The Interactive-Music Network

DE4.1.2
Music Notation Coding

Version: 2.2
Date: 10/06/2005
Responsible: DSI

Project Number: IST-2001-37168
Project Title: The Interactive-Music Network
Deliverable Type: report
Visible to the Working Groups: all
Visible to the Public: yes

Deliverable Number: DE4.4.2
Contractual Date of Delivery: 31-May-2005
Actual Date of Delivery: 20-June-2005
Title of Deliverable: Music Notation Coding
Work-Package contributing to the Deliverable: WP4.4
Nature of the Deliverable: report
Working Group: Music Notation
Author(s): P. Nesi, P. Bellini, J. Barthelemy, G. Zoia, N. Mitolo

Abstract:
This document is on the integration of music notation in multimedia applications, the most interesting future applications are considered and the requirements are defined. The most referenced music notation formats are considered and a preliminary evaluation with respect to the functionalities needed for the new applications is reported.
Moreover this document reports the ongoing activity for the integration of music notation inside the MPEG framework.
Keyword List: music notation, standards, MPEG, XML

Table of Content

1 Executive Summary and Report Scope

The aim of Working Group on Music Notation is to stimulate the realisation of a widely adopted format for music notation representation. A format that will deal with the needs of all the actors (publishers, music editor producers, copists, integrators, etc.) involved in the realisation and the distribution of an "interactive" multimedia music piece. We don't have to see music notation only related to music printed on a sheet, now and much more in the future, music notation will be accessed using different kinds of devices, from the PC to the UMTS terminal, from the classical printed sheet music to the electronic lectern.

Currently the WG on Music Notation and the WG on Standards are working on producing requirements for the integration of music notation in the MPEG framework and on analysing (technically) how the integration could be achieved. This document reports the work performed on these aspects.
In this document, the main problems and the most promising evolution trends of music notation are reported. To this end, a comparison of the most representative music notation formats is included. The comparison has been performed on the basis of a set of aspects which has been related to major problems.

2 Introduction

The modeling of music notation is an extremely complex problem. Music notation is a complex piece of information which may be used for several different purposes: audio coding, sheet music production, music teaching, entertainment, music analysis, content query, etc. In the Internet Multimedia age music notation and associated applications are being left behind. Many other applications are getting the market and most of them will become more diffuse in a short time. A unique and comprehensive format for music incorporating other media is required. In order to satisfy diverse applications several aspects have to be considered, ranging from problems of information modeling to integration of the music notation itself. The main problems are related to the organization of music notation symbols in a suitable and acceptable form. Such a set of rules has never been formalized, although they are informally used in several books and considered part of the experience of professional copy editors (engravers) of sheet music. Other problems include adoption of classical music notation in applications subject to DRM (digital right management), versioning, multimedia integration and navigation, etc.
The evolution of information technology has recently produced changes in the usage of music notation, transforming the visual language of music from a simple coding model for sheet music, to a tool for modeling music in computer programs. Enabling cooperative work on music and other information integration tasks is also necessary. More recently, users have discovered the multimedia experience, and thus, the traditional music notation model is likely to be replaced with something much more suitable for multimedia representation of music. The improved capabilities of computer programs are going to solve also early music notation formatting problems with powerful tools based on artificial intelligence technologies.

The problems of notation modeling and visualization of music scores have been addressed in computer systems several times -- (Anderson and Kuivila, 1991), (Blostein and Haken, 1991), (Dannenberg, 1993), (Dannenberg, 1990), (Rader, 1996), (Gourlay, 1986), (SelfridgeField, 1997), (Blostein and Baird, 1992), and (Pennycook, 1985).
Among many possible computer-based applications of music, the notation editing for professional publishing and visualization is one of the most complex for the intrinsic complexity of music notation itself (Ross, 1987), (Heussenstamm, 1987), (Wood, 1989), (Byrd, 1984). Music publishing requires the production of high quality music scores in terms of number of symbols and their precise placement on the staff. Music scores have to be presented to professional users (musicians) with specific relationships and proportions among symbols. The formal relationships among symbols are often neglected or misused by most commercial computer-based editors. Their aims consist of producing music by means of MIDI (Music Instrument Digital Interface) interfaces or sound-boards, or in acquiring music by means of a MIDI interface, or simply in editing music on the computer so as to manage sound-boards, etc. In such events, a limited number of music notation symbols is needed (e.g., notes, rests, accidentals, clefs and points -- approximately 50 symbols), which makes impossible to reproduce through a MIDI interface the different effects specified by many others music notation symbols.
Sometimes the sets of music symbols available in professional editors (e.g., Score, Finale, Sibelius, etc.) for publishing music include also *instrumental and execution symbols* (symbols for describing the behavior of the musicians while playing a specific instrument -- e.g., up or down bow, with mute, without mute,

fingering, string numbers, accents, etc.) for: strings, harps, drums, flutes, etc.; *orchestral and repeat symbols*, etc. These symbols together with many others, are needed when main scores and parts for classical music, operas and ballets have to be produced, and they are mandatory when scores are used in music schools to train students to perform specific executions and interpretations of music. The number of elementary symbols is close to 300 in the commercial editors which are commonly used to prepare music scores for publishing, Such symbols are frequently treated as components to implement more complex symbols. Typically, musicians have to personalize/prepare a score for the execution; therefore, even this great number of symbols is not powerful enough to meet all their needs. On these grounds, some music editors provide a font editor to help the user in adding/creating symbols; which unfortunately, are treated as simple graphic entities.
Commercial music editors for publishing are mainly inclined to place music symbols on the score page rather than to collect relationships among symbols and then organize the visual information accordingly. They are orientated towards printing music, since this is what they consider their most important application. They provide a complete set of symbols which a user skilled enough in both music notation and computer graphic interface can use in order to produce professional sheet music. In music editors of this kind, the arrangement of many music notation symbols is often left to the users' competence. Music symbols (excluding notes, rests and a very few other symbols) are mainly considered by music editors as graphic elements to be placed in any position on a score page without feeling bound to address any problems related to the music notation like, for instance, the visual and syntactic relationships among symbols. This results in many music editors allowing several notation symbols to be placed in incorrect positions, thus producing strange and incorrect music scores. Eventually it follows that users are left with no technical support on how to place symbols.

For these reasons, professional music editors are powerful, but at the same time they are difficult to handle for non-expert users in music notation. Typically, musicians can read music, but they are pretty unfamiliar with rules for arranging symbols (e.g., when a symbol is associated with symbol B, symbol C has to be moved up one-half unit). Musicians have no problem when asked to read a correctly annotated score, but can be puzzled when they are faced with non-perfect visual constructs. Musicians are in fact artists, not music engravers nor notation technicians. It often occurs that conductors and composers are more sensitive to problems of music notation and visual expressiveness; while archivists in music theaters, music engravers, and music notation teachers are the experts of music notation problems.

2.1 Domains of Music Notation

A deep analysis suggested that the music notation mainly comprises of four domains of representation. Elsewhere, these domains may be referred by different terms, but here we categorize the domains thus:

- **Audio: (also called gestural, or performance)** description of how a given note or chord has to be played in terms of parameters in the frequency domain, orchestration, sampling of audio signals to be used as wave tables to play notes, etc
- **Logical:** description of the music notation structure in terms of sequence of notes, rests, chords, etc. organized into instruments, voices, measures, etc. This also includes the expression symbols that represents how a certain note or sequence should to be executed: markers (general terms used in some programs to identify accents, ornaments, etc.), accents, ornaments, accidentals, key signatures, clefs, dynamics, key signatures, crescendo and decrescendo, phrasings, bowings, instrument indications, metronomic indications, multilingual lyric, etc. Many performance parameters (e.g., MIDI) can also be included in this domain. Also needed is the modeling of rules to generate alternative music notation for specialized consumer groups to be associated with the logic aspects. In this case, strong relationship with Audio Domain is needed. All aspects that are common to Audio and Visual aspects should be placed in the logical domain to create what we calling the bone of the music representation.
- **Visual:** parameters describing, the visual representation of the logical content as absolute and/or relative positions (spacing), colours, stem direction, beaming structure and slope, refrain/chorus, generic rests, slurs shape (knots based or other), 8va or 8ba or 15ma, repetition symbols, barlines, etc., and also the eventual image representation of the score, graphic representation of modern music notation symbols, formatting representation of the main score and individual parts in terms of justification parameters and page layout, ordering for symbols positioning around the figures, etc. (for example, which symbol has to be positioned closer to the notehead if a staccato, a slur, and a fingering are present on the same note) This domain implicitly also includes the formalization of the relationships among music notation

symbols. It also includes the description of the image sheet and its relationships with the music notation structures and elements (parts, measures, etc.)

- **Graphic** representation deals with the mechanisms to produce the music notation on an image sheet either as print output or screen canvas. This is performed by using graphic primitives and fonts, and associated parameters such as size, font and style together with mechanisms to avoid confusions (e.g., collisions, overlapping, etc) among graphic details. Details from the visual domain have to be considered as high level aspects/directives for the graphic representation. Separating graphic and visual aspects gives the possibility of generating alternative visualizations of music notations, for example for visually or print impaired users.
- **Analitical** information: information about the work, which might include bibliographic information, as well as interpretive information that ranges from phrase markers and roman numeral analysis of underlying harmony to a Schenkerian graph.

The above definition domains for music representation has been in some how derived from other definitions of domains. One of the most representative is that provided for SMDL.

Emerging New Applications

Recently, with the spread of computer technology into the artistic fields, new needs for computer-based applications of music have been identified: (i) cooperative music editing in orchestras and music schools, as with the project MOODS ESPRIT (Bellini, Fioravanti and Nesi, 1999) (Bellini, Nesi, Spinu, 2002); (ii) music score distribution via Internet, as with many WWW sites distributing music scores or MIDI files; (iii) multimedia music for music tuition systems: edutainment and infotainment. We have started to investigate these two new fields since 1994. MOODS consists of an integrated system of computer-based lecterns/stands for cooperative editing and visualization of music. MOODS is an innovative solution for automating and managing large amount of information used by (i) orchestras during rehearsals and public performance of concerts, operas, ballets, etc. (ii) students of music during lessons in conservatories and music schools, (iii) publishers during massive editing of music.

The targeted MOODS end-users are theatres, itinerant orchestras, groups of musicians, music schools, television network orchestras, and publishers of scores. MOODS can be used to: (i) reduce the time needed for modifying main scores and parts during rehearsals in a cooperative way; (ii) manage (load, modify, and save) instrumental and personal symbols on main scores and parts; (iii) manage and reproduce the exact execution rate at which each measure of a score has been performed; (iv) automate page turning during rehearsals and final performances; (v) change music pieces quickly or restart from marked points; and (vi) manipulate the main score and all instrument parts together with the full music score in real time. Computerized music lecterns can be used by musicians to avoid transporting heavy paper music scores, to save their work, to manage versions, etc. A distributed system with the above features must be capable of showing music in different formats and at the same time must support cooperative editing of music in different formats, which means showing the changes of one operator to the other ones in real time. A public demonstration of MOODS functionalities was given at La Scala Theatre in Milan. The main difference between classical music editor and MOODS is the availability of cooperative work on music and the presence of integrated multimedia aspects. When cooperative work is relevant and the music has to be visualized at several resolutions regardless of the visualization support features -- for example, low/high resolution monitors and printers -- the following two main functionalities have to be available.

a clear distinction between the music notation model and the visualization rules: automatic reformatting while taking into account the user's needs, transposition, page dimension, etc.

music notation model has to be abstract enough to allow the *interactivity with music at a symbolic level*: adding and deleting symbols, changing symbols features and making those changes available regardless of the visualization details without any reloading of the music score. To this end, the music notation and model has to support an indexing model (Bellini, Nesi, Spinu, 2002). This makes it possible to define policies of versioning, selective and un-linear undo, navigation among music notation symbols, etc.

This feature seems to be very far from being a reality in present Internet applications, even though it is of great interest for cooperative applications -- e.g., virtual orchestras. It is noteworthy thinking of cooperative applications as something which will undergo a strong implementation among Internet Applications in the very near future.

For the second type of application music can be distributed by using images, audio and symbolic files. At present, the distribution of music via Internet consists only of images of music scores or simple symbolic files, while audio files are widespread and better distributed. The music received from Internet can be either *interactive or not*. For interactive music we intend music that can be manipulated in a certain measure: addition/deletion of symbols, transposition, reformatting, etc., without the violation of the copyright law. Images of sheet music do not allow the music manipulation, while the MIDI model is too coarse to meet the needs of professionals, because MIDI provides a reduced set of music notation symbols.

In the past, the language called NIFF (Notation Interchange File Format) was supported by editing/publishing and scanning programs, for creating an interchange format for music notation (NIFF, 1995). In the following years, the NIFF format has been abandoned as an interchange format by several companies and publishers. Recently MusicXML has been proposed with the same aim, but with far greater success, with over 25 implementations to date. (Good, 2001).

The Internet is at present dominated by the distribution of documents through the so-called mark-up languages derived from SGML, HyTime, XML, etc. (Sloan, 1997). A mark-up language consists of a set of constructs to express how text has to be processed with the main aim of text representation and rendering according to the style model. Generalized mark-up languages specify only what you have, rather than how to

visualize it. The visual aspects are specified by using standard tags which state the typographic features: spacing, font change, etc. For example, when using XSL a formatting style should be defined. This means that a clear distinction has to be made between the content and the formatting aspects. These languages are mainly conceived for describing monodimensional pieces of information like texts and fail to model relationships among symbols at a logical level. For this reason, mark-up languages do not fit for any direct support to cooperative work; since they are not structured enough to be independent of the formatting aspects. In fact, mark-up languages were created for formatting textual documents and not for defining relationships among document symbols. This is one of the main problems which prevents from an unreserved adoption /acceptance of SMDL or other mark-up languages. Recently, several other XML compliant mark-up languages for music modeling have been proposed, among them: MNML (Musical Notation Markup Language), MusicML, MML (Music Markup Language), MusicXML, etc. The MNML is a simplified version of SMDL. In MNML, it is possible to describe completely the basic melody and the lyrics of a music piece; on the other hand it fails to describe all the needed details of real music scores (annotations, publishing aspects, etc.). MusicML can only represent notes and rests. Even staccato points and slurs are missing.

The adoption of a symbolic model totally separate from the visualization issues makes the distribution of interactive music a very complex task. In MOODS (Bellini et al., 1999), a cooperative editor of music has been proposed. It was built on a non-XML-based format. The solution was the definition of a specific language called MILLA (Music Intelligent Formatting Language) able to define rules for a formatting engine while letting users also have the possibility of forcing exceptional conditions. WEDELMUSIC model and language can be considered the XML evolution of MOODS format. With WEDELMUSIC several MOODS' early problems for music modeling have been solved, in addition WEDELMUSIC is a multimedia model. WEDELMUSIC provides an effective framework which includes most music symbols and their relationships. On such basis several new and innovative applications can be built and some exceptions and several modeling problems already highlighted in Selfridge-Field (1997), Gourlay (1986). WEDELMUSIC format presents multimedia capabilities and includes identification, classification, symbolic, visual, protection, animations, image score, image, document, performance, video, lyric, aspects, synchronization, etc. It keeps separate visual/formatting from symbolic aspects. WEDELMUSIC format can be profitably used for new applications and as a format for interchanging symbolic description of music scores. Formatting rules and the corresponding MILLA engine of WEDELMUSIC can be found in Bellini, Della Santa and Nesi (2001).

The distribution of sheet music via Internet can be considered one of the emerging new applications of music notation. Presently great part of the sheet music distributed via Internet is in the form of PDF or Image files. In both cases, the music delivered is not interactive since no transposition, or annotation can be performed. Other more interactive solutions have been produced by Coda/Finale, Sibelius and Noteheads. They distribute music notation in their proprietary formats and this can be viewed and played in an Internet Browser via MIDI by using specific plug-ins or ActiveX. For these applications, the Digital Right Management, DRM, capabilities are particularly relevant to protect the distributed content. Typically, the models permit printing the sheet music only a limited number of times. In this field, WEDELMUSIC propose an integrated solution for distributing multimedia content, files in any format, including also music score in a protected manner among archives and these to their attendees (Bellini, et al. 2003).

2.2 Integration of Music Notation in the MPEG framework

Several products are currently available in the market that present some form of integration of music notation and multimedia (see some examples in the following of this document). These products are in the area of music education (notation tools integrating multimedia), music management in libraries (music tools integrating multimedia for navigation and for synchronization), karaoke (mainly synchronization between sound, text and symbolic information), etc.
The integration of music notation in MPEG could completely satisfy the requirements of these tools and much more than that, adding: interoperability, portability for i-TV and mobile devices, scalability of complexity, etc., permitting to these tools to integrate the powerful model of MPEG for multimedia modeling, representation, coding and playback.
The integration of music notation with MPEG will open the way for a large number of new applications and markets, for instance:

1) the multimedia electronic lecterns for classrooms and lecture halls,,
2) music education via i-TV,
3) multimedia content integrated with music notation for multilayer music distribution
4) multimedia content integrated with music notation on music instruments, mobiles, PDAs, etc.
5) The editing of all kinds of music on computer screens using the apropriate notation.

This initiative may increase the present market for music notation that presently is mainly limited to sheet music production, and may open the path to create very interesting new applications, renovating the present applications that already use some integration between multimedia and music notation.

It should be understood that "music notation" in this document means all kinds of symbolic music notation, including different styles of Gregorian Chant, Renaissance, Classic, Romantic and 20th Century styles, simplified notations for children, Braille and other forms yet to be invented.
The importance of a reactivation of writing in music should not be underestimated. The repercussions would have a wide effect on music education and on the development of music as a whole. This, in turn, would be of great long term benefit to the music industry, quite apart from the applications, mentioned in this document, which can be written and sold in the shorter term.

2.3 Music Notation Applications

Music is a complex piece of information, although it is usually considered in its audible representation: many notations has been developed over the years and ages to represent visually the information needed by a performer to play the musical work and produce the music as intended by the author. However music notations are not intended for exclusive use of performers, musical scores are considered works of art as the music they convey, which is especially true for 20th century music.
Symbolic music representation generalizes music notation to model not only the visual aspects of a music score but also audio information or annotations related with the music piece. One of the first Symbolic Music Representation formats is MIDI (www.midi.org), which is good to transport music event information and very poor on the audio and visual representation sides.
The evolution of information technology has recently produced changes in the practice usage of music representation and notation, transforming the music notation from a simple visual coding model for sheet music to a tool for modelling music in computer programs and electronic tools in general. Currently symbolic music representation/music notation is used for several different purposes: sheet music production, music teaching, audio rendering, entertainment, music analysis, database query, performance, etc.

Recently, music representation and associated applications are rapidly evolving with the integration of multimedia and interactivity aspects. Users have discovered the multimedia experience, and thus, the traditional music notation model has been in many cases replaced with more suitable multimedia representations of music. Many new applications are getting the market attention, and most of them will become more diffused in short time. The industry needs a unique and comprehensive representation of music information integrated with other media. To make possible the realization of these diverse applications, we have to take into account of several aspects ranging from problems of information modeling to integration of the music representation. The main problems are related to the organization of music elements and symbols in a suitable and acceptable form to cope with the more general concepts of music elements and relationships with audio visual aspects and content.

In addition to these problems, there is also the need of representing non-western music notation such as music from the Far East countries (China, Korea, Japan…), or Middle East (Arabia, Northern Africa), and Northern Africa, etc. In order to support the inclusion of these music representations, the music model has to be kept sufficiently open and general. Educational aspects have also a strong influence on music representation models, since they need to integrate semantics information for rendering/playing and performance evaluation of music elements. The problem alone of notation modeling and related rendering have been addressed in computer systems several times.

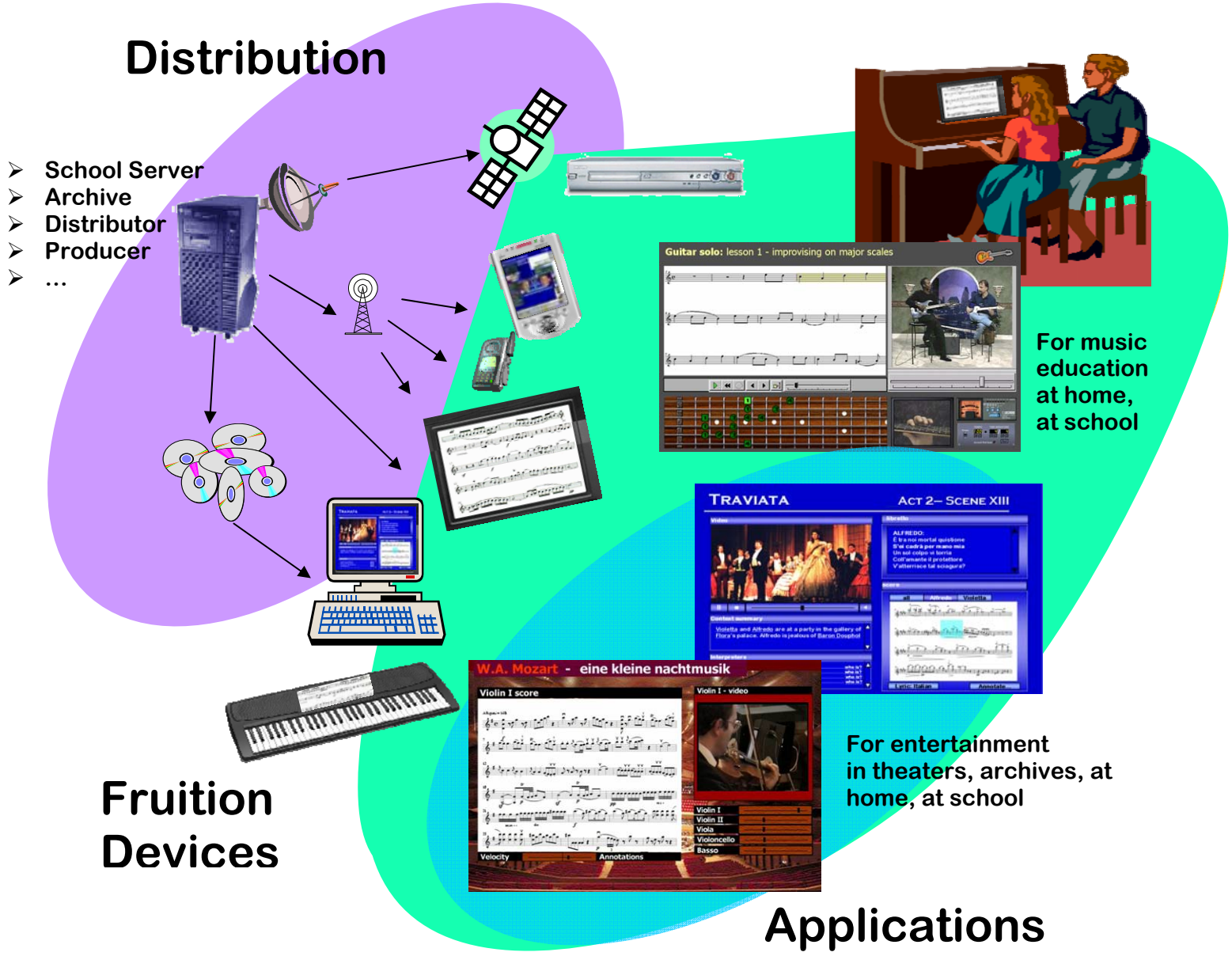


Fig.1 – General scenarios exploiting MPEG SMR for entertainment and education. MPEG-4 with SMR support distribution by means of satellite data broadcast, internet, wireless or traditional media towards theatres, home, archives, mobiles devices such as i-TV, tablet PC, PC, PDA, phones

With the spread of computer technology into the artistic fields, new application scenarios for computer-based applications of symbolic music representation have been identified by several activities (see Fig.1): (i) multimedia music for music tuition systems such as: IMUTUS, MUSICALIS, Freehands, Yamaha tools, etc.; (ii) multimedia music for edutainment and infotainment, in archives such as in WEDELMUSIC or in theaters such as OPENDRAMA. (iii) cooperative music editing in orchestras and music schools, as with the project MOODS. Thus, a new concept of multimedia interactive music is growing also to the thanks to the several innovative products and supported R&D projects of the European Commission, like the above mentioned ones and the MUSICNETWORK center of excellence http://www.interactivemusicnetwork.org All these applications are taking advantages from the multimedia integration and of distribution via Internet, satellite data broadcast, etc., to reach a large number of users and the mass market.

In the following the most relevant applications of music notation and music notation integrated with audiovisual are reported. After the list, the most relevant have been detailed.

2.3.1 Software Applications

Multimedia music notation distribution

- Distribution of music with associated text and musical score
- Editing
- Annotating
- Integrating with multimedia aspects
- Transcoding audio in music notation
- Rendering, interpreting music notation for producing audio

Music education and music courseware construction

- Authoring tools for music notation creation with integration with multimedia objects
- educational tools with virtual reality showing the teacher that is playing the instrument.
- educational tools connecting audio-visual information to music notation symbol
- singer training, see karaoke in the following
- play training, see karaoke in the following
- ear training, see karaoke in the following
- music notation games for kids,
- Direct correction while playing, instantaneous corrections while the pupil plays, corrections can be with animations, video, audio, documents, etc.
- Tools for music education on i-TV
- Creation of tutorials
- electronic lecterns for music education
- lecterns for singer reporting music notation synchronized with the orchestras
- karaoke

Music editing in all domains (for domains see the section 2.6): This application may include multimedia aspects for the production of multimedia music applications, such as for music education and distribution

- music editors for producing sheet music
- printing music notation
- publishing music notation pieces.
- music composition, creation of a new music piece,
- Music notation transposition
- Adding, navigating and playing multimedia music

Karaoke, Entertainment tools, play along, etc. (based on synchronization functionality and play)

- presenting music score while a professional player on video is playing its instrument, also for educational
- presenting music score while a professional real audio is played
- presenting music score while notes under play are tuned on
- presenting music score while animation of a virtual player is shoving the right position of the hand and body posture.
- Presenting music score while the lyrics is shown, lyrics can be multilingual, that can be simply plugged to the music notation model.
- Presenting music score while the scene of the opera is shown in the video, the score can be a piano reduction, or the user can select the view that prefer to follow. The video of the opera can be substituted with a virtual representation of the opera.
- Presenting an image of a music score (may be ancient or handwritten music) linked and synchronized with symbolic music notation

Multimedia Music manipulation in digital library

- Music annotating in all domains
- Music editing, transposing, playing, etc. according to DRM rules

- music reading,
- Music notation navigation with multimedia integration
- comparison of music pieces,
- Multimedia music Navigation in all domains
- Music searching in all domains
- Music formatting

Multimedia Music navigation in all domains, including navigation in the multimedia content that can be connected to music notation,

- o tools for music education via i-TV,
- o educational purpose
- o integrating Music Notation and MPEG4

Music searching in all domains

- o Providing simple queries
- o Structuring and searching
- o modeling music notation as the structure/description of a music piece

Music annotation in all domains

- o Adding personal annotation on a music piece
- o Educational application
- o Entertainment applications

Music formatting

- o From logic to visual domain independently on the device, PDA, Tablet, etc.
- o Production of Braille music,
- o Production of Spoken Music,
- o etc.

2.3.2 Hardware & Applications

Piano keyboards

- with multimedia facilities, graphic user interface, etc.
- showing the right position of the hands, a video, the karaoke,
- synchronization with other keyboards

PDA and Tablet PC tools

- electronic lecterns for music education
- electronic lecterns for cooperative work in orchestras and groups
- lecterns for singer reporting

2.3.3 Multimedia music distribution through physical devices and online

One of the biggest advantages of digital formats and their integration is the possibility of distributing content in a variety of forms in a completely integrated manner. The distribution of multimedia music content including music notation can be used for a large variety of applications of education, edutainment, entertainment, infotainment, etc.

At present with traditional distribution via CDs or sheet music these aspects are neglected, while the 12% of the European Population is capable of reading music and playing an instrument. In addition, the largest market for music notation is closely related to music education. About 80% of young students in Europe are involved in the study of music, that implies that a large part of them is capable of music notation reading and music instrument playing in some measure.

For these reasons, in the same way that a text can be synchronised to images and sounds in commercial DVDs, it is strongly desirable to have the possibility of distributing content in different forms related to the same piece of music, including music notation for various (or separate) several instruments and voices. This would allow on the client side

- manipulation of music score and thus generation of
  - transposition of a music score which has been provided by a server,
  - piano reduction,
  - specifically formatted music for printing music on a different device or for showing music score in a different format: Braille, tablature, etc.
  - guitar tablatures,
- Visualization/playback of
  - Music score synchronized with audio (MIDI or recorded from real performance) or video,
  - multilingual lyrics to the same music notation,
  - lyrics with music notation synchronized with video or audio (music karaoke, vocal and aural training),
  - animated hands (via virtual reality) for showing musical instruments being played,

Most of these aspects have to be automatically performed on the basis of a music piece and not encoded for each specific music piece as in many software programs for music education and entertainment.
DRM, in this case has to be flexible and precise enough to allow:

- specification of symbols that can be added by the user
- specification of features that can be modified by the user
- specification of costs and rules for the above activities, including renting period, etc.
- managing activities such as printing, playing, annotations, transposition, etc. giving the possibility to the user to save their preferences and obtained result.

Presently these aspects are in the MPEG21 while the Music Notation is not in MPEG so it would be impossible to integrate all these aspects in unique MPEG compliant tool.
There is currently no generally available DRM solution for music notation (only WEDELMUSIC and Scorch allow this), but the above aspects are in the MPEG21 standard. Integrating music notation into the MPEG DRM standards would be of benefit to both music notation and MPEG-21 - which would acquire an important new application field.

2.3.4 Music education and courseware construction

For music education and courseware construction music notation needs to be integrated with multimedia capabilities. Music courseware need to have music notation integrated with video, images, documents, animation, audio, video, etc. In addition, a music course may present some exercises which require special music notation symbols (e.g., given by the teacher for the student, user defined) and/or audio processing (e.g., play training assessment, theory training assessment, etc.). For these reasons, for music education and courseware production the users on the client (music content usage) and sever (content generation) sides have a visual representation of the musical score (with full access to audio and logical aspects), with full capabilities of music manipulation in all domains and a support for the synchronisation and establishing relationships with other media.. Therefore, system and model have to provide the possibility of:

- music editing, transposing, playing, formatting, etc.; this implies the dynamic alteration of aspects in the logical, audio and visual domains.
- Recognition of notes played by an instrument or voice, using pitch recognition or other technologies. This means that the audio events have to be compared with what is expected by the notation conventions at the logical/symbolic level, and that an apropriate visualisation is then created.
- synchronization of notes played with the musical score, so as to show in a visual manner the correct point in the score being played. This requires the presentation of the visual aspects synchronized with the audio possibly allowing for changing the velocity of performance.
- assessment with respect to semantic annotations (how to execute a certain symbol, how to assess a certain execution of a music notation piece from the pupil, etc.) related to the above three domains: error recognition, evaluating a performance (scoring the execution for didactical purpose), learning aparticular performance practice.
- playing along: automatic accompaniment of the user by having the computer play selected voices/parts. This can be performed by using real audio recorded, generating MIDI, or by a more sophisticated music

notation rendering. The system must be capable of automatically following the pupil performance and to product comments and corrections on-line or off-line.
- Cooperative work on music notation for classes, for orchestral work on music lecterns (e.g., Tablet PCs), for managing rehearsals, etc.
- Annotation, for adding teacher and pupil annotations that may be textual or in music notation; see also Section 3.2.6 on annotation. This also includes capabilities of versioning.
- Navigation among music notation features and multimedia aspects and content.

I-TV for music distance learning. A Canadian attempt to develop a distance-learning learning module for music theory (for 16/17 year-olds). At the time there was no way for a teacher in Province A to see what the student in Province B was doing. In fact, there was no way to write freehand on a computer screen, coordinate the results with sound output, solve all the latency problems inherent in transmission, and get the teacher's response...in real time. Solid theory software that works on a stand-alone computer is still a very scarce commodity, whether or not the technological components are in place.

2.3.5 Music searching in all domains

On the client side, it is very important to be capable to specify a query in terms of audio, logical and visual aspects, rather than by relying on traditional cataloguing and identification methods. A query to find music pieces could be expressed by:

- Humming (monophonic), playing a MIDI keyboard (polyphonic), playing an acoustic instrument (either);
- writing a sequence of notes with a music editor, and specifying for some of them the range of pitch, duration, markers, audio effects, etc.;
- giving a rhythm with a sequence of note durations;
- selecting (by mouse or other means) a segment of an onscreen music notation score (monophonic or polyphonic, query by example). It could be done even selecting two music pieces as a sort of range;
- requesting the presence of educational aspects, annotations and support.
- integrating cataloguing aspects with query by content: logical, audio and visual representation aspects.
- Combination with logical operators of these queries.

