

## 演繹オブジェクト指向に基づく音楽記述に関する考察

平田 圭二

NTT 基礎研究所

本研究の目的は、楽曲や譜例の中で観察される音楽的な概念や現象を、意味が明確に定義されている計算機科学や数学の言葉で形式的に記述する枠組を与えることである。ここではその第一歩として、五線譜などの楽譜に対し非常に極端な抽象化及び捨象を施したもの(楽曲と呼ぶ)を表現対象とした。次にその抽象化された楽曲に対して、知識表現の分野で演繹オブジェクト指向と呼ばれる手法を適用した。和声学から簡単な例を幾つかとりあげ、和音の響きの複雑さ、和音の性質の継承、和音同士の関係、和音列同士の関係などの形式的な記述を行った。オブジェクトや演繹規則の表現力は高いので、本手法は対位法、編曲法などの音楽理論の記述や認知科学的なモデルの記述にも応用できると考えている。

## Study on Music Description Based on Deductive Object-Oriented Approach

Keiji Hirata

NTT Basic Research Laboratories

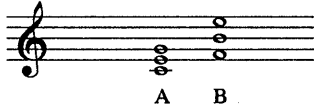
Various musical concepts and phenomena appear in scores, in general. An aim of this research is to provide the framework in which such concepts and phenomena can be formally described by using the terminology of computer science and mathematics. The meanings of the terminology should be definitely prescribed. In this paper, first, we drastically abstract a real score. Then, the method, which is called the deductive object-oriented approach in knowledge representation, is applied to the abstracted score. Some simple examples are picked up from the harmony theory. We describe in a formal way the complexity of chord sounds, the inheritance of chord properties, the relationship among chords and the relationship among chord sequences.

### 1 はじめに

現在まで、芸術性豊かな音楽を認識、理解するメカニズムの研究や、音楽理論をプログラムなどで表現し計算機で取り扱う研究が数多く行われてきた。そのためには、楽曲や音楽理論になんらかの抽象化を施したり、記号で表現したり、情報の一部を捨象したりする必要がある。これは、計算機で操作できる対象が記号的数学的なものに限られるからである。本稿では、現実世界の複雑な音楽的現象を計算機で操作できる程度に数学的なものに記号化、抽象化することを形式化と呼ぶ。実際、認知メカニズムの解明や音楽理論のプログラム化の研究では、様々な立場から多種多様な形式化が行われている [1]。音楽的な概念や関係を表す記号、関数、述語などを明確に定義せずに ad hoc に導入し形式化を行う例もあるが、これでは形式化された音楽記述の持つ意味を正しく理解することは困難になるであろう。

本研究の目的は、楽曲やその構成原理を表現する汎用的で形式的な枠組みを与えることである。抽象化された楽曲や譜例の中で観察される音楽的な概念や現象を、意味が明確に定義されている計算機科学や数学の言葉で記述する枠組と換言しても良いだろう。本研究の第一段階として、一般の音楽理論のテキストで良く見かける「構成音が似ているので似た機能、似た響きを持つ」「この音とこの音が増4度の音程を成すので和音全体では複雑な響きがする」などという文章を形式的に記述する方法について考察を行う。

本稿で表現の対象となるのは、単なる和音の並び(楽曲)であり、これは五線譜などの楽譜に対し非常に極端な抽象化及び捨象を施したものである。本稿では次にその抽象化された楽曲に対して、知識表現の分野で演繹オブジェクト指向(Deductive Object-Oriented Paradigm, DOO)と呼ばれる手法を適用する。その際、和声学から簡単な例を幾つかとりあげ、和音の響きの複雑さ、和音の性質の継承、和音同士の関係、和音列同士の関係などの形式的な記述を行う。



