

ハーヒーフン: 2 段階演奏表情付け法による インクリメンタルな演奏生成システム

平田 圭二
NTT コミュニケーション科学基礎研究所
hirata@brl.ntt.co.jp

平賀 瑠美
文教大学
rhiraga@shonan.bunkyo.ac.jp

あらまし 本稿では、現在開発中の音楽システム「ハーヒーフン」の設計、動作機構、実装等について述べる。従来の演奏表情付けシステムの欠点は、インタラクティブでインクリメンタルにユーザの指示を受け付けて、生成された演奏を部分的に精錬し改良できない点、及び生成された演奏がユーザの意図を適切に反映しない点である。このため、我々は 2 段階演奏表情付け法を提案する。その第 1 段階では、自然言語で記述されたユーザ指示を音楽的に重要な音の発音時刻、音長、音量の変化分に変換する。第 2 段階では、音楽的に重要な音の周囲の音にその変化分を伝搬させる。

キーワード 音楽、演奏表情付け、演奏生成、知識表現、音楽理論、事例に基づく推論

Ha-Hi-Hun: Incremental performance synthesis system based on 2-stage performance rendering method

Keiji Hirata
NTT Communication Science Laboratories

Rumi Hiraga
Bunkyo University

Abstract This paper presents the design principle, the system organization and the implementation of a musical system “Ha-Hi-Hun” that we are developing. We believe that the next-generation performance rendering system should be able to refine and improve a generated performance interactively, incrementally and locally through direct instructions in the natural language of a musician. In addition, the generated performance must reflect the musician’s intention properly. For these purposes, we propose a new framework called two-stage performance rendering. The first stage translates a musician’s instruction in natural language into the deviations of the onset time, duration and amplitude of structurally important notes and second stage spreads the deviations over surrounding notes.

key words Music, Performance Rendering, Performance Synthesis, Knowledge Representation, Music Theory, Case-Based Reasoning

1 はじめに

我々は、これまで演繹オブジェクト指向データベース (Deductive Object-Oriented Database, DOOD¹[17]) の枠組を用いた音楽知²の表現と、事例に基づく推論 [11] を採用した一連の音楽システムを研究開発して来た。これら音楽システムは、一般にハービー君族と呼ばれる (図 1)。

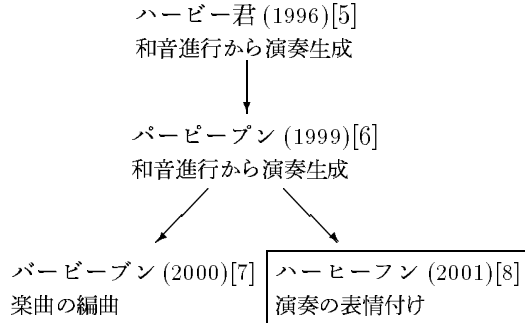


図 1: ハービー君族の系譜

本稿では、演奏生成 (Performance Rendering, PR) システム「ハーヒーフ」の設計、動作機構、実装等について述べる。

まず、ピアノ教室における教師と生徒のレッスンを考える。教師は生徒の演奏をより良くするために、適宜「ここをもっと情熱的に」、「このフレーズをさらにもっと明るく」というような直接的な指示を生徒に与える。理想的な場合、生徒は教師から与えられた指示通りに演奏を変化させる。そして教師は、修正した演奏を聴き、必要に応じて次の指示を出す。ユーザにとって使いやすい PR システムは、このピアノレッスンにおける生徒のような機能やユーザインタフェースを持つべきである。しかし、従来の PR システムは、生成された演奏をユーザの直接的な指示によってインクリメンタルかつインタラクティブに精錬し改良することができない。以下、本稿ではピアノ曲を演奏する PR システムに限定して議論を進める。

ユーザが PR システムに自由に指示を出して、ユーザの意図通りに制御するためには、楽曲の一部だけを修正する指示が出せて、しかも全体として不自然な演奏を生成しないことが必要である。そのため、PR システムは、(a) ユーザの指示を適切に解釈すること、(b) 生成された演奏全体が音楽的な一貫性を維持することの 2 点を満足すべきであると考えられる。

上記 (a) に関して、自然言語で与えられた指示は、通常、主観的で曖昧であり、時間と共に変化する。(a) を満足するためには、システムはカスタマイズ可能あるいは個人の振る舞いに特化可能であり、文脈に敏感な振る舞いができなければならない。

