Abstract—IEEE 1599 is a new standard to encode music with XML symbols. It offers two important original characteristics compared to existing standards of the worlds of music, musicology and computer applications to this art and science. On one side, the encoding is in the form of symbols that can be read both by machines and humans. On the other, it allows the realization of applications in which all aspects of music, such as audio and sound, graphical representation, historical data, performance indications, represented thanks to the new concept of layers, are fully integrated, synchronized and can be accessed both individually and as parts of a whole. The article will give a brief description of the standard and of applications that have been built to show its power.

Index Terms—Music encoding, standards for music, XML, symbolic representation of music, music layers, IEEE.

I. INTRODUCTION

Every music lover has always known that music is much more than audio and cannot be reduced to the act of accessing the encoded file of an I-Pod through tiny earphones. The musical experience is seen, instead, as the act of entering a new world, living a new experience, understanding a narration and recognizing descriptions, as in the case of serious reading of a piece of literary art such as Tolstoi’s War and Peace and Manzoni’s I promessi sposi, which create a parallel reality. In addition, music offers the possibility of investigating how the whole is built from the technical standpoint of music – which is the object of musicology, the science of music.

For such an experience, music must be represented with something that goes beyond unreadable, binary standards for audio, such as WAV and MP3, which are not music standards. Musical aspects beyond audio must be represented in human-readable form, such as symbols and characters. This has always been for music scores in classical music, and for music notation in all civilizations for at least forty centuries (e.g., Babylonian tablets). And with other symbols such as the harmonic grid in jazz, and also in other written codes, as in non-western music.

This is also the case of the new standard for music encoding IEEE 1599, which uses XML clauses [1]. This work represents the culmination of decades of efforts of specialists in the field of computer applications to music and musicology, of which the Plaine-And-Easie Code [2], DARMS [3] and SMDL [4] are worth mentioning, while MusicXML [5] and MEI [6] are direct ancestors of this project. However, IEEE 1599 goes beyond such past efforts, as described in the next paragraphs.

II. CHARACTERISTICS OF IEEE 1599

The main distinguishing characteristics of standard IEEE 1599 are: the use of symbols to represent music, and the concept of layers.

Symbols

Every element of Common Western Notation can be represented by XML clauses that can be nested as needed, as in the example:

```
<clef type="G" staff_step="2" event_ref="c1"/>
```

Treble

This is described in several pages of DTD’s that have been published at the site of the IEEE Standards Association, one of one of the associations of the Institute for Electrical and Electronic Engineers. In addition, thanks to the inherent extensibility of XML, it is possible to add clauses for special needs – such as a proprietary characters used by a particular music publisher, or notation which is not yet standardized [7].

Layers

In recognition of the different aspects of music, the concept of layers has been introduced [8] and is an integral feature of the standard, as shown in Figure 1.

The general layer provides a general description of the
music work and groups information on all related instances, including titles, author, type, number, date, genre, and related items. The logic layer provides music description from a symbolic point of view and represents the core of the format. It contains: the main time-space construct for localization and synchronization of music events, the description of the score with symbols, and information about a graphical implementation of the symbolic contents, as well as the spine with Logically Organized Symbols or LOS, a sorted list of music events.

All layers are related to each other, as shown in Figure 2, which shows that standard IEEE 1599 allows representation, performance, and audible and visual fruition of a piece of music independently of the original standard or format with which it was previously encoded. It therefore supports existing formats and recognizes that music contents that are already encoded in pre-existing file formats can not be ignored. Since no overall description has existed for all aspects of music in one single format, it thus provides a meta-language in XML to both describe all related music elements and link all corresponding media objects already encoded. Thus, music contents are either newly encoded in the proper layers (the General, Logic, and Structural Layer actually store information) or they remain in their original format, with links from the corresponding layers to files (the Notational, Performance, and Audio layers contain mappings to external files). Hence, media files are handled as they are, and media contents are still available in their original encoding. The comprehensive format described in the standard uses the layers to represent the relationships between music events and their occurrence in media files, thus allowing an overall synchronization. In other words, it is possible to navigate music in all its aspects.

