Ontology-Based Annotation of Music Scores

Samira Si-said Cherfi
CEDRIC/CNAM
Paris, France
samira.cherfi@cnam.fr

Christophe Guillotel
IReMus, Sorbonne Universités
Paris, France
christophe.guillotel@cnrs.fr

Fayçal Hamdi
CEDRIC/CNAM
Paris, France
faycal.hamdi@cnam.fr

Philippe Rigaux
CEDRIC/CNAM
Paris, France
philippe.rigaux@cnam.fr

Nicolas Travers
CEDRIC/CNAM
Paris, France
nicolas.travers@cnam.fr

Abstract

Digital music scores are a way to present music notation and lack of semantic information useful for musicology purposes in order to manipulate music concepts. We propose a general approach to extend score encodings with semantic annotations. It relies on an ontology of music notation designed to integrate semantic music elements either extracted or produced by a knowledge process. We illustrate the whole mechanism by extracting RDF facts based on the identification of dissonances in Renaissance counterpoint.

CCS Concepts • **Information systems → Web Ontology Language (OWL)**; *Resource Description Framework (RDF)*; • **Computing methodologies → Ontology engineering**;

Keywords Music Notation Ontology, Knowledge extraction, Annotation, SWRL

ACM Reference Format:
Samira Si-said Cherfi, Christophe Guillotel, Fayçal Hamdi, Philippe Rigaux, and Nicolas Travers. 2017. Ontology-Based Annotation of Music Scores. In *K-CAP 2017: K-CAP 2017: Knowledge Capture Conference, December 4–6, 2017, Austin, TX, USA.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3148011.3148038

1 Introduction

Many digital music scores are now produced in XML-encoded formats such as MusicXML [6], MEI [14, 18], or MNX [15]. Those formats share a common approach to encode the complete content of sheet scores, including rendering instructions. While the initial motivation was mostly to address interoperability concerns. Each note can be described with attributes (pitches, octaves, alterations, durations...). Exciting perspectives are allowed regarding the emergence of new applications: automatic transformations (e.g., transpositions), audio synchronization, multimodal interactions, and, of course computer-aided analysis of the music itself.

Since musical scores encode semantic categories of music, they are subject to an analysis that identifies the linguistic structures (phrases, cadences, dissonances...). Those high-level concepts are implicit in score encoding, thus, they need to be extracted by an analytic process, and linked to the score fragment they are related to. This completion of the notation would help to figure out the score organization and the specificities of its musical language.

A first, immediate application is a visualization mechanism. For instance, someone wanting to see the occurrences of a given harmonic sequence would be proposed a score representation with highlighted relevant fragments. For large digital libraries of music scores, this knowledge can also leverage meaningful searches such as, for instance, finding pieces featuring a particular cadence formula. It constitutes a fruitful framework for digital musicology enhancement.

In the present paper, we propose a framework for managing the knowledge extraction process from music scores. It aims at supplying a generic approach to represent, extract, link and search music language concepts that can be inferred from the music notation. We will illustrate it with the analysis of dissonance in the late Renaissance counterpoint. More specifically, we propose the following contributions:

1. an ontology of music notation content, focused on the core "semantic" information present in a score and subject to the analytic process;
2. a modeling of high-level analytic concepts, in such a way that concept instances can be associated, as annotations, with fragments of the score;

Section 2 reviews the state of the art and we present the notation ontology in Section 3. Section 4 details the knowledge management process. Section 5 concludes the paper.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
K-CAP 2017, December 4–6, 2017, Austin, TX, USA
© 2017 Copyright held by the owner/author(s). Publication rights licensed to Association for Computing Machinery.
ACM ISBN 978-1-4503-5553-7/17/12…$15.00
https://doi.org/10.1145/3148011.3148038

2 State of the Art

Digital musicology Software systems with implications begun with the Humdrum toolkit by David Huron [8]. Melisma system [22] or Music21 [4] extended it as independant modules that can be combined for analytical tasks. They draw on various heuristic approaches, while we rather target fully automated algorithms for structured score encodings.