¹DOOD という用語はもともと国際会議や研究分野の名称として用いられていたが、本稿では知識表現手法の名称として用いる。

²音楽知とは、計算機で生成、変換、解析することを前提とした、音楽全般に関する体系的な知識や知性の総体のこと [1]。

(b) に関して、例えば、ユーザが PR システムに「楽曲中の音 Q を強く演奏せよ」という指示を与えたとする。PR システムが、楽曲中の Q の役割と Q の周囲の音との関係に考慮せず、単純に Q の音量だけを増加させたとしても、生成される演奏は不自然なものになってしまうだろう。そこで、生成された演奏が自然で一貫性を持って聴こえるようにするためには、システムは Q だけでなくその周囲の音に対しても音量を大きくしたり小さくしたり (デュナーミク³)、音の発音時刻をずらしたりする (アゴギク⁴) 必要がある。

これら (a)、(b) の条件を満たすために、本論文では 2 段階演奏表情付け法を提案する。本稿の構成は次の通りである。続く第 2 章では従来システムとその問題点を挙げ、第 3 章では楽曲の構造や演奏表情に関する情報を表現するためのデータ構造を導入し、第 4 章では 2 段階演奏表情付け法を例を交えて説明し、第 5 章では今回作成したプロトタイプシステムの動作について述べ、最後に第 6 章でまとめと今後の課題について述べる。

2 従来システムと問題点

従来の PR システム [4, 2, 16, 15, 14, 9] に楽譜を与えると、規則、数式、事例等を用いて楽曲中に含まれる全ての音のアゴギクとデュナーミクを一度にバッチ的に算出する。本稿では、これら演奏を生成するために用いられる規則、数式、事例等を演奏知識と呼ぶ。演奏知識は、人間による模範演奏から抽出、獲得されたり、音楽学における知見からも得られる。

従来システムには次の 2 つの問題点がある。

1 つめは局所の変更が困難なことである。ユーザが入力として与えた楽曲のある一部分 (例えば特定の 1 つの音や特定の 1 小節) の演奏だけを変更したいとする。従来システムにおいて音楽学の知見をもとにルールや数式を構成した場合、どのルールや数式がどの程度出力に寄与しているかが不明確である。演奏を制御するルールや数式を模範演奏から抽出した場合も、どの模範演奏を入れ換えるとルールや数式がどのように変化し、その結果生成される演奏がどのように変化するのかという因果関係が不明確である。従って、どの模範演奏を入れ換えるべきなのか分からない。よって、従来システムは、出力を試聴してその結果をフィードバックしてユーザの嗜好に合わない一部分だけを修正するというサイクルを実現することが大変困難である。

2 つめは、ピアノレッスンにおける生徒のように、ユーザからの主観的かつ定性的な指示を解釈できないという

³dynamics. 音量の強弱変化による演奏表情付けを意味する。強弱法とも言う。

⁴agogics. 演奏の時、厳格なテンポやリズムに微妙な変化を付けて精彩豊かにする方法。速度法とも言う。

ことである。発想記号⁵や発想標語⁶を含むそのような指示は、通常、ユーザ毎に異なって解釈される。またユーザ指示には、明るく、さらに明るく、もっとそれ以上明るく等の変更の程度を表現する言葉も多く含まれるが、従来システムは、程度の指示に対応することも大変困難であった。この問題点は、普遍的な演奏知識だけでは、ユーザが望む演奏を生成するのは大変困難であることを示唆している。

3 演奏表情を表現するデータ構造

3.1 DOOD による多声旋律の表現

DOOD では、対象をオブジェクトとその属性の集合から成るオブジェクト項として表現する。本稿では、オブジェクト項を $o(\dots, l = v, \dots)$ と記述する。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、基本項 p, q を比較し、 o_2 の全ての属性 l_n について o_1 のそれが全て具体的なならば、 o_1 の方が具体的 (あるいは o_2 の方が抽象的) という条件で決められる。属性の少ないオブジェクト項は包摂関係に関してより抽象的である。

最小上界 (least upper bound, lub) の定義: オブジェクト項 x, y が与えられた時、 x と y の最小上界とは

$\min(\{z | x \sqsubseteq z \wedge y \sqsubseteq z\})$ であり、 $\text{lub}(a, b)$ と書く。□

lub の直観的な意味は、ある 2 つのオブジェクト項の最大の共通部分であり、積集合を計算するイメージである。

3.2 多声旋律の表現

ここでは、一般のピアノ譜程度の楽譜を考える。旋律や和音を構成する複数の音が楽曲の一部あるいは全体を構成する。本稿ではそのような楽曲の一部あるいは全体を多声旋律と呼ぶ。

