

パーピーブン：ジャズ和音を生成する創作支援ツール

平田 圭二[†] 青柳 龍也^{††}

本稿では、現在開発中の創作支援ツール「パーピーブン」の設計方針、音楽知識表現手法、動作原理、実装等について述べる。パーピーブンは、単純な和音進行が与えられると、ケーデンス単位で和声的文脈を考慮しながらユーザの意図を反映した演奏を生成する創作支援ツールである。創作支援ツールは多様な刺激をユーザに与えることが第一義であり、そのために、音楽知識を演繹オブジェクト指向データベースによって表現し、事例ベース推論を行う。ユーザの意図を表現するために既存の楽曲分析理論を援用し、新たに包摂関係を利用した相対的類似度を提案し、音楽的な直観に合致する和音進行の類似検索を実現した。ユーザ意図を表現するのにケーデンス木が有効であり、実際に多様で一貫した曲調の楽曲を効率的に生成できることを確認した。

Pa-Pi-Pun: Creativity-Support Tool for Generating Jazz Harmony

KEIJI HIRATA[†] and TATSUYA AOYAGI^{††}

This paper presents the design principle, musical knowledge representation method, system organization, and implementation of a musical system called "Pa-Pi-Pun". Pa-Pi-Pun is a creativity-support tool that takes a simple chord progression as an input and then generates a new progression that reflects the user's intention and takes into account the harmonic context on a cadence basis. The primary goal of a creativity-support tool is to give diverse stimulation to a user. In order to do so, Pa-Pi-Pun employs a deductive object-oriented database technique for musical knowledge representation and case-based reasoning. Moreover, it invokes an existing music analysis theory to represent the user's intention, introduces a relative similarity using the subsumption relation, and uses a new case retrieval method based on the relative similarity which provides a high affinity for musical introspection and intuition. It has been confirmed that a new data structure, called the cadence tree, is effective in representing a user's intention and various, yet consistently flavored, pieces are created efficiently.

1. はじめに

今後ますますインターネットが普及しマルチメディア技術が身近になるにつれ、人間が物をデザインし作品を創作する行為は文化的、経済的により重要になるであろう。我々は、ユーザの創造性を高める支援ツールとして計算機はどのようにあるべきかに興味があり、音楽の創作という具体例を取り上げ、実際に創作支援ツールを構築することでこの問題に取り組んでいる。

「パーピーブン」は、ユーザが単純な和音進行を与えると、それに対するボイス（和声付け）を生成する音楽創作支援ツールである。パーピーブンの要件は、記述されたユーザの意図を適切に把握し、それを反映し

た多様な刺激（音や楽曲）をユーザに提供することである^{1),6)}。ユーザは予め生成したい楽曲に対する明確な意図を持つことが期待されているので、必然的に音楽的に高いスキルを持つことになる。また、我々はパーピーブンの処理が可能な限り決定的になるよう設計した。これは、確率モデルやアドホックなアルゴリズムを採用すると、パーピーブンの動作予測が困難になり、ユーザの意図通りに制御できなくなってしまうからである。

ユーザの意図を適切に把握するという観点からはインタフェースレベルや内部モデルのレベルが検討課題となり、ユーザの意図を反映した多様な刺激を提供するという観点からは楽曲の生成方式が検討課題となる。

インタフェースレベル: ユーザの意図通りにツールを制御する際、低レベルなインタフェース (I/F) を直接ユーザに提示するか、抽象化して高レベルの I/F をユーザに見せるかという選択肢がある。

I/F レベルを比較的低く設定したツールでは、ユーザに提供するプリミティブの粒度は細かく、ユーザの意図

[†] NTT コミュニケーション科学基礎研究所
NTT Communication Science Laboratories

^{††} 津田塾大学 情報数理科学科
Department of Mathematics and Computer Science,
Tsuda College

をシステムに伝える作業はプログラミングに近い。自由度が高いという利点を持つが、ツールをあるユーザや用途に特化する手間が大きい。一方、I/F レベルを比較的高く設定したツールは、メロディ、和声、リズム等の高次の音楽概念をユーザに提供する。簡便性、自動化の可能性、効率の良い意図伝達等の利点がある。しかし、高レベルな I/F を提供することはツールにあらかじめ特定の音楽知識を埋め込むことに等しく、これがユーザの自由な発想に制約をかける、生成する楽曲にユーザの意図しない偏りを付加する、ユーザの制御範囲を限定するという理由で好ましくないという議論もある²²⁾。

I/F レベルが低いシステムには、パッチ言語 MAX²⁵⁾ や楽音合成システム Csound⁴⁾ 等がある。I/F レベルが高いシステムには、Cypher²¹⁾、UPIC²⁶⁾ 等の作曲システム、Band-in-a-Box 等の自動演奏ソフトなどがある。

パービーブンは、I/F に関して簡便性と高い自由度の両立を目指す。つまり、ジャズピアノ和声に関する最小限の音楽知識を組込むことで、比較的高い I/F レベルを実現し、さらに、ユーザの意図を柔軟に表現するためのデータ構造(ケーデンス木)を新たに導入する。

楽曲の生成方式: 楽曲を生成する方式は事例ベースとルールベースに大別できる。

事例ベースにおける事例は人が実際に演奏した楽曲(の断片)であり、元の演奏や楽曲の雰囲気を持続した最小単位である。そのようなすでに存在している楽曲の断片をうまく再組合せすることで元の雰囲気を保持した新たな楽曲を再構成することができる¹⁷⁾。