2.3.6 Music annotation in all domains

The users from both server and client side need to be able to annotate music interactively by using

- personal or predefined music notation symbols (for example, to explain the semantic of music symbol or a possible comment on some music notation part), or
- simple textual or multimedia information (for example, the addition of an image or a video related to the music aspect).
- Audio details (for example, the example of the interpretation of a music notation symbol or music part as interpreted by different famous artists or by your teacher, etc.)

visual details and representation constraints (for example, the user may decide to hide a given symbol or to change it position with respect to the others, etc.). These annotations should be stored separately, and should be semantically linked to symbolic elements of the score, like measure, beat, note, voice and so on. They could appear differently in the main score and the parts, when the music is executed, printed or visualised, etc.

2.3.7 Music editing in all domains

Music editing is typically done on both the server and client sides. Users need to manipulate the logical, audio and visual aspects of a piece of music, and to relate these aspects to each other within multimedia applications. Editing is nowadays typically carried out on computer screens, but it is still generally oriented towards the production of printed music. Multimedia applications for composing or creating material for educational purposes do not, however, need to maintain this restriction. Multimedia notations can use windowing systems to present linked levels of detail unthinkable within the paper paradigm.

2.3.8 Music formatting in the visual domain

The music formatting is very useful for creating the visual and graphical details from logical representations using rules such as performed in [13] and other languages. These can be loaded as defaults or defined by the user for customising and profiling his needs.

This feature is very important for music editing tools (used by copyists, composers or music lovers) that need to create sheet music from scratch. It is also very important for Music Recognition tools that must recreate the music page from a logical description recognised from an image (e.g., via Optical Music Recognition, transforming paper-based music to symbolic machine representation) or from audio processing approach (e.g., pitch tracking, transforming wave source to symbolic machine representation, audio music recognition or transcription).

This feature also includes the possibility of defining different formatting styles.

2.3.9 Music navigation in all domains

Music notation navigation is a bi-directional feature. Users need to be capable to pass seamless from music notation items to other elements of the multimedia model and vice-versa:

- The user could browse a music score and follow hyperlinks associated to notes or measures, or other entities of the logic, audio and visual domains. The hyperlinked content could be external (a URL) or internal (an MPEG-4 component, element). The content associated with a note/measure could be a video/image/text explaining a difficult passage, for example.
- The navigation could be into the music notation structures and lower-level elements of the logical domain. From an MPEG4 elements it is useful to reach the music-notation content. For example, presenting the associated music notation with a video presenting the hand of a pianist, etc. This is a feature very useful feature for educational tools.

2.3.10 Music formatting and presentation in other representations

Producing enhanced music notation is a laborious activity. The production of these enhancements requires expertise and resources retrieved from all the use-cases described above. Additional expertise is required to combine all the issues mentioned above into a workable and practical set of resources and operations, enabling the translation from the ‘normal’ domain to the ‘accessibility’ domain. Incorporation of these accessibility notions into a widely used multimedia representation format would enable:

- Enhanced browsing experience with specialized functionality for specialized assistive needs (for example, when a non expert perform music editing it could be useful for him to have a verbal description of the context: measure 45, quarter chord of a D and C flat with arpeggio, etc.). This is very useful for impaired, young students, dyslexic people, etc.
- Synchronizing delivery of new content available in such a format for print impaired and sighted (for example, to have a verbal description of music synchronized with its playing).
- customizing the presentation of the musical content to the needs of the individual print impaired user (for example, selecting symbols to be verbally described or printed, selecting a simpler model for the visualization, for example without beaming, etc.).
- preserving specific knowledge and resources to allow the accessibility of musical information for other impairments.

2.4 Application Scenarios for Music Notation integrated in MPEG

The purpose of this section is to present two straightforward application scenarios that introduce Music Notation technology integrated with existing MPEG-4 and possibly MPEG-7 technology. These two application scenarios are useful first to exemplify to the MPEG and Music Notation communities two simple cases where Music Notation and other multimedia object types are integrated resulting in mutual added value; secondly, these scenarios can be useful to better understand and consequently refine the definition of the requirements in order to possibly approach a call for technology as a next step.

The first section introduces a scenario related with interactive multimedia content distribution for an “enhanced karaoke” application; the second section introduces a slightly more complex example (at least in terms of music notation functionality) related with music education and interactive courseware. Of course many other application scenarios could be presented, following the template of this document.

2.4.1 Scenario 1. “Enhanced” karaoke

In this scenario, the purpose is conveying to the user a set of multimedia objects so that he may be allowed to interact with this content by selecting or stopping part of it and replace the stopped components by local performances. Since karaoke is a successful application only dealing with audio and lyrics, we call this application “enhanced karaoke”, since it also involves musical instruments other than voice, and it also involves more interaction with the end-users.

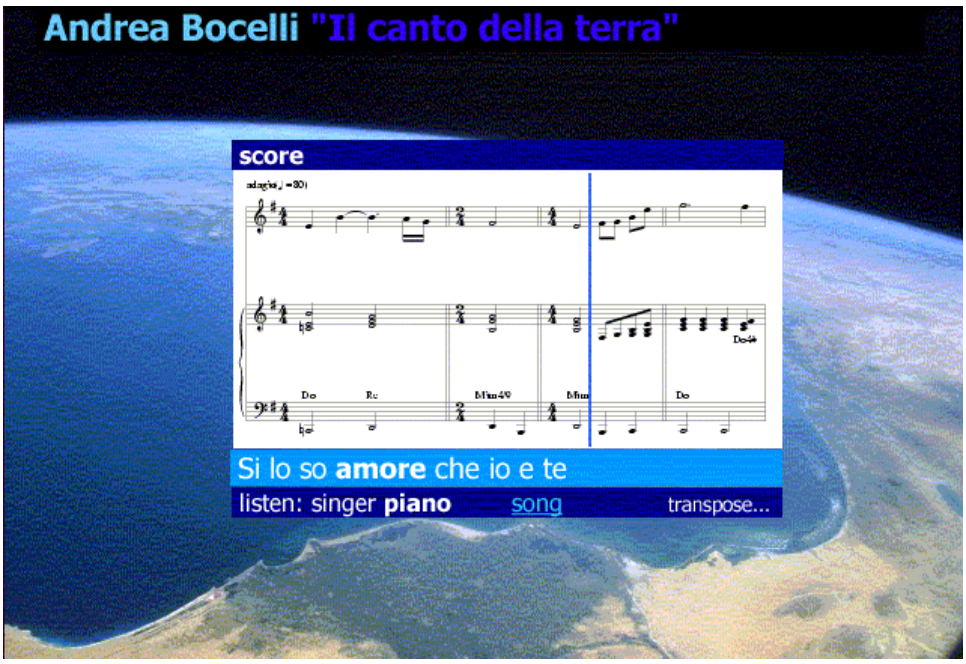
Involved objects and content

In this scenario, several objects are involved in relationship to one song. For what concerns audio, three stereo AAC objects may be used to encode the singer’s voice, a guitar and piano; a fourth object, e.g. an SA object, is used to synthesize in real-time the bass through access units carrying non time-stamped SASL commands (so the decoding time stamp of the access unit is used to synchronize the events). Four other main objects are present, a video accompanying the song (the video may report the scene of the opera or the simple clip of the song), a text containing lyrics, music notation content and a scene description including graphic shapes acting as selection buttons and interaction sensors and routings.

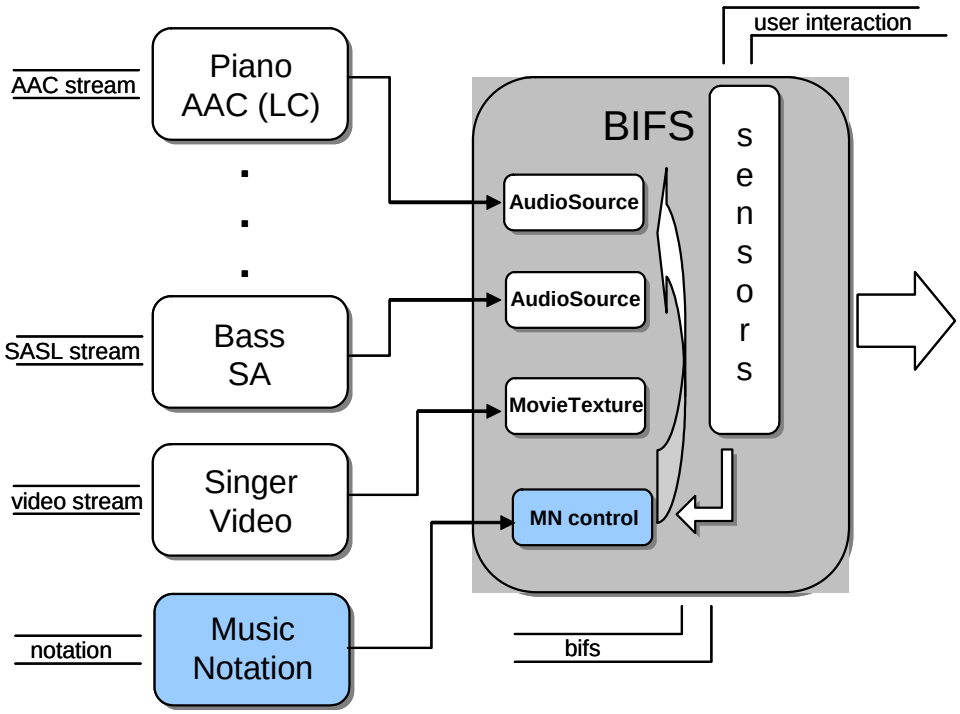
Scene description and interaction

The scene description allows the display of the accompanying video (e.g. a singer), and it contains some icons to be used for the selection of the different instruments and voices and of the text. By default all the AAC and SA objects are active and the text display is not active. Finally, the music notation decoder is active and displays a score with all the parts. If the user does not click on one or more of the icons, a line moves over the visualized score in synchronization with the musical content. A scene mock-up with just voice an one audio track is shown in the following picture.

If the user clicks on the voice icon (“singer”, in the picture), the video is minimized and text is displayed synchronized with the music, so that a normal karaoke application is enabled. If the user clicks on one or more of the instrument icons, whatever is the state of the text display, that instrument is muted and the music notation decoder highlights (either by changing color or by a new window) the part that has been muted, always with a line /cursor moving on it synchronously with the rest of the sounds to highlight which music notation symbol has to be played. If two or more sound parts are muted a similar behavior occurs for all of them. Whenever the user clicks again on the corresponding icon the previous situation is restored in relationship with that particular part or text. The following picture is another mock-up of the same application scenario.



In addition, the user has a button allowing him to transpose (music transposition), since users have not always the same voice as the original singer, or dispose of an instrument slightly different from the original, for example a tenor saxophone instead of an alto saxophone, in which case he has also to see the score part transposed. To this purpose, all music objects must be transposed (not difficult if those objects are SA ones, some processing may be required for AAC in AudioFX or different tracks may be available), and the MN object must be transposed too.
The main blocks for this scenario are summarized in the following block diagram.



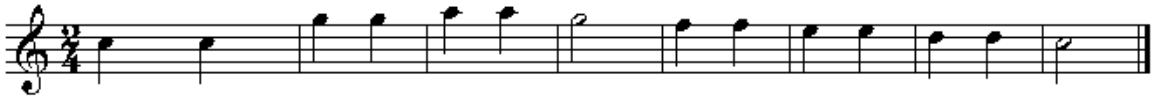
Main required MPEG tools
The tools already available in MPEG required for this simple application scenario are:

- MPEG-4 AAC (LC for instance)
- MPEG-4 SA
- MPEG-4 Video (e.g. Simple Profile)
- MPEG-4 BIFS (any profile supporting 2D graphics, timed text and multiple sounds)

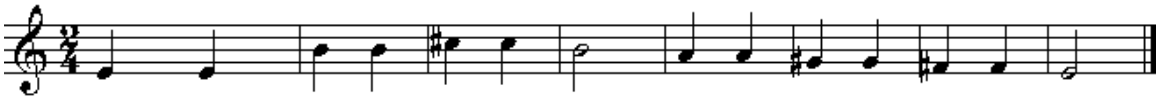
Main requirements for Music Notation
First of all it is necessary to have a proper music notation format with its normative decoding process. This means having a format carrying the music notation and in addition a different chunk offering the possibility to describe proper synchronization between score "events" and times (score alignment with live performance). The music notation format must support all the necessary functionality to correctly display music notation information (different fonts, different justification as a function of note duration and constant, different size, colors, etc.. as in the text), particularly in synchronization with other media in the scene. Further, in music it is needed to add also justification parameters since the spacing among symbols has musical meaning.
The MN object must be able to represent in a synthetic manner music objects. Music objects are essentially notes, but also more synthetic objects such as trills, arpeggios, portandos, and so on which should not be represented as the notes actually played, but as single objects.

Further, interaction is necessary between the user and the media. This means having the music notation decoder interfaced to the scene with one (or maybe two) nodes with suitable fields able to receive necessary information to drive the decoder and at the same time delivering information from the decoder to other fields of relevance. In this particular example a field is required containing on/off state for each of the parts (to be possibly routed to AudioSource nodes or to an AudioSwitch node for the audio object switching).
As mentioned above, in some cases, the user may further wish to use a different instrument, thus the music notation need to be transposed. This will change the visual representation. Transposition MN node shall be able to transpose correctly a part, or an instrument part, according to the rules currently in use in music notation. For example, a correct transposition of a very simple extract.

Original displayed score:



Transposed displayed score:

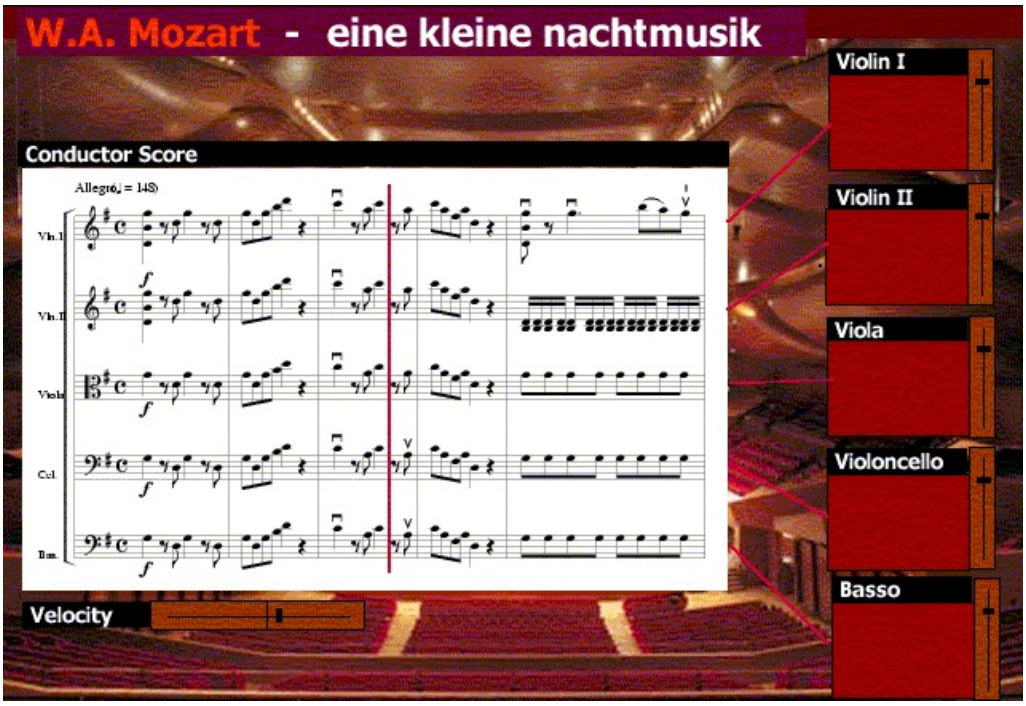


2.4.2 Scenario 2. Interactive music "tutor"

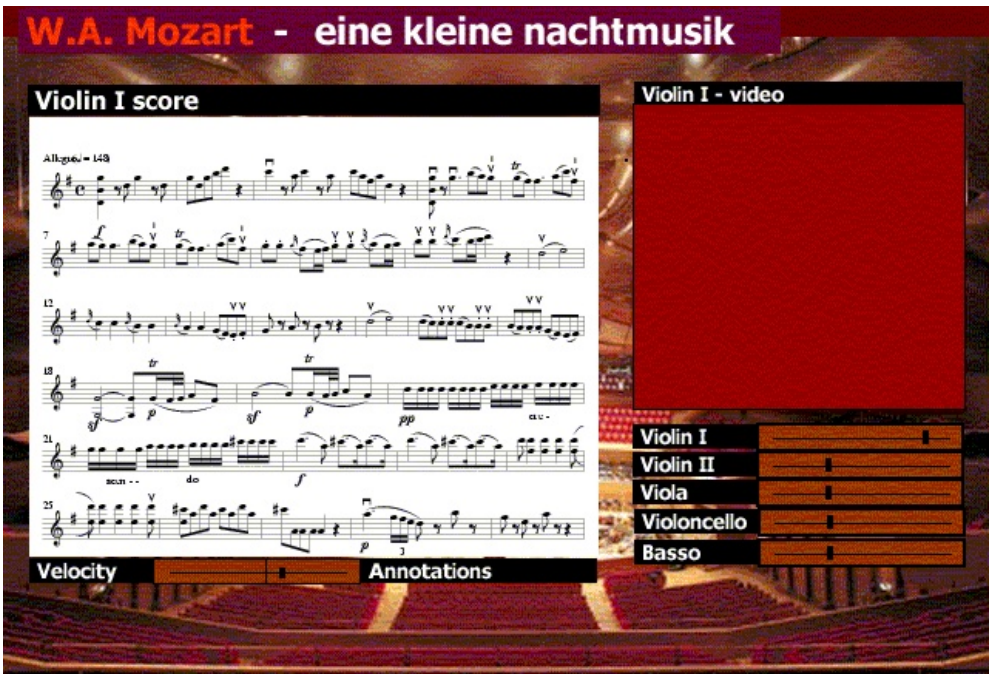
In this scenario the purpose is having the user look for a training category in an archive of courseware and subsequently download multimedia interactive content matching the search criteria. In this case the user has the possibility to access multimedia sequences containing a required feature and interactively work with this content to learn and compare his/her ability by this content. The way in which an eventual live performance may be compared or measured against the downloaded interactive presentation (e.g. scoring) is outside the scope of the standardization and it is related to any individual application that may automate the evaluation process based on the available content. Nevertheless, a suitable model must be available to describe musical notation in a way to allow with the required precision this comparison.

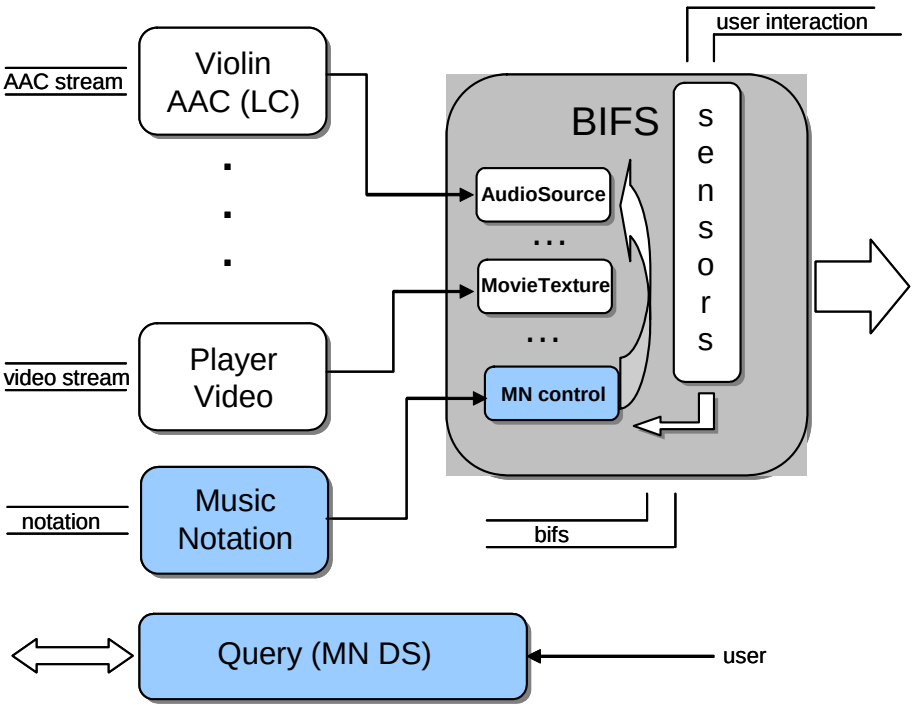
Involved objects and content
In this scenario the user tool has access to a possibly wide library containing performances of music pieces for educational purposes. All the available material is annotated by suitable descriptors according to the MPEG-7 standard with additional features related to notation. The available material is encoded in multimedia files composed by several objects each. Concerning audio, each instrument that is supposed to have a main role in the performance is encoded as an independent audio object (e.g. AAC LC). Each of these instruments also has a close-up video recording. Audio is passing through a processing node offering the possibility to slow or accelerate the performance (factor 0.5 to 1.5) without altering the pitch. Finally music notation is available, and a scene description for content composition and user interaction is provided.

Query, scene description and interaction
The user has the possibility to query the database for the particular skill to exercise he/she is looking for. For instance chromatic scales on the violin, or staccatos on the piano, and so on. The search will provide him links to material available for download and view examples and possibly exercise the desired features.
Each scene description allows the display of the accompanying videos (close-up of the instruments), and it contains different control icons to be used -- e.g. to affect the speed of the performance or completely mute parts in the performance. By default all the AAC objects are active and the different close-up videos are available as small resolution movies (the video may show the movements of the hands of a reference player, or the gesture of the conductor to be followed, etc.). Finally, the music notation decoder is active and displays a score with all the parts. If the user does not click on one or more of the video pictures, a line moves over the visualized score in synchronization with the musical content, like in the following picture.



If the user clicks on the picture of the instrument he/she is interested in, the video is magnified for that instrument, the music notation is reformatted to present only the selected part and not the main score with all the parts anymore (always with a line moving on it synchronously with the rest of the sounds), the sound of that instrument is enhanced in intensity over other instruments. Other parts may also be muted or reduced in volume. The user also has the possibility to control the execution speed of the performance through suitable control icons (like sliders). In this case the sound is slowed down by an AudioFX node implementing a speed_change effect and the music notation tool behaves accordingly maintaining synchronization with the audio. The user is also usually interested in repeating some sections, marking them and restarting from the marked point several times (sound can be buffered, but this is a feature possibly related to a non normative use of the normative file; indeed a precise synchronization between score and other media, especially sound, is a strong requirement). Whenever the user clicks again on the corresponding video the previous situation is restored in relationship with that particular part and instrument. The following pictures show another view (magnified instrument) of this scenario and a block diagram summarizing the main blocks involved.





Main required MPEG tools

The tools already available in MPEG required for this simple application scenario are:

- MPEG-4 AAC (LC for instance)
- MPEG-4 Video (scalable, FGS)
- MPEG-4 BIFS (2D visual profile)
- MPEG-4 AudioBIFS (AudioFX node for processing)
- MPEG-7 DSs

Main requirements for Music Notation

As in the previous example, first of all it is necessary to have a flexible music notation format with its normative decoding process. This format must support all the necessary functionality to correctly display music notation information, particularly in synchronization with other media in the scene. This means, as

said earlier, having a format carrying the music notation and in addition a different chunk offering the possibility to describe proper synchronization between score "events" and times. More than this, a suitable "subset" of the music notation functionality should be "visible" at the MPEG-7 description layer, in order to allow a query on relevant aspects of a score that may be worth searching for. The main requirements are:

- Production of main score and parts from the same synchronized music notation model
- Definition of sections
- Stop and play
- Accelerate and decelerate the execution rate
- Score alignment with live performance (similar to the first case)
- The MN object must be able to represent in a synthetic manner music objects. Music objects are essentially notes, but also more synthetic objects such as trills, arpeggios, portandos, and so on which should not be represented as the notes actually played, but as single objects.
- Description of musical content: it shall include all elements needed to describe music notation at a high level, including details of execution such as dynamics (staccato, pizzicato, legato, slurs, fingering, bowing…), rhythmic and meter details (tempo, rhythm, time signature…).
- Query by example: it shall be possible to select a segment of music notation to search for similar music, at the notation level.

Interaction is necessary between the user and the downloaded media. This means having the music notation decoder interfaced to the scene with one or more nodes with suitable fields able to receive necessary information to drive the decoder and at the same time delivering information from the decoder to other fields of relevance. In this second example a field is required containing on/off state for each of the parts (to be possibly routed to AudioSource nodes or to an AudioMix node for the audio object enhancement). In addition a field is necessary to control the speed of the score line display. To summarize the main interaction requirements:

- Showing selected single part with needed visualization parameters.
- Showing main score with required visualization parameters.
- Transposing the selected parts to be played with a different instrument
- Selecting parts to be muted or reduced in volume
- Accelerating and decelerating the execution rate for the music notation
- Adding some execution annotations such as fingering, bowing etc. that are typically added to the music notation during the rehearsal and during music studying.

2.4.3 Other Application Scenarios of Symbolic Music Representation

Currently many new music-related applications are strongly attracting the market attention. Among the several products currently available in the market that present some form of integration of symbolic music representation and multimedia it is possible to mention tools in the areas of:

- music education (notation tools integrating multimedia, multimedia electronic lecterns for classrooms and lecture halls, music education via i-TV),
- music management in libraries (music tools integrating multimedia for navigation and for synchronization),
- interactive entertainment, such as karaoke-like (mainly synchronization between sound, text and symbolic information),
- piano keyboards with symbolic music representation and audiovisual capabilities,
- multimedia content integrated with music notation on music instruments, mobiles, PDAs, etc.

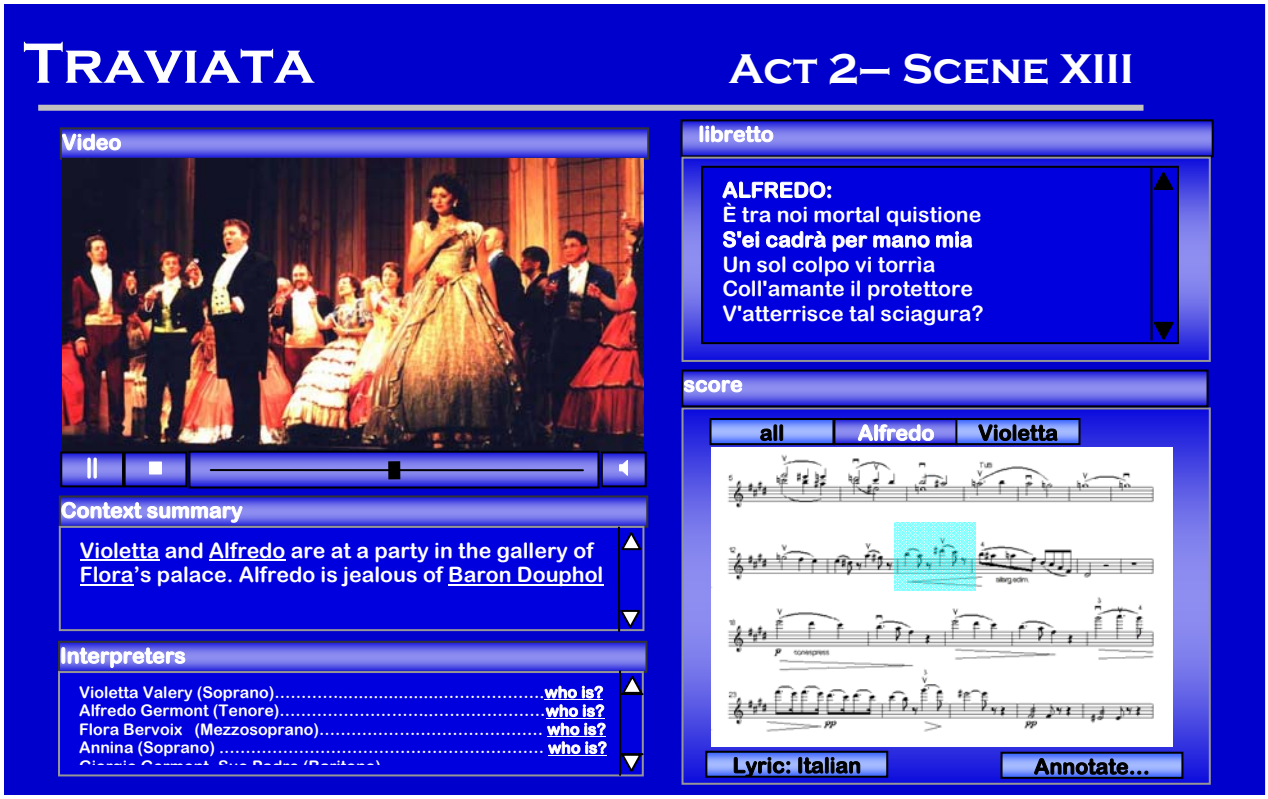
Some of these innovative applications integrating synchronizations, 3D virtual reality, are already in the day by day reality or in very good and powerful tools (VOYETRA, SMARTSCORE, PLAYPRO, PIANOTUTOR, etc.) or prototypes coming from research and development projects (OPENDRAMA, WEDELMUSIC, MOODS, IMUTUS, etc.). All of these applications of interactive multimedia music representation take advantage of MPEG technology for the standard multimedia integration and for the possibility of distributing content in a variety of forms in a completely integrated manner, see for example. The distribution of multimedia music content including music notation can be used for a large variety of applications of education, edutainment, entertainment, infotainment, etc., to reach the mass market on i-TV, mobiles, media centers, and personal computers,

2.4.4 Scenario: Multimedia Music for Entertainment and Infotainment

One of the most interesting applications and products is the multimedia music publication, for supporting entertainment, interactive music tutorials, multimedia music management in libraries, archives, and theatres. Examples of these applications are WEDELMUSIC, OPENDRAMA, etc. In these cases, we can see the integration of some symbolic music representation with audio, lyrics, annotations, different visual and audible renderings (different notations, parts and main scores, ancient and common western notation), synchronization with audiovisual such as video, audio, images, etc.

For example, in the case of a multimedia music application for the fruition of lyric operas, like in OPENDRAMA, the application can show the music score synchronized with a video of the opera or with a virtual reality version allowing the user to (see next Fig):

- follow the music by reading score of the instruments or of the singers and eventually selecting the lyric in a different language.
- read the libretto and its eventual translation,
- read and understand the plot which in some cases is really intricate,
- add personal annotations on the music, to mark a particular passage



A mock-up of a tool exploiting MPEG SMR for entertainment, edutainment, suitable for theatres, i-TV, archives, PC, etc., coded in MPEG-4 with SMR support.

An application like this may be used inside a theater from a palm top installed in the stalls, allowing the user to follow what is happening on the stage or can be used on i-TV or PC for entertainment and edutainment.

2.4.5 Scenario For SMR for Educational and Edutainment

Another interesting application is the exploitation of MPEG SMR to support educational or edutainment applications. Examples of these applications are MUSICALIS, IMUTUS, WEDELMUSIC with authoring support. In these cases, we have the integration of SMR with audio, lyrics, annotations, different semantics for visual and audible renderings (different notations, parts and main scores, ancient and common western notation), synchronization with audiovisual such as video, audio, images, and also the 3D virtual rendering of the position of the hands and body (posture and/or gesture), or of the scene with the actors, etc. (see next Fig).

Music education and edutainment are the largest markets for music representation. In the same way that a text can be synchronized to images and sounds in commercial DVDs, audio, video and music representation can be synchronized as well, for various (or separate) instruments and voices, and with various models or rendering aspects (common western music notation, tablature, Neumes, Braille, etc. Then, it is strongly desirable to have the possibility of distributing content in different forms related to the same piece of music. For music education and courseware construction music notation needs to be integrated with multimedia capabilities. About the 80% of young students are involved in the music study; this implies that a large part of them is capable of reading some music representation and playing an instrument in some measure.

