# NEW FRONTIERS IN MUSIC EDUCATION THROUGH THE IEEE 1599 STANDARD

Adriano Baratè and Luca A. Ludovico

*Laboratorio di Informatica Musicale, Dipartimento di Informatica e Comunicazione, Università degli Studi di Milano*
*Via Comelico 39/41, 20135 Milano, Italy*

Keywords: Music, XML, IEEE 1599.

Abstract: IEEE 1599 is an international standard conceived for a comprehensive description of music. Inside a unique XML document, all the data and meta-data related to a single music piece can be coded, ranging from catalogue to structural, from notational to performance, from audio to graphic information. This format is particularly fit for computer-supported music education, since it allows a number of applications such as evolved score following, real-time comparison among different performances, studies on score transcriptions, musicological analyses, etc. After describing the key features of the standard, a number of case studies will be presented in order to demonstrate its applicability to music education.



## 1 INTRODUCTION

IEEE 1559 is a multilayer music code whose international standardization was achieved in 2008. Its development followed the guidelines known as IEEE P1599 - Recommended Practice Dealing With Applications and Representations of Symbolic Music Information Using the XML Language. The ultimate goal of the IEEE 1599 standard is providing a highly integrated representation of music, where score, audio, video, and graphical contents can be encoded together in a fully synchronized framework.

As Section 2 will explain, the format focuses on single music pieces and is based on the concept of "music event", namely an atomic entity which can be described from different perspectives. In fact, the mission of IEEE 1599 is taking into account the different facets music is made of. Please refer to (Steyn, 2002) and (Haus and Longari, 2005) for further details. Needless to say, multimedia descriptions of music events are supported too, so a number of digital objects can be attached to the XML document in order to provide the symbolic score with graphical representations as well as audio/video recordings.

The information encoded into an IEEE 1599 document can be rich and heterogeneous. In this context, heterogeneity is involved from two standpoints: i) many categories of descriptors (metadata, music symbols, text, still graphics, audio, and video), and ii) the possibility to insert many instances for each category (different score editions, audio performances, etc.). The former item catches the various aspects music is made of, whereas the latter allows comparisons or integrations among different instances of the same type of description.

Since IEEE 1599 was born as a format for a comprehensive, integrated and synchronized representation of music information, it constitutes a good base for advanced applications oriented to music enjoyment and entertainment. But soon the format has revealed its applicability to music training and education as well. This subject has been already addressed in (Baratè et al., 2009). The present paper draws inspiration from the mentioned one and takes into account the most relevant experiences made in the last two years. Sections 2 and 3 will introduce the key features of the standard, Section 4 will show its applicability to the field of music education, and finally Sections 5 and 6 will provide some practical examples.

## 2 IEEE 1599 IN BRIEF

In IEEE 1599 music can be described in all its aspects. In order to support heterogeneous materials, this standard employs six different layers to represent information: *General* (catalogue metadata about the piece), *Logic* (logical description in terms of score symbols), *Structural* (identification of music objects


Baratè A. and A. Ludovico L..
NEW FRONTIERS IN MUSIC EDUCATION THROUGH THE IEEE 1599 STANDARD.
DOI: 10.5220/0003912601460151
In *Proceedings of the 4th International Conference on Computer Supported Education* (CSEDU-2012), pages 146-151
ISBN: 978-989-8565-06-8

and their mutual relationships), *Notational* (graphical representations of the score), *Performance* (computer-based performances), and *Audio* (digital or digitised recordings of the piece). IEEE 1599 adopts XML as the language to encode symbolic contents and multimedia synchronization information.

The main focus of an IEEE 1599 document is the description of a single music piece. The *Logic* layer contains a list and a definition of the music symbols that compose the score. Music events, intended as logic entities, can correspond to one or many instances in other layers. In detail, each music event (note, rest, etc.) can be described in up to *n* layers, in up to *n* instances within the same layer, and in up to *n* occurrences within the same instance.

IEEE 1599 is not a simple container for heterogeneous descriptions related to a music piece. Different logic or multimedia descriptions of the same music events present references to a common logical structure, known as the *spine*. For example, given a note, its logical description (e.g. G♯5), the corresponding area in a graphical file (e.g. a rectangle with vertex coordinates in pixels) and its timing in a number of different audio/video files (e.g. 1200 ms, 2.35 s, 712 frames) are in relationship with the same music event. This aspect creates synchronization among the instances within a layer (*intra-layer synchronization*), and also synchronization among the contents disposed in many layers (*inter-layer synchronization*).