一方、ルールベースにおけるルールで表現されている音楽知識は、文脈情報の多くを捨象し平均化したものである(例えば機能和声法)。そのルールを文法規則あるいはオートマトンと見なして新たな楽曲を生成する。標準的なスタイルが確立された音楽ジャンルの楽曲において、これまで発表された楽曲生成システムの多くはルールベースに基づいていた³⁾。文献 14) は、ルールベースで奏法や曲調を複製したような雰囲気の楽曲を生成するためには、正規文法より表現力の高い文法が必要になること、確率統計的な要素を取り入れる必要があることを指摘している。

一般に、ルールベースの欠点は、知識獲得、明示的記述、推論制御の困難さにあると言われている。例えば、機能と和声に制約されない和音進行の規則、楽曲における文脈や曲の流れといった高次の音楽的情報や概念は、時間、個人、場所等によって多様に変化する。このように主観的、暗黙的かつ曖昧であり、文脈依存性が高い対象に関する知識を専門家から聞き取り、ルールとして明示

的に書き下すのは大変困難である。さらに解を計算する過程を制御するには、事実や知識はそのまま、メタルールや選好ルール (preference) を導入するのが一般的だが、ここでも前述のルール記述の困難性の問題が発生する。

これに対し、事例の収集とその分析は比較的容易であり、事例をより構造化したり複雑化することで、推論制御に相当する操作が実現できる。

以上の議論より、我々は、創作支援ツールの生成方式は事例ベースの方が好ましいと判断した。

パービーブンにおける事例: 適切なレベルの事例としては様々な候補が考えられるが、実際の音楽家が和音列にボイスリングを付加する時は、和音やケーデンス* を事例として解釈することが多い。和音を事例と見なすと元演奏におけるボイスリーディング(隣接した和音の各構成音間の接続関係)に関する情報が失われてしまうと考えると、パービーブンではケーデンスを事例とした。

本論文の構成は以下の通りである。続く第2章ではパービーブンが採用している音楽知識表現について、第3章ではパービーブンのシステム構成および動作原理について、第4章では関連研究について述べ、第5章で本論文を締めくくる。

2. 音楽的概念や関係の表現

ユーザの意図を伝達し記述する際の重要な検討課題は、どのような音楽的概念や関係をどのように表現するかである。

パービーブンが扱う音楽的概念や関係は、1つの音、和音名、ケーデンス、ケーデンスから成る木構造、ボイスリング、ボイスリーディングである²³⁾。これまで様々な音楽知識表現手法が提案されているが^{7),10),24)}、我々はこれまでの一連の先行研究^{8),9),11)~13)}により、演繹オブジェクト指向データベース(Deductive Object-Oriented Database, DOOD)^{15),27)}と音楽知識の親和性の高さを確認している。DOODの枠組を用いると、これら抽象度の異なる音楽的な概念や関係全てをオブジェクト項と包摂関係に自然に対応付けることができ、音楽的な概念や関係を統一的に明示的に表現、操作できる²⁸⁾。これより、音楽的な直観に合致した内部モデルを構成できる。

2.1 DOOD と包摂関係

DOODでは、対象をオブジェクトとその属性の集

* cadence. 終止感や段落感を与えるような特定の和音進行パターン。ジャズやポップスの場合、通常、数個の和音列から成る²³⁾。

** 素性構造¹⁶⁾を用いても同様の枠組を構築できるだろう。

合から成るオブジェクト項として表現する。本稿では $o(\dots, l = v, \dots)$ という表記法を用いる。ここで o は基本オブジェクト項, $l = v$ は属性, l は属性ラベル, v は属性値を表わす。

包摂関係 (\sqsubseteq) とは「情報量が多いオブジェクト \sqsubseteq 情報量が少ないオブジェクト」という意味である。あるいは「具体的なオブジェクト \sqsubseteq 抽象的なオブジェクト」, 「特殊 \sqsubseteq 一般」を意味する。オブジェクト項が表す集合の意味を考えて, 集合の包含を表す記号 (\subseteq) と向きを揃えるために上記のような記法を採用している。

オブジェクト項間の包摂関係は以下の演繹規則によって定義される。今 $o_1 = p(\dots, l_m = v, \dots)$, $o_2 = q(\dots, l_n = w, \dots)$ とすると, o_1, o_2 間の包摂関係は,

$$o_1 \sqsubseteq o_2 \leftarrow p \sqsubseteq q \wedge \forall m \exists l_n (l_m = l_n \wedge v \sqsubseteq w)$$

と定義される。つまり o_1 の基本項 (p) と全ての属性 (l_m) について, それらが全て具体的ならば, o_1 の方が具体的 (あるいは o_2 の方が抽象的) という意味である。

もし o_1, o_2 がオブジェクト項を要素に持つ集合の場合, o_1, o_2 間の包摂関係として様々な定義が考えられる。以下に代表的な 2 つを演繹規則として記述する。

$$\begin{aligned} o_1 \sqsubseteq_H o_2 &\leftarrow \forall s_i \in o_1 \exists t_j \in o_2 s_i \sqsubseteq t_j \\ o_1 \sqsubseteq_S o_2 &\leftarrow \forall t_j \in o_2 \exists s_i \in o_1 s_i \sqsubseteq t_j \end{aligned}$$

\sqsubseteq_H はいわゆる集合の Hoare 順序に, \sqsubseteq_S は Smyth 順序に等しい。例えば, $\{b, d\} \sqsubseteq_H \{a, b, c, d\}$, $\{a, b, c, d\} \sqsubseteq_S \{b, d\}$ のようになる。Hoare 順序は集合要素間に選言の意味がある場合に, Smyth 順序は連言の意味がある場合に用いられる。