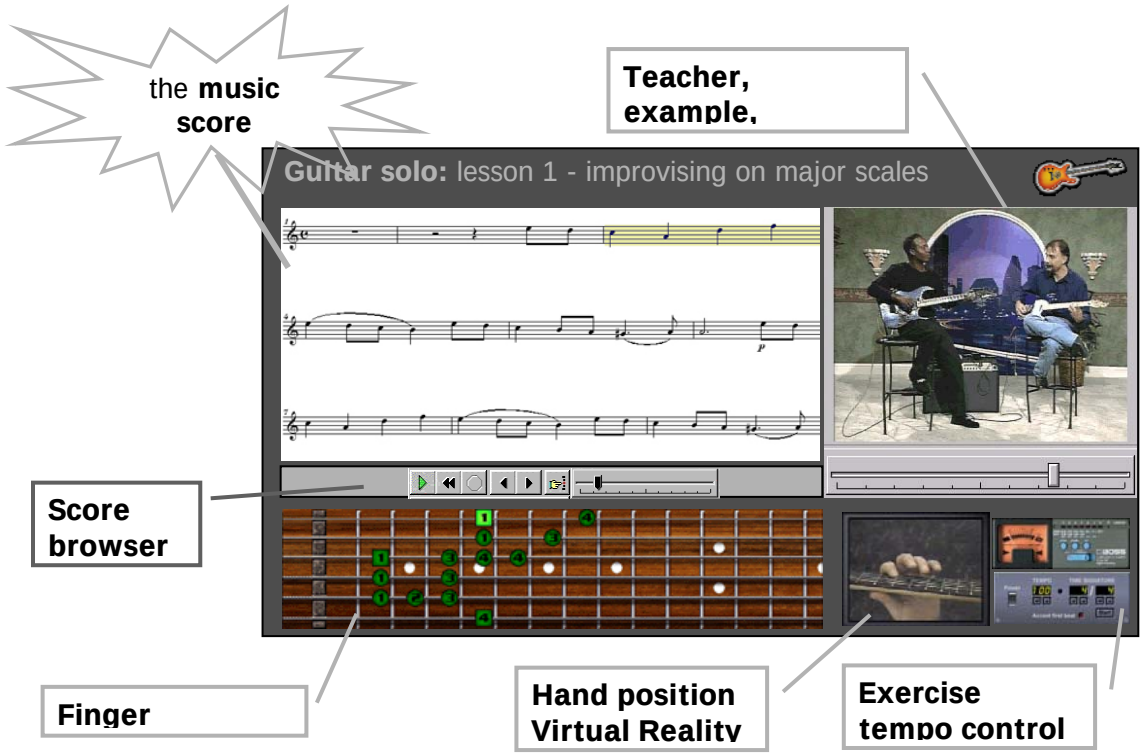


Fig.3 – A tool mock-up for exploiting MPEG SMR for education, edutainment, suitable for assisted learning, self-learning, on i-TV, tablet PC, etc., coded in MPEG-4 with SMR support.

Music courseware needs to have music representation integrated with video, images, documents, animation, audio, video, etc. In addition, a music course may present some exercises which require special music notation symbols (e.g., given by the teacher for the student, user defined) and/or audio processing (e.g., play training assessment, theory training assessment, etc.). For these reasons, for music education and courseware production the users on the client (music content usage) and sever (content generation) sides have a visual representation of the musical model (with full access to logical aspects and additional semantics information for visual and audio rendering), with full capabilities of music manipulation and a support for the synchronization and establishing relationships with other media.. Therefore, system and model have to provide the possibility of:

- navigation among music representation features and multimedia aspects and content. The user could browse a music representation and follow hyperlinks associated with music symbols/entities. The hyperlinked content could be external (a URL) or internal (an MPEG-4 audio visual element). The content associated with a music symbol could be a video/image/text explaining a difficult passage, its semantics, the aim of the composer, etc.;
- music editing, transposing, playing, formatting, piano reduction, etc.; this implies the dynamic alteration of aspects in the logical, audio and visual domains of symbolic music;
- selecting one lyric from the collection of multilingual lyrics for the same music representation; lyrics with music notation synchronized with video or audio (music karaoke, vocal and ear training);
- synchronization of audio visual events with the music representation elements, so as to show in a visual manner the correct point in the music is played, score following, etc. This requires the presentation of the visual aspects of music synchronized with the audio possibly allowing for changing the velocity of

performance, the adjustment of different playing parameters, etc. Executing the audio of the teacher or a famous musician with the visualization of the preferred music representation;
- showing the video of the teachers or showing the 3D rendering of the correct gesture of the hands while playing the instrument synchronously with the music notation;
- playing along: automatic accompaniment of the user by having the computer play selected voices/parts. This can be performed by using real audio recorded, generating MIDI, or by a more sophisticated music notation rendering. The system may be capable of automatically following the pupil performance and to product comments and corrections on-line or off-line;
- specifically formatting music for rendering on different devices, or with different resolution, or for rendering music representation in a different formats and on different devices: Braille, tablature, guitar tablatures, spoken or talked music, Korean music, etc. For example, talked music is useful when non experts work on music notation editing the device may produce a verbal description of the music context such as: "*measure 45, quarter chord of a D and a C flat in arpeggio, ….*". This rendering model is also very useful for impaired, young students, dyslexic people, etc.

More interactive and complex players may integrate additional capabilities to support:
- *annotation*, for adding teacher and pupil annotations that may be textual or in music notation, or audiovisual. This also includes capabilities of versioning and the possibilities of attaching complex information and rendering hints to SMR elements;
- *recognition* of notes played by an instrument or voice, using pitch recognition or other technologies. This means that the audio events have to be compared with what is expected by the notation conventions at the logical/symbolic level, and that an appropriate visualization is then created;
- *assessment* with respect to semantic annotations (how to execute a certain symbol, how to assess a certain execution of a music notation piece from the pupil, etc.). As noticed above, a good level of assessment needs a precise modeling of quality for the main domains of music performance, such as pitch, rhythm and sound. Particularly, sound quality, as well as detailed prescriptions on physical gesture should be made available;
- *cooperative work* on music notation for classes, for orchestral work on music lecterns (e.g., Tablet PCs), for managing rehearsals, etc.;
- *customization of SMR rendering* according to the needs of the individual (for example, selecting symbols to be verbally described or printed, or played, selecting a simpler model for the visualization or playing, for example without beaming, without ornaments, etc.);

3 New Applications Requirements

3.1 General Requirements on Music Notation Model

The following three general requirements have been recognized as essential ones for the music notation model and format to comply with Music Notation functionality in a multimedia context, regardless the specific application scenario:

- **GR1, scalability**: it has to be possible to write simple music notation pieces files without spending time and token to describe all the context. At the same time the model has to support the writing of complex structure of music pieces with multivoices, multistaff, etc., according to the applications identified. This will make it suitable for simple devices and applications such as those supposed for mobile equipment.
- **GR2, interoperability**: this is a typical feature of MPEG applications; interoperability is required among i-TV, Mobile Devices, PDAs, Tablet PCs, electronic musical instruments, special hardware, cars, etc.
- **GR3, extensibility**: to have the possibility to create new symbols with their related rules or code for formatting (conversion to the visual domain and from this in graphical domain) and executing/playing (conversion to the audio domain).

3.2 Current and new application requirements of Music Notation in Multimedia

- **R1**, annotation of music notation with audiovisual content: of the students on the lesson, audio visual annotations of the teacher.
- **R2**, decode of music notation with different decoders to obtain different views of the same music representation (main score, parts, Braille, Gregorian, spoken, etc.). This can be done by using additional information or automatic formatting engines. This is very useful to work on different devices (piano keyboards, PDAs, mobiles, i-tv, sheet, screen, etc.), etc. The music notation integrated with audiovisual could be streamed with video, audio, .. on mobiles, etc. (see MPEG4 model)
- **R3**, supporting query by content integrating audiovisual and music notation, usage of music notation model for verifying the match with MPEG7 descriptors of music that presently are very simple, increasing the expressivity of MPEG7 music descriptors.
- **R4**, distribution of music notation with multimedia aspects integrated in, by using a suitable DRM model based on MPEG21
- **R5**, executing music notation synchronously with audio visual content, animation, for educational purpose, for entertainment, etc., see Karaoke, etc.

During the development of the work of the AHG, and thus the identification of the requirements the following requirements were identified and in a second phase mainly included in the above mentioned:

- navigation on music notation and multimedia aspects of MPEG, as annotations or story path built in the content. The navigation is mainly the activity related to the exploitation of annotation and relationships of music notation with multimedia, thus see R1 and R5.
- showing/playing audio visual content related to Music Notation symbols and context. This activity is mainly related to the synchronous execution of music notation with audio. See R5.
- authoring music notation integrated with multimedia aspects for creating the content used by the above mentioned requirements, synchronization, annotation, etc.. This requirement is mainly included in the description of R1 and R5
- joining the music notation with the MPEG SAOL. A good model for audio performance modeling see R5.

Five of the above nine operational requirements have been discussed and clarified during the AHG on Music Notation Requirements (see m10067) and the Music Network Open Workshop in Leeds, September 2003. Here below is a summary of those collected information; for the complete documents for each of those, including references, hyperlinks and other details, please refer to the AHG web page at http://www.dsi.unifi.it/~nesi/mpeg/ahg-mn-65-66.html

3.2.1 R1 -- Annotation of music notation with audiovisual content

Basically, annotation means the ability to add at the music notation some symbols or text or on this case some audiovisual information. The annotation can be personal or one can use annotations made by others

(search, acquire, download, merge…). Some annotations and related functionalities are mainly oriented for the authoring while others are of interest only for the final user.

Annotations can be:

- traditional: accidentals, fingering, bowing, adding of attention marks, adding of caution accidentals, dynamics, tempo marking, breath marks for wind instruments, special marks such as pedals for piano, harp, tablature for guitar, cue notes…
- hand-produced annotation (lines, frames, specific shapes…),
- multimedia content as audiovisual (audio, video, still images, animations, actions to the external world and on physical devices).
- Accessible to queries and search engines on the basis of metadata and description of content

Audio visual, multimedia annotations should describe in an enhanced manner the music notation:

- showing details of execution (for example images of hands or arms of the teacher, video of movements),
- giving audio details (detail of execution for a particular passage),
- showing manuscript details (original from the composer),
- showing images of different editions
- playing the sound of a piece of music
- showing the semantics of a given symbol (in music notation, text or other visual language)
- visualizing an image related to the music piece or movement
- presenting a multimedia document explaining the music piece or the critical review, or the historical period or similarity with other music pieces, etc.
- linking the music notation to some web page or any other http link
- associating/adding/manipulating metadata to the annotation list of a music piece (author name, purpose, version, …)

Annotations should be

- flexible to be joined almost to all music notation symbols at different levels from elementary symbols such as notes or rests to complex structures such as voices, parts, staves systems.
- shareable among a group of users (in orchestras or ensembles), a community of users (music schools or conservatories), or a wide community of users (Internet…).
- Join with other lists of annotations of other people
- Maintained separate from the original score,
- Identified by their own set of metadata and identification.
- Independent by the original media from which they have to maintain separate.
- Protected to allow their distribution among specific users or groups of them.

R1 -- Rational for its introduction

Annotation of music scores is a traditional activity. Teachers, musicians (professionals or not), pupils, all musicians using scores put currently annotations on these at least to help them playing music (fingering, cue notes, attention marks, breath marks…), or to play it in a slightly different way, what is used to be called "interpretation" (adding bowing, dynamics, tempo…).

By using multimedia systems, annotation systems can be enlarged to other kind of annotations: video or audio (for example showing details of execution), still images (for example manuscripts) and also graphical representation of music scores or handwritten music manuscripts. A lot of interesting new applications could be considered in the field of music education. Several examples of tools present music notation annotated with audiovisual content. Those applications are not flexible enough. For example, they are not supported by any authoring tools and thus the content for the lessons is limited to that provided by the tool producer.

Distribution of traditional scores can also be enhanced by annotations.

R1 -- Basic Functionalities involved

- Annotation production (if you have the rights to do that) and usage:
- Protection: to define the rights which can be applied to the production and usage of annotations. Right to use, merge, save, load, modify, etc.
- Adding audio visual annotations

- Adding traditional annotation
- Distribution: ability to publish a set of annotations in order to make it shareable among a group of users.
- Save present loaded annotation on the disk
- load and apply some annotation referred to the music piece that you have from the disk or from some remote site. Once loaded they are applied to the music notation
- Merging annotations if you have the rights
- Listing and navigating in the annotations
- Searching: searching (on Internet) for annotations related to a particular score using the metadata or the description of the annotations
- Comparison: comparison of different annotations
- Browsing: using hyperlinks, navigate in associated media (audio, video, still images).
- Manipulating the annotation list
- Adding and changing the metadata of the annotation list

R1 -- Applications that needs this requirement

In fact, all applications related to music notation, from publishing to playing on an instrument, to karaoke and live performance, needs more or less this requirement.
The following applications will particularly benefit of annotations:

- Educational software
- Educational : production of material

R1 -- Description of potential users

All musicians, professionals or not, must be able to use annotations and produce them their selves. Publishers will be able to enhance publication of traditional score with multimedia content.
Teachers, professionals will be able to produce annotated scores and use them for education purpose or for the purpose of ensemble execution

Use of annotations will be particularly useful in, but not limited to:

- Music schools and conservatories
- Music player (performers, instrumentalists, students etc)
- Music analysis (analysis on performance practice, for example)
- Music preservation
- Music libraries
    - Public libraries (with public access)
    - Musical institutions libraries (with private or restricted access) (example : Teatro alla Scala, Opera de Paris, many orchestras…)

R1 -- Customer satisfaction and need

The ability to annotate scores is a basic need for musicians playing an instrument. It's particularly fundamental in groups (orchestras, bands, chamber music ensembles…). It's also fundamental in the education field. The customer must benefit at least from the same range of annotations he is accustomed with, and which are reviewed at beginning of section 4.2.1. The customer could greatly benefit of applications offered by audio, video, or image inclusion in annotations, or of applications of links to external material. This way, he could for example hear at the execution of a particular passage from a great master, or from a teacher. He can also view details on how it's done: details on fingering, on position of hands, arm…
NB: the musical example given by the teacher is, with annotation on scores, one of the most prevalent mean for music teaching used in instrument playing education.

R1 -- Presence in current applications

Current applications

Digital music scores distribution in the educational field or to the wide music community is small compared to distribution of traditional music scores (sheet music or music scores). A part of it is based on specialized music software (Finale, Sibelius), and their on-line free viewers Scorch or SmartViewer, or based on Acrobat (PDF format). Some digital music score distribution is based on proprietary software (MusicNotes). These digital scores can be viewed with professional tools (Finale or Sibelius) when the format permit it, but with much more expensiveness.

Available operations

Operations available in these specialized viewers are very limited: view, print… For some of these advanced operations such as hearing (in MIDI) or even transposing are available (Scorch, SmartViewer). Some of these includes DRM (SheetMusicNow). None of these includes the ability to annotate. Analysing Scorch's transposition, it has been detected that it frequently did not restore all the enharmonics correctly on a translation "round trip" (to a new key, back to the original key). That was one of my motivations for discussing the "base 40" system for intervallic invariance at the MusicNetwork meeting. MIDI key numbers are not sufficiently articulate to support transposition for works with more than three sharps/flats, or unusual modulations.

Annotation of music is theoretically possible when distributed in PDF. Nevertheless, a number of specific annotations are difficult or impossible to achieve: beaming, fingering, cue notes…. Annotations are difficultly semantically linked to musical events and figures.
Annotation is actually not possible with specialized software tools (Finale or Sibelius), according to the requirements expressed at the beginning of section 4.2.1. The following requirements are not achieved: separate storage, ability to merge, ability to share.

WEDELMUSIC (www.wedelmusic.org) and OPENDRAMA include mechanisms for managing annotations. Annotations are shareable, dependant on the semantics of the score, and can be extended to links.

R1 -- Dependency with other requirements

Ability to annotate is strongly dependent on a DRM system which should include:

- The right to annotate
- The ability for the user to protect his own annotations (to define the level of protection for these annotations).

Ability to annotate is also dependent on the distribution of the music notation itself (requirement R4).

– Executing music notation synchronous with audio visual content
– Navigation on music and multimedia aspects

R1 -- Conflicts with other requirements

None is foreseen for now.

3.2.2 R2 -- Decode of music representation to obtain different views of one music piece

The music representation related to a music piece contains information associated to different domains (visual & graphical, logical, audio). Using this information, different views of a music piece can be generated:

- The single parts of the instruments playing, notated using Common Western Music Notation
- The main score containing all the instruments or a subset (in CWMN)
- Other notations as the simplified notation for kids or the Gregorian notation (single part)
- Braille notation for blind music (single part)
- Analytical representations (Schenkerian notation)
- The audio encoding in MIDI/SASL of a single part or a subset of the parts
- Spoken music generation (single part)

The view of one or more instrument can be also transposed, this is needed when a part written for a certain instrument has to be played by another instrument.
The generation of the notation view could be automated using algorithms for the proper formatting (justification) and for automatic line/page breaking.
A constraint to be considered regards the limitations of the device used to look at the score, that can be a PDA, a piano keyboard, an i-TV or a Braille pad all with different dimensions and resolutions, for this reason the user should control which symbols are to be viewed (to avoid the display of unnecessary symbols) and should control the zooming factor. To be noted that when zooming in, the display of the score has to be adapted in a way to avoid the use of scroll bars.
Some views, such as the parts or main score of a piece written in CWMN, the Braille and the analytical representation must also be printable, but the formatting for printing is not necessarily the same as the one for visualizing on a computer screen.

R2 -- Rational for its introduction

This is a basic requirement it is mandatory to give the user the possibility to see the score and interact with it.

R2 -- Basic Functionalities involved

- Formatting
- Justification
- Line Breaking, Page Breaking
- Paging
- Visualization on the screen
- Printing Music Score, Printing Braille Music, Printing on the screen
- Speaking Music

R2 -- Applications that needs this requirement

All applications of music notation

R2 -- Description of potential users

Any person interested in music notation - composers, performers, academics, teachers, students. The technology developed for this user group may well turn out to be useful in other areas - for example the development of specialised notations for scientific or technical applications involving the use of time.

R2 -- Customer satisfaction and need

Most of the users of current applications are not completely **satisfied with** the capabilities of **existing** tools in representing music according to different views. In most cases, the production of other views requires much additional work on the part of the user. Power users will probably never be satisfied with the results of automated part extraction, but the enablement of simplified notations (without the overhead of current applications) may allow near perfect part extraction to be achieved at that level.

R2 -- Presence in current applications

In some applications this is possible. For example passing from tablature to music notation, passing from main score to parts, etc.

R2 -- Dependency with other requirements

no

R2 -- Conflicts with other requirements

The synthetic audio can be considered also a “view” that can be generated from the music representation**,** joining the music notation with the MPEG SAOL, deals also with this.

3.2.3 R3 -- Supporting query by content integrating audiovisual and music notation.

This requirement is related to query by content by using combination of traditional query (melody, rhythm, key tonality, ambitious, instrumentation) with audio-visual content description (low-level and high-level description).

Evaluation of similarity, based on statistical approaches and audio-visual descriptors, combined with traditional query could be achieved this way, thus enhancing the traditional cataloguing paradigm. Description of music notation content should be extensible to traditional cataloguing schemes, in order to encompass current practices in libraries, and interoperable with audio and visual descriptions.

R3 -- Rational for its introduction

Description of content has been employed for hundred of years by musicians searching for music. Traditionally, description of content involves the following:

- Theme, incipit. This describes the melodic characteristics of music, together with rhythmic characteristics.
- Tonality (key). This is so frequent that tonality is often used in the name of music works, even when not useful (for example, “Beethoven’s Symphony No. 5 in C Minor, op. 67”)
- Ambitus (that is, description of the pitch range for each voice, as lower pitch – higher pitch). This is particularly useful for choirs and singers.
- Instrumentation, that is, description of the musical instruments needed to play the score.

These descriptions of content are very often placed in front of the music score, in order to give a resume or an overview of the content to the potential user.

These descriptions are not useful in a search by similarity. Defining similarity of melodic descriptors, for example, has been something a stalemate since many years, even if interesting researches has been made on this point (Pierre-Yves Roland). The same for all other kind of traditional descriptors like rhythm.
These kind of traditional descriptions could be enhanced by the possibilities offered by audio descriptors, for which current research has shown interesting results these last years (Peeters and alii). Audio descriptors can define low-level description of audio (brightness, loudness, dynamic of sound), as well as high-level descriptors (timbre – instrument).
Last researches in this area have shown that similarity matrixes based on audio can be established for measuring differences between pieces of music. Combination of traditional descriptors with audio descriptors and similarity matrixes combined with statistical analysis of audio and/or music notation should be an interesting way to achieve new results in search by content.

R3 -- Basic Functionalities involved

Main functionalities are those which are implemented mainly in Libraries:

- Cataloguing, classification
- Comparaison
- Audio indexing
- Querying

R3 -- Applications that needs this requirement

In fact, all applications related to music notation, from publishing to playing an instrument, needs more or less this requirement.
The following applications will particularly benefit of query by content:

- Educational software (classification, search engines)
- Educational: production of material (compilations, inclusion of examples…)
- Music libraries applications (cataloguing, classification, search engines…)
- Procession content production by using technical aspects

R3 -- Description of potential users

All musicians, professionals or not, must be able to search for music by using content-based queries.

Uses of queries will be particularly useful in, but not limited to:

- Music libraries:
    - Public libraries (with public access)
    - Musical institutions libraries (with private or restricted access) (example : Teatro alla Scala, Opera de Paris, many orchestras…)
- Music schools and conservatories
- Educational institutions, outside of music schools (they own very frequently musical activities such as choirs or music ensembles[*])
- Actors of musical institutions (choirs, ensembles, concert organizers…), professional or amateur.
- Music player, professional or not.
- Content Integrators
- Content producers
- Content distributors and broadcasters

R3 -- Customer satisfaction and need

Traditional queries (instrumental material…) as well as traditional cataloguing paradigm must be supported, and the description of content cannot replace the overview of the content itself.
In advanced libraries such as in Music department of Bibliotheque Nationale de France, the need of an "analytical indexation" based on analysis of content is claimed, at the audio level as well as at the logical level.
In the meanwhile new applications of query by content (such as "query by humming") are emerging, which can be integrated in a slight manner to traditional tools.
A slight shift in cataloguing paradigm, induced by the use of content-based description, analytical indexing and similarity evaluation should produce new uses and new practices in Libraries.

[*] In France, there is in principle at least one choir in every educational institution (primary school, secondary school…).

For example the archive in Milan located in the Mattei Foundation has the possibility to create complex query with music notation and multimedia aspects.

R3 -- Presence in current applications

Current applications

Digital music scores distribution does not include actually any form of query by content, neither based on the audio part nor on the logical part (music notation). Some test sites are offering query by humming applications (Fraunhofer).
Digital libraries and archives doesn't currently support any kind of query by content, even if administrators often claim for the need of "analytical indexation" based on content (Elizabeth Giuliani, Music department of Bibliotheque Nationale de France).

Features of query by content has been requested several times by users, see for example the results of HAMRMONICA project and the current activities of the MUSICNETWORK in the area of library and the results of the last IAMIC meeting (International Association of Music Information Centers), and IAML (International Association of Music Libraries) meetings.

Available operations

WEDELMUSIC includes a mechanism for query by content, on the basis of melodic, rhythmic, harmonic and tonality descriptors, available in Digital Libraries Catalogs.
Another solution is that coming out from OPENDRAMA in which the music notation is integrated with 3D virtual reality of the scene in the operas. In this case the query may integrate/union 3D virtual reality descriptors of the scene with the music aspects.

R3 -- Dependency with other requirements

Query by content is dependent on the following: Music distribution

R3 -- Conflicts with other requirements

None is foreseen for now.

3.2.4 R4 -- Distribution of music notation with multimedia aspects

Music notation is the expression of the intention of the composer of a music piece, independently of its rendering, performance and execution. The music notation makes it possible to different users - human users and also machine users, computer systems or even analog systems like piano rolls – to render it in a particular performance. The music notation makes also possible a graphic rendering of the music piece - commonly known as "score". In this view, producers, rights owners and also potential users of music notation are not the same as producers, rights owners and potential users of a musical performance.

Distribution of a multimedia content including music notation and their corresponding relationships and associated files and content. This content as well as the relationships have to be protected by a DRM model for example based on MPEG-21.

The associated content to be synchronized with music notation can be:

- Lyrics of some music notation, multilingual or not
- One or more audio files
- Videos, documents, images, animations, etc.
- Scene description streams

R4 -- Rational for its introduction

When music notation is introduced in a multimedia context such as in educational or entertainment applications it has to be protected.

Uses and DRM for music notation are quite different of uses and DRM of audio performances of the same musical work. As stated at the beginning of this section, users and producers of music notation are different from users and producers of music performance.

- Rights owner of a music notation item could be different from rights owners of a performance of the work, and also different from the rights owner of the graphic rendering of the piece.

- Uses of music notation are different from uses of a performance of a music piece, since music notation is intended to be used for performance (mainly for live performance, but also for machine-based performance) or as well to allow a more comprehensive listening of a performance.

In addition, the music notation cannot be delivered as raw images since the users of music notation may need to interact to the music notation, to change it, etc., for example for piano reduction, formatting, arranging, annotating, printing in different formats, transposing for a different instrument, etc. In the case of distribution with associated media content not all of thes tranformations will be equally important, but of particular interest can be different formatting and arranging, for instance in association to source muting, source switching, etc.; all these cases require indeed a "vectorial" notation representation and are not possible with raw images.

R4 -- Basic Functionalities involved

Since the requirements related to the distribution of cross media content are well know in MPEG we concentrate on the aspects related to Music Notation and their relationship with cross media.

The basic functionalities related to the distribution of music notation are

- **Protection of music notation**, specification of what can be done on the music notation, the price, how many times, (?) can be done, etc. The operations that can be done on Music Notation are reported in the list of functionalities and in the following. According to the publishers, each operation should be traced and counted. For example, the printing could be allowed for three times, while the playback for an infinite number of times, etc. The usage of the Music Notation may implies the use of other related files such as Lyrics in several languages, and cross media content related to the piece to which music notation is associated.
- **Download from Internet music** and related multimedia content
- **Sharing music notation** and related multimedia files among friends
- **Fruition and manipulation** of music notation integrated with multimedia content for educational or entertainment. The end users should have the possibility to view, copy, play, print, transpose, etc. their music notation content as they do presently with the paper versions but using instead digital media storage devices such as CDs, DVDs or even streams. In addition, the symbolic version of music notation allows them to perform a set of innovative operations that open the market for the publishers and media distributors in general; on the other hand, the distribution may result more complex since the DRM model has to be sophisticated with respect to that need for the distribution of sheet music.
- **Simple multilayer music** content distribution can be possible at a lower level of interaction than the one allowed by complete fruition and manipulation above. As simple audio digitally distributed on CDs has been extended by audio + video + lyrics + very simple interaction possibilities on current DVDs, the same it is desirable to add different layers for an even more complete distribution of music-related content with some more, but still limited, interaction possibilities, like the MIDI file and the music score with simple source switching, custom visual appearance formatting, and so on.

Some of the functionality of music notation is mainly useful for the production other for the end users. In any case, this distinction is not always precisely defined. In any case, the activities performed at level of end user are also needed at level of production.

Production:

- analysis, arrangement, editing, cataloguing, cloning,
- converting from image, converting from other formats,
- describing
- synchronization, transcoding

End Users:

- play along, play MIDI, rendering, score following,
- annotating, comparison, correcting, analysis, formatting, justification, navigation, page breaking,
- printing Braille, printing, generating spoken music, screen visualization, paging,
- transcoding
- transposition

DRM aspects of Music Notation
From what I see we need to specify with some DRM Rule language

- **usage of the music notation**: printing, playing, producing Braille, producing spoken music, etc. limiting the number of times. In this case, the usage can be done in some different views: main score and parts, for each part. In all these cases, the user could be interested in producing the music notation with a different view, formatting music, justifying music, line breaking music, substituting the visual aspect with some other from a list of possible (tablature, different fonts, Chinese notation, hand position, etc.) several views: a copy of the screen or a printing page.
- **modification of the music notation**: transposition, annotation, arrangement, addition of some symbols, etc., limiting with these activities with some parameters
- **excerpts and copy**, to copy the whole or part of the MN, limiting the number times, or the size in the case of excepts. The resolution can be also considered in this case.
- **Synchronization**: to join some music notation with other media
- **link definition and navigation**: to establish some links or to exploit some links to navigate on these.
- **processing:** music analysis, music comparison for comparison and query by example, etc.
- **export**: to export the music notation in some other format, for example in MIDI, etc..

R4 -- Applications that needs this requirement

The applications that need of this requirement are those of distribution. They are mainly related to MPEG-21 for the packaging aspects and with the format for modeling music notation and associated cross media content.

- Music distribution for home and mobile systems
- Sheet Music Distribution
- Educational content distribution towards: PC (on line, DC, DVD, etc.), PDA, i-TV, etc. for distance learning

More details on music education can be obtained at the following links in which there are several list of requirements for the creation of cross media content for music education. Some of these have been created as the standard activities that the students have to perform in US music schools from up to K12:

- Music Educator: http://www.musiceducator.org/modules.php?name=Downloads&d_op=MostPopular
- Technology Institute for Music Educators, http://www.ti-me.org/
- http://teams.lacoe.edu/documentation/places/art.html#music

R4 -- Description of potential users

The potential final users of the above applications can be:

- Music students and teachers
- Music lovers
- Music professionals
- Musicians in general. They are about the 12% of the European Population.
- "singing is the activity practiced by 19.8% of European citizens, and playing a musical instrument is an activity practiced by 12.5% of European citizens".

  …

From Europeans' participation in Cultural activities, A Eurobarometer Survey Carried out at the request of the European Commission, by EUROSTAT, April 2002, resulted evident that the 61,3 % of people in Europe listen Music at least every day, while the 20% several times a week. There is then a potential strong interest in delivering to these classes of users music notation content as a significant added value to other media.

The other potential benefit holders of applications capable of protecting music notation and delivering it with one or more other media types are (for a larger list see the www site of the AHG):

- Music Notation Publishers
- Music Notation Distributors, sheet music on demand on-line and off-line

- Music archives, libraries: 25000 in the world, 5000 in Europe
- Music Education company

R4 -- Customer satisfaction and need

The customer satisfaction has to be considered at two levels.

The corporate users are presently not satisfied of the distribution tools for music notation integrated with multimedia content. There are only few tools and these are not based on standard formats. This problem is very relevant for publishers, and the archives, that would invest on stable technologies.

The final users are presently non satisfied by the distribution possibilities and by the marketing in general of the sheet music. In addition, the distribution of educational content based on music notation is practically marginal and limited exactly because of the lack of tools for modeling multimedia music.

R4 -- Presence in current applications

Applications of multimedia content distribution including also Music Notation

- WEDELMUSIC, covering B2B and B2C in an integrated manner and supporting a DRM and certification with a Certification of Authority.
- P2P applications such as Napster, Morpheus, KazaA Lite 2.1.0 K++, eMule 0.28b, WinMX 3.31, Bearshare, LimeWire, Gnutella, iMesh, Ohaha, FileFunnel, Rapigator, Newtella, Music City, etc., are de-facto distributing also music scores or at least music events in several formats: Finale, Sibelius, MIDI, etc.

Typical Applications (see in the reference list)

- Sheet Music Distribution tools
- Content Distribution tools used also for Sheet distribution
- Music Education & Distribution.
- Music Editors
- Catalog and collection management

R4 -- Dependency with other requirements

It is related to the following requirements:

-- Protection aspects, if this will become an independent requirement ????
-- Interoperability of DRM and music notation format
-- Extensibility of the music notation format
-- authoring and production of content including music notation
-- query on the base of music notation. This is a preliminary action for the distribution.

R4 -- Conflicts with other requirements

No conflict is foreseen for the moment.

3.2.5 R5 -- Executing Music Notation synchronously with audio visual content

For *execution* of music notation we intend the interpretation of the music notation in terms of symbolic format for the production of a corresponding audio or for the visualization of the notation in some manner.
The other media that could be synchronized with music notation are:
Audio music:

- produced converting symbolic music notation in MIDI with some simple or better rendering models. In this case, the MIDI production can be directly performed synchronously with the visualization of music notation
- recorded in any audio format. In this case, a synchronization is needed

Video with audio music or not

- The video can report the hands of the instrument players, to demonstrate for educational and edutainment activities. This application was experimented by Yamaha and other high cost piano builders on their pianos some years ago by using DVDs.
- The video can report an opera while the related score (as piano reduction or specific part is shown). An example of this application can be seen in OPENDRAMA Project.
- The video can report images not directly related to the execution of the music notation
- In all cases, the music notation has to be synchronized with the video.

Text with audio music or not

- The text can show the words to be song, the lyric. This can be a sort of Karaoke for the singer and the music can be the music of the singer or a specific reduction
- The text can report some annotations that have to be shown during the execution, such as comments explaining the music or the context. These can be shown on the screen during a representation
- The text can reports some comments for the executors: accelerando, piano, crescendo, etc.
- The text have to be multilingual, such as multilingual DVDs…
- In all cases, the music notation has to be synchronized with the text. The synchronization has to be done at level of single syllable or words, phrases depending on the needs

Synthesized speech

- Generated by the text in several languages

Images and Graphics

- The image can be joined to a music notation symbol to be shown when that symbol is executed.
- 2D and 3D Graphics can be joined to a musical symbols to be shown when the symbol is executed
- Images and Graphics can also be synchronized with the music symbolic notation. This would allow full inter-linking functionality to navigate and render both the symbolic representation and images/graphics of the manuscripts, for example.

Animation

- The animation can be joined to a music notation symbol to be shown at the user when that symbol is executed.
- The animation can report some movements of the hands of the player for educational purpose. This application is very frequent in educational software such as MUSCIALIS. ACE, etc. It can also reports the actions of the actors such as in OPENDRAMA, etc.
- The animations can be also used for didactical purpose to teach a kid, etc.
- In all cases, the animation has to be synchronized with the music notation.

