Can we calculate music like numbers?

Keiji Hirata
NTT Communication Science Laboratories
hirata@brl.ntt.co.jp

The author dares to answer this question with “yes.”

Although it is not possible to treat music in exactly the same way as we do numbers, it will be possible in such a way that we express our thought by writing mathematical formulae. The author would like to present the framework toward the goal.

1 Background

At present, music is considered an art, and composition of music something that only a small number of skillful and knowledgeable people can engaged in. One major epochal change was when Thomas Edison invented the well-known tin foil phonograph almost 130 years ago; the phonograph drastically changed the way people related to music, and the world now familiar to us emerged. Along the way, the modern listening style, music industry and technology have also been established. However, music was originally a medium everyone could enjoy creating before the phonograph; ordinary people could readily hear others’ performances almost anywhere, anytime. The author believes that a computer has the potential to make music a more expressive medium than it was before the phonograph.

2 Formalization Problem

As we notice, the computer is rigid in the sense that unless everything is described precisely, a computer does not work as a programmer intends. But, it is quite difficult to formalize (mathematically represent and rewrite) the perceptual meaning of music on a computer. Formalization looks like a symbolization in Semiotics, where symbolization is understood as the interaction between object, sign, and interpretant (Fig. 1). Object stands for a physical thing in the real world, sign is a corresponding representation in the symbol world, and interpretant means semantics or images of the object that come up into consciousness and mind. It is important that correspondence between symbol manipulation (calculation by computer) and object transformation (events occurring in the real world) is always kept consistent. The function of semantics is often called grounding.

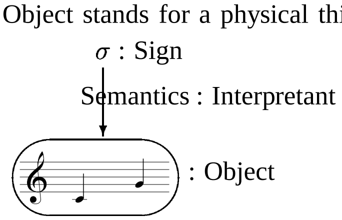

Figure 1: Semiotics

Here the author asserts that the successful correspondence should satisfy the following two conditions. (1) If we hear two melodies a and b as the same, then $a = b$ holds in the symbol world; otherwise $a \neq b$. (2) If we hear melody b similar to and simpler than melody a, symbol of melody a is rewritten to that of melody b by a certain rational rules.

The difficulty to satisfy the conditions is mainly because of music’s inherent features: aesthetics, subjectivity, tacitness, ambiguity, individuality, emotion, and so on. These features make music itself unique and expressive, while they become obstacles to consistent and versatile modeling of the rational rewriting rules in computer programs. After all, many practical music systems often solve the formalization problem by limiting genre and styles and/or introducing heuristics.

3 Music Theory and Reduction

Music is an advantageous medium, since we have music theory. The goal of music theory is to analyze music on a score and to understand its underlying structures. Musical structures here mean tacit, deep information, such as relationships among notes and groups made of relevant notes, in contrast to the visible, superficial information of a score.

Music theory gives the perceptual ground that a different musical structure produces a different kind of feeling (please remember the successful correspondence between the symbol and real worlds), and that feeling includes punctuation, termination, progression, floating, tonality, and so on. These feelings are relatively low-level ones, hence the author uses the word “perception” here, instead. It is known that such low-level feelings are relatively common among people, compared to high-level ones, such as emotion and preference. Music theory proposes several analysis methods for identifying musical structures, which actually imply grouping based on the Gestalt psychology and prominent notes to govern the listener's perception.

In 1983, F. Lerdahl and R. Jackendoff proposed a music theory in their book *A Generative Theory of Tonal Music*; their music theory, called GTTM, is one of a few music theories that involve the concept of reduction. In general, reduction is rewriting an expression to an equivalent, simpler one; it often has the same meaning as abstraction or simplification. The GTTM reduction is designed based on the Gestalt grouping, and the reduction successfully associates a melody with another one sounding similarly. Fig. 2 shows an excerpt from the GTTM book and demonstartes the reduction concept. The best way to read Fig. 2 is to hear the successive levels in a same tempo. If the reduction is done satisfactory, each level should sound like a natural simplification of the previous level. Alternative omissions of notes must make the successive levels sound less like the original.