The interaction and synchronization among these layers will become clear thanks to the examples of the next paragraphs.

### III. EXAMPLES OF APPLICATIONS OF IEEE 1599 TO INCREASE MUSIC ENJOYMENT

These are examples of applications, easily realised thanks to symbolic expression of music and synchronised layers, to open up new ways for music fruition.

- **Pieces by a jazz Big Band.** The harmonic grid is shown with the soloist name for each solo – with didactical tools as in jazz textbooks [9]
- **An opera.** A DVD of an opera allows the user to: see the play, hear the music, see the score, read the libretto; choose alternative renditions
- **Music with a “program” or story.** E.g., Vivaldi’s Four Seasons with poems that refer to the music
- **Music with no apparent meaning.** For instance, free jazz of the 60’s – 70’s is perceived by many as a random collection of meaningless sounds, while an associated video, generated anew each time, may help understand what is meant
- **A fugue.** The theme is highlighted as it gets passed among the different voices
- **A piece of Indian classical music.** The scale of the raga is shown and the melodic development is highlighted
- **A piece of several drums, as in African Drumming,** to show how the rhythmic pulse come from the fact that the hits do not fall together
- **Preservation of the music heritage from the past.** To store documents in any media [10]
- **Musicological study.** Ease of queries – for example, all pieces utilising the lowest note of a grand piano; questions as to why a certain note is used in a given harmonic context
- **Books about the making of a masterpiece,** e.g. Kind of Blue, LP by M. Davis, J. Coltrane [11].
IV. EXAMPLE WITH A SCORE, BUT DIFFERENT JAZZ VERSION: KING PORTER STOMP

This is an application built, like the others, at the Laboratory for Musical Informatics of the University of Milan [13]. The screenshot shown in the figure contains different windows, of which those with the extra caption real time operate in synchronism while the music is being played.

The user starts with the piece selection window: in this case, there are two choices, King Porter Stomp 1924, and King Porter Stomp 1939. They refer to two published scores of that piece, which was famous in the 30’s, by American composer and pianist Jelly Roll Morton, or Ferdinand Joseph La Motte, 1889-1941. In the file selection window, the user can choose among alternate multimedia files, in this case a recording from 1926 in MP3 format, a MIDI rendition of the 1924 score encoded in MP3, and an excerpt from Louis Malle’s movie Pretty Baby of 1977 – a character patterned after Morton is heard composing the piece in the background.

The latter, i.e. the movie, is the one shown here in the window player, display, which for plain music looks instead like a common music player. Two more choices are versions by the jazz orchestra of Fletcher Henderson – one recorded in 1928 and one in 1932 – a bandleader, composer and arranger who popularised the piece for a band with a section of trumpets, trombones, and reed instruments – clarinets and saxophones. This shows how apparently very different music pieces have the same root and structure.

Upon that selection, several synchronised activities start and execute in real time. The music starts playing, in the case of the movie a video segment starts together with its sound. On the score, the running cursor indicates what is being played, here the beginning of the 7th bar. The user can move the red cursor with the mouse and initiate playing from another point in the score, while all other real time windows adjust synchronously, and of course move the player cursor.

The XML code window shows the encoded events, in this case those of the LOS, Logical Organised Symbols of Figure 1 and 2, scrolling with the music. In the command window, the user can select which XML code is displayed: spine, LOS, notation and audio, again those of Figure 1 and 2, and in the same window one can choose the voice the running cursor will follow – there are three voices in this case.

The chords window displays the elements of the music harmony of the piece, again synchronously with the playing, and the window for the multimedia files allows selection of pictures or other, portraits of Morton, of his band, and curiosities, including a map of Storyville at the turn of the 20th century, a New Orleans district that was closed in 1917 and torn down in the 1930’s.
V. A JAZZ PIECE WITH NO SCORE: CRAZY RHYTHM

The figure of above is the screenshot for jazz piece Crazy Rhythm. Instead of a score, it displays the harmonic grid, pointed to by the running cursor, and lets picture and name of each soloist pop up at the appropriate moment. There are four saxophonists taking solos: André Ekyan on alto and Alix Combelle on tenor, both from France, followed by Afro-Americans Bennie Carter on alto and Coleman Hawkins on tenor, both in Europe at the time, top jazz specialists of their era. The rhythm section consists of violinist Stéphane Grappelli on the piano, Django Reinhardt on guitar, Eugène d’Hellemmes on bass and Tommy Benford on drums. Though entirely improvised, the recording, made in Paris on April 27, 1937, counts among the best of jazz history, Number #1 of the new brand Swing, the first record company dedicated entirely to jazz, and everybody plays extremely well.

