Representation of Jazz Piano Knowledge Using a Deductive Object-Oriented Approach

Keiji Hirata

NTT Basic Research Laboratories
hirata@nefertiti.brl.ntt.jp
http://www.brl.ntt.jp/people/hirata

This paper presents a method for representing jazz piano knowledge that is based on the deductive object-oriented (DOO) approach. One of the motivations behind this research is the desire to clarify which elements of musical knowledge can be handled by our current knowledge processing (KP) techniques and which problems should be challenged in order to develop more practical KP techniques. The paper demonstrates how naturally and rationally the DOO framework can capture musical entities, concepts and relations. Also reported are the preliminary results of an experimental system.

1 Introduction

A fundamental problem in designing an intelligent musical system is the construction of a framework that can properly represent and manipulate musical knowledge. Since music is one manifestation of human knowledge activity and is not a special target of knowledge processing (KP), it is supposed that general-purpose KP schemes that have been investigated to date are applicable to building intelligent musical systems.

This research[3][4] aims at formalizing some part of musical knowledge, using computer science and artificial intelligence terminologies as much as possible. It will lead to clarifying which elements of musical knowledge can be handled by our current KP techniques and which problems need to be challenged in order to develop more practical KP techniques. The author thinks that a deductive object-oriented (DOO) approach[6][8] is noteworthy among KP techniques, since this approach provides desirable properties: expandability, flexibility, robustness and a theoretical basis.

Through examining musical knowledge representation based on the DOO framework, we hope to develop a programming language for musical KP. To facilitate design of the language, this paper discusses by what primitives and operations musical knowledge should be represented. Next, an experimental system is built to show the advantages of this approach. Last, the perspectives and problems are described.

2 Representing Musical Knowledge

Musical concepts and entities are linked to each other by various relations in our minds. This network can be viewed as a knowledge structure, which has evolved through knowledge activities such as learning and practice. The DOO framework,in comparison, provides universal methods for representing real-world entities, in which the key notions are object, relation and deductive rule. This section exhibits how naturally and rationally the DOO framework can represent musical entities, concepts and relations.

This research focuses not on the physical aspects of music, but rather on musical KP based on symbolic manipulation; thus, we assume that a note is atomic. The following examples employ jazz piano knowledge on the presumption that this style may be popular and familiar to a large segment of the readership.

2.1 Object

In our DOO framework, every real musical entity and concept is represented as an object; e.g., the note C, the note C5, major scale, a set of notes {G,B,D,F} (which could be called a chord), a chord symbol sequence, such as $D_{m7} \rightarrow G_7 \rightarrow C_{maj7}$, and a polyphonic melody are all considered to be individual objects. Our framework introduces a notation for representing an object: $c(l_1{=}t_1, \cdots, l_n{=}t_n)$, where c stands for a class name or a basic object, l_i an attribute name, and t_i an attribute value. This notation can be regarded as an extension of a predicate $c(v_1, \cdots, v_n)$. For example, the note C5, (the pitch is C, and the octave position is 5th) is represented as note(pitch=C, octave=5), and the chord symbol G_7 by chord_symbol(root=G,name=7). This term representation is used as an object identifier; that is, every musical entity and concept is uniquely identified by the corresponding term representation. Of course, it is possible that C5 itself can serve as a basic object; such a decision depends on the situation and the person describing the knowledge.

Another important and useful data structure in KP is the set. A set is used to represent a collection to which objects having similar properties belong. For example, a chord is usually defined as three or more tones sounded simultaneously, and it is natural that a chord is viewed as a set of note objects; the chord of C, E, G is represented by {C,E,G}. Note that our framework treats a set as an object.

2.2 Subsumption Relation

Example 2.1: Let us consider the representations of notes. A musician may have various images in terms of a note: e.g., a real note C5 (whose MIDI note number is 60), the pitch C, and the notion of a note (a note class) (Fig. 1). It is possible to say

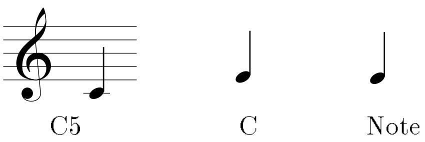


Figure 1: Various notions of a note

that C5 is C and C is a note. In the figure, from left to right, abstraction goes, and, inversely, concretion does. In other words, the abstraction corresponds to climbing up for superclass, and the concretion for subclass. This relation is organized in the so-called *is_a* hierarchy or class hierarchy.