$$A: a[\text{notes} = \{x, y, z\}, x-z = 7, y-z = 4]$$

$$B: b[\text{notes} = \{x, y, z\}, x-r = -2, y-r = 4, z-r = 9]$$

図 1: 簡単な和音の例

## 2 演繹オブジェクト指向について

ここで簡単に DOO について解説する [2][3]. DOO でオブジェクト指向という時の“オブジェクト”は、オブジェクト識別子及び属性と属性値の複数ペアから成っている。メソッドや局所変数などはすべて属性として統一的に扱われる。一方、演繹というのは、知識を一階述語論理で記述して推論を行うことを意味し、その性質は数学的によく研究されている。DOO はこれら 2 つの技術を統合している。述語の現れる場所に、より表現力の高いオブジェクトが書けるよう拡張されている。これにより、一般のオブジェクト指向の特徴である、オブジェクト内への情報隠蔽やオブジェクト階層に沿った性質継承などが、すべて演繹規則で定義したり記述できるようになった。さらに is\_a 関係により表現される知識も演繹規則により定義できる。

まず、本稿では和声学に現れる概念や現象を記述することを考えている。この時、定量化は難しいが順番 (is\_a 関係) を付けることは比較的容易な場合が多い。また情報が部分的にしか分からないような場合も多い。このような対象の表現には DOO の枠組が適していると考ええる。

## 3 和音の記述

まず  $k$  個の音から成る和音は  $\{p_1, \dots, p_k\}$  のように書く。  $p_i$  は音高に対応する整数である。音高の値  $p$  には特に上下限を設けない。各音ごとに異なっている音色、音量、音価などに関する情報はすべて捨象する。和音で演奏された各音が実は微妙にずれて演奏されているといった情報もすべて捨象する。

時間軸に沿った和音の並びに関しては、和音間の前後関係だけ記述できれば十分であると考え、第 4 章で時間の前後関係を表現するオブジェクトを導入する。

### 3.1 オブジェクト識別子と属性

和音は、それを構成する音同士の音高差やある基準となる (仮想的な) 音との音高差などで特徴付けることができる。しかし、それらは部分的な情報に過ぎない。そこで、ある和音 (オブジェクト) に関して、現在分かっている性質 (属性) だけを記述する; つまり、オブジェクト  $c$  の属性  $p_1$  の値が  $v_1$ 、属性  $p_2$  の値が  $v_2$  であるということ、ここでは  $c[p_1 = v_1, p_2 = v_2]$  のように書く。  $c$  は個々のオブジェクトに付与される識別子 (identifier) であり、このような表現手法は identity-based representation と呼ばれる。また、述語を用いて  $p_1(c, v_1) \wedge p_2(c, v_2)$  のように二項関係として書くこともできる。

例: 図 1 に 2 つの和音 A, B を示す。  $x, y, z, r$  はこれらの関係を満たす適当な整数である。ここでのオブジェクト識別子は  $a, b$  である。属性名として  $\text{notes}$  (和音の全構成音) や、  $x-z, z-r$  (ある 2 つの構成音の音高差) などがある。 A の和音は、  $x, y, z$  という 3 つの音から成り、  $x$  と  $z$  の音高差は半音 7 個分、  $x$  と  $z$  の音高差は半音 4 個分であることを表している。同様に B の和音も、  $x, y, z$  という 3 つの音から成り、  $x$  と仮想的な根音  $r$  との音高差が  $-2$ 、  $y$  と  $r$  の音高差が  $4$ 、  $z$  と  $r$  の音高差が  $9$  ということを表している。これら音高差は属性でなく制約と考えることもできよう。

この表現手法では 1 つのオブジェクトに対して、多重ビューを与えることができる。例えば図 1 の A の和音名は  $C$  であるから、  $a[\text{notes} = \{x, y, z\}, a-c = 7, b-c = 4, \text{name} = C]$  のように書くことで、冗長ではあるが、  $a$  というオブジェクトを構成音と和音名という 2 つの側面から表現されている。