The mentioned common data structure, namely the *spine*, is a list of music events that are sorted and labelled in order to allow references from other layers. Against common sense, score symbols do not correspond necessarily to the list of the events contained in the spine. Because if it were, music works with no notation, such as pieces where the performance is improvised, or music whose score is unknown, could not be supported by IEEE 1599, unless the user reconstructs the score. On the contrary, a traditional notated score or a complete encoding of the piece is not required to produce a valid IEEE 1599 document. Moreover, the author of the encoding can choose the definition and granularity of events.

The *spine* has a fundamental theoretical importance within the format, thanks to its "glue" function in such a multilayer environment. However, it simply lists (and does not define) events in order to provide a unique label for them. As a consequence, the mere presence of an event in the *spine* has no semantic meaning: what is listed in the *spine* structure must have a counterpart in some layer, otherwise the music event would not be defined.

For further details about the structure and the syntax of the format, please refer to the official IEEE documentation or to scientific papers such as (Ludovico, 2008) and (Ludovico, 2009).

## 3 CHARACTERISTICS OF THE FORMAT

A number of characteristics applicable to music representation and education can be derived from what stated in Section 2. First, the standard and its applications present no constraint about music genres, cultural areas, historical periods, or different approaches to music composition and analysis. IEEE 1599 is fit to baroque counterpoint as well as jazz improvisation, to operatic arias as well as pop songs. The features of the encoding can change noticeably, but the overall approach preserves its capability to convey information in an integrated, intuitive and immediate way. The case studies presented in the following sections will provide examples of the heterogeneity of sources and materials which IEEE 1599 supports.

As regards multimedia contents, in IEEE 1599 they are represented by adopting in-use formats for digital objects. Thus, multimedia contents are not translated into XML format, as existing formats are more suitable. Rather, multimedia documents are synchronized with the other contents within the appropriate layer and linked to the common data structure, i.e. the *spine*. For instance, IEEE 1599 supports common graphical formats for score scans (e.g. BMP, GIF, JPEG, TIFF, etc.), and well-known audio/video file types for recordings (AIFF, MP3, WAV, etc.). This approach presents some advantages:

- Existing collections and archives of digital objects in standard formats can be reused;
- The design and implementation efforts of the IEEE 1599 format are limited to the logical description of music events, and not to their multimedia counterparts;
- The verbosity typical of an XML-based language afflicts only the strictly necessary part of the encoding. When binary formats are more efficient and effective, they are used to convey the required information. For instance, this is the typical case of compressed multimedia formats.

Another important matter is the possibility of multimodal interaction with music contents. This feature is not properly a characteristic, rather a consequence of the structure and potentialities of the IEEE 1599 standard. Specific implementations apply the concept of "multimodal interaction" with music contents to the fields of music training and education, as explained in Section 6.

# 4 APPLICABILITY TO MUSIC EDUCATION

An IEEE 1599 document is potentially very rich in information. The characteristics illustrated before let us design a wide range of applications. For example, the matter of evolved music enjoyment for entertainment has been addressed by (Baggi et al., 2005), whereas the valorisation of music as a lively cultural heritage through IEEE 1599 has been cited in (Baratè et al., 2011).

Another field where IEEE 1599 can find application is music education and training. In fact, such a format supports a number of features that can be employed to create *ad hoc* implementations, e.g. a guide to music listening or a tool for instrumental and ear training. The heterogeneity of both descriptions and instances within each layer opens up new ways to enjoy a music piece.

Let us explore the different ways to render a "logical" score, i.e. the information encoded within the *Logic* layer in terms of music symbols. First, it can be dynamically reconstructed by a viewer/editor starting from the XML document. Besides, it can be linked to the *Notational* layer, which could contain not only printed or even hand-written scores, but also other forms of graphical description. This feature makes IEEE 1599 support even those scores not belonging to Common Western Notation. Now, let us couple two of the mentioned characteristics of the format, namely i) the possibility to watch and listen to music together, in a context of advanced score following, and ii) the support of non-traditional scores: a tool based on these features can be used to teach e.g. contemporary music, as regards both score reading and performance.

The *Audio* layer can host various performances of the same piece, which in general correspond to many different interpretations. But - once again - the purpose of this layer can be extended and somehow forced. For instance, a multiplicity of tracks can be used to encode cover versions of the piece, which can substantially differ from the original one. It sounds natural to apply this concept to jazz music and improvisation. Another possibility, employed in one of the case studies below, consists in linking as different audio instances the single tracks of a multi-track recording.

Also ear training and instrumental practice can be based on the adoption of IEEE 1599. By using audio source separation techniques or multi-track recordings, some parts and voices can be easily removed from (or differently mixed in) the audio output by a specific software; then a score-following application can highlight the part the student has to perform.

