* , vol. 7.

[3]Benzecri J.P. (1973), *L'analyse des données, Tome 2: L'Analyse des Correspondances*, Dunod, Paris.

[4]Bolasco S. (2004), *Mondo digitale, n. 1.*

[5]Bolasco S., Canzonetti A., Capo F.M. (2005), *Text Mining*, CISU.

[6] Bove, G. (2000), Valutazione nei servizi alla persona, Valutazione della qualità, efficacia ed efficienza nei servizi alla persona, con particolare riferimento all'istruzione e alla sanità", Dipartimento di Statistica "G. Parenti", Università di Firenze, interim paper, (www.ds.unifi.it/iniziative_sc/convegni/w25febbraio/programm.htm), 2000.

[7] Colorni, A. (2002), Mondo digitale, 1, 4-15.

[8]Deerwester S., Dumais S.T., Furnas G.W., Landauer T.K., Harshmanet R. (1990), Indexing by latent semantic analysis, *Journal of the American Society for Information Science*, 41 (6), 391-407.

[9]Fabbris L. (1997),*Statistica multivariata: analisi esplorativa dei dati*, Mc-Graw Hill, Milano.

[10]Fraire, M. (2000) Analisi dei Dati a Tre-vie delle Risposte a Domande Aperte e Indicatori Empirici, JADT 2000 : 5es Journées Internationales d'Analyse Statistique des Données Textuelles.

[11]Rutkowski, J., Moscinska, K., Chruszczyk, L. (2006) Development and Evaluation of Computer Assisted Exam Circuit Theory Example, Proc. Computers and Advanced Technology in Education, 2006.

[12]Scippacercola, S. (2007), Un Sistema di supporto per il controllo automatico di risposte libere a quesiti di esame, a cura di A. Andronico e G. Casadei, Atti Didamatica 2007, Cesena, Società Editrice Asterisco, 2007.

[13]Scippacercola, S. (2008), Un modello di e-exam con item a risposta libera ed a verifica immediata, a cura di A. Andronico, T. Roselli, B. Lamborghini (Eds.), Atti DIDAMATICA 2008.

Design and implementation of a user friendly environment for the creation of Learning Objects: CADDIE (Content Automated Design and Development Integrated Editor)

Giovanni Adorni, Diego Brondo, and Mauro Coccoli
{adorni, diego.brondo, mauro.coccoli}@unige.it
University of Genoa,
Department of Communication, Computer, and Systems Science
E-Learning & Knowledge Management Laboratory
Viale Causa, 13 16145 Genova, Italy

Abstract

The aim of this paper is giving a description of the CADDIE (Content Automated Design and Developement Integrated Editor), a user friendly authoring tool for the automated production of sharable content objects, mainly oriented to networked learning and learning object design. It is a freely available Open Source application to assist teachers in the publishing of web content without the need to become proficient in HTML, XML markup or anyother language. Contents produced within the CADDIE environment can be exported in the form of self-contained web pages, LaTeX files, standard document formats, or SCORM 2004 packages as well, usable on a LMS server or designed for a database SCO repository. This project has been developed by ELKM (E-Learning & Knowledge Management) laboratory and led by the University of Genoa.

1. Introduction

The progress of Information and Communication Technology (ICT) is changing in depth production, sharing, management, and exchange of knowledge: methodologies and tools of communication, investigation and information research are rapidly changing too.

Several years of experience in developing e-learning projects at the University of Genoa [1], have pointed out a scenario of increasing managerial complexity, due to the existence of heterogeneous technology tools. In an open and heterogeneous environment, the interaction among different platforms and legacy systems can only be achieved by means of seamless integration and communication, and the use of standards. Another key point is the usability: researchers, teachers, and anyone else having to develop learning objects, has to cope with the complexity of technology, programming languages and elaborate software tools. Hence the objective of the project is to provide an easy-to-use [2] tool enabling authors of educational materials to produce and deliver documents and sharable objects (compliant to international standards, ADL SCORM [3] in primis) in a short time and with a little effort. In this respect, a specific tool has been designed, which we called "the CADDIE" (Content Automated Design and Developement Integrated Editor). In golf, caddie is the person who carries a player's bag, and gives insightful advice and moral support. A good caddie is aware of the challenges and obstacles of the golf course being played, along with the best strategy in playing it. In a similar way we want CADDIE to be an useful help in design and development of educational contents aiding and driving the authors to a winning strategy.

Furthermore the CADDIE is thought as a part of a comprehensive pipeline of software tools developed at the University of Genoa that covers the whole process of course design, from concept map indexing to publication or storage passing through the creation of learning objects [4].

CADDIE can be adapted to several application scenarios and learning backgrounds (school, university, primary school, business, and corporate). On account of this and on the different skill proficiency of potential users, it was of primary relevance to develop a system respecting two main concepts: being accessible and easy-to-use.

2. Analysis and design of a course

The first activity to be accomplished for an educational project to be successful is performing an accurate analysis of the requirements and of the objectives.

The CADDIE implements a formal yet flexible model to be followed by authors, which we summarised in three main steps (figure 1). This logical steps are easy-fitting to any contents factory process.

2.1. Knowledge representation

The first step is to create a map of the concepts, a logical and well-formed diagram expressing the relationships between concepts of the topics of the lesson. Using such a diagram is a good method for knowledge representation and allows the connection of this encoded knowledge to relevant information resources. In the notions schema, showed in figure 1, topics represent the subjects of discourse (the lesson arguments) and arrows are the associations, representing relationships between the subjects.

As an example of notions map we could consider a set of lessons on operating systems: processes, memory technology, disk supports, file systems, history of OSs, state of the art, process management, and memory allocation could be the topics, related each other by priority and propaedeutic connections (e.g., the lesson on memory allocation must come after the one on memory tecnology, not vice versa). Tracking the associations it is possible to glean a linear sequence of topics obtaining, as result, a well ordered index (chapters, sections, subsections, etc.).

The map of concepts represents in all respects the index of the e-learning course and due to its intrinsically hierarchical nature, it can be easily turned into a data structure such as a XML tree following the above priority rules. The XML file index can be imported into the CADDIE creating the framework of the project.

Primarily the purpose of XML is facilitating the data sharing across different information systems and thus facilitating the software elaboration and integration. Consequentially we consider a proper choice of names for XML elements in order to convey the meaning of the data in the markup structure. This increases human readability while retains the rigor needed for software parsing.

2.2. Content management

The second step is the creation of the content filling the previously designed logical structure and this can be obtained in two different ways. First way is writing from the scratch with the CADDIE internal editors. Second way is considering an empty framework solution (result of the topic linearization) and adding content to it both by writing from scratch or importing from existing libraries.

Any topic can be linked by the author to one or more occurrences, which connect the subjects to pertinent information resources (i.e. PDF documents, images, multimedia files, wed resources, Wikipedia pages, etc). Such contents are organized into the XML file and, from this point on, the file can be re-edited or rendered in different shapes depending on the specific application or device the SCO was designed for.

Users not only can fill up the concept structure previously created, but also they can edit the concept map itself by adding or even removing one or more nodes as new chapters or sections of the course: they are free to modify the framework at any time.

2.3. Final output

The last step is to save the whole work for immediate use or re-use. Depending on specific user needs and educational strategies, different output solutions can be required. The CADDIE offers the possibility of customizing the output format and presentation according to the given requirements: the presentation device, the type of lesson, the accessibility level, and other possible final uses, just like a golf caddie gives the right club to the golfer.

3. Implementation

With the intent of making CADDIE available to any person, anywhere, and on any device, the more suitable solution is that of developing a web application and services. Moreover, being a web application, it can be easily integrated in LMSs or other collaborative and distributed environment for knowledge management and sharing.

From the users' point of view, two possible scenarios arise: using a stand alone application on the local machine or accessing to a remote service provided by a server-side application. In the former case users will have to install on their system a CADDIE environment constituted by a web server with PHP support and MySQL database management system and the web application can run locally without the need of an internet connection. In the latter just a web browser with JavaScript support will be needed for CADDIE and web server reside on a remote machine accessible through a web connection on which CADDIE will be shared among the users as a server-client application.

Both the user interface and the usability concepts it is based on, the CADDIE was developed taking into account an user-centered design [5], with its intended users in mind at all times. Many users neither are expert in technology nor know any programming languages. CADDIE web interface was designed to be as easy as possible, just like an internet page and all the complex operation for contents and outputs manipulation are completely transparent. So that teachers, instructors or business managers have to focus their attention only on the contents of the lesson.

We use a modular programming approach (see figure 2) breaking up CADDIE into manageable units and create

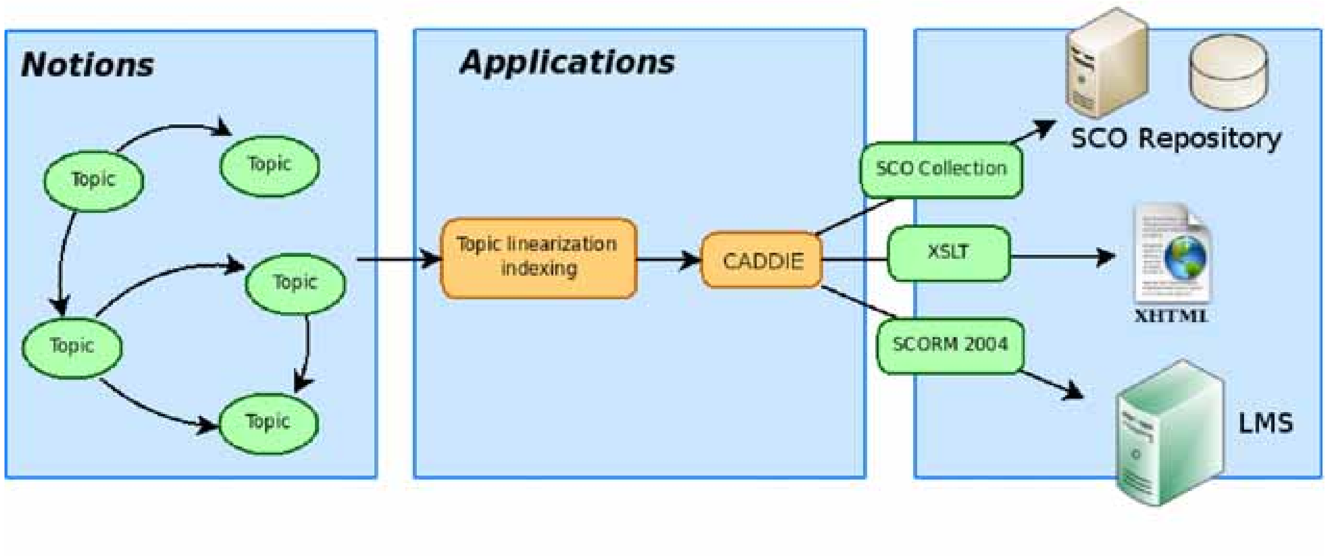


Figure 1. Learning Object production line

code that can be easily re-used or extended. Like a golf caddie that usually carries several clubs during the game to face to the different situations the player may come across, the CADDIE implements different modules that come to the aid of the user in design ad developement of learning contents.

3.1. Web application

We devoleped a web application written in PHP (PHP Hypertext Preprocessor) a reflective programming language and we used XHTML (eXtensible HyperText Markup Language), a markup language that has the same depth of expression as HTML, but with a syntax that conforms to XML syntax, for the web interface [6]. This guarantees documents to be well-formed and a better accessibility.

A significant advantage in building web applications to support standard browser features is that they should perform as specified regardless of platforms or clients. Rather than creating clients for MS Windows, Mac OS, GNU/Linux, and other operating systems only one core web application is written. Additionally, the ability of users to customize many of the display settings of their browser (such as selecting different font sizes, colors, and typefaces) can give an advance to the web application accessibility [7].

3.2. Web Acessibility

The Web Content Accessibility Guidelines [8] cover a wide range of recommendations for making web content more accessible. Following such guidelines will make content accessible to a wider range of people; the WCAG also requires that web applications work when JavaScript is disabled or not supported, but following these guidelines will also often make web application not much functional. There is a trade-off between accessibility and a better usability: to give best performances CADDIE uses AJAX thecnology and thus it will not work in all web browsers. As its name suggests, AJAX requires JavaScript. This alone means that AJAX applications will not work in web browsers and devices that do not support JavaScript.
The current solution to these problems is to either provide a non-AJAX alternative to the application or to allow its AJAX application to continue to work if JavaScript and XMLHttpRequest (required by AJAX to work) are not supported. Such a requirement may be very difficult to achieve. In CADDIE we try to follow WAI (Web Accessibility Initiative) directives as much as possible but we still use JavaScript for client-side dynamic and interactive operations.

3.3. Topic Maps

Despite the generality of considerations about the map of concepts, at the implementation time we chose a specific formalism well suited to our needs. We decided to use Topic Maps [9] for representing the notions schema in figure 1.

Topic Maps is an ISO standard for the representation and interchange of knowledge. A topic map can represent information using Topics (representing any concept real or abstract), Associations (which represent the relationships between them), and Occurrences (which represent relationships between topics and information resources relevant to them) [10]. They are therefore similar to semantic net-

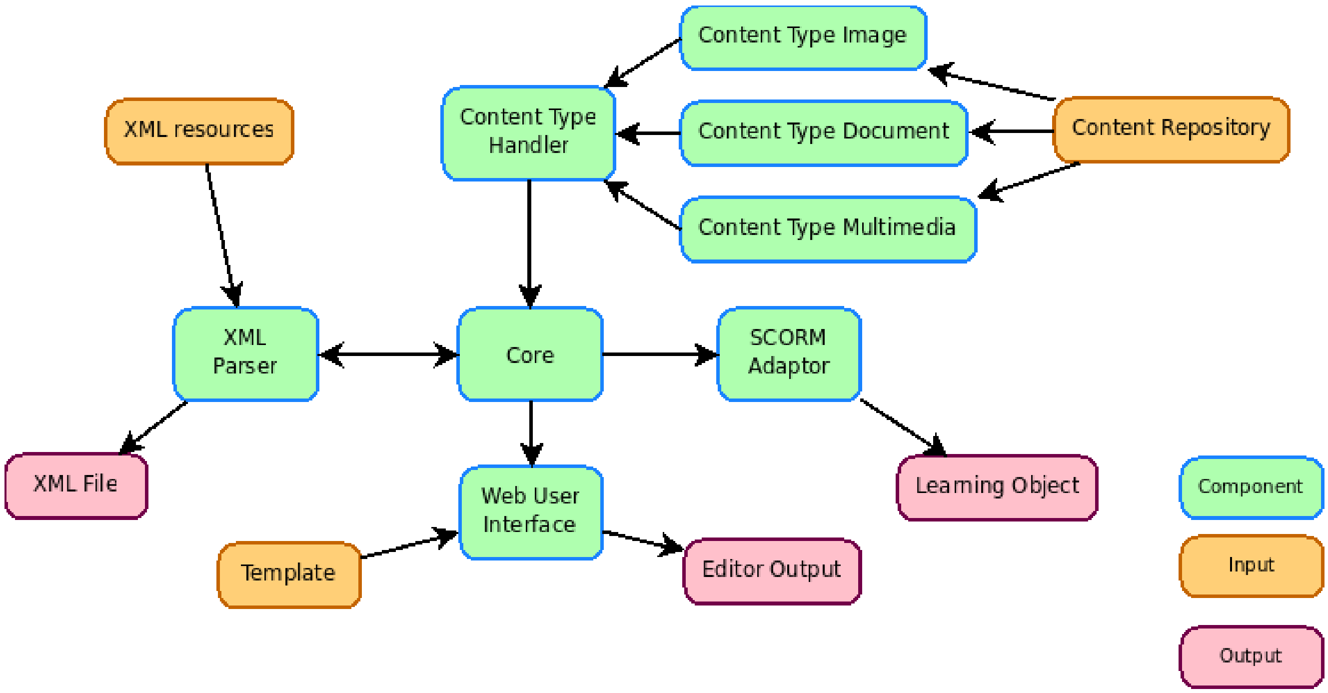


Figure 2. CADDIE schema

works and both concept and mind maps in many respects. In loose usage all those concepts are often used synonymously, though only topic maps are standardized.

Topics, associations, and occurrences can be typed, but the types must be defined by the creator of the topic maps which defines the ontology of the topic map. There are also additional features, such as merging and scope useful for navigation. The concept of merging and identity allows automated integration of topic maps from diverse sources into a coherent new topic map.

Topic maps used as a knowledge representation methodolgy, rely on a standard language, that is XTM (XML Topic Maps) based on XML syntax. For the sake of the integration in CADDIE, the XTM graph is linearized into a plain XML tree (the framework of the project) so that concepts and some of their relationships become an ordered sequence of topics in a course.

3.4. Content managemet

The type of content that can be used within the CADDIE may be heterogeneous: text, images, HTML pages, WiKi resources, multimedia files, etc. are possible assets to be profitably used in teaching activity. Such resources can be both written straight with CADDIE editors if needed, or imported into the CADDIE environment by means of a redirection to an URI.

Each asset (text field, image, multimedia content, etc.) is handled by a specific PHP class. These classes handle both the module interface providing a suitable editor for each content type (i.e. creating the rigth XHTML forms with specific input fields according to the content type itself) and the object manipulation methods (PHP functions). This approach enables developers to easily add new features, fulfilling new educational purposes, just making new modules for handling different type of contents. At the moment CADDIE provides full-featured add-ins for managing text, XHTML and images. Soon it will also fully support multimedia contents like movies and audio files and anything else to come. This modular approach guarantees flexibility and adaptiveness to different technological scenarios as well as to pedagogical methodologies; further extensions of CADDIE will be able to cope with new features or new standards to come.

Once obtained a structure compliant to own intents, the only thing left is the users to save the entire project and get the work done. Thanks to a set of plugins the customers can choose among different output formats. Using a module rather than another one, the learning structure will be converted into different formats for different purposes.
You can see the CADDIE modular schema in figure 2; the advantage of a modular approach to programming [11] are self evident.

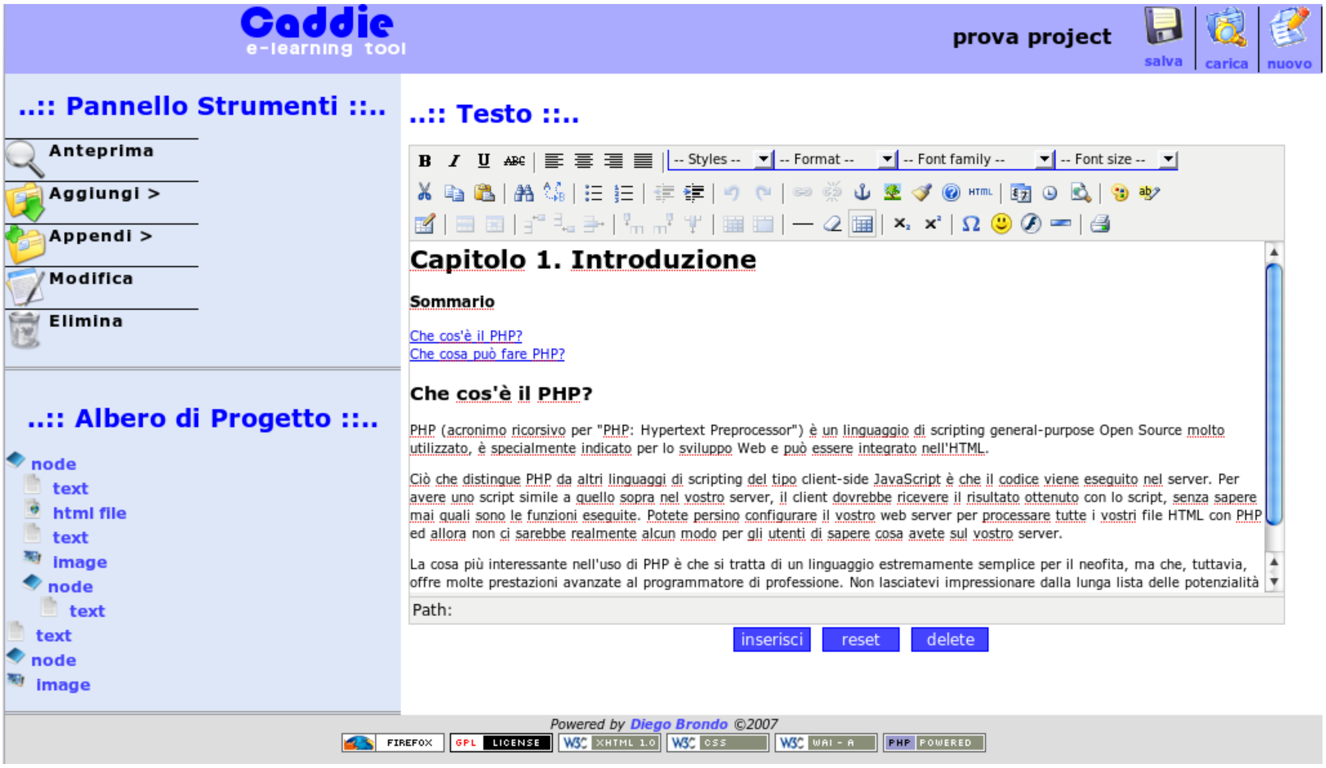


Figure 3. CADDIE (Italian user interface)

3.5. CADDIE outputs

Users can get XML output of the course they are creating for general purpose applications and re-use or should a later time modification be planned. When the CADDIE is integrated within any Learning Management Systems, its main functionality becomes the capability of generating SCORM 2004 compliant Learning Objects. A specific converter module can create a Sharable Content Object with manifest and metadata, portable on any LMS platform taht is SCORM compliant too. Moreover the output can also be redirected to a database with the purpose of adding their work into a shared learning objects repository. Finally users can also create a web version of the given lesson in XHTML pages customizable through suited XML/XSL-Trasformations depending on the client device or application. Furthermore this modular approach gives robustness to the standard changes, or database repository changes.

3.6. Graphical User Interface modularity

The web user interface PHP module uses Smarty libraries a ”Template/Presentation Framework” that provides the programmer and template designer with a wealth of tools to automate tasks commonly dealt with at the presentation layer of an application. Web page designers are not dealing with PHP code syntax, but instead an easy-to-use templating syntax not much different than plain HTML. The templates are a very close representation of the final output, dramatically shortening the design cycle. It is also possible to design different solutions suitable for different media (computer, mobile, PDA, word speech, etc) and different languages (see an example of the italian user interface in figure 3).

3.7. Security

One web applications allows users to enter information, later that information may be stored into files or database and retrieved. Users are capable of using any tecnology in both helpful and harmful ways. Protecting CADDIE against malicious users and automated attacks is a must especially if we are in a client-server scenario [12].

PHP has achieved a stable and solid presence on the web in the last several years, and its popularity as a server-side scripting language is only increasing. PHP is able to access files, execute commands and open network connections on the server. These properties make anything run on a web server insecure by default. PHP is designed specifically to be a more secure language for writing CGI (Common Gateway Interface) programs than Perl or C, and with correct selection of compile-time and runtime configuration options, and proper coding practices, it can give secure applications.

In CADDIE we wrote code in PHP 5.2 that has more secutity features than older versions and we followed the basic PHP security tips: for example disabling `register_globals` configuration option.

A special attention is reserved to user submitted data. Malevolent people can use input form for SQL injection, a technique that relies on a security vulnerability occurring in the database layer. The vulnerability is present when users input is either incorrectly filtered for string literal escape characters embedded in SQL statements or user input is not strongly typed and thereby unexpectedly executed.

CADDIE sanitizes input data, before using it in a SQL query: prepends backslashes to the characters that might trigger SQL injection, filters bad input checking if the input type the fileld in the database type correspond each others (i.e. if database expects a timestamp user input must be in that format).

Moreover the PHP errors which are normally returned can be quite helpful to a developer who is trying to debug a script, but this can expose hidden variables, unchecked syntax, and other dangerous information. In CADDIE the solution is to use PHP's custom error handling functions that report to the user only general error messages, while more detailed informations are logged into a file accessible only by system administrator.

4. Conclusion

A number of web publishing applications for authoring is available, both commercial and open source, with different characteristics and specific features. Most common applications for writing web pages provide a set of powerful and complex functionality whose use implies a somewhat computer knowledge and a fairly steep learning curve. Of course web pages can be used as educational content (even if not SCORM compliant) but too generic software does not give appropriate e-learning solutions [13]. On the other hand specific e-learning resources editors exist such as eXe.

With CADDIE we want to introduce a new working environment in which contents can be considered on their own and depending on their structural organization. Such contents can become learning resources if needed, or they can be used elsewhere for different applications but the focus is centered on the knowledge, not merely on the format of its representation.

Moreover CADDIE is designed to offer an easy and friendly system in which users, in out-and-out learning building, can use its features with extreme ease and often without the need for data conversion, thus saving on potential data losses and incompatibilities.

A further evolution of CADDIE will be the implementation of an efficient system for metadata auto-mining that will bring about two main advantages: the possibility of creating a well-formed repository for finding and reusing resources and the ease for authors in filling metadata forms.

References

[1] Giovanni Adorni, Mauro Coccoli, and Gianni Vercelli, "Integrated management of courses in a distance learning environment", in *Proc. of the Twelfth International Conference on Distributed Multimedia Systems, Bridging the gap between communication and intelligence*, 2006, pp. 222–227.

[2] Eric Butow, *User Interface Design for Mere Mortals*, Addison-Wesley Professional, 2007.

[3] ADL Initiative, *SCORM 2004 2nd Edition Overview in ADL (Advanced Distributed Learning)*, 2004.