A      B      C

A : {65, 68, 71, 74}  
 B : {65, 69, 71, 76}  
 C : {65, 68, 71, 75}

図 2: 3 つの  $G_7$  和音

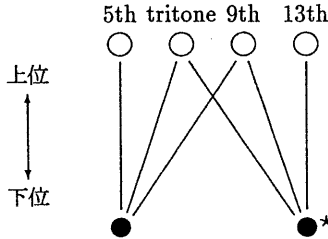


図 3:  $G_7$  の部分音間の階層関係

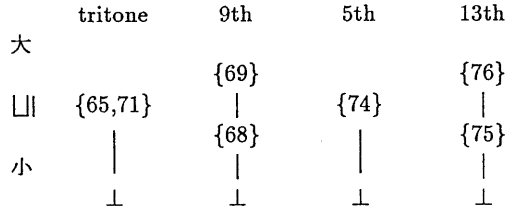


図 4:  $G_7$  の各部分音の属性値の順序関係

### 3.2 和音間の順序関係

例: 図 2 に A ~ C の 3 つの  $G_7$  和音を示す。あるユーザがこれらは似たような響きや機能を持っていると感じたと仮定する。この時、 $G_7$  という和音の持つ構造や、和音同士の関係を形式的に表現してみよう。図 2 のように、以降本稿では音高に MIDI の note number を用いる。

これらの和音を観察したユーザが、例えば次のようなことを考えたでしょう<sup>1</sup>:

- i どの和音も tritone ({65, 71}) と 9th ({68}, {69}) を含んでいる。
- ii 5th ({74}) と 13th ({75}, {76}) は排他的に出現している。
- iii 同じ 13th でも {76} より {75} の方が複雑な響きがする。

i は単純な和音から複雑な和音が合成される、ある和音に対して差分を記述することで新しい和音が生成されるということの意味している。その部分的な和音は階層を構成し和音間に階層を定めることができる。そして、この階層関係に沿って単純な和音の響きが複雑な和音の響きの一部として継承される。ii は選言 (OR 関係の分岐) が生じていることを意味している。iii はある構成音を半音ずらすとより複雑な響きが得られることを意味している。従って半音ずらす前と後に順序がつけられる。

これより、図 2 の和音を構成する部分音間の階層関係は、例えば図 3 のようになる; もちろんこの他の階層関係もユーザ次第でいくつか考えられる。図中、上から下に向かって属性が継承され、どちらの ● のオブジェクトにおいても  $G_7$  と考えられる和音が属性 notes の値として得られる; いずれも tritone と 9th は必ず一緒に継承し、5th と 13th は排他的に継承する。さらに、tritone や 9th という部分音内部の属性値にも順序関係が存在している (図 4)。これらの和音オブジェクト間の階層関係や順序関係をまとめて、包摂関係と呼び  $\sqsubseteq$  という記号で表す。 $\sqsubseteq$  は is.a 関係の拡張とみなすこともできる。その直観的な意味は、複雑  $\sqsubseteq$  単純, 具体的  $\sqsubseteq$  抽象的, 特殊  $\sqsubseteq$  一般的, である。 $\sqsubseteq$  は推移律を満たす。図中、下に向かって和音 (の値) はだんだん小さくなって行き、 $\perp$  は複雑過ぎて矛盾が生じてしまった場合を表している。オブジェクト間の包摂関係, 属性継承という性質はオブジェクト指向の考え方に沿ったものである。

<sup>1</sup> これらの観察はもちろん状況に依存する。この例では状況について言及していないが、状況の論理的な構造をユーザが認識できれば、それも DOO の枠組みの中で陽に記述することができる。

### 3.3 属性の継承

DOO の枠組では、オブジェクトの属性継承を演繹するための規則は一般に以下のように定義されている:

$$o_l \sqsubseteq o_h \Rightarrow v_l \sqsubseteq v_h \text{ s.t. } o_l[p = v_l], o_h[p = v_h]$$

