

Title	延長的簡約を拡張したGTTMによる楽曲の構造解析
Author(s)	井田, 健太郎
Citation	
Issue Date	2002-03
Type	Thesis or Dissertation
Text version	author
URL	<a href="http://hdl.handle.net/10119/1536">http://hdl.handle.net/10119/1536</a>
Rights	
Description	Supervisor:東条 敏, 情報科学研究科, 修士

# 修士論文

## 延長的簡約を拡張した GTTM による楽曲の構造解析

指導教官 東条敏 教授

北陸先端科学技術大学院大学  
情報科学研究科情報処理学専攻

井田健太郎

2002年2月15日

## 要旨

人間の言語活動と音楽活動とは非常に密接に関係していると考えられる。人間の言語は初めお互いの意志を伝えるための手段として発声することから始まった。その後、記録するために言葉として書くようになった。音楽も初めは娯楽や意思の伝達手段として演奏することから始まった。その後、音楽も楽譜という手段を用いて記録するようになった。これらのことから、言語に関する研究は音楽に関する研究に応用することができるのではないだろうかという考えに至る。そのために本研究では音楽認知の初期段階として、楽曲の構造解析を試みる。

本研究では音楽の構造を分析するための理論として、Fred Lerdahl と Ray Jackendoff が 1983 年に提唱した Generative Theory of Tonal Music ( GTTM ) を用いる。GTTM は Schenkerian のウルザッツ概念に基づく構造主義的分析と、Chomsky の生成言語文法理論を理論基盤としている。楽曲を階層グループに分析し、簡約内部の緊張 - 弛緩の階層構造を生成文法的に記述している。この理論は、各々の分析における規則が箇条書きされているため、計算機上への実装に向いている。しかし問題点としては規則の扱いが定式化されていないため、一意性、最善性の問題がある。本研究ではこれら問題点の解決法として、ボイスリーディング、第一声部進行、並列計算の試みを提案する。

また、これら声部に関する理論を拡張した GTTM による構造分析の有効性を実験によって検証し、考察をおこなった。その結果、2 つの構造分析においては有効性が認められ、結果として十分な分析結果がえられた。しかしタイムスパン簡約に関しては、現時点では成果は得られていない。その理由は、GTTM の理論の中には和声に関する分析が欠如している。そのため、タイムスパン簡約をおこなう際に、和声に関する規則が実装できずにいるため精度が低くなってしまっている。よって今後和声分析に関して、GTTM 以外の理論を拡張する必要があると考えられる。

# 目次

<b>1</b>	<b>はじめに</b>	<b>1</b>
<b>2</b>	<b>GTTM 関連の研究</b>	<b>4</b>
2.1	GTTM と他の理論との比較研究 . . . . .	4
2.2	GTTM を実装するために本研究で必要と考えた理論の研究 . . . . .	8
2.2.1	直接 GTTM に関する研究 . . . . .	8
2.2.2	間接的に関係する研究 . . . . .	10
2.3	GTTM の応用に関する研究 . . . . .	13
<b>3</b>	<b>Generative Theory of Tonal Music</b>	<b>18</b>
3.1	GTTM の概要 . . . . .	18
3.1.1	グルーピング構造 . . . . .	19
3.1.2	拍節構造 . . . . .	23
3.1.3	タイムスパン簡約 . . . . .	26
3.1.4	延長的簡約 . . . . .	29
3.2	GTTM における諸問題 . . . . .	33
3.2.1	ポリフォニーへの拡張 . . . . .	33
3.2.2	用語定義の曖昧性 . . . . .	34
3.2.3	規則適用の非一貫性 . . . . .	35
3.2.4	構造，簡約間における解のフィードバック . . . . .	35
<b>4</b>	<b>声部分析を拡張した GTTM による分析法</b>	<b>37</b>
4.1	本研究における問題点の解決 . . . . .	37
4.1.1	改良した理論における楽曲の扱い . . . . .	38

4.1.2	ボイスリーディング . . . . .	38
4.1.3	ヒューリスティクス, 暗黙的パラメータの明示化 . . . . .	39
4.1.4	第一声部進行 . . . . .	40
4.1.5	並列計算の試み . . . . .	40
4.2	本研究が提案するシステム . . . . .	42
4.2.1	【システム1】ファイル形式変換 . . . . .	43
4.2.2	【システム2】音符情報の算出 . . . . .	45
4.2.3	【システム3】ボイスリーディング分析 . . . . .	48
4.2.4	【システム4】第一声部進行分析 . . . . .	48
4.2.5	【システム5】グルーピング構造分析 . . . . .	49
4.2.6	【システム6】拍節構造分析 . . . . .	53
4.2.7	【システム7】タイムスパン簡約分析 . . . . .	56
4.2.8	延長的簡約分析 . . . . .	59
5	実験 . . . . .	60
5.1	グルーピング構造 . . . . .	61
5.2	拍節構造 . . . . .	63
5.3	タイムスパン簡約 . . . . .	65
6	まとめと今後の課題 . . . . .	68
	謝辞 . . . . .	71
	参考文献 . . . . .	72

# 第 1 章

## はじめに

現在，計算機上で音楽を扱う様々な研究が行われている．自動伴奏や楽曲の検索などがその例である．これら研究の最も基礎となっているものに音楽の構造解析が挙げられる．楽曲の編曲や演奏の表情付けなどの研究も，構造を解析したうえでの処理である．ところがこれまでの研究では，楽譜を計算機上で扱うためのデータ構造や，実際に構造を解析するための理論に関する研究はあまり行われてこなかった．しかし，ここで自然言語の視点からこれらを見てみると，自動伴奏は楽曲の構造に対して行うものであり，この構造を求める事が自然言語における構造解析に対応していると考えられる．また，楽曲の検索についても構造解析を行うことによる検索エンジンが考えられる．これらのことから，システムが楽曲の持つ意味を理解して処理をするためには，音楽の理論を用いて楽曲を分析し，その結果得られた構造に基づいて処理を進めるのが最も合理的であると考えられる．このような音楽学の理論に，Generative Theory of Tonal Music (GTTM)[4]がある．

しかし現在，この理論は人間を相手に分析結果を説明することを前提としているため，規則が厳密に定義されていないなどの理由により規則定義の曖昧性や，規則適用の非一貫性などの問題が生じてしまう．これにより，規則の適用が一意に求まらないだけでなく，規則適用と解析結果が対応しない場合もある．これら解析における曖昧性は，人間が解析する際の主観性を処理するために GTTM の長所として挙げられている．しかし，人間ならば曖昧な処理も可能だが，計算機には困難である．そのため，曖昧性の処理は人間が解析する際には長所であるが，計算機上に実装を考えるとプログラムやアルゴリズムとして記述することが困難であるために問題点になってしまう．

本研究では、これら GTTM を実装することにおける問題点に対し、一般的な音楽理論としてボイスリーディング、第一声部進行の拡張を行った。これらを拡張することにより GTTM の一般性を損なうこと無く、曖昧性を処理するために不足している情報を補うことが期待できるからである。

また、もう1つの問題点として各構造、簡約における解のフィードバック問題が挙げられる。GTTM の理論は以下の2つの基本構造と2つの簡約のための規則から構成されている。

- 1 グループ構造
- 2 拍節構造
- 3 タイムスパン簡約
- 4 延長的簡約

実際の解析では、グループ構造と拍節構造の解析結果を用いてタイムスパン簡約、延長的簡約を行う。しかし、GTTM の規則の中には次ステップの解が最も安定するような解を選ぶ規則がある。例を挙げると、タイムスパン簡約の結果が最も安定するような解をグループ構造で求めると言うものが挙げられる。これについても、計算機上に実装を考えると処理の流れが一方向でないために問題となる。

本研究では、各構造、簡約の解析ステップにおいて考えられる解全てを並行して計算し、次のステップにおいて解を絞りこむ手法を用いる。これにより処理の一方向化が期待できるからである。

本研究では、以上の様な仕様で GTTM の実装を試みた。本稿では、2つの基本構造と2つの簡約の相互依存関係を明らかにし、全体のシステムのアルゴリズムを提案する。また曖昧性の処理や、解のフィードバックなど、GTTM を実装する際のこれらの問題点を明らかにし、解決のために拡張したボイスリーディングと第一声部進行について説明し、これらの効果を確認する。

本稿の構成は、まず次章で本研究の関連研究について紹介し、3章では本研究の基礎理論となっている Generative Theory of Tonal Music について説明し、この理論の問題点について明らかにする。4章では本手法における問題点の解決について説明し、新た

に導入した理論の説明をおこないシステムの構成について詳細を述べる．5章では本研究の提案手法の有効性を確認するために実験をおこない，結果について考察する．最後に6章でまとめと今後の課題について述べる．



## 第 2 章

# GTTM 関連の研究

GTTM が 1983 年に提唱されてからこれまでに様々な関連研究が行われてきた。本章ではこれらを【GTTM と他の理論との比較研究】、【GTTM を実装するために本研究で必要と考えた理論の研究】、【GTTM の応用に関する研究】の 3 種に分けて紹介する。

### 2.1 GTTM と他の理論との比較研究

音楽を解析するための理論は GTTM を含めいくつか存在する。

片寄，竹内による研究

片寄，竹内 [11] はこれらの理論を音楽的な視点，応用性の視点から比較している。これらをまとめたものが表 2.1, 2.2 である。これらの理論から本研究の目的である解析の自動化を考えると，Lerdhal&Jackendoff の GTTM が各構造を生成するための規則が簡潔書きされているために実装に最も向いていると考えられる。GTTM は Schenker の音楽解析理論と Chomsky の生成言語文法理論を基盤とし，自然言語の構文解析のように構文規則を用いて木構造を求める。この構文規則にあたる音楽解析規則が簡潔書きされているのである。しかし表 2.2 にあるように，一意性，最善性に問題がある。よって GTTM を用いて実装する場合にはこれらの問題点の解決もあわせて考えなければならない。

表 2.1: 各理論の音楽的な視点からの比較 (文献 [11] より参照)

理論	目的	手法・特徴	分析の内容と課題
Meyer	リズム構造の 認知的分析	詩脚法・拍節アク セントによる分析	アクセントを拍節的な強拍部とすると、認知 構造の分析でなく楽譜の分析になってしまう。
Narmour 村尾	認知構造の数 量化	「暗意-実現のプロ セス」の分析によ る非メトリカルア クセントの数量化	構造主義分析では切り取られるイディオスト ラクチャを反映している点はより認知的・数 量的公式化の変数に問題がある。分析対象を 旋律だけでなくカデンツやバス音・対位旋律 などを含めた解析が必要である。
Lerdhal, Jackendoff	音楽文法の生 成的記述	グループ、拍節構 造、タイムスパン 簡約、延長的簡約。 木構造	シェンカーのウルザッツ概念に基づく構造主 義的分析。2種類の聴取傾向をルール化。厳 格な階層グループに分析。簡約内部緊張-弛 緩の階層構造を生成文法的に記述。認知的に は分析の一義性が問題とされる。グループ化 と緊張-弛緩の関連が未整理。
保科	演奏解釈のた めの楽曲分析	グループ、フレー ズの分析。重心・ 頂点の明示化。	厳格な階層グループ構造に分析。グループ内 部の最強調部を重心(頂点)として明示。複合 グループをフレーズとして分析。グループ構 造と演奏変数との相関性を指唆。分析ルール が経験的。
竹内	演奏家のため の演奏解釈	重心や頂点に対応 した演奏変数の対 応。重心を Jack- endoff 理論より分 析。	アナクルーズとデジナンスに対応した、ディ ナーミクとアゴーギクの適用。重心や頂点は Jackendoff の延長的簡約の2重弛緩構造(2重 の左枝)より求めるが、階層的な拍節構造に 一致しない場合は拍節的に強拍(表拍)部位を 採用する。

表 2.2: 各理論の応用性の視点からの比較 (文献 [11] より参照)

理論	音楽表現/構造解析に対する視野	ストラクチャに対する考え方	自動演奏システムへの応用メリット	問題点
Narmour (Meyer)	認知視点からみた音楽構造にかかわる可能なかぎりの音楽的意図の解析.	イデオストラクチャ指向暗意-実現に基づいたたくさんの関係可能性を見る構造はツリー構造にはならない.	詩脚レベルに対応する表現ルールが分かれば,イデオストラクチャ表現ルールの一般化ができる.	複雑.離れた暗意-実現に関して演奏表現の関係がつかみにくい.メロディしか解析法が示されていない.
村尾	構成アクセントと認知演奏アクセントの分離.計算式に基づいた Narmour 詩脚法の単純化	イデオストラクチャ指向,構成アクセントの大きいところをクロージャ(グループ)スタート.	構成アクセントレベルに対応する表現ルールが分かれば,イデオストラクチャ表現ルールの一般化ができる.	構成アクセント計算式の正当性.代償を行なう部分かどうかの決定.メロディしか解析方法が示されていない.
Lerdhal, Jackendoff	一般的な音楽素養を持った聴取者の言語理解に相当する構造化	タイムスパン簡約:木構造(スタイルストラクチャ),延長的簡約:イデオストラクチャ	フレージング,拍節表現を規則的に表現するシステムの条件節を導く手法としての期待が持てる.	優先規則の扱いが定式化されていない.(一意性,最善性の問題)
保科	演奏者に演奏表現をより分かりやすく指導することを目指す.	階層はグループとフレーズ.グループ(フレーズ)の中のエネルギーポイント(重心・頂点)は一つ.(重心や頂点の分析は音楽経験的)	構造解釈が出来てしまった状態から演奏表現を生成する筋道が分かりやすい.	エネルギーポイント(重心)の決定の仕方がヒューリスティックで人間の主観処理を前提とする.