[4] Giovanni Adorni, Mauro Coccoli, and Gianni Vercelli, "Eiffe-l: e-learning experiences with the platform of the university of genova", *Journal of e-Learning and Knowledge Society*, vol. 1, no. 3, 2005.

[5] Jonathan Lazar, *Web Usability: A User-Centered Design Approach*, 2006.

[6] Terry Felke-Morris, *Web Development & Design Foundations With XHTML*, Addison-Wesley, 2007.

[7] John M. Slatin and Sharron Rush, *Maximum Accessibility: Making Your Web Site More Usable for Everyone*, Addison-Wesley, 2003.

[8] "Web content accessibility guidelines 2.0", 2007, http://www.w3.org/TR/2007/WD-WCAG20-20071211/.

[9] Jack Park and Sam Hunting, *XML Topic Maps: Creating and Using Topic Maps for the Web*, Addison-Wesley, 2002.

[10] "The tao of topic maps", http://www.ontopia.net/topicmaps/materials/tao.html.

[11] Erich Gamma, Richard Helm, Ralph Johnson, and John M. Vlissides, *Design Patterns: Elements of Reusable Object-Oriented Software*, Addison-Wesley Professional Computing Series, 1994.

[12] Chris Snyder and Michael Southwell, *Pro PHP Security*, Apress, 2005.

[13] Giovanni Bonaiuti, *e-Learning 2.0 Il futuro dell'apprendimento in rete, tra formale e informale*, Erickson, 2006.

Semantic authoring of learning paths with Topic Maps

Giovanni Adorni, Mauro Coccoli, Gianni Vercelli, Giuliano Vivanet
{adorni, mauro.coccoli, gianni.vercelli, giuliano.vivanet}@unige.it
University of Genoa,
Department of Communication, Computer, and Systems Science
E-Learning & Knowledge Management Laboratory
Viale Causa, 13 16145 Genova, Italy

Abstract

The rapid development of the World Wide Web in the last few years has brought great opportunities in the way educational materials can be made available to learners. The number of resources available on the Internet is vastly growing, but, on the other hand, some problems emerged as a result of this proliferation of contents, such as the increasingly difficult management and accessibility of these materials. Topic Maps are an ISO standard whose aim is describing knowledge structures and associating them with information resources. Topic Maps are here proposed as a knowledge representation model to describe the semantic relationships among educational resources. Instructional designers and authors could use this model to facilitate the design of learning paths and their delivery in different contexts. In this paper, after a description of Topic Maps standard, a working hypothesis is discussed about its application in the context of learning design and also a short survey of related works is presented.

1 Introduction

The use of Information and Communication Technology (ICT) in learning activities has become so pervasive in the last few years that new models are needed for the process of instructional design, based on environment and tools enabling users to capture represent and share their knowledge [1].

Additionally, more and more often learning management systems are required to have high degree of flexibility, interoperability and personalization of contents and services and, therefore, to provide internal knowledge management and representation systems based on standards for resources, contents, and processes. From a technical point of view, semantic technologies can support both developers and users in achieving such goals.

There are several knowledge representation models, technologies and languages, such as eXtensible Markup Language (XML), Resource Description Framework (RDF), XML Topic Maps (XTM) and Web Ontology Language (OWL) that allow description of resources in a standardized way, enhancing the information sharing, reusability and interoperability.

Topic Maps (TM) [2] is an ISO standard (ISO/IEC 13250) for the representation and interchange of knowledge. It can be regarded both as a promise and a challenge for researchers involved in the learning design as well as in the management of educational resources.

2 ISO Standard 13250: Topic Maps

The TM development process began in 1991 when a UNIX system vendors' consortium founded a research group, known as Davenport Group, to develop a framework that enables the interchange of software documentation. The first attempt at a solution to the problem was called SOFABED (Standard Open Formal Architecture for Browsable Electronic Documents) [3].

In 1993 a new group was created, the CApH (Conventions for the Application of HyTime), whose activity was hosted by the GCA Research Institute. This group elaborated the SOFABED model as topic maps. By 1995, the model was accepted by the ISO/JTC1/SC18/WG8 as basis for a new international standard. In 2000 the Topic Maps specification was ultimately published as ISO/IEC 13250 [4].

In the same year a new independent consortium, TopicMaps.Org, was founded with the goal of specifying topic maps based on the W3C recommendations XML (to enable the applicability of the TM paradigm to the World

Wide Web). The XTM 1.0 specification was published in 2001; then it was passed over to ISO, which approved a Technical Corrigenda of ISO/IEC 13250 making the XTM notation part of the standard [4]. In the following years the TM development process has proceeded in different ways. Three OASIS Technical Committees were formed to promote the use of published subjects (element conceived to identify a single subject in a topic map), while a ISO committee JTC1/SC34 started two further standard initiatives: Topic Map Query Language (TMQL, ISO/IEC 18048), a query language for topic maps, and Topic Map Constraint Language (TMCL, ISO/IEC 19756), a constraint language for topic maps. In 2003 the second edition of ISO/IEC 13250 was released while, three years later, JTC1/SC34 published ISO/IEC IS 13250-2:2006 that specifies the Topic Maps Data Model.

Finally, in 2007, the same committee released XTM 2.0 (a revision of the XTM 1.0 vocabulary) whose syntax is defined through a mapping from the syntax to the Topic Maps Data Model. Thus the ISO standard is now a multi-part standard that consists of the following parts [5]:

Part 1 - Overview and Basic Concepts: provides an overview of each part and how the parts fit together. It also describes and defines the fundamental concepts of Topic Maps (standard under development);

Part 2 - Data Model (TMDM): specifies a data model for topic maps (it defines the abstract structure of topic maps). The rules for merging in topic maps are also defined, as well as some fundamental published subjects (published standard);

Part 3 - XML Syntax (XTM): defines the XML Topic Maps interchange syntax for topic maps (published standard);

Part 4 - Canonicalization (CXTM): defines a means to express a topic map processed according to the processing rules defined in the TMDM in a canonical form (project deleted on December 2007);

Part 5 - Reference Model (TMRM): provides a basis for evaluating syntaxes and data models for Topic Maps (standard under development);

Part 6 – Compact Syntax: defines a simple text-based notation for representing topic maps, it can be used to manually author topic maps, to provide human-readable examples in documents and to serve as a common syntactic basis for TMCL and TMQL (standard under development);

Part 7 – Graphical Notation: defines a graphical notation used to define ontologies and represent TM instance data (standard under development).

As previously said, Topic Maps define a model for encoding knowledge and connecting this encoded knowledge to relevant information resources [6]; in this paradigm emphasis is on information retrieval, not on logical reasoning and this is one of the most relevant difference between topic maps and formal ontologies. Moreover, TM standard defines an XML-based interchange syntax called XTM; the specification provides a model, a vocabulary and a grammar for representing the structure of information resources used to define topics and the associations between topics.

The main elements in the TM paradigm are often referred to by the acronym TAO which stands for *Topic*, *Association* and *Occurrence* [7]. According to ISO definition a topic is a symbol used within a topic map to represent one (and only one) *subject*, in order to allow *statements* to be made about the subject, that can be "*anything whatsoever, regardless of whether it exists or has any other specific characteristics, about which anything whatsoever may be asserted by any means whatsoever*". In substance a subject is anything about which the creator of a topic map chooses to discourse [6]; for instance an object, an event, a place, a name, a concept, etc.

An association represents a relationship between two or more topics. An occurrence is a representation of a relationship between a subject and an information resource. The subject in question is the one represented by the topic which contains the occurrence (for instance an occurrence can be a webpage, a book, an image, a movie depicting the subject).

Therefore two layers can be identified (Figure 1) into TM paradigm: a *knowledge layer* that represents topics and their relationships and an *information layer* that describes information resources [7].

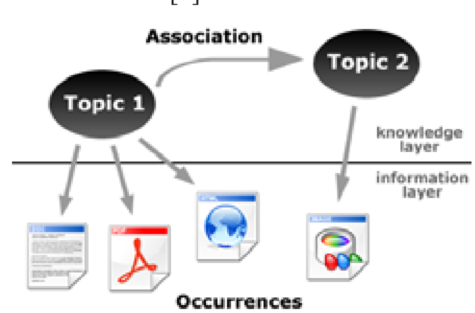


Fig. 1. Topic Maps paradigm: knowledge layer and information layer.

The existence of two different layers is one of the most interesting feature of this model; in fact the same topic map could be used to represent different sets of information resources, or different topic maps could be used to represent the same resource repository, for instance to provide different "views" to different users [7].

Each topic can be featured by any number of *names* (and variants for each name); by any number of *occurrences* and by its *association role*, that is a representation of the involvement of a subject in a relationship represented by an association. All these features are statements and have a *scope* that represents the context within which a statement is valid (outside the context represented by the scope, the statement is not known to be valid). According to ISO the unconstrained scope is the scope used to indicate that a statement is considered to have unlimited validity [6]. Using scopes it is possible to remove ambiguity about topics; to provide different points of view on the same topic (for example, based on users' profile) and/or to modify each statement depending on users' language, etc. [7]. Topics, topic names, occurrences, associations and associations' roles require a *type* element (becoming instances of classes). These classes are also topics, which might, again, be instances of other classes.

Therefore, to solve ambiguity issues, each subject, represented by a topic, is identified by a *subject identifier* (usually a URI, similarly to RDF). This unambiguous identification of subjects is also used in TM to merge topics that, through these identifiers, are known to have the same subject (two topics with the same subject are replaced by a new topic that has the union of the characteristics of the two originals) [8]. This feature could be used in order to share the knowledge and to solve redundancy issues.

3 Use of Topic Maps to design learning paths

Recent evolutions, in education as well in ICT, are leading designers and developers of e-learning systems and services towards the adoption of new criteria and models in the process of instructional design. The proposed scenario is featured by the use of the Topic Maps paradigm as a model for the design of learning paths, exploiting the flexibility and the expressivity of such a paradigm.

Currently, the design of educational paths in the context of web-based courses is mainly oriented to the serialization of teaching materials with the aim of creating self-contained learning objects (according to the standard SCORM) [9]. Despite assets and learning objects are designed with the aim of allowing great reusability, accessibility and interoperability, it is a common users' experience that some criticism may reveal itself, depending on the development process.

In the daily practice, teachers as well as instructional designers have to deal with synopsis definition of their courses, by outlining main subject matters which drive the structure of the lectures and single learning units [10].

Several research projects have been developed to investigate the use of repository systems for collecting and sharing learning objects with characteristics of being standard, re-usable, and searchable by means of suited semantic services based on the use of metadata associated to each of them. Despite good practices and the above criteria are used, depending on the specific needs of teachers, or students, or even of the course itself, it may happen that produced materials are not suitable for different applications. To face to this problem, the possibility of moving the generalization level from the contents to the definition of the contents' scheme is here investigated. It is worth noting that the inner architecture of topic maps is multilayered and thus it implements the same principle so that, within a semantic environment, different resources can be associated to the same concept and in different scenarios the same course can have different contents for even different targets, according to the scope defined within the description of the TM itself [10].

In our opinion, Topic Maps can be profitably considered as a means for describing the structure of a course as well as the outline of a lesson according to the logical structure of the course itself.

In order to support learning paths design process, we propose an ontological model (Figure 2) intended to be implemented in e-learning content authoring environments. In a preliminary step, the following requirements have been defined: formalisation (the model must describe course structure in a formal way, so that automatic processing is possible) [11]; pedagogical flexibility (the model must be able to describe learning contents that are based on different theories and models of instruction) [11]; centrality of student (the process of learning paths design must be based on learners' profile);

centrality of learning objectives (the process of learning paths design must be based also on preliminary specification of instructional objectives); personalization (the model must be able to define learning paths which can be adaptively matched to users' profile); domain-independent (the model must be able to represent learning paths regardless of content domain); reusability (the model must be able to describe contents structures reusable in other contexts); interoperability (the ontological model definition must be independent of specific particular knowledge representation languages, so that it can be applied in different e-learning tools and environments); medium neutrality (the model must be able to describe learning contents regardless of publication formats) [11]; compatibility (the model must be compliant to available learning objects standards) [11].

The *Learner* is the root element of the model. Firstly, it is required to identify all the students that attend a course; they can be defined as individuals or groups (the specification of learners depends on a process of user profiling not described into the ontology).

For each learner, it is necessary to specify the *OverallGoal*, the general learning aim of the *Course*. The learning objectives can be organized into a taxonomical structure (*OverallGoal*, *Objective* and *SubObjective*) which match with a hierarchical structure of the contents (*Course*, *Module*, *UnitOfLearning*). It is important to note that is possible to define propaedeutic relationships among objectives and, as a consequence, among modules and units of learning.

For each unit of learning, it is possible to identify the map of concepts, founded on topics and a limited set of relationships. The model defines two different *TopicType*: *PrimaryType* and *SecondaryType*. The first one includes the concepts that are considered requirement of the unit of learning and that, as a consequence, have no learning resources associated. The second one includes the key-concepts of the unit of learning that have specific learning resources associated. Among these secondary topics, it is possible to establish the followings relationships: *isPartOf* (a part-whole relation that may be used to represent, for instance, a paragraph and its sub-paragraphs within a learning object), *isRequirementOf* (a propaedeutic relation that may be used, for example, to organize the sequence of learning objects), *isRelatedTo* (defines a close-relation among two or more topics that can be used, for instance, to establish a connection among different course contents), *isSuggestedLink* (defines an indirect association among two or more topics that may be used, for example, to connote in-depth link).

Moreover, for each secondary topic we can specify a value of *Effort* (a generic element that may be useful to define informative data, such as the expected learning time, the difficulty, university credits, etc.).

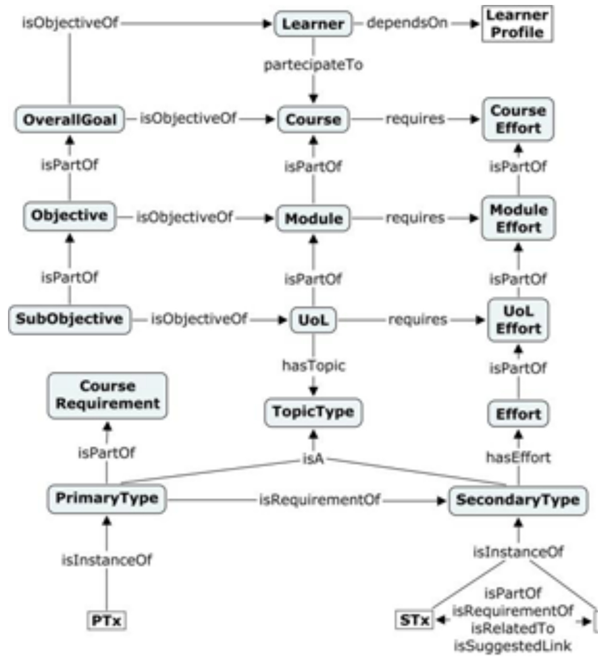


Fig. 2. The learning paths design model.

This structure of instructional content can be stored within a topic map in a well formed form and in a standard language, thus, it can be easily exported over the Internet and many systems can re-use and interoperate with the XTM representation of the topic map.

Moreover, the layered structure also enables authors to define different maps based on a common repository or archive of resources so that personalized learning paths can be defined while the contents at the occurrence level remain the same and different educational strategies can be implemented.

The same application can be investigated looking at the inner structure of a SCORM compliant learning object. One *Organization* can be here considered; the tree shaped structure composed by single items is equivalent to the one that will be represented within the related teaching unit; the hierarchical structure will be translated in terms of a Lesson, divided into Sections, sub-Sections, and so on. Topics and resources will be associated to these elements. The given sketch of this structure is written inside the related Learning Object in the manifest file (imsmanifest.xml file in SCORM).

Based on this standardized layout, the design of Learning Objects can be partially automated when the relevant resources for the educational objectives can be retrieved at the occurrence level of a suited topic map. Moreover, in default of relevant resources at the occurrence layer, a semantic representation of the relationships among educational contents could help the instructional designer to retrieve other materials linked to super-topics (*isPartOf* relation) or to other topics semantically related (*isRelatedTo* relation), facilitating the Learning Objects design.

By means of TM and XTM, and looking at the associations, the designer can build a sequence for the occurrences and the related topics; hence, reasoning (i.e. browsing a TM and making queries at a semantic level) on a given argument is made possible by simply looking at its description and thus automating the production process of retrieval of related contents (trough metadata) into SCORM objects [10].

4 Related Works

In the last years some research projects have been developed to investigate the use of Topic Maps paradigm in e-learning context.

QUIS (QUality, Interoperability and Standards in e-learning) is an EU funded project whose activities are directed towards quality in e-learning, interoperability and reusability of learning material. In the course of project development, a repository of standards in e-learning has been created and a requirement specification for a next generation of e-learning system has been produced. This requirement specification has a holistic pedagogical approach and requires an on-line learning environment that provides possibilities for personalization. The researchers suggest that TM could be used to achieve a personalized user interface, and present a prototype of a Personal Learning Environment (PLE) based on Topic Maps model [12].

According to Koper [13] "*an important question related to the educational semantic web, is how to represent a course in a formal, semantic way so that it can be interpreted and manipulated by computers as well as humans*". The semantic representation of learning courses opens the possibility to solve some problems like the development of flexible, problem-based, non-linear and personalized web-based courses; the building and sharing catalogues of learning and teaching patterns; the adaptation to learners' profile and the semantic research capabilities.

TM4L (Topic Maps For e-Learning) is an e-learning environment that provide authoring and browsing support for creating ontology-based learning content and/or structuring digital repositories. It uses topic map-based overlay structures for encoding domain knowledge and connecting it to instructional resources, which are considered relevant to a specific domain [14].

Ouziri [15] has proposed a TM-based approach to represent learning resources and associated semantics such as metadata. The main goal of his work is to enable more efficient accessibility and reusing of web-based learning resources, given in a common ontology. Karabeg et. al. [16] have proposed an information structuring methodology, called Polyscopic Topic Maps, as basis for flexible and exploratory learning. It consists of a framework for structuring information based on TM data model and the concept of scope. One of the most interesting features of this project is the chance to design personal learning path taking care of the prerequisites.

5 Conclusions

In this paper the Topic Maps standard and a semantic model for learning paths design has been presented. The main goals to achieve are interoperability of educational contents, reusability of both the contents and their knowledge structures, personalization of contents and services.

In order to test the ontology and to verify requirements (especially the compatibility with e-learning standards), currently, we are using it to model some courses at the University of Genoa. At the same time, we are investigating other approaches that use semantic technologies for learning paths design, in order to compare our model with them. One of the most interesting solutions is that implemented into Intelligent Web Teacher (IWT), a learning platform that enables users to create their own personalised learning paths exploiting a knowledge representation tool which allows teachers to define and structure disciplinary domains, by constructing domain dictionaries and ontologies. [17].

Moreover, we are interested in methodologies that allow formalizing the knowledge acquisition process and, as a consequence, the ontology-driven conceptual analysis. In regard to this issue, we are exploring OntoClean, a formal methodology that applies domain-independent notions,

used for ontological analysis in philosophy (such as, essence, identity, and unity), to analyze conceptual modelling process. These notions are exploited to define a set of meta-properties which permit to characterize relevant aspects of the intended meaning of the properties, classes, and relations that make up an ontology [18].
Nevertheless, the proposed scenarios have to be carefully considered about the risk of over-engineering of the knowledge and about a possible wrong interpretation of education which has not to be considered a mere summing of learning contents. Learning is a social process first of all and it cannot be limited with technical and/or formal boundaries [11]. Based on these considerations, it is worthwhile noticing that the proposed approach must be accompanied with the design of well suited educational strategies and supporting services, also in fulfillment of the recent results in the field of socio-constructivist theory.

References

1. C. Petrucco, *Le prospettive didattiche del Semantic Web*, Atti Didamatica 2003, pp.168-176.

2. J. Park and S. Hunting (Eds), *XML Topic Maps: Creating and Using Topic Maps for the Web* (Addison Wesley, Boston MA, 2002).

3. S. Newcomb, *A Perspective on the Quest for Global Knowledge Interchange*, in: Park J., Hunting S. (Eds), *XML Topic Maps: Creating and Using Topic Maps for the Web* (Addison Wesley, Boston MA, 2002).

4. H. Holger Rath, *The Topic Maps Handbook*, 2003 (unpublished), p.45; http://www.empolis.com/downloads/empolis_TopicMaps_Whitepaper20030206.pdf.

5. ISO/IEC JTC1/SC34/WG; http://www.isotopicmaps.org/.

6. ISO/IEC IS 13250-2:2006: Information Technology - Document Description and Processing Languages - Topic Maps - Data Model. International Organization for Standardization, Geneva, Switzerland; http://www.isotopicmaps.org/sam/sam-model/.

7. S. Pepper, *The TAO of Topic Maps*, Proceedings of XML Europe 2000, Paris France (2000); http://www.ontopia.net/topicmaps/materials/tao.html.

8. L.M. Garshol, What are topic maps?, 2002 (unpublished); http://www.xml.com/pub/a/2002/09/11/topicmaps.html.

9. ADL Advanced Distributed Learning, Sharable Content Object Reference Model (Scorm); http://www.adlnet.org.

10. G. Adorni, M. Coccoli, G. Vercelli, and G. Vivanet, *Topic Maps e XTM per l'e-learning*, Journal of e-Learning and Knowledge Society, n°3 September 2007.

11. R. Koper, *Modeling Units of Study from a Pedagogical Perspective: the pedagogical meta-model behind EML*, Heerlen: Open Universiteit Nederland, (2001); http://dspace.learningnetworks.org/handle/1820/36?mode=simple.

12. L. Kolås, A. Staupe, *The PLExus Prototype: A PLE Realized as Topic Maps*, Seventh IEEE International Conference on Advanced Learning Technologies (2007).

13. R. Koper, *Use of the Semantic Web to Solve Some Basic Problems in Education*, Journal of Interactive Media in Education (1), Special Issue on the Educational Semantic Web (2004).

14. D. Dicheva, C. Dichev, *TM4L: Creating and Browsing Educational Topic Maps*, British Journal of Educational Technology - BJET, 37(3), 391-404 (2006).

15. M. Ouziri, *Semantic integration of Web-based learning resources: A Topic Maps-based approach*, The 6th IEEE International Conference on Advanced Learning Technologies, (ICALT), Kerkrade, The Netherlands, July 4-7, pages 875-879 (2006).

16. D. Karabeg, R. Guescini, and T. Nordeng, *Flexible, Exploratory Learning by Polyscopic Topic Maps*, Proc. of the 5th IEEE International Conference on Advanced Learning Technologies ICALT 2005, Kaohsiung, Taiwan, July 2005.

17. G. Albano G., M. Gaeta, and P. Ritrovato, P., IWT: an innovative solution for AGS e-learning model, Int. J. Knowledge and Learning, Vol. 3, Nos. 2/3, pp.209–224 (2007).

18. N. Guarino and C. Welty, An Overview of OntoClean, in S. Staab and R. Studer, eds., The Handbook on Ontologies, pp. 151-172, Berlin:Springer-Verlag (2004).

Dealing with SCORM standard in a virtual communities environment

Luigi Colazzo
University of Trento
Department of Computer and Management Sciences
Via Inama, 5
Tel. +390461882144
colazzo@cs.unitn.it

Andrea Molinari
University of Trento
Department of Computer and Management Sciences
Via Inama, 5
Tel. +390461882344
amolinar@cs.unitn.it

Nicola Villa
University of Trento
Laboratory of Maieutics
Via Inama, 5
Tel +390461882339
nicola.villa@unitn.it

Abstract: *Many Learning Management Systems (LMSs) are available nowadays to accommodate synchronous and asynchronous distance learning. LMSs can play an important role in the evolution from traditional education to distance learning. There are also relevant needs for compatibility and reusability of educational contents exchanged among different LMSs. An attempt of responding to these needs is the "Shareable Content Object Reference Model" (SCORM) standard, which gives a framework for content producers consumers in order to create, distribute and use educational contents. However, creating (or adapting) a learning platform in order to be SCORM-compliant is a complex task, that introduces some rigidities inside the platform itself and (most of all) in the way users can interact with the platform and add content to the platform, following modern approaches to learning. The paper presents these problems, and the solutions adopted in our virtual communities platform to solve the issues deriving from being fully SCORM-compliant.*

1. INTRODUCTION

The paper describes our experience as designers, developers, experimenters and administrators of a virtual communities system (VCS) used by the Faculty of Economics of our University.

The system has been completely rewritten and the new version is at present in use at our university as a technological platform for projects of blended learning. The system is called *On Line Communities* and will briefly be described in chapter two of this paper. At present we are rewriting the software in order to adapt it to the requirements of our local government. In fact, recently *On Line Communities* has been adopted as a model by the Autonomous Province of Trento for developing their own technological platform for Lifelong Learning (LLL) projects.

In this paper we want to discuss what types of problems are connected to the development of a new learning environment for the public employees, in particular regarding the adoption of new types of Learning Object (LO), as a requirement of our commitment. Basically, *On Line Communities* considers LO to be any kind of electronic document that can be used/exchanged through the net: from a text to a complex didactic storyboard. This choice will remain substantially unchanged, but among the requirements of the new application, there is the need to use didactic material complaint to Shareable Content Object Reference Model (SCORM) standard of ADL [3]. The specifications of the structure of a SCORM package is based on the use of meta-data through which it is possible to represent: a) the navigation methods and the fruition sequence of the educational material; b) Learning Objects Metadata (LOM) of a specific course, that is the aggregation model (CAM).