2.2 1 つの音の表現

1 つの音には様々な種類がある。例えば, C5 と言うと, それは特定の音高を指す。C5 に音色や演奏する楽器の情報は含まれていない。C と言うと, オクターブを無視した C の音の集まりや「ト長調のスケールは G A B C ... から成る」と言った時の C などを指す。ここで音の長さや音の絶対的な高さ (周波数) などの情報は含まれていない。さらに, 単なる 1 音と言うと, C や F などのいわゆる音全体を総称する概念を指す。

これらを, 以下のようなオブジェクト項で表現する。

| | |
|---------|-------------------------|
| 単なる 1 音 | note |
| C | note(pitch=C) |
| C5 | note(pitch=C, octave=5) |

ここで, note が基本オブジェクト項, pitch = C が属性, pitch が属性ラベル, C が属性値である。

この時, note(pitch = C, octave = 5) と note(pitch = C) を比較すると, note(pitch=C, octave = 5) の方が octave = 5 の分だけ記述が具体的なので,

$$\begin{aligned} \text{note}(\text{pitch} = \text{C}, \text{octave} = 5) &\sqsubseteq \text{note}(\text{pitch} = \text{C}) \\ \text{note}(\text{pitch} = \text{C}) &\sqsubseteq \text{note} \end{aligned}$$

という包摂関係が成り立つ。

2.3 和音名の表現

和音名とは, 音の集合 (和音) に対する識別記号であり, 例えば C₇ や D^b_{M7} などと書く。一般にジャズの和音名は, 根音, 和音の型, 付加されるテンション音という要素から構成されるので, root, type, tension という 3 つの属性ラベルを設ける。例えば, Eb_{7b5}(9,11) という和音名は chord(root = Eb, type = 7b5, tension = {9,11}) と, V₇ は chord(root = V, type = 7) と表現される。ここで, Eb は音名, V は階名である。

例として, G₇(9) と G₇(9,13) の間の包摂関係を考える。各々のテンション音は {9}, {9,13} という集合で表現され, テンション音どうしの間には連言の意味があるので, Smyth 順序を用いて比較する。よって, chord(root = G, name = 7, tension = {9,13}) \sqsubseteq chord(root = G, name = 7, tension = {9}) となる。

2.4 ケーデンスの表現

パーピープンでは, オブジェクト項として表現されたケーデンスを事例とする。例えば C 調の D_{m7}-G₇-C_{M7} というケーデンスを表現することを考える。このケーデンス全体を 1 つのまとまりと見なし, C_{M7} という和音でそれ全体を代表していると解釈する。すると, D_{m7}-G₇-C_{M7} \sqsubseteq C_{M7} という包摂関係が成り立つべきである。よって, ケーデンスを表現するオブジェクト項が持つ属性ラベルとして, 構造特徴 (prom), ケーデンス列 (seq), 調 (key) の 3 つを設ける。一例を示す。

$$\begin{aligned} \text{cdc}(\text{prom} = \text{C}_{M7}, \\ \text{seq} = [\text{cdc}(\text{prom} = \text{D}_{m7}), \\ \text{cdc}(\text{prom} = \text{G}_7) \\ \text{cdc}(\text{prom} = \text{C}_{M7})], \\ \text{key} = \text{C}) \end{aligned}$$

ここで, cdc はケーデンスを表現するための基本オブジェクト項である。属性 seq は, ケーデンスを構成する実際の和音進行を表す。パーピープンではケーデンスを基本単位とし, それより下の音楽的な構造には着目しな

いので, seq の属性値は和音名を要素とする単純なリスト構造としている. このリストの要素が, chord オブジェクトでなく cdc オブジェクトとなっている理由は 2.5 節で述べる. prom という属性は, この和音進行に含まれる和音の中で最も重要 (構造特徴的, prominent) な和音を表す. seq 属性の値が表す和音進行は, この構造特徴的な和音で代表される. key 属性はそのキーデンスの調を表す.

2.5 ケーデンスから成る木構造の表現

全ての和音は必ずいずれかのケーデンスに属し, 楽曲は複数のケーデンスの連なりから構成される. しかし, ケーデンスは単に一次的に並んでいるのではなく, ケーデンス間にも並置, 従属などの関係がある. 例えば, 同じ和音進行でも曲中の出現場所が異なると, その意味が変わる. 我々は, このケーデンス間の関係を表現するためにケーデンス木という構造を導入する. ケーデンス木とはケーデンスを節や葉とする木構造であり, 曲全体の流れや文脈を表現する. パーピーブンは, ケーデンス木に基づいて前後の文脈を考慮に入れた再組合せを行うことで, 音楽的に一貫性のある解を出力できる. seq の属性値中に現れる和音が cdc の形になっているのは, ケーデンス木を構成するためである.

前節で例として挙げた和音進行に対し, $(D_{m7}-G_7)-C_{M7}$ という解釈も考えられる. ここで, $D_{m7}-G_7$ という部分ケーデンス (C 調) の構造特徴的な和音を G_7 と考える. すると, このケーデンス全体は,

$$\begin{aligned} & \text{cdc}(\text{prom} = C_{M7}, \\ & \quad \text{seq} = [\text{cdc}(\text{prom} = G_7, \\ & \quad \quad \text{seq} = [\text{cdc}(\text{prom} = D_{m7}), \\ & \quad \quad \quad \text{cdc}(\text{prom} = G_7)], \\ & \quad \quad \text{key} = C), \\ & \quad \text{cdc}(\text{prom} = C_{M7})] \\ & \quad \text{key} = C) \end{aligned}$$

というケーデンス木で表現される.

