Reid Swanson
ISE 575

Logical Representation of Musical Concepts

The goal of Somnuk Phon-Amnuaisuk's paper is to describe a suitable form of knowledge representation of music that can be used for complex and abstract musical reasoning tasks. The premise of the paper is that any such representation must be expressive enough to allow reasoning about the desired entities, while abstracting enough information to make the engineering, computation and reasoning efficiency manageable. The approach taken in this paper is a rule based expert system because it allows for a compact set of rules that have large breadth, heuristics are easily added to cover unseen cases and the results returned can be justified by looking at the rules used to produce them.

There are three types of common representations for musical knowledge. The first are notational typesetters that are geared to helping users typeset musical symbols. The second are sequencers which are concerned with the storage and playback of musical events. Synthesizers are the other common representation that mostly deal with the generation of sound. Inferring representations that reason about musical concepts are lacking in the community. Despite the value and deep analysis this type of representation could provide it is lacking from most available musical software packages and is a primary motivation for the work of this paper. For example the paper talks about two types of typical analysis in Musicology. The “narrative and descriptive” and the “analyze extra-musical content”. The first is essentially a process of breaking the piece down into component parts and describing how the parts go together, which is a good candidate task for what a computer should be able to do.

A central theme of this paper is the need to balance between abstraction and expressiveness. To achieve a high level of expressiveness and inferential competence it is also necessary to provide several different representations of the same knowledge. This is because in general no one representation captures all the aspects of the concept desired but is used because of its efficiency or convenience in a particular type of reasoning environment. For example MIDI is an excellent representation for the performance aspect of musical events, but are not able to capture other useful information such as the difference between C4# and D4b. In this paper both declarative and procedural representations are allowed but it is noted that procedural representations are preferred because they can encode the same information as declarative but can contain functions that can be used in the control structure to further constrain the meaning and allow for heuristics.

The paper continues with a description of the logical language used and how traditional musical scores are encoded in the system. Knowledge in the system is encoded as a set of sentences composed of atomic terms conjoined by standard logical operators. A concept is defined as a collection of attribute-value pairs where the attributes and values are predicate terms as defined in the system. These concepts are then represented in the system as an attribute value matrix similar to those used in computational linguistics. However, here values are also allowed to be programmatic objects as well. A score is procedurally defined as composed of two parts; musical materials and interpretations. Musical materials are a collection of concepts. The primary element is that of a musical line and also includes pitch, note events, line and interval. Interpretations are also a collection of concepts just like the musical materials object, however these are reserved for concepts that are added, removed and modified during the inference process. Concepts can be created and modified using some basic rules and ensuring that these new and modified concepts remain well formed (through the checking of a priori constraints defined in advance) and are compatible (when two groupable concepts share an attribute their values must be compatible). Comparison and similarity can be derived from the graphical representation of the AVM, although the weights for each arch in comparison are, in general, not equal and so each concept should define its own similarity metric.

The final system was able to produce harmonizations given an input score as they show for one of Bach's melodies. The final result of their work is still a bit mysterious after reading the paper. They talk a lot about different representations and how they defined concepts as logical predicates. However, they seem to abandon logical predicates in favor of procedural representations and graphical representations. All this would fit well with the thrust of their paper, that they want multiple representations, however despite talking about the need to map between representations they don't discuss how they did it. Nor do they talk about how they actually did any inference. The major themes of the paper were very sound but I didn't feel that the paper went into enough detail to know what kind of representations work, what kind of inferences the system could achieve or how it actually was able to produce the harmonization it did.