The work is organized in the following way: in section two we'll propose a quick overview of our virtual community system, called *On Line Communities*, introducing also the new lifelong learning project called "L3". In section three we will describe the approach used to integrate SCORM packages into our system, and the problems that a standard can create within a pre-existing application. In particular, we want to concentrate on the impact of the standard onto different training experiences. The discussion of this point is crucial, due to our needs to reuse the didactic material not only inside but also outside the platform, for example the lifelong learning project for the public employees, and other future projects that will involve our university team.

2. LIFELONG LEARNING AND ONLINE COMMUNITIES

In this section we will recap the main features of *Comunità On Line* (*On Line Communities*) portal [4], the

VCS used in our projects. The idea to equip the faculty with such a platform goes back to the end of the 90s, and in the meantime the platform has been totally restructured two times, with several partial restructuring processes, starting from being an e-learning system and evolving towards the idea of managing and supporting virtual communities.

The objective of the new On Line Communities system was to create a collaboration space for people connected to the web, where it could be possible to widen the virtual space for relationships among the actors. It was therefore ready for didactic processes, but not only: research teams, recreation groups, boards, friends, colleagues etc.. The core of the application is composed by some abstract entities, i.e., Virtual Communities as aggregation of People to which some communication services are available in order to obtain some objectives. In detail, a virtual community [5], [6], is a space on the web dedicated to a collaboration objective, populated by people who communicate among each other, using a series of communication systems. With this approach it could be possible to represent all the hierarchical relationships between different types of community (like Faculties, Didactic Paths, Master Degrees, Courses, etc.).

On Line Communities had been experienced with limited number of users since 2003, and was finally released on early 2005. As from 2005 it was used by the whole faculty of Economics of our University in all its components (students, teachers, dean, secretaries, administrative staff, external partners) and others faculties are using the system in many courses. At present the system has more than 1400 active communities, 7500 users and more or less 1 million real accesses since November 2005.

L3 (Lifelong Learning) is a project commissioned by the Autonomous Province of Trento (PAT) to our research group, whose aim is to create a collaboration platform for training projects within the PAT itself and connected entities. In the specific case, our intervention in the first phase has a special valence in an organizational and technological support for the fruition of a master in e-procurement available for 100 managers of the Public Administration of our region.

One of the constraints our local government has imposed for the system, concerns the possibility to use material consistent with the SCORM/AICC standard within the platform. Such requirement needs to redesign some of the services available in the original system built for the university, since the metaphor of community contrasts with that of a traditional course contained within the SCORM packages.

3. SCORM AND VIRTUAL COMMUNITIES ENVIROMENTS

On Line Communities is a multi-standard system, that's to say a system purposefully neutral compared to the nature of objects in use. There are various reasons for this decision taken many years ago and which has proved itself most productive. To sum it up:

- A Learning Object (LO) by definition can be a very simple or complex object, from a text to a very complex software: to put a constraint on LOs accessible through the VCS does not seem to be reasonable and recommendable.
- A VCS aimed at university teaching has a high variability in the nature and format of the LO, for example, an LO containing a textual demonstration of a theorem will almost certainly be written in Latex, while the text "*De bello gallico"* could be made accessible in Html rather than in PDF file.

University lecturers, have to change/update their courses from one year to the next, and often prefer to use their own LO rather than those available on the market or on the web. Furthermore, a lecturer would hardly use the LO of another lecturer unless s/he wanted to make a point. This, therefore, eliminates the need for a Repository and also a large amount of the meta-data.

As mentioned before, the extension of our e-learning platform to the new lifelong learning experience, needs new features, explicitly required by the new conditions that lifelong learning settings will create. Starting from simplest services, one of these concerns the integration of SCORM-compatible didactic materials into the platform. Being strictly required by the Italian central government on one side, and being not well supported on the market too, the full implementation of SCORM standard in all its facets inside On Line Community has not been a simple operation, especially on the side of supporting the dynamic metaphor of the virtual learning communities. The lectures included in a SCORM package are, in fact, typically static, which means that the content, once packed into the SCORM "zip" file, may not evolve according to the various users' experiences. This approach differs significantly from our platform, and especially from our aim to offer an high collaborative environment to students, in which content can evolve according to the shared experiences of the communities' users. On the other hand, however, the evolution of our portal to a fully SCORM-compatible system (with a total integration of the SCORM packages) could kill its independence from the various types of educational materials. We want to avoid as much as possible, to be tie to a particular standard; at its end (which is not uncommon in the ICT world) it will be necessary return from scratch to build our system. We want to include the SCORM material, as every other didactic material (PowerPoint slides, word and PDF documents, etc.), and be able to manage them transparently and independently from SCORM (or other) standards.

Our choice has been to proceed leaving the tutors the possibility to publish materials created outside On Line Communities. We adopted the solution of integrating the use of SCORM packages into the download service of material related to a course in that specific service, therefore without creating a specialized “SCORM service”; the motivation was mainly to maintain the user service of didactic material “transparent” to the format of the materials themselves. On the contrary an application widely used as Moodle™ adopt this strategy separating the didactic material from those SCORM-compliant. As an example, imagine some notes regarding a lecture taken with a normal word processing software: in such a case, no one is expecting a specialized service called “word processing material”, or a “presentation material” in case of some PowerPoint slides: how could we manage a PDF file, while we are in the end talking about educational material? So this simple thought lead us to avoid to characterize SCORM-compliant material, have a special section and treatment and have it in a special section different from the download section. So the lesson created through SCORM standard, can be downloaded from the material section just like the notes of the lessons. Of course, the SCORM package have a lot of information inside the manifest, and these information could be very useful for the user and for the system itself. In the download section of our application, near each SCORM file, it is possible to access to a specific interface that provides the user with the tree of the material contained inside the SCORM package. The user, using the package tree, can display the structure of the package (and of the lecture) and download the single resource (Asset o SCO) contained in it, such as, for instance, an image, an html page, a PDF document, a video etc. One relevant problem indeed, especially for a traditional LMS with high user loads as our *On Line Communities*, concerns only in minimal part the access and creation of a tree-view of the material from the IMS manifest file contained inside the package. This fact can be done as a following process, when the author uploads the SCORM package: the real issue is the decompression of the material inside the SCORM package when the user wants to download only chunks of the material itself. The real problem is therefore an architectural problem: the SCORM package that represents the course (or the module, or the lesson) especially with videos and html pages, can contain huge amount of data and a relevant number of files. Having to compress them in real time for many online users, with other many users connected, with many users requesting the same package is a situation that could put the system under an incredible stress. The fact is not, in our tests, a borderline situation: imagine in fact a teacher that finally uploaded the latest, fundamental lecture before the final exam. Last but not least, all this work of analysis, browsing and decompression could be done to extract just one single file. To solve this relevant problem introduced by the structure and the nature of SCORM packages, we analyzed different strategies related to the original upload operation:

- upload the SCORM package normally as a single file and then, during the download phase, carry out the “on-demand” decompression for the index file (IMSMANIFEST.XML) and for the various components of the package
- as above, but with immediate decompression after the upload of the index file, in order to have it ready for the browsing phase, when the user is seeking the desired chunk
- upload a SCORM package and at the same time carry out the decompression into a second repository, where the whole SCORM package will reside, and therefore where all the singles chunks of the SCORM package will be immediately available to every request.

We have chosen the third hypothesis because it reduces the risks of system deadlocks and time-outs caused by decompression workload to serve multiple users. So once the SCORM package is uploaded, the system decompressed it with a background process in a secondary folder, so we have both the manifest and all the content ready for browsing and downloading. The adopted solution sees as main drawback the doubling (at least) of the storage occupation which, however, permits to avoid having to decompress “on the fly” upon request by the user bulky SCORM packages, perhaps for several users at the same time. We have observed, for example, that when the teacher sends a note of new didactic material available, especially during the lecture periods, On Line

Communities has a steep rise of accesses and requests for downloading material averagely of fifteen times, and moreover within few minutes. It is clear that we will never be able to avoid these peaks (like every web application), and therefore SCORM package both in its hierarchic form and in its decompressed form already available makes the task easier.
An open question that we have been asking ourselves concerns the applicability of a “web 2.0”, social networking approach to SCORM-compliant contexts like our On Line Communities platform. This kind of collaborative and cooperative approach unhinges the classical concept of static lesson in which only the teacher could generate contents and knowledge. In such ideal context of web 2.0, unfortunately we notice how the use of SCORM type packages could be a critical element, because it is not suitable to this type of services that require the classic “CRUD” (create, read, update, delete) primitives. A lesson created according to the standard is, in fact completely static, i.e., consisting only of the information that the author had decided to offer to those following the course. Moreover, it is not possible to add new ways of interaction with the contents: the list of the lectures will be adopted by the student in “as is” mode. This creates a relevant limitation because it transforms static what could be an interactive path; in fact the teacher controls the administration of the contents used by the users. These reasons, added to what said above, confirm the choice we have done. The result is that the integration into the system does not mean the complete adoption of the standard but rather to offer to the users the possibility to use their own material quite freely.

5. CONCLUSIONS

Two main problems arise for most educational organizations in sharing educational contents through e-learning platforms: content compatibility and system extendibility. In order to solve content compatibility, the SCORM standard provides a model for content exchange among different learning platforms [1]. The purpose of SCORM is to achieve interoperability, durability and reusability within SCORM compatible content [2]. At present, many development technologies exist to include / attach SCORM compatibility to existing LMS. Many educational organization have built such a system, customizing most of the times to their own needs, from the simplest web site to the most complex adaptive LMS [7]. The problem is how to extend these new functionalities, especially when an extension like the compliance to SCORM standard has such a relevant impact on the platform. This paper necessarily presents a synthetic view of the problem created, within already an existing virtual communities platform, by the introduction of this compliance. SCORM is such an important topic for these systems, and in general for didactics assisted by ICTs, that it cannot be ignored or simply “attached” to a system. The use of SCORM standard has some profound architectural implications on the platform itself (for example, regarding the inaccessibility of some technologies that are used). We discussed these problems and presented the solutions applied to our On Line Communities platform, in which the basic metaphor, i.e. the idea of virtual community, offers advantages of implementation and possibilities for further development, in contrast with the closure created by the simple concept of “course” as a container of SCORM material. The solution we described in this paper reflects this philosophy; we have not created SCORM-compatible communities but we have foreseen the possibility to use to the lectures in an accessible way within the space of the community that we are offering to our users. In future, we want to extend the possibilities offered through the inclusion of a tool that could create SCORM packages, but we don’t want to limit LO to be created through it. In fact, with the solution adopted we are SCORM-compliant and, at the same time, we allow users to manage their materials and contents with no restriction imposed by the standard.

References

[1] A. P. Polyxeni, M. M. Nektarios, C. D. Stavros, *Supporting Interoperability an Existing s-Learning Platform Using SCORM*, The 3rd IEEE International Conference on Advanced Learning Technologies (ICALT ‘03), Athens, Greece, July 09 - 11, 2003.

[2] ADL Technical Team, *SCORM Version 1.3 Application Profile Working Draft Version 1.0*

[3] ADVANCED DISTRIBUTED LEARNING (2004) SCORM 2004 3rd edition - Content Aggregation Model

[4] Colazzo L., Haller H., Molinari A. (2007) WAYS – What Are You Studying? A community tool for learners on the move, ,Proc. of ED-MEDIA 2007, June 25- June 29, 2007; Vancouver BC, Canada

[5] Jones S. G., *Cybersociety*, Sage, London, 1995.

[6] Jones S. G., *Virtual Culture*, Sage, London, 1997.

[7] K. G. Aharalampos, S. P. Demetrios, *Integrating Adaptive Educational Content into Different Courses and Curricula*, Educational Technology & Society, Vol 3, 2001.

A Study of Moodle Learning Platform on the Internet Online Music Course

Wen-Chun Chang[1], Mei-Ju Su[2]*, Heng-Shun Chen[3], Yu-Huei Su[4]
[1] Hukou Elementary School;
[2]Graduate Institute of Electronics Engineering, National Taiwan University;
[3]Department of Medical Informatics, National Taiwan University College of Medicine;
[4]Department of Music, National HsinChu University of Education
Correspondence to: Mei-Ju Su, Graduate Institute of Electronics Engineering, National Taiwan University. No.1, Sec. 4, Roosevelt Rd., Da-an District, Taipei City 10617, Taiwan
E-mail: merrisu123@gmail.com

Abstract

In an era of rapid technological developments and information conveyance, diversified digital learning is no longer confined inside classrooms. Internet technology and appropriate internet platforms are the best channels for diversified teaching Moodle learning platform on the internet is an example of such online teaching course design management and online learning platform not bound by time and space. Similarly, a musical instrument learned by means of information technology may not require face to face teaching. This study has setup an online recorder learning course, recorder E-learning through Moodle learning platform on the internet. The course design uses ADDIE 5-stage course design method with students of Hukou Elementary School, Hsinchu County as study subjects. This study of learning platform on the internet integrated into the music subject provides students a learning environment that allows them to readily and repeatedly listen to music. Hopefully, students' music learning interest will be cultivated and their ability to play the recorder can be enhanced.

Keywords: music education ,moodle, e-learning

Introduction

Moodle learning platform on the internet

Ou Chang-Chia（2006）defines digital learning as a learning method that integrates online and offline learning strategies and teaching activities using digital teaching materials and methods facilitated by users that adopt information communication technology as a medium. The internet learning environment evolved from one-way browsing of internet data to webpage production software design and interactive webpage, and into the teaching platform on the internet as we know it today. The learning mode does more than allowing users to browse data or download files, now, a message board, a discussion forum, a comment section, and homework submission are also available. The two-way file upload and interactive activities serve as an interactive teaching platform between the teacher and the students. [1]

Free Software Moodle (known as "magic beans" by internet users) is a course and software package in the internet-base of PHP language. It was created by Martin Dougiamas. (Moodle company founder, based in Perth, Western Australia) It is a type of internet course management system, CMS and continues to be developed. The purpose of its design is to provide a construct framework of education. Moodle is free software with the source code opened up. (http://www.opensource.org/docs/definition_plain.html) Moodle can be used on any computer that supports PHP (http://www.php.net/) and other types of database. According to the official website statistics data, 36,748 websites from 205 countries in 65 languages registered for Moodle teaching platform as of January 2nd, 2008.（Especially MySQL (http://www.mysql.com/) It shows Moodle is widely prevalent software.[2]

Moodle learning platform on the internet allows the teacher to organize and design teaching activities based on the principles of teaching. The teacher then engages in internet digital teaching material production and collects internet resources related to digital teaching materials such as texts, pictures, voices, videos, and teaching webpages etc. Through related module functions, of Moodle internet teaching platform, they are organized into internet online resources that meet teachers' teaching course requirements. Students are also asked to submit e-homework, participate in internet discussions, and engage in internet homework evaluation learning activities through this learning platform on the internet. In short, teachers and students engage in teaching and learning activities through Moodle learning platform on the internet. [3]

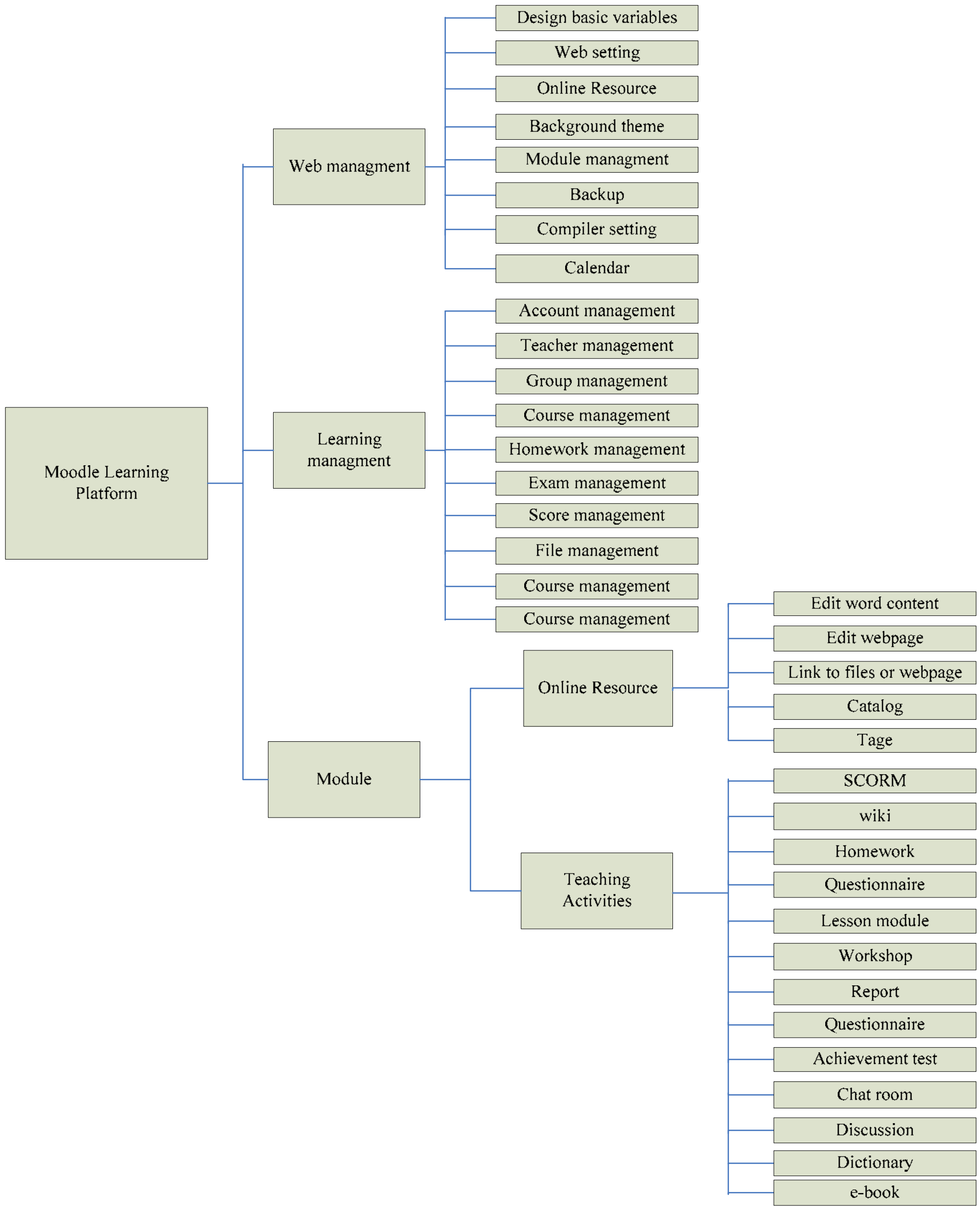


Fig 1. Moodle function architecture diagram

Existing conditions of integrated information in teaching in Taiwan

The Ministry of Education has planned the public internet learning system in its implementation of "Challenge 2008～E-Generation Training Program." It also invested 1.2 billion dollars to develop internet learning contents for elementary schools and high schools, and plan services such as teacher, expert, proprietor in the same trade, and private group participations in the development of internet learning contents, teaching material production, technical support services, platform setup and maintenance etc. Open-format is established under the same standards to promote free software, resource integration and exchange. With contributions from specialists in the national defense and military field together with doctorate degree holders and master degree holders, internet learning contents and platform related works are developed. (Wu Tieh-Hsiung, 2033) meanwhile, to promote free software applications on campus, the Ministry of Education also OSSACC that serves as the guidance provision unit for free software applications in elementary schools and high schools. At present, The Dept. of Education, Hsinchu County is actively involved in the promotion of free software application operations. It is planned and promoted by HCC EDU. Development and Information Institution. The focus lies in prompting Moodle learning platform on the internet. Moodle learning platform on the internet study activities and teaching seminars are also held to promote Moodle learning platform on the internet setup and teaching applications on campus. Teachers integrate information into teaching materials to upgrade teaching efficiency and quality. In this study, Moodle 1.5.3 Chinese version learning platform on the internet software is used. [4]

The purpose of developing "Recorder E-learning "online course teaching design is to make up for the insufficiencies of site teaching. Therefore, vocal performances, recorder playing, listening, and appreciation for music are teaching activities done in the music class. Since the study subjects are fourth grade students, it is a crucial period for them to lay foundation for recorder playing. As a result, it is expected that internet platform based courses be integrated into the versatile music teaching methods so that students' learning interest and recorder playing performance can be enhanced. In addition, recorder playing evaluations, students' learning feedback forms, and intra-school and extra-school Moodle course design evaluations may serve as reference for improving Moodle internet course contents related to "recorder E-learning" and internet teaching strategies.

Methods

This study is a music internet online course based on recorder teaching. Information technology based teachings are incorporated into music teaching. Where applicable, theme related animation productions may be scheduled to help students develop appreciation for the teaching materials. Students will be able to perceive visual art and the beauty of music art with the aid of visuals and sounds. With Moodle internet platform course teaching activities students' learning methods will be information based learning. Students will not only learn about knowledge and skills related to music and visual art, they will also become competent in information technology use.

Moodle Tool:

The software tools used in this study include: Hukou Elementary School Moodle learning platform on the internet, and Yu Jung E-learning." The Moodle website is setup by the Information Section of our school. Moodle 1.8.4 Chinese Version is adopted. The interface is translated into Chinese by the volunteers at Hukou High School, Hsinchu County. Moodle function is illustrated in Fig 1..

ADDIE internet course design method

The internet course development in this study has been facilitated by ADDIE systematized teaching design method. (Molenda, 1991) The five-stage course development covers: analysis, design, development, implementation, and evaluation. The targets and results of each stage are described below:

l. Analysis: It is a process that determines teaching contents including demand analysis, learner analysis, environment analysis, and teaching content analysis etc. The purpose of this process is to confirm teaching goals, contents and learner behavior at start point. A survey is conducted to understand student learning background and condition through "Student Learning Survey."

2. Design: It is a process that forms the teaching activities. The contents include teaching strategies, script writing, and teaching evaluations etc. This process is developed with the aid of Moodle learning platform on the internet of Hukou Elementary School.

3. Development: it is a process that involved teaching material production including a variety of teaching resource productions such as teaching material writing, and internet online course setup. This process is developed through software tools such as teaching demonstration audio/visual files, and internet teaching digital teaching material software.

4. Implementation: it is a process of teaching activity conductions to obtain teaching results. This process is completed through recorder evaluation check lists, and learning feedback forms etc.

5. Evaluation: it is the result of teaching design applied. It is used to determine whether teaching results have reached expected standards including intra-school evaluations and extra-school evaluations. Evaluations in this study are conducted using Moodle internet course design evaluation form.

Related software:

1.Server end：Webpage server（Webserv1.2.10 including Apache 2.0,PHP 5.1.1,MySQL 4.1.16,phpMyAdmin 2.7.0）FTP file transfer software, text edit software, and decompress/compress software etc.

2.User end：Webpage browser and media player etc.

Results

The framework of "recorder E -learning" contents include introduction of the recorder family, recorder solo demonstrated by the teacher, accompanying recorder playing and the music score, student sharing, assessments among peers, and learning feedback forms etc (see Fig.2-6).

Moodle internet online course "Recorder E-learning" integrated into music teaching is as shown in Fig. 7. This "recorder E-learning" is evaluated by ten schools (inter-school assessment) that have applied Moodle in

Hsinchu County. The total score is 455 points with the average of 45.5 points. It is assessed as the best course out of the 20 Moodle online courses developed by elementary schools in Hsinchu County.