Some layers, intentionally ignored till now, may allow many other applications. This is the case of the *Structural* layer, which permits the identification of music objects and their relationships in a score. The possibilities of the format are wide enough to embrace harmonic grids, segmentation, and different kinds of musicological analysis, as shown in (Dalmonte and Spampinato, 2008).

Other interesting applications can emerge, particularly for those cultures and music genres far from Western tradition, a field where there has been little investigation and methodological research.

In general terms, the IEEE 1599 standard becomes more and more interesting when its contents are increasingly rich. In fact, the applications cited above are based on the contemporaneous presence of heterogeneous media descriptions. An IEEE 1599 document would be valid even when only the *Logic* layer has been compiled, but the most advanced characteristics of its applicability would be lost.

Creating a rich IEEE 1599 document presents some problems related to media linking and synchronization. Since a key concept of the format is the mutual synchronization of all associated media, the process of adding a certain kind of material cannot be viewed merely as linking a file, but it needs an automatic, semi-automatic, or manual synchronization procedure. A promising research field consists in the automatic discover of symbols inside digital objects such as score scans (optical music recognition) or tracks (audio-to-score applications). However, in IEEE 1599 all the materials have an explicit relation only with the *spine*. As a consequence, adding a new media has a linear cost, and both intra-layer and inter-layer synchronization with the other objects are automatically achieved.

# 5 IEEE 1599-BASED MUSIC VIEWERS

After the creation of an IEEE 1599 document, the goal becomes playing it appropriately. The characteristics of the format suggest not only the implementation of a traditional viewer, but the design of a tool to deeply interact with music contents. The guidelines of such an application have been described in (Baratè and Ludovico, 2008).

In our opinion, a format such as IEEE 1599 supports three models of music fruition. The first model involves score following, possibly with advanced features. In this case, users concentrate on the synchronization among audio and graphical contents. This activity can be a way either to enjoy music in a rich context or to learn score following. If we introduce some kind of annotation, this application can be a means of musicological analysis, too.

A more advanced model provides the possibility to switch in real time the media currently playing or appearing in the interface. This lets the user compare *on the fly* different performances and scores of the piece. The alignment between the old and the new current media is easily reconstructed thanks to the information contained in the *spine*; similarly, the overall synchronization between audio and score is not affected, and the effect from the user's perspective is a real-time substitution for a kind of media. Such an application can be useful for instrument players, singers, and musicologists.

Finally, a model supporting interaction with music contents can be implemented. In fact, thanks to the mappings encoded in the *Audio* and *Notational* layers, graphic files' areas as well as time sliders can be sensible to mouse clicks and cause a prompt resynchronization of music contents.

A number of applications following the mentioned guidelines have been designed, implemented and presented during international conferences, symposia and exhibitions in order to demonstrate the applicability of IEEE 1599 standard to different fields, ranging from education to dissemination, from music information retrieval to entertainment, from advanced fruition of multimedia contents to music cultural heritage. Among the most relevant experiences related to music education, let us cite the opening of 2006/07 season at Teatro alla Scala (*Celeste Aida - Percorso storico e musicale tra passato e futuro*, Teatro alla Scala, Milan, Italy, December 2006 - January 2007) and 2007 Salzburg Festival (*Napoli, nel nobil core della musica*, ResidenzGalerie, Salzburg, Austria, May - June 2007).

As regards IEEE 1599 players, currently there are two research fields. On one side, our staff at *Laboratorio di Informatica Musicale* is developing the first multi-platform player using C♯ and GTK♯ technologies, whereas all the applications released till now were custom installations. On the other side, a strong interest is rising towards an IEEE 1599 Web player based on HTML5. See (Baldan et al., 2010) for technical details. The case studies shown in Subsections 6.1 and 6.2 are examples of offline viewers, whereas the one presented in Subsection 6.3 works over the Internet.

# 6 CASE STUDIES

The present section describes three recent approaches to music education based on the IEEE 1599 standard.

## 6.1 Advanced Score Following

A recent example of application developed for both entertainment and education is *That's Butterfly*. This IEEE 1599-based installation focuses on the famous aria "Un bel dì vedremo" from *Madama Butterfly* by Giacomo Puccini. The project has been developed in the framework of an itinerant international exhibition sponsored by the *Archivio Storico Ricordi*.

The purpose of the application, whose interface was conceived for touch-screen systems (see Figure 1), was allowing easy interaction with music contents also by untrained people. Multimedia materials included score scans of the autographic score and three historical recordings. This application implements the three fruition models described in Section 5.