Example 2.2: Minor scales appear in three forms: natural minor, melodic minor, and harmonic minor. For example, when we speak of a natural minor scale, that merely means a proper series of whole and half steps, and not a series of note symbols. By providing a key tone to the natural minor scale, a series of note symbols is obtained. (Fig. 2). Here,

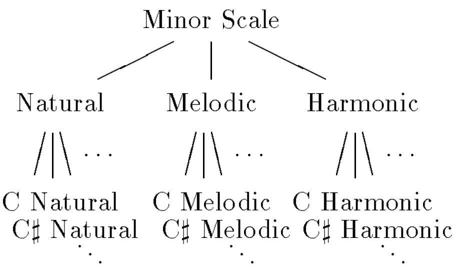


Figure 2: Hierarchy of minor scales

the hierarchy of the minor scales is represented by the subclass relation or class membership.

Example 2.3: Suppose that there is a chord of C, E, G, B♭. Since this chord includes a tritone (the interval of the augmented 4th between E and B♭), it sounds like a dominant 7th. Actually, this

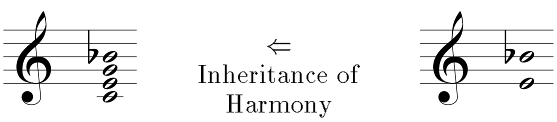


Figure 3: Inheritance of dominant 7th harmony

chord may be represented by the chord symbol C_7. Thus, we can say that the chord of C, E, G and B♭ *inherits* the dominant 7th property from the chord of E and B♭ (Fig. 3). In that sense, the chord of C, E, G and B♭ can be regarded as a subclass of the chord of E and B♭.

Notation: The relations that we have examined so far are: *is_a*, subclass, class membership, and inheritance. A common property of these relations is that they work between two objects that are quite similar to each other, but differs slightly; that is, the two objects belong to similar categories. Therefore, to represent these relations as a whole, our framework introduces a subsumption relation denoted by $\sqsubseteq$. Intuitively, "*instantiated* $\sqsubseteq$ *abstract*" is stated. The examples are as follows: note(pitch=C,octave = 5) $\sqsubseteq$ note(pitch=C) $\sqsubseteq$ note, C Natural Minor Scale $\sqsubseteq$ Natural Minor Scale $\sqsubseteq$ Minor Scale, and {C,E,G,B♭} $\sqsubseteq$ {E,B♭}. Note that $\sqsubseteq$ is a partial ordering.

In the above example, the so-called Smyth ordering [8] is used between chords. But the author thinks that, from the musical KP point of view, our framework poperly should use not only the Smyth ordering but also the Hoare ordering.

2.3 Extrinsic Attribute Relation

Unlike the subsumption relation, an extrinsic attribute relation connects two objects that belong to distinct categories.

Example 2.4: The statement a chord symbol of the chord of F and B is G_7 or $D\flat_7$ is represented by the expression

chord(notes={F,B}).chord_symbols={G_7,$D\flat_7$},

and is depicted in Fig. 4. In general, the expression

$$\boxed{\text{chord(notes=\{F,B\})}} \xrightarrow[\text{chord\_symbols}]{} \boxed{\{G_7, D\flat_7\}}$$

Figure 4: Extrinsic attribute relation

$p.l = q$ means that the value of the p's l attribute is the q object. Note that extrinsic attribute relations do not change its object identity.

Fig. 5 shows a more complicated example. For space efficiency, in the figure, n is an abbreviation of *note*, *p pitch*, *o octave*, *d degree*, and *cs chord_symbols*. Chord (A) is relativized with respect to its root note F3, and, as a result, (B) is obtained. While (A) is a set of notes represented by pitch and octave, (B) consists of the root note and a set of notes represented by degree. Since chords (A) and (B) are different, the extrinsic attribute relations, relative and absolute, link them to each other. The extrinsic attribute relations, chord_symbols and voicing, link (B) and (C).

An experimental system described later actually uses the extrinsic attribute relations for associating with quartal, tritone, scale, triad and so on.

In the context of Ref. [8], the attributes used for part of an object identifier (Section 2.1) are distinctively called intrinsic attributes, and only an object

(A)		(B)		(C)
$\{n(p{=}F,o{=}3)$, $n(p{=}E\flat,o{=}4)$, $n(p{=}A,o{=}4)$, $n(p{=}B,o{=}4)$, $n(p{=}D\flat,o{=}5)$, $n(p{=}G\flat,o{=}5)$, $n(p{=}B,o{=}5)\}$	relative $\longrightarrow$ $\longleftarrow$ absolute	$\{n(d{=}0)$, $n(d{=}10)$, $n(d{=}16)$, $n(d{=}18)$, $n(d{=}20)$, $n(d{=}25)$, $n(d{=}30)\}(root{=}n(p{=}F,o{=}3))$	chord symbols $\longrightarrow$ $\longleftarrow$ voicing	$\{cs(root{=}F$, $name{=}7th$, $tension{=}$ $\{\flat 9,\sharp 11,\flat 13\})\}$

Figure 5: Extrinsic attribute relations between chords and chord symbols

occuring at the position of an extrinsic attribute is inherited.