従ってこの定義から

$$o_l \sqsubseteq o_h \wedge o_h[p = a] \Rightarrow o_l[p \sqsubseteq a]$$

が導かれる。  $o_l[p \sqsubseteq a]$  は  $o_l[p = v] \wedge v \sqsubseteq a$  の略記である。これを言葉で表現すると、上位オブジェクト  $o_h$  の属性  $p$  の値  $a$  がその下位オブジェクト  $o_l$  に継承されると、 $o_l$  における属性  $p$  の値  $v$  は  $v \sqsubseteq a$  となることを意味している。また DOO における多重継承は一般に以下のように定義される:

$$o_l \sqsubseteq o_{h1} \wedge o_l \sqsubseteq o_{h2} \wedge o_{h1}[p \sqsubseteq a] \wedge o_{h2}[p \sqsubseteq b] \Rightarrow o_l[p \sqsubseteq \text{meet}(a, b)]$$

ここで  $\text{meet}(a, b)$  は  $a$  と  $b$  の greatest lower bound を取る演算で、集合の場合は union ( $\cup$ ) に相当する。

例: 図 3 の  $\bullet^*$  の和音 ( $G_7$  オブジェクト) で得られる属性値を調べる。和音間の継承関係と各部分和音の属性値は以下のように記述できる。

和音間の継承関係:  $G_7 \sqsubseteq \text{tritone}$ ,  $G_7 \sqsubseteq 9th$ ,  $G_7 \sqsubseteq 13th$

各部分和音の属性値:  $\text{tritone}[\text{notes} = \{65, 71\}]$ ,  $9th[\text{notes} = \{69\}]$ ,  $13th[\text{notes} = \{76\}]$

すると、上述した多重継承の定義から、 $G_7$  で得られる属性  $\text{notes}$  の値は  $G_7[\text{notes} \sqsubseteq \{65, 71, 69, 76\}]$  となる。従って実際の  $\text{notes}$  の値は、 $\{65, 71, 69, 76\}$  を部分集合に含むような集合になる。少なくともシステムが最低限演繹しなければならないのはここまでであり、 $\{65, 71, 69, 76\}$  より実際に要素の多い集合をユーザに与えるかどうかはシステム依存である。

### 3.4 オブジェクトの包摂関係

属性間の包摂関係はそれを含むオブジェクト間の包摂関係にも影響を与える。以下に DOO における一般的な包摂関係の演繹規則を与える:

$$o_l \sqsubseteq o_h \Rightarrow o_l[p = v] \sqsubseteq o_h \quad \text{S1}$$

$$o_l \sqsubseteq o_h \wedge v_l \sqsubseteq v_h \Rightarrow o_l[p = v_l] \sqsubseteq o_h[p = v_h] \quad \text{S2}$$

ここで  $o_h, o_l$  はオブジェクト、 $p$  は属性名、 $v, v_l, v_h$  は属性値である。ここで  $o[\dots][p = v]$  と  $o[\dots, p = v]$  は同じ意味であるとする。

例: 図 4 より  $13th[\text{notes} = \{75\}] \sqsubseteq 13th[\text{notes} = \{76\}]$  なので、図 3 の  $\bullet^*$  で得られる  $G_7$  オブジェクトに関して、S2 より、


$$G_7[\text{notes} = \{65, 71, 69, 75\}] \sqsubseteq G_7[\text{notes} = \{65, 71, 69, 76\}]$$

という包摂関係を演繹することができる。この関係は、他の構成音が等しければ  $13th$  より  $b13th$  を含む和音の方が複雑な響きを持つ、ということを表現している。

## 4 楽曲の記述

楽曲を構成するには、和音を時間軸方向に並べたり、旋律を重ね合わせたり、2 個以上の楽曲をつなげる、重ね合わせるなどの方法がある。まず、和音を時間軸方向に並べて楽曲を構成する場合を考える。


Tune 1



$t_{1a}[\text{notes} = \{59, 65\}, 65-59 = 6]$   
 $t_{1b}[\text{notes} = \{60, 64\}, 64-60 = 4]$

図 5: 2つの和音からなる楽曲の例

Tune 2



$t_{2a}[\text{vl} = \{72 \rightarrow 74\}]$   
 $t_{2b}[\text{vl} = \{64 \rightarrow 65\}]$