本節では、DOOD に基づいて多声旋律を表現するためのオブジェクト項を導入する。オブジェクト項の抽象化 / 具体化が、多声旋律の単純化 / 複雑化にうまく対応するようにオブジェクト項を設計しなければならない。

⁵アクセント、スラー、スタッカートなどアーティキュレーションを指示する記号と、 $f, p, <, >$ など記号をもって示されるデューナミック記号のこと。

⁶appassionata, dolce, espressivo のように言葉で説明されるもの。

多声旋律は様々な音楽的な構造を持っているが、ハービーフンではその内、グルーピング構造と時間構造を取り扱う。グルーピング構造は GTTM[12] の Time-Span Reduction (TS 簡約) に基づく。時間構造は、音長を分析合成するために新たに提案した音楽理論に基づく。

多声旋律のグルーピング構造 まず、TS 簡約とは、ボトムアップに楽曲構造的に重要な音 (和音) と隣接する重要でない音 (和音) をまとめて (グルーピングして) いくことである、その様子は Time-Span Tree (TS 木) として表現される (図 2 上半分)。図中 $e1 \sim e6$ はイベントである。ここでは簡単のため TS 木は二進木とし、重要な

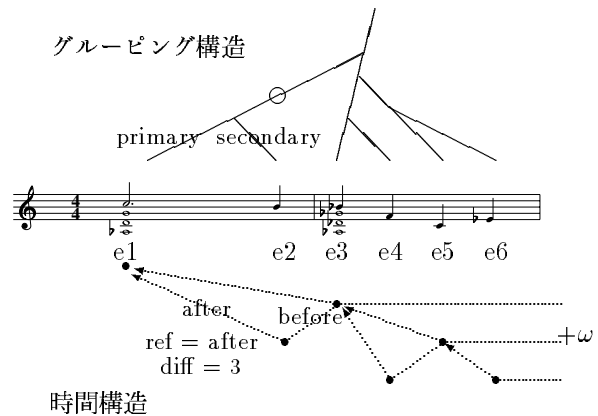


図 2: グルーピング構造と時間構造の例

音 (枝) を primary, そうでない音 (枝) を secondary と呼ぶ。グルーピングされた primary と secondary はある 1 つの音 (和音) で代表され、それを構造特徴音 (salient あるいは prominent note) と呼ぶ。図中、○ の付いた枝を実際に event オブジェクト項と chord オブジェクト項で表現すると下のようになる。

```
event(hd = e1,
      at = t1,
      primary = event(hd = e1,
                      at = t1),
      secondary = event(hd = e2,
                        at = t2))
```

```
e1 = chord(notes = {56,62,67,72},
           duration = 360,
           velocity = 60)
```

ここで、構造特徴音は hd 属性によって表され、notes 属性値は MIDI ノートナンバーの集合であり、duration 属性値は 1 小節 = 480 ticks とした時の値である。tn ($n=1 \sim 6$) はイベント e_n の発音時刻を表現するオブジェクト項であり、次に説明する。

多声旋律の時間構造 図 2 下半分は、同じ多声旋律の持つ時間構造を表している。まず和音 $e1$ の発音時刻 $t1$ を基準とする。次に音 $e3$ に着目すると、 $e3$ は $e1$ と $+\omega$ (無限大時刻) の間に発音し (順序に関する情報)、 $e1$ から 4 拍めの時刻に発音する (時刻に関する定量的な情

報). e2 の発音時刻 t2 は下の temp オブジェクト項で表現される.

```
temp(after = t1,
      before = t3,
      ref = after,
      diff = 3)
```

e2 は e1 (after) と e3 (before) の間に発音し, e2 にとってより重要な音は e1 なので (ref = after, 図中 ←..... で示す), e1 から 3 拍めの時刻で発音する (diff = 3).

ここで, グルーピング構造における hd 属性値と時間構造における ref 属性値は矛盾なく設定されている必要がある. 多声旋律のグルーピング構造と時間構造の詳細は文献 [7] を参照されたい.

3.3 演奏表情の表現

アゴーギクとデュナーミクを実現するための発音時刻のずれや音量変化を逸脱 (deviation) と総称する. chord オブジェクト項と dev オブジェクト項で逸脱を表現すると下のようになる.

```
e1 = chord(notes = {56,62,67,72},
           duration = 360,
           velocity = 60,
           deviation = d1) ←
```

```
d1 = dev(onset = 0.0,
         duration = 1.0,
         velocity = 1.0)
```

前節の chord オブジェクト項と比較すると, この chord オブジェクト項には新たに deviation 属性が付加されており (← の行), dev オブジェクト項が逸脱自身を表している. 発音時刻 (onset) のズレは, その音の音長に対する比の場合と 1 拍の長さに対する比の場合が考えられる. 第 5 章で述べるプロトタイプシステムでは, ユーザがどちらでも選択できるようになっている. 音長 (duration) のズレと音量 (velocity) のズレも同様に比で表されている.

