Logical Representation of Musical Information

Tetsuya Mizutani
The University of Tsukuba
1-1-1 Tennodai Tsukuba City,
Ibaraki 305-8573 Japan
mizutani@ailab.is.tsukuba.ac.jp

Abstract—Logical representation of musical information is introduced. Especially, expression and fingering of piano performance based on musical structures and structural functions of music are given as logical rules written in the first- or higher-order predicate logic. The representation will help us to study musical informatics.

Keywords—Musical Informatics, Logical Representation, Agogic Rules, Fingering Rules

I. INTRODUCTION

THE musical informatics is a field of study for investigating and analyzing higher level fusion of intellect and sensibility of human via music. One of the main aims of this study is to understand how to play artistic performance of music by computers, hence, logical or mathematical representation of musical artistic expression is important. In this paper, some logical representation of musical information such as expression and fingering rules for performances of western classical piano pieces based on musical structures and structural functions of music [2] is presented. The representation will help us to study musical informatics, i.e. to design, implement, verify and analyze musical information systems.

II. MUSICAL STRUCTURES

For analyzing music scores, the theory of musical structures by Riemann [7] is adopted. From the phrasing analysis by the theory, the smallest unit of a score representing musical motivation is called a motif that is usually composed of two measures. A longer unit usually composed of two motives is called a phrase and a longer one of two phrases a sentence. Each motif, phrase or sentence is called a segment for fingering “at a stretch”, e.g. a semimotif, phrase or sentence is called a part and a longer one of two phrases a sentence. Each structure σ_{i} is again composed of a lower level structures

\[ \sigma_{i} = (\sigma_{i1}, \sigma_{i2}, \ldots, \sigma_{im}), \]

or, composed of notes \( v_{ij} \)

\[ \sigma_{i} = (v_{i1}, v_{i2}, \ldots, v_{in}) \]

if \( \sigma_{i} \) is the local lowest structure. Each note \( v_{ij} \) consists of pitch, volume, etc. as follows:

\[ v_{ij} = (num_{ij}, ptc_{ij}, vol_{ij}, lng_{ij}, itv_{ij}, prt_{ij}, msr_{ij}, nvl_{ij}, fgm_{ij}), \]

where

- \( num_{ij} \in \mathcal{N} \) is the index of the note, which is the serial number in the score,
- \( ptc_{ij} \in \mathcal{KN} + \{ \text{rest} \} \) for \( \mathcal{KN} = \{ x | 21 \leq x \leq 108, x \in \mathcal{N} \} \) and \( vol_{ij} \in \mathcal{N} \) are the pitch and the volume of the note, respectively, each of which is represented by MIDI standard, and each key number is usually expressed by a constant called the pitch name, i.e. \( C_{4} (=60), C_{5} (=61), E_{4} (=64), A_{4} (=69) \), etc.,
- \( lng_{ij} \in \mathcal{N} \) is the length of time between its note-on time, i.e. the time when the sound of the note starts, and note-off time, i.e. that when it stops,
- \( itv_{ij} \in \mathcal{N} \) is the length between its note-on time and that of the previous note,
- \( prt_{ij} \) and \( msr_{ij} \in \mathcal{N} \) are the part and the measure which the note belongs, respectively,
- \( nvl_{ij} \in \mathcal{Q}^+ \) is the note value represented by a fraction, e.g. \( \frac{4}{5}, \frac{9}{8} \), etc., and
- \( fgm_{ij} \in \mathcal{FN} \) is a finger number.

\( D_{\text{NOTE}} \) denotes the domain of notes, defined by

\[ D_{\text{NOTE}} \subset \mathcal{N} \times (\mathcal{KN} + \{ \text{rest} \}) \times \mathcal{N}^{5} \times \mathcal{Q}^+ \times \mathcal{FN}. \]

Definition 1: \( \mathcal{FN} = \{ 1, \ldots, 5 \} \) is the set of finger number of the right hand, where 1, \ldots, 5 indicate a thumb, \ldots, a little finger, respectively.

Definition 2: Let \( \mathcal{S} \) be a score and \( F_{\mathcal{S}} \) be a sequence of fingers of the score \( \mathcal{S} \). \( \mathcal{S}[F_{\mathcal{S}}] \) is obtained from \( \mathcal{S} \) by assigning \( F_{\mathcal{S}} \) to it.