図 6: 2つの旋律 (楽曲) からなる楽曲の例

例: 図 5 では,  $\{59, 65\}$  と  $\{60, 64\}$  という 2 和音を連結 (juxtaposition) して楽曲を構成している. 機能 and 声では, 増 4 度の不安定な音程が長 3 度の安定した音程に解決し,  $V \rightarrow I$  というケーデンスが形成されていると解釈される. 図 5 の楽曲の特徴 (部分的な情報) を, 和音の時と同様にオブジェクト  $t_1$  とその属性  $(s, \text{vl})$  という形式で表現すると,  $t_1[s = t_{1a} \rightarrow t_{1b}, \text{vl} = \{65 \rightarrow 64, 59 \rightarrow 60\}]$  のように記述できる. 属性  $\text{vl}$  は voice leading の意味である. また  $c_1 \rightarrow c_2$  という記法は, ‘ $\rightarrow$ ’ [ $l = c_1, r = c_2$ ] というオブジェクトの略記である. つまり, 新しく  $\rightarrow$  というオペレータを導入する代わりに,  $l, r$  という属性を持った  $\rightarrow$  というオブジェクトを導入して和音オブジェクト間の前後関係を表現している. また  $\rightarrow$  は和音オブジェクトの合成とも考えることができる.

例: 図 6 では, 旋律 (楽曲) を重ね合わせて (superimposition) 楽曲を構成している. Tune 2 では  $72 \rightarrow 74$  という楽曲 ( $t_{2a}$ ) と  $64 \rightarrow 65$  という楽曲 ( $t_{2b}$ ) が重ね合わされていると考える. superimposition は多重継承を用いて,  $t_2 \sqsubseteq t_{2a}, t_2 \sqsubseteq t_{2b}$  のように表現できる. これより  $t_2[\text{vl} \sqsubseteq \{72 \rightarrow 74, 64 \rightarrow 65\}]$  が演繹される.


次に, 楽曲も 1 つのオブジェクトであるから, 第 3.4 節で述べた包摂関係の演繹規則 S1, S2 を適用することで, 楽曲間に  $\sqsubseteq$  に関する順序をつけることができる.

例: 図 7 に Tune 1 と似た楽曲 Tune 3 を示す. Tune 3 は, Tune 1 の各和音にそれぞれ A (69) と G (67) の音を付加したものである. この楽曲  $t_3$  は,  $t_{3a}, t_{3b}$  を用いて以下のように表現できる:

$$t_3[s = t_{3a} \rightarrow t_{3b}, \text{vl} = \{65 \rightarrow 64, 59 \rightarrow 60\}]$$

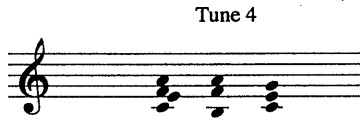
上述したように  $c_1 \rightarrow c_2$  がオペレータによる演算ではなく  $\rightarrow$  というオブジェクトの略記法であることから,  $t_{3a} \sqsubseteq t_{1a} \wedge t_{3b} \sqsubseteq t_{1b} \Rightarrow t_{3a} \rightarrow t_{3b} \sqsubseteq t_{1a} \rightarrow t_{1b}$  という関係が演繹され, 従って  $t_3[\dots] \sqsubseteq t_1[\dots]$ , つまり  $t_1$  より  $t_3$  の方が小さい (複雑な和音列である) ことが分かる.