## 平賀による研究

また、平賀 [6] は音楽認知研究で扱われる楽譜レベルへの記号化、調性・拍節構造、グループへの分節、階層構造、類似性の認識などの課題について批判的見地から検討を行っている。

記号化に関しては音楽的な「意味」を生み出す要因はすべてシステム内で定義しなければならないと述べている。記号表現系では個々の記号自体は無意味(無定義語)であって、「意味」は記号どうしの関係(やその上の操作)として存在するのである。

調性・拍節構造については、音の時間パターンだけでなく、様々な要因間の相互作用が組み合せ論的に効いて、その調性も困難であると述べている。このことから、音楽の持つ曖昧性、構造の複雑さが明らかになり、どのような相互作用があるかを探ることにもつながる。また、逆に処理結果の「正解率」だけをアドホックに追究するのは、あまり意味がないことになると述べている。

グループへの分節に関しては、グループ化自体はごく自然な概念であるが、具体的には不整合が発生すると述べている。これは GTTM を例にとり示すと、GTTM にはグループを決定するために選好規則 (preference rules) を与えている。この選好規則を最大限満たすグループ構造が選択される。しかし、当然考えられる問題として、グループ化が一意に決定できるのかと言う疑問がある。GTTM では選好規則間の相互作用に関しては意図的に書かれておらず、分析者の判断に委ねられている。このため分析が食い違うことが起こり得るのである。実際にはこのような不整合をどのように扱うかが問題になる。

階層構造に関しては、GTTM の「簡約構造 (reduction)」を例に述べている。簡約とはグループ中から代表となる 1 つの要素 (単音や和音) を撰択することを示す。この簡約操作を再帰的に適用することにより、階層的な簡約構造が得られる。GTTM には隣接するグループどうしの構造的な重要性に基づく簡約と、調性的な緊張-弛緩関係に基づく簡約の 2 つがある。しかしこれらに対し著者は以下のように述べている。

曲全体が緊張から弛緩(解決)へ向かう事を解析することにどれだけの意味があるのか、むしろ簡約が結果として切り捨てるもの(メロディラインやリズムのような曲の微細構造)の方に認知的な顕著さ (salience) があるのではない

だろうか．

類似性の認識に関しては，どのような要因のもとで類似性が成立するのか，認知システムの中でどのように処理され，どういう役割を果たしているか，について述べている．前者は類似をどのように認識させるかということであり，リズムや音程，音符の平行移動など様々な類似をどのように認識するのかという問題を挙げている．具体的には類似性チェックの為の照合の計算量問題と，直感的に類似性を感じないゴミまで拾ってきてしまう可能性のある照合問題である．

これらの問題に関する考え方は本研究において，入力ファイル形式，GTTMの各構造の相互関係，曖昧性の処理，簡約の扱い，パラリズムの扱いを考える際に参考になっている．

## 2.2 GTTMを実装するために本研究で必要と考えた理論の研究

本研究で必要と考えた関連研究は大きく分けて，直接 GTTM に関係するものと，入力ファイルのデータ構造などの間接的な研究の 2 つに分けられる．ここではこれら 2 つに分類して紹介する．

### 2.2.1 直接 GTTM に関係する研究

#### MAURO BOTHELHO による研究

MAURO BOTHELHO[12] による研究では，グルーピング構造をリズム的なグルーピングと調性的なグルーピングに分けている．リズム的なグルーピングはデュレーションと輪郭のパターンによって 1 つの時間の組織化であり，調性的なグルーピングはボイスリーディングのパターンによって作られた 1 つの時間の組織化である．そしてこの調性的なグルーピングが GTTM に取りこまれるなら，グルーピング選好規則の GPR7 を弱め，GPR6 を強めるであろうと述べている．

## Péter Halász による研究

Péter Halász[13] による研究では、タイムスパン簡約をコンピュータ上でシミュレーションしている。Péter Halász はこの際、優先規則の TSRPR<sub>4,5,6</sub> を省いている。これらはコンピュータの限られた能力のために省いたが、経験的にこれらを省いても分析過程には影響は無いと言っている。また、シミュレーションの結果からタイムスパン簡約は以下の2つの段階に分けることができると述べている。

第1段階 簡約は、異なった拍節構造をもっている異種の和音の継承から成り立つ音楽の構成に適用される。この構成は、重要な拍節の音を所有している基礎の和音に次第に単純化されていく。

第2段階 第一段階で残った和音は、異なるグループの初めと終わりの両端である。これらは、各グループの重要性に関して次第にお互いを除外する。

つまり簡約の初期は、拍節構造を参照して和音単位のグループに簡約していく。その後、グループ間に対して和声的に重要な方へ簡約をおこない、木構造を生成していくという事である。

## 上符による研究

上符裕一 [16] による研究では、タイムスパン簡約までの結果が出ていると仮定し、延長的簡約について評価実検、考察をおこなっている。この実検では、音楽構造解析理論を計算機モデルとして構築するために、人手による解析結果を集めたコーパスを利用して選好規則を自動抽出し、曖昧性を定量的に解消する手法を提案している。また実験結果から、少量のコーパスにおいてもスムージング手法により有効な規則を推定でき、未知事象を多く含む場合には有効であることを示している。

この研究報告からも分かるが、実際にはタイムスパン簡約までの分析が非常に難しく、実装が困難である。

## 2.2.2 間接的に関係する研究

### Gerd Castan, Michael Good, Perry Roland による研究

Gerd Castan, Michael Good, Perry Roland[5] による研究では、音楽を Extensible Markup Language(XML)<sup>1</sup>を用いて表現している。プログラムを作成するうえで、そのプログラムを扱う環境での互換性を考えると、出力するデータの形式は重要であると考えられる。この場合、文字を1バイトか2バイトの文字コードのまま扱う TEXT 形式が適当である。また、データにおける検索を考えると、単語ごとに区切り子をいれた Markup を採用すべきである。

初期のマーク付け言語 GML から発展した SGML(Standard Generalized Markup Language) が開発されている。これは ISO, JIS 共に登録されている規格である。

インターネットの初期の時期には、SGML をメタ言語とした HTML が作られた。しかし、HTML は文章をブラウザに表示することを目的としているので、既成の文章の構造タグ付けや、内容タグ付けができない。このため、W3C により Extensible Markup Language(XML) が検討された。この XML ver1.0 は <http://www.w3.org/TR/REC-xml> で公開されている。

本研究では将来的に、解析した結果のデータベースにネット上から検索できるようなシステムにも対応できるように XML を採用する。

### 鷺坂による研究

鷺坂 [17] による研究では標準 MIDI ファイル<sup>2</sup> を入力としてメロディを自動抽出している。標準 MIDI を入力とした理由は以下のとおりである。

- 大量の音楽データをインターネットなどで手に入れることができる。このため大規模データベースの構築が容易である。

---

<sup>1</sup>XML はインターネットの標準として W3C より勧告されたメタ言語である。メタ言語とは、言語を作る言語という意味である。XML は電子的な文章を自動管理するために、論理的な文章構造を記述するようにした言語である。

<sup>2</sup>MIDI(Musical Instrument Digital Interface) とは、楽器とコンピュータを相互接続する際のプロトコルである。この MIDI には、鍵盤が押された、ノブが回された、ジョイスティックが操作された等の情報を送信する方法が定められている。

- 演奏時間の記述が明瞭で，音楽データから演奏を再現できる．
- 標準化された楽譜記述言語であるため，検索の結果得られたデータを他のマルチメディアソフトに容易に取りこむことができる．

MIDI を用いて音楽を扱う場合，それは昔のピアノのロール紙に対比することができる．それは，MIDI メッセージが音の波形ではなく，制御データを表現しているからである．制御データには，例えば「音を鳴らせ」「音を止めろ」「音色を何番にしる」などがある．表 2.3 に代表的な MIDI メッセージを挙げる．

本研究でも既存の MIDI データに対応できるように，入力として MIDI を用いる．



表 2.3: MIDI メッセージの型

---

• **noteOn**

鍵盤上で一つの音を弾いた時，シンセサイズはその音を鳴らすと同時に 3 バイトのメッセージを MIDIOUT ポートから送信する．

---

[noteOn]

channel: 1

noteNumber: 60

velocity: 116

---

• **noteoff**

その音の発音が終了する時，鍵盤はもう一つの 3 バイトメッセージを送出する．

---

[noteOff]

channel: 1

noteNumber: 60

velocity: 40

• あるシンセサイズでは，ノートオフメッセージの代わりにベロシティ 0 のノートオンメッセージを送信する．

---

• **programChange**

このメッセージは，チャンネルセレクトバイトとプログラムセレクトバイトを含む．このメッセージを受信すると，そのボイスやパッチが指定されたプログラム番号に切りかわる．

---

[programChange]

channel: 1

number: 32

---

## 2.3 GTTMの応用に関する研究

平田，青柳による研究

平田，青柳 [7] による研究では，楽曲の構造を GTTM のタイムスパン簡約により求めている．隣り合った 2 つの音を楽曲構造的に重要な音とそうでない音に分け，重要な音を重要な音に集約するように 2 分木を作り，これを再帰的に行なうことにより木構造を構築していく．これにより求められた木構造に対し，最小上界，最大下界を用いて楽曲を編曲する．実際には以下の手順により求める．

編曲を行なうときの音楽家の内省に基づき，以下のステップで  $M$  から  $P$  を合成する．

- S1.  $M$  に類似した  $M_c$  を持つ事例を検索する．
- S2.  $M$  と  $M_c$  の共通部分を  $M_a$  とする．
- S3.  $M_c$  から  $M_a$  への変換  $F$  を模倣して， $P_c$  を  $P_a$  に変換する ( $F'$ )．
- S4.  $M$  から  $M_a$  への変換  $G$  を模倣して，さらに逆転させて  $P_a$  から  $P$  を作る ( $G'^{-1}$ )．

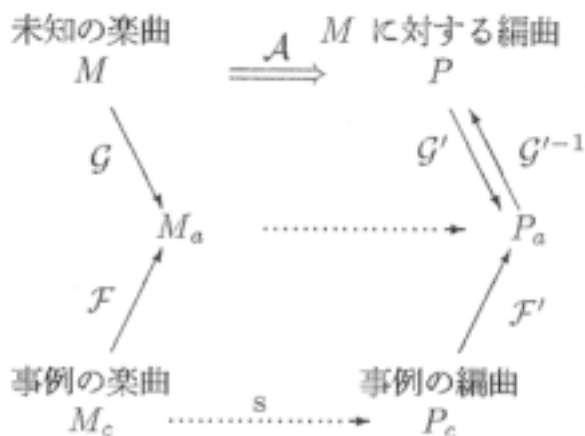


図 2.1: 事例に基づく編曲の形式化 (文献 [7] より参照)

S1. の最類似事例の検索は，パーピーブン [8] で提案した方法 (相対的類似度) を援用する．図 2.1 中， $s$  の点線は，音楽家が  $M_c$  という簡単な譜面を見たときに  $P_c$  という演奏を行なったという対応関係を表す．

編曲アルゴリズムの設計に際し以下の仮定を置く．

A1.  $M$  と  $M_c$  は似ている .

A2.  $\mathcal{F}$  と  $\mathcal{F}'$  は似ている .  $\mathcal{G}$  と  $\mathcal{G}'$  も似ている .

ここで , A1. は S1. で保証される . A2. は音楽家の内省より得られる .  $M, M_c, M_a, P, P_c, P_a$  は DOOD<sup>3</sup>[18] のオブジェクト項として表現されるので , これらオブジェクト項間には包摂関係が成り立つ .

編曲のアルゴリズムは , まず  $\mathcal{F}$  を実現する具体的な関数について考える .  $\mathcal{F}$  を特徴づける関係は ,

$$\mathcal{F}(M_c) = M_a \quad (F1)$$

$$\mathcal{F}(M) = M \quad (F2)$$

である .  $\mathcal{F}$  には様々な関数が考えられるが ,  $M$  と  $M_c$  の共通部分  $M_a$  を計算するには最小上界を用いるのが自然なので ,

$$\mathcal{F}(X) = \text{lub}(X, M)$$

という定義が考えられる .  $P_a$  は  $P_a = \mathcal{F}'(P_c)$  として求められるが , A2. より

$$P_a = \mathcal{F}(P_c) = \text{lub}(P_c, M) \quad (2.1)$$

を得る .

次に , 関数  $\mathcal{G}$  を特徴づける関係は

$$\mathcal{G}(M) = M_a \quad (G1)$$

$$\mathcal{G}(M_c) = M_c \quad (G2)$$

なので , 同様に

$$\mathcal{G}(X) = \text{lub}(X, M_c)$$

とすることができる . ここで  $\mathcal{G}$  を特徴付ける関係 G1, G2 の重要度について考える . A1. から , G1 において  $M$  を  $M_c$  で置き換えた場合 1 と , G2 において  $M_c$  を  $M$  で置き換え

---

<sup>3</sup>Deductive Object-Oriented Database(演繹オブジェクト指向データベース) 属性の欠落や属性の型宣言を記述するために一階述語論理を拡張した手法である . さらに演繹規則によって項間の包摂関係を定義することができ , この包摂関係によって曖昧な項とそうでない項を形式的に結びつけることが可能となる .

た場合 2 を比較する．場合 1 からは， $\mathcal{G}(M_c) = M_a$  かつ  $\mathcal{G}(M_c) = M_c$  より  $M_a = M_c$  が得られる．一方，場合 2 からは， $\mathcal{G}(M) = M_a$  かつ  $\mathcal{G}(M) = M$  より  $M_a = M$  が得られるが，G1 も考慮すると  $\mathcal{G}$  は恒等写像になり， $\mathcal{G}$  の特徴が消失する．よって，G2 の方がより重要な関係であることが分かる．

逆写像  $\mathcal{G}^{-1}$  を特徴付ける関係は，G1，G2 から， $\mathcal{G}^{-1}(M_a) = M$  及び  $\mathcal{G}^{-1}(M_c) = M_c$  となる．上の G1 と G2 の重要度に関する考察と A1. から，前者の関係を  $\mathcal{G}^{-1}(M_a) = M_c$  と置き換えて， $\mathcal{G}^{-1}$  を実現する関数を考える．単純なものとして，

$$\mathcal{G}^{-1}(X) = glb(X, M_c)$$

が考えられる．ゆえに A2. より，

$$P = \mathcal{G}^{-1}(P_a) = glb(P_a, M_c) \quad (2.2)$$

である．

以上の考察をまとめると， $M$  から  $P$  を合成する編曲アルゴリズム  $\mathcal{A}$  は式 2.1，2.2 より

$$\mathcal{A}(X) = glb(lub(P_c, X), M_c) \quad (2.3)$$

と表すことができる．

しかし，この編曲システムのためのタイムスパンは専用のエディタを使用して人手により入力する．このため入力者の音楽的なスキルが要求される．

## 東条による研究

東条 [15] による研究では、和音と和声の進行を HPSG<sup>4</sup>を用いて文法的に表している。楽曲からカデンツ<sup>5</sup>を求め、そのカデンツから次のような3つの基本進行を見いだす。

- T-D-T
- T-S-T
- T-S-D-T

この進行を以下の生成規則により定義する。

$$\bar{T} \rightarrow S, \bar{T}.$$

$$\bar{T} \rightarrow D, T.$$

$$\bar{T} \rightarrow T.$$

$T$  に付随する任意回数の  $\bar{T}$ :

$$\bar{\bar{T}} \rightarrow T, \bar{T}^*$$

これを HPSG で表すことにより、DOOD では自分で定義していた階層的な文法概念やその中のヘッドの概念などを組み込むことができる。

この研究が対象とする入力は音符列であるため、自動化を試みる場合には音符列から和音を検出しなければならない。この検出処理に GTTM のグルーピング構造が利用できると思われる。

---

<sup>4</sup>Head-driven Phrase Structure Grammar. 日本語では、主辞駆動句構造文法と呼ばれる。属性構造の考え方にさらにヘッドの概念を中心に据えた文法理論。

<sup>5</sup>楽曲の終わりおよび大部分の段落点などに特有の和声構造をいう。効果としては、完結的なもの、中間的に一段落して後続部分を期待させるもの、軽いひとくぎりなどさまざまである。1個のセンテンスを組み立てる上での単語の働き（構文規則）に「主語」「述語」「補語」の3種があるように、1個のカデンツを組み立てる上での和音の働き（和音機能）に「(T) トニック」「(D) ドミナント」「(S) サブドミナント」の3種がある。

## 東, 森, 小杉による研究

東, 森, 小杉 [2] による研究では GTTM のグルーピング構造を用いて, 計算器上でメロディを自動生成している. そのシステムの概要は, ランダムに発生させた音符列に幾つかの制約ルールを与え聴取実験を行なうことによって, 人間が旋律と認識するために必要な「音の規則」を明らかにしている. この際用いる制約ルールの 1 つが GTTM におけるグルーピング構造である.

## 第 3 章

# Generative Theory of Tonal Music

### 3.1 GTTM の概要

GTTM は、音楽に関して専門知識のある聴取者の直観を形式的に記述するために Fred Lerdahl と Ray Jackendoff によって 1983 年に提唱された [4]. 楽譜に表された楽曲を表層構造として、これを分析することにより内在する階層構造を深層構造として明らかにする。この理論は Chomsky の変型生成文法の枠組にならったと言われている。楽曲を分析するための理論は他にも幾つか提案されているが、GTTM は計算器上に実装を考えた場合に最も適していると考えられている [11].