In all cases, a synchronization is necessary with the music notation. The synchronization can be done on the basis of the time line. On the other hand, one of the requirements of music notation execution is the possibility to change the execution rate. This functionality is very frequent for musicians since one of the parameter of music execution if the metronomic indication, that is a way to define the number of beats per minutes (considering also, that the time signature of the music piece is known).

This implies that the synchronization cannot be done at level of time, but has to be done at level of symbols or measures. Better if at level of symbols. In this way, any shrinking and stretching of the music execution time is replicated in the execution rate of the other media and vice versa.

Algorithms of shrinking and stretching for audio, video, etc., are already available.

R5 -- Rational for its introduction

When a music notation is included into a multimedia tool and integrated with other multimedia content it has to be synchronized.

A very interested market is growing on the related applications.

There are relevant applications in the fields of:

- Music education software tools, courseware
- Tools for distance learning of Music, virtual teacher, etc.
- Entertainment, Infotainment, edutainment, etc.

Specific devices:

- Portable and stable keyboards and pianos
- Fix Personal computers
- Distance learning on i-TV, by using MPEG4
- Tablet Computers for teaching and/or learning music
- Special hardware for recording studios
- Lecterns for classrooms, cooperative or not
- Monitors for people in theatres or cinemas

- Monitors for people on the aircraft
- Other wearable and portable devices

R5 -- Basic Functionalities involved

Since the requirements related to the presentation and execution of cross media content are well-known in MPEG we concentrate on the aspects related to Music Notation and their relationship with cross media.

The basic functionalities related to the execution of music notation synchronously with audio visual content are

- Synchronization of music notation symbols to audio, video, text and animation
- Execution of music notation symbols to audio, video, text and animation
- Authoring to define relationships among music notation symbols to audio, video, text and animation
- Streaming of the whole synchronized media
- Protecting of the whole synchronized media
- Copy of the whole synchronized media

R5 -- Applications that needs this requirement

The applications that need of this requirement are those of fruition of music notation synchronized with audio visual. They are mainly related to MPEG21 for the packaging aspects and with the MPEG4 for synchronization of music notation with cross media content.

- Music education software tools,
- Tools for distance learning of Music, virtual teacher, etc.
- Entertainment, Infotainment, edutainment, etc.

They can be on specific devices:

- Portable and stable keyboards and pianos
- Fix Personal computers
- Distance learning on i-TV, by using MPEG4
- Tablet Computers for teaching and/or learning music
- Special hardware for recording studios
- Lecterns for classrooms, cooperative or not
- Monitors for people in theatres or cinemas
- Monitors for people on the aircraft

More details on music education can be obtained on the following links in which there are several list of requirements for the creation of cross media content for music education. Some of these have been created as the standard activities that the students have to perform in US music schools from up to K12:

- Music Educator: http://www.musiceducator.org/modules.php?name=Downloads&d_op=MostPopular
- Technology Institute for Music Educators, http://www.ti-me.org/
- http://teams.lacoe.edu/documentation/places/art.html#music

R5 --Description of potential users

The potential users of the corresponding technology that see music notation synchronized with audio visual are:

- Music and multimedia content Publishers, distributors, integrators
- Electronic instrument builders: pianos, lecterns for singers, etc.
- Music Education tools companies
- Distance learning companies
- Digital video Broadcasters, satellite TV, STB builders
- Recording studios
- theatres and cinemas
- music schools and conservatories, music universities
- music and multimedia archives, libraries, and information centers. etc.
- Sheet Music distributors on line

For some lists see the home page of the AHG and the R4 requirement.
The potential final users of the above applications can be:

- Music students and teachers
- Music lovers
- Music professionals
- Musicians in general. They are about the 12% of the European Population.
- singing is the activity practiced by 19.8% of European citizens, and playing a musical instrument is an activity practiced by 12.5% of European citizens

R5 --Customer satisfaction and need

The customer satisfaction has to be considered at two levels.

The corporate users are presently not satisfied of the simple music notation since it is not interactive. This has been highlighted several times in past activities of analysis such as HARMONICA and also in requirement analysis performed by EC projects such as MOODS, CANTATE, WEDELMUSIC, IMUTUS. Those analyses have been done by using a large set of experts coming from different sectors.

The same results have been obtained by the MUSICNETWORK analysis of the state of the art. This problem is very relevant for educational institutions and companies especially those that work on distance learning.

The final users are presently non satisfied by the multimedia capabilities of the music notation. All the media players in the market do not support the music notation and in that field there is still the lack of standard and a very high competition among few actors of the market.

R5 --Presence in current applications

Applications of music notation that integrate multimedia content are:

- WEDELMUSIC, WEDELMUSIC: Web Delivering of Music Scores,
- MUSICALIS; www.musicalis.com music editor and educational tools
- www.onlineconservatory.com
- OPENDRAMA, music tools for opera show, http://www.opendrama.com
- IMUTUS, Interactive music tuition system, distance learning tools
- http://www.rainmusic.com/pianomusic/lessons.htm
- http://www.musiclearning.com/ Online interactive Music Education
- more or less all the other educational music tools that you can recovered in the main page of the AHG

R5 -- Dependency with other requirements

It is related to the following requirements:

-- authoring and production of content including music notation
-- supporting query by content integrating audiovisual and music notation
-- distribution of music notation with multimedia aspects integrated in

R5 --Conflicts with other requirements

No conflict is foreseen for the moment.

4 Music Notation Problems vs new Applications

This section deals with the most important modeling problems adopting the perspective of the emerging applications of music notation. As stated by several authors in the past, the modeling of music notation presents several problems: (i) the intrinsic complexity of formalizing both *music notation and the relationships* among music symbols; (ii) the necessity of providing different visualizations/views of the same music as it occurs with main score and parts; (iii) the complexity of any automatic formatting of music symbols; (iv) the necessity of adding new symbols in order to expand the music notation model and make it fit for modern music and users' needs; (v) the necessity of presenting new functionalities for multimedia applications; (vi) the necessity of producing high quality sheet music in terms of tiny adjustments of symbols so as to avoid collisions and produce very clearly written and well recognizable sheet music at a first glance, (vii) the formatting of music around the page end.

The modeling of all possible relationships among notation symbols is a complex task. As the number of symbols grows, the number of relationships among them grows with a ration more than proportional. The syntax and semantics of music are strictly related and cannot simply be modeled by using only *non-visual* (relationships among symbols) or *visual* aspects of music (positions with respect to the staff). The modeling of music symbol **relationships** is not enough to describe the music notation: for example, a *diminuendo* starting between two notes (and ending on the second) depicts the instant in which it has to be applied. The description of music as a collection of **graphical symbols** on the page regardless of their relationships relies too much on the visualization device (automatic reformatting is really complex). In most commercial music editors, the music notation symbols are simply managed as graphical symbols that can be placed on the screen.

The automatic generation of the main score from the different parts and the extraction of parts from a main score are complex tasks for real music, since the conductor's main score is typically formatted in a different way from the musicians' scores. Parts are obtained by using compressed versions of music notation symbols, and also present several *instrumental, personal symbols*, which are typically omitted in the main score. The main score may present some specific textual notes or symbols for the conductor. The main score may show more than one part on a unique staff, in this case, parts are treated as distinct voices/layers. Other indications are added in the parts, so as to make the coordination among musicians easier; for instance, the so-called grace/cue notes. The main score and parts present different justifications, line breaking and page breaking. In some music editors, even the insertion of a measure in a part results in a huge manual work of reformatting in order to provide the replication of the performed changes on the main score, or vice versa. In Dannenberg (1990), a solution was given, suggesting the adoption of a unique model and distinct views for showing music on the basis of each specific need and context.

The positioning of music elements on the score presents many critical conditions. A music editor should help the users to produce well-formatted music scores. In most of the music editors, this task is addressed by considering only the visual representation of music symbols and not their relationships. As above highlighted, some visual aspects of music scores cannot be formalized in terms of symbol relationships. For this reason, the visual arrangement of music notation symbols makes the description of the music scores complete. Music notation can be considered a visual language with its own rules. On this ground, the modeling of visual rules for formatting music is a requisite step to the completion of the model (Bellini, Della Santa and Nesi, 2001).

Typically, there is the need of adding new symbols for stating specific instrument aspects, and/or expanding the music notation model to take in modern music as well. The addition of new symbols means to add fonts, relationships among other symbols, execution rules and formatting rules. For example, the addition of a symbol modeling a special figure (such as a cloud of notes) implies a deep change in the music notation model; whereas the addition of a new marker for notes can be performed by adding positioning rules and fonts, since the relationship with the note is already established. Several other symbols must be added to fulfill the needs of all instruments of a large orchestra (specific accents, jargon symbols, heel, toe, bridge, etc.). Most of them put on different visual representations depending on the instrument they are used for.

At a first glance, the new applications related to Internet distribution of music scores seem to be not too much different from the music editors currently on the market. The applications of the next few years will be

mainly based on: (i) cooperative work on music notation, (ii) interactivity of music notation in the respect of the owner rights, (iii) the availability of different integrated information associated with the music piece, (iv) music as support for tuition systems. In all these instances, music has to be independent from any visualization support features, formatting features and resolution aspects. Then, the following functionalities have to be available.

- a clear distinction between the music notation model and the visualization rules: automatic reformatting on the account of the user's needs, etc.
- music notation model has to be abstract enough to allow interactivity with music at the symbolic level: adding and deleting symbols, changing symbols features, selective and unlinear undo, versioning, etc.
- music model has to integrate several aspects: audio, symbolic music notation, images of music score, documents, video, animations, www pages, multilingual lyric, etc.
- mechanisms for distributing music by using shops, libraries, conservatories, etc., as local distributors.
- a refined protection model for Digital Right Management.

The first two problems may be managed by formal models supported by a separate engine for the automatic formatting of music according to precise rules. Ideally, this work is infeasible (Bellini, Della Santa, Nesi, 2001), but positive and promising compromises can be obtained. The new multimedia applications are bringing music in a new era. Simple audio files or sheet music are undergoing important changes in order to be included in more complex multimedia objects.

5 Music Notation Models and Languages

In the last years many music notation models and languages have been developed, among them MIDI, Finale, Score, MusiXTEX, NIFF, MusicXML, and SMDL are the most referenced. In this section, a short discussion and comparison among these formats is reported. The comparison does not claim to be exhaustive; it is focussed on the aspects relevant for the adoption of these languages and models in the new applications.

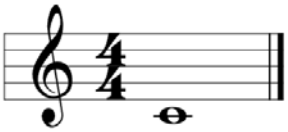
Music notation formats were developed for different purposes, at the beginning one of the main objectives was to design a format suitable for the production of sheet music, since manual typesetting of sheet music is a time/money consuming task. At this thread belong formats as DARMS, Score, MusicXTeX, Finale and all the formats of music notation programs that are used to print music. With the coming of electronic music and electronic music devices the design of a protocol allowing devices interconnection was needed. Since devices were produced from different manufacturers a common agreement on the communication protocol was necessary, on this the MIDI standard was born in early 80s. MIDI was supported by music notation program producers for the benefits that it was carrying, such as the direct listening of a music piece using MIDI devices and the direct input using a MIDI keyboard. The use of Standard Midi Files for interchange of musical information was a consequence and it was not felt as a problem for the low expressivity of MIDI.

At the very beginning of the 90s the optical music recognition tools began to be not only research prototypes. These tools were produced from companies that didn't want to be bound to a specific music notation program, on this the NIFF format was born with the support of the most important music notation companies.

After the beginning of the Information Technology era when digital information exchange has become a daily matter, interoperability are felt as necessary not only for telecommunication devices and data-interchange through XML applications has become quite common, many XML languages for music notation interchange has been developed. On this thread MusicXML is the most referenced example.

In recent years we are seeing a growing interest in multimedia applications, on this thread SMDL was a pioneer taking into account also multimedia aspects in music modelling and the EC WEDELMUSIC project was the first to produce a working prototype.

In the following the models and formats are briefly described, for some of them an example of textual (XML) encoding of a simple "do" whole note is reported.



5.1 MIDI, Music Instrument Digital Interface

(http://www.midi.org/)

MIDI is a sound oriented language. It is unsuitable for modeling the relationships among symbols and for coding professional music scores: in MIDI accents, ornaments, slurs, etc. are missing or very difficult to recognize. MIDI is the most widespread coding format among those that can be found on Internet since (i) it can be easily generated by keyboards, (ii) it has been adopted by the electronic music industry for representing sounds in a computers form, (iii) MIDI music can be played and manipulated with different synthesizers. A very large percentage of these files have been generated without the intention of recovering the music score and therefore they are do not fit as music scores due to the presence of several small rests or notes. There is also a limited space for customizing the format. This space has been used for defining several versions of enhanced MIDI formats (SelfridgeField, 1997). In the literature several extensions of MIDI have been proposed but no one has really got the large diffusion to substitute the classical MIDI format. MIDI format is mainly used as an interchange format, on the other hand, its capability in modeling music notation is very limited. Most of the music editors are capable of loading and saving music notation in MIDI.

5.2 SCORE

(http://www.scoremus.com/, http://ace.acadiau.ca/score/LINKS3.HTM)

SCORE is probably the most widespread notation editor among publishers for the high quality and professional Postscript printing (Smith, 1997). In SCORE each symbol can be precisely arranged on the page according the engravers' needs. Complex symbols can be produced by using graphic elements on the sheet music. Several minor accompanying symbols of the note can be placed on the staff in any position and thus the relationships with the note are non-defined. This means that a movement of the note or its deletion does not influence the life of these symbols. SCORE presents no distinction between slur and ties and no automatic conversion of sequences of rests into generic rests. SCORE is a page-oriented editor, in the sense that music information is collected only related to the preparing page: the editor is capable of managing only a printable page at time. Since the music is created page by page, parts (each page a file), are very hard to automatically extract because the match is based on part numbering and the numbering of staffs may be different in successive pages (some parts can be missing, the numbering is not ordered from top to bottom). When symbols are managed in the middle of a line or page break, they are drawn with two distinct graphic lines. This makes complex the extraction of parts and the reorganization of music measures along the main score and parts. The insertion of some measures in a page or the changing of the page dimensions may need the manual manipulation of all pages. This is totally unacceptable for music which has to be distributed via Internet and which has to be viewable on several different devices for both size and resolution, with changes to be performed in real-time.

5.3 MusiXTEX

(http://icking-music-archive.org/software/indexmt6.html)

MusiXTEX is a set of macros for LaTEX and/or TEX for producing sheet music (Taupin, Mitchell, Egler, 1997), (Icking, 1997). The language is interesting since its structure is substantially symbolic while graphic commands can be added to provide precise positioning. The relationships among symbols depend on the order symbols appear in the code. The language is printer-oriented and thus it allows the placement of graphics symbols anywhere on the page.
Some simple rules for the insertion of symbols are available (definition of up and down of the stems for notes). With MusiXTEX specific rules for the visual organization of symbols on the page could be defined exploiting the power of LaTEX and TEX. Classification features could be implemented by using a similar technique. In MusiXTEX, the work has to be manually performed during the score coding. MusiXTEX does not support: (i) the automatic beaming (identification of the groups of notes to be beamed together in the measure), (ii) the automatic definition of stem direction of notes in beams, (iii) the automatic management of positioning for accents.

5.4 NIFF, Notation Interchange File Format

(http://www.musique.umontreal.ca/personnel/Belkin/NIFF.doc.html)

NIFF was developed with the perspective of defining an interchange format for music notation among music notation editing/publishing and scanning programs (NIFF6, 1995). Its present version if 6b. NIFF was derived from design rules of the Resource Interchange File Format (RIFF). NIFF design resulted from a quite large group of commercial music software developers, researchers, publishers and professionals. For this reason, the model was defined with the aim of supporting the most relevant aspects of several models. This represented a great limit in the expressiveness of the languages and models for describing exceptions. The main features of NIFF are: (i) a feature set based on SCORE, (ii) division of visual information in page and non-page layout information, (iii) extensible, flexible and compact design, and (iv) inclusion of MIDI data and format. Since 1995, the notation has not been improved. It is currently supported by LIME editor and few others. Since 1996, a kit for implementing a loading module for NIFF files is available.
NIFF language includes in a unique model both visual and logical aspects. This makes it difficult to deliver music regardless of the visualization details, which is a feature needed in cooperative applications. Relationships among symbols are defined via specific links (anchorage). They can be set according to default rules but specific actions can be performed for managing some changes in the editors. In NIFF, no support is provided for either versioning or cooperative work, since each logical element in the format cannot be univocally identified. NIFF presents a limited number of symbols but grants the possibility of defining new symbols.

NIFF and XML

Links: http://www.music-notation.info/en/niffml/niffml.html (NIFFML)
http://niffty.sourceforge.net/xml/index.html (NIFFXML)

Recently NIFFML, a XML version of NIFF has been presented by Gerard Castan and a similar attempt has been made by Jeff Thompson who developed an applet (Niffty) to visualize NIFF files in web pages. The most evident difference between the two encodings of NIFF in XML are the tag names, NIFFML from G. Castan uses the names of tags defined for NIFF (4 chars) while NIFFXML from Thomson uses a more long and understandable names, i.e. tag **`tmsl`** in NIFFML is tag **`TimeSlice`** in NIFFXML.

The following it is an example of NIFFXML from Jeff Thomson:

```
<NIFF xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
      xsi:noNamespaceSchemaLocation="NIFFML.xsd">
  <Setup>
    <NiffInformation niffVersion="6b" standardUnits="points" writingProgramType="engraving program"
                     absoluteUnitsPerStandardUnit="20"/>
    <PartsList>
      <PartDescription name="UnnamedPart" partID="0" midiChannel="0" abbreviation="UnnamedPart"
                       numberOfStaves="1"/>
    </PartsList>
  </Setup>
  <Data>
    <Page>
      <PageHeader>
        <Width value="10944"/>
        <Height value="14544"/>
      </PageHeader>
      <System>
        <SystemHeader>
          <Width value="10400"/>
          <AbsolutePlacement vertical="1800" horizontal="400"/>
          <Height value="400"/>
        </SystemHeader>
        <Staff>
          <StaffHeader>
            <Width value="10400"/>
            <AbsolutePlacement vertical="1800" horizontal="400"/>
            <AnchorOverride value="pghd"/>
            <Height value="400"/>
            <NumberOfStaffLines value="5"/>
          </StaffHeader>
          <TimeSlice type="measure-start" startTimeD="4" startTimeN="0">
            <AbsolutePlacement vertical="400" horizontal="0"/>
          </TimeSlice>
          <Clef shape="G clef" staffStep="2"/>
          <TimeSignature topNumber="4" bottomNumber="4"/>
          <TimeSlice type="event" startTimeD="4" startTimeN="0">
            <AbsolutePlacement vertical="400" horizontal="1020"/>
          </TimeSlice>
          <Stem>
            <Height value="0"/>
            <LogicalPlacement vertical="above"/>
          </Stem>
          <Notehead shape="whole" durationD="1" durationN="1" staffStep="-2">
            <PartID value="0"/>
          </Notehead>
          <TimeSlice type="event" startTimeD="4" startTimeN="4">
            <AbsolutePlacement vertical="400" horizontal="10400"/>
          </TimeSlice>
          <Barline type="thin line" extendsTo="bottom of bottom staff" numberOfStaves="1"/>
          <TimeSlice type="measure-start" startTimeD="4" startTimeN="5">
            <AbsolutePlacement vertical="400" horizontal="10400"/>
          </TimeSlice>
        </Staff>
      </System>
    </Page>
  </Data>
</NIFF>
```

5.5 SMDL, Standard Music Description Language

(http://xml.coverpages.org/smdlover.html)

SMDL is a mark-up language built on SGML and HyTime standards (SMDL10743, 1995). The aim was the definition of an interchange abstract model. SMDL model includes the following aspects: logical, gestural, visual and analytical. The logical aspect includes the music content (pitches, rhythms, dynamics, articulations etc.), that is, the abstract information common to both gestural and visual domains. The gestural domain describes specific performances of the logical domain. It includes dynamic symbols etc. and MIDI information. In SMDL, the visual domain describes the musical typographic details of scores (the symbol, placing, fonts etc.). The analytical domain consists of music analyses for classification purpose and a support for defining more sophisticated analysis in the logical and gestural parts. SMDL was analyzed in depth in the CANTATE project (CANTATE, 1994). The result of the analysis was as follows: SMDL cannot be used for modeling scores. It can only produce visually visible scores by using other formats such as FINALE, SCORE, NIFF, etc. or images of sheet music. SMDL cannot be used as a standard interchange format for the visual aspect but only for the logical aspects of music: a sort of container in which several different music formats and related information can be collected and organized (NIFF, MIDI, animations, textual descriptions, etc.) Currently there is no commercial SMDL software for editing music or producing music digital objects. In the CANTATE project an application of SMDL for the distribution of music for libraries was developed.
SMDL presents a neat distinction between the visual and the logical parts that should be modeled by different languages and formalisms. For this lack of integration among music aspects, SMDL cannot be used for either distributing interactive music via Internet or as a basis of cooperative work on music scores. More recently, several other mark-up languages for music modeling have been proposed. Unfortunately, they are weakened by the lack of managing slurs, accents, etc. The adoption of a structural model totally separate from the visualization issues makes the production of professional music a very complex task.

5.6 MusicXML

(http://www.musicxml.com/)

MusicXML is a XML format for music notation interchange developed by Recordare (Good, 2001). It is based on two other textual formats for music notation representation, the MuseData and the Humdrum formats (Selfridge-Field, 1997). It represents the music in a time-wise (parts nested within measures) or part-wise (measures nested within parts) manner, a XSLT allows the transformation from one to the other format. The format covers the western musical notation from 17th century onward, it is mainly oriented to describing the logical structure of music even if some graphical detail could be added. A plug-in for Finale allows it to load and save files using this format, the Sharpeye OMR application uses it as a interchange format with Finale, a plug-in for Sybelius allows to export in MusicXML thus allowing for much more accurate exchange from Sibelius to Finale than MIDI provides. Also PhotoScore Professional has also joined SharpEye Music Reader as a major OMR application exporting MusicXML. MusicXML includes a subset of the most commonly used formats. It can be useful for interchanging music notation. At the XML level, MusicXML is strongly based on Tag rather than on Attributes.

The following is the encoding in MusicXML of a simple “do” example:

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<!DOCTYPE score-partwise PUBLIC
    "-//Recordare//DTD MusicXML 1.0 Partwise//EN"
    "http://www.musicxml.org/dtds/partwise.dtd">
<score-partwise>
  <part-list>
    <score-part id="P1">
      <part-name>Music</part-name>
    </score-part>
  </part-list>
  <part id="P1">
    <measure number="1">
      <attributes>
        <divisions>1</divisions>
        <key>
          <fifths>0</fifths>
        </key>
        <time>
          <beats>4</beats>
          <beat-type>4</beat-type>
        </time>
        <clef>
          <sign>G</sign>
          <line>2</line>
        </clef>
      </attributes>
```

```
      <note>
        <pitch>
          <step>C</step>
          <octave>4</octave>
        </pitch>
        <duration>4</duration>
        <type>whole</type>
      </note>
    </measure>
  </part>
</score-partwise>
```

5.7 FINALE, Enigma format

(http://www.finalemusic.com/)

Music produced by FINALE program is coded in Enigma format. This format is partially documented. The editor and format are mainly oriented towards page preparation rather than defining relationships among symbols. Clear evidence of this is found in several symbols being non linked to figures and simply placed on the page, assuming any position without any bound to follow the sort of the note whenever moving or deleting. The format has been recently improved and allows the definition of some relationships among music notation symbols. This is not its internal philosophy but it is left to the users. The Finale model does not present a clear trace for voices that pass from one staff to the other in multi-staff parts, such as those for Piano, Harp and Organ. In several cases, the arrangement of music notation symbols is quite hard since the automatic mechanism for completing the measure is quite disturbing.

5.8 GUIDO format and tools

(http://www.informatik.tu-darmstadt.de/AFS/GUIDO/)

It is a textual format (human readable) for symbolic music description. The description is extremely compact and it seems to be optimized for direct user entry rather than to have an editor to produce it (like MusiXTex). A set of tools are present to transform this description to MIDI and to render it as PostScript or GIF, or to convert to it MIDI and FINALE files to GUIDO. GUIDO language has been designed in three layers, *Basic GUIDO* describes the basic musical symbols of western music notation (notes, rests, slurs, etc.) and their structure (staffs, voices, chords), *Advanced GUIDO* extends the Basic GUIDO to support exact score formatting and more sophisticated musical concepts, *Extended GUIDO* introduce concepts beyond conventional music notation. The tools currently support Basic GUIDO and Advanced GUIDO. However the specification of Advanced GUIDO and Extended GUIDO are not available. Automatic formatting rules are encoded in the renderer and automatic beaming are supported, however precise positioning and beaming can be specified forcing the position. The symbols supported in the Basic GUIDO are few and cover only the most important expressive indication (staccato, tenuto, accento, marcato…) and ornaments (trillo, mordente, gruppetto, tremolo, glissando) and no symbols for specific instrument are present (violin, piano, arpa, etc.). Moreover GUIDO lacks in the possibility to introduce new user defined symbols.

The simple “do” example can be encoded in GUIDO as:

```
[\clef<"treble">\meter<"4/4"> c/1 ]
```

Recently **GUIDO XML** language has been introduced. The music notation model behind GUIDO XML is the same of GUIDO. GUIDO XML improves interoperatibility with other XML languages and tools (i.e. XSLT, XQuery). Lossless converters between GUIDO and GUIDO XML has been produced, in this way tools developed for GUIDO can be used also with GUIDO XML.
The encoding of the “do” example in GUIDO XML is:

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE guido SYSTEM "guido.dtd">
<guido>
   &lt;segment&gt;
      <sequence>
         <clef s="&#x22;treble&#x22;"/>
         <meter s="&#x22;4/4&#x22;"/>
         <note name="c" octave="1" duration="1/1"/>
      </sequence>
   &lt;/segment&gt;
</guido>
```

5.9 MEI, Music Encoding Initiative

(http://dl.lib.virginia.edu/bin/dtd/mei/)

The MEI DTD is modeled on the one created by the Text Encoding Initiative (TEI). The primary goal of TEI was the creation of a comprehensive yet extensible standard for the encoding and transmission of textual documents in electronic form. The MEI shares the same goal for music documents.
No tools are currently present to aid the production of MEI xml documents, exists only a tool for the production of mup files from xml documents. Mup files can then be used to generate printable postscript files.
MEI has rich structuring constructs and metadata fields, since it has been developed for library use.

The simple “do” example is encoded in the following way:

```
<?xml version="1.0"?>
<!DOCTYPE mei SYSTEM "mei.dtd">
<mei>
   <meihead>
      <meiid>do-example</meiid>
      <filedesc>
         <titlestmt>
            <title>DO Example</title>
         </titlestmt>
      </filedesc>
   </meihead>
   <music>
      <body>
         <mdiv type="sample">
            <score>
               <scoredef meter.count="4" meter.unit="4" clef.shape="G" clef.line="2">
                  <staffdef id="s1" octave.default="4" />
               </scoredef>
               <bar n="1">
                  <staff def="s1">
                     <note pname="c" dur="1"/>
                  </staff>
               </bar>
            </score>
         </mdiv>
      </body>
   </music>
</mei>
```

5.10 MML, Music Markup Language

(http://www.musicmarkup.info)

Music Markup Language (MML) is an XML-based language for describing music objects and events. MML consists of several layered modules that could be called upon. Music is defined as a function of time and frequency and thus the Time and Frequency modules are the core modules. Additional modules that can be “plugged” on the core ones are:

- *Organization Module*: for describing metadata and other functions such as playlists and albums;
- *General Module*: focusing on marking repetitions in music pieces;
- *Texture Module*: focusing on music production, specifically intensity, harmonics and envelope;
- *Effects Module*: describes effects that are applied to the Music Core
- *Performance Module*: describes performance related objects and processes, such as Controls (breath, foot, hand), band/orchestral related elements, and singing related events;
- *Control Module*: describes events typically used in electronically triggered environments, such as the start and stopping of music events, volume, sound sources, and channel controls in multi-channel environments.
- *Notation Module*: this module attempts to provide a framework that allows the description of music objects and events in any notational language, not only CWN (Common Western Notation);
- *Lyrics Module*: provides a framework for synchronizing lyrics with music events;
- *MIDI Module*: describes interfacing MML and MIDI

- *Synthesizer Module*: describes a virtual abstract and universal synthesizer that could be mapped to manufactured real synthesizers from different manufacturers.

Applications supporting MML are currently the following:

- A perl script can be used to transform a MML file into a MIDI file.
- MML was used in a synthesized singing voice project.
- Students projects rendered MML in 2-D graphic displays.

The simple “do” example is encoded in MML in the following way:

```
<?xml version="1.0" ?>
<mml xmlns="http://www.musicmarkup.info/dtd">
  <head>
    <title>A Test...</title>
    <meta http-equiv="Content-Type"
          content="text/mml; charset=iso-8859-1" />
    <meta name="description" content="Test" />
    <meta name="composer" content="MusicNetwork" date="2004" />
  </head>
  <song>
    <div noteset="4" note="4">
      <notation clef="C">
      </notation>
      <bar barid="treb-1">4C:1</bar>
    </div>
  </song>
</mml>
```

To be noted that basic pitch/duration of notes it is not encoded in XML but as a <octave><pitch><duration> character sequence i.e. “4C:1” represents a 4th octave C whole note. Also synchronous notes (i.e. chords) are not represented using XML but by enclosing them in “[“, “]” i.e. [4C:2 4A:2] is a chord of two half notes C and A in the 4th octave.

5.11 MusiXML

(http://www.music-notation.info/en/musixml/MusiXML.html)

MusiXML is a XML application to model music notation proposed by Gerd Castan, one main goal is to store each data only once instead of holding them consistent, as in the case of main score/parts representations, thus separating content from style.

Main score and single parts have different instances of page, system and staff, but they share parts of <content>. The solution proposed is to store <content> in a separate place, the logical domain, and to refer to it from the visual domain <filter> section. Thus the work containing both the main score and the single parts is structured like:

```
<work>
  <body>
      … content …
  </body>
  <filter>
    <extract>
       (rendering information for main score with reference to content)
    </extract>
    <extract>
       (rendering information for a single part with reference to content)
    </extract>
  </filter>
</work>
```

The <filter> section contains <extract> elements defining each different printout. The <extract> section contains declarative instructions on how to process <content>.
Currently no tools are available to produce MusiXML files.

5.12 Lilypond

(http://lilypond.org)

Lilypond is a scalable and extensible compiler for producing high-quality music notation. It is available under the GNU license. The program produces a PostScript or Pdf file and takes as input a text file encoding a logical representation of the music. The system is partially implemented in Scheme and includes the GUILE scheme interpreter which allows users to override and extend some functionalities. The language of the input format can describe simple music expressions such as sequences of notes that can be combined to form more complex expressions such as chords and voices. The language provides keywords for controlling the combinations. For instance, sequences of notes can be inversed, or music expressions can be given a name in order to reduce the size of large combinations. Once the input file is defined, the program parses it, interprets it for visual representation and formats it for producing an image sheet. The separations between the different steps makes possible the representation of the same piece in different forms.

The main advantage of Lilypond is the high-quality of the music notation that can be automatically outputed. Several spacing rules and formatting procedures are taken into consideration which makes the final sheet more intuitively readable. The different domains of music notation that are concerned by the project (logical, visual and graphic domains) are considered independently which authorizes for instance different notations for one musical description. Several music formats (MusicXML, MIDI, Finale's ETF, ABC) can be partially converted as input for the system. The language is scalable and extensible.

However, the program has no user interface which makes it difficult to handle for non-expert users. It offers bad interactivity with music at symbolic level: once one event is modified, all the image sheet has to be re-compiled. Last, it is not clear if new musical objects can be defined by considering their visual or syntactic relationship with the other objects.