ケーデンス木を用いると, ケーデンス木全体のどこにどのような部分ケーデンスが出現しているのかを表現することができる. 先の例に関しては

$$\begin{aligned} D_{m7}-G_7-C_{M7} & \sqsubseteq C_{M7} \\ (D_{m7}-G_7)-C_{M7} & \sqsubseteq G_7-C_{M7} \sqsubseteq C_{M7} \end{aligned}$$

という包摂関係が成り立つ.

ここで提案したケーデンス木は, Generative Theory of Tonal Music (GTTM)¹⁹⁾ の Time-Span Reduction (TSR) を和音進行に応用したものと考えることができる. TSR による楽曲分析は次のように行われ

る. まず楽曲に含まれるすべての音や和音を 2-3 個ずつのグループにまとめる. 各グループ中の構造特徴的な音や和音でそのグループを代表させ, 1 段上位レベルでも同様に 2-3 個ずつのグループにまとめる. 最終的に楽曲全体が 1 つのグループとなるまでこのグルーピングを繰り返す.

我々は, TSR によるグルーピングで作られる階層構造 (つまりケーデンス木) が曲の解釈に相当し, 曲の流れや和声に関する文脈 (和声文脈) を表現していると思なす. 同じ和音進行でも異なるケーデンス木を構成することが可能であり, それは異なる解釈を表現していることに相当する.

2.6 ボイシングの表現

ボイシングとは, ある和音列やケーデンス, モチーフやテーマが与えられた時, それに対して実際に演奏すべき和音を作り出すことである. 音楽家は, ある和音やケーデンスを見ると可能なボイシングを連想し, 数多くのボイシング候補の中から最も適していると思えるものを選択する.

パーピーブンは, 項の識別に影響を及ぼさない属性, 即ち非固有属性を用いてこの連想機能を実現する. ここで新たに '/' という構成子を用いて, $o(\dots, l = v, \dots) / (\dots, m = w, \dots)$ のようなオブジェクト項を導入する. '/' の右辺に現われる $m = w$ が非固有属性である (m は属性ラベル, w は属性値). 人間は一般に, ある和音列に対し複数通りのボイシングを連想したり, 異なるケーデンスに対して同じボイシングを連想する. 従って, ボイシングの情報から元の和音進行を一意に識別できず, ボイシングは非固有属性として記述する.

ボイシングは $\text{cdc}(\text{prom} = \dots, \text{seq} = \dots, \text{key} = \dots) / (\text{play} = \{\text{voicing}(\dots), \text{voicing}(\dots)\})$ のようなオブジェクト項で表現する. 非固有属性 play の属性値は, 実際に演奏された様々なボイシング (voicing オブジェクト) の集合となる. '/' の左辺は $\text{cdc}(\dots)$ という和音進行等を表すオブジェクト項であり, 識別子として機能する. voicing オブジェクトの属性には実際に演奏された和音の列 (図 3 (2) 参照) が現われるが, voicing オブジェクトは検索対象にはならないので, この和音の列の表現は適当な方法で構わない.

2.7 ボイスリーディングの表現

ボイスリーディングとは, 隣接した和音の各構成音間の接続関係のことである. 各構成音間が等音高, 半音差, 全音差のいずれかの時に, 構成音間には接続があると考えられる. 例えば, 2 つの和音 C_1 ($\{57, 60, 65\}$),

C_2 ($\{55, 60, 64\}$) を考える^{*}。この時、 $C_1 \rightarrow C_2$ という和音進行には、 $57 \rightarrow 55, 60 \rightarrow 60, 65 \rightarrow 64$ というボイスリーディングが見い出せる。これ以外の組み合わせは、等音高、半音差、全音差という条件に合致しない。ここで、集合と新たな構成子 \downarrow を用いて上例のボイスリーディング全体を $\{57 \downarrow 55, 60 \downarrow 60, 65 \downarrow 64\}$ と表現する。 $i \downarrow j$ は、音高 i から音高 j への個々のボイスリーディングを表す。

パーピーブンでは、 $i - j$ の絶対値が小さいほど円滑なボイスリーディングと定義する。この定義は以下 6 つの演繹規則により表現される。

$$\begin{aligned} i \downarrow i \sqsubseteq i \downarrow i+1 & \quad i \downarrow i \sqsubseteq i \downarrow i-1 \\ i \downarrow i+1 \sqsubseteq i \downarrow i+2 & \quad i \downarrow i+1 \sqsubseteq i \downarrow i-2 \\ i \downarrow i-1 \sqsubseteq i \downarrow i+2 & \quad i \downarrow i-1 \sqsubseteq i \downarrow i-2 \end{aligned}$$

これより、最も円滑なボイスリーディングを選ぶためには、包摂関係に関して最小（極小）のボイスリーディングを選べば良い。

3. システム構成と動作

本章では、DOOD に基づく音楽知識表現をどのように用いて事例ベース推論を実現しているかについて述べる。

3.1 全体構成

図 1 にパーピーブンのシステム構成を示す。図中、

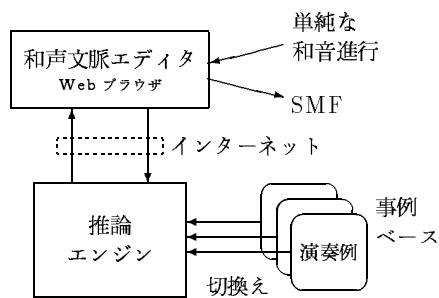


図 1 パーピーブンのシステム構成

和声文脈エディタ (Context Editor) は、パーピーブンをインタラクティブに意図通りに操作するための GUI であり、Web ブラウザから起動する Java アプレットとして実装した。ユーザは和声文脈エディタを用いて和音進行をケーデンス木として入力する。このケーデンス木の入力編集によって、和声文脈や曲の流れに関