構造特徴音を表すオブジェクト項に新しく deviation 属性を設けたことで, TS 簡約における各階層のグループ毎に逸脱を指定できるようになった.

4 2 段階演奏表情付け法

第 1 章では, ピアノレッスンにおいて教師が生徒に指示を出す様子を検討した. ここで我々は, 演奏表情付けの指示は構造特徴音に対して発行されるという仮説を立てる. つまり, 教師が生徒に「この音に注意を払って演奏せよ」と指示した時の「この音」とは構造特徴音であるという意味である. 2 段階演奏表情付け法はこの仮説に基づいて設計されている.

4.1 構成

2 段階演奏表情付け法の全体構成を図 3 に示す. 第 1 段階は, ユーザ指示を構造特徴音の逸脱に変換する. 第 2 段階は, 第 1 段階で変更を加えた構造特徴音の周囲の音に逸脱を加える.

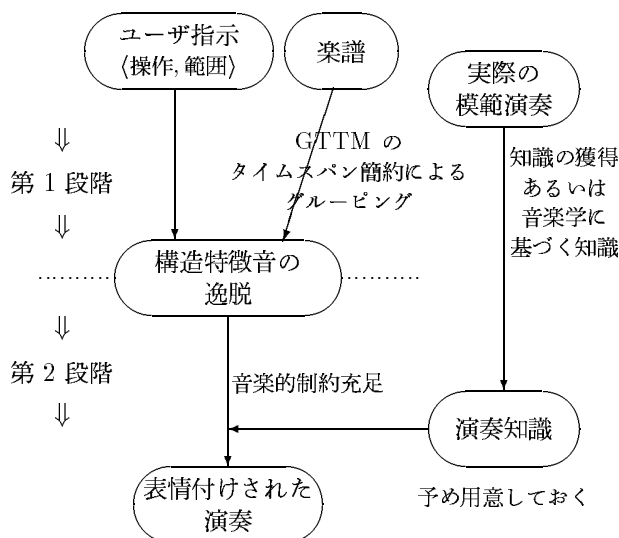


図 3: 2 段階演奏表情付け法

2 段階演奏表情付け法の入力は, ユーザから与えられる指示, 演奏すべき楽譜, 演奏知識を抽出, 獲得するための模範演奏である. ユーザ指示には, 実行したい操作とその操作を適用する譜面の範囲が含まれる. 操作の例としては, 速く, 明るく, より情熱的に等がある. 模範演奏は, 音楽学の知見に基づく規則や数式等で置き換えることも可能である. 出力は, 入力楽譜をユーザ指示に従って表情豊かに演奏したものである.

第 1 段階は, 自然言語で表現されたユーザの主観的な指示と範囲から, 逸脱を加えるべき構造特徴音を特定し, ユーザ指示をその構造特徴音の逸脱に変換する (図 4). 逸脱を加えるべき構造特徴音は, ユーザ指示に含まれる

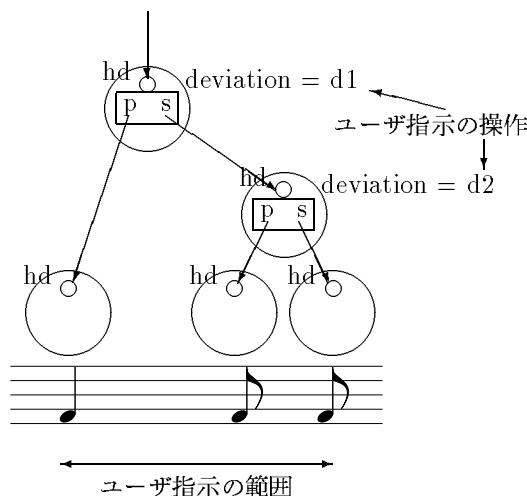


図 4: 構造特徴音への逸脱の設定

範囲をカバーする TS 木の部分木によって決まる. 図中,

小さい丸が hd 属性を表し、大きい丸が event オブジェクト項を表し、p は primary 属性を、s は secondary 属性を表す。deviation 属性だけ丸外に書き加えた。第 1 段階において、ユーザ指示の操作から各 deviation 属性値 d1, d2 を求める。この逸脱の値を求める方式は種々考えられる。例えば、予めシステムが提供するデフォルトで決めたり、事例に基づく推論から得たり、各ユーザや作業環境に対し学習を行うことで獲得したりする。従って、逸脱の値の決定は通常、ユーザの主観に大きく依存し恣意的である。