\( \mathcal{S}[F_{\mathcal{S}}] \) is composed of a finite number of musical structures

\[ \mathcal{S}[F_{\mathcal{S}}] = (\sigma_{1}, \sigma_{2}, \ldots, \sigma_{m}). \]
B. Keyboard Information

Each element in a note $v_{ij}$ is represented by a superscript, i.e. $v_{ij}^{num} = v_{ij}^{num}_k$, $v_{ij}^{ptc} = v_{ij}^{ptc}_k$, etc. Additionally, $S(k)$ indicates the $k$th note of the score $S$, i.e. $S(k) = v_{ij}$ if and only if $v_{ij}^{num} = k$.

III. Functions and Predicates

Some functions and predicates useful to represent agogics [9] and fingering [1] are defined.

A. Agogics Information

Definition 3: A function $tpb : D_{NOTE} \rightarrow \mathbb{Q}^+$ denoting time per beat of a note is defined by

$$tpb(v) = \frac{x_{stv}^{num}}{x_{num}^{num}},$$

and $tpbs$ denotes a sequence of times per beat of a sequence of notes, by

$$tpbs(\tau) = (tpb(v_1), \ldots, tpb(v_m)),$$

where $\tau = (v_1, \ldots, v_m)$.

Definition 4: For a finite subset $X$ of $\mathbb{R}^2$ having the uniqueness condition of functions

$$\forall xyz \ (x, y) \in X \land (x, z) \in X \implies y = z),$$

$aprx(X) : \mathbb{R} \rightarrow \mathbb{R}$ is defined as a function satisfying that a map represented by

$$\{(x, aprx(X)(x)) | (x, y) \in X \}$$

is an approximation of $X$.

Definition 5: A function $loc : (D_{NOTE})^n \times \mathbb{N} \rightarrow \mathbb{Q}^+$ denoting the location, as note values in a score, of the $i$th note in a sequence of notes of length $n$ ($1 \leq i \leq n)$ is defined by

$$loc(\tau, i) = \begin{cases} 0 & i = 1, \\ \sum_{k=1}^{i-1} (nth(\tau, k))^{rel} & i \geq 2, \end{cases}$$

where $nth(\tau, i)$ is the $i$th element of $\tau$.

Definition 6: A function $agc$ denoting the set of pairs of the location of a note and its time per beat, i.e. the agogic of notes, is defined by

$$agc(\tau) = \{(loc(\tau, i), tpb \circ nth(\tau, i)) | 1 \leq i \leq |\tau|\},$$

where $|\tau|$ is the length of $\tau$ and $\circ$ is the composition operator of functions.

B. Keyboard Information

Definition 7: A function $kd : KN \rightarrow \mathbb{N}$ denoting the distance between two keys $x, y \in KN$ on a keyboard is defined by

$$kd : KN \rightarrow \{1, 2\},$$

$$kd^*(x) = \begin{cases} 2 & \text{if } x \mod 12 \in \{4, 12\}, \\ 1 & \text{otherwise}, \end{cases}$$

$$kd(x, y) = \sum_{i=\min(x, y)}^{\max(x, y)-1} kd^*(i).$$

A function $skd : KN \rightarrow \mathbb{N}$ for signed key distance is defined by

$$skd(x, y) = \begin{cases} kd(x, y) & \text{if } x \leq y, \\ -kd(x, y) & \text{otherwise}. \end{cases}$$

For example, $kd(C_4, C_4^\sharp) = skd(C_4, C_4^\sharp) = 1$, $kd(C_4, E_5) = skd(E_5, C_4) = 18$, but $skd(E_5, C_4) = -18$.

The notion of the spans that two fingers can stretch within some limitations is useful for the investigation of the fingering rules. Some functions for spans between two fingers [1], [6] are introduced below.

Definition 8: Functions $maxpr, minpr : FN \rightarrow \mathbb{Z}$ are defined as the maximum and the minimum practical stretch span, respectively, between two fingers to play two notes simultaneously or consecutively in legato, defined by the signed key distance between these notes.

$maxrel, minrel : FN \rightarrow \mathbb{Z}$ are regarded as the maximum and the minimum relaxed spans, respectively, that two fingers fall without tension onto the keys.

$maxpr(x, y)$ represents the span between $x$ and $y$, where $x$ is the finger number of the first note and $y$ that of the second. If $C_4$ can be played by a thumb and $E_5$ by a little finger simultaneously with the practical limitation, then $maxpr(1, 5) = skd(C_4, E_5) = 18$. If $x = y$ then $maxpr(x, y)$ is 0, else if $x < y$ then it is positive, otherwise it is equal to $-maxpr(y, x)$.

Other functions are similar. And moreover, if the first finger is a thumb, then $minpr$ etc. are negative since the next finger can be passed over the thumb.