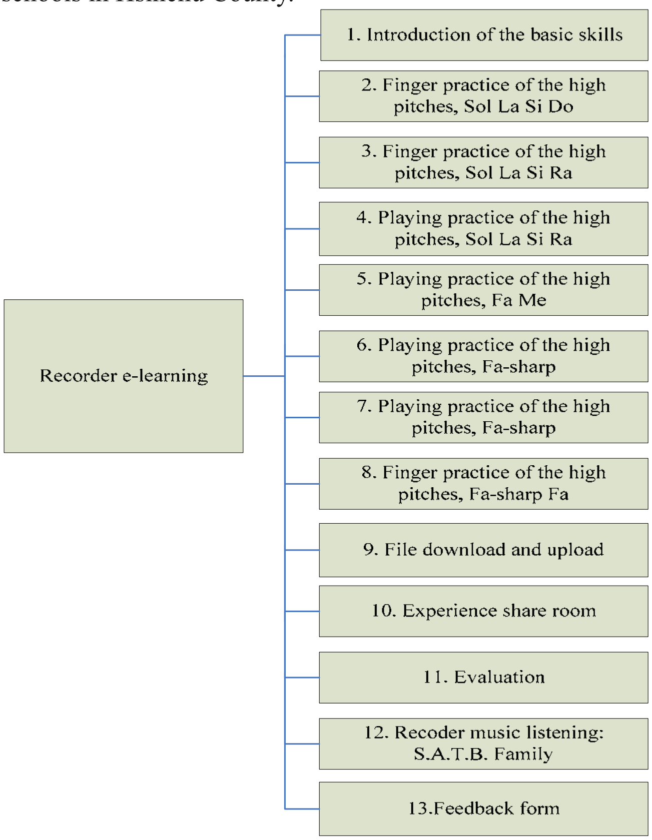


Fig2.Recorder e-learning function architecture

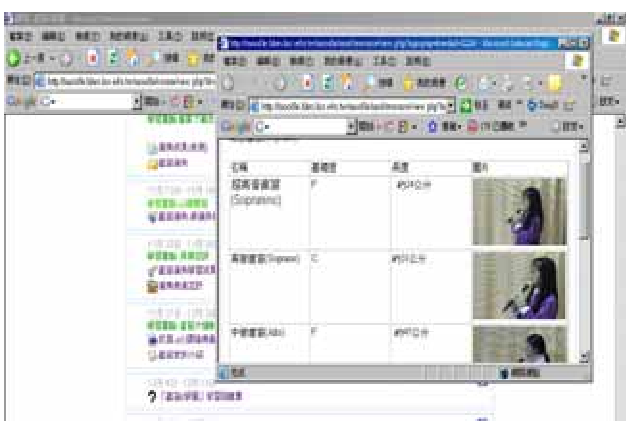


Fig.3: Introduction of the recorder family

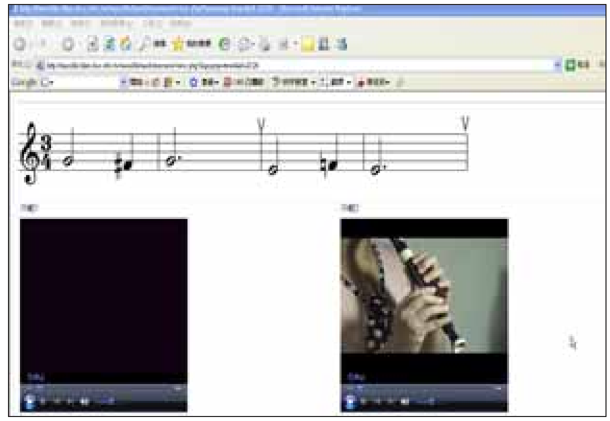


Fig.4: solo demonstrated by the teacher

Fig.5: recorder music score

Fig.6: student sharing among peers

Fig.7: integrated into music teaching

Discussion

After implementing internet course "recorder E-learning" in music teaching, it is found that this classroom teaching mode helped students to view the teacher's demonstrations in big classes that had not been possible in the past. During "recorder E-learning" implementations, students were able to see the teacher's instructions on their personal computers. They were also free to choose and practice repeatedly music verses that were not familiar to them. Meanwhile, since the teacher had first collected and produced many internet online resources, the students were free to click them for learning purposes. By using the computer, naughty students seemed to be able to concentrate on computer learning as compared to that of classroom teaching. [5]

Nevertheless, during the internet teaching process, the students were asked to pay attention to intellectual property. Unauthorized public pasting and distribution were forbidden. Digital teaching resources from the "Learning Fueling Station"（http://content1.edu.tw/） and "six learning websites"（http://learning.edu.tw/sixnet/） setup by the Ministry of Education are authorized for teaching use in the nation. Teachers need not worry about intellectual property related issues when accessing these sites.

From the learning feedback form of "recorder E-learning", it is found that although the students are fond of "recorder E-learning" and believe that it helps them to learn, they do not agree they will be able to learn to play the recorder faster. The reason for this is because during recorder playing, the teacher's supervision is required because the teacher will be able to correct the students immediately. When a student plays well, the teacher will also encourage the student right away. Therefore, "recorder E-learning" requires interactive teaching of the teacher in order to achieve maximum results.

References

[1]Collins, B. Applications of Computer Communicationsin Education: An Overview. IEEE Communication Magazine, 1999:37, 82-86.

[2]Ou Chang-Chia, Moodle digital learning course management system, CMS, Taipei: Wen Kuei Information, 2006

[3]Molenda, M. Philosophical critique on the clains of "constructivism.", Educational Technology, 1991:31（9）,44-48。

[4]Wu Tieh-Hsiung .Existing Conditions and Development Trends of Digital Learning in our Country, Extracted from 2003 Annual Information Education Seminar Dissertation Collections, Tainan: National Tainan University, 2003

[5]Collins. B, Goal-based Scenarios and the Problem of Situated Learning: A Commentary on Andersen Consulting's Design of Goal-based Scenarios. Educational Technology, 1994: 34(9),30-32.

Digital Art and Quality Living─Study of the Feasibility of Computerizing Senior Citizen Art Therapy

Serina Lai[1],Yen-Shan Lin[1], Yen-Ting Chen[2], Yu-Huei Su[3], Mei-Ju Su[4], Heng-Shun Chen[5*]
[1] Angel Art Gallery; [2]Seattle Central Community College; [3] The Innovation Research Center for the Arts and Culture Industry, National HsinChu University of Education; [4]Graduate Institute of Electronics Engineering, National Taiwan University; [5]Department of Medical Informatics, National Taiwan University
[*]Correspondence to: Heng-Shun Chen , N1.Jen-Ai Road, Taipei ,Taiwan,
chenhs@ntu.edu.tw

Abstract

With a rapidly ageing society in Taiwan, care of the elderly has received attention from both academics and industry to include the themes of integrating art and healthcare. In order to counter the barriers of care for ageing citizens Angel Art Gallery uses digital art which beyond the boundaries of time and space to develop their enjoyment of art. It offers a new path of art therapy for senior citizen with the collection resources of the Angel Art Gallery and its collaborated famous artists of different fields. After careful planning, onsite trials and evaluation by many consultants, three sections of artistic creation, nature scenes, and animal and plant ecology were implemented using digital video contents. Through a U-Care system, the digital art contents were streaming through broadband Internet and play onsite via a Babybot, which is a telecare device at remote site. User requirement and awareness survey were performed on 150 subjects along with preliminary evaluation the effectiveness of the digital art therapy at the Angel Art Gallery and nursing homes, senior apartments, and art classrooms in Northern Taiwan.

Keywords: Senior Citizen, Computerized Art, Art Therapy

Introduction

Need for Art Therapy in an Aging Society

With the progress of the standard of living and medicine and health, the average lifespan of humanity has gradually extended. According to statistics concerning Taiwan's elderly population from the Council for Economic Planning and Development (CEPD), the percentage of the population 65 or older was 7% in 1994. Taiwan has already started to become an "aging" society as defined by the United Nations. [1]

In recent years, care of the elderly has received attention from the government and the private sector, and resources have been invested to provide benefits or to found related industries. Of these, humanistic art therapy, as opposed to physical therapy, has gradually gained notice. How can the mental life of the elderly be improved? How can more opportunities and choices for beneficial cultural participation be provided for the elderly? What specific kinds of culture and arts do the elderly need or prefer? Can opportunities to enjoy art be transmitted to senior citizens without regard for the boundaries of time and space in order to develop their enjoyment of art? Great importance has already been placed on an aging society's need for art therapy.

The nature of art therapy is based on concern, love, and respect for human nature. Art therapy is psychological therapy using art forms as a tool or media that allows people to explore personal problems and potential through the experience of verbal and non-verbal expression and artistic creation, in order to assist people to have a harmony of their inner world and exterior world.[2] Two kinds of methods can be applied in art therapy: focusing on the creative process and focusing on appreciation and sharing. 1. Focusing on the creative process: artistic creation itself is the therapy, and the process of artistic creation can ease the conflicts or annoyances that a person is feeling, and is helpful to self recognition and self growth. 2. Focusing on appreciation and sharing: Using art as psychological therapy, as some associations that arise from the created artwork and within the artwork can help individuals to maintain a harmonious balance between their inner world and the exterior world.[3]

Computerized Art Offers a New Path for Senior Citizen Art Therapy

For senior citizens, using art forms as a tool or medium of psychological therapy and changing and finding a release for emotions through the conversation between art and the spirit, achieves the goals of self exploration, self recovery, and self actualization.[4] On one hand, senior citizens' lack of a feeling of being which causes loneliness can be consoled, and the discomfort caused by physical deterioration and weakness can also be lessened on a certain level due to a relaxed mood. On the other hand, culture and art stimulation also can give senior citizens a new goal to pursue, raising their quality of life by deepening their mental world.

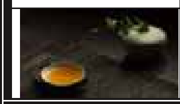
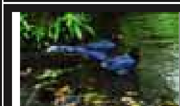
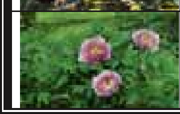
Artistic creation film	
	Staying With and Above the Nature──Short Animated Film on Dai Wu-Guang's Paintings Duration: 8 min. 35 sec. Making modern ink and color paintings into interesting animation using digital production.
	The Mind Changes with the Scene──Short Film on Guo Dong-Tai Still scene photography
Natural scenes film	
	Seeing Taiwan from Above──Short Film on Qi Bo-Lin's Aerial Photography Duration: 15 min. 20 sec. Editing aerial photographs into an exciting film.
Animal and plant ecology	
	Formosan Blue Magpie Duration: 20 min. 55 sec. Film on the Art and Ecology of the Formosan Blue Magpie.
	Enriching Your Life with Bloosom Duration: 11 min. 31 sec. Making a collection of peony photographs.

Figure 1: Digital Art Films the Angel Art Gallery Has Developed for Senior Citizen Art Therapy

As stated previously, art therapy can be divided into the two types of focusing on the creative process and focusing on appreciation and sharing. Many researches have been done over the years in the former, which has always stressed the inspiring quality received in the immersion of the participant in the artistic creative process.[5] As for the later, art sharing and appreciation, it places the focus on the interactive relationship between the participant and the artwork, and in the past, it was often oriented toward spaces of art display such as museums and art galleries. Furthermore, because the concept of elderly participants and art therapy has not yet been universally accepted, the need for artistic content has long been seen as a non-essential need, and was thus ignored. In addition, senior citizens do not have the same mobility as younger people, and they have a limited range of activity, so the ease of acquisition of artistic content has often become the main factor for the success of the promotion of art therapy focusing on appreciation and sharing. [6]Therefore, using the development of computer technology to computerize art and break through the limitations of time and space has become a new path that is worth trying.[7]

Methods

Angel Art Gallery attempts to combine creativity, art, and handicrafts with technology to promote enjoyment of life. Currently, they are also trying to get involved in the senior citizen art therapy business and exploring its feasibility. We used the collection resources of the Angel Art Gallery and cooperated with famous artists of different fields. After careful planning, with the onsite examination and evaluation by many consultants and with the goal as senior citizen art therapy, we divided it into the sections of artistic creation, nature scenes, and animal and plant ecology, to implement the planning of the creation of digital video content. The preliminary results included five digital art films (Figure 1):

After making the short art films with digital technology, we further used the method of field interviews and streaming through Babybot (Figure 2)which is telcaring for the elderly(Figure3) to evaluate the effectiveness of the computerization of art therapy within the field of senior citizen art therapy. Researchers brought videos and questionnaires and started at nursing homes, senior apartments, art classrooms, the Angel Art Gallery and telcare members, all in northern Taiwan. We implemented onsite interviews, collecting the raw data from the ones engaged in art appreciation to act as evidence for later analysis organization. The methods of interview investigation included audio recording and filling out questionnaires (Figure 4). This research summarizes the recordings of conversations during the visits and the participants' answers to the questionnaires into the experiences and viewpoints of each elderly participant to evaluate whether or not this batch of film contents achieved effectiveness for senior citizen art therapy.

Figure 2: Babybot for Senior Citizen Art Therapy

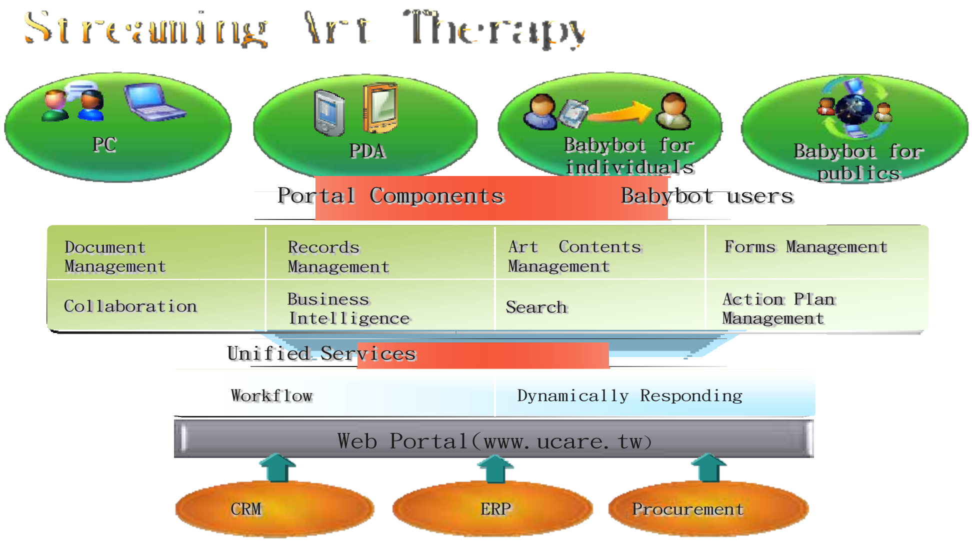


Figure 3: Streaming Service for Art Therapy

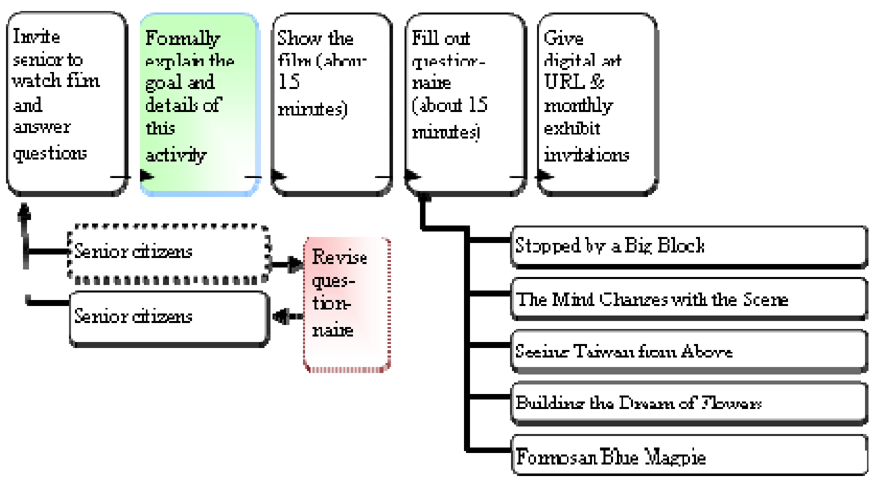


Figure 4: The Onsite and Evaluate Steps for Senior Citizen Art Therapy

Results and Conclusions

When evaluating the effectiveness of digital video art on senior citizen art therapy, the questionnaire we designed had three main areas: survey indexes of personal data, art characteristics, and digital video content. Furthermore, the great amount of specific data we collected shows how much time senior citizens are willing to spend on artistic activity programs, the motivation and goal for doing artistic activities, and the preferences for different types of art. This will help as a reference for content design when computerizing senior citizen art therapy.

The results shows: (1) In the North of Taiwan, there are 50% of senior citizens participating in artistic activities is about 1 week. (2) 60% of senior citizens determine their art appreciation itinerary by themselves. Through this survey, it can be seen that the interests of 70% of seniors are nature scenes and ecology art. (3) The motivation and goal for doing artistic activities depended upon themselves of which making up a percentage of 35. The 25% of senior citizens think of the convenience of transportation is the foremost factor. (4) 41% of senior citizens think of how to increase the beauty in life, purifying the spirit as the usual reason in participating in art-related activities. (5) After viewing the films (digital film content), the feeling of Increasing the beauty in life, purifying the spirit are produced by 46% of senior citizens. Furthermore, 31% of senior citizens created the feeling of Calming down, relieving tension.

London psychologist Bernie Wood once said "Allowing the audience to be absorbed in the enjoyment of a film can heal wounds buried deep in the psyche.[8] Film therapy is a breakthrough in medicine."Making an analysis on the types of art therapy, digital films based on art can at least provide the effects of art therapy and music therapy, especially the two parts of art therapy of art appreciation/sharing and color therapy. Angel Art Gallery has used digital films in senior citizen art therapy and its practical results have confirmed the importance of culture and art in the lives of seniors and proved the feasibility of bringing art into the lives of seniors through computerization of art. These analysis results will act as an important reference for the next stage of plan improvement.

References

[1]Heath, Sharon. Fan Rui-Zhen(2007), trans. ART@TUDES: Mastering the Art of Your Attitude. Taipei: Wunan Publishers.

[2] E. Ulman; Art therapy_ problems of definition; American /journal of Art Therapy;2001,vol.40

[3]Jiang Han-Sheng(2002). Banquet of the Four Seasons: Jiang Han-Sheng's Musical Prescriptions. Taipei: China Times Cultural Publishing.

[4]Wu Jie-Zhen(2005). Space Scenery and Public Art—Spirit, Healing, and Art. Taipei: Artist Publishing.

[5] W. E. Fredrickson; Perception of tension in music_ musicians versus nonmusicians; Journal of Music Therapy, 2000, p40-50

[6] C. Larson; The Fine Art of Healing the Sick; U.S. News & World Report, 2006, 140 no21 54-6 Je 5

[7]Zhang Tian-Jun(2001). Zhang Tian-Jun's Famous Painting Treatment Room. Taipei: China Times Cultural Publishing.

[8]Huang Long-Jie(2004). Poems of the Psychological Treatment Room. Taipei: Living Psychology Publishers.

Virtual Reality, VR Digital Art Website Setup and Teaching Applications

Chuan-Cheng Chang[1*], Yu-Liang Chang[1], Ya-Chen Huang[1], Yu-Huei Su[2]
[1]Department of Arts and Design, National Hsinchu University of Educationl;
[2]Department of Music, National HsinChu University of Education
*Correspondence to: Department of Arts and Design, National Hsinchu University of Educationl,521 Nan-Da Road, HsinChu ,Taiwan
E-mail: chengc@mail.nhcue.edu.tw

Abstract

Virtual Reality, VR is an innovative technique that provides rapid and precise perspectives and realistic mapping effects. Users are introduced to the computer simulated virtual world. As VR technique reached maturity including interface design, interactive technique, and software/hardware devices etc. allow users to interact more naturally and concretely. Its application aspect has expanded to digital art, design, and education. The Virtual Reality Laboratory, National Hsinchu University of Education has applied techniques adopted at the virtual art museum to demonstrate the simulated virtual display space. At present, the simulated interactive tour to "Shou Shan Yan Guan Yin Temple" has been completed. Two types of digital learning modes, "Auto Tour" and "Free Exploration" have been developed. They integrate with internet teaching and elicit students' learning interest and enhance their knowledge through human-computer interaction, HCI. Meanwhile, students' aesthetic cognitions toward 3D space can also be cultivated. In the "Lin Pen Yuan Family Garden" VR setup in this laboratory, other than entering the interactive simulated space using the keyboard, a plug-in pedal called the dance machine has also been researched and developed to allow users to be immersed in the scene while moving his lower limbs.

Keywords: Virtual Reality, VR, traditional art, architecture

Introduction

VR technique has made remarkable progress in recent years. This computer simulated physical technique has been applied in many high-value industries including computer games, recreational industries such as interactive movies, flight simulations, educational training such as sports simulations, virtual surgeries, virtual 3D supersonics in the medical area, virtual museums, visual art design in virtual stages, acoustic simulations, and sound technology for noise prevention tests. (Huang Shih-Yu, 2008).

Since 2002, "Two Trillion Double Stars Facilitation Plan", a six-year program, has been promoted in Taiwan. Digital contents such as 3D games, internet multimedia and information related to 3D animations, communication, and biotechnology etc. have been listed as leading-edge industries. (Kuai Liang, 2008)Other than 3D/VR digital contents that have become the current industries of focus, instantaneous interactive virtual and animated games have also become the mainstream industries. They are extensively applied in museums and large gaming centers such as the Gold Museum in jinguashih, the 921 Earthquake Museum of Taiwan in Wufeng, Taichung, the Sword Lake Mountain (Shock Theater) in Kukeng, Yunlin, and the 3D, 4D and 4D hands-on exhibit of Nice to Meet You. (Huang Shih-Yu, 2008).

National University of Education VR Laboratory applied the technique adopted by the virtual art museum to display the virtual space of architecture. Currently, the VR interactive tour to Shou Shan Yan Guan Yin Temple, Guishan Township, Taoyuan County has been completed. Digital learning modes have also been integrated to conduct traditional temple architecture teaching. The VR development of National Historic Site Lin Pen family Garden is also underway in the laboratory. Other than the use of interactive virtual space using the keyboard, a plug-in pedal called the dance machine has also been researched and developed to allow users to be immersed in the scene while moving his lower limbs. The beauty of the traditional architecture will be introduced through game playing.

Methods

The laboratory is not bounded by traditional tour modes such as narrations, texts, pictures, and commentaries at fixed locations; instead, the following methods are used to establish VR , which is applied in teaching:

1.Setup 3D environment and exploration contents of VR:Using 3D VR technique, the three dimensional form of the architecture is presented. The topics are classified to shift the images before the tour items. The decorations, bracket sets, horizontal inscribed boards, history, and background of enshrined Buddhist Gods are introduced with picture aids and film-showing.

2.Design virtual exploration scenarios based on digital learning modes: By means of Auto Tour and Free Exploration modes, traditional architecture and art teaching is conducted while explorations on space and interactions also take place

3.Conduct internet teaching experiments to determine results:By imparting simultaneous or non-simultaneous internet teaching not confined by time or place, the architecture curriculum is adopted to conduct the teaching experiment. A students' knowledge test and a students' attitude survey are also conducted to determine results.

Results

1.VR website setup of Shou Shan Yan Guan Yin Temple completed

A website has been setup for this VR. After entering the homepage, left click the mouse on tour to enter the 3D environment menu. Press the link in the red box to enter the 3D homepage menu as shown in Fig. 1.

Fig.1: Website homepage of Shou Shan Yan Guan Yin Temple
http://www2.sses.tyc.edu.tw/homes/upload/VR/templeArts/index.htm

2.Auto Tour and Free Exploration digital learning mode setup

Shou Shan Yan Guan Yin Temple VR interactive learning is conducted through two learning modes namely: Auto Tour and Free Exploration. Press the 3D button of tour teaching mode on the left right of Fig. 2 to enter Auto Tour. It is played similar to slide showing. Users simply click on the image to switch frames in sequence as to classification. Users may not explore as they wish. Press Free Exploration on the lower left of Fig. 2 to move freely in the virtual space. The users may explore the 3D space from different angles. The teacher may also help students simulate a tour or present an oral tour with the aid of a projector.

Fig2.Menu of Auto Tour and Free Exploration

3.Complete the knowledge content setup of Shou Shan Yan Guan Yin Temple

The content setup of the website is completed based on the knowledge structure of temple art to allow students to cognize and learn about temple arts and cultures including the auspicious architectural decorations, the architectural forms, the horizontal inscribed boards, and the enshrined Buddhist Gods etc. (As shown in Fig. 3) Students will understand the layout of the temple and the symbolic significance of the constituents. The architectural forms, features, history, and folk beliefs of the temple will also be introduced to cultivate students' knowledge for local arts and cultures and the connotations behind them.

Fig. 3Auspicious decorations

4.Develop the 3D VR of National Historic Site Lin Pen Yuan Family Garden

The Lin Pen Yuan Family Garden in Banciao is the only well-preserved and intact Lin Yuan architecture in Taiwan. The carvings of the architectural decorations are exquisite and grand. The wooden carvings, stone carvings, and brick carvings are all considered masterpieces. It is the place where the treasured traditional architecture is located. This architecture is presented in 3D VR by our laboratory. (Fig. 4 and Fig. 5) Other than using the keyboard to input the location of the architecture the user wishes to look at, the dance machine also allows him to simulate bodily movement by moving his lower limbs in all directions. The user will be immersed in the scene. It is particularly suitable for children as the wall-less beauty of art from the museum will be brought before the user that allows him to maneuver.

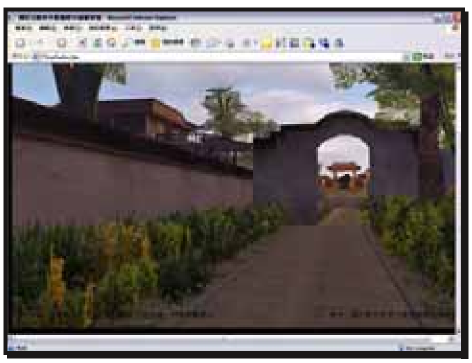


Fig.4 Enter《3DVR of National Historic Site Lin Pen Yuan Family Garden》

Fig. 5 Maneuver《3D VR of National Historic Site Lin Pen Yuan Family Garden 》

5.Digital learning of traditional architectural art combined with internet teaching

Through internet teaching, in this research, the arts and cultures of "Shou Shan Yan Guan Yin Temple" are cognized and learned.《3D VR of National Historic Site Lin Pen Yuan Family Garden 》serves as the course design intended for Lin Yuan explorations. After the teaching experimentation, the cognition test and the attitude survey will determine the results of the VR teaching applications.

Discussions

Since VR is able to present the real virtual world and overcome the problem of cancelled fieldtrips due to weather, environment, and geography, it is considered highly suitable when used to conduct culture tours. When using VR as a tour tool, personnel expenditures can be cut down and visitors' learning efficiency and interest can be enhanced. More importantly, it brings the public closer to art and historic relics. When applied in elementary school and high school art teaching, it will effectively enhance students' artistic accomplishments.

In terms of user interface design, VR is an interactive interface that is easy to use. Students may use it at home or in school in between classes. They may preview or review on the internet or search online to find answers they need. This laboratory has combined digital art presentation and digital learning mode to setup a virtual exploration space for traditional architecture digital art. Students may develop learning interest, increase knowledge for traditional architecture art, and cultivate aesthetic cognitions for 3D space e through game playing. Students' understanding and love for traditional architecture knowledge will be further enhanced.