第 2 段階は、第 1 段階で構造特徴音に設定された逸脱を周囲の音に伝播させる(図 5)。その際、構造特徴

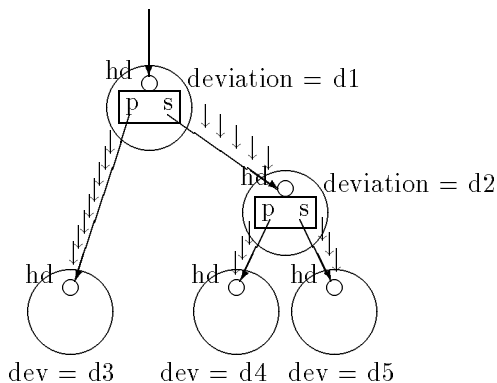


図 5: 音楽的制約充足

音 d1, d2 に設定された逸脱は不変であり、周囲の音の逸脱 d3~d5 のみが影響を受ける。図中、d1, d2 の逸脱がより下位の event オブジェクト項に伝播して行く様子を ↓ で示した。伝播のための演奏知識は、GTTM や暗意-実現モデル [13] のような音楽理論に基づいて模範演奏を分析することによって獲得したり [10]、先行研究が提案しているルール [4] を用いる。第 2 段階で用いられる演奏知識は、構造特徴音及びその周囲の音の逸脱に関する制約と見なすことができるので、第 2 段階の処理を音楽的制約充足と呼ぶ。

4.2 例 1: 生き生きと

ユーザ指示〈生き生きと、第 1 小節から第 2 小節〉を発行した時、システムが出力演奏を生成する様子の例を図 6 に示す。楽譜は予め TS 簡約に基づいて分析しておく(3.2 節)。図中、 α , β , γ が付いた破線の四角はそれぞれ木の一部を表し、 γ という木は α , β という枝を持つと解釈する。

図 6 の (a), (b), (c) は、ある楽曲の一部分に表れる音や和音の逸脱をピアノロール形式で表したものである(ただし音高は表されていない)。横軸は時間を、線分の太さは音量を表している。(a) は譜面通りの機械的な(ニュートラルな)演奏である。(b) は第 1 段階が終わった後の状態であり、 α や β に含まれる構造特徴音 p, r の逸脱だけが設定されている。(c) は、第 2 段階が終わった後の状態であり、(b) で設定された p の逸脱が q

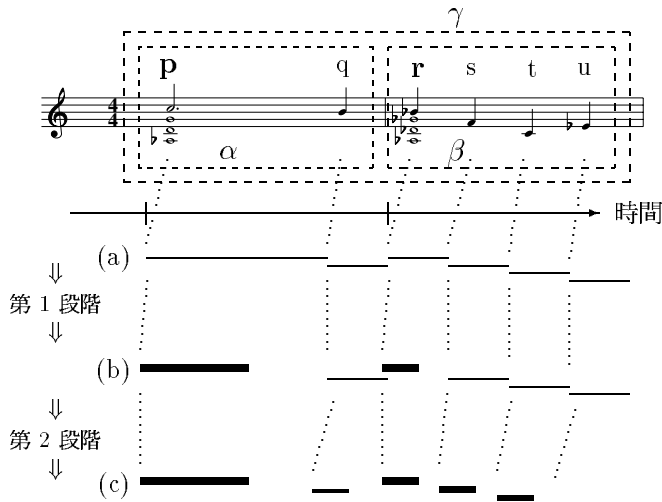


図 6: 例 1: 生き生きと

に伝播し、同様に r の逸脱が s, t, u に伝播する。

本例では『生き生きと』という指示なので、(b) において p, r の発音時刻を、それぞれの音長に対する 10% 分だけ前方にずらし、音長を 40% 短縮した。この 10% や 40% という値は恣意的であり、ユーザや状況が異なる場合は異なる値が選択される。音楽的制約充足の際、様々な演奏知識を用いることができるが、この例では簡単のため、音長に関しては $y = 0.6$ という定数関数を用い音長を一律に 60% に減少させた。発音時刻と音量に関しては図 7 のような一次式に従って変化させたので、各音の発音時刻のズレは音の場所に正比例して増加し、逆に音量は場所に正比例して減少する。