The ensemble exposes the theme once, in four voices, for all 32 bars. At each solo, image and name of soloist appears: thus, it is possible to compare styles by jumping on the image of another soloist, and even to compare the sound of the alto saxophone with that of the tenor. Bars and grid are of course synchronised. Each soloist takes 32 bars, except for Hawkins who, in the middle of the development of sentences that keep building up, takes an unplanned second chorus, after the encouraging shout by Django “Go on, Go on”, which is automatically displayed at the 31st bar. It is details like this that would be totally lost to a casual listener, and which instead constitute the key for understanding improvised music. The ensemble takes over at the 30th bar of Hawkins’s second chorus.

In this application, the standard and its browser are used to teach a would-be jazz expert what to look for: such as distinguishing among soloists, instruments, following the improvisation at each bar, understanding what everybody, from horn to drums passing through the whole rhythm section, is doing. This is a tool to learn jazz appreciation.
VI. AN OPERA ON THE COMPOSER’S MANUSCRIPT: TOSCA, BY PUCCINI

This example demonstrates how it is possible with this standard to preserve documents from the past and make them live. Instead of simply digitising an old document subject to degradation, this technology allows both preservation and realisation of an application that allows exploration of music with everything it contains.

The opera is Tosca, with music and original manuscript by Giacomo Puccini. The application has been presented at the Exhibition Tema con variazioni: musica e innovazione tecnologica, Parco della Musica, Rome, December 2005 – January 2006. The manuscript has been made available by publishing house Ricordi, which has the manuscripts of practically every great Italian composer of the last two centuries.

The usual synchronisation of events applies here, as in the other examples. The user can select between the original manuscript and a printed score with the words, also between clarinet and canto. Three tenors can be selected, among which Giuseppe di Stefano and Enrico Caruso from a 1909 Victor Recording owned by the author, and also a video. There are also various images of the time, of Rome, Castel Sant’Angelo, posters of the time and the like.

It shows how easy it is to gain knowledge of a complex piece like an opera, including libretto and different performances that differ considerably. The switch from Di Stefano and Caruso is illuminating, especially because of the difference in voice and style, but also to highlight the recording technique of different epochs, which was entirely acoustic – neither electricity nor electronics were available for recording before 1925.
Free jazz is a musical current popular in the 60’s and 70’s that broke with jazz tradition. Among others, it rejected tonality, the diatonic scale, the regular tempo and meter, the structure in measures, it used sounds forced outside of the range of the instrument, and any device that was not part of a learned repertoire. It was, therefore, hard even for seasoned jazz lovers to make sense of such music, which sounded to many like a set or random sounds.

However, experience shows that, since most public has to be told the meaning of something, true or supposed, often just a simple gesture, image or explanation opens up something akin to a revelation. Figure 6 shows a proposal for a “creative” display that automatically shows varying images meant to convey the mood of the piece – in this case, the segment Aum in the suite containing also Venus and Capricorn Rising from record Tauhid, by Pharoah Sanders, on Impulse, 1969. While the music, represented by the central (blue) stripe, scrolls with musical symbols over an imaginary landscape, pictures appear that represent ancient Egypt – pride of the ancient past heritage claimed by the composer – rising cosmological events, astrological charts, and the fallacy of a modern urban landscape that shakes while it hides social injustice. A generative grammar in XML can be used to that end, and it ensures that the user would approach the system with renewed curiosity, to discover new aspects of the music, since the display would look different every time it is accessed.

ACKNOWLEDGMENT


REFERENCES

[1] http://www.w3.3.org

For CV etc. of the authors Denis L. Baggi and Goffredo M. Haus, please consult the Editorial.