References

[1]Bell, John T. & H. Scott Fogler (1996). Vicher: A Virtual Reality Based Educational Module for Chemical Reaction Engineering. Computer Applications in Engineering Education, 4(4).

[2]Byrne, C.M. (1996). Water on tap: The Use of Virtual Reality as an Educational Tool. PhD. Dissertation, department of Industrial Engineering, University of Washington, Seattle, WA..

[3]Chang Ching-Shu(2006).Shou Shan Yan Guan Yin Temple "VR" Tour 【 Online Sea 】 http://www2.sses.tyc.edu.tw/homes/upload//VR/temple_Arts/index.html

[4]Chang Chuan-Cheng(2005).A Study of Art Teachers' Viewpoints on the Present and Future Internet Art Learning Trend Survey.

[5]Cheng Wei-Jung（2003） "VR, Video Studio" Contemporary Internet Art Appreciation, Yang Ying-Feng Digital Art Museum Coverage, Mei Yu Bimonthly, 134th issue.

[6]Huang Shih-Yu（2008）A Preliminary Study of VR Applied in Architectural Space Design, Graduate

Institute of Spatial Design, National Yunlin University of Science and Technology http://teacher.yuntech.edu.tw/yangyf/hh/hhc3.html

[7]Kuai Liang（2008）A Preliminary Study of Innovation Culture Industry-Exemplified by YAOX EDUTAINMENT CO., LTD, Humanities and Arts and Innovation E-Newsletter, 5th issue http://www.cc.nctu.edu.tw/~cpsun/innovation-e-paper-5%20(conference-kuailiang).htm（8 Feb 2008）

[8]Lin Cheng-Hung (1997). An Insight of VR, Taipei, Gotop Information Inc.

[9]Lin Pei-Chun, Chuang Hao-Chih(2002).A Study of Digital Art Aesthetics, Mei Yu Bimonthly, 130th issue

[10]Lin Pen Yuan Family Garden(2006).http://www.linfamily.tpc.gov.tw/_file/1373/SG/23099/D.html.（10 Feb 2008）

[11]Shou Shan Elementary School(2006). School-based Curriculum on line.http://www2.sses.tyc.edu.tw.

[12]Taiwan Ship Net Online http://www.ship.org.tw/vr/index.asp（26 Jun. 2007）

[13]Tim Guerra, & Dan Heffernan (2004). The Guerra Scale, Learning Circuits, http://www.learningcircuits.org/2004/mar2004/guerra.htm

[14]William Winn(1993). A Conceptual Basis for Educational Applications of Virtual Reality(R-93-9). Seattle: University of Washington, Human Interface Technology Laboratory.

[15]Wang Ting-Ming (2003). New Ideas for Art Instruction and Internet Learning, Mei Yu 134th issue Wang Hsiu Hsiung(1998). Appreciation, Cognition, Interpretation, and Judgment: Theories and Practice of Art Appreciation Instruction

[16]Wang Nian-Tsan, Shen Wan-Ru(2007). Introduction of Virtual Art Museum, Mei Yu 157th issue, P.26-33.

A RULES-BASED SYSTEM TO ACHIEVE TAILORING OF SCORM STANDARD LOS

Pierpaolo Di Bitonto, Teresa Roselli, Veronica Rossano
Department of Computer Science – University of Bari
Via Orabona 4, 70125 Bari
{dibitonto, roselli, rossano}@di.uniba.it

Informatics systems supporting distance learning do not yet fully exploit the available technological potential. They often do little more than present information without subjecting the teaching content to any form of intelligent processing to tailor it to the needs of the user. Aim of the present work is to define an agent program (in other words the rules that allow logical reasoning to be executed) to set up an intelligent agent that can change the sequencing of a SCORM standard LO according to the student profile, offering the student different learning paths within the same LO even if such alternatives are not inherently present in the LO. This diversification process is based on the known characteristics of the user, her/his prior knowledge, learning style, etc. After consulting this information, the agent can decide what teaching strategy to adopt, in what sequence to present the available teaching content, and therefore in what way to interact with the user. To carry out this guiding action, the agent obtains the information on the user profile, processes it, infers what strategy to adopt and chooses what changes to make according to the LO on which it is operating.

1. Introduction

Informatics systems supporting distance learning do not currently exploit to the full the potential (methodological and technological) of the information technologies now available. These tools often do little more than present the information content, without making any attempt to subject this content to intelligent processing to tailor it to the specific needs of the user. In fact, they do not allow personalization of the teaching to the true needs and abilities of the individual, nor the adoption of the many different existing teaching strategies. They often remain tied to a mechanistic view of the teaching process, using the results of exercises only to make reports on the students' progress rather to directly influence the subsequent learning experience. They are unable to make an independent assessment of the cognitive characteristics of each user, an essential requirement for true tailoring of the learning process.

The present article illustrates one part of a wider research project that aims to create an intelligent agent [1, 2] that will be able to infer the individual student's learning styles and change the presentation of a SCORM standard Learning Object (LO) to adapt it to these styles [3].

The focus in this work is on those aspects related to the intelligent agent reasoning processes, i.e. the rules present in the knowledge base (KB) that dictate the intelligent behavior; the architecture has already been illustrated in [4].

Aim of this article is to define the intelligent agent program (in other words the rules that allow the execution of the logical reasoning) that allows the LO sequencing to be tailored to the student profile, offering different learning pathways inside the same LO, even if the management of such alternatives is not inherently present in the LO itself. The tailoring process is based on the known characteristics of the user, her/his prior knowledge, way of learning etc. After consulting this information the agent can decide what teaching strategy to adopt, in what sequence to present the available teaching content, and hence in what way to interact with the user.

To carry out this guiding action, the agent obtains the information on the user profile, processes it, infers what teaching strategy to adopt and chooses what changes to make, according to the LO it is operating on.

The second section of this article presents the technological and methodological premises underlying the architectural proposal, the third section describes the agent program in detail, and the fourth section makes some conclusive remarks and outlines future lines of development.

2. Technological-methodological premises

To gain a full understanding of the work described, a brief introduction needs to be made to the methodological and technological issues faced. In particular, the agent-based approach must be dealt with and contextualized to the present work; and a brief description of the IEEE LOM [5] (Learning Object Meatadata) language made, as well as of the essential features of the SCORM standard allowing interoperability between LOs and the LMS (Learning Management System).

2.1 Agent-based approach

According to Russell and Norvig an agent is any software agent that can perceive its surrounding environment by means of sensors and modify it by using actuators. One of the most common techniques used to define an agent is the PEAS (Performance Environment Actuator Sensor) description. It is important to define the agent's performance so as to provide it with heuristics on which to choose the strategy to be actuated from among those available. The environment supplies the context of use, characterizing the problem that the agent is designed to solve. The actuators are the tools through which agent carries out actions on the environment. Finally, the sensors are the sources from which the agent acquires the information on the environment that it needs to work on.

Agents are structured in two distinct parts: the architecture and the agent program. The architecture makes the perceptions available to the program, carries out the program and passes the actions it has chosen to the actuators as they are generated. Instead, the agent program relates the perceptions to the actions.

In this particular context, the environment the agent operates on is the e-learning platform, and the performance is measured in terms of the adaptation of the resource to the teaching strategy. This adaptation is based on the information contained in the LOM, in the user profile and in the file manifest.

In other words the sensors are the parsers of the file manifest, metadata and user profile. The actuator is the component that deals with modifying the file manifest.

2.2 The LOM Standard

The LOM standard is a set of IEEE specifics that serve to describe teaching resources or their component parts. It includes more than 80 descriptive elements subdivided into the following 9 categories:

- *General*: this includes all the general information that describes the resource as a whole. The descriptors in this group include: title, structure, aggregation level
- *Lifecycle*: this groups the descriptors of any subsequent versions of the LO and of its current state
- *Meta-metadata*: these include the information on the metadata themselves
- *Technical*: this indicates the technical requisites needed to run the LO and the technical characteristics of the LO itself, such as the format or size of the file
- *Educational*: this contains the pedagogical and educational features of the LO. This is the most important category and contains elements like: Interactivity type, Learning resource type, Semantic density, Typical learning time, that supply indications on how the resource can be used in the teaching program
- *Rights*: this indicates the intellectual property rights and any conditions of use of the LO, such as cost, as well as the information on copyright.
- *Relation*: this describes any relations (of the type: "is a part of", "requires", "refers to") with other LOs.
- *Annotation*: this allows the insertion of comments about the use of the LO in teaching, including an identification of who wrote the annotation.
- *Classification*: this makes it possible to classify the LO according to a particular classification system.

The strong point of the LOM is that it can be used to describe many closed vocabulary items. This facilitates automatic processing of the information.

2.3 SCORM

This defines the LO-LMS (Learning Management System) interaction. It is organized in the form of four books: Overview, that presents the standard in general, Content Aggregation Model (CAM) that describes how the LO components must be packaged, Run Time Environment (RTE), that describes how the LMS manages running of the contents and memorization of the results of a test; Sequencing and Navigation (SN), that indicate how the LO components can be sequenced to create differential learning pathways. Only the CAM needs to be described herein, to aid understanding of this article.

The CAM describes the components used in a learning experience, how to create the content package, describe the resources so as to guarantee their retrieval and define the information relative to the sequencing of the components. It consists of four main sections: Content Model, that indicates how to define the components contained in a learning experience; Content Package, that indicates how to represent the behavior of a learning experience (Content Structure) and how to group the activities of a learning resource (Content Packaging); Metadata, that indicate how to describe the learning resources; Sequencing and Navigation, that indicate a rule-based model for dynamically defining the sequences and order of the activities.

In the Content Model there are various levels of aggregation, namely the asset, the SCO (Sharable Content Object) and the Content Organization. The asset is the most elementary form of a learning content. It is different from the others because it does not exchange information with the LMS. Some examples of assets are HTML pages or multimedial files. The SCO is the smallest unit that communicates through the API with the LMS. It is generally made up of several assets. Finally, the Content Organization is a map that represents how the content can be used in the form of structured subunits (activities). The activities represented in a content organization can also consist of other activities (sub-activities) that could in their turn consist of yet other activities. There is no limit to the number of levels of nesting of the activities.

The Content Package indicates how the content package must be created, paying attention to two essential elements: the physical files that make up the learning resource (file .gif, .html, …) and the file imsmanifest.xml (manifest), that describes the structure of the content and the resources associated with the package. The file manifest includes four main sections: metadata describing the package elements, organization containing the content structure and the organization of the learning resources used in the learning package, (sub)manifest(s) that describe any logically nested content unit. The file manifest indicates to the platform how the assets, SCO and Content Organization are to be navigated, by means of the Sequencing and Navigation rules.

3. Working logic

To illustrate how the agent program works, abstracting from the details that will be examined in depth later on, we briefly describe the logic according to which this intelligent agent carries out computations.
The agent's task can be summarized in terms of four fundamental tasks: defining the student's learning style, matching the LO with the student profile, to infer the teaching strategy to be adopted; presentation of the LO in a framework of adaptation of the navigation to the student profile and feedback obtained by the system from the student to improve the profiling and content personalization activities. The presentation and feedback activities are managed by the agent architecture, and the student profile-LO matching phase by the agent program.
In more detail, after filling in an initial questionnaire the system infers the user's learning styles (verbal, visual, …) and memorizes them in the student profile. The information on the learning styles is needed to choose the educational actions best suited to the individual student's characteristics. In fact, during the matching phase the system infers the teaching strategies to be adopted (lesson, tutorial, experiment, …) according to the student profile, the description of the LO (LOM file) and the information contained in the file manifest. The next step is to execute a personalized presentation of the LO, depending on the inferred characteristics, by means of the modified file manifest.
The main problem in these activities is that inside the LO to be adopted, the data needed to achieve compete tailoring of the resource, in line with the teaching strategies inferences gathered, may be lacking. Because the system may have several LOs on the same topic and the student model does not have a unique learning style for each student model, the agent will assess which LO is best suited to adaptation to the possible learning styles and vice versa, choosing the best relation between the learning style and the educational resource.
After presenting the LO, the agent processes the feedback collected by the system about the experience of use, so as to confirm or modify the information present in the student profile, with the aim of dynamically adapting its behavior during the course of the interactions to the student's characteristics.
The student profile-LO matching phase is described in more detail below, together with some of the other aspects outlined above.

3.1 Learning styles and teaching strategies

Personalized learning can be designed on the basis of various cognitive characteristics that affect the learning process of any student, and the efficacy of the learning process can be improved by taking these characteristics into account. The term learning style refers to the set of cognitive characteristics that discriminate the way a student goes about learning, determining her/his preference for particular learning strategies, resource typologies, etc. So, to structure an efficacious learning experience, these different learning styles (verbal, visual, field-dependent or independent, systematic, intuitive, impulsive, reflective, analytical, global), that are not mutually exclusive, need to be given due consideration [6].
The system under study derives the cognitive styles of each learner from the results of a questionnaire (drawn up by the Gruppo MT of the University of Padua (Italy) [7]) and memorizes them in the student model. The study conducted by the Gruppo MT defined the learning styles, characterized the typical actions of students belonging to each category and illustrated the most suitable teaching strategies for each of them.
The teaching strategy consists of the sequence of educational activities that the teacher sets up during the interaction with the learner with the aim of facilitating learning. Each strategy can be associated with one or more cognitive styles, based on the defined user profile. Below, two learning styles are illustrated, and the most suitable teaching strategies are reported.
The visual style is typical of subjects who propend for a visual-spatial language code, and tend to learn by means of images and activities based on visualization (e.g.: design, images, diagrams, tables, graphs, conceptual maps, etc.). The typical behavior of subjects of this type is: to pay attention particularly to the graphs and images in a text; memorize by means of schemes, maps and graphs; associate words to images. In general, students with a visual learning style retain a better memory of what they have seen, and tend to need to gain an overall view of the topic. For this reason the most suitable activities for this type of learner are: listing key points, using graphs, films and slides to illustrate complex processes and mathematical functions, using figures, schemes, graphs, diagrams and physical demonstrations, using graphic production in problem-solving so as to gain a visual understanding of the effects on a system of varying the parameters.
Instead, the reflective style is typical of students who analyze the study topic from every perspective and closely ponder each aspect before deciding. While this attitude can be an advantage from some viewpoints, it also means that these subjects need more time to answer or to carry out a given task.
Normally, a reflective student hesitates to intervene in a discussion because s/he prefers to be sure before giving an opinion, fears to make a mistake, does not like to be obliged to take part and takes a fairly long time to process information. The reflective student tends to be more cautious, more anxious, less tolerant of ambiguity and less willing to take risks.
A useful teaching strategy for this student typology is to present the teaching material by focusing on the practical aspects and leaving brief intervals to allow reflection on the materials presented.

3.2 Student profile - LO Association

The association between the student profile and the LO is what allows adaptation of the LO to the specific student's characteristics. In this article it is assumed that one and only one LO exists for each topic in the learning environment the agent is working on. This simplification was necessary because although a comparison among several LOs to determine the best relation between the learning style and the educational resource is currently under study and will be the focus of future developments, it has not yet been defined in detail.

The activity described below consists of three main actions: acquiring knowledge, choosing the educational activities and applying the chosen educational activities. Acquiring knowledge means transforming the information content of the metadata and file manifest into assertions and rules of KB, so as to be able to modify the manner of navigation of the LO according to the student's learning styles. After choosing the educational activities to be presented, the system instances the facts and rules that will actuate the modifications of the file manifest that will then be applied during the presentation phase.

3.2.1 Acquiring knowledge

The knowledge acquisition phase allows the data present in the LOM files and file manifest to be asserted as facts inside the KB. The rules shown in the table are an example of this.

Rule: condition-action	Description
∀ LO Structure (LO, X) and is(X, Atomic) => AggregationLevel (LO, 1)	Structure is an element in the General category; if it has the value "Atomic" then it has a low level of aggregation
∀ LO Structure (LO, X) and is(X, Collection) => AggregationLevel (LO, 3)	If a resource is a collection of more simple elements it can be inferred that it has a medium level of aggregation
∀ LO LearningResourceType(LO, X) and is (X, questionnaire) => interactivityLevel(LO, 3)	LearningResourceType is an element in the Educational category. If a resource is a questionnaire then it will have a medium level of interactivity
∀ LO InteractivityType(LO, X) and is(X, ActiveDocument) and InteractivityLevel(LO, 3) => LearningResourceStyle (LO, Simulation)	InteractivityType is an element in the Educational category. If a LO is defined from the interactivity standpoint as an "ActiveDocument" and it has a medium level of interactivity (3) then the learning resource style will be of simulation type, i.e. suitable for the study of students who have a preference for simulation type activities
∀ LO InteractivityType(LO, X) and is(X, Expositive) and InteractivityLevel(LO, 1) => LearningResourceStyle (LO, NarrativeText)	If a LO is defined from the interactivity standpoint as "Expositive" and it has a medium level of interactivity (1) then the learning resource style will be of simulation type, i.e. suitable for the study of students who have a preference for explanations of textual type and in any case with low interaction and little display.

Table 1. Examples of rules to state the facts inside the KB

The rules presented above to state the facts inside the KB are only a very small instance of those implemented. The resulting assertions help to discover the granularity of the educational resources making up the LO or to determine their characteristics.

On the basis of this information the agent can decide whether to present complex material or text or multimedial material.

The rules for asserting the information contained in the file manifest work in the same way. The underlying idea is to infer the way the resources can be navigated, on the basis of some key elements.

It should be noted that the rules presented provide a model that can be extended to any metadata language.

3.2.2 Choosing the educational activities

The choice of the educational activities to be done is guided by the learning styles. Templates[1] in the CLIPS language[2] have been used to describe the different cognitive styles, learning strategies and ways of navigating the LO.

The names and descriptions of some of the templates inserted are as follows:

- learningStyle: contains all the learning styles that could be attributed to a student
- strategy: contains all the learning strategies that can be tailored to each student
- sequencing_controlMode: contains the information on the control methods of the navigation sequencing of the LO

---

1 Template: a complex, abstract fact that defines in a single pattern a group of inter-related slots. A slot is a structure that associates one or more values to a name. The order in which a slot appears in the template is immaterial

2 CLIPS (C Language Integrated Production System): a tool for developing knowledge-based systems

- sequencing_attemptLimit: contains the maximum number of attempts the user is allowed to make of the activity being carried out
- sequencing_attemptAbsoluteDurationLimit: contains the maximum duration for each single learning attempt

The values present in the templates are processed by the rules that modify the file manifest.

To understand how the KB works, we can take as an example the rule that manages the navigation of the LO for a student classified as having a reflective learning style.

Those who have a reflective learning style tend to examing several alternative answers very carefully before answering. They proceed with care, taking a task step by step and taking a decision only after they have closely pondered the different outcomes. For this reason the most suitable teaching strategy is "tutorial", subdividing the activities to be carried out into small sequences. When considering the tasks in a sequential manner, this stype of student needs to navigate the learning resource by following the flow of activities. Moreover, requests to back-track must not be accepted. For this reason the values of the element *sequencing_controlMode*, that manages the navigation sequencing of the resource, are set at "*flow*" and "*forwardOnly*". Moreover, because these students need more time to esamine and reflect on the steps to take, the value of the element *attempt_AbsoluteDurationLimit* is increased.

3.2.3 Applying the educational processes

It is important to note that for each learning style a rule is set that can potentially activate given educational actions. However, since a student may come under the heading of several cognitive styles at the same time it is necessary to understand which criterion should underlie the choice of the educational process to activate.

For this reason, a rule has been inserted that calculates, for each cognitive style shown by the student, the set of changes to be made to the file manifest, and estimates their degree of applicability.

The strategy with the highest degree of applicability is chosen, which enables the rules modifying the file manifest to be instanced, and hence a new version generated. This version is sent to the LMS, that allows the tailored navigation of the LO.

3.3 Presentation of the LO and Feedback phase

The part related to the presentation of the LO and tracing the student activities is entirely managed by the system architecture and the LMS. In particular, the architecture writes into the file manifest the changes inferred by the agent program and passes the file manifest to the LMS.

When the LMS reads the file manifest it generates the activities tree and goes on, in accordance with what is present in the LO, to trace the contents. The collected data are memorized in the user profile and used to improve the tailoring of the learning activities.

In order to evaluate the effectiveness the questionnaire it has been administered to a larger sample of users. The data analysis will allow us to draw more detailed conclusions. Conclusions and future works

The paper presents a methodology for defining a quality questionnaire for the assessment of collaborative writing environments on line. The starting point of the study was the wide range of collaborative environment developed in latest years. The problem is in evaluation of such environments the variables to consider are many and have different nature because it is necessary to consider both pedagogical and informatics aspects.

For this reason in our study, aiming to assess collaborative writing environments, integrates ISO standard, for evaluating the software quality, and some other quality dimensions for evaluating the constructivist approach to learning. The questionnaire was defined using the GQM approach that allows to define 89 questions that should support expert (teachers, instructional designers, researcher, …) and non-expert users in choosing the collaborative writing environment that best meets their own learning/teaching goals.

4. Conclusions and future developments

The proposed work starts from an observation of the available teaching technologies that, although in some senses now highly mature and reliable, in others still lack full integration. For this reason the present contribution describes how the design of an intelligent agent can combine the currently available technologies in order to personalize the educational products on the market. In particular, we focus on the reasoning adopted by the agent to attain the goals. It is important to note that the proposed method for creating the KB is entirely general and can be applied to any description language and any compatible SCORM LMS. The expected result is that it will increase the adaptability of LOs to the user profile.

Many different future scenarios can be developed from the present proposal. In fact, as stated above, it is assumed that in the LMS there is a LO designed precisely for each topic. This is obviously a simplification and so a search agent is needed to individuate what LOs should be considered and adapted for use. Another future research direction is the creation of an intelligent agent that can infer, on the basis of the LO tracing, whether the student model is appropriate and modify it if not. A multiagent system taking into account such aspects is currently under study.

5. Acknowledgements

This work is supported by the GENOMENA project, funded by the Apulian Region.

References

[1] Doug Riecken: Conversation with Marvin Minsky about agents. In: Communications of ACM. volume 37, Number 7, pp. pp. 23-30.

[2] Stuart J. Russell, Peter Norvig: Artificial Intelligence: A Modern Approach, Prentice Hall, NJ: Upper Saddle River, 2003.

[3] Sharable Content Object Reference Model (SCORM) 2004 3rd Edition Documentation Suite http://www.adlnet.gov/downloads/DownloadPage.aspx?ID=237
[4] Di Bitonto, P., Roselli, T.: Dynamic Tailoring of Scorm Sequencing and Navigation Rules. Submitted to: IEEE International Conference on Advanced Learning Technologies (ICALT 2008), IEEE Computer Society Press, Santander, Cantabria, Spain.
[5] IEEE Learning Object Metadata (LOM) http://ltsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf
[6] Davis E.C., Nur H., RuRu S.A.A.: Helping teachers and students to understand learning styles " English teaching forum, Vol.32, No.3, 1994.
[7] Cornoldi C., R. De Beni e Gruppo MT, Imparare a studiare. Strategie, stili cognitivi, metacognizione e atteggiamenti nello studio, Trento, Erickson, 1993

Synchronous Distance Lectures in Second Life: Teaching Fundamentals of Informatics

Andrea De Lucia, Rita Francese, Ignazio Passero and Genoveffa Tortora

Dipartimento di Matematica e Informatica, Università degli Studi di Salerno,via Ponte don Melillo 1, Fisciano (SA), Italy

adelucia@unisa.it, francese@unisa.it, ipassero@unisa.it, tortora@unisa.it

Abstract

Second Life successfully supports synchronous communication and social interaction. Instead of participating in a distance classroom by watching videos and submitting homework by email, Second Life enables students to attend lectures in a classroom with the teacher and the other fellows. For this reason, the adoption of Second Life for synchronous distance learning is largely growing.

In this paper we present an experience of using Second Life for taking a synchronous lecture. A virtual setting is created in such a way to reproduce the department structure where the students are accustomed to have lectures. The evaluation of the experience provides encouraging feedbacks and outcomes on the student perceptions of presence, performance and satisfaction.

1. Introduction

Several Learning Management Systems, such as Moodle [14], Blackboard [1], etc., offer a wide range of functionalities to support e-learning. Often all these features are not fully exploited and these systems are simply used as a supplement, as content repositories, in Web Enhanced modality [2]. Multimedia is rarely used and the contents are mainly static, structured as HTML pages, PowerPoint™ or Microsoft Word™ documents.

The main goal to reach is to exploit the new technologies to effectively support collaborative learning. Computer Supported Collaborative Learning (CSCL) is a research area aiming at enabling students to work together, to reciprocally help, and to reinforce their social dimension. CSCL technologies can be grouped into two categories: asynchronous and synchronous. In asynchronous modality the students work at different times. With this interaction modality there is no problem with time zones or holidays falling at different times when the learning actors are spread around the world. Thus, international collaborative work is well supported. The most of the features offered by LMS, such as forum, wiki, calendar, homework, are asynchronous.

Synchronous communication enables student to work simultaneously together. Videoconferencing, chat or whiteboarding are some examples. Collaborative Virtual Environments (CVEs) are another communication mechanism used to collaborate [8][15][17]. CVEs enable users to share a virtual environment and to do highly synchronous collaborative work. In particular, virtual worlds such as Second Life (SL) [19] may represent the future of the human-computer interaction. The main advantages of SL is the low cost of the technology, the wide user community and the new web navigation metaphors, which naturally favors the social dimension of their residents. Indeed, SL is the most popular Internet-based virtual world, if measured by number of subscribers and money exchange [8]. It is largely adopted for academic, social and business purposes.