Interestingly, strict counterpoint is one of the first musical idioms to be investigated by means of computer software. The Palestrina program [12], designed for pedagogical purposes, performed a detailed analysis of two-part species counterpoint. The first species counterpoint module was implemented in Music21 and the `cint` extension module in Humdrum. Drawing on Jeppesen's [9] distinction of five main dissonance categories in strict counterpoint, Ceulemans [2, 3] proposed a fully implementable model, allowing for the unambiguous identification of these dissonance types. Our paper takes this model and the dissonance schemata developed by Sigler et al. [21] as a point of departure.

Ontologies, knowledge extraction & annotations Ontologies are characterized by their ability to describe concepts at a fine-grained level and in a structured manner. Ontology-based information extraction (OBIE) is as a subfield of information extraction (IE) [24] in which ontologies play a central role. The general idea is to use information extraction to retrieve instances and values related to the classes and properties of the ontology.

Semantic annotation is the process of adding information to digital content. According to [23], several factors should be taken into account when choosing an annotation strategy. First, developing knowledge is a heavy investment. Semantic web technologies [13, 19, 20] are thus recommended for publishing, sharing, syndicating, and distributing web annotations. They suggest to formally identify concepts and relations in documents which could help for information retrieval, inferences and integration of sources.

Second, annotations are generally usage oriented. One could annotate a music score for harmony (merging ontologies) or performance analysis (like in KIM [11]). The second solution is preferred for evolution as the annotation should preserve its consistency. We will exhibit modularity in our approach, with the core *notation ontology* that captures semantic information from the notation, and *analytic ontologies* that cover the specific concepts to a particular application.

Musical Ontologies [5] proposed an ontology on musical works but with very high level concepts without notations. While the *Music Ontology* [17] describes a more content oriented usage for basic information about musical works and artists which integrates whole scores dedicated for many music-oriented web services. Similar works [1, 10] propose music vocabularies which describes classical music and performances. These ontologies describe music at a work level and are suitable to provide general information such as music categories (*e.g.,* Chamber music) and performances (*e.g.,* musicians). More content oriented ontologies are suitable.

We do believe that such a model would be of invaluable help to let the community formalize discussions and proposals and address issues related to the topic. We hope that the present work, although focused on its scope of application, can serve as an encouragement to initiate such an endeavour.

3 The MusicNote ontology

Our core *notation ontology*, called MusicNote, aims at capturing the part of music notation that is necessary and sufficient to support a music analysis process. It therefore abstracts the "semantic" music content that is explicitly found in a score encoding without the graphical layout of scores.

3.1 Structural aspects

The `MusicNote`[1] ontology relies on the structural aspects of a score. Essentially, a score is modeled as a *hierarchical structure*, where leaves consist of *voices*, and inner nodes of *parts*. A voice is a sequence of *events*, occurrences of the abstract class `Event` which is refined in several sub-classes.

A score is made of *parts*. Class `Part` represents an abstract concept which is refined in: (i) A *group* (of parts) consists of a set of subparts, and mostly serves the organizational aspect of the score (*e.g.,* two groups for wind instruments and string instruments for a concerto), and (ii) a *single part* encapsulates the music notation elements assigned to an individual performer (instrument or vocal). Fig. 1 shows for instance parts for the piano, the violins, cellos, etc. A single part contains one or several *voices*.

3.2 Voices and events

A *voice* is essentially a time series of *events*, where an event denotes the production of a sound artifact at a specific timestamp (the "onset"). Thus, a voice is modeled as a function from the time domain where there is exactly one event at each instant. In other words, events cannot overlap.

Fig. 2 shows a voice with its time series of events. The non-overlapping constraint implies that the *duration* of an event covers the time range between its onset (included) and the onset of the next event in the series (excluded).

The event concept can be refined: sound (`SoundEvent`), text or syllables to be sung (`SyllEvent`), or actually any piece of information attached to a time period. The `SoundEvent` itself is decomposed in `Note` (non-decomposable), and `Chord`, composed of at least two notes with the same duration.

The measures are represented at the part level, which means that voices from a part share the same temporal range from measures. A property of a measure is the *time signature*, the value of which can vary from one measure to another. Each event is linked to the measure from its onset (Fig. 2).

