An IEEE Standard for Symbolic Music

Denis L. Baggi, University for Applied Sciences of Southern Switzerland

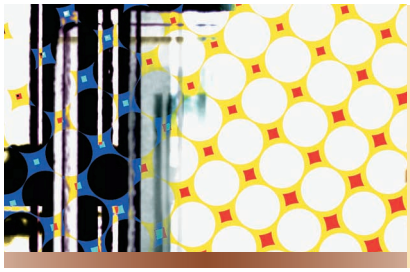
Music consists of more than *sound*. Popular audio standards, such as WAV and MP3, are encoded in binary form, unreadable by humans, and strongly dependent on current technology. In contrast, P1599, a new IEEE music standard, recommends representing and encoding with symbols expressed in XML, thus creating a readable, open, extensible, and durable format.

This standard strives to represent all aspects of music, namely, the *whole musical experience*, which is similar to understanding a narration and recognizing images, and may also consist of exploring how the music itself is built. The IEEE Computer Society's Standards Activity Board sponsors this work, which Intelligent Manufacturing Systems (www.ims.org), a global research fund, supports.

Music has always incorporated the newest technology of a given epoch—from medieval carillons to more recent fine mechanics and electronics. Its marriage with computer science is at least five decades old, and can further be traced back to early composing algorithms by Mozart and Haydn. Proposals for symbolic music standards have existed for more than 30 years.

SYMBOLIC REPRESENTATION EXAMPLES

The following examples illustrate the power of both this IEEE standard and of symbolic representations of music.



IEEE Standard P1599 opens a new world to music listeners.

- *An opera.* The viewer of a DVD of an opera can see the play, hear the music, view the score, read the libretto, and select excerpts of alternative renditions.
- *A jazz Big Band piece.* The user sees the harmonic grid and the soloist's name at the beginning of each solo in the piece, can examine the structures within the selection as well as alternate renditions of the piece (D. Baggi, "Understanding Jazz: The Structures of Swing," *Surveys of iCIMSI/SUPSI*, 2001).
- *A fugue.* The *theme* can be highlighted aurally and graphically as it moves across the different voices.
- *Music with a "program" or story.* Segments of Antonio Vivaldi's *Four Seasons*, for example, correlate with specific poems written by the composer.
- *Indian classical music.* As the melody unfolds in the *scale* of the raga, a diagram shows the "notes" of the melody.
- *An African drumming piece with several drums.* The user sees a graph illustrating the interplay of five drums to show, for example, that the various hits do not fall exactly on the beat.
- *Preservation of music heritage.* Symbolic encoding allows document storage in any media (G. Haus, "Rescuing La Scala's Audio Archives," *Computer*, Mar. 1998, pp. 88-89).
- *Musicological study.* Users can pose queries about a given musical characteristic—for example, a search for pieces that use the lowest note of a grand piano.

New applications based on P1599, among others in *ethnomusicology*—an area that has seen little investigation and methodological research—open a new world of research and enjoyment.

PAST MUSIC STANDARDS PROPOSALS

The idea of representing music with symbols is not new—music notation has existed for several centuries. Computer symbolic notations specifically have existed for several decades, as in the *Plaine-And-Easie Code* (B.S. Brook, *Musicology and the Computer*, City University of New York Press, 1970, pp. 53-56) and the DARMS project (R.F. Erickson, "The DARMS Project: A Status Report," *Computers and the Humanities*, June 1975, pp. 291-298).

More recently, developers have attempted to use SGML, a subset of which—the Standard Music Description Language—has been defined for music (S.R. Newcomb, "Standard Music Description Language Complies with Hypermedia Standard," *Computer*, July 1991, pp. 76-79). Though well defined, SMDL failed to attract much attention because of its lack of applications.

Two de facto standards that use XML inspired the development of P1599, namely, *MusicXML* and *MEI*. Dozens of existing applications already support Recordare's MusicXML (www.recordare.com/xml.html). Further, to encode music, the University of Virginia's Digital Library created the Music Encoding Initiative (www.lib.virginia.edu/digital/resndev/mei/mei_ocve.pdf).