In SL "Residents" interact each other through emotional avatars, reaching an advanced level of social interaction. SL also proposes interaction patterns typical of video games and web communities metaphors. To obtain best learning results it is crucial to take advantage, in a fully technological enabled environment, of the abilities of today's students which can be considered technology power users [16]. Indeed, today's students have grown with Internet and video games and are usually very practice in multi-player on-line games and instant messaging and it is natural and pleasant for them to use it as learning environment.

SL adoption for distance learning is growing [20][21]. In particular, many universities, see [9][13] as an example, and other kind of organizations such as [3][6] are already using the SL Grid to support distance learning [18]. In [4] we described how to exploit the 3D programmable virtual world provided by Second Life to create an environment and a location for collaborative Learning. The experience of a course jointly offered in Second Life by the Harvard Law School and the Harvard Extension School is reported in [25]. No evaluation of the experience was provided. On the contrary, several research works have been devoted to evaluate learning in virtual environments. It is well established that presence is an aspect influencing learning. Witmer and Singer [24] propose a questionnaire for measuring presence, which is useful when all subjects experience the same type of environment [22].

In this paper we adopt the Witmer and Singer questionnaire for evaluate a synchronous lecture in a Second Life environment, ad-hoc created for giving lectures. Also student perceptions concerning performance and satisfaction have been considered and discussed.

The paper is organized as follows: Section 2 describes the communication support offered by Second Life, while the proposed experience is presented in Section 3. The evaluation of the experience is discussed in Section 4. Finally, Section 5 concludes.

2. Second Life

Second Life (SL) is one of the most popular 3-D Online Virtual Worlds. A Virtual World is a computer-simulated persistent environment similar to the real world. In SL users access the online system with a client and interact with content and other “residents”, represented by avatars, a user representation in the virtual world. The SL environment enables to construct 3D objects and provides a scripting language for developing interactive content. SL successfully supports synchronous communication and social interaction [9]. Instead of participating in a distance classroom by watching videos and submitting homework by email, SL enables students to attend lectures in a classroom with the other fellows.

Exploiting SL inter avatar communication features, animations and gestures are offered to augment face to face communication. As a drawback, gestures are somewhat counter intuitive to use and may even cause disruptions to conversations as one searches for the appropriate gesture or animation to augment his/her communication. Once in the environment, people have a first person perspective, they participate, do not only watch. Situational awareness, “who is there”, is well supported, as well as awareness on “what is going”. The possibility of seeing when a student does not actively participate is also an interesting feature. Indeed, it is easy to understand if a student is not active: his/her avatar visibly sleeps.

SL offers many tools for participating in the emerging social network: it is possible to join or create a group, send an IM to a group to get help from those logged in, send and receive group notices, add people to the friends list and more. Using these features it is possible to locate advice or help, or publicize events. It also enables residents to leave messages and other communications for other users to find when they log in later. SL allows for using web, video, audio streaming, slide presentations projected as images, and VOIP. People can privately as well as publicly chat on an open channel.

3. The proposed experience

In this section we describe an experience of usage of Second Life for synchronous distance learning. In particular, we illustrate the virtual environment we have created and the Second Life objects designed for carrying out the didactical activities. We also illustrate how this experience has been conducted and the results of the evaluation.

3.1. Designing a virtual lecture setting

To start a learning activity in SL, as in the real life, it is necessary to have a place where teaching a lecture. Thus, it is necessary to buy a land, which metaphorically means to pay Linden Lab organization, owning SL, for the hosting service. On such a land we recreated the building of the Department of Mathematics and Informatics (DMI) where the students were used to have lectures. This is a relevant factor, because the realism and presence sensation grows and let students to have the impression to be in their usual didactic setting. The obtained result is impressive: it seems to really move around in the DMI building. Figure 1 and 2 show, respectively, the DMI reception and one of the DMI classrooms.

Figure 1. The DMI hall

The setting has been enriched by the active behavior of some objects, obtained by using the programming language offered by SL, namely Linden Scripting Language [12]. As an example, doors are programmable objects which can be opened and closed or chairs are objects on which avatars can seat. We also developed a slide shower object, depicted in Figure 3. Slides are well readable and can be changed using a button interface. It is also possible to give only to the teacher the right to act on the buttons.

3.2. A lecture in Second Life

The lecture was given to students of the course Fundamentals of Informatics of the bachelor program in Environment Evaluation and Control School of Sciences Faculty of the University of Salerno.

26 students participated to the experimentation. A pre-experiment questionnaire has been proposed to them to assess their background in using computer and video games. The novices players received a quick training session about basic SL operations before the lecture.

Figure 2. A virtual classroom

The participants were allocated into two laboratories, one setting is shown in Figure 4 (a), while the teacher was in a separate room, as shown in Figure 4 (b). The lecture lasted one hour and concerned basic statements of the C programming language.

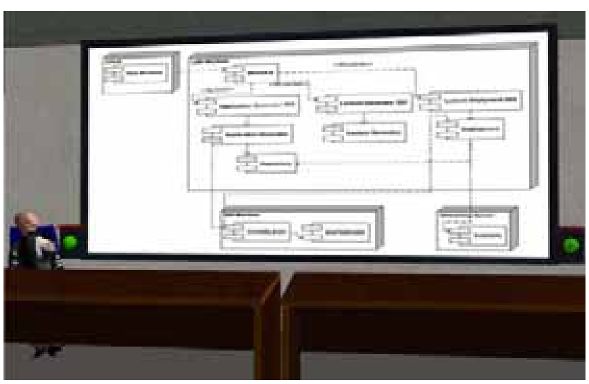


Figure 3. The slide presenter

Each student had at his/her disposal a computer and was able to interact with the others and with the teacher only using a textual chat, while the teacher uses the vocal chat. This kind of communication was chosen to better control the student interventions: they wrote a question on the chat and the teacher vocally answered. In Figure 2 the virtual classroom, the avatars of the students and of the teacher participating to the lecture are shown. Also two tutors, near the slide presenter in Figure 2, participated to the experience. Their task in the real setting was to supervise that the students communicated only with the textual chat.

4. Evaluation

The evaluation of the proposed experience starts by examining the pre-experiment questionnaire to assess subjects skills in using computer and video games. Also an estimation of the student inclination to involve himself in analogous experiences is collected.

Figure 4. The real setting: the students (a) and the teacher (b) in different labs

The data collected during the pre-experiment questionnaire on the subject background are summarized in Figure 5.

The results have been further investigated considering the genre, because we thought that males had more attitude to computers and videogames. On the contrary, results revealed that there was no significant difference.

It is well established that presence is an aspect influencing learning. In this paper the meaning of presence is used as defined in Witmer and Singer [24] as "the subjective experience of being in one place or Environment". Another factor increasing presence is immersion, "a psychological state characterized by perceiving oneself to be enveloped by, included in, and interacting with an environment". When the interaction with the environment happens in a natural way the perceived sense of immersion should increase and thus presence. Immersion is also increased when outside distractions are minimized and active participation is increased when the user feels to effectively control the events in the environment.

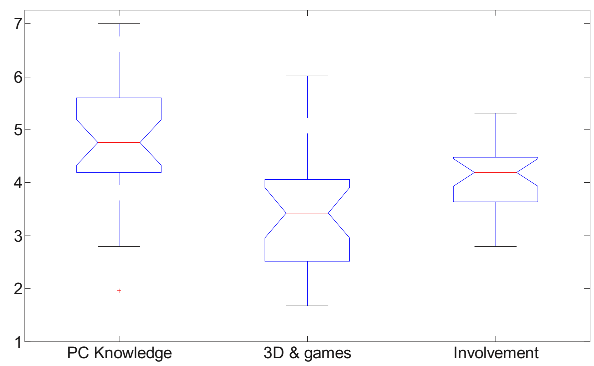


Figure 5. The pre-experiment questionnaire results

To evaluate the perception of presence we proposed to the involved students the presence questionnaire in Table 1, adopted by [24]. Witmer and Singer stated that four factors contributes to presence: Control Factor (CF), Realism Factor (RF), Sensory Factor (SF) and Distraction Factor (DF). A higher degree of presence is obtained if:

- a user has more control over his/her actions in the virtual world (CF);
- the scene realism is high (RF);
- the visual aspects of the device extent the Sensory Factor (SF);
- the user is not distracted while performing the activity - Distraction Factor (DF).

In Table 1 the presence questionnaire is depicted. It consists of 32 questions using a seven-point scale. A total presence score is calculated by taking the sum of all question outcomes.

Table 1. The Witmer and Singer Presence Questionnaire

1.	How much were you able to control events? CF
2.	How responsive was the environment to action that you initiated (or performed)? CF
3.	How natural did your interactions with the environment seem? CF
4.	How completely were all of your senses engaged? SF
5.	How much did the visual aspects of the environment involve you? SF
6.	How much did the auditory aspects of the environment involve you? SF
7.	How natural was the mechanism which controlled movement through the environment? CF
8.	How aware were you of events occurring in the real world around you? DF
9.	How aware were you of your display and control devices? DF
10.	How compelling was your sense of objects moving through space? SF
11.	How inconsistent or disconnected was the information coming from your various senses? RF
12.	How much did your experiences in the virtual environment seem consistent with your real-world experiences? RF, CF
13.	Were you able to anticipate what would happen next in response to the actions that you performed? CF
14.	How completely were you able to actively survey or search the environment using vision? RF, CF, SF
15.	How well could you identify sounds? RF, SF
16.	How well could you localize sounds? RF, SF
17.	How well could you actively survey or search the virtual environment using touch? RF, SF HAPTC **(not applicable)**
18.	How compelling was your sense of moving around inside the virtual environment? SF
19.	19. How closely were you able to examine objects? SF
20.	How well could you examine objects from multiple viewpoints? SF
21.	How well could you move or manipulate objects in the virtual environment? CF APTC **(not applicable)**
22.	To what degree did you feel confused or disoriented at the beginning of breaks or at the end of the experimental session? RF
23.	How involved were you in the virtual environment experience? INV/C
24.	How distracting was the control mechanism?
25.	How much delay did you experience between your actions and expected outcomes? CF
26.	How quickly did you adjust to the virtual environment experience? CF
27.	How proficient in moving and interacting with the virtual environment did you feel at the end of the experience? CF
28.	How much did the visual display quality interfere or distract you from performing assigned tasks or required activities? DF
29.	How much did the control devices interfere with the performance of assigned tasks or with other activities? DF, CF
30.	How well could you concentrate on the assigned tasks or required activities rather than on the mechanisms used to perform those tasks or activities? DF
31.	Did you learn new techniques that enabled you to improve your performance? CF
32.	Were you involved in the experimental task to the extent that you lost track of time? INV/C

The questions numbered 17 and 21 are left out because they concern haptic interfaces, and, as a consequence, are not applicable.

The questionnaire results are summarized in Table 2. We can observe that the presence perception of the users is very high, in spite of the lack of audio chat for the students.

Table 2. Questionnaire results

	Total	Scaled
mean score	150,0	5,0
std. Dev.	14,0	
max	186,0	6,2
min	116,0	3,8

The histogram in Figure 6 (a) shows that the 96% of the students expressed a presence sensation higher than the expected value (120).

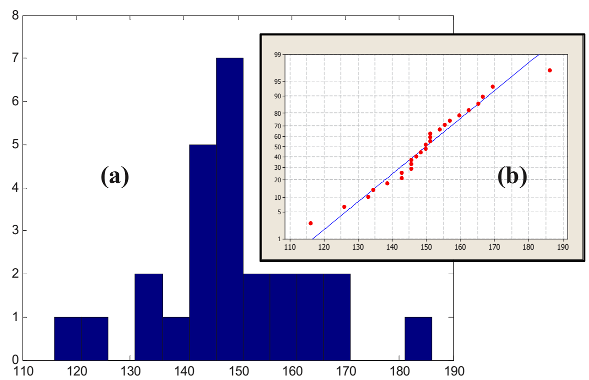


Figure 6. Presence score frequency distribution and Kolmogorov-Smirnov test

The first feature to point out about the histogram is its single-peaked and symmetric nature.

The obtained results have been quickly verified by a Kolmogorov-Smirnov [4][22] test for normality, graphically resumed in Figure 6 (b). In this figure the empirical cumulative distribution function of data and a normal distribution with mean 150,0 and standard deviation 14,0, are represented by scattered points and continuous line, respectivelly. The Kolmogorov-Smirnov test, performed with a significance level α=0.05, does not enable us to confute the normality hypothesis.

Presence questionnaire results are reorganized and further investigated in Figure 7, where data are categorized with respect to the factors contributing to presence. In particular, the CF box reveals that the user perceives a good degree of control on the environment. Similar results have been obtained for SF and DF, even if with a greater dispersion. Let us note that the DF results have been reversed. In addition, the RF factor has been affected by the audio which was reserved to the teacher (Question 15 and 16). These results have been also very appreciable from the point of view of the quality and the number of intervention of the students. In addition, a free comment section was proposed together with the questionnaire to collect free personal impressions and comments on the activities. Analyzing these comments, we had confirmation of the positive results. Indeed, most of comments have been enthusiastic. As an example, some students have written: "It has been an exceptional experience" or "why do not we have always lectures in this way?".

This has been further supported by the following comment: "the distance between student and teacher is reduced: it seems to be more natural, spontaneous and easier to communicate in Second Life".

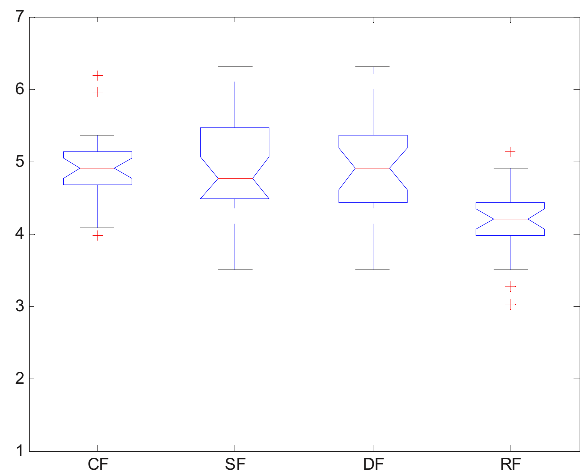


Figure 7. Results grouped by presence factors

Additional questions concerning the user satisfaction and the student opinion on their productivity have been proposed at the end of the experience. The results are summarized in Table 3. Let us note that the standard deviation is reduced and the values are concentrated around the mean, which is very high in both cases.

Table 3: The result obtained formulating two direct questions about the experience.

	Productivity	*Satisfaction*
mean	6,2	6,515
std. dev.	1,6	0,9
min	1	4,2
max	7	7

Another interesting consideration concerns the correlation between the student background and the presence scores, depicted in the scatter plot diagrams in Figure 8. It is reasonable to argue that the questionnaire results are not influenced by the sample characteristics:

user impressions are positive also when the student was not particularly expert in computer, videogames or with particular inclination to be involved.

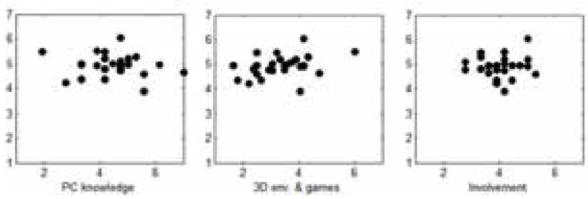


Figure 8. Correlations between presence and student background

5. Conclusion

In this paper we have presented the first experience of synchronous distance lecture in Second Life at the Department of Mathematics and Informatics of the University of Salerno. The results of a student evaluation concerning presence, satisfaction and productivity are very encouraging even for subjects with little experience in computer games and 3D environments. Let us note that students were not evaluated after the experience and the questionnaire was anonymous.

Currently, we are further investigating the students perceptions and their performances comparing face to face learning with synchronous learning in Second Life.

Acknowledgments

The authors would like to thank Dr. Luciano Ragazzi, *ANS Ricerca e Innovazione S.r.l.* research and development, for the support provided.

References

[1]. Blackboard, www.blackboard.com.
[2]. McCreanor, P., T., Developing A Web-Enhanced Course: A Case Study, *In the proc. Of 2000 IEEE 30th ASEE/IEEE Frontiers in Education Conference*.
[3]. Cochrane, K., Case Study: International Spaceflight Museum, *in the proceedings of the Second Life Community Convention*, San Francisco, USA August 20th, 2006, pp. 2-5.
[4]. R.B. D'Agostino and M.A. Stephens, Eds. (1986). Goodness-of-Fit Techniques, Marcel Dekker.
[5]. De Lucia, A., Francese, R., Passero, I., Tortora, G. (2008). Supporting Jigsaw-Based Collaborative Learning in Second Life. To appear *in the Proceedings of The 8th IEEE International Conference on Advanced Learning Technologies* ICALT 2008.
[6]. Doherty, P., Rothfarb, R., Barker, D., Building An Interactive Science Museum In Second Life, *in the proceedings of the Second Life Community Convention*, San Francisco, USA, 2006, pp. 19-24
[7]. Edwards, C., Another World. *IEEE Engineering & Technology*, December 2006.
[8]. Gerhard, M., Moore, D., Hobbs, D. (2001) Continuous Presence in CVEs: Towards the Evaluation of a Hybrid Avatar-Agent Model for User Representation, in Proceeding of the International Conference on Intelligent Virtual Agents (IVA 2001), Madrid, Spain.
[9]. Harvard University, http://slurl.com/secondlife/Berkman/105/74/35.
[10]. Kemp, J., Livingstone, D., Putting A Second Life “Metaverse” Skin On Learning Management Systems. Learning Communities, New Directions for Teaching and Learning: No. 97, San Francisco: Jossey-Bass, 2004, pp. 5-23
[11]. Light, J., and Miller, J., D. Miramar: a 3D workplace. *In the IEEE Proc. IPCC*, 2002., 271-282.
[12]. Linden Scripting Language, http://wiki.secondlife.com/wiki/LSL_Portal.
[13]. MIT,http://slurl.com/secondlife/MIT/51/35/23/?title=Massachusetts%20Institute%20of%20Technology
[14]. Moodle, www.moodle.org.
[15]. Nijholt, A., Zwiers, J., Peciva, J., The Distributed Virtual Meeting Room Exercise, In the *Proceedings of the Workshop on Multimodal multiparty meeting processing (ICMI 2005)*, Trento, Italy.
[16]. Powerusers, http://powerusers.edc.org/
[17]. Rosenman, M., Merrick, K., Maher, M., L., and Marchant D., Designworld: A Multidisciplinary Collaborative Design Environment Using Agents In *a Virtual World. Design Computing And Cognition,* Springer Netherlands (2006).
[18]. M. Rymaszewski, W. J. Au, M. Wallace, C. Winters, C. Ondrejka, B. Batstone-Cunningham, and S. L. residents from around the world. Second Life: the office guide. *Wiley Press*, 2007, pp.4.
[19]. Second Life. http://secondlife.com
[20]. Proceedings of the Second Life Education Workshop at the Second Life Community Convention San Francisco, 2006.
[21]. Proceedings of the Second Life Education Workshop (2007). Part of the Second Life Community Convention. Chicago Hilton, 24th-26th August 2007.
[22]. S.S. Shapiro and R.S. Francia (1972). "An Approximate Analysis of Variance Test for Normality," Journal of the American Statistical Association, 67, 215-216.
[23]. Usoh, M., Catena, E., Arman, S. & Slater, M. (2000). Using presence questionnaires in reality. Presence: Teleoperators and Virtual Environments, 9(5), pp. 497-503
[24]. Witmer, B. G. & Singer, M. J., (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments*, 7(3), pp. 225 – 240.
[25]. Zhu, Qing; Wang, Tao; Jia, Yufu, (2007). Second Life: A New Platform for Education. In the Proceedings of Information Technologies and Applications in Education, 2007. ISITAE '07. First IEEE International Symposium on 23-25 Nov. 2007 Page(s):201 – 20.

How to use Grid technology for creating contextualized and adaptive learning experiences: the European Learning Grid Infrastructure software architecture

Matteo Gaeta *Member, IEEE*, Francesco Orciuoli, Pierluigi Ritrovato

***Abstract*— One of the widely recognised reason of the continuous fail of e-learning adoption has been the hydraulic view (content centred) of the learning process, supported by monolithic software systems. Supporting content delivery delegate the learner to a passive role (sink of the information transfer) with very limited personalisation capabilities, lacking in motivate them. Service Oriented Architectures are going to be widely adopted in the business world. The new forms of applications creation and delivery they provide enable a user centred and process oriented approach to e-learning system development. In this paper we present a concrete example of a SOA for Learning developed and experimented in the frame of The European Learning Grid Infrastructure (ELeGI) research project. In particular, we show how SOA implemented using Grid technologies allows creating adaptive learning systems that implement complex learning processes considering pedagogical aspects that are implemented as workflows of activities supported by services.**

I. Introduction

The dominant learning approach in Technology Enhanced Learning has been for many years the Information Transfer, which is based upon the central figure of the teacher whose primary activity is the provisioning of educational contents to be transferred to learners that consume these contents in a passive way. This hydraulic view of learning have found its perfect technical mirror in the "page oriented approach to the Web" where the goal is to produce more and "better" static pages for the consumption of interested students. E-Learning is thus considered to be an activity which helps teachers to produce and students to consume multimedia books over the Web. This paradigm has been popular in many e-Learning projects, not because it is effective, but owing to the fact that it is easy to implement with basic Internet facilities and does not require any change in the traditional roles of the actors. As a consequence distance learning platforms support, mainly educational resources management that is only an input of the whole learning process and on their presentation (delivery). This aspect, together with other relevant lacks, as the support for pedagogies, the contextualization of the learning experience, the learner's centrality and activity, lowers the level of effectiveness of the above learning solutions. Moreover, The aforementioned solutions are focused on a specific educational model (often built-in in the content) and this means: i) to superimpose to learners how they have to learn without taking into account learner's dispositions or preferences, and ii) to develop pedagogically closed e-Learning solutions constraining learners to learn and teachers to teach following a predefined approach. Entering the so-called knowledge society, e-learning solutions must be subject to a real evolution where the main efforts must be also addressed to support the whole learning process, not only on specific part of the process.

In order to advance effective learning, current research trends promote a shift from contents to processes focused on the learner, on socio-constructivist forms of learning and on the definition and support of different pedagogies. In this approach the learner has an active and central role in the learning processes aimed at facilitating the construction of learner's knowledge and skills, instead of the simple memorisation of information.

The new approach to learning finds life blond in the Web 2.0 wave [1]. Within this new form of learning, technologies have to support the creation of high dynamic personalised distributed virtual environments where different actors provide resources and services, and well defined learning processes are consumed in a personalised, collaborative and ubiquitous way [2].

This work aims at reporting the experiences and results achieved in the context of the European Learning Grid Infrastructure (ELeGI) [3] research project described in [4]. ELeGI is an Integrated Project funded by the European Commission (February 2004 – June 2007) that has been conceived in order to provide a solution to this new approach through Learning models and related processes for formal/informal learning supported by a software architecture based on Grid technology. The ELeGI software architecture allow the creation of dynamic and contextualised environments for each learner (taking into account his/her aptitudes and social context, providing tutoring and enhanced presence), supporting him/her during the whole learning process.

Matteo Gaeta, Pierluigi Ritrovato and Francesco Orciuoli are with the Information Engineering and Applied Mathematics Department, University of Salerno and with the CRMPA – Research Centre in Pure and Applied Mathematics Via Ponte Don Melillo 84084, Fisciano (SA) - Italy (e-mail: {gaeta, orciuoli, ritrovato}@)crmpa.unisa.it).

This work has been partially supported by the European Commission under the Information Society Technologies (IST) programme of the 6th Framework Programme for RTD - project ELeGI, contract IST-002205. This document does not represent the opinion of the European Commission, and the European Commission is not responsible for any use that might be made of data appearing therein.

The rest of the paper is organised in four sections: Section 2 provides an overview of the ELeGI theoretical models and related processes that define how the contextualised, active, and personalised learning experiences have to be created following real pedagogical principles. Section 3 describes the service oriented, Grid based, software architecture supporting the theoretical models and providing advanced capabilities in terms of dynamic adaptation (to context, learner preferences, didactic methods). Section 4 is devoted to describe how the ELeGI software architecture works. Finally Section 5 presents the related works.

II. The ELeGI models and related processes

A. The Learning Model

The ELeGI approach to formal learning consists in the definition of a general Learning Model (depicted in Figure 1), able to support different pedagogical models, whose purpose is to enable the automatic generation of Unit of Learning (UoL) and to dynamically adapt it during the learning process according to the learner's behaviour and performance. For the Learning Model presented in [6], a UoL is something delimited as educational object, such as a course, a module or a lesson structured as a sequence of Learning Activities that include Learning Objects and/or Learning Services.

In order to achieve the expected adaptation capability, the learning model uses other three specific models: the Knowledge Model, the Learner Model and the Didactic Model, which are managed by specific processes that interact to define the personalised learning experience. The three models allow to take into account: the knowledge that formalise the cognitive domains through the definition of concepts and relations between the concepts; the context, where the educative process occurs; the learning method and style; the learner's preferences and demands.

The Knowledge Model (KM) describes, in a way that the machine can process, the subject-matter to be attained using Ontology. Three different kind of ontology con be created: Macro Ontology (MO), Generic Contextualised Ontology (GCO), and Specific Personalised Ontology (SPO); In the automatic construction and personalisation of the student's learning path the KM play an important role since make possible to relate the Learning Objects to the concepts of the knowledge domain using specific metadata.