するユーザの意図や解釈を記述する。推論エンジン部は、和声文脈エディタで作成した問合わせを受けて事例ベース推論^{**}を行い、その結果を標準 MIDI ファイル (SMF) として出力する。推論エンジン部は並行論理型言語 KL1¹⁸⁾ で実装した。和声文脈エディタから MIDI Plug-in を起動し、その SMF を試聴することができる。

和声文脈エディタのウィンドウを図 2 に示す。図中央

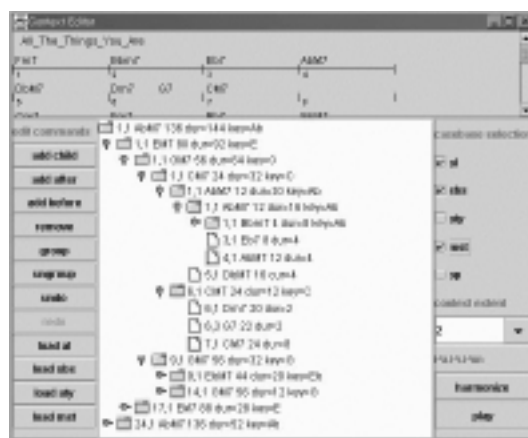


図 2 和声文脈エディタのウィンドウ

の白い部分には、ケーデンス木がその部分木の深さに対応した段下げとともに表示される。ユーザの指示で部分木の表示を畳んだり展開したりできる。ウィンドウ左側には、ケーデンス木の操作コマンド (add child, undo など) と組込みの問合わせをロードするボタン (load all, load sbs など) があり、ユーザはこの操作コマンドを用いて問合わせのケーデンス木を自由に編集する。ウィンドウ右側には、事例ベースの選択ボタン (al, sbs など)、和声文脈の深さ指定 (context extent, 3.6 節)、harmonize ボタン、play ボタンがある。現在のプロトタイプシステムでは、4 人のピアニストの長調 3 曲、短調 1 曲の演奏から採譜した 5 種類の事例ベースが提供されている。

パーピーブンの大まかな操作手順は次のようである。まず最初に事例ベースを作成し (3.2 節)、ユーザはパーピーブンと対話しながら、与えられた和音進行に対してユーザの意図を反映させ、楽曲の流れやボイスリーディングを考慮に入れて演奏を構成していく (3.3 節)。以下、順に説明する。

^{*} 本節では可読性のため、和音を MIDI ノートナンバで表現された構成音の集合として表現する。

^{**} パーピーブンでは、一般の事例ベース推論に含まれる事例ベース更新は行っていない。

(1) Original musical score showing two measures. The first measure contains a Cm7 chord and the second measure contains an F7 chord. Dashed boxes group the notes and chords.

(2) The grouped notes and chords from (1), showing the voicing of each chord.

図3 実際の演奏のグルーピング

3.2 事例ベースの作成

以下のようなステップで事例ベースを人手により作成する。パーピーブンにおける1つの事例とは、非固有属性 play を持つ cdc オブジェクト項 (2.6 節) である。

ステップ 1. 人間の実際の演奏を採譜する (図3 (1)) .
 ステップ 2. 実際に演奏された音や和音の発音タイミングの近いものどうしをグルーピングする (図3 (2)) .

ステップ 3. 演奏された曲の和音進行をケーデンス単位まで分析する (図4) .

ステップ 4. 上の 3. で分析した各ケーデンスに, 2. で作成した対応する和音列を play 属性として付与し, 全ケーデンスを事例ベースに格納する.

ケーデンス木の途中ノードに対応するケーデンスも事例として事例ベースに格納する.

ステップ 2 におけるグルーピングの例を図3に示す. 実際の演奏 (1) において, 音楽的に重要な音を破線のようにグルーピングして (2) のような和音の列を得る. 個々の和音が voicing オブジェクト (2.6 節) として表現される.

ステップ 3 の例を示す. $((C_{m7}-F_7-B_{M7}^b)-E_{M7}^b)-(A_{m7}-D_7-G_7)$ のように解釈された和音進行は, 図4のようなケーデンス木として表現される. いかなる調の曲にも対応できるように, ケーデンスや構造特徴和音に含まれる音名はすべて調に対して相対化し階名で表す. そして, 実際のケーデンス木では key 属性の代わりに d 属性を設けて, 直上のケーデンスとの key の差を相対的に表す. 図中, 1つの台形が1つのケーデンスを表しており, 最上段太字が prom 属性 (構造特徴的な和音名) を, 中段 $d = N$ が d 属性を, 最下段が seq 属性を表す. 和音進行とケーデンス木の対応を上向きの矢印で示した. (2) 中の $d = 3$ は, (3) の key G に対する (2) の key B^b の音程差が半音3個分 (短三度) で

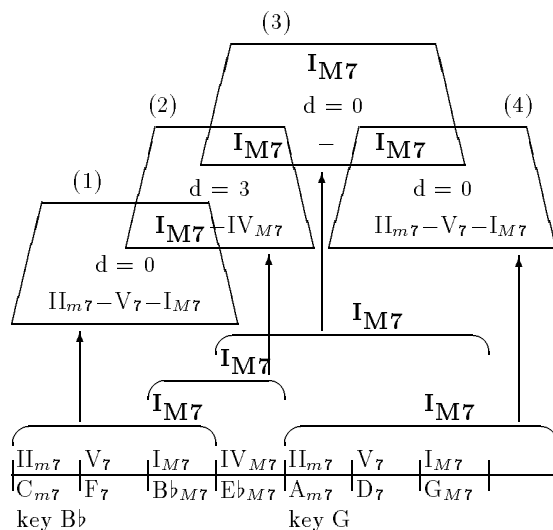


図4 和音進行の分析

あることを表わしている. $d = 0$ は, 直上のケーデンスと同じ key を持っていることを表わしている.

