XML Music Performance Description

Extended Abstract

Axel Berndt
Center of Music and Film Informatics, University of Music Detmold, Detmold, Germany
berndt@hfm-detmold.de

Benjamin Wolff Bohl
Center of Music and Film Informatics, University of Music Detmold, Detmold, Germany
bohl@edirom.de

Motivation and Related Work

Haus & Longari (2005) distinguish six layers of music information: general, structural, music logic, notation, performance, and audio. Regarding the performance layer, most representation formats provide either very technical information (e.g., MIDI) or very rough information (MusicXML, MEI). While the technical information is rich in details, precise and unambiguous it lacks a musically meaningful abstraction. Abstract representations, on the other side, are easy to read and write into music scoresbut ambiguous and incomplete. How loud is forte? How fast is allegro?

Throughout the past two decades, several research projects contributed to these questions in order to bridge the gap between abstract information and technical implementation. Solely analytical projects investigated the phenomena of music performance and their shaping by human musicians (Hähnel 2013). Generative projects were developing systems that (semi-)automatically create expressive performances (Mazzola, Göller & Müller 2002, Widmer & Goebl 2004, Flossmann, Grachten & Widmer 2009, Friberg, Bresin & Sundberg 2006). The data models for describing these performances were rarely a matter of deeper discussions. Hence, they never reached any relevance beyond the scope of their corresponding systems.

Our XML-based performance description format was likewise solely dedicated for usage within our performance rendering engine (Berndt 2011). However, this engine had several different applications throughout the past years, e.g. in the context of games music research, historically informed performance research, listening studies, and music production. This motivated us to make some deeper reflections on the format and its perspectives which marks the main concern of this contribution.

General Concept

An expressive performance of a musical piece transforms the musical raw material (pure notes) into a sounding output. From this a central design decision of our approach derives: the clear separation of the musical raw material (MIDI data, in our case) and the performance description (as XML format). The MIDI file contains only the notes of the music, exactly as notated, with no performance related information. The XML file provides one or more performance descriptions of the piece which we call *performance styles*.

Our software implements the techniques to render this information into expressive MIDI sequences that are output directly or sent to a *Digital Audio Workstation* for music production. The software can make a seamless transition between different performance styles which is, for instance, used for adaptive game scoring. In a listener study we used this interactivity for an analysis-by-synthesis approach where the participants could adjust certain performance parameters according to the tasks they were given (reference omitted).

The XML format is based on the formal models that were introduced by (Berndt 2015). These cover a wide range of performance phenomena, including tempo and dynamics curves, articulation, metrical accentuation, asynchrony, and rubati.

Basically, the format groups all data into global information, applying to all musical parts, and local information, applying to single parts. Local information dominates the global. The advantage of this rule pays off in solo-plus-accompaniment constellations. Global performance instructions are executed by the whole ensemble; only the solo part features its own local instructions that "overwrite" the global. Local information can be defined for each part individually. Technically, the parts correspond with MIDI channels.

Both, global and local information, are subdivided into header and dated information. Header information applies to the whole movement, e.g. the definition of articulation styles and MIDI controller mappings. Dated information is organized in sequential lists, so-called *maps*, one for each type of performance features. Each map element demands a date attribute (in MIDI ticks). Discrete performance features, e.g. articulation instructions, are applied at the respective time position in the MIDI sequence. Most performance features, however, extend over a certain timeframe. A dynamics instruction, for instance, reaches from its date until the date of the succeeding instruction in the same dynamics map. Furthermore, rubato instructions and metrical accentuations define a certain timeframe; in case of metrical accentuations typically one measure. The schemes are then repeatedly applied (e.g., measure-wise) until the succeeding instruction in the same map defines a new scheme and framelength.

A more extensive introduction of the XML model cannot be given in this extended abstract, though, but will be subject of the presentation.

Reflections and Perspectives

The format enables very rich, nuanced music performances. The interplay of global and local information is the key to complex polyphonic performance structures. Each part can feature its own performance plan.

But where do these information come from? The performance instructions that the format conveys are explicit and numerically precise. The models underlying the format act as interfaces, the user adjusts high-level parameters of the features while the details are generated. However, the creation of a fully fledged performance is, nonetheless, an excessive amount of work. This must be supported by editing tools, such as music notation software plugins. Reduced editing workload can also be achieved by semiautomatic approaches. Rough performances may be generated automatically. The human editor then corrects and incorporates further details.

The format's close connection to the MIDI standard reflects the direct coupling to music production as the core application domain. This may imply some disadvantages such as bad readability. A better compatibility with other formats also demands a flexible numeric basis of the attributes. This can also help overcoming the inherited numeric restrictions of MIDI.

Our performance description format will be complemented by a score representation that follows the same basic concepts of global, local, header, and dated information and provides information about the musical macro structure (sections, phrases, figures, and time signature). This is already used in the context of adaptive games music. MEI support will be implemented via converters translating MEI into our new format which is developed with the focus on a better computer processability and serves as intermediate format for MEI-to-MIDI conversion.

References

Berndt, A. (2011), *Musik für interaktive Medien: Arrangement- und Interpretationstechniken*, Verlag Dr. Hut, Munich, Germany.

Berndt, A. (2015), Formalizing Expressive Music Performance Phenomena, *in* A. Berndt, ed., 'Works in Audio and Music Technology', TUDpress, Dresden, Germany, chapter 4, pp. 97–128.

Flossmann, S., Grachten, M. & Widmer, G. (2009), Expressive Performance Rendering: Introducing Performance Context, *in* 'Proc. of the 6th Sound and Music Computing Conf. (SMC)', Porto, Portugal.

Friberg, A., Bresin, R. & Sundberg, J. (2006), 'Overview of the KTH Rule System for Musical Performance', *Advances in Cognitive Psychology, Special Issue on Music Performance* **2**(2–3), 145–161.

Haus, G. & Longari, M. (2005), 'A Multi-Layered, Time-Based Music Description Approach Based on XML', *Computer Music Journal* **29**(1), 70–85.

Hähnel, T. (2013), *Baroque performance—a research study on characteristic parameters of 18th Century music*, Vol. 2 of *Studies in Cognitive Musicology*, Electronic Publishing Osnabrück, Osnabrück, Germany.

Mazzola, G., Göller, S. & Müller, S. (2002), *The Topos of Music: Geometric Logic of Concepts, Theory, and Performance*, Birkhäuser Verlag, Zurich, Switzerland.

Widmer, G. & Goebl, W. (2004), 'Computational Models of Expressive Music Performance: The State of the Art', *Journal of New Music Research* **33**(3), 203–216.