The simple "do" example is encoded as:

```
\score {
   \notes {
      \time 4/4
      \clef violin
      c'1
   }
}
```

5.13 CAPXML

(http://www.capella-software.com/capella.htm)
CAPXML is an XML format used by capella music editor as interchange format. Capella 2004 is a full music editor for sheet music production. The main focus of capella is to produce music scores for printing. The score rendering is at a good quality and it allows fine tuning for positioning of symbols on the score page.
Other tools are available to:

- scan scores from sheet music (capella-scan)
- produce audio cd from scores (capella-play-along)
- compose music (tonica)

The simple "do" example is encoded as:

```
<?xml version="1.0" encoding="ISO-8859-1" standalone="yes"?>
<score xmlns="http://www.capella.de/CapXML/1.0">
  <layout>
    <pages top="20" right="20" left="20" bottom="20"/>
    <distances>
      <staffLines small="1.2" normal="1.8" pageObj="1.8"/>
      <systems top="4" between="0"/>
    </distances>
```

```
      <instrumentNames>
         <font face="Times New Roman" height="12"/>
      </instrumentNames>
      <staves>
         <staffLayout description="layout1">
            <notation defaultClef="treble">
               <barlines to="7" from="3" mode="internal"/>
            </notation>
            <distances top="4" bottom="4"/>
            <instrument abbrev="" name=""/>
            <sound volume="80" instr="0"/>
         </staffLayout>
      </staves>
      <spacing abs="16" rel="75"/>
      <beamFlattening to="29" toMax="55" from="49" fromMax="148"/>
   </layout>
   <barCount y="3.5" x="0">
      <font face="Times New Roman" height="8"/>
   </barCount>
   <systems>
      <system justified="true" instrNotation="long">
         <staves>
            <staff layout="layout1" defaultTime="4/4">
               <voices>
                  <voice>
                     <lyricsSettings lineDist="1.5" firstLine="5">
                        <font face="Times New Roman" height="12"/>
                     </lyricsSettings>
                     <noteObjects>
                        <clefSign clef="treble"/>
                        <chord>
                           <duration base="1/1"/>
                           <heads>
                              <head pitch="C5"/>
                           </heads>
                        </chord>
                        <barline type="end"/>
                     </noteObjects>
                  </voice>
               </voices>
            </staff>
         </staves>
      </system>
   </systems>
</score>
```

5.14 WEDELMUSIC Model and Format

(http://www.wedelmusic.org)

WEDELMUSIC is an XML compliant format which includes constructs for the description of integrated music objects. Digital music objects compliant with the WEDELMUSIC format are called WEDELMUSIC or simply WEDEL objects. The model is supported by a full set of tools for multimedia music packaging and distribution. They are focused on a specific music piece or concept. Each WEDEL object presents sections about its: identification, classification, protection, printing, symbolic music (fonts, formatting rules, versions), image score, performance, documents, animations, lyric, audio, video, and color image. Hereafter these aspects are discussed with the rational for their inclusion.

- **Identification** section allows the identification of the music piece, ISMN, ISBN, etc., are included.
- **Classification** section allows the classification of the music piece according to multilingual archive mechanisms integrating Z39.50 and UNIMARK fields, plus other fields. Distinct classification records may be set up for the whole object and for its components.
- **Protection** section models details about the encryption of the WEDEL object and the watermarking of music (audio files and sheet music). According to a sophisticated Digital Right Management module, each operation which can be performed on the WEDEL object can be either permitted or inhibited; more than 150 different multimedia functionalities are managed. A permission table is available to define DRM.
- **Symbolic Music** section describes the scoring information, musical notation symbols, and their relationships. Symbolic music can include main score and parts. The symbolic description includes specific sections for classification and identification of the music score (main score and parts).
- **Image of Sheet music** section allows to integrate images of music scores into the WEDEL object without converting them into symbolic format. Therefore, in the same WEDEL object both symbolic notation and original images can be present.
- **Audio** section may contain none, one or more audio files in any format. Audio files can be watermarked according to the WEDEL object's parameters.
- **Performance** section describes the synchronization aspects between each audio file and the music score of the WEDEL object. The synchronization of audio files allows the contemporaneous visualization of the score and listening to the music with the selected audio and execution rate. This can be done by using symbolic music notation or simple images of music score. A performance can also be a sequence of slides with associated an audio file.
- **Documents** section may include documents in any format. Documents may be author biography, critical description of the piece, description of the music piece, etc.
- **Lyric** section includes multilingual lyrics associated with the music score and therefore with the WEDEL object. Several multilingual lyrics can be associated with the same symbolic part. This is performed by using the symbolic indexing of each music notation symbol.
- **Video** section may contain video files in any format.
- **Animations,** they can be in Flash included in HTML pages, or in PPT, or video, etc.
- **Image** section may include color images in any format. For example, the portrait of the author/performer, the picture related to the music or opera, or performer, the home/city of the author/performer, a picture of the instrument, a picture of the live performance, the CD cover, etc.

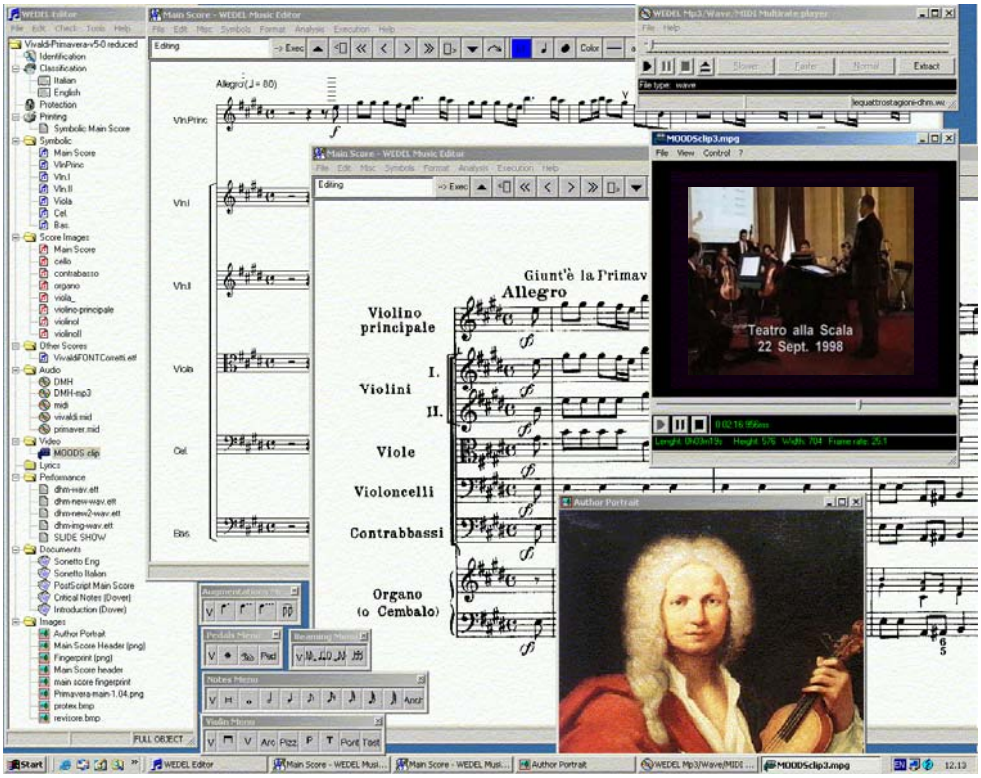


Fig.1 – An Integrated Multimedia Music Object

In each WEDEL object, several relationships among its components can be established as depicted in Fig.1. For example, the following relationships can be established:

- Symbolic-lyric: different lyric files can be referred to the same symbolic file.
- Symbolic-image score: relationship performed thanks to the specific number associated with each measure
- Image score to symbolic: relationship performed thanks to the specific number associated with each measure.
- From Symbolic to video, images, documents, audio files: relationships performed via http links which can be assigned to music notation elements.
- Images of music score to lyric:
- From symbolic or images of music score to the audio via performance section: each audio file can be synchronized with the music score.

Music Notation Aspects

The relationships among visual notation constructs can only be formalized considering syntax and semantic aspects. As far as music is concerned, the definition of a visual grammar is a highly complex task. To create the WEDELMUSIC formal model, we first modeled music notation components with a unified object-oriented formal model and language. In WEDELMUSIC and MOODS, the object-oriented model has been used as a music representation model and coding language, and also as the network message interchange protocol among lecterns. In WEDELMUSIC, the logical structure of music notation is modeled in XML while the graphic details are generated in real time by using the style coded in MILLA (Bellini, Della Santa and Nesi, 2001).

MILLA can be used for defining default and exceptions rules in order to manage the already described two kinds of problems. Simple conditions can be also forced. Several symbols strictly related to notes are managed: accidentals, ornaments, fingering, accents, instrumental symbols, etc. In several instances, the

mere presence of more than one of such symbol around figures changes the figures' order, orientation and position. For instance, accents must be placed on the opposite side of the stem (if any). In the event of a slur on the same note, some accents must leave space close to the note at the slur, thus they have to be placed just a bit farther from the notehead than usual (Ross, 1987), (Heussenstamm, 1987). The slur itself has to make room to other symbols when they are contemporaneously present on the same note.
In MILLA, the users can specify the following types of rules for the:

- insertion of symbols, such as the estimation of the direction (up/down with respect to the staff or with respect to the notehead) of stems, beam, slurs, etc. (they can be manually imposed if needed).
- positioning of symbols, such as the estimation of stem length, distance with respect to staff lines, beam angle, position of the ornaments, expressions, etc. with respect to the notehead.
- ordering symbols while considering the presence of neighboring symbols, ensuring precedence among symbols with respect to the notehead when depicting slurs, accents, markers, etc.;
- justification of measures according to specific algorithms, such as linear, logarithmic, and different scales; to consider alignment among voices and parts in main scores, line breaking and page breaking;
- beaming rules for beaming notes on the basis of time signature, etc.; and
- compressing symbols for the activation and deactivation of rules which display symbols in a compressed format, including generic rests and repeat symbols.

The set of MILLA rules associated with a music piece is a sort of formatting style, like the styles usually adopted in word processors. It is more sophisticated than those styles since MILLA rules can be activated according to the conditions estimated when considering local and/or general visualization contexts. Such rules typically change on the basis of the context: presence of other voices or parts, presence of other symbols, up/down the staff, etc. Most of the defined rules prevent the generation of collisions, others may permit the production of a specific collision under certain conditions (stems and beams, beams and barlines, stems and slurs, etc.). Once a new visual problem is detected, a specific rule can be added or the change can be manually performed.

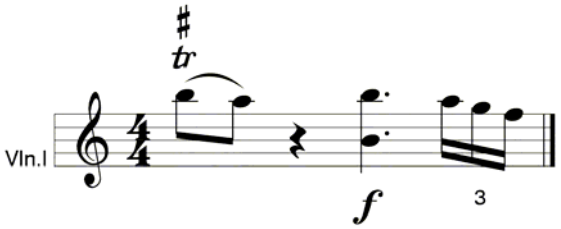


Fig. 2. An Example

In Fig.2, an example, which has been automatically formatted by WEDELMUSIC taking into account specific rules, is reported.
The slope of the two beams is determined by the following MILLA rule:

```
RULEPOS BeamInt BEAM RELNOTES ANYWHERE ANGLE=[0.15;0.3];
posIF STAFFNUMBER=1.0 AND STAFFLINES=5.0 AND DELTA<2.5 AND DELTA>0.5 THEN BeamInt;
```

The DELTA value is defined as (highest – lowest)/(number of notes -1) where highest-lowest is the distance in spaces (the distance between two staff lines) from the highest note to the lowest note of the beam.
The value of DELTA for the first beam is (11-10)/(2-1)=1 and for the second one is (10-8)/(3-2)=1, then both beams satisfy the condition of being greater than 0.5 and less than 2.5. This leads to the appliance of rule BeamInt. According to the former, the slope of the line connecting the higher and the lower note is calculated as follows: if this slope is between 0.15 and 0.3 this slope is applied; otherwise, when less than 0.15, slope 0.15 is assumed and 0.3 when greater. .
The rules used to determine the length of the stem for a chord are as follows:

```
RULEPOS Stem3_5 STEM RELNOTA LENGHT=3.5;
RULEPOS Stem4 STEM RELNOTA LENGHT=4;
RULEPOS Stem4_5 STEM RELNOTA LENGHT=4.5;
RULEPOS Stem5 STEM RELNOTA LENGHT=5.0;
RULEPOS Stem5_5 STEM RELNOTA LENGHT=5.5;
RULEPOS Stem6 STEM RELNOTA LENGHT=6.0;
RULEPOS Stem6_5 STEM RELNOTA LENGHT=6.5;
```

```
…
RULEPOS StemHeight STEM RELNOTA LENGHT=STEMHEIGHT;

posIF NOTE CROMA AND INCHORD AND NUMUP>0.0 AND NUMDOWN>0.0 THEN Stem3_5;
posIF NOTE SEMICROMA AND INCHORD AND NUMUP>0.0 AND NUMDOWN>0.0  THEN Stem4;
posIF NOTE BISCROMA AND INCHORD AND NUMUP>0.0 AND NUMDOWN>0.0  THEN Stem4_5;
posIF NOTE SEMIBISCROMA AND INCHORD AND NUMUP>0.0 AND NUMDOWN>0.0  THEN Stem5_5;
posIF NOTE FUSA AND INCHORD AND NUMUP>0.0 AND NUMDOWN>0.0  THEN Stem6_5;
posIF INCHORD AND NUMUP>0.0 AND NUMDOWN>0.0 THEN Stem3_5;

posIF INCHORD AND NOTE STEMUP AND UPPERD>7.0 THEN StemHeight;
posIF INCHORD AND NOTE STEMDOWN AND LOWERD>7.0 THEN StemHeight;

posIF NOTE CROMA AND INCHORD THEN Stem3_5;
posIF NOTE SEMICROMA AND INCHORD THEN Stem4;
posIF NOTE BISCROMA AND INCHORD THEN Stem4_5;
posIF NOTE SEMIBISCROMA AND INCHORD THEN Stem5_5;
posIF NOTE FUSA AND INCHORD THEN Stem6_5;
posIF INCHORD THEN Stem3_5;
```

The rule applied to the chord of the example is the one evidenced, since other rules do not apply. The UPPERD and LOWERD values are the distance from the center line of the staff in spaces of the upper and the lower note of the chord. The NUMUP and the NUMDOWN values are the number of notes of the chord above and below the center line of staff.
In this case the rule applied establish that the stem has to be 3.5 spaces.

The XML encoding of the example reported in Fig. 2 is the following (some details are missing, though):

```
<?xml version="1.0" encoding="UTF-8"?>
<SWF_Part>
   <WDFCID>00000100000K</WDFCID>
   <Identification>
      <WDFID>00020000H</WDFID>
      <Publisher PUBLISHING_STATUS="OTHER">Test</Publisher>
      <Preparation_date MAJOR_VERSION="0" MINOR_VERSION="1">20021015</Preparation_date>
      <Music_Geographic_Area>Europe</Music_Geographic_Area>
   </Identification>
   <Classification xml:lang="en" Description="English">
      <Author>N/A</Author>
      <Unique_Short_Name>Example</Unique_Short_Name>
      <Title>Example</Title>
      <Genre>Classical</Genre>
      <Style>Classical</Style>
      <Original_language/>
      <Composition_date>00000000</Composition_date>
      <Epoque>
         <Start_year>0</Start_year>
         <End_year>0</End_year>
      </Epoque>
   </Classification>
   <score ID="1" TYPE="NORMAL" INSTRUMENT="Vln.I">
      <origin FROM="WEDELED"/>
      ...
      <measure PROGRESSIVE="1" ID="1">
         <justification MAINTYPE="LOG" MAINJUST="2.0" PMAINTYPE="LOG" PMAINJUST="2.0"/>
         <header>
            <clef TYPE="TREBLE"/>
            <keysignature TYPE="DOM"/>
         </header>
         <timesignature TYPE="FRACTION" NUMERATOR="4" DENOMINATOR="4"/>
         <layer NUMBER="1">
            <beam ID="10" STEMS="DOWN">
               <note ID="1" DURATION="D1_8" HEIGHT="11">
                  <ornament TYPE="TRILL" UPDOWN="UP" ACCSUP="SHARP"/>
               </note>
               <note ID="2" DURATION="D1_8" HEIGHT="10"/>
            </beam>
            <rest ID="3" DURATION="D1_4" HEIGHT="1"/>
            <chord ID="5" DURATION="D1_4" STEM="DOWN">
               <chordnote ID="5" HEIGHT="4"/>
               <chordnote ID="6" HEIGHT="11"/>
               <augmentation DOTS="1"/>
               <dynamictext DYNAMIC="F" UPDOWN="DOWN"/>
            </chord>
            <beam ID="11" STEMS="DOWN">
               <note ID="7" DURATION="D1_16" HEIGHT="10"/>
```

```
            <note ID="8" DURATION="D1_16" HEIGHT="9"/>
            <note ID="9" DURATION="D1_16" HEIGHT="8"/>
         </beam>
      </layer>
      <barline TYPE="END"/>
   </measure>
   <horizontal ID="1" TYPE="TUPLE" UPDOWN="DOWN" TUPLENUMBER="3" TUPLELINE="FALSE">
      <address MEASURE="1" LAYER="1" FIGURE="11" CHORD.OR.BEAM="7"/>
      <address MEASURE="1" LAYER="1" FIGURE="11" CHORD.OR.BEAM="9"/>
   </horizontal>
   <horizontal ID="2" TYPE="SLUR" UPDOWN="UP">
      <address MEASURE="1" LAYER="1" FIGURE="10" CHORD.OR.BEAM="1"/>
      <address MEASURE="1" LAYER="1" FIGURE="10" CHORD.OR.BEAM="2"/>
   </horizontal>
</score>
</SWF_Part>
```

The symbolic description of a part (the score tag) is decomposed in a sequence of measures followed by the sequence of horizontal symbols (slurs, tuples, crescendo, diminuendo, etc.). Each single measure contains the header information as clef, key signature and the time signature followed by one or more layers with figures (notes, rests, beams, chords, etc.)
It should be observed that horizontal symbols refer to the starting/ending figure using a path of identifiers (measureID/layer/figureID/figureID…) for example the second note of the first beam is identified by path:1/1/10/2.
It should be observed that the HEIGHT attribute of notes/rests refers to the line/space where the note has to be positioned, 0 is the first line of the staff, 1 the first space, 2 the second line, 3 the second space etc.

WEDELMUSIC presents several innovative features that will be extremely useful for the new generation of music notation applications:

- Symbolic indexing of music notation. This makes it possible to implement selective and non-linear undo, symbolic navigation among symbols and multimedia documents (Bellini at al, 2002), management of versioning, mounting multilingual lyric.
- DRM support, including rules and watermarking support of sheet music and audio/video files. It is very important to integrate this aspect into the music model since it allow to define specific rules for versioning, music fruition, music notation printing, multimedia integration and fruition, etc.
- Integration with multimedia components such as described above. Multimedia integration can be very useful for creating relationships among music notation symbols and multimedia content for educational purpose and content valorization.
- Definition of custom and new symbols with their formatting rules, starting from 500 symbols.
- Management of information related to formatting the music in two different cases/views. For instance, computer view for music score reading on the computer screen during tuition and edutainment and print view as a graphic user interface for producing professional sheet music. This is performed for both the main score and each single part. This makes it possible to have, in a unique format, both main score and parts with their formatting information.

5.15 Other models and formats

XML based

- **ChordML** is a simple XML application that focuses on lyrics and chord symbols, not scores. (http://www.cifranet.org/xml/ChordML.html)

- **EMNML** (Extensible Music Notation Markup Language) is described in “Developing A New Way To Transfer Sheet Music Via The Internet” by Eric Mosterd. (http://sunburst.usd.edu/csci/research/theses/graduate/sp2001/emosterd.pdf)

- **JScoreML** is a music notation markup language and a score editor (http://nide.snow.utoronto.ca/music/index.html).

- **MHTML** is music notation extension of HTML (http://www.shigeta.net/mhtml/index.htm).

- **MusiqueXML** derivates from MusiXML. Used in a project to transform a music notation XML format to SVG.
  (http://francois.chastanet.free.fr/musiquexml/MusiqueXML.htm).
- **VMML** (Virtual Musician Markup Language) is described in a paper by Damien Tardieu.
  (http://vrlab.epfl.ch/research/S_VMML.pdf)
- **XMusic** is an XML Music notation language.
  (http://www.palserv.com/XMusic/)
- **XScore**
  (http://grigaitis.net/xscore)

Text based

- **ABC**
  (http://abc.sourceforge.net/, http://www.gre.ac.uk/~c.walshaw/abc/)
- **CMN Common Music Notation**
  (http://www-ccrma.stanford.edu/software/cmn/)
- **DARMS**
  (http://www.ccarh.org/publications/books/beyondmidi/online/darms/)
- ****kern – Humdrum Toolkit**
  (http://www.lib.virginia.edu/dmmc/Music/Humdrum/)
- **Mup**
  (http://www.arkkra.com/)
- **Muscript 2.0**
  (http://www.pjb.com.au/muscript/muscript2.0.html)
- **Musedata**
  (http://www.musedata.org/)
- **Nightingale Notelist**
  (http://www.ngale.com/index_02.html)
- **OMNL** – Open Music Notation Language
  (http://www.todesco.com/paperchord/omnl/omnlcontent.html)
- **Scot**
  (http://www.recordare.com/good/scot.html)
- **Sharp-Eye**'s .mro format
  (http://www.visiv.co.uk/tech-mro.htm)

6 Music Notation Models and Languages Comparison

Tab.1 is a scheme summary for the comparison of some of the above-mentioned languages and models, considering the aspects discussed in previous sections. In the table, "(Y)" means that the answer is not completely true or satisfactory since these features are potentially available and their presence depends on the visual/graphic editor used. In most cases, the positioning of symbols is made a bit more complex by using a set of Visualization Parameters. This approach results in imposing a single rule for the positioning of the whole score, and turns out to be a coarse simple and unsuitable solution for managing automatic positioning.

	WEDEL MUSIC	**MIDI**	**SCORE**	**MusiXTEX**	**NIFF**	**SMDL**	**Music XML**	**FINALE**	**GUIDO**	**CAPXML**
Logic model	Y	N	Y	Y	Y	Y	Y	Y	Y	Y
Classification	Y	N	N	(Y)	(Y)	Y	(Y)	(Y)	N	N
Identification	Y	N	N	(Y)	(Y)	Y	(Y)	(Y)	N	N
Visual	Y	N	Y	Y (print)	(Y)	N	Y	Y	Y	Y
Visualization Rules	Y (Milla)	N	N	N	N	N	N	(N)	N	N
Visualization Parameters	Y	N	Y	N	Y	N	N	Y	(Y)	Y
Performance or MIDI play in real time	Y	Y	N (SCORE MIDI)	N	Y	(Y)	Y	Y	Y	Y
Synchronization with real audio	Y	N	N	N	N	N	(Y)	N	N	N
Music scores images	Y	N	N	N	N	N	N	N	N	N
Animations	Y	N	N	N	N	N	N	N	N	N
Versioning support	Y	N	N	N	N	(Y)	N	N	N	N
Selective and un-linear undo	Y	N	N	N	N	N	N	N	N	N
DRM support	Y	N	N	N	N	N	N	N	N	N
Multimedia Integration (video, etc.)	Y	N	N	N	N	(Y)	N	N	N	N
Symbolic indexing of music notation	Y	N	N	N	N	N	N	N	N	N
Multilingual Lyric	Y	N	N	N	N	N	N	N	N	N
Graphic Lyric	Y	N	Y	Y	Y	Y	Y	Y	N	Y
Stream Lyric (a sequence of syllables)	Y	N	N	N	N	Y	N	Y	(Y)	Y

Tab.1 -- Coverage of Music Aspects, Y and N represent if a given feature is supported or not, when they are in round brackets means that incomplete or non satisfactory coverage is provided.

In Tab.2, the results of the comparison are reported. The information has been obtained by reviewing tools related to the languages. When a number is given, it was obtained by means of questionnaires distributed to musicians and computer scientists. In the Table, Interactive Music Editing states the availability of a visual editor; Adding/Editing Graphic Entities is the possibility of using graphic primitives (like lines) superimposed on the score; Print support, Extraction of Parts, Extensibility of Symbols and Fusion of Parts do not need any comment; Main Score editing full length is the possibility of editing the main score as continuous information not interrupted by page breaks; Music Distribution for Cooperative work is the possibility of cooperative working during music editing or execution; Number Notation Symbols is a vote about the number and the relevance of the available notation symbols; Logic and Visual Independence is a vote about the independence of logic and visual aspects. This last vote has been obtained by analyzing the

music formats and models and observing the behavior of symbols on the editors (when available) during insertion, moving and deletion.

	WEDELMUSIC	**MIDI**	**SCORE**	**MusiXTEX**	**NIFF**	**SMDL**	**MusicXML**	**FINALE**	**GUIDO**	**CAPXML**
Interactive Music Editing	Y	Y	Y	N	Y	N	Y	Y	(Y)	Y
Adding/Editing Graphic Entities	(N)	N	Y	Y	N	N	N	Y	Y	Y
Print support	Y	Y	Y	Y	Y	N	N	Y	Y	Y
Extraction of Parts	Y	N	Y	N	Y	N	Y	Y	N	Y
Fusion of Parts	Y	Y	Y	N	(Y)	N	N	(N)	N	Y
Main Score editing full length	Y	Y	N	Y	Y	N	Y	Y	N	Y
Automatic formatting of music, symbol positioning	Y	N	Y	Y	N	N	Y	Y	(Y)	Y
Automatic line breaking	Y	N	Y	Y	N	N	Y	Y	Y	Y
Managing Formatting information for main score and parts in the same model and file	Y	N	N	N	N	N	N	N	N	N
Music Distribution for computer supported cooperative work	Y	N	N	N	N	N	(Y)	N	N	N
Music Distribution tools	Y	Y	N	N	N	N	N	Y	(Y)	N
Music Analysis	Y	N	N	N	N	N	Y	Y	N	Y
Braille Music	Y	N	N	N	N	N	N	Y	N	Y
Vote on the number of notation symbols: from 1 to 10 (very good)	9	4	10	8	4	4	7	9	5	
Extensibility of Symbols	Y	--	Y	Y	Y	Y	(Y)	Y	N	Y
Logic and Visual Independence	Y	N	N	N	Y	Y	Y	N	N	
Developing tools	Tool Kit	Y	N	N	N	N	N	Y	Y	Y
Liveness of editor	monthly	stable	none	none	none	none	monthly	yearly	none	monthly

Tab.2 -- Editor, Languages and Model Comparison, Y and N represent if a given feature is supported or not, when they are in round brackets means that incomplete or non satisfactory coverage is provided

The main problems of the considered languages in meeting the requirements of the new application are mainly due to the lack of formalization in the language and in the model for storing/coding music. Concerning the definition of relationships among symbols, the most flexible and complete language is NIFF. Even this language is not fully satisfactory since it does not model all the relationships. The real problem of these languages is the lack of versioning support and the management of visualization rules. These two aspects are also missing in the several new mark-up languages which have been recently proposed on the WWW. They are presently strongly limited compared with the above models. Most of them do not present slurs, accents, instrumental symbols, justification, etc. There are also several other types of languages for coding music, as described in Selfridge-Field (1997), but unfortunately, none of them can be held as a complete support for the requirements of the new incoming applications.

On the account of the work performed on several projects and products, some tools for operating the conversion among these formats have been produced. Tab.3 summarizes the available converters among the considered languages. Some of them exist only at a hypothetical level because they are for instance based on claimed prototypes but the converter is not distributed or the conversion has been only studied and analyzed.

In all these conversions, several details are lost since different formats treat the information in a different manner, others present a limited number of symbols and relationships, thus being forced to eliminate information during the transformation.

From\ to	WEDEL MUSIC	MIDI	SCORE	Musi XTex	NIFF	SMDL	Music XML	FINALE	GUIDO	CAP XML
WEDEL MUSIC	--	Y	N	N	N	N	N	N	N	N
MIDI	Y	--	Y (MIDI SCORE)	N	Y (Lime)	Y	Y	Y	Y	Y
SCORE	(Y) via Sibelius	Y	--	N	N	N	(Y) via Sibelius	N	N	N
MusiXTEX	N	N	N	--	N	N	N	N	N	N
NIFF	N	Y (Lime)	N	N	--	(N) (Cantate)	Y	N	N	N
SMDL	N	N	N	N	(N) (Cantate)	--	N	N	N	N
MusicXML	N	Y	Y (SipXML)	N	N	N	--	Y	Y	N
FINALE	(Y)	Y	Y (Final SCORE)	N	N	N	Y	--	(Y) only few symbols	N
GUIDO	N	Y	N	N	N	N	N	N	--	N
CAPXML	N	Y	N	N	N	N	N	N	N	--

Tab.3 -- Available Format Converters, Y and N represent if a given feature is supported or not, when they are in round brackets means that incomplete or non satisfactory coverage is provided

From the above comparison it evident that WEDELMUSIC is the best candidate for managing innovative and new emerging multimedia applications. It contains several innovative aspects at level of music notation such as symbolic indexing, strong distinction between logic and graphics, MILLA formatting engine, multilingual lyric, etc.

6.1 Future Trends of Music Notation

Music notation and music notation programs are at present passing through the same evolution that text and text editors went through several years ago. Music publishers are still focused on obtaining good quality (in terms of visual formatting) of music scores. On the other hand, the market would like to have lower prices and considers the quality not so important. Publishers state that they sell only high quality music scores, which is quite obvious when prices are so high. Musicians are used to reading very poor music scores in terms of both formatting and visual resolution. To grasp this, it is enough to observe the quality of sheet music used by musicians in the conservatories and music schools. Their quality is really poor, they are frequently photocopies of photocopies of old originals. Sometimes it can even occur that the music is unreadable in some very ancient pieces.

The market is ready for massive production of sheet music while publishers are not ready for it. The profit margins are in the high quality since the numbers are small. At the same time students and music lovers prefer to make photocopies. This is the same story as it happens with books. Publishers have recovered their market only by producing cheap books with low quality paper and formatting shapes.

For these reasons, the future of this field is in the automatic formatting of sheet music and in the delivering of this via Internet. Huge problems have to be solved for modeling music and for integrating this with the multimedia aspects of music.

7 Integration of Music Notation in the MPEG framework

Several products are currently available in the market that present some form of integration of music notation and multimedia. These products are in the areas of

- music education (notation tools integrating multimedia),
- music management in libraries (music tools integrating multimedia for navigation and for synchronization),
- entertainment (mainly synchronization between sound, text and symbolic information),
- piano keyboards with music notation and audiovisual capabilities,
- mobile applications,
- i-TV distribution
- etc.

The integration of music notation in MPEG could completely satisfy the requirements of these tools and much more permitting to these tools to integrate the powerful model of MPEG for multimedia modeling, representation, coding and playback. The integration of music notation with MPEG will open the way for a large number of new applications and markets related to the above applications.
This initiative may increase the present market for music notation that presently is mainly limited to sheet music production, and may open the path to create very interesting new applications, renovating the present applications that already use some integration between multimedia and music notation.

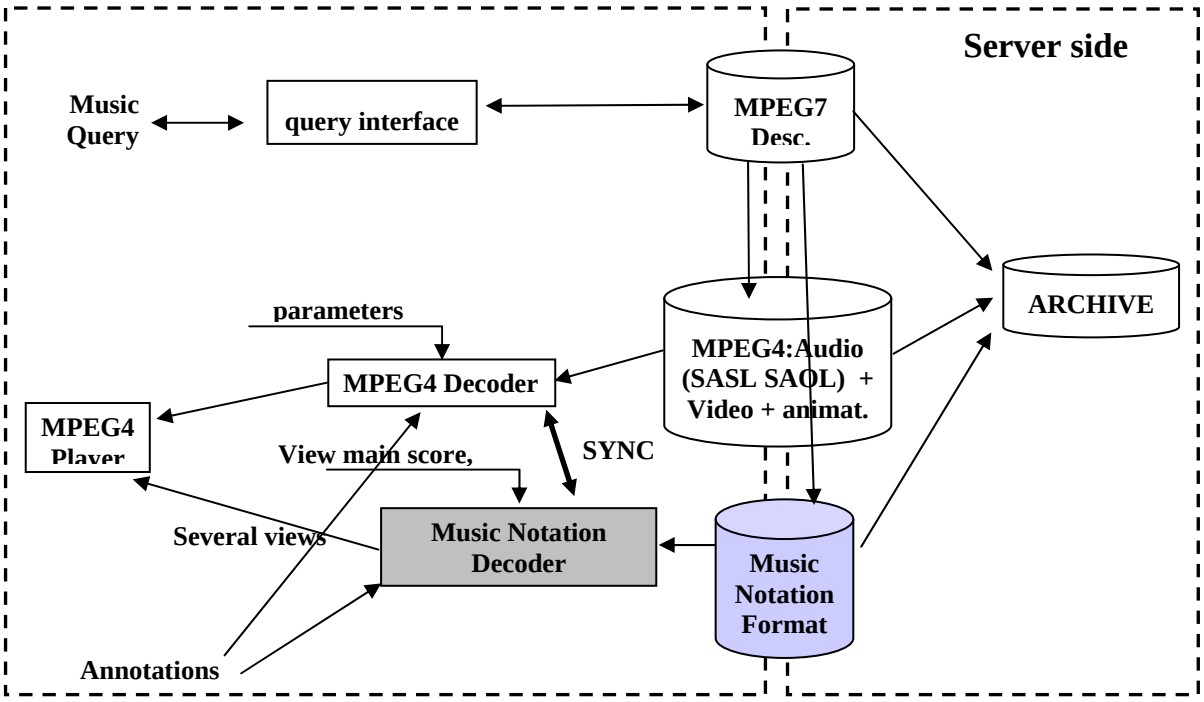
The repercussions would have a wide effect on music education and on the development of music as a whole. This, in turn, would be of great long term benefit to the music industry, quite apart from the applications, mentioned in this document, which can be written and sold in the shorter term.