3.3 ユーザとのインタラクション

引き続き, 以下のようなステップで和声文脈エディタによる問合わせ作成と事例ベース推論を行う.

ステップ 5. ユーザは和声文脈エディタを用いて入力 (未知) 曲の和音進行をケーデンス木で表現する.

ステップ 6. 和声文脈エディタで作成したケーデンス木に対し, ボイシングを付加したい部分木を指定し, 問合わせを作成する.

ステップ 7. 問合わせのケーデンス木に最も類似したケーデンス木を事例ベース中から検索する.

ステップ 8. 検索結果のケーデンス木 (一般には複数) から, 問合わせで指定された部分木を抜き出し, その play 属性を取り出し, 最も滑らかなボイスリーディングを持つボイシングを選んで順次つなげて行く.

ステップ 5, 6 はユーザの意図を表現するプロセスであり, ステップ 7, 8 はシステムがユーザの意図を反映した解を生成するプロセスである. 実際にパーピーブンを使用する時は, この二つのプロセスを交互に繰り返すことでユーザとパーピーブンのインタラクションが実現される.

3.4 問合せの作成

ステップ 5 でユーザが作成する問合せのケーデンス木は, 図4のケーデンス木と全く同じ構造をしている. ステップ 6 では, 例えば図5のようなケーデンス木とそれに対するボイシング指示を作成する. 図中, ケーデンスを囲む2重線が「ケーデンス全体の中でこの部分

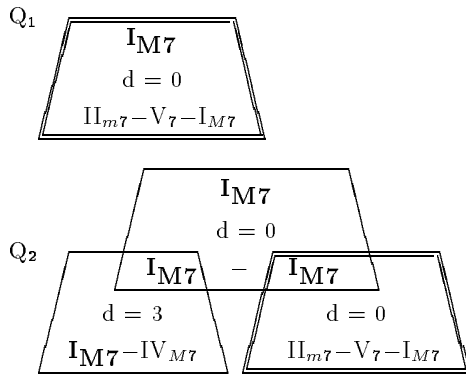


図5 2種類の問合せ例

にボイスिंगを付与せよ」という指示を表す。

3.5 相対的な類似度

ステップ7のため、ここで相対的な類似度という概念を提案する。

オブジェクト A, B, C がある時、A, B にだけ着目して「A, B は類似しているか」という問いは無意味であり、「A から見て、A, B と A, C のどちらがより類似しているか」という問いには意味がある。また、この問いに対して、A, B (あるいは A, C) の方がより類似している、同じだけ類似している、判定できない、の3通りの答えを出すのが自然である。例えば、p と p(l = a), p と p(m = b) を比べると、どちらがより類似しているかは判定できない。

パービーブンは、包摂関係が半順序であることを利用して、相対的な類似度を次のように定義する。まず判定のための論理式 RS を定義する。

$$\begin{aligned} RS(A,B,C) \\ \equiv \text{lub}(A,B) \neq \top \wedge \text{glb}(A,B) \neq \perp \\ \wedge (\text{lub}(A,B) \sqsubseteq \text{lub}(A,C) \\ \wedge \text{glb}(A,C) \sqsubseteq \text{glb}(A,B)) \end{aligned}$$

ここで lub は least upper bound (最小上界), glb は greatest lower bound (最大下界) の意味である (lub と glb も包摂関係を用いて標準的な方法で定義されている)。オブジェクト項 A, B, C に関して RS(A,B,C) が成り立つ時、A, B 間の方が A, C 間より類似していると定義する。定義より明らかのように、同一のオブジェクトが常に最類似である。相対的な類似度は次のように直感的に理解できる。オブジェクトをその属性から成る集合と見なした時、集合 A, B がより類似しているということは、A と B の積集合がより大きくなりかつ A と B の和集合がより小さくなるということに対応する。

我々の類似度に対して、lub の比較だけで相対的な類似度を計算する方式²⁰⁾ が提案されている。実際のパービーブンの処理では、lub 同士は等しくても glb 同士が異なる場合があるので、20) の方式では不十分である。

3.6 類似事例の検索

ステップ7を実現するために、パービーブンは先に定義した相対的な類似度 RS(A,B,C) を用いて最類似事例を検索する。今、事例ベース中に n 個の事例 (C₁, …, C_n) が存在し、ユーザが問合せ Q というオブジェクト項を提示したとする。ここで、見易さのために次の記法を導入する: A ≯ B ≡ ¬(A ⊆ B ∨ A ⊇ B ∨ A = B)。すると、

$$\begin{aligned} Q \text{ に最も類似した事例の集合} \\ \equiv \{C_i \mid \forall C_j \text{ RS}(Q,C_i,C_j) \\ \vee (\text{lub}(Q,C_i) \not\equiv \text{lub}(Q,C_j) \\ \wedge \text{glb}(Q,C_i) \not\equiv \text{glb}(Q,C_j)) \\ \vee (\text{lub}(Q,C_i) = \text{lub}(Q,C_j) = \top \\ \wedge \text{glb}(Q,C_i) = \text{glb}(Q,C_j) = \perp)\} \end{aligned}$$