GTTM ではすべての楽曲を本質的にホモフォニー<sup>1</sup>であるとして扱い、楽曲の全ての声部に対して単一のグルーピング解析を行えば十分であると仮定している。

この理論は以下の 2 つの基本構造と 2 つの簡約の為の規則から構成されている。

- 1 グルーピング構造
- 2 拍節構造
- 3 タイムスパン簡約
- 4 延長的簡約

---

<sup>1</sup>Homophony. 主声部の旋律に対し簡単な伴奏を付した作法のようなもの。単旋律的に解釈できる。

グルーピング構造と拍節構造の結果からタイムスパン簡約を求める。タイムスパン簡約は構造的に重要な音とそうでない音の2つを枝に持つ2分木構造として表される。このタイムスパン簡約が、GTTMの表す楽曲の構造であり、楽曲のもつ意味である。

また、GTTMではタイムスパン簡約の拡張として延長的簡約を定義している。延長的簡約はタイムスパン簡約の結果を用いてトップダウンに行われる。

上記の各々の構造は以下の2種類の規則によって定義されている。

#### 1 構成規則 (Well-formedness rule)

構造を生成するためのルール。

#### 2 選好規則 (Preference rule)

複数の構造が構成規則を満たす場合、好ましい構造を示すためのルール

2つの基本構造と2つの簡約について、構成規則と選好規則を文献 [1][14] を参考に解説する。

### 3.1.1 グルーピング構造

グルーピング構造は、楽曲をより小さいまとまり (グループ) に分ける構造である。その Group 一つ一つは、あるまとまりを持って聞こえる単位である。例えば歌を歌う時、一息で歌う範囲の様なものである。さらにそのグループを重ね、より大きなグループを作っていく、最終的には楽曲全体が1つのグループにまとまる。図 3.1 はグルーピング構造の一例である。楽譜の下に表されている括弧はグループを表す。グルーピング構造における各々の規則は、グループの境界となる場所の候補を特定するものである。



図 3.1: グルーピング構造の例

グルーピング構造は以下の構成規則と選好規則によって与えられる。



## グルーピング構成規則 (Grouping Well-Formedness Rule)

GWFR1 構成要素が連続している場合のみグルーピングを形成することができる。

GWFR2 1つの曲は1つのグループである。

GWFR3 グループはより小さなグループ (サブグループ) を内部に含んでもよい。

GWFR4 グループはサブグループの一部だけを含むことは許されない。サブグループ全体を内部に含まなければならない。

GWFR5 グループがサブグループを含むなら、グループ構造が交差しないサブグループ群によって内部を埋めつくされねばならない。

GWFR1 は連続した音符列からのみ、グループは作られる事を示している。

GWFR2 は曲全体が1つのグループグループとしてまとまる事を示している。

GWFR3,4,5 はグループは、さらに小さなサブグループに分ける事ができる事を示す。この際、サブグループ同士は図 3.2 のように交差したり間隔をおいたりしない。

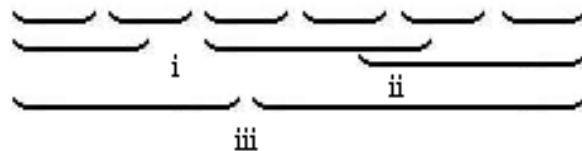


図 3.2: サブグループの例

図 3.2 の i はサブグループ同士が密接せず離れてしまっているために誤りである。ii はサブグループ同士が交差してしまっているために誤りである。iii はサブグループの境界とその上のレベルにおけるグループの境界が一致していないために誤りである。

## グルーピング選好規則 (Grouping Preference Rule)

GPR1(alternative form) 非常に小さいグループへの解析は避ける。特に単音をグループとすることは避ける。

GPR2(proximity) 4つの音符 ( $n_1, n_2, n_3, n_4$ ) が連続しているとする。以下の条件のどれかが成立すれば  $n_2$  と  $n_3$  の間がグループの境界と認識される ( $n$  に休符は含まない)。

- a. slur/rest  $n_2$  の終わりから  $n_3$  の始まりまでの時間間隔が、 $n_1$  の終わりから  $n_2$  の始まりまでの時間間隔および、 $n_3$  の終わりから  $n_4$  の始まりまでの時間間隔よりも長い (スラーの終わりの音はスラーの中の音より、発音されている長さは短くなる)。
- b. attack-point  $n_2$  の始まりから  $n_3$  の始まりまでの時間間隔が、 $n_1$  の始まりから  $n_2$  の始まりまでの時間間隔および、 $n_3$  の始まりから  $n_4$  の始まりまでの時間間隔よりも長い。

GPR3(change) 4つの音符 ( $n_1, n_2, n_3, n_4$ ) が連続しているとする。以下のどれかが成立すれば  $n_2$  と  $n_3$  の間がグループの境界と認識される。

- a. register  $n_2 - n_3$  間の音高差が  $n_1 - n_2$  間の音高差および  $n_3 - n_4$  間の音高差よりも大きい。
- b. dynamics  $n_2 - n_3$  間でダイナミクスの変化があり  $n_1 - n_2$  間、 $n_3 - n_4$  間ではそれが無い。
- c. articulation  $n_2 - n_3$  間でアーティキュレーションパターンの変化があり、 $n_1 - n_2$  間、 $n_3 - n_4$  間ではそれが無い。
- d. length  $n_2$  と  $n_3$  が異なった音長をもち、 $n_1$  と  $n_2$  もしくは  $n_3$  と  $n_4$  が同じ音長である。

GPR4(intensification) GPR2,3 で示される効果が比較的明白なところは大きなレベルにおいてもグループの境界がそこで位置付けられる可能性が高い。

GPR5(symmetry) グループの分割が長さの等しい2つの部分からなるようグルーピングすることを優先する。

GPR6(parallelism) グループ間で並行した部分を形成することができる2つもしくはそれ以上のグルーピングは、並行性のあるグルーピングを行う。

GPR7(Time-Span and Prolongational Stability) タイムスパン簡約や延長的簡約がより安定するグルーピング構造を優先する。

GPR1は優先度的には低いが、小さいグループがより小さいサブグループになることを避ける規則である。

GPR2と3は本質的には同じ性質を表している。これらは以下のような限定された知覚の変化が起こる場所にグループの境界があると示す。

- 演奏された音符の間における休符。
- アタックポイントの間の長さの変化。
- 音域の変化(連続した音符の間の相対的な音程の変化)。
- 連続した音符の音の大きさ(ダイナミクス)における変化。
- 調音の変化(レガートやタイで演奏された音符とスタッカートでの演奏)。
- 音符長の変化。

GPR4は、これらの規則の幾つかが同じポイントにおいて適用可能なら、そのポイントがグループの境界を表すための優先が高くなることを示す。

GPR5は、グループの中のサブグループは等しい長さで2当分されるものを優先することを示す。

GPR6は、節が並行していると解釈できるときは並行したグルーピングを行うことを示す。

GPR7はグルーピング構造を求めた後で解析される2つの簡約との相互作用を示す規則である。より安定したタイムスパン簡約あるいは延長的簡約を作りだすグルーピング構造が、より優先されることを示す。

### 3.1.2 拍節構造

拍節構造は指揮者が指揮棒を振ったり，聞き手が足でリズムをとることに似ている．この構造は楽曲中の最も短い音符から，より長い音符のレベルすべてに対し，強拍と弱拍を同定する分析である．図 3.3 は拍節構造の一例である．楽譜の下に表されている括弧は拍節構造を表す．



図 3.3: 拍節構造の例

拍節構造は以下の構成規則と選好規則によって与えられる．

#### 拍節構成規則 (Metrical Well-Formedness Rule)

MWFR1 すべてのアタックポイントは，曲中の各部分における最小の拍節レベルの拍でなければならない．

MWFR2 あるレベルのすべての拍は，より小さなレベルでの1つの拍でもある．

MWFR3 各々の拍節的レベルで，強拍は2または3拍の間隔をもつ．

MWFR4 タクトゥス<sup>2</sup>や大きな拍節レベルは，同等な間隔をもつ拍によって構成される．

MWFR1 は，すべての音符に拍節構造の拍がなければならず，拍が音符の開始点において起こると示すことで，拍節構造に音符のアタックポイントを関連付ける．

MWFR2 は，あるレベルにおける拍が，それぞれのより低いレベルにおいて同様に拍でなければならないと示す．

MWFR3 は，それぞれのレベルにおいて強い拍が2拍あるいは3拍離れてなければならないと示す．

<sup>2</sup>拍を意味するラテン語．基準音価のことであり，楽曲中の音符の最小の長さである．

MWFR4は、タクトゥスにおいてレベルや上記の拍が等しく間隔を置かれなくてはならず、より低いレベルのために存在しているとき、弱拍が強拍の間に等しく間隔を置かれなくてはならないと定義する。

#### 拍節選好規則 (Metrical Preference Rule)

MPR1(parallelism) 複数のグループ、またはグループの各部を並行的と解釈できる場合、並行的な拍節構造を優先する。

MPR2(strong beat early) 最も強い拍がグループ内で比較的早く現れる拍節の構造を優先する。

MPR3(event) 拍点に音符がある(さらに強拍となる)拍節構造を優先する。

MPR4(stress) 強く演奏された拍が強拍である拍節構造を優先する。

MPR5(length) 以下の a~f のような「より長い」という条件を満たす拍を強拍とする拍節構造を優先する。

- a. 相対的に長い音
- b. 相対的に長く続く一定の音量
- c. 相対的に長いスラー
- d. 相対的に長い同じアーティキュレーションパターンの繰り返し
- e. タイムスパン簡約による相対的に長く続く1つの音高(同一音高音の連続)
- f. タイムスパン簡約による相対的に長く続く1つの和声(同一和音の連続)

MPR6(bass) バス音が拍節的に安定した拍節構造を優先する。

MPR7(cadence) カデンツでは拍節的に安定した構造を優先する。つまり他の場合よりもカデンツ内での局所的な選好ルールの違反は避けなければならない。

MPR8(suspension) 掛留音はその解決よりも強拍である拍節構造を優先する。

MPR9(time-span interaction) タイムスパン簡約における競合が最小になるような拍節構造を優先する .

MPR10(binary regularity) 各レベルにおいて , 強拍が1つおきにくる拍節構造を優先する .

10 の MPR は 3 つのタイプに分類できる [1] .

- MPR3,4,5 は知覚の考慮によって動機を与えられる .  
MPR3 は , 音符開始点である拍は強拍であるものを優先すると示す .  
MPR4 は , 緊張させられる音符 (より大きな音で演奏される) が強拍である方をより優先すると示す .  
MPR5 は , その文脈で比較的長いイベントの開始点が強拍であると思われるはずであると示す .
- MPR6,7,8,10 は調性音楽に直接相互参照をする規則である .  
MPR6 は , 低音のための優先規則の違反を防ぐものである .  
MPR7 は , カデンツが拍節上安定するべきであると示す .  
MPR8 は , 掛留音がその解決より強拍であるべきであると示す .  
MPR10 は , あるレベルの全ての 2 つ目の拍が強拍であるべきであると示す .
- MPR1,2,9 は , いろいろな他のモジュールと相互に作用する規則である .  
MPR1 は , 音符の並行したシーケンスが , 並行した拍節構造を与えるべきであると示す .  
MPR2 は , 強拍がグループの始めの方で優先されるべきであると示す .  
MPR9 は , もし拍節構造がタイムスパン簡約において矛盾を最小にするなら , それが望ましいと示す .

### 3.1.3 タイムスパン簡約

タイムスパン簡約は、あるメロディはより小さいメロディの精緻化 (elaboration) であるという考えを表すものである。つまりある楽曲は、その楽曲の中で最も重要と思われる音からの派生で構成されているという考えである。図 3.4 を用いて説明すると、2 つの連続する音符のうち曲のイメージを構成するために、より重要と思われる音符を選択する (簡約する)。これを繰り返し行い、最後に残った 1 音がその楽曲中で最も重要な音符である。

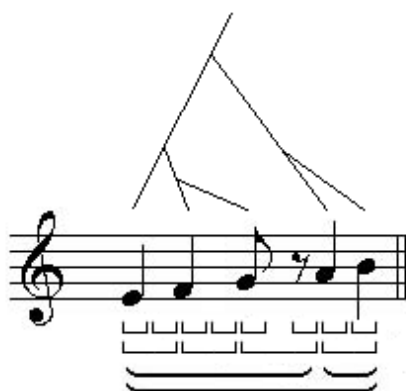


図 3.4: タイムスパン簡約の例

タイムスパン簡約は以下の構成規則と選好規則によって与えられる。

タイムスパン簡約構成規則 (Time-Span Reduction Well-Formedness Rule)

TSRWFR1 全てのタイムスパン  $T$  は、 $T$  のヘッドとなるイベント  $e$  (もしくは音列  $e_1e_2$ ) をもっている。

TSRWFR2  $T$  が他のタイムスパンを含んでいない (最小レベルのタイムスパン) ならば、 $e$  は  $T$  のイベントである (イベント  $e = T$ ) 。

TSRWFR3  $T$  が他のタイムスパンを含んでいるなら、 $T_1, \dots, T_n$  は  $T$  のすぐ近くで含まれるタイムスパンとして  $e_1, \dots, e_n$  をそれぞれのヘッドにすると、以下である。

a. (Ordinary Reduction)  $T$  のヘッドはイベント  $e_1, \dots, e_n$  の 1 つであるかもしれ

ない。

b.(Fusion) もし  $e_1, \dots, e_n$  がグループ境界によって分けられないなら,  $T$  のヘッドは 2 つあるいはさらに多くの  $e_1, \dots, e_n$  の重ね合わせであるかもしれない。

c.(Transformation) もし,  $e_1, \dots, e_n$  がグループ境界によって分けられないなら,  $T$  のヘッドは  $e_1, \dots, e_n$  から撰択される相互に調和したピッチの組み合わせであるかもしれない。

d.(Cadential Retention)  $T$  のヘッドはカデンツで, その最後の  $e_n$  と最後から 2 番目の 2 つからなるタイムスパン  $T_n$  のヘッドであるかもしれない。

TSRWFR4 もし 2 要素のカデンツが直接タイムスパン  $T$  のヘッド  $e$  に従属するなら, ファイナルは直接  $e$  に従属する。そして最後から 2 番目は直接ファイナルに従属する。

TSRWFR1 は, 全てのタイムスパンがヘッドを持っていると示す。

TSRWFR2 は, タイムスパンがひとつのイベントを含んでいるとき, そのイベントがヘッドであると示す。

TSRWFR3 は, 以下のようなヘッドだと思われるものを生成するための制約と方法を与える。

a.(Ordinary Reduction) タイムスパンにおけるイベントの 1 つがヘッドとして選ばれることがありえる。

b.(Fusion) ヘッドが融合によって多くの他のものから組み立てられるかもしれない。

c.(Transformation) ヘッドが変換によっていくつかのイベントから作られるかもしれない。

d.(Cadential Retention) カデンツでの最後のイベントについて最終のイベントとその前の両方がヘッドであるかもしれない。

TSRWFR4 は, 2 つのカデンツの要素がタイムスパンのヘッドに従属しているなら, 最後の要素は直接ヘッドに従属し, 最後から 2 番目の要素は最後の要素に従属することを示す。



## タイムスパン簡約選好規則 (Time-Span Reduction Preference Rule)

TSRPR1 より強い拍の部分の優先 .

TSRPR2 協和部や局所的な主音に関連がある部分の優先 .

TSRPR3 旋律の高い音 , より低いバス音の優先 .

TSRPR4 並行的な部分は , 並行したヘッドとなる .

TSRPR5 より安定した拍節構造部をヘッドとして優先 .

TSRPR6 可能な候補がいくつもある場合 , 延長的簡約において安定度の高い結果を与えるものをヘッドとして優先 .

TSRPR7 カデンツ進行部の優先 .

TSRPR8 開始部の優先 .

TSRPR9 開始部よりも終結部 (カデンツ) の優先 .