[1] http://cedric.cnam.fr/isid/ontologies/files/MusicNote.html

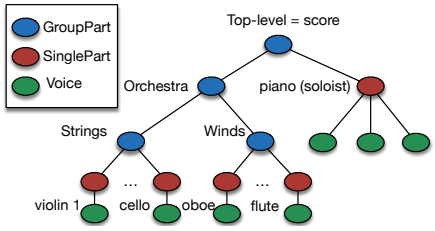


Figure 1. Structure of a score

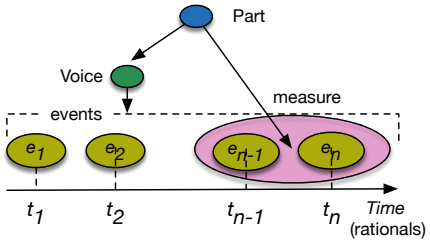


Figure 2. Voice as a time series of events

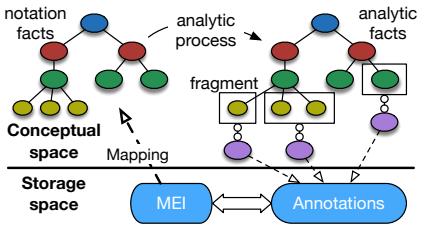


Figure 3. Annotation mechanism

4 Modeling analytic concepts

Assume now that we aim at developing an analytic model of some musical language. An expert analyst has to express this model as concepts that need to be integrated in our core ontology, as a new module called *analytic ontology*.

To better understand this modular aspect, consider the generic analytic process illustrated by Fig. 3. Starting from a score encoding such as MEI, the ontological space is populated with *notation facts* thanks to a mapping that extracts the "semantic" part of the score. They are instances of the *core notation ontology* and cover the necessary information to carry out an analytic procedure. The process produces *analytic facts*, instances of concepts taken from the *analytic ontology*. They qualify fragments of the notation, e.g., as dissonances, or any other semantic artifact.

Note that the formalization of an analysis acts as a proof of completeness of the core ontology. It might happen that a concept qualifies some aspect of the music notation which is not captured by the ontological representation. Moreover the algorithm might need an information that has been omitted. We illustrate these ideas with the concrete example of dissonance modeling in Renaissance counterpoint.

4.1 Dissonance in Renaissance counterpoint

Our dissonance typology is derived from Knud Jeppesen's [9] distinction of five main dissonance types in Renaissance counterpoint: suspension, anticipation, passing note, neighbor note and escape note.

To resume the model of these patterns in two-voices counterpoint, we take into account harmonic, melodic and metric criteria at three distinct moments in the contrapuntal stream: (1) the event immediately preceding the dissonance, (2) the dissonance's impact and (3) the event immediately following the dissonance. The dissonant notes are highlighted in red in the examples of Fig. 4.

The *suspension* implies that voice 1 remains in place at the impact while voice 2 moves subsequently and then conversely. The *anticipation* constitutes the exact mirror of this pattern with a weak beat. Constant melodic movements at voice 1 while voice 2 remains in place implies different dissonances: *passing note* (unidirectional and parsimonious), the *neighbor note* (bidirectional and parsimonious) and the *escape note* (unidirectional and non-parsimonious).

4.2 The Counterpoint module

The dissonance analysis can be expressed in terms of our ontology concepts as follows: a dissonance pattern is a relationship between two events e_1 from voice 1 and e_2 from voice 2. By considering each pair of events (e_1, e_2) such that their time range overlaps.

The *MusicNote* ontology gives `Events` and `Voices` concepts to support the rule requirements. However, the `Event` concept needs to be extended with a *metric weight* property, and a *consonant* property to the `Interval` concept.

Then, we create the Counterpoint module which contains the concepts of this analytic model, imported on demand. It consists in a `Dissonance` concept that identifies a dissonance pattern occurrence as a relationship between two events of distinct voices, and `Dissonance` subtypes. This "analytic" ontology can be integrated to our notation ontology thanks to standard mechanisms [16].

