Cooperative content and metadata generation in music archives

Adriano Baratè, Luca A. Ludovico, Alberto Pinto, Paolo Tagliolato
LIM – DI*Co* - Università degli Studi di Milano
Via Comelico 39/41 – I-20135 Milano, Italia
{barate, ludovico, pinto, tagliolato}@dico.unimi.it

1 IEEE 1599, a New Standard for Music

Our proposal is based on the pre-existence of a centralized set of symbolic descriptions of music pieces, whose data and metadata are available for the user through a Web interface. The goal is not only allowing the access to such contents, but also providing the possibility to add further information, thus enriching the original corpus in a collaborative way. A particular XML format is proposed in order to encode the central database of music pieces, namely the new IEEE 1599 standard.

This project aims at representing music in a comprehensive way, thus opening up new ways to make both music and music-related information available to musicologists and performers on one hand, and to non-practitioners on the other. Its ultimate goal is to provide a highly integrated representation of music, where score, audio, video, and graphical contents can be appreciated together. In fact, in a single file, music symbols, printed scores, audio tracks, computer-driven performances, catalogue metadata, text and graphic contents related to a single music piece are linked and mutually synchronised within the same document. Heterogeneous contents are organised in a multilayered structure that supports different encoding formats and a number of digital objects for each layer [1].

Not all layers must, or can, be present for a given music piece. Of course, the higher their number, the richer the musical description. Richness has been mentioned in regard to the number of heterogeneous types of media description, namely symbolic, logic, audio, graphic, etc. But the philosophy of the proposed format allows one extra step, namely that each layer can contain many digital instances. For example, the *Audio* layer could link to several audio tracks, and the *Structural* layer could provide many different analyses for the same piece. The concept of *multilayered description* – as many different types of descriptions as possible, all correlated and synchronised – together with the concept of *multi-instance support* – as many different media objects as possible for each layer – provide rich and flexible means for encoding music in all its aspects.

In this architecture, the central repository of music pieces is not a standard relational database, rather it is a collection of IEEE 1599 files. Relationships among different music entities can be retrieved by parsing XML files. In particular, the relationships which are internal to a single music piece can be discovered by analyzing a specific layer of the file. For instance, the description of all the available analyses associated to a composition can be found inside the *Structural* layer of the corresponding file. Other relationships, which typically relate many pieces, can be retrieved by parsing the same element inside the whole music repository. For example, in order to extract all the pieces written for the same ensemble and belonging to the same historical period, it is sufficient to query properly the *Logic* and the *General* layers. For further details about IEEE 1599 format and clarifying examples, please refer to [1] [2].

2 Framework Architecture

The framework we propose is based on a central repository constituted by a set of IEEE1599 files, whose information contents include both descriptions of scores and media objects. In this approach, all the data and metadata related to a single music piece are embedded into one XML document, that is, in a certain sense, a database itself, with its entities and internal relationships. We propose to consider a large repository of interacting IEEE1599 files, in order to provide a very rich environment for Music Information Retrieval.

In our proposed architecture one of the central ideas is the double role of users, that could be either passive users, accessing the repository only to view its contents, or active users, who interact with repository contents in a collaborative way.

Typically, a passive user queries the system to retrieve data and which someone else has entered. Even if passive users do not add contents to the repository, their experience has not to be limited to a mere exploration of metadata, but they can also download a complete package that connects to the system and contains some of the IEEE1599 applications we proposed in past works (see [3], [4], [5]).

The active user, after logging into the system, interacts with a number of Web applications, entering new data and metadata automatically incorporated to the IEEE1599 file that describes a music piece already present in the repository.

In this manner the concept of social tagging is expanded, not only including the tagging of objects and concepts, but extending the field of application to new data association in a cooperative environment.

3 Social Tagging

The concept of *social tagging,* a part of the Web 2.0 generation of cyberspace, permits users to play a big part in data and metadata creation.

As described, in our framework there is a number of already created MX files that represents music pieces, with a set of information necessarily fixed but also with other data and metadata that can be integrated by so-called active users.

As an example, let us consider the IEEE1599 file that describes a jazz piece named “Jean Pierre”, by Miles Davis: while this file must contain the title, the author and a logic representation of the piece itself in terms of notes and pauses, it can be also catalogued in terms of genre as *jazz*, or *fusion*, or *jazz-rock*. Since this information depends on the users’ perspective, it has to be modifiable. This is a typical example where social tagging can provide a flexible and effective way to support analysis-oriented applications.

4 Music Analysis and Social Tagging

MX Analysis element, within structural layer, is devoted to the storage of analytical interpretation of music. This element and its sub-elements follow a formal model [9] based on the existing different analytical approaches. The model sketched in this section is composed of a basic structure, which is a sort of skeleton of an analysis, and of an extensible part devoted to contain the aspects of a specific analytical approach.