9 の TSRPR は 3 つのタイプに分類できる [1] .

- TSRPR3,8,9 は知覚の関心によって動機を与えられるものである .  
TSRPR3 は , 部分における最も高いか最も低い音符に優先を与える .  
TSRPR8 は , グループの始めの方にあるイベントが重要である可能性がより高いと示す .  
TSRPR9 は , 初めが終了よりどちらかという部分全体のヘッドに対して優先されるはずであると示す .
- TSRPR2,7 は調性音楽において , 和音を扱うものである .  
TSRPR2 は , もしそれが局所的な和音と一致しているなら , イベントが優先されるはずであると示す .  
TSRPR7 は , ある部分がカデンツを含んでいるならカデンツはヘッドとして選ば  
れるべきであると示す .

- TSRPR1,4,5,6 は理論の他のモジュールを参照するものである．  
 TSRPR1 は，高い拍節の強さを持っているイベントに優先を与える．  
 TSRPR4 は，並行した部分がヘッ드의並行した配置を持つべきであると示す．  
 TSRPR5 は，より安定した拍節構造に向かうようにするヘッドを撰択するように示す．この規則は，タイムスパン簡約で衝突を最小にするタイムスパン簡約を選択するように示す MPR9 規則と相互に作用する．  
 TSRPR6 は，延長的簡約につながって，タイムスパン簡約が延長的簡約をより安定する方を優先するように作られると示す．

### 3.1.4 延長的簡約

延長的簡約は，タイムスパン簡約では表現できない，以下の 2 つを表すものである．  
 1 つは，グループの終わりの音と同じ音が次のグループの始まりである場合．つまりグループにまたがる音の持続である．もう一つはメロディの進行における緊張と弛緩の構造である．この構造は和声の進行に関する構造分析である．図 3.5 は延長的簡約の例である．延長的簡約は以下の構成規則と選好規則によって与えられる．

#### 延長的簡約構成規則 (Prolongational Reduction Well-Formedness Rule)

PRWFR1 延長的なヘッドとして機能するイベントは，基礎となるグルーピング構造の中にひとつ存在する．

PRWFR2 イベント  $e_i$  は，以下の方法全てに対してもうひとつのイベント  $e_j$  の直接の精緻化でありえる．

- もしルート，バス音符と 2 つのイベントの旋律的な音符が同一であるなら， $e_i$  が  $e_j$  の強い延長である．
- もし 2 つのイベントのルートが同一であり，バス音と旋律的な音符の両方あるいは片方が異なるなら， $e_i$  が  $e_j$  の弱い延長である．
- もし 2 つのイベントの和声のルートが異なっているなら， $e_i$  が  $e_j$  に，または  $e_j$  から進行する精緻化である．

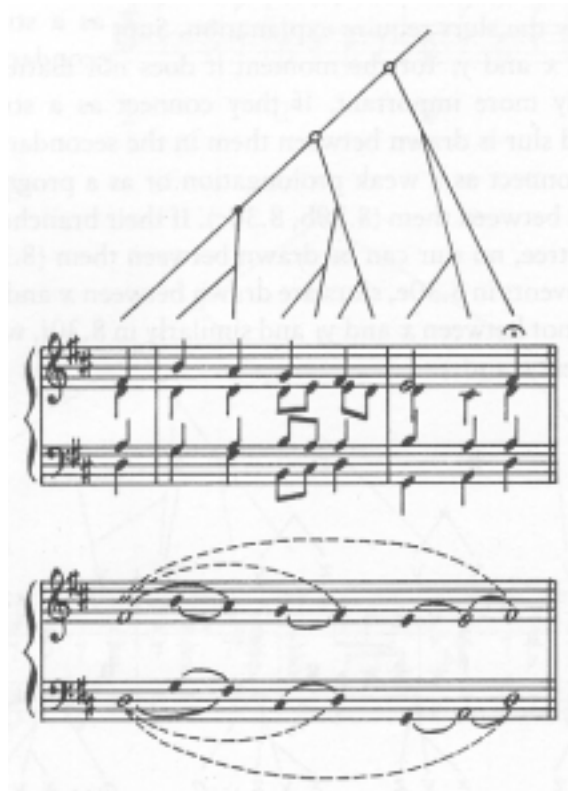


図 3.5: 延長的簡約の例 (文献 [4] より参照)

PRWFR3 基礎をなしているグルーピング構造の全てのイベントが延長的なヘッド，あるいは延長的なヘッドの再帰的な精緻化である．

PRWFR4 もしイベント  $e_i$  がイベント  $e_j$  の直接の精緻化であるなら， $e_i$  と  $e_j$  の間の全てのイベントが  $e_i, e_j$  あるいはそれらの間のあるイベントの直接の精緻化であるに違いない．

PRWFR1 は，全ての小片には延長的なヘッドとなるイベントが存在することを示す．これは，小片における全ての他のイベントが（多かれ少なかれ間接的に）このイベントの精緻化であることを意味する．

PRWFR2 は，分岐によって2つのイベントを結びつける理由を示す．それは，調子の分析（ルート音，バス音，旋律的な音符）から与えられた調和的なパラメータから，弱い精緻化と強い精緻化を定義する．2つのイベントのルートが同じでバス音が異なるとき，弱い延長が起こる．

ルートとバス音両方が同じならば，強い延長が起こる．

PRWFR3 は，ある小片における全てのイベントが，あるレベルにおいて延長的な木の一部にならなくてはならないと示す．

PRWFR4 は，延長的な木における枝は交差しないかもしれないと示す．

### 延長的簡約選好規則 (Prolongational Reduction Preference Rule)

PRPR1(Time-Span Importance) 延長的領域  $(e_i - e_j)$  の延長的に最も重要なイベント  $e_k$  を選択することにおいて， $e_k$  が比較的タイムスパン的に重要である選択の方を強く優先する．

PRPR2(Time-Span Segmentation)  $e_k$  を延長的領域  $(e_i - e_j)$  での延長的に最も重要なイベントであるようにする．もし  $e_i, e_k$  が同じタイムスパン領域に属し， $e_j$  が他のタイムスパン領域に属していれば， $e_k$  は  $e_i$  に属することを優先する．

PRPR3(Prolongational Connection) 延長的領域  $(e_i - e_j)$  での延長的に最も重要なイベント  $e_k$  を選択することにおいて，その領域の終点と共に最大に安定した延長的接続を形成するように付加する  $e_k$  を優先する．

PRPR4(Prolongational Importance)  $e_k$  を領域  $(e_i - e_j)$  での延長的に最も重要なイベントであるようにする。  $e_k$  が延長的により重要な終点の精緻化である延長的簡約を優先する。

PRPR5(Parallelism) 並行した節は並行した解析を行う延長的簡約を優先する。

PRPR6(Normative Prolongational Structure) カデンツのグループが、その延長的構造の中に以下の4つ(5つ)の要素をなるべく含むようにする。

- a. 延長的な始点。
- b. カデンツの1つの要素から成り立っている延長的な終点
- c. 右に分岐する延長、最も重要なのは延長的な開始の精緻化を管理することである。
- d. 右に分岐する延長の連続、(次に)最も重要なのは延長的な開始の精緻化を管理することである。
- e. 左に分岐する延長の連続、カデンツの最初の要素の最も重要な精緻化「サブドミナント<sup>3</sup>」

最初の2つの優先規則は、タイムスパン簡約に延長的な構造を関連付ける。

PRPR1は、タイムスパン簡約における重要なイベントは延長的にも重要なイベントであると示す。

PRPR2は、同じタイムスパンの部分である2つのイベントを付加することを弱く優先すると示す。

残りの優先規則は、これらに矛盾するかもしれない規則である。

PRPR3は、さらに付属を管理する規則の集合である。付属は、ヘッドに向かって進行する要素の左にあるかもしれない。あるいはヘッドの精緻化である要素の右にあるかもしれない。

PRPR4は、イベントをより重要なイベントとして扱うための規則である。

PRPR5は、整合性をもって使用される、不確定な並行規則である。

PRPR6は、分岐のある特定のパターンが、音楽のある特定のスタイルにとって代表的であると示す。

---

<sup>3</sup>Subdominant. カデンツの下属音のこと。

## 3.2 GTTMにおける諸問題

これまで説明してきた GTTM は本来、人間の手によって分析し説明することを目的にしている理論であるため、各構造分析の規則の定義が厳密にされていないという問題点がある。このため、これまで GTTM のコンピュータ上への実装は困難であるとされてきた。また、本システムでは入力として Standard MIDI File(SMF) を用いているため、スラーや休符などの情報が得られない。このため、少ない情報における GTTM の規則適用を考えなければならない。本章ではこれらの問題点についてまとめ、次章でこれら問題点に対し本研究のシステムがどのように解決を試みるかを明らかにする。

### 3.2.1 ポリフォニーへの拡張

GTTM では、全ての楽曲を本質的にホモフォニー (homophony) であるとして扱っている。そのためポリフォニー (polyphony)<sup>4</sup>の楽曲でも分析された結果の構造は1つであるとされている。しかし、図 3.6 のようにポリフォニーの楽曲に対しては、旋律をまたいで並行性が存在したりするので、それを検出するためにも各旋律ごとに解析する必要があると思われる。このため、ポリフォニーはホモフォニーが時間的に重畳しているとみなして、各旋律ごとに解析する必要があると考える。



図 3.6: GTTM における分析の例

<sup>4</sup>polyphony. 多声音楽。複数の声部をもつ音楽のこと。

### 3.2.2 用語定義の曖昧性

GTTM の各規則の中には、規則の定義に使用されている用語自体が曖昧なものがある。拍節構造を例にとって見てみると、Metrical Preference Rule 5(MPR5) の定義中に以下のようなものがある。

- MPR5(length) 以下の a~f のような「より長い」という条件を満たす拍を強拍とする拍節構造を優先する。
  - a. 相対的に長い音
  - b. 相対的に長く続く一定の音量 他

ここで相対的に長いと書いてあるが、その比較対象が定義されていない。そのため比較する範囲は、この定義を適用する分析者もしくは実装者の主観が入ってしまう。同様に曖昧な定義が MPR6, MPR7 にも見られる。

- MPR6(bass) バス音が拍節的に安定した拍節構造を優先する。
- MPR7(cadence) カデンツでは拍節的に安定した構造を優先する。

ここでは拍節的に安定したという表現が使用されているが、どのような状態が拍節的に安定しているのかということが定義されていない。そのため、ここでも主観が入ってきてしまう。

このような曖昧な表現が GTTM の規則のなかに多く現れる。もともと GTTM は人間を対象に解析規則を表す理論であるため、Lerdahl, F. & R. Jackendoff はこの曖昧性を、音楽特有の曖昧性を処理するために GTTM において優れている点であるとしている。しかし人間ではなく計算器上に実装する際には、この曖昧性が実装を難しくする原因になる。

### 3.2.3 規則適用の非一貫性

グルーピング構造の各規則は、グループの境界の候補を特定するものである。そのため基本的には各規則が適用された音符間でグループは切れる。しかし、GTTM のグルーピング例を見てみると同様の条件でも切れている所と切れていない所が存在する。以下の図 3.7 は人手によって GTTM の規則を適用した結果である。

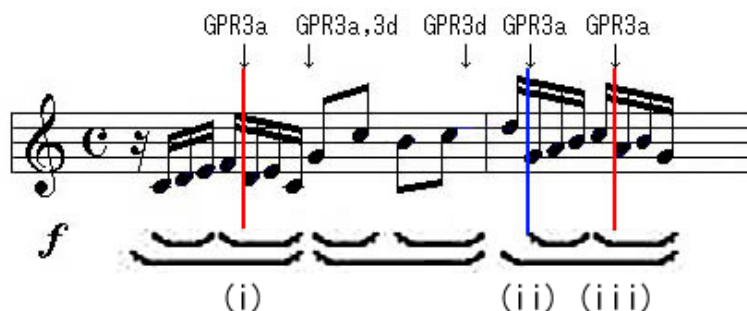


図 3.7: グルーピング構造における曖昧性の例

この例では、(ii) の GPR3a ではグループは切れている。しかし、同様に GPR3a が適用されている (i),(iii) ではグループは切れていない。またこの例以外でも、規則が全く適用されていない場所でグループが切れている場合がある。このように規則の適用を見ただけではグループの切れ目を一意に求めることはできない。

このように、各優先規則の優先順位、扱いが定式化されていないために実装が困難になっている。

### 3.2.4 構造、簡約間における解のフィードバック

GTTM の構造、簡約の参照関係は以下の図 3.8 のようになっている。

図 3.8 において実線は処理の流れに沿った解の参照を示す。点線は処理の流れに反した解の参照を示す。グルーピング構造を例にとって説明すると、グルーピング選好規則に以下のような規則がある。

- GPR7(Time-Span and Prolongational Stability)

タイムスパン簡約や延長的簡約がより安定するグルーピング構造を優先する。



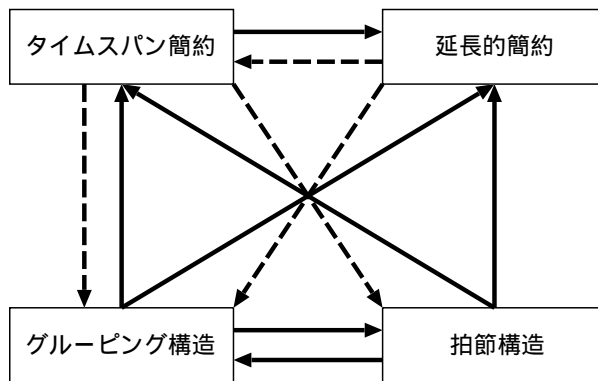


図 3.8: GTTM の各規則の参照関係

プログラム化する際、グルーピング構造を求めてからタイムスパン簡約や延長的簡約をおこなう。しかし、このグルーピング構造を求める際にはタイムスパン簡約の解や延長的簡約の解を用いるのである。処理の流れが一方向でないために実装が困難となってしまうのである。

## 第 4 章

# 声部分析を拡張した GTTM に よる分析法

本章では，前章に挙げた問題点への対処として本研究が提案する手法の説明をおこない，それらの手法の有効性を確認するために本研究が提案するシステムの詳細について説明する．

### 4.1 本研究における問題点の解決

前章で挙げた問題点に対し，プログラムの主観をできるだけ入れずに音楽的に妥当であると考えられる規則のみで実装をおこなう必要がある．本手法では，以下の表 4.1 ような対応で問題点を解決した．

表 4.1: 本手法における問題点への対応

ポリフォニーへの拡張	→	ボイスリーディング分析
用語定義の曖昧性	→	ヒューリスティックスの導入
	→	暗黙的パラメータの明示化
規則適用の非一貫性	→	ボイスリーディング分析
	→	第一声部進行
構造，簡約間における解のフィードバック	→	並列計算の試み

#### 4.1.1 改良した理論における楽曲の扱い

本手法では、ポリフォニーをホモフォニーの時間的重畳として扱う。即ち、まずホモフォニーレベルでのグルーピングを行い、次にホモフォニー間のグルーピングを行う。最終的に一つのグループに集約する。(図 4.1 に例を示す)



図 4.1: 本手法におけるポリフォニー的分析の例

これにより、異なる旋律間での音符の並行性(同じ様な音列)を検出すること等が期待できる。

#### 4.1.2 ボイスリーディング

GTTM の規則は全てグループやフレーズの切れ目を見い出す規則である。従ってポリフォニーに GTTM の規則を適用すると、重畳したホモフォニー間に多数の切れ目を見い出してしまう。そこで GTTM の規則適用を適切な範囲に制限するために、グループやフレーズの連続性を見い出す規則を付加する。また、3.2.3 節で述べた問題点は、グルーピングに関する情報が不足しているものと考え、それを補うためにやはり連続性に関する情報を検出する。それがボイスリーディングと次節で述べる第一声部進行である。

ボイスリーディングは音の進行を表したものである (図 4.2) .