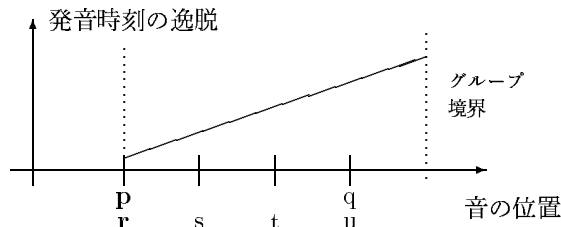


図 7: 音楽的制約充足に用いられる関数の例

もちろん、演奏知識に関する音楽的制約は上述のように単純なものに限定する必要はなく、さらに複雑な数式やアルゴリズムを用いることもできる。

4.3 例 2: 優雅に

図 8 は、例 1 と同じ譜面に対して〈優雅に、第 1 小節から第 2 小節〉という指示を発行した場合を示している。ここで (a), (b), (c) は前節と同様の意味である。優雅な演奏を得るためレガート奏法を適用する。レガートとは音と音の間に切れ目を感じさせないように演奏することなので、各音の音長を全体的により長めにし、さらに、ゆったりした感じを出すために発音時刻を若干後方にずらす。図中 (b) では、そのような逸脱が構造特徴音 p, r に設定されている。続く第 2 段階では、音量と音長

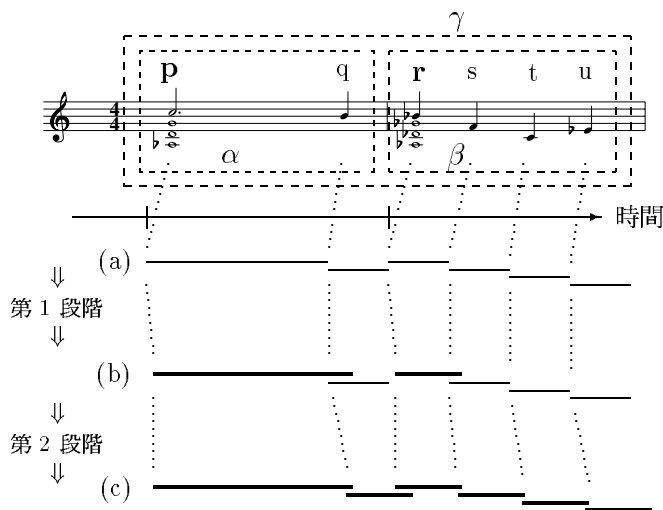


図 8: 例 2: 優雅に

に関して p と r の逸脱を一定の割合で q, s, t, u に伝搬させ、発音時刻の伝搬に関しては図 7 の関数を用いた。

4.4 特長

2 段階演奏表情付け法は文献 [3] における楽曲の構造的分解から大きな示唆を受けている。この構造的分解はユーザの主観に大きく依存していたのに対し、2 段階演奏表情付け法は GTTM の TS 簡約に基づいているので、構造特徴音という概念が使えるようになり、ユーザは出力の演奏を意図通りに制御できるようになった。

2 段階演奏表情付け法の特長を以下にまとめる。

- 入力楽曲をグルーピング構造と時間構造に関して分析するので、ユーザは PR システムに楽譜の一部に対する局所的な指示を発行することができるようになった。局所的修正が可能になったので、インタラクティブかつインクリメンタルな動作が実現する。
- ユーザ毎に異なる主観的な知識と、事例から獲得されたり音楽学から得られるような普遍的な知識を分離して管理、適用している。前者の知識は第 1 段階で使用され、ユーザの嗜好や個性に柔軟かつ適切に対応できるようになった。後者の知識は第 2 段階で使用され、音楽的に自然で一貫性のある演奏を実現できるようになった。また、この分離により知識管理が容易になった。

5 プロトタイプシステム

我々は 2 段階演奏表情付け法を採用したプロトタイプシステムを実装した。本プロトタイプシステムは、グルーピングエディタ [7] と PR エンジンから成り、音楽ジャンルとしてはソロピアノ曲を対象としている。グルーピングエディタは簡潔な GUI を持ち、入力楽曲や事例模範演奏のグルーピング構造を操作、指定する。簡単のため、現在の PR エンジンの第 2 段階の音楽的制約はシ

ステム組込みとし、システム稼働中は変化しないこととした。

event オブジェクトは再帰的な構造をしているので、事例の粒度は単和音から楽曲全体まで変化させることができる。本プロトタイプシステムでは、簡単のため、1 つの和音でカバーされる範囲を 1 つの事例とした。