The benefits of solving the problems highlighted above are remarkable. A unified model for music notation integrated with an MPEG standard could provide the solutions for several classes of applications including:

- distribution of music content and courseware on media devices, i-TV and other devices such as PDAs,..
  - multimedia music distribution on multimedia devices, mainly visual and audio support;
  - fruition of composite music-related information, mainly audio and simple visual representation;
- creating innovative applications incorporating multimedia integration with:
  - music editing and formatting (an example might be WEDELMUSIC editor with MILLA engine)
  - cooperative work on music score (an example might be MOODS)
  - producing opera content with VR-3D, synchronization of music notation, lyric and virtual modeling (an example can be OPENDRAMA)
  - synchronization of symbolic music score and real audio (an example might be WEDELMUSIC)
  - 3D hand representation and synchronous score visualization (an example might be MUSICALIS, EMEDIA, PLAYPRO, GUITAR Magic)
  - synchronization of images of music score and real audio (an example might be WEDELMUSIC)
- creating content and courseware for music education
  - examples of multimedia music content can be MUSICALIS, SMARTSCORE, WEDELMUSIC
  - cases in which the pupil can do
    - play training (examples can be MUSICALIS, PIANO TUTOR),
    - compositional training,
    - play along (see Master Play along, MUSICALIS, PIANO TUTOR, VOYETRA),
    - sing training,
    - conduction training (see example of Vienna Philharmonic),
    - improvisation training, etc.
- production of music notation and synchronization with other media
  - optical music recognition applications
  - music transcoding: from audio to music notation via audio processing
  - establishing relationships among images of sheet music and symbolic representation
- creating and managing content in music archives
  - examples can be WEDELMUSIC tools
- creating the support for the interoperability among the above mentioned different applications.
- creating musical information for consumers with different accessibility needs.

The missing parts to cope with these applications and requirements are mainly:

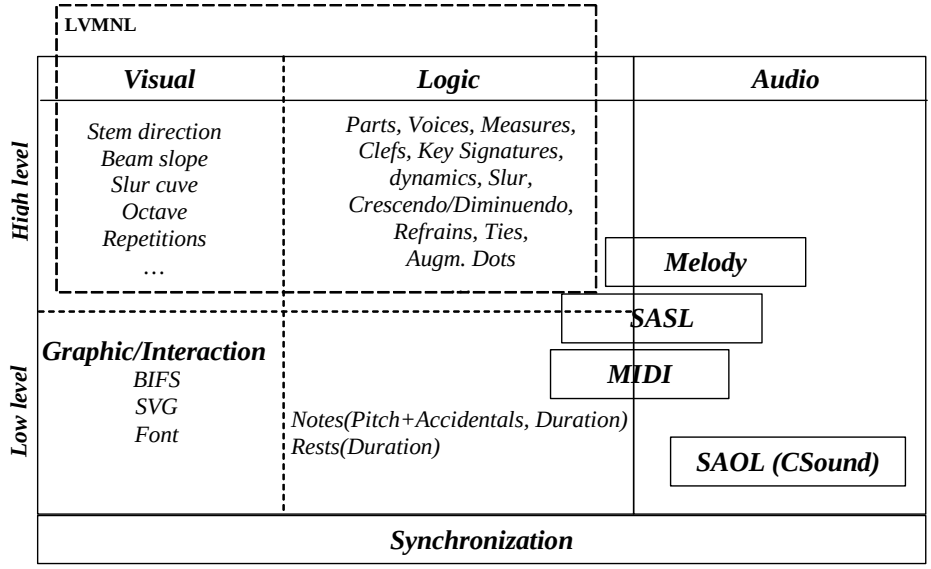
- Music notation modeling
- Music notation decoder



Applications and companies that may provide these technologies are listed below and many others are reported in the www page of the AHG http://www.dsi.unifi.it/~nesi/mpeg/ahg-mn-65-66.html :

- ARCA, Italy
- Berlioz, http://www.berlioz.tm.FR
- Capella, Music Editor, Educational Tools, audio processing tools, etc., Germany http://www.whc.de
- Encore music editor and company
- Finale 2003, CODA, www.codamusic.com , CODA music, US
- Graphire Music press, **Errore. Riferimento a collegamento ipertestuale non valido.**
- IGOR, NOTEHEADS, http://www.noteheads.com
- Lime music editor and company
- Mozart, Mozart, http://www.mozart.co.uk
- MusEdit, http://www.musedit.com
- MUSICALIS; HTTP://www.musicalis.com
- MusicEase, it includes an automatic accompaniment tool.
- Nightingale (Mac), note-processor, from Adept Music Notation Solutions (USA)
- Recordare with MusicXML, for the Music Notation format
- SCORE: http://ace.acadiau.ca/score/LINKS3.HTML , music editor
- Sibelius Software, UK, www.sibelius.com
- VIVALDI, Amadeus, http://www.vivaldistudio.com
- WEDELMUSIC, WEDELMUSIC Authoring, http://www.wedelmusic.org, Italy

The following figure reports some of the Music Representation domains and the relationships with MPEG solutions. This has been discussed at the first workshop on Music Representation and more in deep in a document visible on:



The requirements reported in this section are mainly related to the Visual and Logical parts.

7.1 Requirements of music notation model in MPEG

The following list of requirements is referred to the Music Notation requirements for including music representation inside the MPEG. These aspects are mainly focused on:

- Defining the most relevant requirements for modeling Logical and Visual aspects described above,
- Graphical aspects will be supported by BIFS and FONTS
- Audio aspects will be integrated with those supported in MPEG in SASL and SAOL.

The requirements are divided in:

- General requirements,
- Logical requirements
- Visual Requirements
- Extended Functionalities requirements
- Integration Requirements

7.1.1 General Requirements

- **scalability**: it has to be possible to write simple music notation pieces files without spending time and token to describe all the music notation **context** (for instance, it has to be possible to write simple music pieces describing only the needed parts avoiding to spend time and language tokens in describing the full context, that could be: collection, sections, parts, voices, etc.). At the same time, the model has to support the writing of complex structure of music pieces with multivoices, multistaff, etc., according to the applications identified. This will make it suitable for simple devices and applications such as those supposed for mobile equipment. This requirement implies that the music notation decoder has to consider for each aspect that can be missing a default value and behavior.
- **interoperability**: interoperability is required among different devices: i-TV, Mobile Devices, PDAs, Tablet PCs, electronic musical instruments, special hardware, cars, etc.; and also among software applications: music editors, music tools for educational purpose, music viewers on the i-TV or PDAs.
- **extensibility**: includes the possibility to create new symbols with their related rules or code for formatting (conversion to the visual domain and from this in graphical domain) and executing/playing (conversion to the audio domain). Please see the requirements of **formatting**.
- **support for domains:** the language has to support a sharp distinction between music notation domains. In the language has to be clear which aspects are logical, visual and graphics and possibly audio and analytical.
- **decodable**: the model has to permit the decode of music notation with different decoders to obtain different views of the same music representation (main score, parts, Braille, Gregorian, spoken, etc.). This can be done by using additional information or automatic formatting engines. This is very useful to

work on different devices (piano keyboards, PDAs, mobiles, i-tv, sheet, screen, etc.), etc. The music notation integrated with audiovisual could be streamed with video, audio, .. on mobiles, etc. (see MPEG4 model). Different decoders can be realized according to different models of music reproduction and visualization. This requirement has been detailed in the document M10355.

- **distribution and Protection:** the music notation model has to support the distribution on Internet integrated with any kind of audiovisual content to satisfy the needs of realizing the above mentioned applications. To this end, the several activities that can be performed on the music notation model have to be classified to identify the Digital Rights Management rules, for example by defining specific rule with the MPEG21 to control all the possible functionalities of a music notation content:
    - playing, printing, annotation, execution, formatting, transposing, etc.
- **editable**: The music notation has to be editable with some music editor. The format has to be supported by some specific editors for changing in deep all the details of the symbols involved.
- **readable**: The music notation model should be readable by humans when will be in a non binary form. This probably means that we need to have some XML model.

7.1.2 Extended Functionalities requirements

- An **Univocal reference ID** for each music notation symbol in the music piece. This is useful for Multilingual and Annotation addition.
    - **Multilingual Lyrics:** Each symbol or at least the most relevant should be univocally identified for associating to them the lyric syllables. The model in this way may include the possibility of loading, on the same music representation, lyric texts of different languages. This is a very important feature for realizing multilingual CD-DVDs without constraining the producer to create different versions of the same music notation model. This requirement is also useful for the production of multilingual Karaoke.
    - **Annotation:** the model has to permit the insertion and the creation of annotations by the final users. The annotation may be musical entities, text, and HTTP documents or links. They can be useful for adding personal comments to a digital object, to find in a second phase the specific point of the entertainment content, etc. This means that each annotation has to be associated with each specific element of music notation univocally, symbol by symbols. The list of annotations may be navigated and listed. Annotation coming from several users may be merged and distributed, saved and loaded. Examples of annotations are:
        - **Reference Link**, as described in the following
        - Traditional music notation elements and symbols: accidentals, fingering, bowing, adding of attention marks, adding of caution accidentals, dynamics, tempo marking, breath marks for wind instruments, special marks such as pedals for piano, harp, tablature for guitar, cue notes…
        - hand-produced annotation (lines, frames, specific shapes…),
        - multimedia content as audiovisual (audio, video, still images, animations, actions to the external world and on physical devices).
- **Reference links** for each music notation symbol. These are useful for realizing:
    - **User Navigation (relationships or links):** the model has to permit the insertion and the creation of relationships between each notation symbol (or on at least the most important, such as the **Music Notation Events and the structural aspects**) and audiovisual elements (text, documents, audio pieces, video, music pieces, http links, etc.). The relationships-links should be established during the production of the content and can be navigated by the final user. They can be regarded as ***relationships with internal content elements or external links***. They can be useful for:
        - explaining the music notation,
        - navigating from the music notation elements to other audiovisual aspects in a MPEG4 file,
        - connect the content with web resources.
    - **Automatic Execution**: the music notation can be executed without playing/rendering it for the simple purpose of activating specific actions associated with execution of a specific music notation symbol. This can be useful for activating the execution of some action when a music note is executed (this can be realized by including in the model some specific annotations) for:
        - educational purpose
        - creating applications in the field of art and theatrical activities.

        - playing some audiovisual or
        - activation of some external events.
- **Formatting:** the decoder of music representation has to properly manage all the information for reproducing the music notation on different formats and page size (computer windows, sheet music, etc.), different type of decoders may be considered (e.g., main score, parts; Braille or CMN, etc.). The information needed to reproduce the music on different formats have to be included into the music notation model, the rest has to be part of the context.
- **Query by content:** the model has to support the query by content by integrating audiovisual and music notation, usage of music notation model for verifying the match with MPEG7 descriptors of music that presently are very simple, increasing the expressivity of MPEG7 music descriptors.

7.1.3 Logical requirements

The detailed requirements of a music notation model are not needed for creating a call for technology and for enforcing the music notation into MPEG. On this basis, a call for technology could be prepared. The present requirements regarding the logical aspects of music notation are those which have been identified to be strictly necessary to:

- guarantee the above mentioned general requirements.
- have a music notation format correctly working in the MPEG environment
- guarantee the exploitation of the functionalities which are needed to support the innovative applications listed in the WWW site of the AHG.

The scope of this model has been preliminary limited to include:

- **Musical structure representation**
  - Structuring the logical representation
  - Including visual aspects as separate and well identifiable aspects
- **Western music notation**
- **Tablatures and Percussions**

Recently, the logical part that we are going to define is absolutely broader.

The most relevant Requirements of the Logical aspects of Music Notation are:

- **Structure**: the structure of music notation model has to model the following aspects
  - **Collection:** a set of operas or collections plus related **metadata**.
  - **Opera:** a set of movements plus specific **metadata**.
  - **Movement**: a non interrupted (of musical time) single division of music, which is a collection of MVP, from 1 to N; plus specific **metadata**.
    - **Visual**: main score, size of the page sheet and printing parameters, computer window visualization parameters.
  - **MVP, Multi Voice part**: a part with one or more voices, a set of voices that are synchronized, they can be arranged in different manner on staffs/staves, plus specific **metadata**.
    - **Logical**: executable or not.
    - **Visual**: hidden or not, number of staves and their scale/size, size of the page sheet and printing parameters, computer window visualization parameters.
    - **Graphic**: size of the lines, etc.
  - **Voice**: a sequence of **Music Notation Events, MNE**, which can navigate on different staves, etc.
    - **Logical**: executable or not, type of voice (pentagram, tablature (number of strings), percussions (number of lines), etc.)
    - **Visual**: hidden or not
  - **Ossia**: an alternative or an explanation of music representation symbols or set of them. It is an alternative MVP associated with the referred MVP by using a starting and an ending MNE.
    - **Logical**: executable or not, executable in the place of referred music section.
    - **Visual**: hidden or not
  - **Music Notation Event, MNE**: An MNE is something that occurs along the MVP. Each event has:

        - a starting instant that can be started in conjunction of the end of one the previous MNEs. Thus the events can be overlapped.
        - a duration expressed in time units. The time units are capable of describing all note duration plus augmentation dots and touples duration. The visual features of the MNE include also the visual duration, etc. The duration can be 0.
    - **Segment:** portion of a Movement identified with a starting and ending Music Notation Event.
    - **Refrain**: manage the modeling of compressing music model that presents segments that have to be executed more that one time or in a different order in which they have been initially encoded. In the Music Notation model in MPEG is convenient to have the music linearly exploded and not compressed to avoid the multiplication of references. Thus a refrain results to be mainly a visual aspect.
        - **Visual**: A Refrain is represented by specific visual aspect on barlines, with labels and/or jump symbols associated with barlines.
- **Music Notation Events:** Some examples are reported in the following to give more evidence at the needed music notation requirements
    - **Barline**: a barline can be regarded as the marking of a measure and it is regarded as a special event. This solution simplifies the problems of music analysis, execution, and justification of main scores that have parts with different time signatures (such as the Don Giovanni piece). A barline is a MNE with duration equal to the sum of the durations of the MNEs that are present until the next barline. The Barlines have a unique ID as the other music notation symbols. Each barline as a:
        - **Visual**:
            - there are several types of barlines: simple, double, hidden, dashed, etc. Some of them are associated with the refrains, and thus their visual representation has a double meaning, as segment end, as delimitation of measure, as end-start of refrain.
            - End line barline or not
            - End page barline or not
            - Justification parameters for:
                - Formatting the measure in several formats such as: main score on resize-able computer windows, single part on resize-able computer windows, main score on a sheet music single part on a sheet music. The justification include a type (linear, log, etc., and a scale factor) for each of the formats identified.
    - **Context** and the **Transposed Context** with a duration equal to the next change of context The **Context** is defined as:
        - **Clef or clef change**, they are typically smaller in size that normal size. They are typically placed before the MNE from which the clef has to be intended changed.
            - **Visual:** position
            - **Graphic**: font
        - **Key signatures changes**, they are typically accompanied by accidentals adjusting the tonality from the previous to the next. They are typically placed before the MNE from which the key signature has to be intended changed.
        - **Time signature changes**, they are typically placed before the MNE from which the time signature has to be intended changed.
            - **Visual:** position
            - **Graphic**: font
    - **Notes**, with all their features and markers that are discussed in the sequel:
        - **Logical**: pitch, cue, accidental, Marker (see in the following), active (to be played or not), logical duration (user for the execution), reference to the audio section of MPEG.
        - **Visual**: visual duration, number of dots, note head, size (small or normal), colour, direction of the stem, visible or hidden, etc. On the basis of the visual duration and on the basis of the presence of dots the correct visualisation is produced.

        - *When logical and visual durations are missing the note/rest is ghosted, when logical duration is present and visual is not present it is hidden, when visual is present and logical is not present it is graced. When both are present they are used for different purpose.*
        - **Graphic**: font
    - **Rests**, their features that will be discussed in a second phase are
        - **Logical**: active (to be played or not), logical duration (user for the execution), reference to the audio section of MPEG**.**
        - **Visual**: visual duration, number of dots, position (duration based or central), highness, size (small or normal), colour, etc. On the basis of the visual duration and on the basis of the presence of dots the correct visualisation is
        - **Graphic**: font
    - **Chords**, Vertical groups of simultaneous MNE with the features of the notes. Some of the features are associated with the chord other to each single note. Those associated with the whole chord are:
        - **Logical**: pitch, cue, accidental, Marker (see in the following), active (to be played or not), logical duration (user for the execution), reference to the audio section of MPEG.
        - **Visual**: visual duration, number of dots, direction of the stem, etc. the chord can be represented in several manners, by using a freatboard or by means of textual annotation. Each note of the chord may have different colour, and specific accidentals and marker such as fingering, etc. On the basis of the visual duration and on the basis of the presence of dots the correct visualisation is
        - **Graphic**: Each note of the chord may have different font
    - **0 duration MNEs** such as:
        - **Anchorages**: hidden points for attaching: crescendo, diminuendo, external events of synchronisation, lines, markers not associated to notes and rests but that have a temporal position, etc. The anchorage temporal position is given in percentage with respect to the duration of the previous MNE in the MVP. For example, thet may have associated with breath and/or attention markers.
- **Horizontal Groups, HG**: are structural elements that present a starting and an ending MNE. Each of them has some additional features. For example:
    - **Beams**: to see beams as HG permits to realize beams crossing barlines
        - **Visual**: the slope of the beam, the thickness of the beam, the presence or not of flags in some cases, the direction of the beam, the position of the beam in the middle or out of the staves in a multistaff part. etc.
    - **Tuplets**: they can be nested and, they have a number associated with. On the basis of the number and of the duration of the MNEs included the Logical Duration of these MNEs is changed.,
        - **Visual**: the number can be written or not. can be drown in different manners (squared and rounded), up or down.
    - **Tie**: connection between to successive notes or chords of the same pitch, it has a double meaning as visual and structural description of the harmony
        - **Visual**: shape of the tie, up or down, etc.
    - **Slurs**:
        - **Visual**: up and down, shape of the slur, 2 or more knot points for creating different type of slurs.
    - **Bend**: also called brackets,
        - **Visual**: they can be continuous, dotted, dashed, up or down, etc.
    - **Lines**:
        - **Visual**: they can be continuous, dotted, dashed, up or down, etc.
    - **Double event Piano pedals**:
        - **Visual**: up or down, etc.
    - **Dynamics**: crescendo, decrescendo, etc.
        - **Visual**: they can be continuous, dotted, dashed, etc.; up or down, etc.
    - **Double Events Ornament:** glissando, wave, etc.
        - **Visual:** up or down, etc.

    - **Octave and 15esime**: change of octave, typically only visual aspect.
        - **Visual**: up or down, etc.
- **Markers**: are music notation symbols associated with MNE, they may have a logical meaning that may impact on the music execution and a visual representation. The most relevant categories of Markers are reported in the following. For each of them:
    - **Logical**: some of them may have a specific music notation meaning expressed in terms of music notation to be used during the execution for better rendering of music influencing the music note playing. In this case, a semantics has to be associated with in terms of a music **Segment**.
    - **Graphic**: fonts, shape in terms of graphic primitives in SVG.
    - **Visual**:
        - Default position given up or down with respect to the note and its stem when applicable
        - Specific displacement with respect to the default position given in music units

  For example:
    - **Accents-expressions**: staccato, tenuto, portato, etc.
    - **Generic User expressions**: added by the user or for custom markers
    - **Simple event Ornaments**: turn, double turn, arpeggio, arpeggio up, arpeggio down, turn back, turn up, etc. They may have one or two accidentals.
    - **Instrumental indication**: bow up, bow down, etc.
    - **Mutes**: with the E, with +-, with “con sordina”, “senza sordina”, etc.
    - **Refrain Marks**: they can be placed only on barlines
    - **Tempo marking:** it can be associated with any MNE. It includes a reference duration note, a number and a text.
    - **Dynamic Text:** f, p, ff, fz, pp, ppp, etc.
    - **Text:** any textual indication
    - **Harmonics**: round and diamant
    - **Fingers**: to state the finger that has to be used and how
    - **Single event Pedal and holes**: Arpa pedal, clarinet holes, recorder holes, etc.
    - **Percussion markers**: stick, snap pizzicato, maracas, soft, wire, damp, etc.
    - **Strings**: specification of the string to be played.
    - **Change of tonally**: for percussions, etc.
    - **Spaces: vertical and horizontal spaces.** Vertical spaces are sized in terms of music unit. Horizontal spaces are sized in terms of percentage of duration of the MNE at which they are associated with. They are an **only Visual** aspect.
    - Etc.
- **Lyric**: lyric test is comprised of Syllables and has to be multilingual. Each syllable can be associated with a MNE and in some cases with two MNEs (one for the start and one for the end).
    - **Graphic**: font
- **Observations:** They are external observations of the music representation that can be used by the DECODER to present the music in different manners or to access to the music in differently structured manners.
    - **Vertical Section** (also called harmonic section in some cases): a set of MVP, to describe as the orchestra is decomposed, for example the group of strings in the operas. A VS may have a visual representation or not: staff brackets of different type: rounded and graph.
    - **Horizontal Section:** an arbitrary piece of MVP. A HS can be used for identifying segments in the music piece and not for structuring the information. It is a high level descriptor associated with the music by using a starting and ending point.
- **Transposing**: the model has to permit the production of transposed music notation content without losing the initial information and tone. This means that in each barline, it has to be stored the original **Context** and the **Transposed Context**. The transposition should be performed on all music piece of a single part or on a smaller part of it.
- **Metadata:** have to include
    - classical models such as UNIMARK for classification,
    - some and multiple identification codes and:
    - incipit description,

  - tonality,
  - Ambitus,
  - Instrumentation,
  - critical edition and comments, etc.

General remark: each music notation symbol, MNE, observations, markers, HG and structural aspect have to provide:

- **Univocal reference ID** for each music notation symbol for syllables and annotations
- **Reference links from each music symbol** for user navigation and automatic execution

7.1.4 Visual Requirements

Visualization of:

- **Dynamic changes of visual** aspects according to the device on which the visual representation is performed: windows, page layout, Braille page, etc.
- **Justification** information for each measure and music segment, no direct graphical information have to be included into the music notation model to allow the reformatting according to the device size. Possibility to
  - Justify main score and parts, including MNEs without duration, grace notes, etc.
  - justify on the basis of the information contained in each measure (barline) and by considering the visual duration of MNEs, simultaneities, the size of the fonts, etc.
  - impose the values of justification of each measure and simultaneities, and/or
  - delegate the DECODER its definition
- **Vertical Staff Arrangement**: the arrangement of the staves along the vertical axes, distance among parts or staves in general: Possibility to
  - impose the values, and/or
  - delegate the DECODER its definition
- **Line breaking**: possibility to
  - impose the symbol/barline for placing the Line breaking, and/or
  - delegate the DECODER its definition
- **Page breaking**: possibility to
  - impose the symbol/barline for placing the Page breaking, and/or
  - delegate the DECODER its definition
- **Position of symbols around the event based symbols (such as notes, rests, etc.): position and ordering of the symbols. T**heir position can be made on the basis of
  - specific measures given in musical units (the distance between two staff lines), this means that the arrangement of the symbols is fixed and can be changed and imposed only in the phase of authoring; or can be
  - automatically produced by the DECODER by using formatting Rules and table in for positioning symbols such as: order with respect to the note/chord, position and bounding box for each font element
- **Guitar Chords** in different formats: fretboards and textual, if notes (direction of the stem, position of the notes on left and right, etc.)
- **Continuous Bass** in different formats
- **Note Stem direction according to**
  - Formatting Rules, or
  - As imposed by the code
- **Slur, dynamics, tie, etc. direction,** according to:
  - Formatting Rules, or
  - As imposed by the code
- **Slur, dynamics, tie, glissando, shapes** according to:
  - Formatting Rules, or
  - As imposed by the code
- **Beam grouping, beam position, beam direction (up, down or middle) and slope** according to:
  - Formatting Rules, or
  - As imposed by the code

7.1.5 Integration Requirements

- **Synchronization**: the model has to allow realizing synchronizations among the music notation model and other elements of the audio visual file. The synchronization has to be possible between the music notation and the audio, video and also with the music reproduction from SAOL-SASL.
    - For instance establishing that a note will have to be played when a given audio sample is played or a video segment is shown.
    - When the music playing change the execution rate the synchronization of music notation execution has to follow the changing in the rate. This is very useful for education purpose. For instance to see in details the movement of the musician hands and at the same time to read the music notation and hear the corresponding produced notes.
    - This requirement can be useful for the production of several different types of Karaoke.
    - A typical application of synchronization is the realization of a score follower that permits to show the music notation in synchronous with a played music. It can be useful for educational purpose.
- **Playing/Rendering**: The music notation has to be executable in conjunction with SAOL and SASL or with MIDI. Deep relationships have to be identified and defined among the Music notation representation and the coding of audio in MEPG4. This requirement also is related with the synchronization. For example, the visualization of the music notation synchronous with the playing of the music in SAOL or MIDI.

7.2 Music Notation Decoder and Formatting capabilities

One of the main duties of the decoder is the production of the visual aspects from the logical description of the music notation. The requirements of the visual aspects have also a strong impact on those of the Decoder. The decoder can be realized with one or more internal decoders and have to be capable to produce from Logical and Visual music notation information some output according to some parameters, for instance to produce:

- Main score or parts, selection of a part, specification of formatting modes
- Visualization of music on windows or page layout
- Braille notation, on printer and screen
- Verbal description of music notation, spoken music

Other representation of music produced on the basis of the information contained in the Logical and Visual parts, and other produced by the decoder itself at run time.
We can have two kinds of decoder: one which uses only the visual information provided in the music notation data (for low resources platforms) and one which can provide a visual representation starting from the logical information and few visual information using some kind of formatting engine. This is necessary when the content visualization has to be adapted to a specific context.

The formatting rules should be placed on some style file, as well as the fonts have to limited on the graphical parts of the music notation model. Some examples of formatting rule includes:

- **Positioning:** rules for placing music notation symbols around the notes and rests.
- **Ordering or priority:** rules for stating the ordering by which the markers have to be drawn around the notes. For instance, staccato, tenuto, slur, fingering, can be a good order and not the opposite.
- **Justification**: inclusion of justification parameters, liner or logarithmic for each measure or music segment or interval
- **Line breaking**: rules for identifying the point in which the music notation may finish the line. The end of line can be also imposed by the music notation producer.
- **Page breaking**: rules for identifying the point in which the music notation may finish the page. The end of page can be also imposed by the music notation producer.
- **Compression:** rules for depicting simple symbols in the place of their complex semantics in music notation. For example, rules for stating when a generic rests (rest stating that the whole measure is mute) or multirest (rest stating that there are several measures of rest) has to be written instead of placing several detailed rests to reach the corresponding value.
- **Beaming:** rules for grouping notes and chords in beams on the basis of the time signature and the context. Rules for deciding the slope of the beam, rules for deciding if the beam has to be up or down, or in the middle in a multistaff part.

7.3 Music Notation functionality and integration in MPEG

Two fundamentally different layers have been identified regarding the introduction of music notation into the MPEG framework. The first layer is the interface between existing MPEG frameworks and the music notation functionality; defining this interface means adding suitable node and fields, if and when required, to MPEG-4 BIFS (Systems) or defining new description schemes in MPEG-7. The second layer is the one dealing with music notation content itself. This layer requires the specification of a suitable format and associated bitstream and possibly object type identifications and descriptions. There are several ways to edit and otherwise interact with notation formats and content. The following (low level) functionality has been identified as dealing with common western music notation presentation and interaction in the multimedia representation and description contexts of MPEG. Many of these functions would also be useful for a wide set of other notations.
The different methods and features have been initially grouped by purpose, as for defining a general music notation API interface:

Score general information

- get the title, author and other basic metadata.
- get the number of measures
- get the duration
- get an array with part names
- get an array with current bookmarks
- get the list of languages for which the lyric is present
- get the decoder supported visualization formats
- get the link present under a given point

Score Visualization

- select the parts to be shown:
  - 1 part → single part view,
  - some parts → main score with the selected parts,
  - 0 parts → full main score

  the parts may be identified by name, by position, by id, by group (violins, flutes)
- select the visualization format: CWN, Tablature, ...
- set for each part if the lyric has to be displayed or not and possibly set the language to be used (English, Italian, French…)
- set the one/two page view
- zoom factor +/-/set/fit
- enable/disable hyperlinks view
- set the background color
- set the default music color
- set the default lyric color
- set the default text color
- set the hyperlink color
- set the visited hyperlink color
- set the overall transparency
- set the background transparency
- ...

Score Browsing

- page +/-
- go to page
- go to measure
- go to beat
- go to time
- go to bookmark
- go to first page
- go to last page

- set the turn page animation (however it may be done with animation schemes)

Score modification
- Transpositions
- Insert/delete personal textual annotations
- Add/delete bookmarks
- Insert fingering, bowing, breath indications, attention marks, caution accidentals, dynamics, tempo markings, specific marks like piano pedal, cue notes…

Score Synchronization
- synchronization with a SASL/MIDI source
- synchronization with a natural Audio/Video
- synchronization with animations (singing head, synthetic ballets, operas)
- synchronization with Text-to-Speech (narrating voice, text-to-sing)
- set synchronization granularity (measure, note)
- set the synchronization position showing mode: a ball showing the current playing measure, a bar indicating the execution position, coloring current playing notes, no indication or merely generate an event saying the position
  coordinates of the measure/note currently playing.

This overall functionality should be possibly split into three main groups in relation to an MPEG-related block scheme:

1) functionality supported by the existing MPEG frameworks
2) functionality specific to the notation player tool and its interface with the systems layer
3) functionality specific to notation authoring tools and/or left open to implementers in players.

For instance, all (or nearly) the synchronization points can be satisfied by the MPEG-4 systems, once the timing/AU issue for music notation is defined. This built-in synchronization between notation and other media is one of the major issues for integrating notation in MPEG and not having it as a standalone format. Moreover, some of the information related to the musical score is already part of MPEG-7 descriptors, as far as those descriptors can be put in tight relationship with content in MPEG-4 streams.
The line between groups 2 and 3 is more "subjective", as it defines what is the functionality to be delivered to the user, associated to a normative behavior, and what is to be left open. Please note that this is at the Systems layer, one layer above the notation format. It deals with interaction with other media and with the end user, and it also and especially deals with existing and future applications and their specific functionality.
It is very important that 2) includes all the interactivity features the user may desire for user applications in general and that 3) includes the features that are typically part of specific applications, like authoring capabilities, interoperability with custom formats and so on. Typical interaction points (like e.g. node fields) to be “exposed” to users may include: selection of the part(s) to be show, selection of the visualization format, zoom in/out, go to measure/time/bookmark, transposition, etc. The remaining features may be supported by authoring tools and should be supported in the notation format. To be further noticed that some of these parameters (like those related to colors, general formatting, etc.) are intended to be supported by players in standard way, while some others (like e.g. paging options) are typically decoder implementation-dependent (i.e. platform-dependent) and should be left open to implementers.

The interaction with music notation (score modification) may be performed using two fields in the Music Notation node:

1. one string field indicating the command to be performed when the user clicks on the music score.
2. one 2d coordinate eventIn field indicating the position where the command have to be applied.

In this way using the TouchSensor node and routing the clicking position to the MusicNotation node the commands can be performed.
The command string field format may be standardized for some general/basic commands (i.e. for transposition, to add some basic symbols) and left open for more specific needs.

7.4 Interface and Synchronization issues

7.4.1 Access Units for music notation

As defined in ISO/IEC 14496-1, an Access Unit (AU) is an individually accessible portion of data within an elementary stream. An access unit is the smallest data entity to which timing information can be attributed at the Systems (i.e. also media synch) layer.

For Music Notation an Access Unit could be:

- a set of music notation objects starting all in a temporal window such as a measure (a column in the score)
- a set of music notation objects on a particular instrument (a row in the score)
- a set of events starting all in a temporal window
- a score page (whose definition depends on the viewing area size)

The definition of an Access Unit is necessary to provide the streaming capability associated to music notation content; it is also needed for score synchronization and for score browsing.