Definition 9: A predicate $Adj : D_{NOTE} \times D_{NOTE}$ denoting that two notes are adjacent in the score, but some rests may occur between them, is defined by

$$Adj(v_1, v_2) \equiv v_1^{ptc} \neq rest \land v_2^{ptc} \neq rest \land v_1^{num} < v_2^{num} \land \forall i(v_1^{num} < i < v_2^{num} \implies S(i)^{ptc} \neq rest).$$

Definition 10: A predicate $Reach : D_{NOTE} \times D_{NOTE}$ denoting that two keys can be played simultaneously for two fingers with a practical span, i.e. the fingers are reachable for the keys, is defined by

$$Reach(v_1, v_2) =$$

$$\minpr(v_1^{num}, v_2^{num}) \leq skd(v_1^{num}, v_2^{num}) \leq maxpr(v_1^{num}, v_2^{num}) \land v_1^{ptc} \neq rest \land v_2^{ptc} \neq rest.$$
IV. Structural Functions and Agogic Rules

A. Structural Functions of Music

Each note in a musical structure could have a different meaning and be played by its own rôle for each different occurrence. Structural functions of music [2] represent a certain emotion to listen or play the music reflecting the meaning and the rôle. In each motif or phrase, there are 6 kinds of functions as follows:

1. inceptive: the starting point of the structure,
2. anacrusis (anticipative): a range where the performance could be getting tenser to the initiative,
3. tiara: a couple of notes just before the initiative in which area is a climax of the tension,
4. initiative: an ideal climax in the structure which must be emphasized,
5. desinence (reactive): a range where tension is diminished following the initiative, and
6. conclusive: the end point of the structure.

Noted that a tiara may not exist in a motif or a phrase.

A structure $\sigma$ can be divided into 6 parts:

$$\sigma = \text{incp}(\sigma) \cdot \text{anac}(\sigma) \cdot \text{tiar}(\sigma) \cdot \text{init}(\sigma) \cdot \text{desn}(\sigma) \cdot \text{cncl}(\sigma),$$

where $\text{incp}$ etc. are functions expressing the inceptive etc. of $\sigma$, while $\cdot$ is the concatenation operator of structures. We assume that the length of $\text{init}(\sigma)$ is 1, that of $\text{tiar}(\sigma)$ is between 0 and 2, and those of other elements are greater than 0. $\text{mid}(\sigma)$ abbreviates $\text{anac}(\sigma) \cdot \text{tiar}(\sigma) \cdot \text{init}(\sigma) \cdot \text{desn}(\sigma)$.

Figure 1 shows structural functions of the first motif of Chopin’s Polonaise Militaire Op. 40, No. 1.

B. Agogic Rules

The logical representation of some primitive agogic rules [9] based on structural functions are as follows:

1. the performance of the inceptive starts slowly:

$$\text{ave} \circ \text{tpbs} \circ \text{mid}(\sigma) < \text{ave} \circ \text{tpbs} \circ \text{incp}(\sigma),$$

2. the performance accelerates in the anacrusis, while it decelerates in the desinence:

$$\text{Dcr}(\text{aprx} \circ \text{agc} \circ \text{anac}(\sigma))$$
$$\wedge$$
$$\text{Inc}(\text{aprx} \circ \text{agc} \circ \text{desn}(\sigma)),$$

3. the tiara is played longer than both the initiative and the last note of the anacrusis:

$$\text{tpb} \circ \text{nth}(\text{anac}(\sigma), |\text{anac}(\sigma)|) < \text{ave} \circ \text{tpbs} \circ \text{tiar}(\sigma),$$

4. the initiative is played longer than the “neighbor” notes other than the tiara:

$$\text{tpb} \circ \text{nth}(\text{anac}(\sigma), |\text{anac}(\sigma)|) < \text{tpb} \circ \text{init}(\sigma)$$
$$\wedge$$
$$\text{tpb} \circ \text{nth}(\text{desn}(\sigma), 1) < \text{tpb} \circ \text{init}(\sigma),$$

and

5. it ends slowly in the conclusive:

$$\text{ave} \circ \text{tpbs} \circ \text{mid}(\sigma) < \text{ave} \circ \text{tpbs} \circ \text{cncl}(\sigma),$$

where $\text{ave}(a) = (\sum_{k=1}^{m} a_k)/m$ for $a = (a_1, a_2, \ldots, a_m), a_i \in \mathbb{R}$, and $\text{Inc}(f)$ and $\text{Dcr}(f)$ are functions that a function $f : R \rightarrow R$ is increasing and decreasing, respectively.