2.4 Deductive Rule

Roughly speaking, a deductive rule is similar to Prolog clauses, replacing predicates with object terms. In general, a user codes his own musical knowledge by means of deductive rules that play the role of methods. Moreover, in our framework, deductive rules define the subsumption relation, the orderings for sets such as Hoare and Smyth, and inheritance. Thus, the entire behavior of the system is easy to understand and control.

2.5 Temporal Objects

Assume that temporal structures are organized as a tree, where the root is the starting point. Thus, our framework represents the temporal structure with nested sets having intrinsic attributes that stand for its displacement in time from the event preceding it.

Example 2.5: Suppose X=$\{e_1,e_2\}$(disp=d_x) and Y=$\{e_3,e_4\}$(disp=d_y). Please bear in mind here that a set can be a basic object. Fig. 6 shows the temporal structure of the object term {X,{Y}(disp=d_2)} (disp=d_1). Note that time displacement can be zero

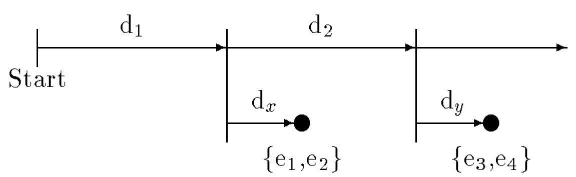


Figure 6: Temporal events represented by a tree

or negative, where the root point means the termination. Moreover, let us add a deductive rule with respect to the subsumption relation of sets: {S} $\sqsubseteq$ S if S is a set.

The abstraction of musical temporal objects consists of three kinds of operations, removing, relativizing and grouping, and can be captured by the representation introduced above. For abstraction, notes that are not prominent can be removed. This operation corresponds to removing an object that is an element of a set. When timing information is abstracted, only the relative position (or temporal ordering) of each note remains. This operation corresponds to dropping an intrinsic attribute for its time displacement. The operation which makes the notes having different onset times a group of simultaneous notes can be also regarded as abstraction. This operation is achieved by the additional deductive rule, {S} $\sqsubseteq$ S.

Example 2.6: The Figure 7 demonstrates the abstraction of a sample melody. In the figure, the abstraction from (A) to (B) uses removing; that from (B) to (C) relativizing and grouping; that from (C) to (D) grouping.

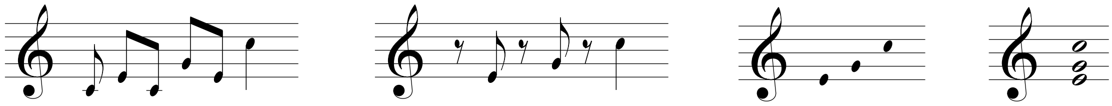
Example 2.7: Let us consider the chord progression $C\sharp_{m7}\rightarrow F\sharp_7\rightarrow C_{m7} \rightarrow F_7 \rightarrow B\flat_{maj7}$. This notation merely indicates a sequence between chord symbols. It is natural that the temporal structure is a tree, the root of which is $B\flat_{maj7}$; the time displacement is negative. Thus, we get $\{B\flat_{maj7},\{F_7,\{C_{m7},\{F\sharp_7,\{C\sharp_{m7}\}\}\}\}\}$, and let this term be P. Next, suppose that the chord progression P is derived by instantiating a chord progression, either $C_{m7}\rightarrow F_7\rightarrow B\flat_{maj7}$ or $F\sharp_7\rightarrow F_7\rightarrow B\flat_{maj7}$. That is, the former chord progression is more instantiated than the latter two. Our framework captures this situation as follows: the formulas P $\sqsubseteq$ $\{B\flat_{maj7},\{F_7,\{C_{m7}\}\}\}$ and P $\sqsubseteq$ $\{B\flat_{maj7},\{F_7,\{F\sharp_7\}\}\}$ hold.

3 Prototype System

The author has prototyped an experimental system. First, the data for building a database is given to the system. These data comprise a transcription of an actual solo performance by Herbie Hancock (one chorus of "Autumn Leaves") and comments on the performance. Then, a total of 374 objects have been stored in the system. The data in the system are of the form described in this paper. No background knowledge, e.g., the harmonic theory based on chord functionality, was given the system in advance.