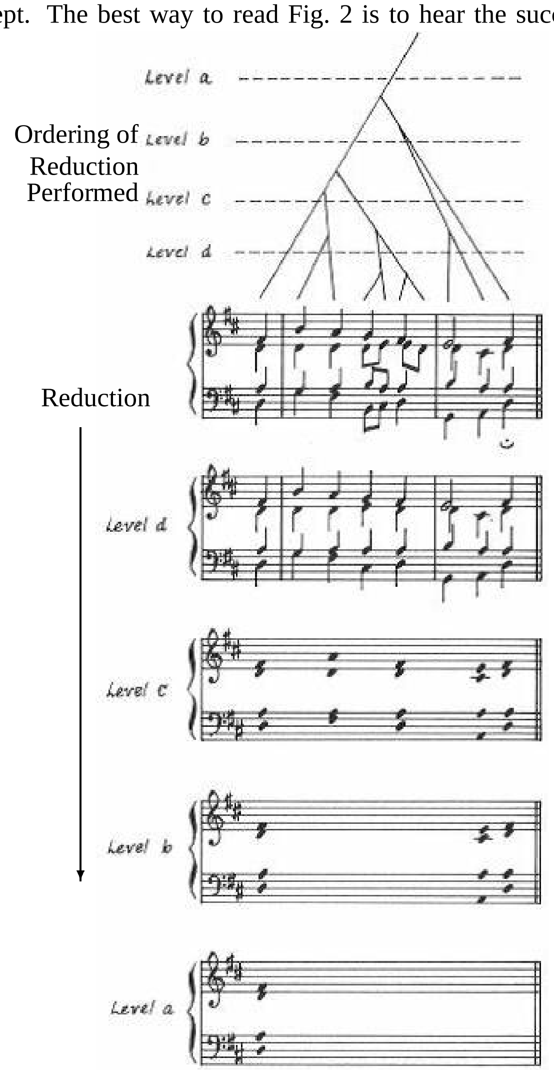

Figure 2: Reduction Concept

The key idea of the framework is that the GTTM reduction is adopted as the rational rewriting rules. Although GTTM certainly provides the set of rules for determining the correct group boundaries and its hierarchy in a melody, the rule definitions are unfortunately too ambiguous and incomplete to be implemented as computer programs, because music theory has been developed and presented only to humans, not to computers. This is a specification problem.

4 Music Algebra

We now have to discuss the issue of term representation, but the author would like to omit it due to space limitation. Let us move to an algebraic system for calculating music based on the reduction concept.

A new notation $a \subseteq b$ denoting "a is reduced to b" is introduced[1]. By $\subseteq$, we can place melodies one by one in the ordering according to the GTTM reduction. The ordering is not a total order, examples of which include the integer and real numbers, but rather a partial order. Mathematically speaking, the domain of melodies formalized in the framework makes a complete lattice with respect to the $\subseteq$ relation. A complete lattice is a set in which a partial order is defined and all infinite subsets can have an intersection (least upper bound, lub) and a union (greatest lower bound, glb).

Since relation $\subseteq$ is the only operator defined on the domain, we have to define the other operators using $\subseteq$. For example, $a = b$ is defined as $a \subseteq b \wedge b \subseteq a$. Basic operators lub and glb are also defind using $\subseteq$; $lub(A, B)$ extracts the largest common part or the most common information of A and B in the bottom-up mannter, and $glb(A, B)$ joins A and B in the top-down manner.

5 Present and Future

Apparently, the computation capability of the music algebra is more or less different from that of the ordinary arithmetic. To investigate and elaborate the music algebra, the author has been applying the music algebra technology to implementing several music tasks, such as reharmonization, music retrieval, arrangement, performance rendering, music summarization, and, at present, music social ware (e.g. online collective composition systems). Through the developments, the author is convinced that the music algebra becomes a basis for both practical and theoretical foundations of an expressive music medium for everyone.

Finally, future issues contain enhancement of the music algebra itself and application of the framework to other media, such as vision, sound and natural language.

[1]The direction of this symbol is possibly opposite the reader's intuition. For the rationale of the direction, consider "C5 note is reduced to C note." What C note implies includes what C5 note, because C note may imply, for instance, C4 and C6 as well as C5. Thus we have C5 $\subseteq$ C.