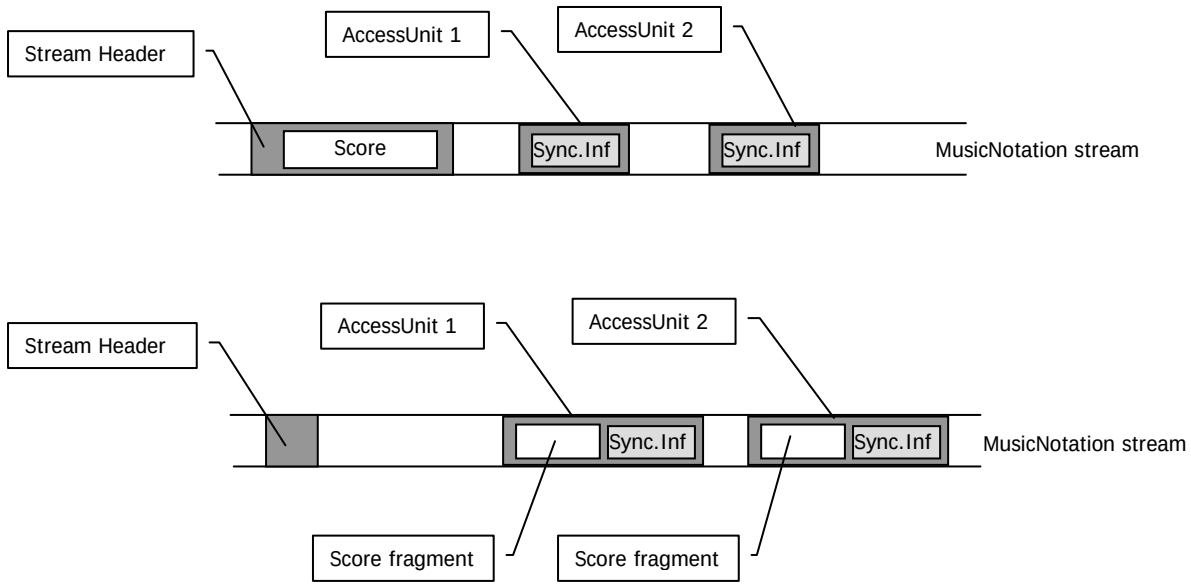


Figure 1 – Music notation stream and access units

The first problem is that music notation, like a text, is not by itself time-stamped. None of the currently used music notation formats includes such timing information.
Different time stamps are to be considered:
- The time corresponding to the time of the musical event (the time at which the musical event occurs in one of the several possible corresponding audio items)
- The time at which the musical event (in its graphic form), and other elements like staves, clefs, tempo markings and all graphical elements must be available for decoding and composition.

Music notation could be marked with some sort of "virtual-time" markers, for example based on measures, beats and so on... This kind of information can easily be deduced from current music notation formats. The HyTime model could also be considered here.
But for Access Units, we need a real-time marker, that is time stamps expressed in something like seconds or milliseconds. The same is true for synchronizing musical events with other streams.

To cope with the problem of synchronizing with different performances, there are two solutions:

- To produce as many time-stamped music notation streams as it is needed for synchronization.
- To time-stamp the music notation stream on the server side on delivering time, on the basis of virtual-time markers, and on the basis of a synchronization description available on the server side.

This specific problem can be considered a problem of authoring. Its solution may be implemented by defining two independent chunks of information, one dealing with all the music notation objects and the other with synchronization of music notation events with other media events. In this way, a particular score (the notation) can be associated with several different performances, and a particular performance can be associated with several different scores. Scores can be delivered independently of audio files, and vice versa.

Regarding the encoding of music notation into the bitstream and in the AUs, we could try to define something similar to the SASL/MIDI case. SASL information can be encoded in MPEG-4 in three different ways:

1) part (or all) of the score file can be included in the header (object type specific info) and lines shall be time-stamped
2) score lines can be inserted in AUs with timestamps (so I can deliver certain AU "events" for a later use)
3) score lines can be inserted in AUs without timestamps, in which case the decoding time stamp of the AU is the reference time and events are scheduled immediately

Each of these three possibilities is associated with best use cases. For notation information the same, or similar, structure may be followed. Regarding time and synchronization information with other media, this could well be a different chunk of information in the header (if known in advance) or separated type of data in the AU, as mentioned above.
These three cases separate situations in which content is more or less available (reasonably) before the fruition of the same from those situations when this is not possible/reasonable. Say, if a MIDI or SASL file is ready, it can be put in the header, if instead the author is playing a MIDI or SASL device in real-time and sending it in a stream, this must be done through AUs. Furthermore, if all the content (e.g. the complete SASL file or, in this case, the complete score and notation info of a symphony) is ready for use and broadcasted, it is not wise putting this all in the header; users may actually switch on "TV" at any time, and the server will be required to periodically retransmit at reasonable intervals (carousel) all that chunk of info. If instead the author is producing an mp4 file for a DVD, that's another case; and so on. (cases when the score is of course available before, but the timing info to synch with other media is not, can be easily imagined).
Moreover in the case of music notation the synchronization information may be provided also using a specific field (an array) of the MusicNotation node. In this way the synchronization info is carried with the BIFS stream and will allow the synchronization of the same music notation stream with different audio sources.

Talking about AUs means then talking about defining the possibility to break content down into AUs in a meaningful way, other than defining the format of the content. Each AU has associated timing information (the decoding time stamp/composition time stamp), but in some cases, like SASL, the AU may contain additional time stamping for the SA scheduler so that it allows, if possible, to include e.g. in one AU at time 10 all the events from 10 to 13, and so on. Notation, and its synchronization with other media, can be a very similar case. A graphical representation is shown in Figure 1.

7.4.2 Additional issues on synchronization with other media

Still on synchronization, MPEG-4 also provides the FlexTime nodes allowing the synchronization of streams coming form different servers. The TemporalGroup and TemporalTransform nodes can be used for this task. A way to use MusicNotation node with this nodes should be identified and it is related to the possibility to refer to a segment of the score (i.e. od:10#page2).

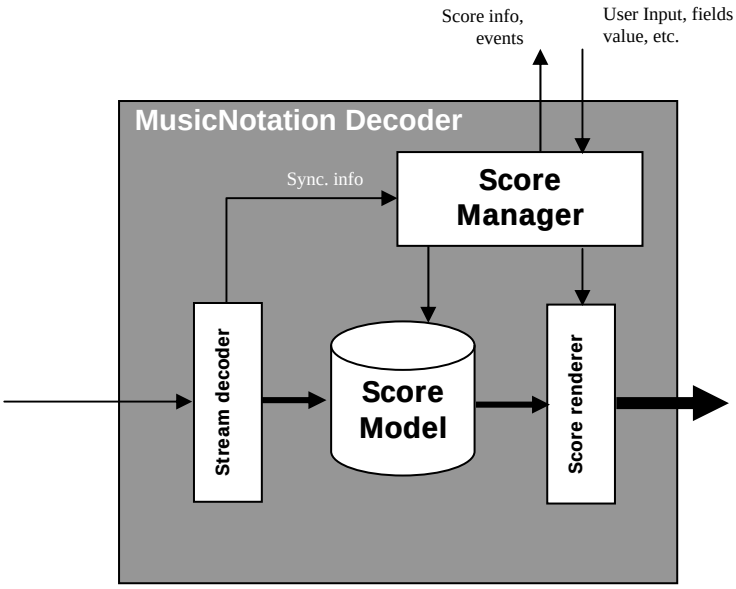


Figure 2 – A possible structure of a Music Notation decoder

In Figure 2 a possible scheme for a Music Notation decoder is reported. The **Stream Decoder** receives the stream and decodes it and it puts the score information in the Score Model and delivers synchronization info to the **Score Manager** component. The Score Manager manages:

1. the presentation of the proper pages at the proper time controlling the **Score Renderer** module,
2. the modification of the score due to the user interaction,
3. the generation of events and access to the score information.

The **Score Renderer** renders a score page using the **ScoreModel** and the visualization info supplied in the scene definition or from other nodes through routes (i.e. colors, transparency, etc.) and produces the image to be delivered to the compositor.

A particular case is the synchronization/relation with SASL/MIDI; some specific adaption tools can be considered to:

1. transform a MIDI/SASL stream to Music Notation to only visualize the sequence of notes in the stream
2. transform a MusicNotation stream to MIDI/SASL to only listen the music automatically produced by the score

MIDI and SASL, chunks of data in the MPEG-4 SA stream, contain in fact structured event information that can be converted to a subset of the information necessary for the music notation support (at the audio layer only). In a way conceptually similar to text and additional bookmarks driving a face (lips) animation, MIDI/SASL streams may be used to drive the music notation decoder tool (as well as a subset of the music notation information may be made compatible to the SA scheduler engine to provide structured synthesized sound feedback from a musical score). According to the overall functionality, quality and flexibility these conversion boxes (see Figure 3) may be completely non normative, completely normative, or just specify some basic rules leaving the conversion process open to implementers. This feature is in any case highly desirable, since it will permit, for instance, to immediately play back music notation by a MIDI device or a SA decoder with associated orchestra, as well as graphically display the notes included in a MIDI/SASL bitstream.

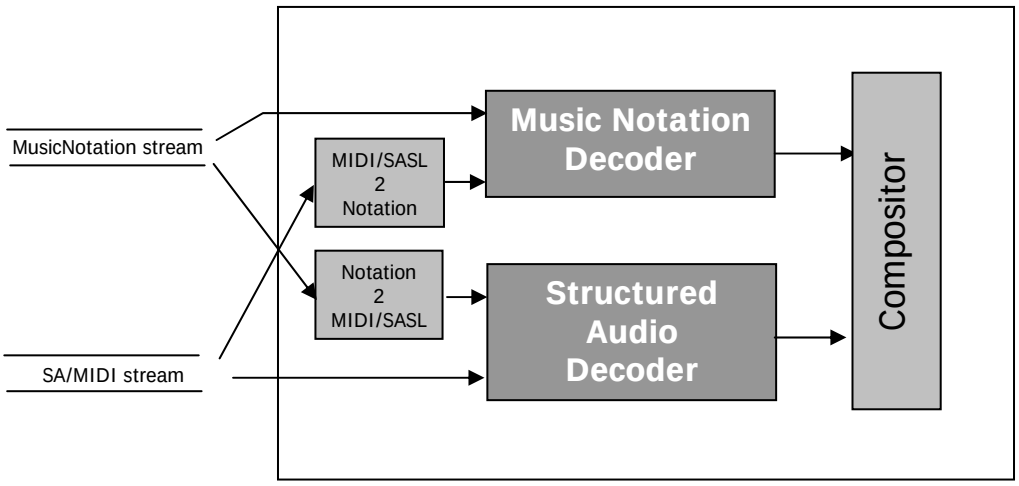


Figure 3 – Relation of music notation with Structured Audio

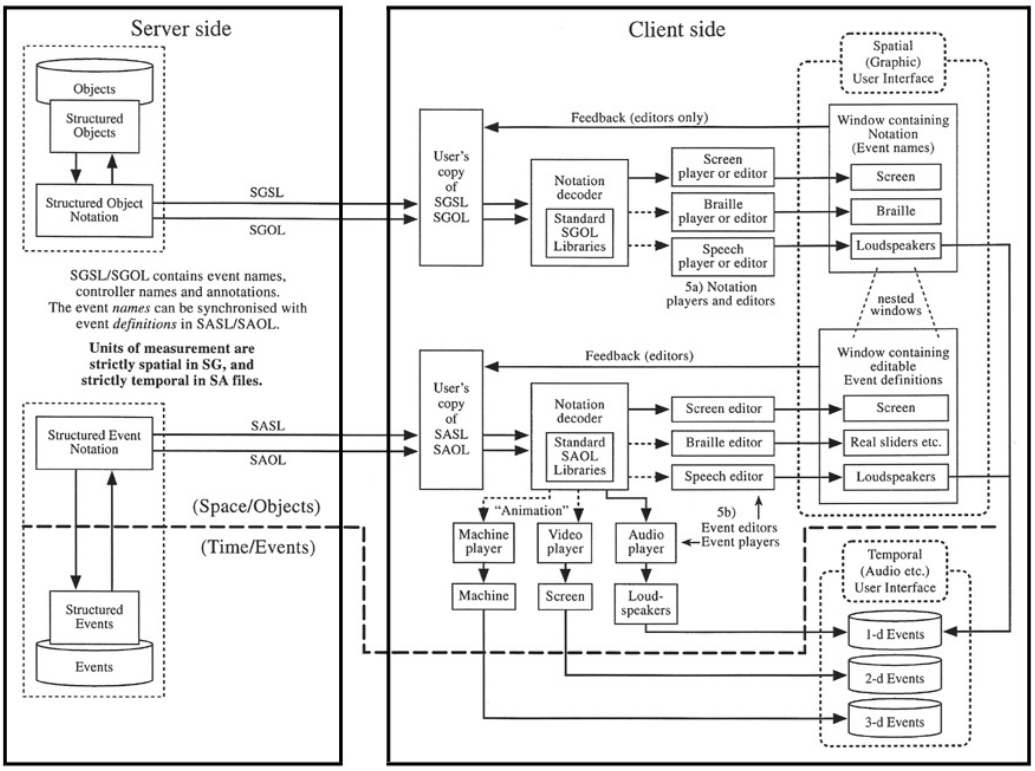
7.5 Additional analogies with Structured Audio: a proposal

One of the possibilities that have been analyzed is to further take as "template" the MPEG-4 Structured Audio toolset to create a parallel Music Notation structure with timing and graphic information to deal with the visualization of music score not bound to a particular music notation framework as the common western music notation (CWN). Analogies mainly lie on the capability to map events to 1-d or 2-d sonic or visual corresponding media representations (sound for SA, graphic scores for Notation).
This proposal has been supported by James Ingram, and consists of two languages:

1. the Structured Graphics Score Language (SGSL) is used to model the logical and visual part of the score, the model is composed of Objects (notes, rests, etc.), which have a graphic position in the score and Controllers (crescendo, diminuendo, slur, etc.), which manage a set of Objects.
2. the Structured Graphics Orchestra Language (SGOL) is used to manage the rendering of Objects specifically for a certain notation framework. A standard library for CWN on the client side will avoid the delivery of SGOL for CWN, which is envisioned to be frequently used (the same as standard core opcodes exist in SA for most frequently used processing functions). Moreover the SGSL objects can be bound to the related SASL audio event for synchronization purposes.

The overall scheme of this proposal is summarized in the diagram below.

If it is maybe a little premature to talk about formats and such detailed architectures at this stage; it is nevertheless useful to exercise potential overall solutions in order to better analyze requirements and problems that may arise from the proposed integration of music notation in MPEG. In particular, this proposal highlights even more some relationships with existing MPEG tools and it contains an approach to information partitioning for music notation that reflects those tools and structure (including libraries for standard notation symbols).



7.6 Overall process including Music Notation

The overall process from the producer to the consumer is depicted in Figure 4 where a MusicNotationXML language is used to support interoperability with other music notation formats in the same way as XMT is interoperable with SMIL, SVG, VRML. A MPEG-4 authoring tool can integrate the music notation info with the scene description and the scene content and produce the binary representation of MusicNotation and the synchronization information to be stored in an mp4 file. It is the task (as usual) of the authoring tool to synchronize the MusicNotation presentation with the overall scene which may contain natural audio/video playback, animations, etc. etc. Finally the presentation produced in the authoring phase may be sent in broadcast to TVs or streamed through Internet or accessed off-line.

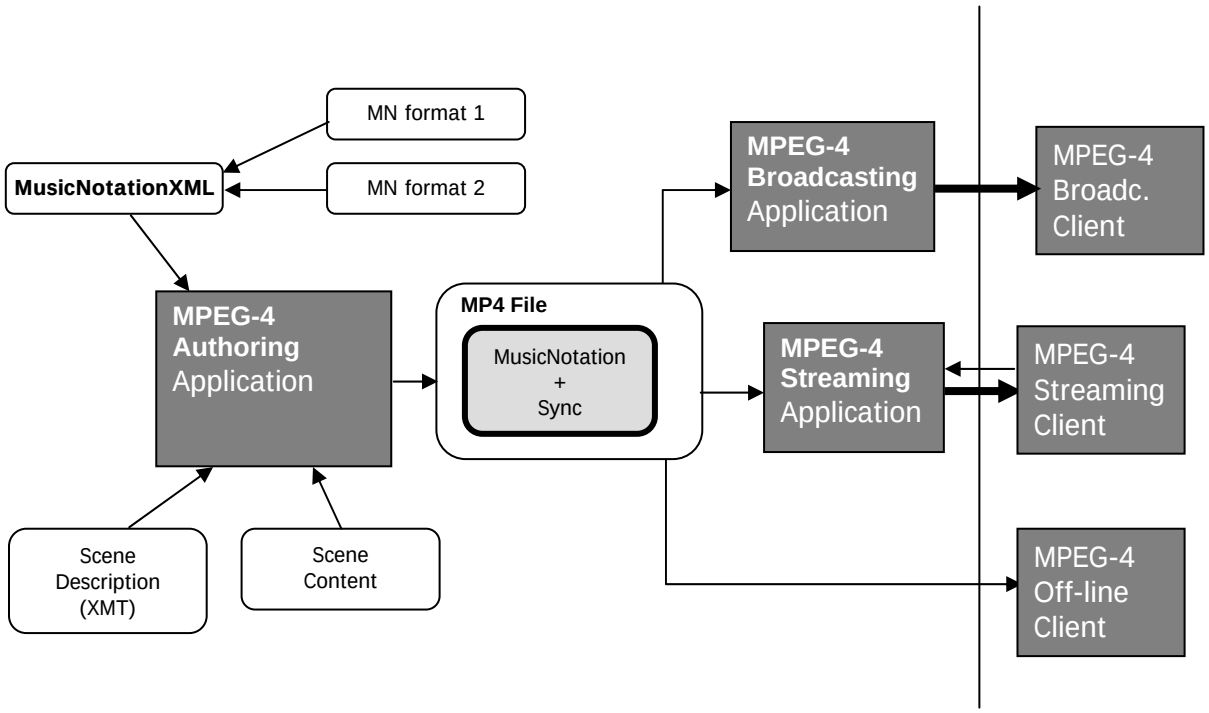


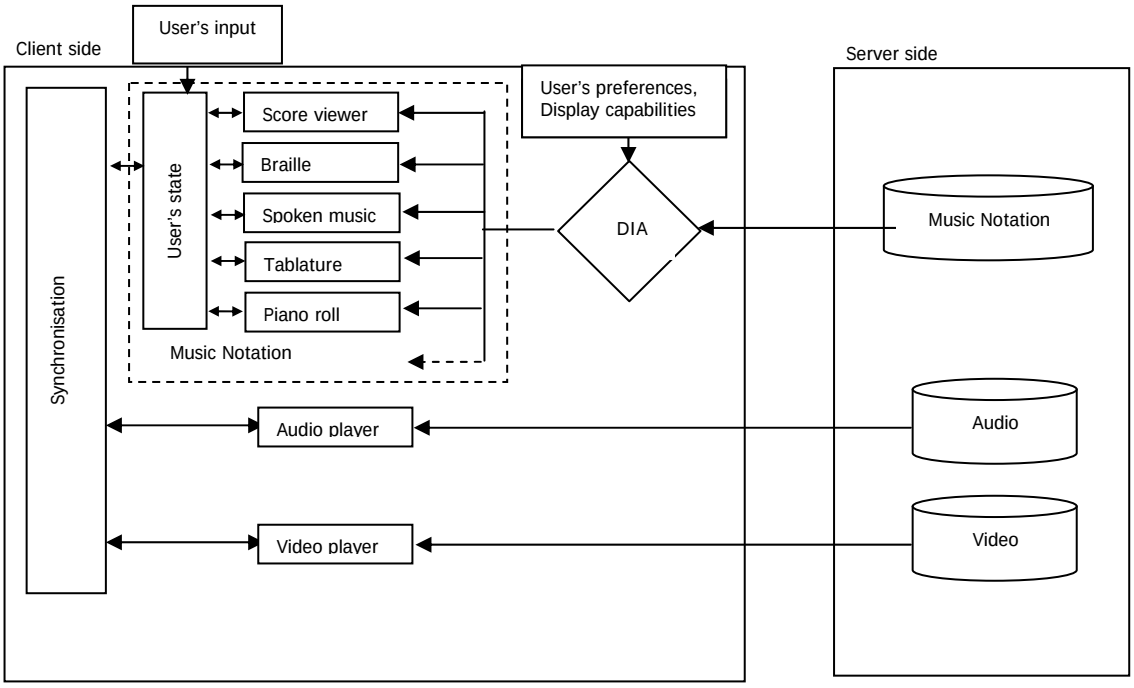
Figure 4 – Overall process

7.7 Different decoders and viewers

An important requirement of Music Notation is that different decoders, particularly for visually impaired persons, should be made available on the client side.

The different views must be “synchronized”, that is to say that any user’s interaction with one of these views (like transposition, annotation…) must be reflected in all other possible views available. To this end, a set of compatible functions must be made available for all decoders, and user’s interaction must be implemented by the means of a common module.

These different views must also be made available on the client side conforming to user’s preferences and capabilities of display.
In an MPEG 21 environment, this adaptation could be done by using the MPEG 21 DIA mechanisms.
In an MPEG 4 environment, the adaptation could only be made internally to the Music Notation node.



7.8 MPEG 7 relationship: MPEG-7 Content Description Segments and Music Notation

It would be interesting to study how MPEG 7 concepts and ideas could be used to model music notation. As suggested by Jacques Steyn, with reference to music, music in audio format obviously belongs to the MPEG-7 audio related segments of Content Description while music notation belongs to a visual representation. MPEG-7 defines a Graph DS "that allows the representation of complex relations between segments. It is used to describe spatio-temporal relationships, between segments that are not described by the tree structures." (José Martinez, 2001:32, Overview of the MPEG-7 Standard)

8 The MUSICNETWORK and MPEG

Since the beginning, one of the main aims, if not “the aim” of MUSICNETWORK has been to focus on stimulating the realisation of widely adopted formats for music notation representation. These formats, which must be integrated with multimedia applications and models, will deal with the needs of all the relevant actors (including publishers, music editor producers, copyists, integrators, etc.) involved in the realisation and the distribution of an "interactive" multimedia music piece. Music notation representation is an important issue, and an open standard format allowing exchange and cooperation with other multimedia formats still does not exist. We should not be limited to the applications only related to the printing process. It is clear that music notation needs to be and will be in the future accessed more and more using different kinds of devices, from the PC to Tablet PCs and UMTS terminals, from the classical printed sheet music to the electronic lectern.

Great part of the work described in this document has not been economically supported by the MUSICNETWORK. The MUSICNETWORK has performed only the coordination of the reported activity.
On this regard, the MUSICNETWORK has worked with the support of many working-groups, including Music Notation, Standards, Imaging, Accessibility, Education, Library, etc.
The actual work has been performed with the coordination of a large number of experts coming from several areas of Europe and worldwide. Currently, the MUSICNETWORK has more than 850 participants, from more than 50 different countries.

The first task was the identification of the requirements for the integration of music notation with multimedia applications. This work has been performed with the support of several experts at the MUSICNETWORK open workshops by means of the discussion forums and the mailing lists provided by the MUSICNETWORK portal. This analysis has been performed with considerations of the past experiences of several European Commission projects, including CANTATE, MOODS, WEDELMUSIC, IMUTUS, PLAY, PLAY2, etc., that worked in the area of music notation, and in some cases, on the integration of music notation with some multimedia content and features.

The second step was to study the state of the art technology in the area of music, computer music and electronics, to better understand the music notation/representation formats, their integration, and aspects on all the WGs involved in the MUSICNETWORK, and their usages in multimedia applications. These activities have been described in a number of deliverables and reports of the MUSICNETWORK that have been downloaded by thousands of participants from the MUSICNETWORK web site. It is evident that this work has received very strong attentions and interests and it has been strongly appreciated by all related communities.

This second step has allowed us to understand the state-of-the-art and the real needs of the users and of the companies that produce music and computer music applications for the market, mainly in the areas of education, entertainment, content distribution, archiving and cultural valorisation, electronic consumer equipments (such as i-TV, PDA, cellular phones, and others), etc. At the same time, the experts of the MUSICNETWORK have identified the major problems that are preventing and/or limiting the exploitation of the present technical and technological solutions in those applications.

As mentioned before, one of the main limitations for the full exploitation of music notation/representation integrated with multimedia, is the lack of a common standard for representing the notational information. On the other hand, the presence of a standard is not the unique problem since presently there are some *de facto* standards that cannot be used and are not used for solving the above mentioned problem. These include representations from FINALE, SIBELIUS, etc. In fact, in most cases, these *de facto* standards are capable of supporting more than 95% of the global production of music notation pages. However, they remain unacceptable and incapable to be fully exploited in multimedia music applications. Another interesting case is that of the XML-based music notation formats. In the last 5 years, we have seen about 15 different XML-based models proposed by several groups and companies. Among these only MUSICXML of Recordare has gained interest demonstrating a quite interesting interoperability with several applications, including in some measure, FINALE and SIBELIUS. In this case, the MUSICXML remains a subset of their models, and it is not capable of modeling multimedia music concepts. The same problem is evident from the potential new

effort for starting a standardisation process in IEEE. If the idea is to model only the music notation, the format will remain without a real usage in the area of multimedia music. This is confirmed by other efforts in the past in which we have seen other proposed standards as interchange formats for music notation such as NIFF and SMDL (SMDL was an ISO draft, now cancelled, to produce a standard on music notation).

On the basis of this experience and rationales, the effort of the MUSICNETWORK has been concentrated to a different direction, that can lead in a unique step to create a standard for the representation of music notation which allows complete integration with a multimedia framework. The path to reach this goal has been identified in realizing a standard for music notation/representation within the MPEG framework, with a special attention to MPEG-4. MPEG-4, which is the emerging standard on multimedia distribution, is composed by several parts, some of which are already supported by Microsoft Media, Quick time, Interactive TV, and of many digital TV distribution channels, etc.

The scope of this report is to document the activity performed by the MUSICNETWORK to coordinate this effort of bringing new functionalities regarding the music notation and representation into the multimedia world. The work performed in MPEG has been identified as a mean to reach the above mentioned objectives. Other details regarding the work performed to this aim can be found in the documents produced and listed at the end of this document. Please note that the activity of MUSICNETWORK in the MPEG is not devoted to the simple consolidation of technology into a standard, the activity of MUSICNETWORK as described in this document, is of harmonization and coordination of the activities of several different actors (research institutions and industries and content providers) to take into account and make available a large set of NEW REQUIREMENTS and FUNCTIONALITIES into the present state of the art of the products, thus opening the way to new market for a large number of new products with music notation in the area of culture and education, mainly. Moreover, those requirements have been identified by the experts coordinated by the MUSICNETWORK according to the users' needs.

MPEG is the largest standardisation group that works on multimedia coding. It is a working group of ISO (International Standard Organisation, WG 29) and it is the producer of all the MPEG standards: MPEG-1 (coding for distribution at small bitrates, including the core of the mp3 format), MPEG-2 (audiovisual coding for digital TV and higher bitrates), MPEG-4 (multimedia coding for audiovisual objects, to be used also by satellite distribution, etc). At the forum, there are participants from all the major companies including IBM, MICROSOFT, HP, SAMSUNG, SONY, THOMPSON, PANASONIC, YAMAHA, SANYO, PHILIPS, SHARP, etc. under the umbrella of their respective National Bodies. Overall, MPEG includes more than 400 major companies involved in Consumer Electronics devices and technologies, and all the major research centers in the area. At each meeting, typically about 300 partners are represented. It is to be noted that a standard accepted in MPEG is an ISO standard accepted by National Bodies world-wide. Any activity and document produced by an MPEG meeting is accepted in plenary session and thus it obtains acceptance by the important and large companies, which are present at the meeting, such as those mentioned above and many others. This is why MPEG has been considered by the MUSICNETWORK as the best forum to propose and to create a Music Notation/representation standard with multimedia integration.

The MUSICNETWORK activity aiming at the integration of Music Notation in MPEG started in May 2003 (month 10), with the elaboration of a joint proposal for the MPEG meeting in July 2003, in Trondheim. The MPEG group has agreed on setting up an Ad Hoc Group, which is, in MPEG parlance, a specific group aimed to study a particular topic. In this case, the Music Notation and its possible integration in MPEG. A mailing list (a reflector in MPEG parlance) has been setup, together with a web site to support this activity. The chairs of this AHG have been designated by MPEG to be Paolo Nesi (also chair of the MUSICNETWORK) and G. Zoia of EPFL, Switzerland. The result obtained so far has been really exceptionally good, as commented by MPEG and ISO:

As a general assessment of the work performed by MUSICNETWORK for integrating Music Notation in the MPEG framework, *Leonardo Chiariglione (Digital Media Strategist and chair of MPEG and of DMP and of concertation group on DRM of the EC)* have publicly stated at the MPEG Forum in front of hundreds of experts of the world (in Munich 2004) and reported in an email of appreciation to Nesi and Zoia that

«*The topic of this proposed MPEG work item and the approach followed by the Ad Hoc Group are an excellent demonstration of the synergy that can be created between advanced research topics and*

standardisation as carried out by MPEG in integrating existing technology with potential industrial impacts. In the case of music notation the added value is in bringing to life application fields where education, culture and entertainment blend.

Recently the colocated Symbolic Music Representation Ad Hoc Group and MUSICNETWORK meetings resulted in the production of the Draft Call for Proposal on Symbolic Music Representation. The work is being done at a high level of professionalism and great enthusiasm and can sure become a guide for other future topics.

I congratulate on your achievements so far and look forward to seeing your successful reaching of the next milestones."

The above statement has been also reported and public available in the Assessment document of the MUSICNETWORK, DE6.1.1 which is publicly available on the web site.

The activities performed are all public. All the documents, test cases, web pages, etc., and even emails exchanged, can be accessed via free web access. Please see at the end of this document, for a complete list of references and relevant http links.

8.1 Motivations and Rationales

Currently, the music market is characterized by several products that present some forms of integration between music notation and multimedia. Examples include music education related areas (integrating multimedia), music management in library (integrating with multimedia for navigation and for synchronization), karaoke (synchronization based integration), etc.

The integration of music notation in MPEG can cover the needs of these tools and can provide much more features by adding interoperability, porting them on i-TV and on mobiles, supporting scalability of format complexity, etc., permitting to these tools to integrate the powerful MPEG model for the multimedia modelling and play.

The integration of music notation with MPEG can open the way for a large number of new applications and markets. Examples include multimedia electronic lecterns, music education via i-TV, multimedia content integrated with music notation on piano keyboards, mobiles, PDAs, etc. This action could enlarge the present market for music notation, nowadays limited mainly to music-sheet production; it may open the path to create very interesting new applications, and also renewing existing applications that already use some integration between multimedia and music notation.

An AHG has been created by MPEG on the basis of the request from the members of the MUSICNETWORK to explore further this integration.
The MUSICNETWORK had previously started an activity on collecting detailed requirements for the music notation and its integration on multimedia. These requirements include: format analysis, and activities for defining the requirements and the functionalities of the new standard.

The requirements developed have been taken into account for producing a call for technology to include a Music Notation Model into MPEG. In addition, they are used for determining the criteria for assessing the proposals, to verify that the potential standard meets the specified requirements, and to specify the conformance point for the standard. This will be applicable to all tools that use similar technology, involving:

- Music notation format; and
- Music notation decoder (formatter of music notation, piano roll, Braille production, etc.)

MUSICNETWORK participants as well as external companies, institutions, organisations and research centers are very keen and interested in working with the AHG in MPEG, in an effort to guarantee that their needs will be included into the MPEG standard, so that their tools that will be compliant with the standard. There is also a strong market impact since these tools could be integrated in a wide range of multimedia applications for I-TV, PDA players, Piano Keyboards, etc., and integrators will have the possibility of

selecting among several different standard implementations of Music Notation Decoders, all compliant with the same Music Notation format.

The MUSICNETWORK has started an activity on collecting detailed requirements for the music notation and its integration on multimedia. These requirements include: format analysis, and activities for defining the requirements and the functionalities of the new standard.

8.2 Expressions of Interests

A request for Expressions-of-Interest (EoI) has been produced, in order to show the interests of music community in this topic. A wide range of EoI has been received, from the academic sectors as well as from industries. From the academic world, EoI have been received from institutions including libraries, research centres and universities. From industries, EoI have been received from companies dealing with audio processing (automatic generation of scores from audio), or dealing with optical recognition of music scores, or even dealing with e-learning and education.

The AHG has also solicited EoI to the ongoing activity including informal letter to the AHG chairs and formal input to MPEG. Several formal EoI have been received, so far, from the following companies:

- GIUNTI, educational and multimedia, Italy
- NOTISSIMO, Music publisher, France
- EXITECH, Music Editor Builder, archive and library area, music education, Italy
- SIGTONE, Music educational tools, France
- SPACE SPA, Music entertainment tools, OPENDRAMA, Italy
- CAPELLA, Music Notation editors, CAPXML, Germany
- RIGEL, Music specialisation systems, Italy
- MUSIC IMAGINING Ltd, music document digitization and restoration, UK
- FNB, Music Notation tools for Accessible music, The Netherlands
- ARCA, Music notation tools for impaired, Italy

Some Institutions:

- DSI-DISIT, University of Florence, Florence, Italy
- University of Osnabruk, Germany
- MTG, UPF, Barcelona, Spain
- IRCAM, Paris, France
- ICSRiM, University of Leeds, Leeds, UK
- NEWI, Institute of Higher Education, UK
- EPFL, Lausanne, Switzerland

Other companies interested in this activity and participating to the general discussion (either by participation to the Munich meeting or to the previous MUSICNETWORK meetings related to the AHG activity), which may be interested in a call for proposals include:

- SUN Microsystems GmbH, Germany
- Digital Sheet Music, Denmark
- Sheet Music Now, Germany
- Music Group Netherlands, The Netherlands
- IBM Deutschland, Germany
- TISCALI, Italy
- Recordare, USA
- MakeMusic, CodaMusic, USA
- coolHeads, USA
- MIDI.ORG, USA
- Dancing Dots Braille Music Technology, USA
- MIDIILLUSTRATOR,
- Middle-C Ltd, UK
- Etc..