図 4.2: ボイスリーディングの例

図中 1 音目と 2 音目, 4 音目と 5 音目は音程が 2 半音以下になっているので, 連続性があると考えボイスリーディングが成立する. 3 音目の前後においては 3 半音以上の音程が見られるので, ここには連続性がないと考える.

このボイスリーディングを求めることにより, GTTM の規則に, ボイスリーディング中の規則とそうでない規則とに強弱をつけることができる. また, 規則が存在しない場所でもボイスリーディングが切れていれば, そこはグループやフレーズの切れ目の候補とすることができる. 本手法では, ヒューリスティクスとして, ある音と次の音との音程が 2 半音以下のもの [10] をボイスリーディングと定義している.

### 4.1.3 ヒューリスティクス, 暗黙的パラメータの明示化

曖昧な用語の定義を適当に補填してしまうと, GTTM の文献や人手によって分析された結果に即した出力が得られない. そのため, 本システムでは用語の定義の曖昧性を解消するために, ヒューリスティクスを導入した. その典型的な例を以下に示す (表 4.2). 表中, MPR5(Length) の相対性を解消するために文献 [4] から例題を 20 個用いて, ヒューリスティクスとして導入したパラメータの値を調整した.

表 4.2: ヒューリスティクスの例

SearchRangeMPR5a	=	2;
・ 音長を比較する音符は前後 2 個		
RelativelyLongMPR5a	=	1.1;
・ 比較する音符の音長が 1.1 倍以上になっていれば規則を適用		

#### 4.1.4 第一声部進行

第一声部進行は主旋律の抽出を意図している．主旋律は，楽曲のイメージの中でも最も強い音の進行であると考えられる．今回は，この主旋律を捉えるために楽曲中の各音符につき第一声部（最も高い音）を抽出した（図 4.3 の太線）．ボイスリーディングは音程が2半音以上離れている音に関しては連続性が無いとして扱うのに対し，第一声部進行ではそのような音にも連続性があると考えられる．



図 4.3: 第一声部進行

#### 4.1.5 並列計算の試み

構造，簡約間における解のフィードバックを実装するために，本研究では全ての解を次のステップへ並行して解析していく手法を提案する．各構造，簡約における規則を適用した結果，解として考えられるもの全てを次の解析ステップへの入力とする．次のステップでは，各フィードバック規則によって適切な解の絞りこみをおこなう．これによりフィードバックを実現する．この手法を図示すると図 4.4 の様になる．

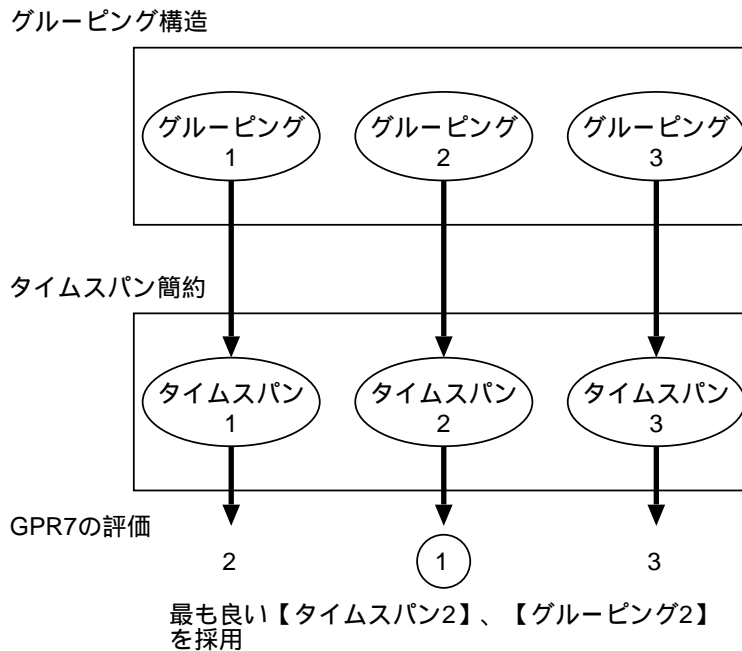


図 4.4: フィードバックの実装

この手法は，マルチエージェントの効用主導エージェントに似た手法である．図 4.4 の場合，グルーピング構造の解析時に解候補が 3 つあったと仮定する．これら 3 つの解候補全てに対しタイムスパン簡約の解析をおこない，ここで，グルーピング優先規則の GPR7 を評価する．GPR7 の評価が最も高かったグルーピング (図ではグルーピング 2) を最適解とし，そのタイムスパン (図ではタイムスパン 2) を最適なタイムスパンとする．

この手法は，一般的な音楽解釈として捉えることができる．音楽はそれを聴いた人によって解釈が異なる．そのため，各構造，簡約の段階で解が一意に求まるというのは不自然な分析であり，意味のある分析であるとは思えない．

また，これまでの研究でも同様の疑問が挙げられている [11][6]. よって，考えられる全ての解候補をすべて次の分析過程に残して分析していく手法が，より人間の分析過程を忠実にモデル化していると考えられる．

## 4.2 本研究が提案するシステム

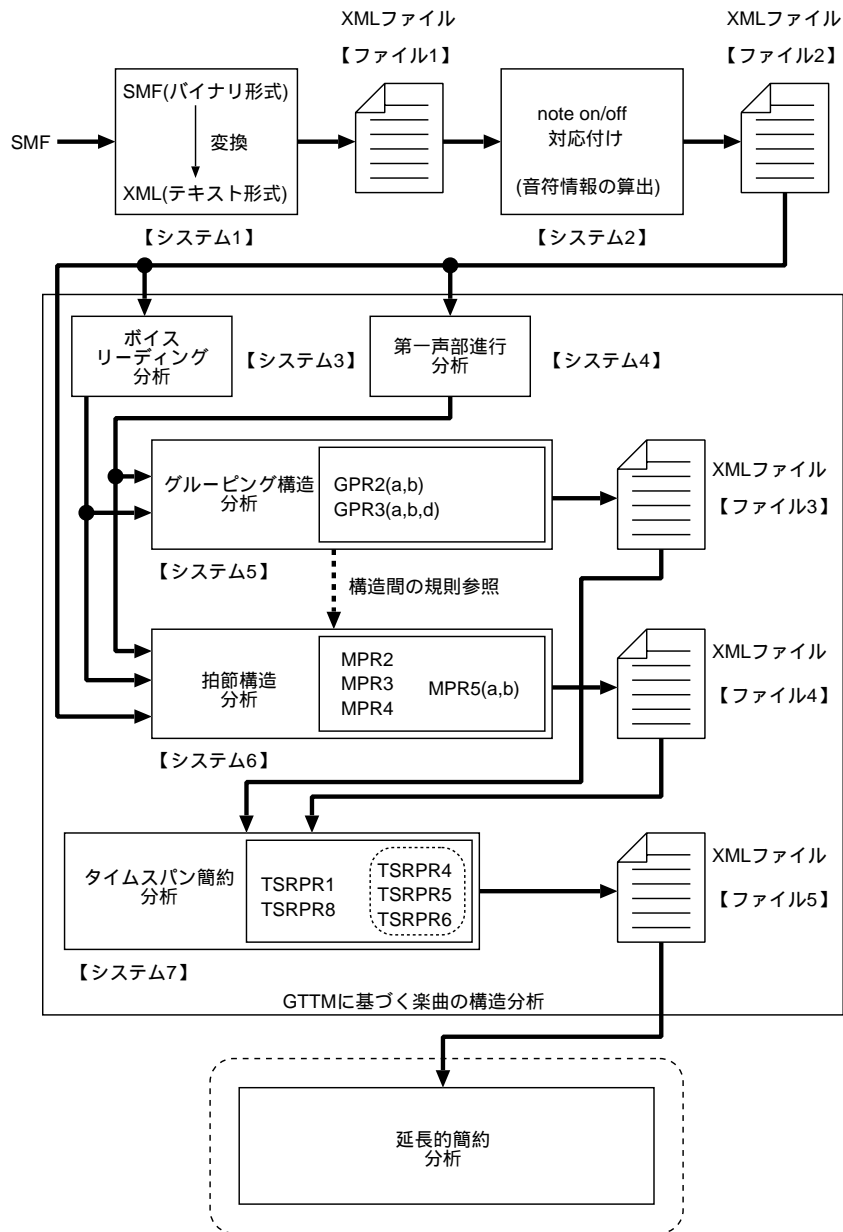


図 4.5: 本研究が提案するシステム

図 4.5 は本研究が提案するシステム構成である。システム全体の入力は標準 MIDI ファイル (SMF) である。システムの各モジュールは中間ファイルとして XML ファイルを出力する。これにより各モジュールの出力である中間ファイルを他のアプリケーションに入力することが可能となる。各モジュールの応用としては、グルーピング構造を入力として和声解析をおこなったり、グルーピング構造と拍節構造を入力として自動伴奏付けをおこなったり、拍節構造を入力として音楽と映像のアンサンブルなどが考えられる。

本システムは 7 個のサブシステムから構成されている。システム 1, 2 は GTTM による解析の前処理であり、システム 3-7 は GTTM に基づく解析システムである。これらシステムについて順に説明する。

また XML ファイルの説明のために、本論文では Microsoft 社から配布されている XML-notepad という viewer を用いる。

#### 4.2.1 【システム 1】ファイル形式変換

SMF<sup>1</sup>形式のファイルを TEXT 形式のファイルに変換する。SMF はバイナリ形式であるため、そのままではプログラム上で扱いにくい。このため前処理として、プログラム上で扱いやすい TEXT 形式のファイルへ変換する必要がある。本システムでは、入出力には XML ファイルを採用しているため、本サブシステムではバイナリ形式のファイルを XML ファイルに変換する。本サブシステムは共同研究者<sup>2</sup>のものを使用した。

本システムが出力するファイルのデータ構造を図 4.6 に示す。

---

<sup>1</sup>Standard MIDI File. MIDI ファイルをソフトや機種に依存しないように一般化した MIDI ファイル形式

<sup>2</sup>青柳龍也，津田塾大学助教授



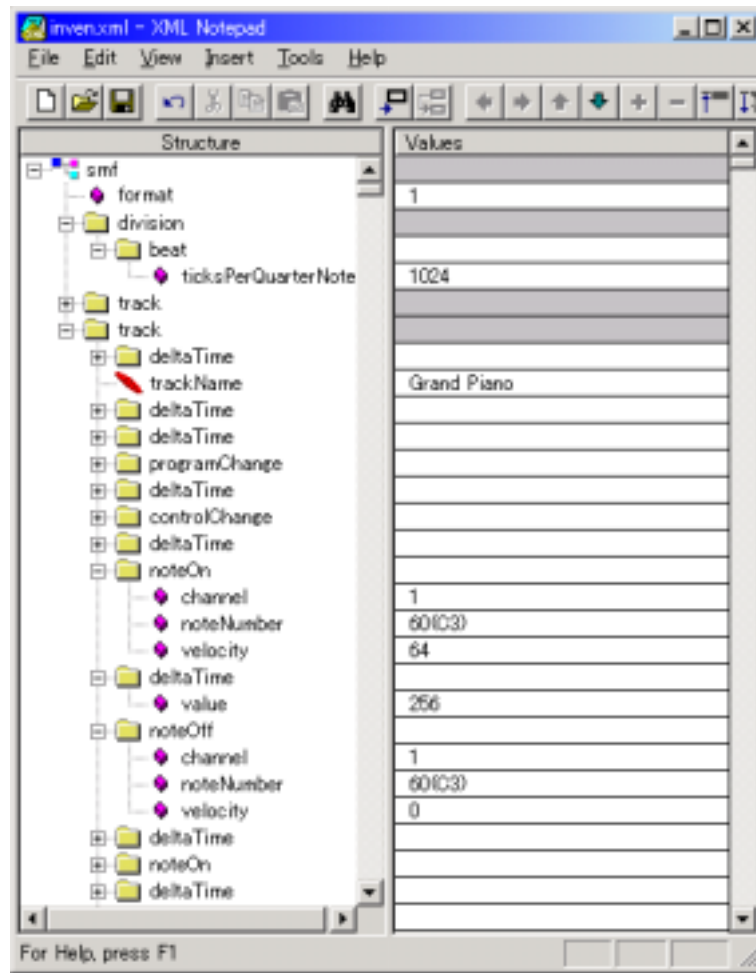


図 4.6: テキスト形式に変換した MIDI ファイル【ファイル 1】

## 4.2.2 【システム2】音符情報の算出

MIDI(Musical Instrument Digital Interface) とは、楽器とコンピュータを相互接続する際のプロトコルである。このMIDIには、鍵盤が押された、ノブが回された、ジョイスティックが操作された等の情報を送信する方法が定められている。

MIDI を用いて音楽を扱う場合、それは昔のピアノのロール紙に対比することができる。それはMIDIメッセージが音の波形ではなく、制御データを表現しているからである。制御データには、例えば「音を鳴らせ」「音を止める」「音色を何番にしる」などがある。表 4.3,4.4,4.5 に本サブシステムでおもに処理する代表的なMIDIメッセージを挙げる。

表 4.3: MIDI メッセージの型 (noteOn)

---

### noteOn

鍵盤上で一つの音を弾いた時、シンセサイザはその音を鳴らすと同時に3バイトのメッセージをMIDIOUTポートから送信する。

---

[noteOn]

channel: 1

noteNumber: 60

velocity: 116

---

表 4.4: MIDI メッセージの型 (noteOff)

---

### noteOff

その音の発音が終了する時、鍵盤はもう一つの3バイトメッセージを送出する。

---

[noteOff]

channel: 1

noteNumber: 60

velocity: 40

---

・ あるシンセサイザでは、ノートオフメッセージの代わりにベロシティ0のノートオンメッセージを送信する。

---

表 4.5: MIDI メッセージの型 (deltaTime)

---

<b>deltaTime</b>
このメッセージは、MIDI 内での時間の推移を表す。単位は tick.
[deltaTime]
value: 1024

---

本サブシステムは、(noteOn,noteOff) の組から duration を求め、新たに表 4.6 の様な note を作成する。

表 4.6: 新たに作成したタグ (note)

---

<b>note</b>
noteOn,noteOff から音符一つ分の長さを求め、以下の情報と共に出力する。
[note]
noteNumber: 60      音符の KeyNumber
velocity: 116      音の強さ
onset:960      その音符が曲の始めから何 tick 目で鳴るか
duration:960      音符の長さ (tick)
noteId:3      その音符は曲の先頭から何個目か
[position]
bar:2      その音符は何小節目で鳴るか
notePosition:1      その音符は小節の何拍目か

---

この処理により複数の MIDI メッセージにより表現されている音符が、note というタグ 1 つで表現できるようになる。入力ファイルにおけるデータの最小単位を音符 1 個にすることにより、楽譜データに近いデータとして解析システムで扱える。