4.3 Knowledge extraction and score annotation

Regarding the implementation of the analytic process, we rely on a knowledge-based approach with a rule-based language. As an example of these benefits, we express in the following our dissonance model as rules in a standard language of the Semantic Web, SWRL [7]. RDF facts are extracted by an XSLT[2] process which computes every concept, relation and event positions from the MEI file.

The general *Dissonance* concept can be specified as follows: if we consider the list of event pairs (e_i, e_j), taken from two distinct voices v_1 and v_2, such that their time range overlaps, then (e_1, e_2) is an instance of a canonical dissonance pattern if interval $< e_1, e_2 >$ is *dissonant* and the pairs (e_m, e_n) and (e_p, e_q) that respectively precedes / follows (e_1, e_2) are both *consonants*. Intuitively, the following rule expresses three intervals, the pair itself (dissonance), its successor and its predecessor. Recall that the `Cons` property is introduced in our ontology to qualify intervals.

[2] http://cedric.cnam.fr/vertigo/MEI2MusicNote.xsl

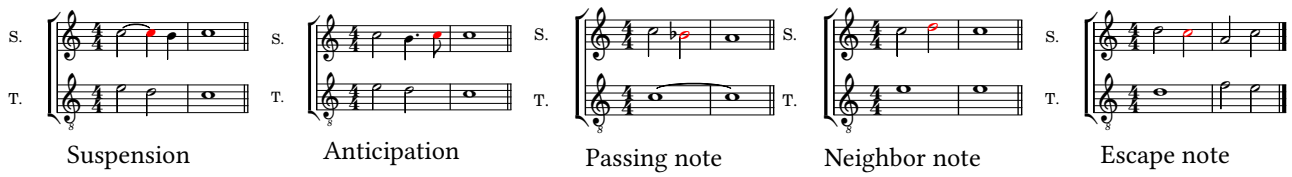


Figure 4. Dissonances examples

```
Dissonance(v1[pos], v2[pos]) :-
        not Cons (Interval(v1[pos], v2[pos])),
        Cons (Interval(v1[pos-1], v2[pos-1])),
        Cons (Interval(v1[pos+1], v2[pos+1]))
```

The *Dissonance* rule can be used as a "macro" to define the five subtypes that we model. A canonical *Suspension* for instance is a dissonance such that the dissonant interval occurs on an accentuated beat and the voice v_1 remains in place while voice v_2 moves downward:

```
Suspension(v1[pos], v2[pos]) :-
        Dissonnance (v1[pos], v2[pos]),
        v1[pos] = v1[pos-1], v2[pos] != v2[pos+1],
        Weight(pos) >  Weight(pos-1)
```

These examples show the advantages of a rule-based approach, with a clear, concise and unambiguous expression that serves both as a definition of the *Dissonance* concept, and generation of the facts. Clearly, some analytic procedures will fall beyond the rule language expressiveness, in which case a pure algorithmic approach can be used.

5 Conclusion

We developed a generic methodology for knowledge modeling, extraction and exploitation. It relies on the core *MusicNote* ontology, and a specialized analytic model. We created one such model, for Renaissance counterpoint analysis, and implemented our methodology in our digital library[3].

This first achievement can now be extended to enable a collaborative annotation system, opened to external contributions. This endeavor hopes to contribute, in the context of music notation, to the semantic web vision of an open digital space of knowledge-annotated documents.

References

[1] Manel Achichi, Rodolphe Bailly, Cecile Cecconi, Marie Destandau, Konstantin Todorov, and Raphaël Troncy. 2015. DOREMUS: Doing Reusable Musical Data. In *International Semantic Web Conference*, Vol. 1486. CEUR-WS.org.

[2] Anne-Emmanuelle Ceulemans. 1996. Les dissonances dans la polyphonie des 15e et 16e siècles : Quelques indications en vue d'une analyse assistée par ordinateur. *Musurgia* III, 2 (1996), 7–21.

[3] Anne-Emmanuelle Ceulemans. 1996. *Une approche rythmique, mélodique et harmonique de la polyphonie vers 1500. Les Kyrie de Josquin des Prez.* Ph.D. Dissertation. Université Catholique de Louvain.

[4] Michael Scott Cuthbert and Christopher Ariza. 2010. Music21: A Toolkit for Computer-Aided Musicology and Symbolic Music Data. In *Proc. Intl. Conf. on Music Information Retrieval (ISMIR)*. 637–642.