8.3 Steps towards the standardization of the SMR

The MPEG Audio Group will continue, through the work of the SMR AHG coordinated by the MUSICNETWORK, in the process of standardization of the SMR with a plan specified until the establishment of the Reference Model 0 (RM0), given that the Call will be successful. During the 72nd meeting in Busan two important decisions have been taken. The SMR is an official subpart of the Audio standard; secondly if and how a Core Experiment process is to be defined and run in the following meetings.

Thus, the ISO/IEC sc29/wg11 working group, better known as MPEG (Moving Picture Experts Group of ISO/IEC) has recently approved a new work item for the standardization process of Symbolic Music Representation (SMR); this process has been stimulated, promoted and developed in the course of the DigiCult project MUSICNETWORK (http://www.interactivemusicnetwork.org/ , http://www.interactivemusicnetwork.org/mpeg-ahg/index.html).

Following an open Call for Proposals, the WEDELMUSIC format (developed in the past from another IST FP5 project: http://www.wedelmusic.org/) has been selected as MPEG SMR Reference Model 0. Fore coming planned integrations will cover applications such as Braille representations of music, Spoken Music and Korean Music.

The creation of MPEG SMR is very important since it opens the way for a large set of new applications and research/development activities in several domains like entertainment, edutainment, infotainment, etc.: from education and distance learning, to rehearsal and musical practice at home, from consumer electronics such as set top boxes for interactive TV to personal computers and mobiles systems. Examples of applications include: Interactive music tutorials, Play training, performance training, Ear training, Compositional and theory training, Multimedia music publication, Software for music management in libraries (music tools integrating multimedia for navigation and for synchronization), Software for entertainment (mainly synchronization between sound, text and symbolic information), Piano keyboards with symbolic music representation and audiovisual capabilities, Mobile devices with music display and editing capabilities.
The first Working Draft of the specification document of the standard will be available in July 2005. The schedule of the standardisation process has been approved at the 72nd MPEG plenary meeting (Busan, Korea, April 2005) with more than 300 participants from all over the world, including several of the most important players in the multimedia scene http://www.chiariglione.org/mpeg/meetings.htm.

The SMR schedule defined inside the MPEG Audio Subgroup, plans to proceed by means of Core Experiments to arrive at the CD (Committee Draft) in April 2006 and from this to the International Standard by the ballot mechanism involving all the National Bodies internationally. Stay tuned on MPEG and MUSICNETWORK to keep informed about future meetings (the next one is in Poznan, Poland July 24-29, 2005).
See http://www.chiariglione.org/mpeg/meetings/busan05/busan_press.htm for the official press release of ISO/MPEG.

9 Bibliography

- Anderson, D. P. and Kuivila, R. (1991). Formula: A Programming Language for Expressive Computer Music. IEEE Computer, pp.12-21, July.
- Bellini, P, Della Santa, R., Nesi, P. (2001). Automatic Formatting of Sheet music. Proc. of the First International Conference on WEB Delivering of Music, WEDELMUSIC-2001, IEEE Press, 23-24 November, Florence, Italy, pages 170-177.
- Bellini, P., Bethelemy, J., Bruno, I., Nesi, P., Spinu, M. B. (2003). Multimedia Music Sharing among Mediateques, Archives and Distribution to their attendees. Journal on Applied Artificial Intelligence, Taylor and Francis, in press, www.wedelmusic.org.
- Bellini, P., Fioravanti, F., Nesi, P., (1999). Managing Music in Orchestras, IEEE Computer, pp.26-34, September, http://www.dsi.unifi.it/~moods/.
- Bellini, P., Nesi, P. (2001). WEDELMUSIC FORMAT: An XML Music Notation Format for Emerging Applications. Proceedings of the 1st International Conference of Web Delivering of Music. 23-24 November, Florence, Italy, pages.79-86, IEEE press.
- Bellini, P., Nesi, P., Spinu, M. B. (2002). Cooperative Visual Manipulation of Music Notation. ACM Transactions on Computer-Human Interaction, September, 9(3):194-237, http://www.dsi.unifi.it/~moods/.
- Blostein, D. and Baird, H. S. (1992). A Critical Survey of Music Image Analysis. Structured Document Image Analysis, (H. S. Baird and H. Bunke and K. Yamamoto, ed.), Springer Verlag, NewYork, USA, pp.405-434.
- Blostein, D. and Haken, L. (1991). Justification of Printed Music. Communications of the ACM, 34(3):88-99, March.
- Byrd, D. A., (1984). Music Notation by Computer. Department of Computer Science, Indiana University, USA, UMI, Dissertation Service, http://www.umi.com.
- CANTATE project, (1994). Deliverable 3.3: Report on SMDL evaluation, WP3.
- CAPXML: http://www.whc.de/capella.cfm
- Dannenberg, R. B. (1990). A Structure for Efficient Update, Incremental Redisplay and Undo in Graphical Editors. Software Practice and Experience, 20(2):109-132, February.
- Dannenberg, R. B. (1993). A Brief Survey of Music Representation Issues, Techniques, and Systems. Computer Music Journal, 17(3):20-30.
- F. Pereira, and T. Ebrahimi (Eds.), The MEPG-4 Book, IMSC Press, 2002.
- Good, M. (2001). MusicXML for Notation and Analysis. In W. B. Hewlett & E. Selfridge-Field (Eds.), The Virtual Score Representation, Retrieval, Restoration (pp.113-124). Cambridge,MT: The MIT Press.
- Gourlay, J. S. (1986). A Language for Music Printing. Communications of the ACM, 29(5):388-401, May.
- Heussenstamm, G. (1987). The Norton Manual of Music Notation, Norton \& Company, New York, London.
- Icking, I. (1997). MuTEX, MusicTEX, and MusiXTEX Beyond MIDI - The Handbook of Musical Codes, (E. Selfridge-Field, ed.), The MIT Press, London, pages 222-231.
- IMUTUS: http://www.exodus.gr/imutus/
- MPEG SMR AHG web page: http://www.interactivemusicnetwork.org/mpeg-ahg
- MUSICALIS: http://www.musicalis.fr/
- NIFF Consortium, “NIFF 6a: Notation Interchange File Format”, 1995.
- OPENDRAMA: http://www.iua.upf.es/mtg/opendrama/
- Pennycook, B. W. (1985). Computer-Music Interfaces: A Survey, ACM Computing Surveys, 17(2):267-289, June.
- Pope, S. T., (1991). The Well-Tempered Object: Musical Applications of Object-Oriented Software Technology. MIT press.
- Rader, G. M. (1996). Creating Printed Music Automatically, IEEE Computer, pp.61-68, June.
- Ross, T. (1987). Teach Yourself. The Art of Music Engraving. Miami, London: Hansen Books.
- Selfridge-Field, E. (Ed.). (1997). Beyond MIDI - The Handbook of Musical Codes. London, UK: The MIT Press.

- Sloan, D. (1997). HyTime and Standard Music Description Language: A Document-Description Approach. E. Selfridge-Field (Ed.), Beyond MIDI - The Handbook of Musical Codes. (pages 469-490). London, UK: The MIT Press.
- SMDL ISO/IEC, Standard Music Description Language, ISO/IEC DIS 10743, 1995.
- Smith, L. (1997). SCORE. Beyond MIDI - The Handbook of Musical Codes, (E. Selfridge-Field, ed.), The MIT Press, London, pages 252-282.
- Taube, R. (1998). CCRMA, Common Music.' CCRMA, Stanford University, California, USA.
- Taupin, D., Mitchell, R., Egler, A. (1997). Using TEX to Write Polyphonic or Instrumental Music ver T.77,'hprib.lps.u-psud.fr.
- Wood, D. (1989). Hemidemisemiquavers...and other such things. A concise guide to music notation, The Heritag Music Press, Dayton, Ohio, USA.

9.1 Main web pages

- web page of the MUSICNETWORK: http://www.interactivemusicnetwork.org
- web page of the MUSICNETWORK working group on Music Notation: http://www.interactivemusicnetwork.org/wg_notation/index.html
- web page on the glossary on music notation: http://www.interactivemusicnetwork.org/glossary/index.htm
- web page of the MPEG AHG on Symbolic Music Representation hosted by the MUSICNETWORK http://www.interactivemusicnetwork.org/mpeg-ahg/

9.2 OFFICIAL MUSICNETWORK DOCUMENTS, deliverables

- **DE4.1.1,** MUSICNETWORK Deliverable, Working Group on Music Notation, “Music Notation Coding”, accessible at: http://www.interactivemusicnetwork.org/wg_notation/upload/musicnetwork-de4-1-1-mn-music-notation-coding-v1-4.pdf
- **DE4.2.1,** MUSICNETWORK Deliverable, Working Group on Music Library, “Music Representation for Music Libraries”, accessible at: http://www.interactivemusicnetwork.org/wg_libraries/upload/musicnetwork-de4-2-1-lib-music-representation-for-music-library-v1-1.pdf
- **DE4.3.1**, MUSICNETWORK Deliverable, Working Group on Music Standards, “Multimedia standards for music coding”, accessible at: http://www.interactivemusicnetwork.org/wg_standards/upload/musicnetwork-de4-3-1-multimedia-standards-for-music-v1-5.pdf
- **DE4.4.1**, MUSICNETWORK Deliverable, Working Group on Music Distribution, “Distribution of coded Music”, accessible at: http://www.interactivemusicnetwork.org/wg_distribution/upload/musicnetwork-de4-4-1-wg_distribution-v1-5.pdf
- **DE4.5.1,** MUSICNETWORK Deliverable, Working Group on Music Protection, “Protection of coded Music”, accessible at: http://www.interactivemusicnetwork.org/wg_protection/upload/musicnetwork-de4-5-1-protection-of-coded-music-v1-4.pdf
- **DE4.6.1,** MUSICNETWORK Deliverable, Working Group on Music Accessibility, “Music Coding for Print Impaired People”, accessible at: http://www.interactivemusicnetwork.org/wg_accessibility/upload/musicnetwork-de4-6-1.pdf
- **DE4.7.1,** MUSICNETWORK Deliverable, Working Group on Music Imaging, “Coding Images of Sheet music”, accessible at: http://www.interactivemusicnetwork.org/wg_imaging/upload/musicnetwork-de4-7-1-coding-images-of-music-v2-8-20040208.pdf
- **DE5.1.1,** MUSICNETWORK Deliverable, All working groups, “Models Integration”, accessible at http://www.interactivemusicnetwork.org/overview/MUSICNETWORK-DE5-1-1-Models-Integration-V1-0.pdf
- **DE6.1.1,** MUSICNETWORK Deliverable, All working groups, “External Assessment and Evaluation Report” : accessible at: http://www.interactivemusicnetwork.org/overview/MUSICNETWORK-DE6-1-1-External-Assessment-and-Evaluation-V1-9.pdf

9.3 OFFICIAL INPUT MPEG ISO DOCUMENTS provoked by the MUSICNETWORK

All the information and documents that have been used to create the following documents are accessible as single or multiple pages on the MPEG AHG main web page (MPEG, ISO/IEC JTC1/SC29/WG11):

http://www.interactivemusicnetwork.org/mpeg-ahg/index.html

- **M9731,** Paolo Nesi, DSI, University of Firenze, Italy, Giorgio Zoia, EPFL, Switzerland, Jerome Barthelemy, IRCAM, France, Pierfrancesco Bellini, DSI, University of Firenze, Italy, David Fuschi, ILABS, GIUNTI, Italy, David Crombie, FNB, The Netherlands, Francesco Spadoni, RIGEL, Italy, Kia Ng, University of Leeds, UK, Martin Schmucker, FHGIGD, Germany; "**Proposal for Music Notation Modeling and its Integration within MPEG-4 and MPEG-7**", MUSICNETWORK IST Network and related working groups (DSI, IRCAM, EPFL), Throndenheime, Norway, July 2003.
- **M10067 Brisbane report**
- **M10068,** Paolo Nesi (DSI, University of Firenze), Giorgio Zoia (EPFL), Pierfrancesco Bellini (DSI, University of Firenze), Jerome Barthelemy (IRCAM). "**Music Notation Application Requirements and MPEG Technology**", AHG on Music Notation Requirements, Brisbane, Australia, October 2003.
- **M10235 Wikaloa report**
- **M10355,** Paolo Nesi (DSI, University of Firenze), Pierfrancesco Bellini (DSI, University of Firenze), Giorgio Zoia (EPFL), Jerome Barthelemy (IRCAM), "**Music Notation Functionality and Interface to MPEG**", AHG on Music Notation Requirements, Waikoloa, USA, December 2003.
- **M10357,** Paolo Nesi (DSI, University of Firenze), Pierfrancesco Bellini (DSI, University of Firenze), Giorgio Zoia (EPFL), Jerome Barthelemy (IRCAM), "**Music Notation Technical Requirements**", AHG on Music Notation Requirements, Waikoloa, USA, December 2003.
- **M10466 Munich report**
- **M10568,** Jerome Barthelemy, Benoit Meudic, Marc Texier, "Proposal for Music Instrumentation and WeightedScales Descriptors and Descriptor Scheme", as a result of the workshop on Music Libraries.
- **M10622,** Giorgio Zoia (EPFL), James Ingram, "**A new Application Scenario for Music Notation in MPEG**". This document presented the "synthetic opera" application scenario, introducing new requirements in terms of interaction with other media, including visual SNHC, and in terms of content protection (MPEG-4 IPMP/IPMPX). AHG on Music Notation Requirements / MUSICNETWORK, Munich, Germany, March 2004.
- **M10654, MPEG2004, ISO/IEC JTC1/SC29/WG11,** Paolo Nesi (DSI, University of Firenze), Pierfrancesco Bellini (DSI, University of Firenze), Giorgio Zoia (EPFL), Jerome Barthelemy (IRCAM), **"Music Notation Technical Requirements and Integration in MPEG-4**", AHG on Music Notation Requirements / MUSICNETWORK, Munich, Germany, March 2004.
- **M10661: MPEG2004, ISO/IEC JTC1/SC29/WG11,** Paolo Nesi (DSI, University of Firenze), Pierfrancesco Bellini (DSI, University of Firenze), Giorgio Zoia (EPFL), Jerome Barthelemy (IRCAM) "**Proceedings of AHG on Music Notation Requirements / MUSICNETWORK**". This document includes all together most of the contributions by different experts; it includes further examples of music notation related applications and examples of music notation xml formats and tools. Munich, Germany, March 2004.
- **M10769, Redmond report**
- **M11002,** James Ingram, **"Position Paper about the SMR Draft Call for Proposals",** Containing some comments on document M11083 provided by James Ingram. AHG on Symbolic Music Representation, Redmond, Washington state, USA, July 2004,
- **M11019,** Paolo Nesi (DSI, University of Firenze), Pierfrancesco Bellini (DSI, University of Firenze), Jerome Barthelemy (IRCAM), James Ingram, David Crombie, "**Draft Evaluation Criteria for Assessing SMR Proposals**", AHG on Symbolic Music Representation, Redmond, Washington state, USA, July 2004,

- **M11021,** Paolo Nesi (DSI, University of Firenze), Pierfrancesco Bellini (DSI, University of Firenze), Jerome Barthelemy (IRCAM), James Ingram, David Crombie, Neil McKenzie, "**Examples of Matching SMR aspects and available technologies**", AHG on Symbolic Music Representation, Redmond, Washington state, USA, July 2004,
- **M11022,** G. Zoia, Jerome Barthelemy, Pierfrancesco Bellini, Paolo Nesi and Mikael Bourges Sevenier (Mindego inc.), "**Graphic functionality in MPEG-4 and Symbolic Music Representation",** AHG on Symbolic Music Representation, Redmond, Washington state, USA, July 2004,
- **M11025,** Paolo Nesi, Giorgio Zoia, James Ingram, Pierfrancesco Bellini, "**Study on Draft CfP on Symbolic Music Representation**", AHG on Symbolic Music Representation, Redmond, Washington state, USA, July 2004,
- **M11083,** Michael Good, **"Response to Draft Call for Proposals N6457",** containing comments on the last version of the Draft Call for Technology on SMR, from Michael Good, Recordare, US; AHG on Symbolic Music Representation, Redmond, Washington state, USA, July 2004,
- **M11175, Palma Report**
- **M11250,** David Crombie, Roger Lenoir, Neil McKenzie, **"Accessible Information in MPEG"**, Palma de Mallorca, Spain, October 2004
- **M11307,** Paolo Nesi, Giorgio Zoia, James Ingram, Pierfrancesco Bellini, Jerome Barthelemy, M. Campanai, Kia Ng, Gioseppe Nicotra, David Crombie, "**Proposed SMR Evaluation Model and Procedure**", AHG on Symbolic Music Representation / MUSICNETWORK, Palma de Mallorca, Spain, October 2004, accessible at: http://www.interactivemusicnetwork.org/mpeg-ahg/Proposed-SMR-Evaluation-Procedure-v1-0.doc
- **M11313,** Jerome Barthelemy, Gregoire Carpentier, "Proposal for a Core Experiment of WeightedScalesDS", Palma de Mallorca, Spain, October 2004.
- **M11354,** Giorgio Zoia, Pierfrancesco Bellini, Paolo Nesi, Jerome Barthelemy, "**MPEG-4 and SMR: report on available functionality for graphics**", AHG on Symbolic Music Representation / MUSICNETWORK, Palma de Mallorca, Spain, October 2004,
- **M11496, Hong Kong Report**
- **M11542 –** David Crombie, Roger Lenoir, Neil McKenzie: "**Proposed Technology for accessible SMR decoders**", Hong Kong, Cina, January 2005.
- **M11554 –** Maurizio Campanai, Pierfrancesco Bellini: "**WEDELMUSIC as SMR proposa**l", Hong Kong, Cina, January 2005.
- **M11560 –** G. Bertoni: "**Proposal for Braille Music Symbolic Representation**", Hong Kong, Cina, January 2005.
- **M11601 –** Tillmann Weyde, Hartmut Ring: "**Proposal for Symbolic Music Representation Format**", Hong Kong, Cina, January 2005.
- **M11630 -** Hyoung-Joong Kim, Yong-Soo Choi, Yong-Ju Cho: "**SMR on the Korean Symbolic Music Representation**" , Hong Kong, Cina, January 2005.
- **M11806, Busan Report**

9.4 OFFICIAL PUBLIC and OUTPUT MPEG ISO DOCUMENTS provoked by the MUSICNETWORK

All the content that has been used to create the following documents are accessible as single or multiple pages attacked to the MPEG AHG main web page:

http://www.interactivemusicnetwork.org/mpeg-ahg/index.html

- **N6049, MPEG2003, ISO/IEC JTC1/SC29/WG11,** Paolo Nesi (DSI, University of Firenze), Giorgio Zoia (EPFL), Pierfrancesco Bellini (DSI, University of Firenze), Jerome Barthelemy (IRCAM), "**Application Requirements of Multimedia and Music Notation**", AHG on Music Notation

Requirements / MUSICNETWORK, Brisbane, Australia, October 2003. (ISO public Document ISO/IEC JTC 1/SC 29/WG 11 N6049)

- **N6149**, **MPEG2003, ISO/IEC JTC1/SC29/WG11,** Giorgio Zoia (EPFL), Paolo Nesi, Pierfrancesco Bellini (DSI, University of Firenze), Jerome Barthelemy (IRCAM), "**Application Scenarios for Music Notation integrated in MPEG**", AHG on Music Notation Requirements / MUSICNETWORK, December 2003, Waikoloa, Hawaii, USA, (ISO public Document as ISO/IEC JTC 1/SC 29/WG 11 N6149).
- **N6454, MPEG2004, ISO/IEC JTC1/SC29/WG11,** "**Workplan for Core Experiment on Music Instrumentation and Weighted Scale Type**", March 2004, Munich, Germany.
- **N6457, MPEG2004, ISO/IEC JTC1/SC29/WG11**, "**Draft Call for Proposals for Symbolic Music Representation",** AHG on Music Notation Requirements / MUSICNETWORK, March 2004, Munich, Germany, (ISO public Document as ISO/IEC JTC 1/SC 29/WG 11 N6457).
- **N6689, MPEG2004, ISO/IEC JTC1/SC29/WG11**, "**Call for Proposals on Symbolic Music Representation**", Audio Subgroup, Public document, Redmond, USA, July 2004. (ISO public Document ISO/IEC JTC 1/SC 29/WG 11 N6689) accessible at: http://www.interactivemusicnetwork.org/mpeg-ahg/w6689.zip
- **N6690, MPEG2004, ISO/IEC JTC1/SC29/WG11,** "**DRAFT SMR Evaluation Procedure**", Audio Subgroup, Public document, Redmond, USA, July 2004. (ISO public Document ISO/IEC JTC 1/SC 29/WG 11 N6690) accessible at: http://www.interactivemusicnetwork.org/mpeg-ahg/w6690%20%28DRAFT%20SMR%20Evaluation%20Procedure%29.doc
- **N6812, MPEG2004, ISO/IEC JTC1/SC29/WG11**, "**SMR Evaluation Model and Procedure**", Audio Subgroup, Public document, Palma de Mallorca, Spain, October 2004. (ISO public Document ISO/IEC JTC 1/SC 29/WG 11 N6812) accessible at: http://www.interactivemusicnetwork.org/mpeg-ahg/
- **N7038, MPEG2005, ISO/IEC JTC1/SC29/WG11, Workplan for the Evaluation of Responses to the CFP on SMR,** Hong Kong, January, Cina, (ISO Document ISO/IEC JTC 1/SC 29/WG 11)
- **N7048, mandate of the MPEG AHG on SMR,** Hong Kong, January, Cina, (ISO Document ISO/IEC JTC 1/SC 29/WG 11)
- **N7152, MPEG2005, ISO/IEC JTC1/SC29/WG11, Audio Subgroup, Report on Symbolic Music Representation RM0 Selection,** April 2005, Busan, Korea (ISO Document ISO/IEC JTC 1/SC 29/WG 11)
- **N7153, MPEG2005, ISO/IEC JTC1/SC29/WG11**, **Audio Subgroup, Workplan for Symbolic Music Representation,** April 2005, Busan, Korea (ISO Document ISO/IEC JTC 1/SC 29/WG 11)
- **N7162, mandate of the MPEG AHG on SMR,** April 2005, Busan, Korea (ISO Document ISO/IEC JTC 1/SC 29/WG 11)

9.5 Direct access to document that contains the above content and much more

1. Application Scenarious
   - ApplicationRequirements-MusicNotation-1-0.html
   - ApplicationRequirements-MusicNotation-1-0.doc
   - Application Scenariuos: ApplicationScenarius.html
   - The corresponding slides presented at the Audio-System group: ApplicationRequirements-slides.pdf
   - The corresponding slides presented at the Audio-Requirements group: ApplicationrRequirementsSummary.pdf
   - What can be done with music notation: Functionalities and their Relationships of Music Notation, definitions
   - Present and innovative Applications of Music Notation
2. Describe technical requirements for Music Notation format (i.e. coded representation) and decoder.
   - Appendix on Music Notation Symbols
   - Glossary of music notation http://www.interactivemusicnetwork.org/glossary/

3. Assess how these requirements can be met with existing MPEG technology and identify missing technology
4. Possible Integration into MPEG
    - Describe the system architecture that supports integration of the Music Notation decoder with existing MPEG technology.
    - Graphic functionality in MPEG-4 and Symbolic Music Representation (graphic-mpeg4-smr.doc) The example is can be visualized Haydn - Menuet (This SVG document shows some of the needed graphic functionality with b-splines, lines and text).
5. Verify that proposed requirements and system architecture satisfy the commonly used formats for and usage of Music Notation.
    - Appendix on Music Notation Comparison, to be downloaded from MUSICNETWORK or to be requested to the AHG chairs since it is covered by copyright and cannot be made public. It will be posted on MPEG:
6. Press releases
    - from March 2004: w6303_Press Release.pdf
7. Some Presentation in slides
    0. Slides presented at the MPEG meeting in Trondheim supporting the above document and position.
    1. SMR with SLIDES: SMR-March-2004-v1-0.pdf

9.6 Links of MUSICNETWORK related events

- First MUSICNETWORK Workshop, Darmstadt, Germany, November 2002: http://www.interactivemusicnetwork.org/workshop/workshopprog.html
- Second MUSICNETWORK Workshop, Leeds, UK, September 2003: http://www.interactivemusicnetwork.org/events/Second_OpenWorkshop_2003/second_openworkshop_2003.html
- Third MUSICNETWORK Workshop, Munich, Germany, March 2004: http://www.interactivemusicnetwork.org/events/Third_OpenWorkshop_2004/MUSICNETWORK-OW-March-2004-Description-v1-4-clean.htm
- Forth MUSICNETWORK Workshop, Barcelona, Spain, September 2004: http://www.interactivemusicnetwork.org/events/Fourth_OpenWorkshop_2004/MUSICNETWORK-4th-Open-Workshop-plan-v1-2.html
- Fifth MUSICNETWORK Workshop, Palma de Mallorca, Spain, October 2004: http://www.interactivemusicnetwork.org/events/modelassessment/MODELAssessment-for-SMR-v1-0.html
- Special section at IAML IASA Conference in July 2004: http://www.interactivemusicnetwork.org/events/iaml_iasa_2004/resonance.html
- Workshop on Editorial Technologies for Music and Multimedia October 2003 Paris. http://www.interactivemusicnetwork.org/events/resonance.html

9.7 Additional http Links

Some links to related tools/products and company (in alphabetic order):

- Berlioz http://www.berlioz.tm.fr , music editor
- Finale, CODA, MakeMusic, music editor and educational tools
- Graphire Music press, http://www.graphire.com
- Igor, NOTEHEADS, http://www.noteheads.com Music editor, sheet music, ...
- MIDIllustrator: MidiIllustrator Music Notation Software: Print, Practice and Play Midi Sheet Music, http://www.MidiIllustrator.com MidiIllustrator is music notation software: Notate, print, practice & play free Midi sheet music. Bring your favorite songs to life. Convert 1000s of Midi & Karaoke files freely available on the World Wide Web into dynamic, interactive scores. User friendly learning tools boost your musical abilities. Learn new music step by step, sight reading notes from the staff OR without reading a note using lead sheets with chord notation, guitar frets & the on-screen piano.
- Mozart http://www.mozart.co.uk , music editor, ...

- MusEdit http://www.musedit.com
- MUSICALIS, music editor and educational tools
- MusicXML, interchange format
- Nightingale http://www.ngale.com/
- NoteAbility http://debussy.music.ubc.ca/~opus1/index.html
- SIBELIUS, music editor and educational tools
- Vivaldi, Opus, Amadeus http://www.vivaldistudio.com , Music editor, OMR, educational tools, sheet music
- WEDELMUSIC: Web Delivering of Music Scores, EXITECH, music editor, multimedia integration and distribution
- Capella, music editor and educational tools, http://www.whc.de/
- MIDISCAN: optical music recognition tool
- SmartScore:
- SharpEye: optical music recognition tool
- SCORE: http://ace.acadiau.ca/score/LINKS3.HTM, music editor

Some Links to Music Learning tools

- eMedia Bass (http://www.emedia.org/)
- ClicknPlayMusic (http://www.cnpmusic.com)
- eMedia Guitar (http://www.emedia.org/)
- Evolution (http://www.evolution.co.uk)
- Flute in play (http://www.musicalis.com/)
- Guitar Magic (http://www.sdgsoft.com/)
- http://guitar.about.com/library/blguitarlessonarchive.htm?PM=ss11_guitar
- http://seattlepi.nwsource.com/pop/less11.shtml
- http://www.geocities.com/Broadway/4131/teachyourself/guitar.html
- http://www.guitar-online.com/
- http://www.musiclearning.com/
- http://www.people.fas.harvard.edu/~desmith/guitar/acoustic/songs.htm
- PlayPro (http://www.intelliware.com.au/software/playpro.htm)
- Voyetra (http://www.voyetra.com/site/)
- ACE, MUSIC ACE series of tools
- Play-along-Recorder, http://www.play-after-me.com
- Play the recorder, http://www.frontdsk.com/lptr/
- Palatine initiative: software for used in universities http://www.lancs.ac.uk/users/music/research/sware.html
- MacGamut, music theory: http://www.macgamut.com/
- Practica Musica:http://www.ars-nova.com/practica.html
- Note Ability: http://debussy.music.ubc.ca/~opus1/NoteAbility/NAwelcome.html
- emusicTheory: http://www.emusictheory.com/
- Musica Analytica: I know absolutely nothing about this: http://www.ertechsoft.com/pages/ma.html
- Music Interval: http://www.musicinterval.com/
- Pitch-class set analysis http://www.palatine.org.uk/directory/index.php/Music/TheoryAndAnalysis/pitch/
- Ear Training tools http://www.mhhe.com/helpdesk/benward.html,
  Ear Training tools http://www.rising.com.au/auralia/
- MIDIllustrator: MidiIllustrator Music Notation Software: Print, Practice and Play Midi Sheet Music

Some referred Projects and tools:

- **CANTATE: music notation, library, SMDL,** projects.fnb.nl
- CCARH and MuseData: http://musedata.stanford.edu/databases/index.html
- CUIDADO: Processing of Music and Mpeg7: http://www.ircam.fr/produits/technologies/multimedia/cuidado-e.html
- Guido http://www.informatik.tu-darmstadt.de/AFS/GUIDO/
- HARMONICA, network http://projects.fnb.nl/harmonica
- IMUTUS: Interactive music tuition system

- MOODS: Music Object Oriented Distributed System
- MUSICWEB educational, teaching, object oriented, CORBA. Computer aided music education http://sun1.rrzn.uni-hannover.de/musicweb
- MUSICAL (eContent): Multimedia Streaming of Interactive Content Across Mobile Networks, http://musical.intranet.gr
- NIFF, http://www.musitek.com/niff.html
- OPENDRAMA: music tools for opera show, http://www.opendrama.com
- SMDL, a closed ISO Standard
- Variations 2 at the Indiana University - Bloomington http://www.dml.indiana.edu/.
- Xemo project http://www.xemo.org
- MUSITECH project: http://musitech.fmt.uos.de at the University of Osnabrueck, department of Music and Media Technology http://www.fmt.uos.de
-

Links to list of Music Notation resources

- Document list of Working Group on Music Notation of MUSICNETWORK. The addition of new document can be done by all MUSICNETWORK participants. The page also contains a list of examples of music notation, just to show which are the most relevant requirements for the modeling of music notation, at structural level. This is very interesting for modelling the logical part of music notation. The addition/upload of new documents can be done by all MUSICNETWORK participants. The registration is free and can be performed via the main www site of the network.
- Music Notation Codes: http://www.music-notation.info/en/compmus/notationformats.html, http://www.music-notation.info/en/compmus/musicnotation.html , http://www.music-notation.info/,
- Notation Scoreing Zone, http://www.synthzone.com/notation.htm
- MusicMoz, http://musicmoz.org/Computers/Software/Notation/
- AudioMIDI, http://www.audiomidi.com/software/notation.cfm
- DMOZ: http://dmoz.org/Computers/Multimedia/Music_and_Audio/Software/Notation/
- Music Notation and Engraving, http://www2.coloradocollege.edu/dept/MU/Musicpress/Default.htm
- Musical Information in Musicology and Desktop Publishing, Music Notation example of the most frequend hardness, http://www.ccarh.org/publications/reprints/ieee/ by Eleanor Selfridge-Field and Edmund Correia, Jr.
- MML http://www.musicmarkup.info/ MML (Music Markup Language)
- An article on music notation XML for interchange format http://emusician.com/ar/emusic_xml_music/index.htm
- An article from IRCAM on Music Notation aspects and Synchronisation (mention of the AHG activities) mn-req-alignment.pdf
- music notation glossary, from David Pawson, RNIB: http://www.dpawson.co.uk/smrf/
- virginia tech glossary of Music notation http://www.music.vt.edu/musicdictionary/
- Glosasry of music notation http://www.interactivemusicnetwork.org/glossary/, thanks to Giuseppe Nicotra

References to Organisations:

- MUSICNETWORK: The Interactive Music Network
- MIDI.ORG: http://www.midi.org
- Music Educator: http://www.musiceducator.org/modules.php?name=Downloads&d_op=MostPopular
- Technology Institute for Music Educators, http://www.ti-me.org/
- EICTA: European Industry Association, Information Systems Communication Technologies Consumer Electronics, http://www.eicta.org
- EL.pub, Interactive electronic publishing R&D news and resources home page, http://www.elpub.org
- EMF, European Multimedia Forum, http://www.e-multimedia.org
- MEA: Mobile Entertainment Analyst, http://www.wirelessgamingreview.com/mea/
- MEF: Mobile Entertainment Forum, http://www.mobileentertainmentforum.org/