5.1 lub に基づく逸脱の転写

本プロトタイプシステムの第 1 段階では、ユーザ指示として模倣すべき事例を与える。その事例はオブジェクト項として表現されている (3.2 節, 3.3 節)。第 1 段階では、模倣すべき事例から逸脱を取り出して、それを入力楽譜の構造特徴音に転写する。事例中の逸脱を取り出す構造特徴音と転写すべき入力楽譜中の構造特徴音を特定するために、事例を表現するオブジェクト項と入力楽譜を表現するオブジェクト項との間の lub (3.1 節) を計算する (図 9)。

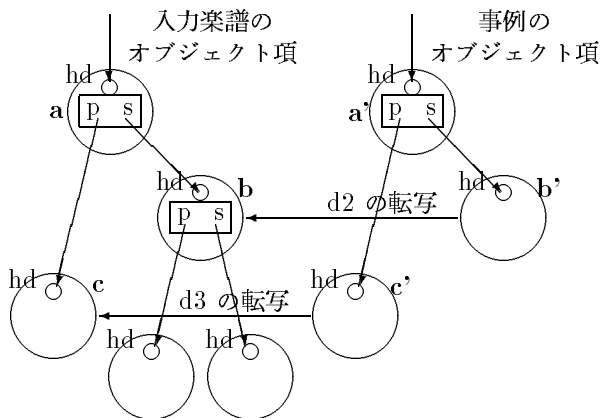


図 9: lub に基づく逸脱の転写

算することであり、図中、 \leftarrow で結ばれたノード同士が張る部分木がその共通部分である (a, b, c から成る部分木と、 a', b', c' から成る部分木)。事例の共通部分の末端ノードに存在する構造特徴音 b', c' から逸脱を取り出し、入力楽譜の対応するノード b, c に転写する。第 2 段階では図 7 のような一次関数を用いて音楽的制約充足を行なう。

5.2 試聴の結果

プロトタイプシステムが生成した演奏を試聴して、事例演奏の持っている雰囲気が入力楽譜の演奏で模倣されているかどうかを確認した。本節ではその試聴結果を簡単に報告する。

サンプル 1 では、入力曲に “All of Me” を、事例曲に “Autumn Leaves” を用いた。事例の Autumn Leaves は標準的なジャズピアノトリオの演奏であり、その中のピアノの右手で演奏されるテーマ部分を事例として用いた。生成された All of Me の演奏は、各フレーズの頭のタイミングが若干遅れ気味になり、フレーズの後半に行くに従って速くなり音量が下がった。これは事例に含ま

れているジャズらしさを出すための逸脱の特長と同じである。

サンプル 2 では、入力曲に “Autumn Leaves” を、事例曲に “J. Strauss の曲” を用いた。事例は標準的な三拍子のピアノ曲の演奏であり、サンプル 1 同様、右手で演奏されるテーマ部分のみを事例として用いた。生成された Autumn Leaves の演奏は、フレーズの中央部で音量が下がり、全体的にスタッカート気味になっている。この場合、軽快なワルツ演奏という事例の持つ特徴を模倣していると考えられる。

現時点では試聴の主観的判断しか述べることができないが、生成されたいずれの演奏も意図通り事例曲の雰囲気や反映していると言えよう。

さらに、事例曲と生成曲の TS 木の secondary 枝 (部分木) を 1 段階刈り取ったものと 2 段階刈り取ったものの二種類を用意した。より多段の secondary 枝を削除し TS 木が小さくなると、より重要な primary 枝 (構造特徴音) だけが残ることになる。そして、secondary 枝を削除したものと削除する前のものと聴き比べた。この試聴実験により、入力楽曲と事例曲の TS 木が異なっているような場合、事例曲のどの primary 枝の逸脱が入力楽曲のどの primary 枝に転写されているかがわかる。この結果、第 1 段階において、事例曲の構造特徴音の逸脱を出力に転写する過程がある程度うまく動作していることが確認できた。

6 まとめに代えて

本論文は、高い制御性を持つ演奏表情付けの新しい枠組を提示した。この枠組はモジュラな構造を持っているので、出力の質を改善するためには、第 1 段階の構造特徴音への逸脱設定のアルゴリズムや第 2 段階の演奏知識や音楽的制約の精錬化や改良、学習機能など種々のモジュールを個別に精査、改良すれば良い。前述したプロトタイプシステム以外にも、これらモジュールの様々な組み合わせを考えることができる。

次に今後のシステム改良について述べる。大きく、知識表現手法の改良と 2 段階演奏表情付けアルゴリズムの改良に分けられる。知識表現手法の改良については文献 [7] を参照されたい。2 段階演奏表情付け法の今後の課題には以下のようなものがある：