The Learner Model (LM) allows representing in a way that the machine can process the learners' characteristics (profile), preferences, knowledge and styles. In particular, the LM includes a cognitive state and a preferences state. Cognitive state represents the knowledge reached by a learner at a given time and is composed by a list of concepts each with an associated score between 0 and 1. Preferences state is a set of couples (property; value). Specific studies have been conducted in order to identify the learner's styles. We have investigated the models defining learning styles (e.g. VARK, Kolb, Kersey, etc.) and the different methodologies for the Learning Styles identification in order to adapt the learning path and define personalised remedial work accordingly. The different learning styles models are in general representative for one of the following four macro dimensions:

- instructional preferences
- social interactions
- information processing
- cognitive personality

The analysis of the different learning styles led us to identify the following four learning styles families: Pragmatic, Active, Holistic, Analytical.

This analysis helps us to identify the kind of activities/services needed and also the nature of contents that are more suitable for the different families. Using a questionnaire (specifically developed) it is possible to identify what is the level (expressed in percentage) of liking for each learning style family by the learner.

The Didactic Model (DM) allows to formalise different pedagogical approaches to be used in a learning experience, both at macro-level (general structure of the UoL) and at micro-level (i.e. didactic guidelines, according to pedagogical theories, to enable the student creating knowledge). In, particular, the micro-level allows to take into account both the didactic approach (e.g. problem based learning, Case-based approach, Jigsaw model, and so) and the concrete flow of learning activities of an UoL. The flow of activities is prepared in accordance to the selected didactic approach and allows implementing it. This flow of activities is called ***Didactic Method***.

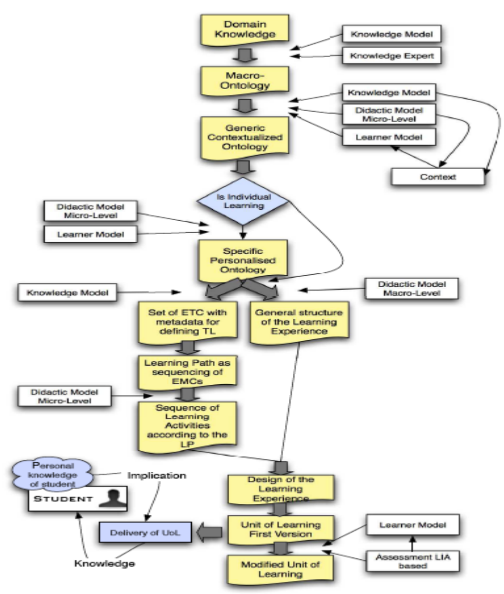


Figure 1: ELeGI Learning Model

III. THE ELEGI SOFTWARE ARCHITECTURE

A. *The Learning Model's processes*

In order to adopt the ELeGI Learning Model and generate personalised, contextualised and adaptable (to learner preferences, knowledge and didactic approaches) UoLs several processes (Figure 2) have to be specified and implemented. In the following we describe the more relevant ones.

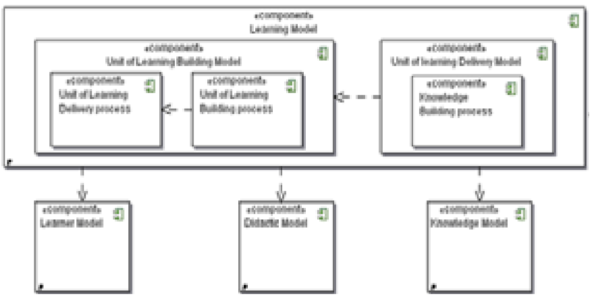


Figure 2: The process view of the Learning Model

i) *knowledge building process*, whose purpose is the formalization of knowledge related to a domain using the three types of ontology The MO formalizes the knowledge about a domain. The GCO takes into account a particular context and/or target group. GCO derives from the MO by adding metadata to concepts. In the same way, the SPO takes into account the features of a single learner. SPO can be obtained from a MO or from a GCO while the information needed for the metadata annotation are gathered from the Learner's Model.

ii) *UoL building process*, whose purpose is to assemble a UoL by using the ontologies produced in the knowledge building process. Firstly, the learning objectives to be achieved, identified by some target concepts within a GCO or SPO, and the skeleton of the structure to be used for the whole learning experience (i.e. a lesson plan), called Learning Experience Model (LEM), must be specified. After that, the automatic construction of the UoL can be started. Through the target concepts and the cognitive state of the Learner's Profile, it is possible to generate from the ontology a contextualised and/or personalized ordered sequence of concepts (Learning Path) needed to explain the target concepts. Then, a workflow of learning activities is produced by sequencing the learning activities corresponding to the Didactic Method (DM) associated (specified through metadata fields) to each concept of the Learning Path (in case an OGC is available) or according to the preferred learning approach expressed by the learner (if an SPO is defined).

iii) *UoL delivery process* entails the run time execution of the UoL. It performs the operations needed to discover the resources satisfying the metadata specifications contained in the learning resources of the UoL and to bind them on-the-fly to the learning activities of the UoL. Specific assessment activities are introduced in the UoL in order to check the progress, to update the learners' profile, and to provide remedial work if needed.

B. *The ELeGI Software Architecture*

In order to instantiate the Learning Model it is necessary to rely upon a set of functionalities allowing the actors of the learning process to create, manage, manipulate, and annotate the structures used in the processes, to collaborate among them, to join a learning experience, etc.

The main purpose of the ELeGI software architecture [7], depicted in Figure 3, is to allow, from one side the creation and the operational management of a Virtual Organisation (VO) based environment, where it is possible to execute the learning processes, and from the other side to allow the production of the appropriate operational process related to the learning model and its conceptualisation. From the software engineering point of view The ELeGI software architecture can be considered as a framework for the development of e-learning applications. In fact, it provides both a rich set of available services and also skeletons, practices, procedures supporting the construction of solutions for definition, creation and delivery of learning experiences. Detailed information about the approach followed for the software architecture design and implementation are described in [8].

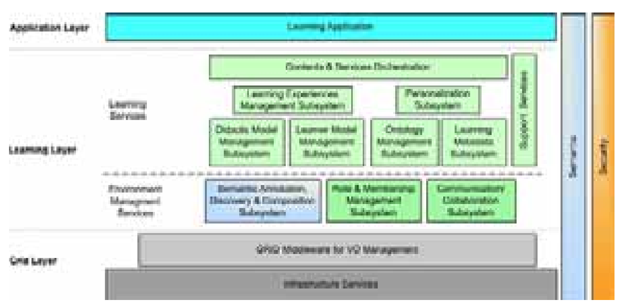


Figure 3– The ELeGI Software Architecutre

In order to speed-up the implementation, we have made recourse to the GRASP middleware [19] and to the Grid aware version of an innovative learning platform named Intelligent Web Teacher [15] for the implementation of the Grid layer services and the Learning layer services respectively.

The Grid layer provides a set of Infrastructure Services and the services to create and manage resources federated in Virtual Organisation (VO) according to [9]. This layer is devoted to the VO operational management and it provides an implementation of the specification of the Open Grid Service Architecture (OGSA) Web Service Resource Framework (WSRF) Basic Profile [10] and the services defined in the OGSA V1 [11]. In particular OGSA provides a stateful and dynamic service model based on a set of WS-Specification able to manage the state associated to a service/resource particularly relevant in learning contexts.

The most relevant ELeGI services belong to the Learning layer, that is mainly devoted to the execution of the processes related to the Learning Model. This layer can be logically divided in two sub-layers.

The first one, the Environment Management Services sub-layer, provides services and tools to support the creation, operation, evolution, and maintenance of a learning community. Functionalities for semantic annotation, discovery and composition of educational contents and services are provided in the Semantic Annotation, Discovery & Composition subsystem, while functionalities to allow intra and inter community asynchronous and synchronous communications are grouped in the Communication/Collaboration sub-system.

The second one, the Learning Services sub-layer, provides services and tools to support the execution of the three processes of the Learning Model. Of course, there are services and tools to create and mange the Ontologies (Ontology Management sub-system), the Learner's Profile (Learner Model Management sub-system) and the Didactic Model (Didactic Model Management sub-system), that represent the three basic structures of the Learning Model. The Personalization sub-system aims at dynamically adapting and delivering educational contents and services, matching the learner's needs and preferences according to his/her profile.

The Learning Experience Management sub-system allows applications or other services to access and manage courses, modules, and other learning experience (e.g. allocating student, staff, etc.), while Contents & Services Orchestration sub-system deals with the issues of execution of Units of Learning that are described using the IMS Learning Design (IMS-LD) constructs [12]. IMS-LD is a cornerstone of our vision. It is a specification used to formalize learning scenarios that can describe a wide variety of pedagogical models, including group work and collaborative learning. It does not define individual pedagogical models; instead, it provides a high-level language, or meta-model, that can describe many different models. Without entering in too much details about the language, we want to emphasize that our approach to investigate the integration of the IMS-LD specification in Grid systems in order to achieve our goal, it is motivated by the convictions that: i) dynamicity and adaptiveness of Grid technologies can bridges some gaps of the current frameworks based on IMS-LD, providing effective benefits from the viewpoint of reuse and repurposing of learning activities, and ii) semantic and knowledge technologies can provide advantages in the binding phase during the deployment of a learning scenario enabling on the fly tailoring with respect to the learner's preferences and expectation making the learning experience much more attractive from the user's perspective.

Finally, the Application Layer uses the services provided by the underlying layers or their composition to implement application in the e-learning domain. To this layer also belongs the portal that, according to the researches on Grid portals, is designed exploiting the Web Services for Remote Portlet (WSRP) standard [13].

C. The Grid added value

Service Oriented Architectures rely upon the service's concept and provide a suitable approach to implement a composition paradigm or building block approach, where high level services/applications are built exploiting already existing services and/or functionalities. This is clearly useful in our learning approach that strongly relies upon a composition paradigm for the creation of (personalized) learning experiences (re-)using data, UoLs, knowledge, tools virtualized as services and distributed across different organizations. The ELeGI software architecture has been conceived as a verticalisation for the e-Learning domain of the Semantic Grid [14].

The key ELeGI advancements can be classified as follows: *i) Personalisation and Contextualisation*: learning experiences can be personalised and contextualised at the level of the knowledge to be created/transferred by/to a learner (e.g. the creation of a GCO and SPO), at the pedagogy level (e.g. the selection of a suitable Didactic Method), at the content level (e.g. the selection of suitable contents on the basis also of the Learner Profile); *ii) Dynamicity and effective re-use of resources: it is* possible to re-use independently all "building blocks" of a UoL (ontologies, LEMs, DMs, etc.), that are resources semantically annotated and discoverable over the Grid infrastructure; *iii) Separation of roles/responsibilities*: there is a real separation of roles/responsibility among didactic domain experts (ontologies), educational designers (DMs and LEMs), content developers, service developers, tutors (on-line support), teachers (target of learning, context information). Each of them is provided with specific tools/services; *iv) Extensibility wrt services integration*: It is possible to use, in the UoL, any kind of service or educational resource virtualised as a service and rendered by WSRP Portlets.

IV. ELeGI AT WORK

In the following paragraphs we shortly describe how the services of the learning layer are used for generating personalised and contextualised learning experiences (UoLs) in a dynamic and adaptable way.

A. Personalization of the learning experience

The functionalities provided by Personalisation Management Subsystem are to create personalised learning paths, to generate the sequencing of learning activities according to the learning path and the learner's profile, and to calculate remedial work.

The generation of the Learning Path requires an ontology (GCO or MO) and the selection of one or more target concepts (on that ontology). The GCO or MO can be created and uploaded to the Metadata Services using the KRT (Knowledge Representation Tool provided by IWT). Each concept of the Learning Path includes a metadata description, according to the Concept Metadata Schema.

The generation of the sequencing of learning activities requires to sequentially scan the whole Learning Path and to take into account the context information associated to each concept. Then the learning activities of the Didactic Method of each concept are contextualized merging the information included in the resource files (that contain the parameter values allowing the search of the right resources and services), with the other information associated to the concept and ordered in a sequence according to the order of the Learning Path.

The remedial work can be considered as a branch in the main execution flow of a UoL and is invoked after an assessment phase (when the system realise that some concepts of the learning path have not been understood properly – bad score). The Remedial Work Service calculates a new sequence of learning activities, in order to better explain the concepts that have not been understood in the main UoL. The service takes in input the learner cognitive state, updated on the basis of scores obtained after the self-assessment phase; GCO and target concepts of main UoL and a fixed LEM. The resulting sequencing of LO, at run-time, will be presented inside the delivery window of the main UoL as a Remedial Course sub-UoL (as portlet created and assembled at run-time). The learner will perform and navigate among the remedial work activities before re-starting with the main flow of the UoL execution, that was frozen after the self-assessment point, waiting for the completion of the remedial work.

In the UoL building phase for remedial work, the Learning Path of main UoL is re-processed using the updated learner cognitive state, in order to purge the concepts acquired and already known.

For remedial work the following didactic strategy has been applied:

1. In the first cycle of remedial work, Didactic Methods and concepts metadata don't change in order to re-propose to the learner the same LOs (content) and activities.
2. After the first cycle the learner is presented with another assessment phase. In case of failure, the learner will be redirected to the second remedial work cycle, otherwise he/she will return to the main flow.
3. In the second remedial work cycle, different metadata values will be applied to the concepts of the Learning Path, in order to find different resources in delivery phase, and to explain the same concept in a simpler way. To change the difficulty level, three fields of metadata are used and modified: Difficulty, Semantic Density and Educational Context.
4. In the third and following cycles, the Didactic Methods are changed in order to try a different didactic approach. Different Didactic Methods taken from the learner's preferred Learner Style will be tried.

B. *Content & Service composition and orchestration*

In order to execute the personalised and contextualised UoL, to control the learner performances and to adapt the learning process accordingly are used three services:

- Assemble Learning Experiences Service
- Learning Experience Orchestration Service
- Learning Experience Assessment Service

The Assemble Learning Experiences Service is in charge of generating the UoL, by merging the sequence of learning activities and the LEM (Learning Experience Model) structure. It provides an IMS-LD package that maintains the structure of the LEM and includes for each concept of the learning path the structure defined in the specified (using the metadata present in the CGO) Didactic Method. More in detail, the sequence of concepts of the Learning Path has to be mapped on the LEM, namely the concepts are distributed in a suitable way on the different activities/phases of the LEM. To generate the sequencing of learning activities of the concepts it is invoked the functionality for generating the learning activities sequencing of the personalization subsystem (above described).

The Learning Experience Orchestration Service coordinates all phases of the UoL execution. This service manages and orchestrates the assignment of users to the roles of the UoL, the fragmentation of the Learning Experience into Learning Phases, the dynamic binding of the resources (LOs and services –discovery them on the Grid infrastructure), the management and the execution of the Learning Phases, including the related assessment, evaluation, update of the profiles, and remedial work.

The Learning Experience Assessment Service allows looking for a minimal set of resources (namely AssessmentLO) needed for the evaluation of learner understanding of the concepts exposed. An AssessmentLO can cover different concepts, and one concept can be covered by several AssessmentLO. In order to solve this problem and have a minimal set of LO covering all the concepts, the service uses a Plant Location Solver algorithm. Each AssessmentLO execution leads to a learner cognitive state update according to the results obtained. The Learning Experience Assessment Service stores in its state that list of AssessmentLO and its execution results. This information will be very useful for the Remedial Work Service.

Concerning the execution of the Learning Experiences, namely the running of the UoL, it has to be considered that the UoL package file produced by the Assemble Learning Experiences Service can be processed and executed using a common IMS-LD engine (CopperCore [16], in our case). However, due to the fact that the resource files of the package contain only information (semantic descriptions of services and didactic resources) – useful to perform the resources and services searching and binding over the Grid infrastructure - the UoL cannot be delivered using a common IMS-LD player. Therefore, a suitable player has been developed updating the player released with the

CopperCore engine that is able to "query" the Grid infrastructure, perform the dynamic binding of resources and services, and deliver the UoL as described in [17]. A detailed learning experience scenario's execution using the EleGI grid based infrastructure is provided in [18].

V. RELATED WORKS

Few solutions in the e-learning area can be compared with ELeGI, as it encompasses and merge many technological and pedagogical aspects. In the following we briefly recall the most noticeable, pointing out the distinctive features of ELeGI. The Sakai project [21] aims at defining and developing a Collaboration and Learning Environment by exploiting the OSIDs, defined in the frame of Open Knowledge Initiative (OKI), and service-oriented portals. ELeGI is less content oriented with respect to OKI/SAKAI, provides support for pedagogies and is more focused on knowledge management and collaboration through social interactions (not only collaborative tool).

The JISC ELF [22] is a framework for e-learning applications. It provides a service-oriented factoring of the core services required to support e-learning applications, portals, and other user agents. Two key aspects distinguish ELeGI from ELF: i) the adoption of the OGSA model and, hence, the exploitation of the VO paradigm to manage the operational issues of a distributed learning environment (e.g. heterogeneous resources localisation, instantiation, etc.) and to address specific issues like selection of the most suitable computational resource, QoS management, community creation support, etc., and ii) the adoption of a methodology, as results of scientific studies on the theoretical models, to address personalisation, contextualisation and dynamic adaption in a pedagogically driven approach.

SLED (from JISC e-learning programme) is an implementation of a learning solution based on a service-oriented architecture integrating a build-in set of services. The ELeGI software architecture, which is based on OGSA, has the same advantages of SLED in terms of extensibility but it moreover provides a set of general-purpose services that can be composed for the construction of more specific services. Furthermore, SLED has the great disadvantage, from the architectural viewpoint, that when one integrates a new service it is required to update the UoL player in order to make available the new service to the users. ELeGI, instead, exploits the mechanism to dynamically localise services/resources in a VO and the mechanism of WSRP enabling the integration of new services without updating the UoL player. This is exactly the advantage provided by the IoC (inversion of Control) pattern adoption.

REFERENCES

[1] Stephen Downes (October 2006), "E-Learning 2.0" [online], http://www.elearnmag.org/subpage.cfm?section=articles&article=29-1

[2] D. Dagger, Al. O'Connor, E. Walsh, S. Lawless and V. P. Wade "Service-Oriented E-Learning Platforms: From Monolithic Systems to Flexible Services" *in the special issue on Distance Learning IEEE Internet Computing* May-June 2007 Issue pp. 28-35.

[3] The EleGI project web site [online] available at http://www.elegi.org

[4] M. Gaeta, P. Ritrovato and S Salerno "Making e-Learning a Service Oriented Utility: The European Learning Grid Infrastructure Project" In the book *Towards the Learning GRID: advances in Human Learning Services.* Publisher: IOS Press in the series: "Frontiers in Artificial Intelligence and Applications", August 2005 ISBN: 1-58603-534-7, pp. 63-75

[5] I. Foster and C. Kesselman, "*The Grid: Blueprint for a New Computing Infrastructure*", Morgan Kaufmann eds, 1999

[6] G. Albano, M. Gaeta, S. Salerno: E-learning: a model and process proposal in the *International Journal of Knowledge and Learning* – Volume 2 issue 1/2 2006 Inderscience Publisher, ISSN (Online): 1741-1017 - ISSN (Print): 1741-1009, pp73-88.

[7] A. Gaeta, M. Gaeta, F. Orciuoli, and P. Ritrovato: "Enabling Technologies for future learning scenarios: The Semantic Grid for Human Learning". Proc. of the 2nd *International Workshop on Collaborative and Learning Applications of Grid Technology and Grid Education CLAG + Grid.edu*, in conjunction with the *IEEE Intern. Symposium on Cluster Computing and the Grid* May 9 - 12, 2005, Cardiff, UK

[8] A. Gaeta, M. Gaeta, F. Orciuoli, P. Ritrovato "Defining a Service Oriented Architecture for e-Learning: The European Learning Grid Infrastructure Project Experience", to appear in the *International Journal on Technology Enhanced Learning* – Inderscience publisher ISSN (Online): 1753-5263 - ISSN (Print): 1753-5255

[9] I. Foster, C. Kesselman and S. Tuecke, "The Anatomy of the Grid: Enabling Scalable Virtual Organizations." *International Journal of Supercomputer Applications*, 2001, Vol. 15, No. 3, 200-222

[10] I. Foster et al.: OGSA™ WSRF Basic Profile 1.0. Recommendation GWD-R GGF, OGSA WG, September, 7, 2005 Available at https://forge.gridforum.org/projects/ogsa-wg/document/draft-ggf-ogsa-wsrf-basic-profile/en/38

[11] I. Foster et al.: The Open Grid Services Architecture, Version 1.0. Report GFD-I.030. GGF, OGSA WG. (January, 25, 2005). [online] http://www.ggf.org/documents/GFD.30.pdf

[12] IMS Global Learning Consortium, IMS Learning Design v1.0 Final Specification, http://www.imsglobal.org/learningdesign/index.cfm, 2003

[13] T. Schaeck and R. Thompson: Web Services for Remote Portlets (WSRP) Whitepaper, 28 may 2003. Available at http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=wsrp

[14] D. De Roure, N. R. Jennings, and N. R. Shadbolt: The Semantic Grid: A future e-Science infrastructure. In *Grid Computing - Making the Global Infrastructure a Reality*, F. Berman, G. Fox, and A. J. G. Hey, Eds.: John Wiley and Sons Ltd., 2003, pp. 437-470

[15] G. Albano, M. Gaeta, P. Ritrovato, "IWT: an innovative solution for AGS e-Learning model" *International Journal of Knowledge and Learning* Vol. 3, No.2/3 2007 pp. 209 – 224; Inderscience Publisher.

[16] Coppercore project, [online] available at http://coppercore.org/.

[17] N. Capuano R Iannone, S. Miranda and M Rosciano "A Grid Based IMS Learning Design Player: the ELeGI Case Study" *Proceedings of SWEL'07: Ontologies and Semantic Web Services for Intelligent Distributed Educational Systems*, p. 74-79, held in conjunction with AIED 2007, July 9 2007, Marina del Rey, USA.

[18] P. Ritrovato, M. Gaeta, A. Gaeta "A grid based software architecture for delivery of adaptive and personalised learning experiences" to be published in *ACM Personal and Ubiquitous Computing Journal* ISSN: 1617-4909.

[19] T. Dimitrakos, M. Gaeta, G. Laria, D. Mac Randal, P. Ritrovato, B. Serhan, F. Yuan, S. Wesner and K. Wulf: An Emerging Architecture Enabling Grid Based Application Service Provision., Proc. of 7th IEEE inter. conference EDOC - September 16-19 2003, Brisbane, Australia pp. 240-250

[20] M. Humphrey and G. Wasson, "Architectural Foundations of WSRF.NET", Intern. Journal of Web Services Research. 2(2), pp. 83-97, April-June 2005

[21] Sakai project, [online] available at http://www.sakaiproject.org/

[22] JISC ELF, [online] available at http://www.elframework.org/

Valle del Sarno experience: a didactical hypermedial walk from present to past

Ignazio Passero°, Luciano Ragazzi*

°Dipartimento di Matematica e Informatica, Università degli Studi di Salerno,
via Ponte don Melillo 1, Fisciano (SA), Italy
*Agenzia Sviluppo Valle del Sarno S.p.A., Via Libroia, 52, Nocera Inferiore (SA), Italy
ipassero@unisa.it, lragazzi@pattoagro.it

Abstract

In this work we present a multi-sensorial didactical walk aiming at involving students and people in general into the experimentation of a mix of sensorial solicitations and technological approaches concerning basic archeology.

The walk has been projected in such a way to interest a broad age range by simply adding or removing steps or details and by continuously alternating presence solicitations to virtual environment immersions.

The experience has been articulated in different steps a simulated on field archeological excavation, 3D interactive demonstrations and video projections, and it is concluded by a virtual tour of the Second Life reconstruction of the Nuceria theatre (Nocera Inferiore, Salerno Italy).

A first look on a survey, collected during an international archeological exhibition, confirms very good user impressions and interest about the experience.

1. Introduction

Second Life (SL) [12] is an interactive virtual world in which players create their own avatars that represent themselves in the interaction with the simulated world and with the alter egos of other people. Exploring SL drives users in a mixture of worlds ranging from extreme realistic to very fantastic, sometimes bizarre, places.

Many academic institutions [1],[3],[5],[10] and commercial enterprises [6],[8] are involving themselves in SL activities with a broad band of interests ranging from the simple needing to occupy new spaces and experiment new paradigms, to the interest toward a very big user community. Distance learning [3],[5],[10],[14],[15], scientific [2],[4], historical [13] or commercial simulations [6], are some example of the power of SL technology in satisfying almost every user interest.

It seems natural to adopt SL metaphors to reconstruct historical situations where avatars may explore excavations, visit historical reconstructions or take a first look on founds.

Valle del Sarno is a project (and a SL island [17]) focused to explore and exploit the potential of SL for developing and providing informal learning activities in enriching a learning path that propose a mix of experiences able to involve with archaeology a broad scholar community or exhibition visiting people.

Nuceria Alfaterna, was founded around the middle of the VI century. B.C.; it was at the head of the Lega Nocerina (a sannitic confederation including also Pompeii, Herculaneum, Stabia and Sorrento).

In Nuceria Alfaterna the Hellenistic-Roman theatre of Pareti was one of the greatest play buildings of ancient Campania [7]. It was built in the second century b.C., and was lately rebuilt in the Augustan Age. It was among the major monuments documented in Hellenistic Campania.