このサブシステムにより、図 4.6 のファイルは図 4.7 のようなデータ構造に変換される。

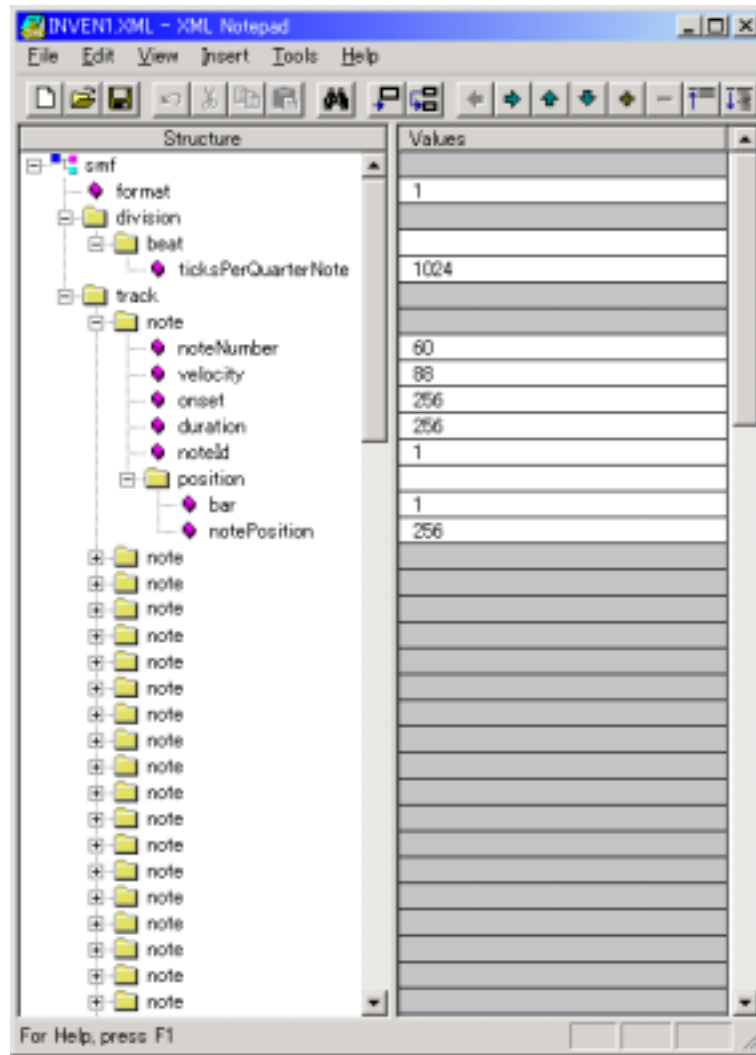


図 4.7: noteOn/noteOff の対応を付けた note 情報【ファイル 2】

### 4.2.3 【システム3】ボイスリーディング分析

GTTMは隣り合った音符間に規則の適用を試みる。GTTMは楽曲をホモフォニーとして見ているため、このような処理になんら問題点はない。しかし、実際の楽曲ではポリフォニー的解釈を行なう方が妥当なものがある。本システムにおいてもこのような楽曲に対応するように、楽曲をポリフォニーとして解釈している。ポリフォニー的な解釈を行なう場合隣り合った音符の選び方には複数の可能性がある。図4.8はこの様子を图示したものである。ボイスリーディングは、4.1.2節で説明した効果意外にも、この音符の組み合わせ候補を絞りこむ効果がある。

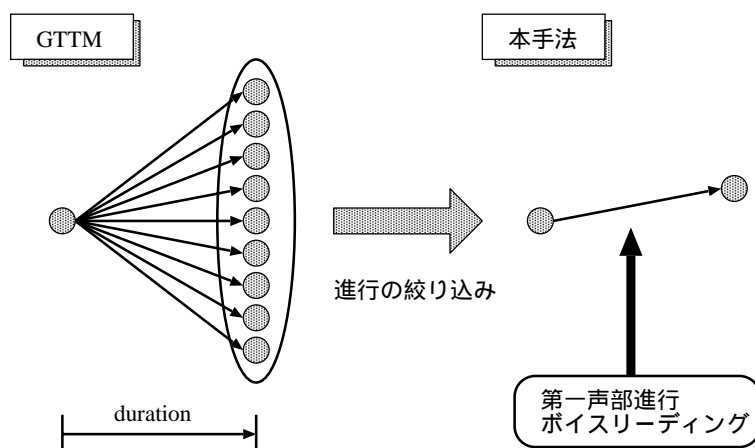


図 4.8: 隣り合った音符の撰択候補

本サブシステムでは基準となる音符から、その音符の長さ (duration) 進んだところにある音符すべてとの音程差を求め、音程差が2半音以下のものに対してボイスリーディングを設定した。ボイスリーディングは全ての音符を基準として分析し、ボイスリーディングを設定した音符に関してはリストとして次の処理に渡した。

### 4.2.4 【システム4】第一声部進行分析

第一声部進行は主旋律の抽出を意図している。また、ボイスリーディング間の接続関係を見いだすことも意図している。第一声部進行は楽曲中の First voice (最も高い音) を抽出したものをを用いている。(図4.3参照) これは各音符情報から、同じ onset 位置 (onset)

のものを検索し，最も音程 (noteNumber) が高いものを順次つなげていくという手法をとっている．

#### 4.2.5 【システム5】グルーピング構造分析

グルーピング構造分析は，図 4.9 のアルゴリズムで実装されている．

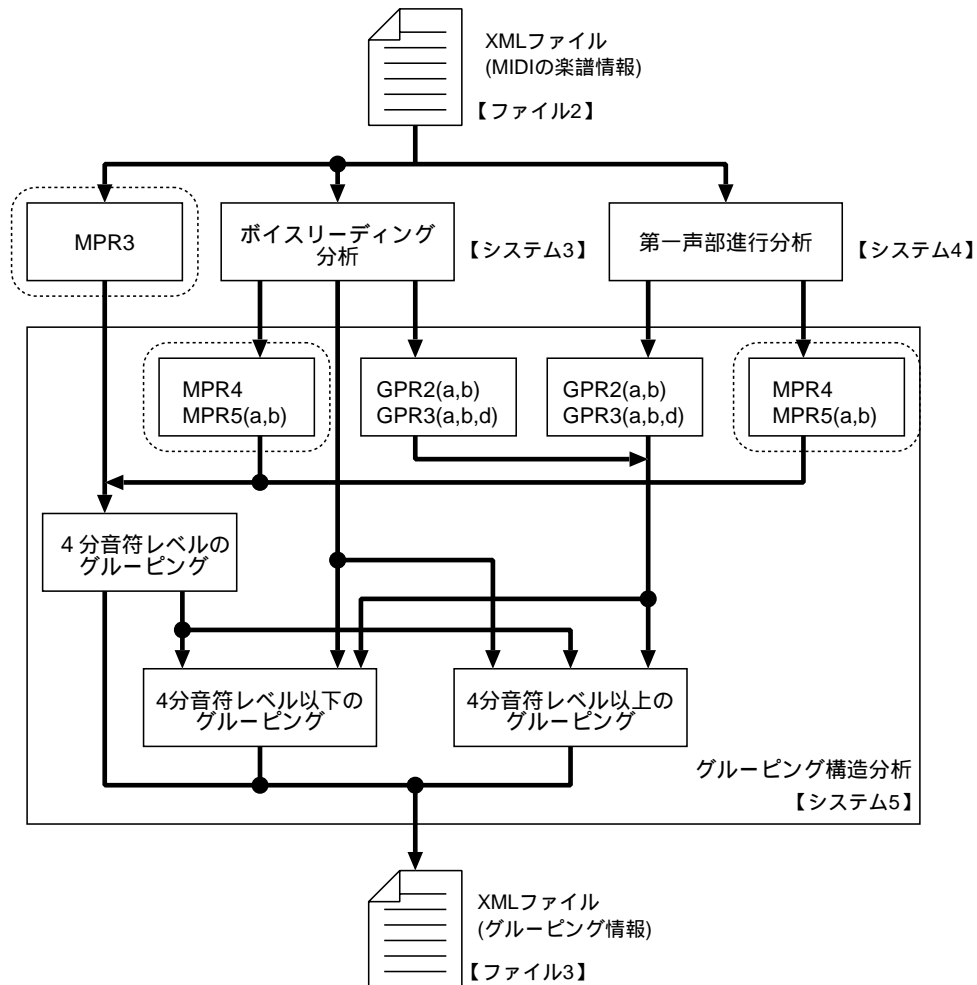


図 4.9: グルーピング構造分析アルゴリズム【システム5】

図 4.9 において，破線で囲んでいる部分は拍節構造を参照する部分である．

GTTM の理論では，より小さなグルーピングをおこない，そのグループを含むような，より大きいグループを重ねていくことにより，楽曲全体を 1 グループとする．しかし本

システムでは、グルーピング問題の簡略化のために4分音符レベルのグルーピングを開始点とする。4分音符レベルのグルーピングの結果を用いて、より小さなグループ、より大きなグループを分析していく。

本システムにおける規則の実装状況は以下のとおりである。

実装済 実装済の規則は以下のとおりである。

GPR2(a,b) GPR2(a) に関してはMIDIファイルにはスラーが無いために、レストの実装のみである。

GPR3(a,b,d) 各評価はヒューリスティクスの値による。

省略 (必要無いと判断) 必要が無い、もしくは実装不可能な規則は以下のとおりである。

GPR1 グルーピングの最小単位として1音を1グループにしても、そのグループを使用しなければ問題無いとの判断。

GPR3(c) MIDIファイルを入力としているため、アーティキュレーションの変化は検出できないので省略。

GPR4 ボイスリーディングや無音状態の検出により暗黙的に実装。

未実装 現在検討中のため、未実装なものは以下である。

GPR5 GPR6 との相互依存関係のアルゴリズム化を検討中。

GPR6 並行性の定義(どの程度一致)を検討中。

GPR7 タイムスパンの規則適用待ち。

本システムがグルーピング時に参照する情報には2種類ある。1つ目は入力される【ファイル2】から拍節構造の選好規則MPR3の適用状況である。この情報を参照することにより、4分音符レベルの拍節がどの音符に依存しているかが確認できる。2つ目は、ボイスリーディングと第一声部進行の分析をおこなった結果に対するグルーピング構造、拍節構造の選好規則の適用状況である。この分析結果は4分音符レベルよりも小さなグルーピングをおこなう際と、より大きなグルーピングをおこなう際に参照する。

4分音符レベル以下のグルーピング 4分音符以下のグルーピングでは、4分音符レベルのグループに対し、以下の条件が成り立つ境界線をグループの境の候補とする。

- ボイスリーディングの境界 .
- GPR が成り立つ境界 .
- MPR が成り立つ境界 .

4分音符レベル以上のグルーピング 4分音符以上のグルーピングでは, 4分音符レベルのグループに対し, 以下の条件が成り立つ2つのグループを結合する .

- ボイスリーディングによってつながっている .
- グループ間に GPR が成立しない .
- グループ間に無音状態 (休符) を挟まない .

以上のグルーピング構造分析によって出力された【ファイル3】が図 4.10 である .



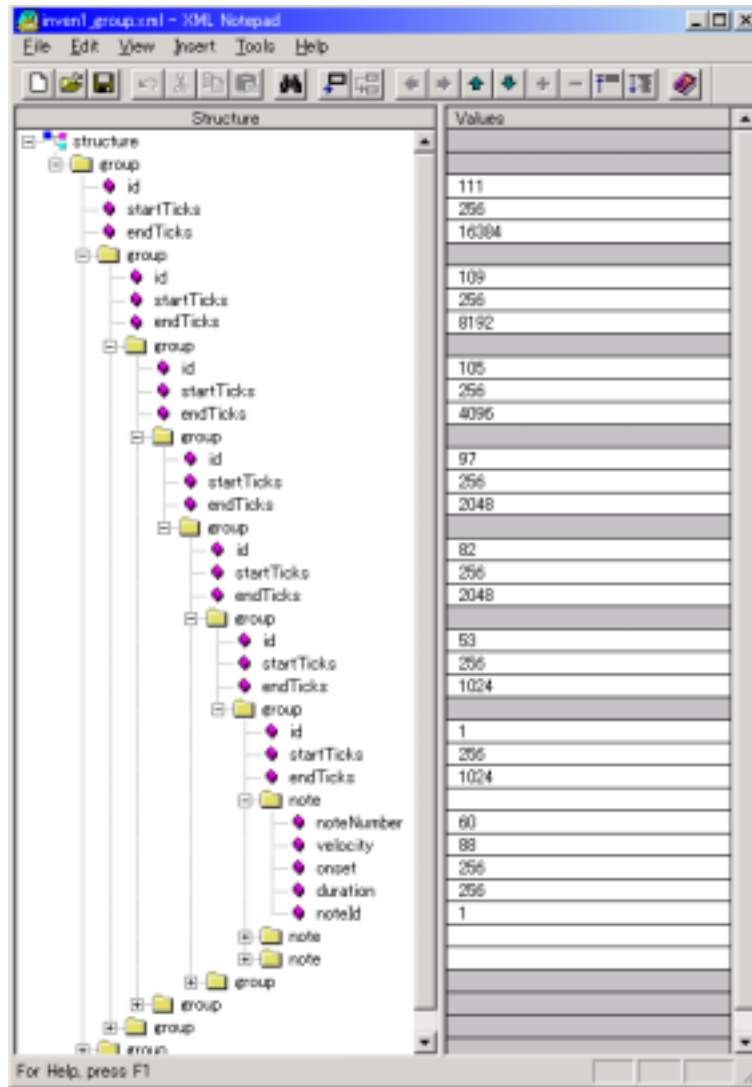


図 4.10: グルーピング構造分析の結果【ファイル3】

## 4.2.6 【システム6】拍節構造分析

拍節構造分析は、図 4.11 のアルゴリズムで実装されている。

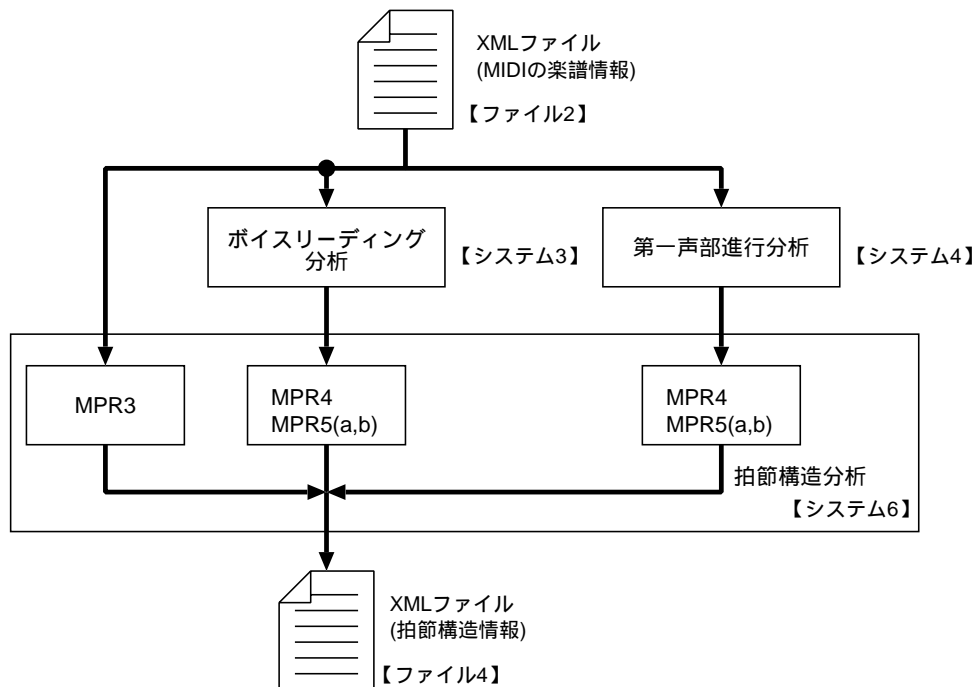


図 4.11: 拍節構造分析アルゴリズム【システム6】

拍節構造の分析はMPR3の分析が最も重要であると考え、なぜなら、MPR3によって音符に拍点を設定することにより、各音符に拍節的な重要度を付加することができ、また強拍・弱拍の分析もおこなえるからである。