Tune 3



$t_{3a}[\text{notes} = \{59, 65, 69\}, 65-59 = 6]$   
 $t_{3b}[\text{notes} = \{60, 64, 67\}, 64-60 = 4]$

図 7: Tune 1 より小さい楽曲の例



$t_{3a}[\text{notes} = \{59, 65, 69\}, 65-59 = 6, s = t_{4a} \rightarrow t_{4b}]$   
 $t_{3b}[\text{notes} = \{60, 64, 67\}, 64-60 = 4]$   
 $t_{4a}[\text{notes} = \{60, 64, 65, 69\}]$   
 $t_{4b}[\text{notes} = \{59, 65, 69\}]$

図 8: Tune 3 より小さい楽曲の例

例: さらに Tune 3 より小さい 3 と音から成る楽曲を図 8 に示す。DOO による表現は、この楽曲を  $(II_{m7} \rightarrow V) \rightarrow I$  というケーデンスであると解釈した場合に対応している。図 7 と図 8 に現れる  $t_{3a}$  を比較すると、S1 から、 $t_{3a}[\dots, s = t_{4a} \rightarrow t_{4b}] \sqsubseteq t_{3a}[\dots]$  という関係が演繹でき、Tune 4 は Tune 3 より小さいことが分かる。また図 7 の  $t_{3a}$  を  $t_{3a}[\text{notes} = \{59, 65, 69\}, 65-59 = 6, s = T \rightarrow t_{4b}]$  のように変更しても同じ結果が得られる。ここで T は全てのオブジェクトに対して最も単純かつ抽象的な要素 (最大値) を表す。

## 5 おわりに

和声学に関する音楽的な概念の一部が、オブジェクト指向的手法 (継承, 包摂関係) や演繹規則によって形式的に記述できることを示した。本稿で取り上げた例はいずれも和声学からの簡単なものにとどめたが、演繹規則の表現力は非常に高いので、本手法は、対位法、編曲法などの音楽理論の記述や、認知科学的なモデルの記述にも応用できると考えている。他の応用として、音楽理論のデータベースシステムなども考えられるだろう; 例えば、ある部分と音を含むような和音や与えられた和音に似た和音を問い合わせたり、与えられた楽曲の構造を解析したり類似度を計算することなどである。

音楽的な概念や現象の形式化に対して、その正当性を議論することは大変困難であると思われる。例えば自然言語文を形式化する場合は、その論理的な意味や構造を比較的明確に認識できるので、形式化された表現との対応を取り易い。しかし、音楽に関しては、対象の論理的な意味や構造を明確に認識できないことが多いからである。一方、本稿で試みた形式化の手法や枠組に関する正当性に関しては、認識できた論理的な意味や構造をその手法で効率的に表現できるかどうかを議論することになる。本稿は DOO の枠組を用いて、和声学に関して認識されている論理的な意味や構造の一部を、自然な対応の元で 1 種類の順序 ( $\sqsubseteq$ ) により表現できることを示したと言えよう。

今後の課題として、他の知識表現手法との比較、他の音楽的な概念の形式化、本稿で提案したシステムの実装、本形式化の枠組みに基づいた帰納学習などを考えている。

謝辞: ベルリン自由大学 渋野昌氏、電気通信大学 青柳龍也氏には、本稿の草稿段階から議論をして頂き大変有益な示唆を得ることができました。NTT 基礎研究所 小坂直敏氏、原田康徳氏から貴重な意見を頂きました。NTT 基礎研究所 勝野裕文氏、ICOT 横田一正氏から演繹オブジェクト指向データベースに関して多大なるご教示を頂きました。本研究を行うにあたり NTT 基礎研究所 尾内 K 理紀夫氏は常に暖かく励まして下さいました。楽譜の印刷に関して九州大学 菺田敏行氏 作成の PSCO Tiny を用いました。PSCO Tiny のカスタマイズは ICOT 久門耕一氏にお世話になりました。

## 参考文献

- [1] 平田圭二, 音楽の表現手法, 人工知能学会誌, Vol.9, No.4, pp.392-398 (1993).
- [2] 横田一正, 演繹オブジェクト指向データベースについて, コンピュータソフトウェア Vol.9, No.4, pp3-18 (1992).
- [3] 横田一正, 安川秀樹, 統合知識ベース管理システム - FGCS プロジェクトにおけるデータベース・知識ベースの研究開発の概要 -, 第五世代コンピュータの研究開発成果, FGCS'92, ICOT, pp80-102 (1992).