The basic structure of an analysis, common to almost all analytical approaches (e.g. set-theory [6], Paradigmatic analysis [8], automated computational analyses like in [7], etc.) is the segmentation of a piece. In each school of music analysis, segmentation has its own specificities in term of when and how it is conducted, but if we define "segmentation" the result of the process, we can say that it is simply a collection of segments, a segment being a set of music events in the music piece.

The exhibition of a network of relationships between segments is the second aspect to model. This is done, during an analysis, accordingly the specific methodology in action, but at a first step we can simply notice the presence of relations between segments. Segmentation, segments and relations between segments are the very basic structure contained in MX format to represent music analyses. We will intend this structure as the formal shadow of an analysis of a piece of music.

To analyze music, a particular musical sensibility is needed. One has to concentrate on particular music features. First of all it is necessary to decide which musical feature we are interested in. To this purpose, many approaches associate feature vectors to segments (like in [10]). We think that the more general concept of object could offer a more powerful possibility: in an object oriented perspective, one can define his own feature object class in order to model his own music attitude. In a limit case a feature object could be just a textual description (not so naturally formalizable within a vector). If we have a segmentation Σ we can imagine an arrow into an abstract feature object class Ω:

$$f: \quad \Sigma \to \Omega$$
$$\sigma \mapsto \omega$$

Depending on the particular decisions for a concrete derived class of objects, let's call it S (Space of objects), a part of an analysis will be based on a new association:

$$f_{my}: \quad \Sigma \to S$$
$$\sigma \mapsto s$$

Now we can predicate something on the segments $\sigma \in \Sigma$ in our particular point of view by means of the featureObjects $s \in S$. In this perspective it is possible to relate segments by means of relating objects associated to them.

Definition. Given a binary relation $\triangleright$ between objects of class S, we can define a binary relation $\triangleleft$ induced by f_{my} on the segments of Σ:

$$\sigma_1 \triangleleft \sigma_2 \Leftrightarrow f_{my}(\sigma_1) \triangleright f_{my}(\sigma_2)$$

In certain cases there will be formal rules to establish such relations. In certain cases there will be methods to say if a segment is related to another and in which way it is.

Within Analysis element, many interpretations of the same music piece can be separately stored.

The implementation of the extensible aspects of the formal model lets an analyst to represent the methodological specificities of his approach, and to share them by means of the object classes that can be re-used by himself or other theorists for future works. At the same time the model structures all the analyses in a homogeneous way, which is a clear advantage for the comparison of different works. The benefits for a work of comparative analysis are quite clear.

From an educational and popularizing perspective, the inclusion of analyses in MX gives rise to new possibilities for users interested in music analysis. Thanks to synchronization of MX layers, it will be possible to follow an analysis (or to compare more than one) of a music piece looking at the segments and their relationships annotated on the score, and at the same time hearing their sounds in some audio versions of the piece.

References

1. Ludovico, L.A.: Outline of the MX Standard. In: Proceedings of the 2nd conference on Semantic And digital Media Technologies (SAMT 2007), Genova, Italy, 2007
2. Pinto, A., Haus, G.: A novel XML Music Information Retrieval Method using Graph Invariants. ACM Transactions on Information Systems, Vol.25, Issue 4, 2007
3. Baggi, D.L., Baratè, A., Haus, G., Ludovico, L.A.: A computer tool to enjoy and understand music. In: Proceedings of EWIMT 2005 – Integration of Knowledge, Semantics and Digital Media Technology (2005), 213–217
4. Baratè, A., Haus, G., Ludovico, L.A.: MX Navigator: An Application for Advanced Music Fruition. In: Proceedings of AXMEDIS 2006 - 2nd International Conference on Automated Production of Cross Media Content for Multi-channel Distribution, pp. 299-305, Leeds, UK, 2006
5. Baratè, A., Ludovico, L.A.: An XML-based Synchronization of Audio and Graphical Representations of Music Scores. In: Proceedings of the 8th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS 2007), Santorini, Greece, 2007.
6. Allen Forte. The Structure of Atonal Music. Yale University Press, New Haven and London, 1973
7. Benoit Meudic. Musical pattern extraction: from repetition to musical structure, in Proceedings of CMMR (Computer Music Modeling and Retrieval)- Montpellier, May 2003
8. Jean-Jacques Nattiez. Analyse musicale et sémiologie: à propos du prélude de 'pelléas'. Musique en jeu, (10), 1973
9. Paolo Tagliolato. Music structure representation: a formal model. In Proceedings of Sound and Music Computing 2006, Marseille, France, May 2006
10. Christina Anagnostopoulou and Gert Westermann. Classification in music: a computational model for paradigmatic analysis. In Proceedings of the International Computer Music Conference, 1997