MPR3の分析は、楽譜情報から最長の音符を検索するところからはじまる。最長の音符にその長さの拍点を打つ。その後タクトゥスに至るまで各音符長の拍点を打っていく。

本システムにおける規則の実装状況は以下のとおりである。

実装済 実装済の規則は以下のとおりである。

MPR2 グルーピング構造との相互参照により実装。

MPR3 拍点の最大は、楽曲における最長の音符レベルまでである。

MPR4 ヒューリスティクスで設定した範囲で、相対的に強い音にマークする。

MPR5(a,b) 各評価はヒューリスティクスの値による。

省略 (必要無いと判断) 必要が無い, もしくは実装不可能な規則は以下のとおりである。

MPR5(c,d,e,f) c, d に関しては, MIDI ファイルには各情報が含まれていないために実装不可能である。e,f に関してはタイムスパンの規則適用待ち。

未実装 現在検討中のため, 未実装なものは以下である。

MPR1 並行性の定義 (どの程度一致) を検討中。

MPR6 和音認識の実装待ち。

MPR7 和音認識の実装待ち。

MPR8 掛留音認識の実装待ち。

MPR9 タイムスパンの規則適用待ち。

MPR10 楽曲の開始点を導出できないために, 開始点認識の方法を検討中。

以上の拍節構造分析によって出力された【ファイル4】が図 4.12 である。

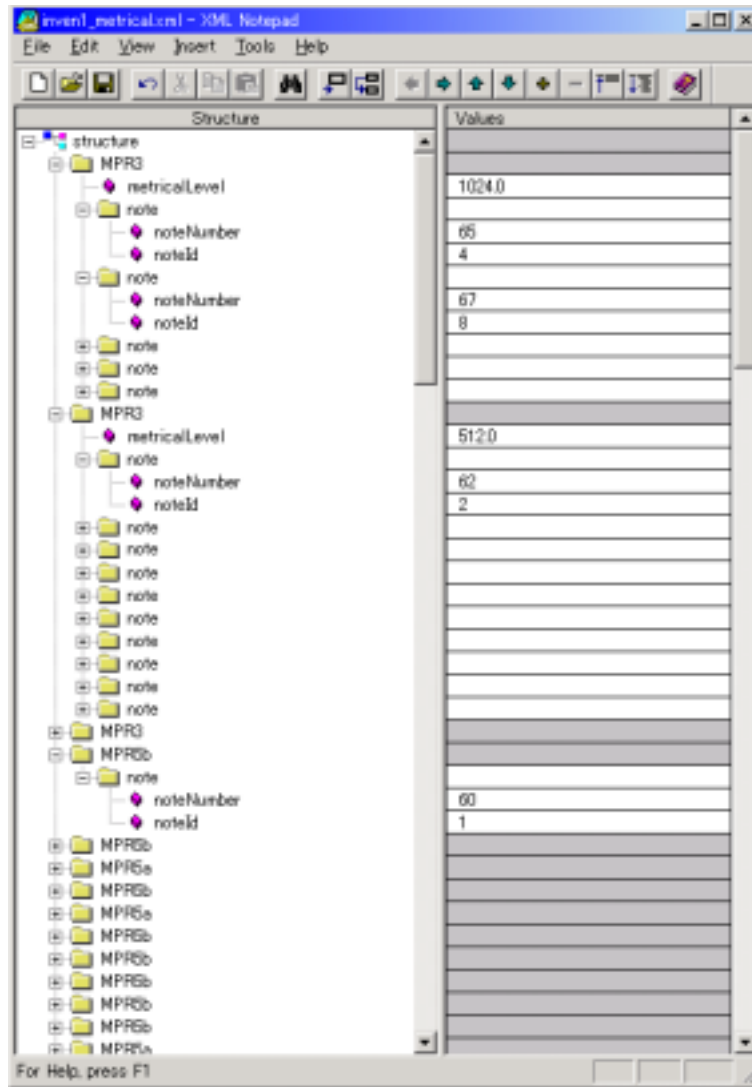


図 4.12: 拍節構造分析の結果【ファイル4】

#### 4.2.7 【システム7】タイムスパン簡約分析

タイムスパン簡約分析は、図 4.13 のアルゴリズムで実装されている。

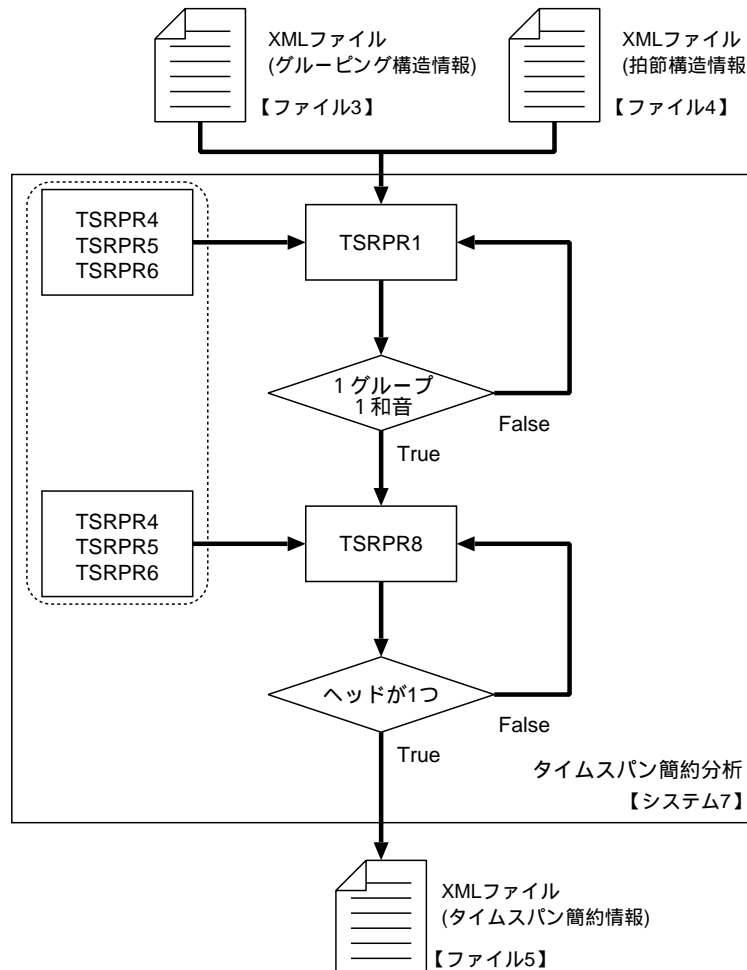


図 4.13: タイムスパン簡約分析アルゴリズム【システム7】

タイムスパン簡約の分析は2つの段階に分けることができる [13] .

第一段階は、グルーピング構造情報と拍節構造情報を入力として1グループ1和音になるまで簡約を繰り返す段階である。この段階では、主に拍節構造の情報を参照する [14] . この段階で必要となる規則には以下のようなものがある。

- TSRPR1
- TSRPR2

- TSRPR3

第二段階は、第一段階で簡約された情報をもとにグループ間において簡約を繰り返す段階である。この段階では和音の情報やカデンツ情報が重要となる。この段階で必要となる規則には以下のようなものがある。

- TSRPR7
- TSRPR8
- TSRPR9

また両段階において参照する規則には以下のようなものがある。

- TSRPR4
- TSRPR5
- TSRPR6

これらの規則は文献 [13] のシミュレーション結果より、省略しても簡約への影響はあまり無いようなので、本システムでは省略した。

以上の事をまとめると、本システムにおける規則の実装状況は以下のとおりである。

実装済 実装済の規則は以下のとおりである。

TSRPR1 拍節構造の選好規則 MPR3 の拍点を数えることにより実装。

TSRPR8 グループ間における簡約では前半部を優先するように実装。

省略 (必要無いと判断) 必要が無い、もしくは実装不可能な規則は以下のとおりである。

TSRPR4 文献 [13] より、省略。

TSRPR5 文献 [13] より、省略。

TSRPR6 文献 [13] より、省略。

未実装 現在検討中のため、未実装なものは以下である。

TSRPR2 和音認識のモジュールを検討中 .

TSRPR3 TSRPR1 との相反関係の扱いを検討中 .

TSRPR7 カデンツ認識のモジュールを検討中 .

TSRPR9 カデンツ認識のモジュールを検討中 .

以上のタイムスパン簡約分析によって出力された【ファイル5】が図 4.14 である .

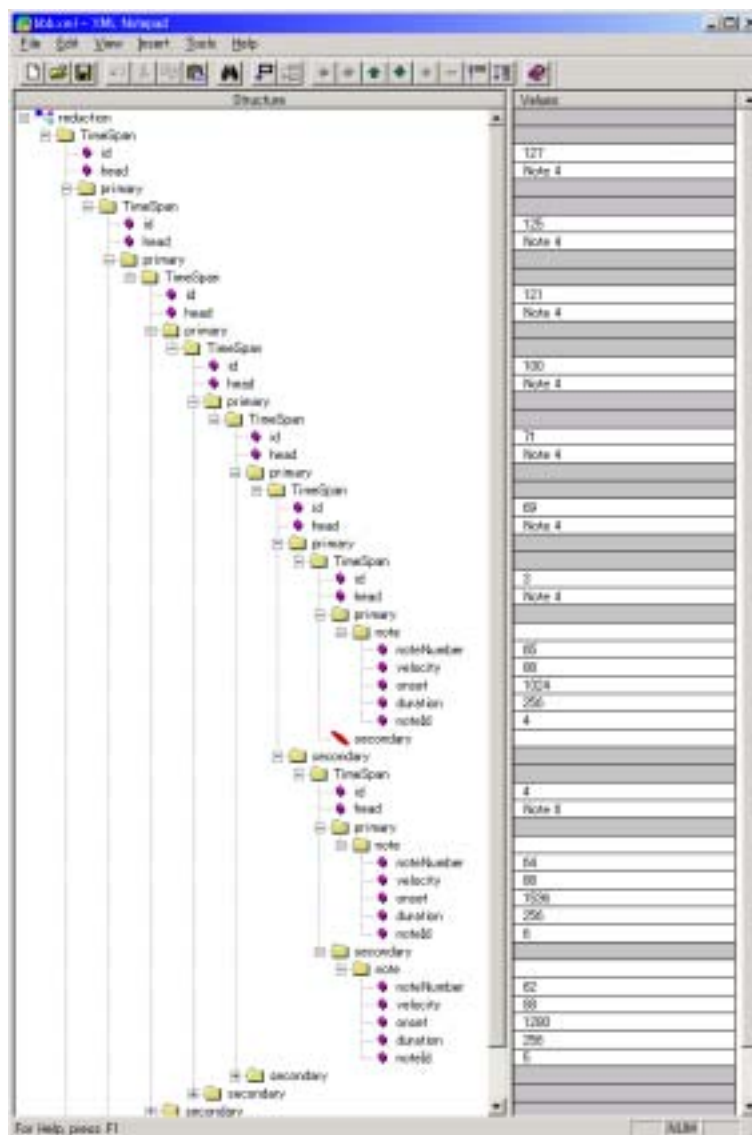


図 4.14: タイムスパン簡約分析の結果【ファイル5】

#### 4.2.8 延長的簡約分析

延長的簡約分析については先行研究の検討の結果，GTTM に基づく楽曲の構造分析を計算器上の実装する際の 1 番の問題点は，タイムスパン簡約分析までの自動化であり，延長的簡約分析についてはすでに実装例がある．このため，今回新たに実装することは見送ったため，文献 [16] を参照して頂きたい．



# 第 5 章

## 実験

本研究で提案したボイスリーディング，第一声部進行，ヒューリスティクスの各手法の有効性を確認するために，文献 [4] や一般的な曲から数曲撰択し，実験をおこなった．本稿ではその中から以下の 2 曲について実際に結果を示しながら検証をおこなう．

表 5.1: 実験に採用した楽曲

テスト曲 1 J.S.Bach / Invention No.1	
入力小節	先頭より 4 小節
最小音符長	16 分音符
最大音符長	4 分音符
音符数	86
テスト曲 2 W.A.Mozart / G Minor Symphony	
入力小節	先頭より 5 小節 (弱起の曲)
最小音符長	8 分音符
最大音符長	4 分音符
音符数	24

テスト曲 1 はよく知られたポリフォニーとして採用し，テスト曲 2 は文献 [4] から採用した．各構造における出力は文献 [4] もしくは専門家による結果を正解とした．以下これらを入力とした各システムの有効性について検証していく．

## 5.1 グルーピング構造

本システムが出力したグルーピング構造を図 5.1, 5.2 に示す．図中の音符間を結んでいる実線はボイスリーディングを表す．五線譜の下に括弧は各グループを表し，括弧上の各数字は各進行上の同一の括弧を表している．

The image shows a musical score for 'Invention' with two staves. The treble staff is marked with a forte 'f' dynamic. Above the treble staff, there are labels for GPR (Grouping Rule) and 3a, 3d, 2a, 2b, 3d. Below the staves, there are red numbers 1 through 15, with brackets indicating groupings. Blue lines connect notes across staves, representing voice leading. The treble staff has a dynamic marking 'f'.

図 5.1: Invention(グルーピング構造)

図 5.1 はテスト曲 1 のグルーピング構造である．図中の最小グルーピングにおいて 1 音を 1 グループとしているグループがあるが，これは GPR1 の実装を省略したからであり，これらは実際のグルーピングに影響を持たないため誤りとはしない．

図から最小グルーピングから 3 レベル上までのグルーピングは，高音部と低音部がそれぞれ分離してグルーピングがおこなわれていることがわかる．また，4 レベル目から高音部と低音部が同一のグループにまとめられていることがわかる．

ボイスリーディングの効果については，高音部と低音部が完全に分かれていることと，小さなレベルにおけるグルーピング規則の補足効果から有効性が見られる．グルーピング規則の補足効果は，高音部 1 小節目の 5 音目と 6 音目の 16 分音符グルーピングからわかる．このグルーピングはボイスリーディング情報からグルーピングがおこなわれてい

る．また高音部 2 小節目の 1 音目がその後のグループから分離してグルーピングされている．これにより，出だしが 16 分音符と 16 分休符と異なっているが平行性がある（同じ音進行）と考えられる高音部の 1 小節目と 2 小節目の出だしの 16 分音符の進行と，低音部の 1 小節目と 2 小節目の後半部の 16 分音符の進行．また高音部 3 小節目と 4 小節目の 16 分音符の進行も，同様の進行が逆さに出てきていることがグルーピングに反映されている．

第一声部進行の効果については，ボイスリーディングが成立していない 2 音間に適用されている規則から確認できる．例を挙げると，高音部 2 小節目の 16 分音符の進行から 8 分音符の進行への変化点に GPR3a,3d が適用されている．これは第一声部進行に対して適用を見た結果である．これがないと，次の 8 分音符 2 つの間と同じ強さでグループの境界候補となってしまう．つまり，進行の変化点より次の 8 分音符 2 つの間でより大きなグループが切れてしまう可能性がでてしまう．

The image shows a musical score for G Minor Symphony with GPR annotations and groupings. The top staff has GPR labels: GPR 2b, GPR 2b, GPR 3a, GPR 3a, GPR 3a,2b,3d, GPR 3a, GPR 2b, GPR 2b, GPR 3a,3d. Below are two staves with red numbers 1-15 and brackets indicating groupings.