P1599 will be compatible with these established standards and with the Reference Model 0 (RM0) music core of ISO/IEC JTC1/SC29/WG11—the coding of moving pictures and audio—now in development.

HOW P1599 WORKS

P1599 considers a musical structure to consist of different layers, each of which represents an aspect of the piece. Figure 1 shows that, in symbolic music information (SMI), the group that the general, structural, and logic layers constitute points to graphical notations expressed, for example, as TIFF, JPEG, or GIF files for a score or a diagram; to audio formats with MP3 or WAV; and to performance standards such as MIDI. The logic layer comprises the logical organized symbols and spine, both denoted in XML.

This means that, for musical representation, the standard chosen to express its audio content is irrelevant because SMI uses the same logical symbol file. Depending on the kind of music, not all layers may be present. For example, a harmonic grid makes no sense for a Gregorian chant, while a score has little relevance for a jazz piece from a jam session.

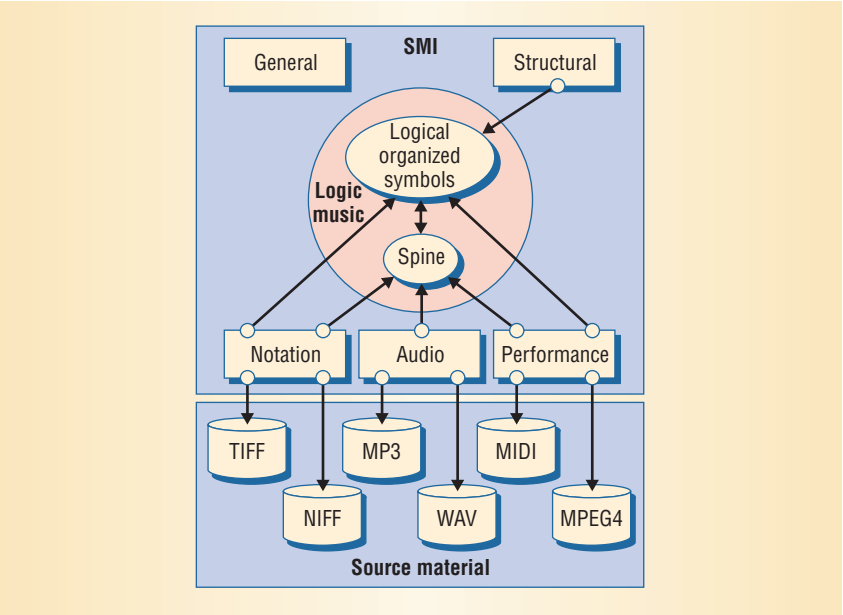


Figure 1. Relationships among symbolic music information layers and between SMI and the source material.