The possibility to compare in real time a number of historical performances is interesting both for people keen on music and for experts. But from the educational point of view, this installation lets further aspects emerge. First, it implements an advanced score follower where the standard vertical line marking the current notes is substituted by a number of coloured rectangles. This evolution is required by the characteristics of the handwritten score by Puccini, where all the revisions and adjustments cause several misalignments among simultaneous notes. From a didactic standpoint, this allows: the exploration and analysis of the compositional process; the comparison among handwritten symbols and printed ones, as they appear in the official score version; and finally the listening of audio with reference to the autographical document.


Figure 1: An interface for advanced score following.

### 6.2 Musical Practice and Ear Training

The case study shown below is an application aiming at musical instrument practice and ear training (see Figure 2). For this kind of projects, a good integration among logic, structural, audio and video contents is fundamental.

The example focuses on the *Messa da Requiem* by Giuseppe Verdi. Its multimedia materials embrace only one printed score and many audio/video recordings of a single evening, i.e. 6 multi-angle video-takes and 24 independent audio tracks. Since the piece is quite long, it has been segmented into movements and the granularity of the encoding is represented by the measure instead of the single note.

The interface puts the score in great evidence, and presents a number of floating panels to enable particular functions. The left panel lists the movements the whole piece is made of. Let us recall that each movement corresponds to a different IEEE 1599 file. The upper panel lets the user select which instrument to listen at. This is made possible by the availability of as many tracks as the instrumental groups of the orchestra, including solo parts. The complete ensemble - i.e. the standard stereo recording - can be selected as well. The right panel, finally, contains the multimedia player for videos. Six takes are available for each movement, including a front view of the conductor, a full-orchestra overview and most instrumental groups.

Let us concentrate on musical instrument learning. In an interactive multimedia environment, the student can see the score to perform, listen to a pre-recorded version of his/her part, and even watch at the performance made by expert musicians. This aspect is relevant not only for players and singers, but also for orchestra conductors. IEEE 1599 standard allows also comparisons among different performances, a feature important for education and not exploited in this case study.

The support of multiple audio tracks is another key feature in this field. An *ad hoc* mixing of such single-instrument recordings is useful for both instrument and ear training. For instance, a trumpet student and a double bass practitioner could decide to play together with the orchestra by removing only the audio tracks referred to their parts. As regards ear training, a student could try to follow a given part in an orchestral recording, and then validate his/her reading capabilities by listening to a demixed audio track.

### 6.3 Valorisation of Gregorian Chant

The last case study we present is an online viewer to

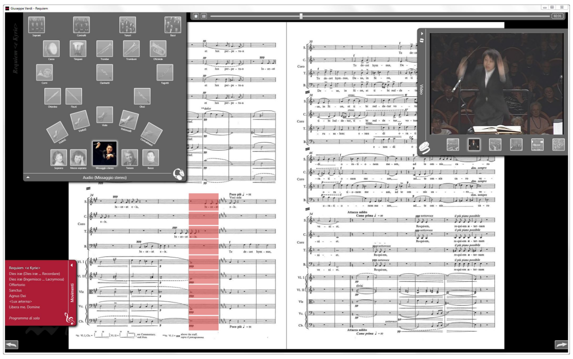

Figure 2: An interface for musical instrument practice and ear training.

enjoy Gregorian chant (see Figure 3). This application is the public outcome of a digitization campaign held at the *Certosa of Pavia* (Italy): 13 graduals dating back to the XVI century and containing masses in neumatic notation have been digitized. In this project, IEEE 1599 has been employed to relive Gregorian chant and to make it readable also by non-experts.

The key features of the format let us implement those fruition models described in Section 5 and already used in the other case studies. However, the peculiarities of this music genre are unique from many points of view.

First, the notation in use substantially differs from the current one. There are methods and transcription rules, well-known in literature, to convert neumes into modern notation, so that also a symbolic encoding could be provided inside the XML document. Moreover, a part of the IEEE 1599 *Logic* layer is devoted to the explicit description of neumes. In this way, not only single notes but also their original aggregations into neumes could be encoded.

As regards the *Notational* layer, this time it is very rich in heterogeneous digital objects. At least, both the original score and its modern transcription can be linked to music events. Moreover, in this sense an aspect of further richness and interest emerges: neumatic notation, due to the extremely wide diffusion of the repertoire and to the hand-made process of copying, often presents small variants as regards note pitches or grouping into neumes. For this reason, it has been interesting to compare other score versions.

For Gregorian chant, an IEEE 1599-based application provides not only a way to make scores enjoyable by untrained people through its score-following features, but also a professional tool to explore in real time melodic and notational variants.