Thus equipped, a user can issue queries to retrieve knowledge regarding the solo jazz piano. The queries are formed by using the subsumption relations and the extrinsic attribute relations. For example, when a query is "provide a voicing (possibly a sequence) of G_7 in a chord progression," the answer is a set of eight voicing sequences. Here,

(A) (B) (C) (D)



{C5,{X,{Y,{C6}(d=2)}(d=2)} (d=1)} $\sqsubseteq$ {{{E5},{{G5},{C6}(d=2)} (d=2)} (d=1)} $\sqsubseteq$ {E5,{G5,{C6}}} $\sqsubseteq$ {E5,G5,C6}

X $\equiv$ {E5,{C5}(d=1)}
Y $\equiv$ {G5,{E5}(d=1)}

Figure 7: Abstraction of a melody

the notion of musical similarity are represented using $\sqsubseteq$. When a query is "provide a realization of a common abstract voice leading of the given three chord-symbol sequences," the system first finds corresponding voicing sequences of the three and, then, calculates their least upper bound with respect to $\sqsubseteq$. This operation implements multiple inheritance along $\sqsubseteq$-ordering.

The system stores only the data for database construction, the size of which is not large. This experimental system does not have deductive rules predefined by a user, and there are only the built-in rules described in Section 2.4. Moreover, it does not provide a learning capability.

However, the answers generated by the system are of good quality. That is because musical knowledge is deliberately translated to the DOO terminology and, as a result, musical knowledge can be embedded into the representation itself naturally and rationally.

4 Related Work

Many attempts have been made toward representation of musical knowledge[2]. A main concern of the previous work was to invent new primitives and operations dedicated to music through observing musical phenomena and deliberating upon models behind the phenomena, where the models are not necessarily computational. Refs. [7] and [5] have attempted to capture a musical piece and jazz improvisation using a natural language (NL) framework, but the NL framework employed can handle only a small part of the knowledge activity and is not universal. As stated earlier, our work concerns the extent to which universal KP techniques can succeed in representing musical knowledge. Universal KP techniques are advantageous, since they have received much study as a computational system and have a theoretical basis. Ref. [1] employs the framework of a semantic network based on KL-ONE, which is a universal technique.

5 Concluding Remarks

Since musical knowledge is actually complicated, the author does not think that only the techniques described in this paper can handle the entire musical knowledge. However, in my experience with representing jazz piano knowledge, use of DOO is very advantageous in that it can represent various elements of musical knowledge naturally and rationally. It will lead to that the affinity of DOO with musical knowledge is quite high. We are further examining our framework by taking into account the other types of examples.

Future work is proper segmentation and clustering of input data for database construction, the efficient treatment of sets, and inductive learning.

It is to be hoped that new techniques developed in this research will be applied to other application fields of KP.

Acknowledgment: The author thanks Dr. Yasunori Harada of NTT, Mr. Masataka Goto of Waseda University, Prof. Kazumasa Yokota of Kyoto University, and Dr. Hiroshi Tsuda of Fujitsu Laboratories for their valuable discussion and comments. Also the author is grateful to Dr. Ken'ichiro Ishii of NTT for his continuing encouragement.

References

[1] A. Camurri et al., Music and Multimedia Knowledge Representation and Reasoning: the HARP System, *CMJ*, 19:2, 1995.

[2] R. Dannenberg, Music Representation Issues, Techniques, and Systems, *CMJ*, 17:3, 1993.

[3] K. Hirata, Prototyping A Jazz Piano Knowledge Base System With a Deductive Object-Oriented Approach, *In SIGMUS*, 95-MUS-11, Info. Processing Soc. of Japan, 1995.

[4] K. Hirata, Towards Formalizing Jazz Piano Knowledge with a Deductive Object-Oriented Approach, *In Proc. of IJCAI95 Workshop on AI and Music*, 1995.

[5] P. N. Johnson-Laird, Jazz Improvization: A Theory at the Computational Level, *In Representing Musical Structure*, P. Howell et al. (eds), Academic Press, 1991.

[6] M. Kifer, Deductive and Object Data Languages: A Quest for Integration, *In Proc. of DOOD'95*, 1995.

[7] F. Lerdahl and R. Jackendoff, *A Generative Theory of Tonal Music*, The MIT Press, 1983.

[8] K. Yokota and H. Yasukawa, Towards an Integrated Knowledge-Base Management System – Overview of R&D on Databases and Knowledge-Bases in the FGCS Project, *In Proc. of FGCS'92*, ICOT, 1992.