これは、半順序の階層において上界の集合を計算することに等しい。上式から明らかなように、Q に最類似の事例は一般に複数個存在する。

例として、図5の Q₁ は図4の(1)と(4)に照合し、Q₂ は(4)のみに照合する。周囲の和声文脈も含めた検索を行うと解が絞りこまれることが分かる。

実際に和声文脈エディタ上で組み立てられる問合せは、可読性を考慮して、ボイスिंगを計算したい部分木と0以上の整数 N のペアという形をしている(図2中の context extent という欄)。これは、ボイスिंगしたい部分ケージンス木を中心に考えて、それより N 段上に昇ったケージンス木を和声文脈とせよという意味である。

3.7 ボイスिंगの選択

ステップ8で、最終的に1つのボイスिंगに絞り込む。ボイスिंगの候補が複数になるのは、最も類似した事例が複数個検索された場合と、検索された事例において play 属性の値である voicing オブジェクトの集合要素が複数個の場合である。今、隣接先行する楽曲部分のボイスिंगはすでに決定済みとし、その最後の和音を T とする。T に対して後続するボイスिंग列の候補が n 個あるとし、それらボイスिंग列の最初の和音を B_i (i = 1, …, n) とする。パービーブンは T と B_i のボイスリーディング(2.7節)に対して、包摂関係に関する最小のボイスリーディングを探し出し、それを解とする。

4. 関連研究

4.1 他の事例ベース推論システムとの比較

ハービー君⁸⁾はパービーブンの前身となった和音の雰囲気複製する事例ベースシステムである。ハービー君における事例（再構成の最小単位）と文脈は、それぞれ和音とケーデンスである。ハービー君では、ケーデンス間に特に関係を付けなかったため、ケーデンス間のボイスリーディングに不自然な部分が多少残っている。また、パービーブンではケーデンス木をユーザが編集することでユーザの意図をシステムに伝えることが可能となったが、ハービー君はそのような仕組みを提供していなかった。これより、パービーブンでは、ハービー君に比べて楽曲全体を通して一貫性のある演奏が得られる。

EMI⁵⁾は、元の多声楽曲の雰囲気複製して新しい多声楽曲を自動生成するシステムである。EMIにおける事例 (signature と呼ばれる) は 1 小節である。文脈は、Cope が提案した和声分析理論 (SPEAC*) に基づいて認識される。類似事例の検索は、各 signature に付与された特徴量ベクトルどうしのアドホックな比較で実現する。解を生成する時は、SPEAC による分析結果と拡張遷移網を用いて signature を再組合せする。EMI では、ユーザとのインタラクションを行わずに膨大な数の楽曲を生成させ、ユーザは実際にそれらを試聴して嗜好に合った楽曲を選択するので、効率が非常に悪い。パービーブンでは、ユーザの意図をインタラクティブにツールに伝えて解を生成するので、効率良く希望する解が得られる。

4.2 セッションシステムとの比較

一般に、セッションシステムにおける処理は相手の演奏を知覚し、理解し、応答を生成するという 3 段階から成り²⁾、応答の生成を実時間で行うか否かという相違点を除けば、パービーブンと共通点が多い。

多くのセッションシステムにおいて、知覚プロセスは相手の演奏の幾つかの特徴（音の大きさ、音の高さの範囲、和音）に着目し、それらを定量化し、特徴量ベクトルに変換することであり、理解は特徴量空間において適当な距離尺度を用いて特定のパターン（相手の意図）を認識することであり、生成は理解で得られた特定のパターンを元に適切な演奏を作曲編曲することである。

セッションシステムでは、元の入力データを一度特徴量ベクトルに変換してから認識するが、パービーブンでは元の入力データ（和音進行）に対し直接的に認識（類

似検索）を行う。特徴量ベクトルからは元の入力データが復元できないということから分かるように、特徴量ベクトルに変換するということは情報の捨象を意味する。これより、音楽の評価基準はオープンなため捨象した情報が将来必要になることがあり、柔軟な処理の妨げになる、特徴量空間における何らかの意図を持つ処理が元の入力データにおいて同じ意図を持つ保証がないという欠点につながる。

例えば、3.5 節で提案した相対的類似度を改良して属性に関する重み付けを行う場合を考える。ここで採用した DOOD やケーデンス木による表現は、復元できるという意味で入力データの情報を完全に保存しているので、音楽知識の表現方式を一切変更する必要はなく、ユーザの意図に沿った重み付けを入力データに直接適用することができ、上記の欠点を回避できる。

5. おわりに

実際にパービーブンを試用して、ユーザ意図を表現するのにケーデンス木が有効であること、多様で一貫した曲調の楽曲を効率良く生成できることを確認した。

ケーデンス木によるグルーピングが GTTM の TSR と共通点を持つということは、相対的な類似度等を含むパービーブンの処理が音楽家の直観に合致しているということを示唆している。しかし、GTTM は多角的な楽曲分析体系であり、TSR 以外の分析手法も含まれている。実際、隣接したケーデンス間の関係のみでは、すべての明示的、暗黙的な制約が記述できるわけではない。

DOOD に基づく音楽知識表現は入力データの情報を完全に保存しているので、本表現法の上で GTTM をはじめとする音楽理論の援用をさらに進めるのは比較的容易であると考えられる。そして、さらに効率的なインタフェースを提供し操作を簡便にすることとユーザの意図をより高い自由度をもって表現することの両立、及び処理対象を和音進行から多声の旋律（ジャズソロピアノ）に広げることを目指したい。