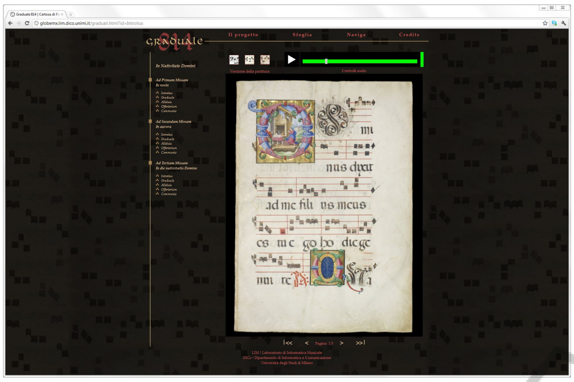

Figure 3: A web interface for Gregorian chant.

## 7 CONCLUSIONS

This paper has shown the applicability of the international XML-based standard known as IEEE 1599 to the field of computer-supported education, paying particular attention to music.

After describing the key features of the format, our work has introduced several innovative applications recently realized, and it has suggested new ways to use the standard in this context.

The adoption of XML constitutes an important step towards interoperability, since it allows the implementation of open and free languages to encode score information. Moreover, IEEE 1599 not only provides organized structures for music symbols, but it is also a wrapper to bind multimedia information to music events within a synchronized framework.

Heterogeneity as regards both the cardinality and the type of music descriptors is supported and even encouraged.

Such a rich environment allows the implementation of a number of offline and online tools to enjoy, teach, learn and analyse music through a computer system. The present paper has shown only three case studies, focusing on different periods and genres of vocal music. Nevertheless, many other applications can easily emerge. In conclusion, thanks to the adoption of IEEE 1599 standard, computer-based methods prove to be particularly effective in the music educational field.

## ACKNOWLEDGEMENTS

The authors wish to acknowledge the members of the Laboratorio di Informatica Musicale for their participation in the IEEE 1599 initiative. This work has been partially funded by the *Enhanced Music Interactive Platform for Internet User* (*EMIPIU*) project.

## REFERENCES

Baggi, D., Baratè, A., Ludovico, L. A., and Haus, G. (2005). A Computer Tool to Enjoy and Understand Music. In *The 2nd European Workshop on the Integration of Knowledge, Semantics and Digital Media Technology (EWIMT 2005).*, pages 213–217. IET.

Baldan, S., Ludovico, L. A., and Mauro, D. A. (2010). Algorithms for an Automatic Transcription of Live Music Performances into Symbolic Format. In *Genesis of Meaning in Sound and Music. 6th International Symposium, CMMR/ICAD 2009, Copenhagen, Denmark, May 18-22, 2009, revised papers*, volume 5954 of *Lecture Notes in Computer Science*, pages 422–438. Springer.

Baratè, A., Haus, G., Ludovico, L. A., and Mauro, D. (2011). A Web-oriented Multi-layer Model to Interact with Theatrical Performances. In *International Workshop on Multimedia for Cultural Heritage (MM4CH 2011), Modena, Italy*, volume 247 of *Communications in Computer and Information Science*. Springer.

Baratè, A. and Ludovico, L. A. (2008). Advanced Interfaces for Music Enjoyment. In *Proceedings of the working conference on Advanced visual interfaces: AVI 08: Napoli, Italy, May 28-30, 2008*. ACM Press.

Baratè, A., Ludovico, L. A., and Haus, G. (2009). IEEE 1599: a New Standard for Music Education. In *Proceedings of the 13th International Conference on electronic publishing, Milan, Italy, June 10-12, 2009*, pages 29–45. Edizioni Nuova Cultura.

Dalmonte, R. and Spampinato, F. (2008). *Il Nuovo in Musica: Estetiche, Tecnologie, Linguaggi. Atti Del Convegno, Trento, 18-20 Gennaio 2008.* Libreria Musicale Italiana.

Haus, G. and Longari, M. (2005). A multi-layered, time-based music description approach based on XML. *Computer Music Journal*, 29(1):70–85.

Ludovico, L. A. (2008). Key concepts of the IEEE 1599 Standard. In *Proceedings of the IEEE CS Conference The Use of Symbols To Represent Music And Multimedia Objects, IEEE CS, Lugano, Switzerland*, pages 15–26. IEEE CS.

Ludovico, L. A. (2009). IEEE 1599: a Multi-layer Approach to Music Description. *Journal of Multimedia*, 4(1):9–14.

Steyn, J. (2002). Framework for a Music Markup Language. In *Proceedings of the First International IEEE Conference on Musical Application using XML (MAX2002)*, pages 22–29.