[5] Leigh Dodds. 2004. MusicBrainz Metadata Vocabulary. (2004). http://www.ldodds.com/projects/musicbrainz/schema/mb.html

[6] Michael Good. 2001. *MusicXML for Notation and Analysis.* W. B. Hewlett and E. Selfridge-Field, MIT Press, 113–124.

[7] Ian Horrocks, Peter F. Patel-Schneider, Harold Boley, Said Tabet, Benjamin Grosofand, and Mike Dean. 2004. SWRL: A Semantic Web Rule Language Combining OWL and RuleML. W3C Member Submission. (2004). http://www.w3.org/Submission/SWRL/

[8] David Huron. 2002. Music Information Processing Using the Humdrum Toolkit: Concepts, Examples, and Lessons. *Computer Music Journal* 26, 2 (July 2002), 11–26. https://doi.org/10.1162/014892602760137158

[9] Knud Jeppesen. 1992. *Counterpoint The Polyphonic Vocal Style of the Sixteenth Century* (third ed.). Dover.

[10] Masahide Kanzaki. 2003. Music Vocabulary. (2003). http://www.kanzaki.com/ns/music

[11] Atanas Kiryakov, Borislav Popov, Ivan Terziev, Dimitar Manov, and Damyan Ognyanoff. 2004. Semantic annotation, indexing, and retrieval. *Web Semantics: Science, Services and Agents on the World Wide Web* 2, 1 (2004), 49–79.

[12] Peter M., Eleanor Selfridge-Field, Suzel A. Reily, and Anthony Pople. 2017. Computers and music. *Grove Music Online.* (2017). http://www.oxfordmusiconline.com/subscriber/article/grove/music/40583.

[13] Deborah L McGuinness, Frank Van Harmelen, et al. 2004. OWL web ontology language overview. *W3C recommendation* 10, 10 (2004), 2004.

[14] MEI 2015. Music Encoding Initiative. http://music-encoding.org. (2015). Accessed March 2017.

[15] MNX 2017. MNX Proposal Overview. https://w3c.github.io/mnx/overview/. (2017). Accessed Oct. 2017.

[16] Jeff Z Pan, Luciano Serafini, and Yuting Zhao. 2006. Semantic import: An approach for partial ontology reuse. In *Proc. of Int. Conf. on Modular Ontologies-Volume 232.* CEUR-WS. org, 71–84.

[17] Y. Raimond, S. Abdallah, M. Sandler, and F. Giasson. 2007. "The Music Ontology". In *Proc. Intl. Conf. on Music Information Retrieval (ISMIR).* 417–422.

[18] Perry Rolland. 2002. The Music Encoding Initiative (MEI). In *Proc. Intl. Conf. on Musical Applications Using XML.* 55–59.

[19] R Sanderson, P Ciccarese, and B Young. 2015. Web annotation protocol. (2015).

[20] Robert Sanderson, Paolo Ciccarese, and Benjamin Young. 2017. *Web annotation data model.* Technical Report. Technical report, W3C Recommendation, 23 February.

[21] Andie; Jon Wild; Eliot Handelman Sigler. 2015. Schematizing the Treatment of Dissonance in 16th-Century Counterpoint. In *Proc. Intl. Conf. on Music Information Retrieval (ISMIR).* 645–651.

[22] Daniel Sleator and Davy Temperley. 2001-2003. Melisma Music Analyzer. http://www.link.cs.cmu.edu/music-analysis/. (2001-2003).

[23] Victoria Uren, Philipp Cimiano, José Iria, Siegfried Handschuh, Maria Vargas-Vera, Enrico Motta, and Fabio Ciravegna. 2006. Semantic annotation for knowledge management: Requirements and a survey of the state of the art. *Web Semantics: science, services and agents on the World Wide Web* 4, 1 (2006), 14–28.

[24] Daya C. Wimalasuriya and Dejing Dou. 2010. Ontology-based information extraction: An introduction and a survey of current approaches. *J. Information Science* 36, 3 (2010), 306–323.

[3] http://neuma.huma-num.fr/quality/