参考文献

- 1) Ames, C.: *Protocol: Motivation, Design, and Production of a Composition for Solo Piano, Machine Models of Music* (Schwanauer, S. M. and Levitt, D. A.(eds.)), The MIT Press, pp. 357-382 (1993).
- 2) 青野裕司: ジャムセッションシステム, コンピュータと音楽の世界 (長嶋洋一, 橋本周司, 平賀譲, 平田圭二 (編)), 共立出版, pp. 283-305 (1998).
- 3) Balaban, M., Ebcioğlu, K. and Laske, O.(eds.): *Understanding Music with AI*, The

* SPEAC とは, Statement, Preparation, Extension, Antecedent, Consequent の頭文字である。

- MIT Press (1992).
- 4) Boulanger, R.(ed.): *The Csound Book*, The MIT Press (2000).
 - 5) Cope, D.: *Experiments in Musical Intelligence*, A-R Editions, Inc. (1996).
 - 6) Cope, D.: The Composer's Underscoring Environment: CUE, *Computer Music J.*, Vol. 21, No. 3, pp. 20-37 (1997).
 - 7) Dannenberg, R. B.: Music Representation Issues, Techniques, and Systems, *Computer Music J.*, Vol. 17, No. 3, pp. 20-30 (1993).
 - 8) 後藤真孝, 平田圭二: ハービー君: 演繹オブジェクト指向に基づいてジャズらしいコードにリハーモナイズするシステム, 情報処理学会研究報告, Vol. 96, No. 75, pp. 33-38 (1996). 96-MUS-16.
 - 9) Hirata, K.: Representation of Jazz Piano Knowledge Using a Deductive Object-Oriented Approach, *Proc. of Int'l Computer Music Conf.*, ICMA, pp. 244-247 (1996).
 - 10) 平田圭二: コンピュータ上での音楽知識の表現, コンピュータと音楽の世界(長嶋洋一, 橋本周司, 平賀譲, 平田圭二(編)), 共立出版, pp. 163-181 (1998).
 - 11) Hirata, K. and Aoyagi, T.: Musically Intelligent Agent for Composition and Interactive Performance, *Proc. of Int'l Computer Music Conf.*, ICMA, pp. 167-170 (1999).
 - 12) 平田圭二, 青柳龍也: パービーブン: 誰でもどこでもインタラクティブに使える知的ジャズ和音生成システム, 情報処理学会研究報告, Vol. 99, No. 68, pp. 7-12 (1999). 99-MUS-31.
 - 13) 平田圭二, 平賀瑠美: 演繹オブジェクト指向データベースを用いた演奏生成アプリケーションオントロジー, 人工知能学会誌, Vol. 14, No. 6, pp. 1033-1040 (1999).
 - 14) Johnson-Laird, P. N.: Jazz Improvization: A Theory at the Computational Level, *Representing Musical Structure* (Howell, P. et al.(eds.)), Academic Press, pp. 291-325 (1991).
 - 15) Kifer, M.: Deductive and Object Data Languages: A Quest for Integration, *Proc. of 4th Int'l Conf. on Deductive and Object-Oriented Databases* (1995). Also in LNCS 1013.
 - 16) 小暮潔: 素性構造 (1) 及び (2), 人工知能学会誌, Vol. 8, No. 2&3, pp. 184-191&305-311 (1993).
 - 17) Kolodner, J.: *Case-Based Reasoning*, Morgan Kaufmann (1993).
 - 18) KLIC 協会ホームページ: <http://www.klic.org/index.ja.html> (1998).
 - 19) Lerdahl, F. and Jackendoff, R.: *A Generative Theory of Tonal Music*, The MIT Press (1983).
 - 20) Plaza, E.: Cases as Terms: A Feature Term Approach to the Structured Representation of Cases, *Lecture Notes in Artificial Intelligence Vol. 1010*, Springer-Verlag, pp.265-276 (1995). Proc. of ICCBR-95.
 - 21) Rowe, R.: *Interactive Music System*, The MIT Press (1993).
 - 22) Rowe, R. and Garton, B.: Editor's Note: Putting Max in Perspective, *Computer Music J.*, Vol. 17, No. 2, pp. 3-11 (1993).
 - 23) Sadie, S.(ed.): *The New Grove Dictionary of Music and Musicians*, Macmillan Publishers Limited (1980).
 - 24) Wiggins, G., Miranda, E., Smail, A. and Harris, M.: A Framework for the Evaluation of Music Representation Systems, *Computer Music J.*, Vol. 17, No. 3, pp. 31-42 (1993).
 - 25) Winkler, T.: *Composing Interactive Music*, The MIT Press (1998).
 - 26) Xenakis, I.: *Formalized Music*, Pendragon Press, pp. 329-334 (1992). Appendix III: The New UPIC System.
 - 27) 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, pp. 3-18 (1992).

(平成12年4月3日受付)

(平成13年1月11日採録)

平田 圭二 (正会員)

昭和 33 年生. 1987 年 東京大学情報工学専門課程博士課程修了. 工学博士. 同年 NTT 基礎研究所. 1990 年から 1993 年 (財) 新世代コンピュータ技術開発機構 (ICOT). 1999 年より現職. 音楽情報処理及び並行論理型プログラミングの研究に従事. 日本ソフトウェア科学会会員, Int'l Computer Music Assoc. 理事, KLIC 協会理事

青柳 龍也 (正会員)

昭和 35 年生. 1988 年 東京大学情報工学専門課程博士課程修了. 同年電気通信大学. 音楽情報処理及び CAI の研究に従事. 1998 年より現職.