- 模範演奏から演奏知識を抽出する作業は非常に手間がかかる。従って、事例獲得や楽譜分析を効率化するためのツールが必要である。
- ユーザインタフェースの問題。例えば、長い曲にも対応できるような作業環境、的確にユーザ指示が実行できるような GUI、生成した演奏内容の適切な可視化等である。
- グループは階層構造を作るので、通常、1 つの音は複数のグループに属する。しかし、演奏生成においてある音の逸脱を計算する際、どのグループ階層がその音にどの程度影響を及ぼすのかは一般に自

明ではない。従って、どのグループ階層がどの程度影響を及ぼすのかを明らかにし、そのような演奏生成モデルを構築することが必要である。

- プロトタイプシステムの第 1 段階では、ユーザが模倣させたい事例を与えてそこから逸脱を転写していた。事例が多数蓄積されれば、データマイニングや学習などにより法則性や規則性を見出すことが可能になろう。特に、獲得した知識が第 1 段階、第 2 段階のどちらで用いられるべきものかを識別する手法を開発しなければならない。

最後に、本システムをどのように評価すべきかということも大きな課題である。少なくとも次の 4 通りの観点から評価を行う必要があると考えている。(1) 観察に基づく一般ユーザに対するシステムの全体評価、(2) ユーザ意図の把握と反映に関する評価、(3) 開発当事者がシステムを改良するための評価、(4) 音楽知表現手法や事例に基づく推論等の個々の要素技術に対する評価である。評価に関しては稿を改めて報告したいと思う。

謝辞: ハービーフンを Java で実装するにあたり、青柳龍也助教授 (津田塾大学) より Java プログラミングについて多くを教えて頂きました。

参考文献

- [1] 青柳, 小坂, 平田, 堀内, 後藤, 引地, 平野, 松島 (訳), コンピュータ音楽 - 歴史 アート テクノロジー, Curtis Roads 著, 東京電機大学出版局 (2001).
- [2] Arcos, J. L., de Mántaras, R. L., and Serra, X., “SaxEx: a case-based reasoning system for generating expressive musical performances”, *Proc. of ICMC*, pp.329-336, ICMA (1997).
- [3] Desain, P., and Honing, H., “Towards a Calculus for Expressive Timing in Music”, in *Music, Mind and Machine*, pp.173-214, Amsterdam: Thesis Publishers, (1992).
- [4] Friberg, A., “Generative Rules for Music Performance: A Formal Description of a Rule System”, *Computer Music Journal* (15)2:56-71, The MIT Press (1991).
- [5] 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 音楽情報科学研究会 研究報告 96-MUS-16, pp.33-38 (1996).
- [6] 平田圭二, 青柳龍也, パービーブ: ジャズ和音を生成する創作支援ツール, 情報処理学会論文誌, Vol.42, No.3 (2001).
- [7] 平田圭二, 青柳龍也, パービーブ: 音符レベルでユーザ意図を把握して編曲を行う事例ベースシステム, 情報処理学会 音楽情報科学研究会 研究報告 2000-MUS-37, pp.17-23 (2000).

- [8] Hirata, K., Hiraga, R. and Aoyagi, T., Next Generation Performance Rendering – Exploiting Controllability, *Proc. of ICMC*, pp.360–363, ICMA (2000).
- [9] 五十嵐滋, 小池宏幸, 水谷哲也, 音楽の構造的機能とそれに基づく演奏創造, 第 14 回日本人工知能学会 全国大会 (2000).
- [10] Ishikawa, O., Aono, Y., Katayose, H. and Inokuchi, S., Extraction of Musical Performance Rules Using a Modified Algorithm of Multiple Regression Analysis, *Proc. of ICMC*, pp.348–351, ICMA (2000).
- [11] Kolodner, J., *Case-Based Reasoning*, Morgan Kaufmann (1993).
- [12] Lerdahl, F. and Jackendoff, R., *A Generative Theory of Tonal Music*, The MIT Press (1983).
- [13] Narmour, E., *The Analysis and Cognition of Basic Melodic Structures*, The Univ. of Chicago Press (1990).
- [14] 野池賢二, 乾伸雄, 野瀬隆, 小谷善行, 演奏情報と楽譜情報の対からの演奏表情規則の獲得とその応用, 情報処理学会 音楽情報科学研究会 研究報告 97-MUS-26, pp.109–114 (1998).
- [15] 鈴木泰山, 徳永健伸, 田中穂積, 事例に基づく演奏表情の生成, 情報処理学会 論文誌 Vol.41 No.04 - 035 (2000).
- [16] Widmer, G., “Large-scale Induction of Expressive Performance Rules: First Quantitative Results”, *Proc. of ICMC*, pp.344–275, ICMA (2000).
- [17] 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).