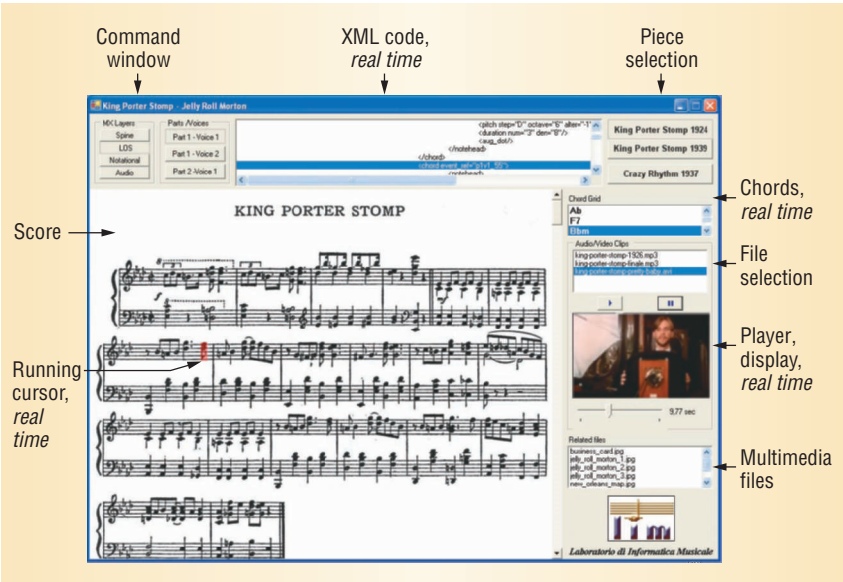


Figure 2. Music browser. Windows with real-time captions update synchronously while the music plays.

P1599 IN A MUSIC BROWSER

Built at the University of Milan's Laboratory for Musical Informatics, this browser application illustrates the power of the standard and how it works. The screenshot in Figure 2 contains different windows—those with the caption *real-time* update synchronously while the music plays.

Beginning with the *piece selection* window, users have three choices: *King Porter Stomp 1924*, *King Porter Stomp 1939*, and *Crazy Rhythm*. The first two refer to two published scores of that

piece, composed by American musician and pianist Jelly Roll Morton, aka Ferdinand Joseph Lamothe (1890-1941). The third is an improvised jazz piece that has no score.

In the *file selection* window, the user can choose among various multimedia files. In this case, the choices include a 1926 recording and a MIDI version of the 1924 score, both encoded in MP3. An excerpt from the 1978 Louis Malle film *Pretty Baby* provides a third choice—the movie features a character modeled after Morton, who composes the piece in the background. The *player, display* window shows the movie; for plain music, this window would instead show the image of a common nonvideo player.

When the user selects the third option, several synchronized activities start and execute. The movie segment begins and the music plays. On the score, the *running cursor* indicates what is being played—in this case, the end of the seventh bar. The user can move the red cursor and restart play from another point in the score, while the other real-time windows and the player cursor adjust simultaneously.

The *XML code* window shows the encoded events of the logical organized symbols scrolling with the music. In the command window, the user can select which XML code is to be displayed—spine, LOS, notational, or audio—and select any of the three voices for the running cursor to follow.

The *chords* window displays the piece's harmonic elements, again synchronously with the playing, while the *multimedia files* window allows featuring pictures of Morton, his band, and other curiosities.

The screenshot for the jazz piece *Crazy Rhythm*, not shown here, displays each soloist's name and picture following the running cursor on the harmonic grid. Four saxophone solos, including one each from top American jazz musicians Coleman Hawkins and Benny Carter, are bracketed between the theme exposure by the whole band.

To conclude, the browser displays music represented with accessible, readable symbols and by the score. Since a single XML file is used for several renditions of the same piece, the format used for reproducing the audio becomes irrelevant.

Progressing from closed, binary audio standards to natural, open symbolic music representation is timely and appropriate. IEEE P1599 will provide an enabling technology that can serve as the basis for realizing an unending stream of new applications.

The attentive listener senses that a whole world lies beyond just the sound of music, waiting to be explored. Formerly, a listener would need years of formal, disciplined study to advance from music enjoyment to music education, but applications using this standard will make the progression easier and faster for every listener.

P1599 offers an example of how technology lets both users and specialists discover the deep meaning of music and generate connections to all aspects of culture, from art enjoyment to contemporary life.

A first draft has been written, and P1599's creators expect a complete standard by 2007. ■

***Denis L. Baggi**, a member of IEEE CS TAB and SAB, is a faculty member at the University for Applied Sciences of Southern Switzerland near Lugano as well as the Working Group Chair for P1599. Contact him at denis.baggi@supsi.ch. Additional information available at www.computer.org/portal/pages/ieeecs/communities/standards/1599/par.html.*

Editor: Jack Cole, US Army Research Laboratory's Information Assurance Center, jack.cole@ieee.org; http://msstc.org/cole