In locality Pareti (Nocera Inferiore, Salerno Italy), the archeological excavations places are not open to students or interested people and constitute an important historical witness that is crucial to reveal to the world.

To overcome the current unavailability of a museum collecting founds and presenting the theatre, the *Agenzia Sviluppo Valle del Sarno S.p.A.*, organized for the *10th* edition of *Borsa Mediterranea del Turismo Archeologico*, an informal learning path explicitly designed to involve the meeting visitors in an archaeological experience focused on the Hellenistic theatre, fine tuning, respect to visiting groups age and instruction, kind and technical level of experiences proposed. The learning path was composed of several 3D experiences ranging from 3D stereoscopic projection to a guided tour of the SL reconstruction of Nuceria theatre.

The experience have also been assessed by a survey, which results are very encouraging.

The paper is organized as follows: Section 2 describes the proposed experience while Section 3 better details the presence/virtual points of coupling of the didactical walk.

The evaluation of the experience is discussed in Section 4. Finally, Section 5 concludes.

2. The Valle del Sarno experience

The key point of the proposed didactical walk in Valle del Sarno is to involve visiting student and people with a mixture of solicitation often presenting the same objects or characters in both their real than virtual essences. By scheduling explicit breaks of the fixed narration and returning to the present, the interest of the visitors is always kept focused on the experience.

Figure 1. Nuceria Theatre archaeological excavation in SL

The Valle del Sarno archaeological journey started with a first look on present situation of the excavations. A 360 degree picture of the actual view of Nuceria theatre is used as internal texture of an empty cylinder so that SL avatars, standing in the center, as shown in Figure 1, may have a full first person look simply left/right turning.

Figure 2. Simulating an archaeological excavation

As all other SL experiences, also in the preliminary first look, the SL client view was shown to people on a cinematographic screen, letting visitors interact with the tutor controlling the avatar.

The experience continued with a practical on field activity, explicitly reserved to younger students. To simulate an archaeological excavation, broken earthenware were scattered in a sand area, inviting younger students to look for and reconstruct the original piece. Figure 2 depicts the sand archaeological area and a moment of the excavations.

Figure 3. Netes avatar 3D design demonstration

The Valle del Sarno journey, as learning experience also included moments of technical explanation end deepening. The students were involved in a 3D technological interactive demonstration. Figure 3 shows this interactive moment. In addition to some details on stereoscopic vision and rendering, this phase presented to younger students the 3D modeling of Netes avatar, the Valle del Sarno mascot.

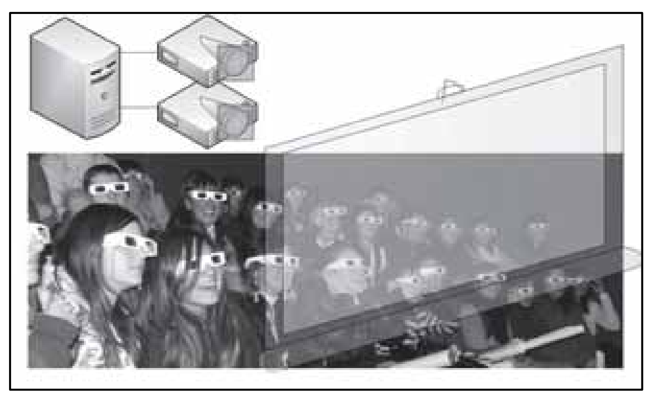


Figure 4. 3D projection room

Some details of animation techniques were also provided through samples that aided students to better understand the following phases.Once conscious of the process and the underlying technology, students continued their journey in the 3D projection room. This room was built adopting cheap technologies, but the obtained 3D immersion effect was surprising.

The 3D effect has been obtained by using a simple desktop PC, just equipped a video-card capable of stereoscopic projection, two standard projectors each mounting a polarizing filter, and distributing to visitors low cost polarizing glasses.

In Figure 4 an image taken during a projection, is augmented with a sketched representation of the adopted technology. During this projection students, equipped with polarizing glasses, had a 3D tour in an historical reconstruction. They had a look on the stereoscopic story told by Netes, a little crayfish living in Sarno river, involved in a 3D reconstruction of the first building of the theatre.

After the 3D video reconstruction of Nuceria theatre the students were ready to come and visit the SL Valle del Sarno island [17].

The experience took place by involving student into the SL experience of an expert user, let them freely interact with questions and drive the navigation by looking at the room screen. The expert adapted the details exposed and the exploration style of SL island to the characteristics of each visiting group, such as the student age, their needs and questions.During this demonstration the students were involved in a virtual tour enhanced with the "physically-accurate atmospheric rendering & lighting" implemented by the SL viewer WindLight [18].

Fully exploiting features offered by SL it was possible, not only to examine the theatre by different point of view (as for example spectator, actor, etc.), but also to change the time of day of the simulation or the weather, the water properties and other details enabling to look at different kind of ambient lighting and to imagine the setting of some antique show.

The theatre is very fine reconstructed with the adoption of all facilities provided by the SL environment. As an example, it was very interesting for visiting people when the expert drive them quickly from eye bird view to fine details examination. Using the advanced camera control features in addition to the availability to fly in the reconstructed world, deeply enriched the experience conferring it a cinematographic dimension, very precious if we consider it was a highly interactive activity and that the journey is always individually reproducible by a visit to *Valle del Sarno* island in SL [17].

3. The Netes Avatar and Nuceria Theatre in SL Valle del Sarno

The interests of visiting students has been always kept focused on the experience by using the multi-essence character Netes. Netes is a little crayfish living in Sarno river and is the mascot of the Valle del Sarno. Netes, present in all the virtual experiences and in the real life as rubber mask (see left part of Figure 5), is the common element of all the proposed learning path.

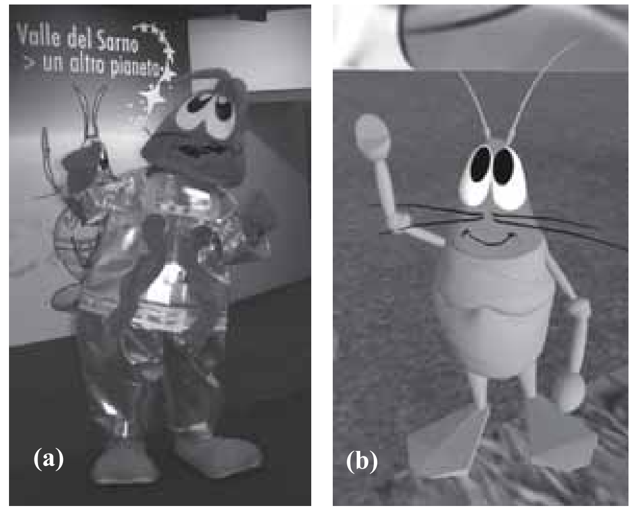


Figure 5. The Netes mascot and avatar.

Figure 5(b) depicts the Netes avatar in SL. It has been built as a SL avatar outfit [11]. An SL outfit is a complete list of appearance configurations of the avatar and wearable objects.

Figure 6. The landscape of *Valle del Sarno* in SLwith the reconstruction of the Nuceria theatre

During the SL theatre virtual tour, the SL expert, could easy capture the attention of visiting students by wearing the Netes outfit because the SL explorations have been done exploiting both shooting modalities available: third person and first person view. Visitors were coming into the stand accompanied by the Netes rubber mask (Figure 5(a)) and it was natural for them to continue the journey immersing in the SL world always following Netes, but in its avatar form.

After the *Borsa Mediterranea del Turismo Archeologico*, the SL Netes outfit has been put at disposal of avatars visiting the Valle del Sarno island [17]. The outfits are distributed by an ad hoc LSL [9] programmed dispenser which records affluence data by sending avatars generalities to an external web server. A landscape of Valle del Sarno is depicted in Figure 6. At the moment more than 350 outfits have been dispensed (let us point out that not all avatars visiting the SL theatre requested the outfit or found the dispenser) and are populating the SL simulated environments with Netes avatars.

The added value obtained adopting the SL technology concerns the persistence and availability of the Nuceria theatre 3D reconstruction.

The motivations behind the choice to reconstruct Nuceria theatre may be resumed in the following points:

- The nature of Archaeology: this science is a visual subject and it is natural to adopt SL as a teaching medium.
- The archaeological maturity of excavations and studies on the monument itself, let us to visually formulate a believable reconstruction of the theatre.
- The advantages of recreating the monument in SL: it allowed us to quickly obtain a 3D setting we exploited organizing virtual tours under the guidance of an expert. These kind of tours, by means of interactions between student and expert, added a kinaesthetic dimension to the exploration because students were able to virtually move, fly and *'walk'* around the monument and take deep look on its details.

Figure 7. Two sculpted objects

The SL theatre reconstruction is very detailed: a quick metric enabling to estimate the building effort may be inferred by simply counting the number of elementary blocks used. The Nuceria theatre has been composed by using more than 5000 elementary blocks (namely prim, which is the abbreviation for primitive) [11] that constitute about the third part of the prim number a complete island may host.

Finest details, like column capitals, are built up by using an alternate technology, the sculpted prim [16] that enables coding articulate block shapes in the standard RGB image space using textures.

Figure 7 shows how sculpted prim technology enables to obtain realistic and very complex 3D objects.

4. The Valle del Sarno experience evaluation

The experience has been evaluated with a questionnaire we proposed at visiting students of the Valle del Sarno journey at the *10th* edition of *Borsa Mediterranea del Turismo Archeologico*. In the case students were too young for fully understand or evaluate the experience we asked to compile the survey only to their tutors.

Table 1. The submitted questionnaire

Hot impressions survey
1. How much were you involved in the experience ?
2. How much was the proposed experience interesting ?
3. Do you think that the experience could be improved ?
4. Was every content presented clear?
5. Did you like the practical on field activity ?
6. Did you like the 3D interactive demonstration ?
7. Did you like the 3D projection ?
8. Did you like the SL guided tour?
9. Did the practical on field activity be involving?
10. Did the 3D interactive demonstration be involving?
11. Did the 3D projection be involving?
12. Did the SL guided tour be involving?
13. Were sounds appropriate for the experience?
14. Were images appropriate for the experience?
15. Were videos appropriate for the experience?
16. Were SL environments realistic?

The questionnaire is composed by 16 questions that ask to evaluate several aspects like involvement of the experience, quality and detail of content proposed, quality and realism of 3D environments hosting the experience.

We ask to judge each question on a seven point Likert scale ranging from *1 for nothing* to *7 very much.*

The proposed experience and the questionnaire enable us to compare standard 3D technologies, like the stereoscopic projection, with SL always keeping in mind the differences among proposed experiences and technological availability and instruments offered. We must consider that, also if not accompanied by a SL expert avatar, the Nuceria theatre is always available in SL, while the stereoscopic 3D technology does not, requiring a specific hardware.

To avoid biasing judgment, we regularly exchanged experts between each sub-phase of the experience. In that way, comparison of SL technology with 3D stereoscopic and modeling experiences is not influenced by expert personality or technical skills.

Table 1 reports all the questions of the *Hot impressions survey.*

Some questions concern evaluation of involvement and interest of the overall experience while other questions require more specific evaluation of each visited environment. The last four questions are formulated to judge the historical and archeological value of the proposed reconstructions and environments.

The students and involved people found easy and quick to compile the questionnaire and the results are very encouraging. As first point we want to underline that the experience, exposing SL technology to a big number of users, has been also a way to observe a diffuse enthusiasm among visiting people.

While it could be obvious the realism and the immersion of 3D stereoscopic projection, it is an important result to point out the strong feel of involvement perceived by users during the SL assisted experience and the seeming realism of SL environments when used for historical reconstructions.

Figure 8 reports the box-plots of obtained results, in particular the upper part describes questions from 1 to 8 and lower one depicts the box-plots of answers to questions from 9 to 16. We adopt notched box plot: the notches represent a robust estimate of the uncertainty about the medians. Boxes whose notches do not overlap indicate that the medians of the two groups differ at the 5% significance level.

A first important observation on result can be that all median values are very high denoting an enthusiastic visiting students response in general.

The sample medians are depicted in the thin part of boxes.As shown in Figure 8, all questions received a median score *6* or above with the exception of question 3 and 13 whose median score is respectively *3* and *5.5*. The questions concern the improvability of the experience and the quality of sounds. While the evaluation of audio has been a little influenced by interferences among parallel sessions, the low median obtained by question 3 may be justified only with a wrong formulation of the question. Indeed, if we consider also other questions it does not seem the experienced to be evaluated as for question 3.

As an example, the quality of content is addressed also by questions 4, 14 and 15 that state a good quality perception for content clearness, images and videos.

A comparison of the SL experience with 3D projection or 3D technological excursus is possible by putting side by side the answer to questions 8 and 12 with respectively 7 and 11 or 6 and 10.

SL reconstruction has been evaluated also by question 16 specifically requesting feedback about SL theatre realism.

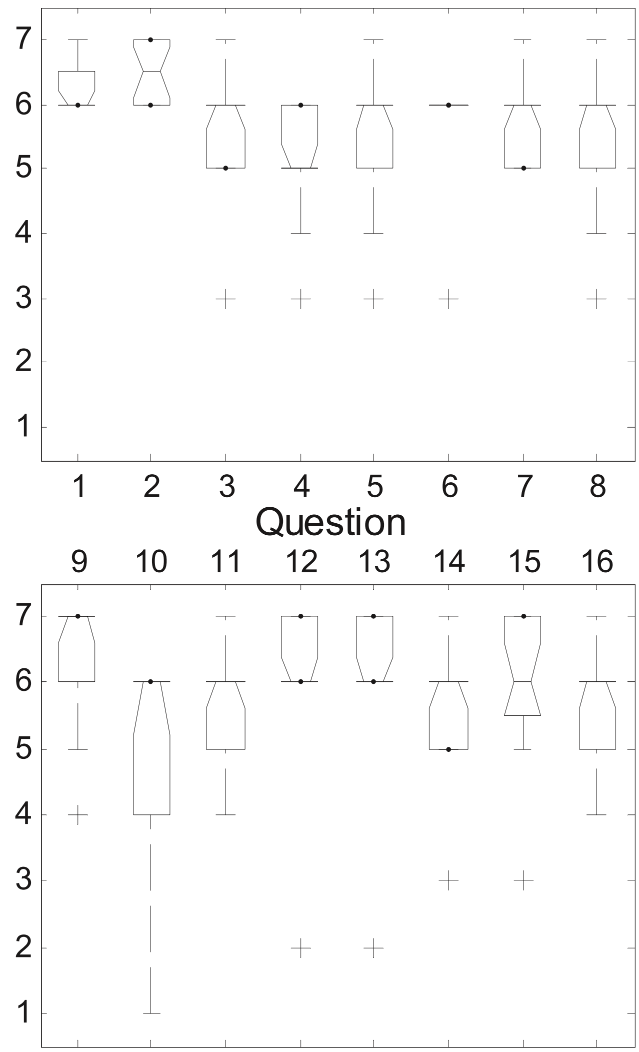


Figure 8. The questionnaire results

In Table 2 some statistic comparing SL with other 3D experiences are reported: the average results of question regarding 3D projection and 3D demo are subtracted from the average of questions 8 and 12 investigating SL and sample statistic are computed on results.

Lower row of Table 2 reports statistics about question 16 explicitly asking a judgment about Nuceria theatre reconstruction realism.

Table 2. SL compared with other 3D experiences

	min	max	avg	median	Std dev
vs 3D video	-3	1	-0,28	0	0,93
vs 3D demo	-1	1	-0,28	-0,5	0,51
SL realism	5	7	6,12	6	0,5

The comparison between SL and other 3D technology adopted during the experience denotes very little differences between the 3D environments when evaluated on user likeness and involvement. Indeed both differences SL/3D video and SL/3D demonstration reveal an average value of *-0,28* denoting that SL performs just a little poorer than other technologies. However the strong sense of realism perceived during the SL tour and the great number of contacts we recorded after the the *Borsa Mediterranea del Turismo Archeologico*, comfort us and give sufficient reason to continue exploring SL, and to enrich the Valle del Sarno island [17] with new archaeological and historical reconstructions.

5. Conclusion

In this paper we presented an hypermedial learning experience focused on the SL reconstruction of the Nuceria theatre (Nocera Inferiore, Salerno Italy).

The experience has been presented and evaluated during the *10th* edition of *Borsa Mediterranea del Turismo Archeologico*. The survey proposed to users after the journey, provided us with enthusiastic user feedbacks which strongly encourage similar activities.

In particular, it is remarkable to note that the student and people involvement obtained during the proposed SL excursion, even if the experience is assisted, is comparable, to analogous experiences adopting sophisticated technologies, such as 3D stereoscopic projections.

This perceived sense of involvement is a remarkable result particularly considering the large availability of SL technology which, associated to the resulting deep sense of realisms of SL reconstructed environments, suggests to extend the experience beyond and later the 10th *Borsa Mediterranea del Turismo Archeologico*.

Acknowledgments

The authors would like to thank all people of *Agenzia Sviluppo Valle del Sarno S.p.A.* for their indispensable contribution to the success of the experience.

References

[1]. Bifrost island, University of Southern Denmark, retrieved march 2008 from http://slurl.com/secondlife/bifrost/101/74/23/
[2]. Cochrane, K., Case Study: International Spaceflight Museum, *in the proceedings of the Second Life Community Convention*, San Francisco, USA August 20th, 2006, pp. 2-5.
[3]. De Lucia, A., Francese, R., Passero, I., Tortora, G. (2008). Supporting Jigsaw-Based Collaborative Learning in Second Life. To appear *in the Proceedings of The 8th IEEE International Conference on Advanced Learning Technologies* ICALT 2008.
[4]. Doherty, P., Rothfarb, R., Barker, D., Building An Interactive Science Museum In Second Life, *in the proceedings of the Second Life Community Convention*, San Francisco, USA, 2006, pp. 19-24
[5]. EDTEC Virtual Campus, San Diego State University, retrieved march 2008 from http://slurl.com/secondlife/Meadowbrook/220/82/23/
[6]. FIAT Ilha Oficial, FIAT Latino America retrieved march 2008 from http://slurl.com/secondlife/FIAT%20LATINO%20AMERICA/57/125/26/
[7]. Istituto Geografico De Agostini, Guida alla Valle del Sarno retrieved march 2008 from http://capri.valledelsarno.it/vds/guidadeagostini/vds_deagostini1-51.pdf
[8]. IBM's Innovation Island retrieved march 2008 from http://slurl.com/secondlife/IBM/106/4/23/
[9]. Linden Scripting Language, http://wiki.secondlife.com/wiki/LSL_Portal.
[10]. MIT, retrieved march 2008 from http://slurl.com/secondlife/MIT/51/35/23/?title=Massachusetts%20Institute%20of%20Technology
[11]. M. Rymaszewski, W. J. Au, M. Wallace, C. Winters, C. Ondrejka, B. Batstone-Cunningham, and S. L. residents from around the world. Second Life: the official guide. *Wiley Press*, 2007
[12]. Second Life. retrieved march 2008 from http://secondlife.com
[13]. Stonehenge - A full scale replica of this historical monument, retrieved march 2008 from http://slurl.com/secondlife/tefnut/28/31/61/
[14]. Proceedings of the Second Life Education Workshop at the Second Life Community Convention San Francisco, 2006.
[15]. Proceedings of the Second Life Education Workshop (2007). Part of the Second Life Community Convention. Chicago Hilton, 24th-26th August 2007.
[16]. Sculpted Prim, retrieved march 2008 from http://wiki.secondlife.com/wiki/Sculpted_Prims
[17]. Valle del Sarno, retrieved march 2008 http://slurl.com/secondlife/Valle%20del%20Sarno/200/30/30
[18]. Windlight retrieved march 2008 from http://wiki.secondlife.com/wiki/WindLight

Reviewers' Index

Piero Mussio

N
Paolo Nesi

O
Vincent Oria

P
Sethuraman Panchanathan
Antonio Piccinno
Sofie Pilemalm
Fabio Pittarello
Giuseppe Polese

R
Syed M. Rahman

S
Monica Sebillo
Timothy K. Shih
R.M.K. Sinha
Jaideep Srivastava
Peter Stanchev

T
Genny Tortora

U
Joseph E. Urban

V
Athena Vakali
Son T. Vuong

W
Ellen Walker
Chun-Der Wu

Y
Chi-Lu Yang

Z
Kang Zhang

Authors' Index

DMS 2009 Call For Papers

The Fifteenth International Conference on Distributed Multimedia Systems

Hotel Sofitel, San Francisco Bay, USA
September 10 - September 12, 2009
Organized by
Knowledge Systems Institute Graduate School

SCOPE

The DMS conference is an international conference series, which covers a wide spectrum of technique discussions, demonstrations, and student program/paper contests in the fields of distributed multimedia computing. Started in 1994, the series of conference has been held at Hawaii, Hong Kong, Vancouver, Taipei, Aizu-Wakamatsu, Japan, San Francisco, Miami and Canada. This time, the conference will be held in the Hotel Sofitel, Redwood City, California, USA.

TOPICS

DMS 2009 focuses on techniques, systems, applications, and theories in the fields of distributed multimedia computing. The conference organizers seek contributions of high quality papers, panels or tutorials, addressing various aspects of distributed multimedia systems and applications, for presentation at the conference and publication in the proceedings.

Topics of interest include, but are not limited to:

audio and video compression
MPEG, Quicktime, Windows API standards
image, video, audio content analysis and indexing/retrieval
image, video and audio watermark
3D audio and video
computer graphics and animation
modeling and analysis of distributed multimedia systems
OS support for distributed multimedia systems
distributed multimedia databases and computing
distributed object management
multi-paradigmatic information retrieval
multimedia human-computer interaction
multimedia communications and network architectures
mobile networks and mobile computing
multimedia stream synchronization and QoS control/scheduling
multimedia software engineering
multimedia processors and ASIC
multimedia technologies for people with disability
intelligent multimedia computing
intelligent multimodal interaction
multi agents, mobile agents and their applications
communication encryption
security of distributed computer systems
sensor networks for multimedia computing
Web servers and services
XML applications
Java, VRML and multimedia languages
visual and multidimensional languages for multimedia applications
multimedia digital libraries and mail systems
multimedia applications for CSCW
multimedia authoring tools and intelligent tutoring
cultural heritage multimedia applications
tele-conferencing, tele-medicine and tele-lecture
virtual reality, distributed virtual environment, and their applications
virtual school, virtual university, and virtual learning community
distance learning methodology, tools and systems
e-commerce, e-education and e-entertainment

The use of prototypes and demonstration video for presentations is encouraged.

WORKSHOPS AND SPECIAL SESSIONS

This year, DMS 2009 will be held in conjunction with workshops, conferences and special sections. Papers submitted to workshops or special sessions are invited by the program committee chairs of the workshops/sections. This joint organization aims to collect research results from different perspectives. The following workshops and conferences are being planned:

International Workshop on Visual Languages and Computing
International Workshop on Mobile Systems, E-Commerce and Agent Technology
International Workshop on Distance Education Technology

Please contact the conference program co-chairs if you are interested in organizing a workshop or a special session. A one-page proposal with the following items is required:

Title of the workshop/special session
Name of the program committee chair(s)
A list of program committee members
E-mail address of the corresponding program committee chair
A brief description of the theme

Each special session should include at least 5 invited (or accepted) papers. Each workshop should include at least 2 special sessions (10 papers). Paper review process of the workshop/special session will be arranged by the individual program committee chair. It is the responsibility of the individual chairs to ensure the quality of papers in the workshop/special sessions. The schedule of workshops/special sessions will be the same as the main conference (See Important Dates below). Each individual program committee chair can set up the Web site of the workshop/special session. However papers will be printed in the same volume as the main conference.

SPECIAL ISSUE OF JOURNALS

The DMS conference is closely coordinated with the International Journal of Software Engineering and Knowledge Engineering, the Journal of Visual Languages and Computing and the International Journal of Distance Education Technologies. Usually the best ranked papers from the conference, after rigorous reviews, extensive revisions and further enhancements, will appear in one or more special issues of the above journals. Papers suggested for the special issue(s) will be reviewed by external reviewers following the standard procedure of review stipulated by the respective journal.

CONFERENCE SITE (HOTEL INFORMATION)

The DMS 2009 Conference will be held in the Hotel Sofitel, Redwood City, California, USA. The hotel has made available for these limited dates (9/9/2009 - 9/13/2009) to DMS 2009 attendees a discount rate of $89 US dollars for single/double, not including sales tax.

INFORMATION FOR AUTHORS

Papers must be written in English. An electronic version (Postscript, PDF, or MS Word format) of the full paper should be submitted using the following URL: http://conf.ksi.edu/dms09/submit/SubmitPaper.php. Please use Internet Explorer as the browser. Manuscript must include a 200-word abstract and no more than 6 pages of IEEE double column text (include figures and references).

INFORMATION FOR REVIEWERS

Papers submitted to DMS'09 will be reviewed electronically. The users (webmaster, program chair, reviewers...) can login using the following URL: http://conf.ksi.edu/dms09/review/pass.php.

If you have any questions or run into problems, please send e-mail to: dms@ksi.edu.

DMS 2009 Conference Secretariat
Knowledge Systems Institute Graduate School
3420 Main Street
Skokie, IL 60076 USA
Tel: 847-679-3135
Fax: 847-679-3166
E-mail: dms@ksi.edu

If you cannot submit electronically, please send four copies of the complete manuscript to the above postal address.

IMPORTANT DATES

March 1, 2009 Paper submission due
May 1, 2009 Notification of acceptance
June 1, 2009 Final manuscript due
June 1, 2009 Early conference registration due