V. Fingering Rules based on Structure

Some fingering rules [1] are given with logical representation as follows:

1. rule $S_{\text{I}}$: for every adjacent pair of notes in a group $\sigma$, a reachable pair of fingers must be used if they are performed legato:

$$S_{\text{I}}(v_1, v_2) \equiv \text{Adj}(v_1, v_2) \wedge \text{Lgt}(v_1, v_2) \supset \text{Reach}(v_1, v_2),$$

$$S_{\text{I}}(\sigma) \equiv \forall i(1 \leq i \leq |\sigma| - 1 \supset S_{\text{I}}(\sigma(i), \sigma(i + 1))),$$

where $\text{Lgt} : D_{\text{NOTE}} \times D_{\text{NOTE}}$ is a predicate representing that a performance plan for two adjacent notes is legato,

2. rule $S_{\text{G}}$: for each pair of adjacent groups $\sigma_1$ and $\sigma_2$, the reachable fingers must be used on the pair of notes of the end of the preceding group and the beginning of the following one if they are performed legato:

$$S_{\text{G}}(\sigma_1, \sigma_2) \equiv S_{\text{I}}(|\sigma_1|), \sigma_2(1)),$$

3. rule $S_{\text{t}}$: a pair of fingers for every adjacent pair of notes must not cross in a group $\sigma$ if it is performed smoothly:

$$S_{\text{t}}(\sigma) \equiv \text{Phrs}(\sigma) \supset$$
$$\forall i(1 \leq i \leq |\sigma| - 1 \wedge \text{Adj}(\sigma(i), \sigma(i + 1)) \supset$$
$$\neg \text{cross}(\sigma(i), \sigma(i + 1)),$$
where Phrs : $D_{NOTE}$ is a predicate representing that a performance plan for a group is smooth, i.e. emphasizes its phrasing, and

4. rule $St_3$ : two sequences of fingers for two groups $\sigma_1$ and $\sigma_2$ must be the same if two scores and two performance plans are both similar:

$$St_3(\sigma_1, \sigma_2) \equiv \lvert \sigma_1 \rvert = \lvert \sigma_2 \rvert \land Sim(\sigma_1, \sigma_2) \land S_{\text{max}}(\sigma_1, \sigma_2),$$

where $S_{\text{max}} : D_{NOTE} \times D_{NOTE}$ is a predicate representing that two performance plans for two groups are similar, while $Sim : D_{NOTE} \times D_{NOTE}$ is a predicate denoting that two sequences of pitches are similar, defined by whether a certain distance between two sequences is less than a threshold or not.

Figure 2 is a score of the first 2 measures of Burgmüller's Innocence with groups and fingering rules to be applied to it. By these rules, we have some candidates of its fingering sequences including the correct answer

$$\begin{align*}
G_1 & : (4, 3, 2, 1), \\
G_2 & : (4, 3, 2, 1, 4, 3, 2, 1), \\
G_3 & : (4, 3, 2, 1), \\
G_4 & : (4, 3, 2).
\end{align*}$$

Adding the notions of penalty point, difficulty of fingering and weight of rules [1], [6], the correct answer will be obtained. The penalty point is assigned to a fingering sequence if it violates a rules. Each rule has its own value for the penalty point. The difficulty means that if a fingering is more difficult technically than the other fingering, e.g. a larger span of fingers is understood as more difficult than a smaller one, then a larger penalty point is assigned to the former than the latter. The weight of rules is the priority over rules. Especially, the similarity rule $St_1G$ is understood as the most important, thus if a fingering violates this rule then the larger penalty point is assigned. To obtain the correct answer, the sequence of fingers which total penalty points is the minimum is sought.

VI. Conclusions

This paper is our first step that we attempt to represent artistic expression of music in a formal logical notation. Since logical representation is a higher level abstraction of information, it must be quantified when one want to perform music along with the logical rules. The quantification is one of interesting and profound fields of informatics, especially AI.

We have already designed, implemented, verified and analyzed an automatic ensemble program [5], [8] by $\oplus$-calculus [3] which is a smallest formal system applicable to verification and analysis of cooperative real-time systems on natural number time. The automatic ensemble program is an example of higher level intellectual and realtime control systems, but the object program for the analysis did not include any specific expression of music. Therefore, we will design and analyze an ensemble program with artistic.

Another important and interesting point of our further study is investigating North African music for the logical representation, and moreover, for the music informatics.

Acknowledgment

The author wish to express his gratitude to Ms. Michiybo Matsuzaki (Pf.) who gave many useful advice for music.

References