図 5.2: G Minor Symphony(グルーピング構造)

図 5.2 はテスト曲 2 のグルーピング構造である．この結果に対しても図 5.1 同様に 1 音を 1 グループとしているところがあるが，同様の理由により誤りとはしない．

図から最小グルーピングから 2 レベル上までのグルーピングは，高音部と低音部がそれぞれ分離してグルーピングがおこなわれていることがわかる．また，3 レベル目から

高音部と低音部が同一のグループにまとめられていることがわかる。

ボイスリーディングの効果については、主旋律を部分的に抽出するかたちになっている。この曲は3小節目中間の4分休符を境に左右に並行性があるが、ボイスリーディングの結果が並行性検出の手がかりになっている。これにより、グルーピングの結果も休符を境に対照的になっている。

第一声部進行の効果については、3小節目中間に位置する4分休符を挟む2音間に適用されている規則から確認できる。この2音間にはボイスリーディングは成立していないが、第一声部進行に対してはGPR2a,2b,3dの3つが成立している。これにより、この2音間は全5小節の中で最も強いグループ境界となっている。

## 5.2 拍節構造

本システムが出力した拍節構造を図5.3,5.4に示す。図中の五線譜の下の各点は、各拍節レベルの構造を表す。横方向に対しては同一レベルの構造を表し、縦方向は同じ音符の各レベルにおける構造を表す。

図 5.3: Invention(拍節構造)

図 5.3 はテスト曲 1 の拍節構造である．本システムにおいて最大の拍節レベルは，楽曲中の最長音符としている．楽曲中の最長音符に最大レベルの拍点を打ち，徐々に小さなレベルへと分析を重ねていく．最小レベルは，楽曲のタクトゥスである．図 5.3 では最長は 4 分音符レベルであり，最小は 16 分音符レベルである．よって中間の 8 分音符レベルを合わせて全部で 3 段の構造が分析される．

拍節構造はタイムスパン簡約をおこなう際に，木構造の最も葉に近い分析からある程度上のレベルまで参照される．しかし，中間レベルから上のレベルに関してはグルーピング情報を参照するので，拍節構造を 1 小節や楽曲全体のレベルまで重ねる必要は無いと判断し，楽曲中の最長音符のレベルまでに限定した．

また，グルーピング構造と拍節構造を相互参照させたことにより，MPR2 を実現している．これは最も強い拍がグループ内で比較的早く現れるという規則である．

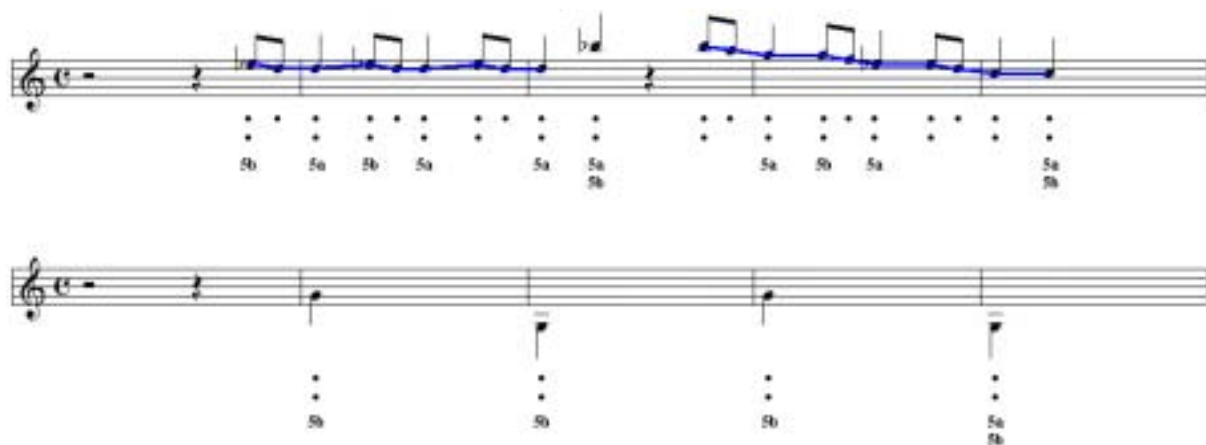


図 5.4: G Minor Symphony(拍節構造)

図 5.4 はテスト曲 2 の拍節構造である．この曲の最長は 4 分音符レベルであり，最小は 8 分音符レベルである．よって全部で 2 段の構造が分析される．

この曲では 8 分音符と 4 分音符の 2 種類しか音符がないため，本システムでは構造が 2 段しか分析できない．このため 5 小節目においては，4 分音符 2 音に対して拍節的な優劣をつけることができない．

### 5.3 タイムスパン簡約

本システムが出力したタイムスパン簡約を図 5.5,5.6 に示す．本システムはポリフォニー的な分析をおこなっているため，同時期に鳴っている音符に関してもそれぞれの木が作られる．このため，木構造のあるレベルにおいて他の木との接続がおこなわれている点がある．それが図中の各数字で書かれている点である．

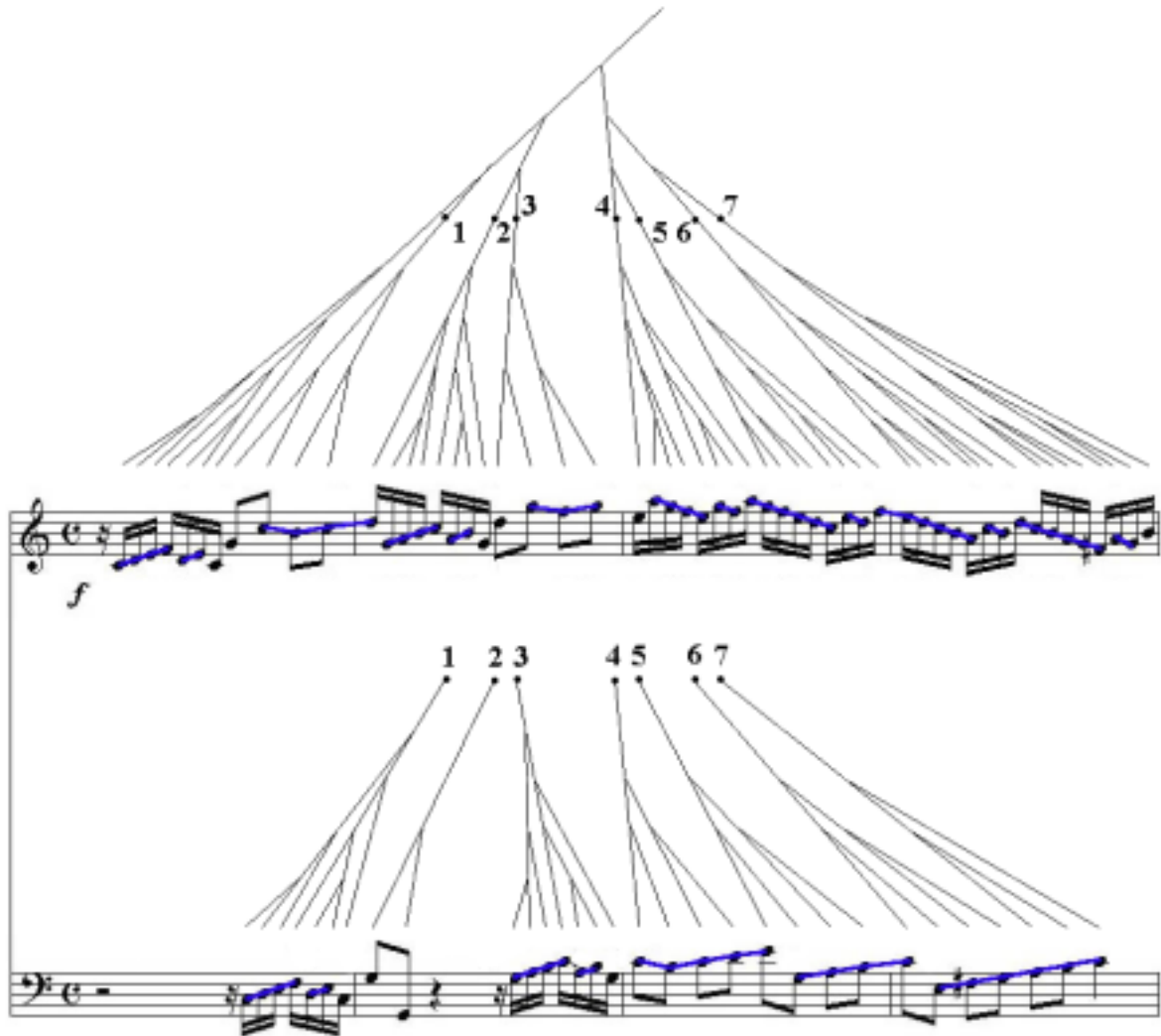


図 5.5: Invention(タイムスパン木)

タイムスパン簡約に関しては，大きく分けて2つの分析がある [1]．最初の分析は，拍節構造を参照して和音レベルのグループまで木をまとめる段階である．この段階につい

ては、TSRPR1により実現している．しかし、実際には拍節構造の参照だけでは足りず、和声の構造についても考慮しなければならないことが結果から見て取れる．高音部 1 小節目の初めの 16 分音符 3 音について見てみると、最初に C と D が簡約され、D が優先される．その後 D と E が簡約され D が優先される．しかしこの 3 音を和音的に見てみると、CDE なので I の和音 CEG に D が入っていると認識する方が正しいと考える．しかし和声の認識ができていない現時点では拍節的に安定した D が優先されてしまう．

第 2 段階では、グループ間において簡約をおこなう．この段階ではカデンツの分析をおこない、その情報を参照して規則を適用し簡約をおこなう．しかし、現時点ではカデンツの分析を実装していないため、幾どの規則が未実装である．このため正解例とはほとんど一致しない木構造になってしまっている．カデンツの分析を実装するために、前段階でも必要となる和音の解析をおこなう必要があると思われる．

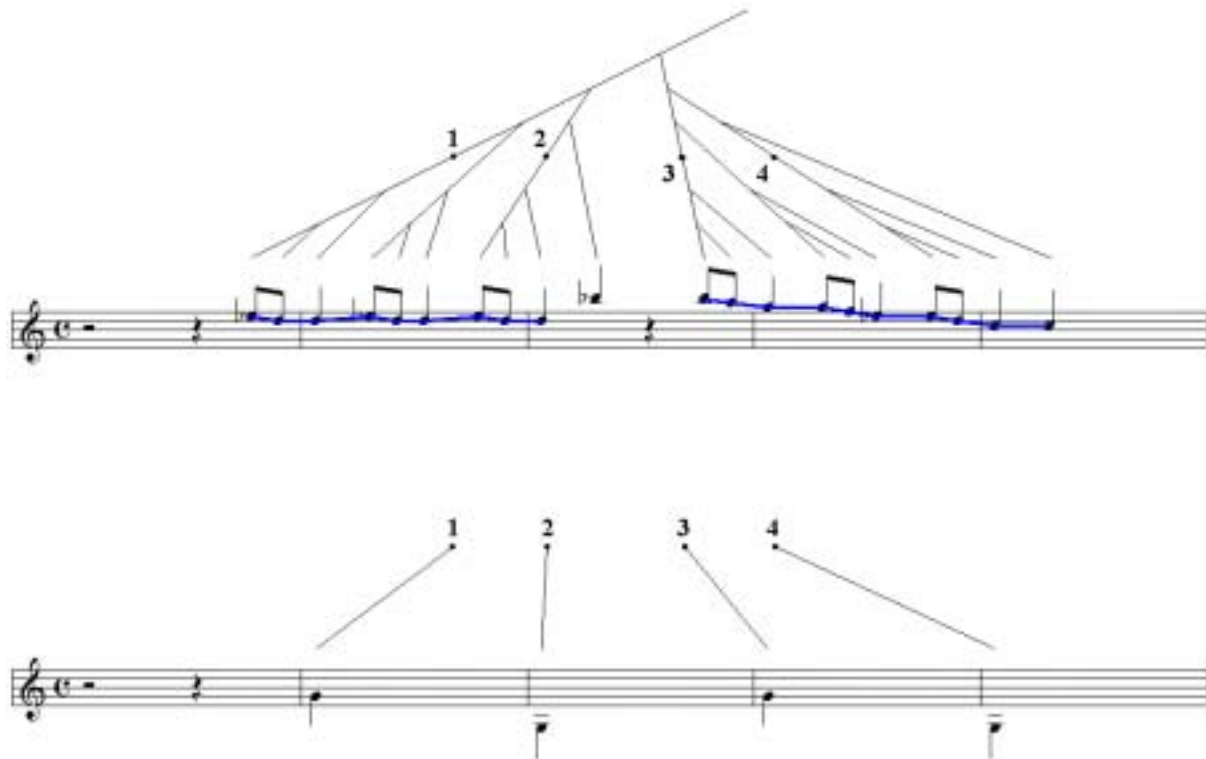


図 5.6: G Minor Symphony(タイムスパン木)

図 5.6 のテスト曲 2 に対しても，和音の解析，カデンツの分析がおこなわれていないために木構造としては GTTM の意図している構造とかけ離れた結果になってしまった．今回の検証にはこれら 2 曲以外の楽曲も入力して実験をおこなってみたが，やはりタイムスパン木に関しては良い結果が得られなかった．今回の検証により，ボイスリーディング，第一声部進行に関しては十分な有効性が確認できた．またこれらの拡張によって，グルーピング構造分析，拍節構造分析の両分析は，十分な結果が得られることが確認できた．しかし，タイムスパン簡約に関しては，実際には GTTM の規則だけでは不足で，GTTM の理論 [4] にはスッポリ抜けてしまっている和声に関する規則及び分析が，実際にタイムスパン簡約をおこなう際には必要不可欠であることが確認できた．



## 第 6 章

### まとめと今後の課題

本研究では，楽曲の構造を自動分析するシステムを提案した．本システムでは楽曲を分析する理論として，各規則が箇条書きされていて最も計算機上にも実装しやすいと考えられる Generative Theory of Tonal Music を用いた．しかし，この理論は人間を相手に分析結果を説明することを前提としている理論であるため，規則が厳密に定義されていないなどの理由により規則定義の曖昧性や，規則適用の非一貫性などの問題が生じてしまう．またこれらの問題点があるために，これまでの研究でも純粋に GTTM を実装して，楽曲の構造を表すタイムスパン簡約までを自動で分析するシステムは構築できずにいた．

本研究では当初延長的簡約までを目的にしていたが，先行研究の調査と共同研究者との打ち合せの結果，延長的簡約は先行研究 [16] で結果が出ているため，これまでの研究で困難とされているタイムスパン簡約までを自動化することを目的とした．

本研究で提案するシステムでは，これらの問題点の解決法としてボイスリーディング，ヒューリスティクス，暗黙的パラメータの明示化，第一声部進行，並列計算の試みを提案した．ボイスリーディングは独立に存在する音符間に旋律的な繋がりを見いだすための構造であり，GTTM における各規則の適用対称の明示化や，GTTM の規則では不足している各旋律を繋げるための規則を補う効果がある．またボイスリーディングを採用することにより，楽曲が部分的に繰り返されていると思われる楽曲の並行性を見いだす手がかりになることが期待できる．ヒューリスティクス，暗黙的パラメータの明示化は GTTM 特有の規則定義の曖昧性を解消する目的で採用した．各々のパラメータは GTTM の文献 [4] から例を 20 曲選択し，調整した結果を用いた．第一声部進行は主旋律の抽出

を意図して採用した。主旋律は最も強い音の進行であると考えられる。第一声部進行を採用することによって、ボイスリーディング間に対する GTTM の規則適用や、楽曲全体を通して規則を適用する対称を特定できる。並列計算の試みは GTTM の各段階の構造、簡約間の解のフィードバックを実現するために採用した。各構造、簡約において考えられる解候補を全て出力し、次の段階における分析で最も適切な解に絞りこむことにより、解のフィードバックを実現している。この考え方は、楽曲に対する分析においては自然の考え方である。音楽は、それを聴いた人によって解釈が異なる。よって、楽曲を分析した結果が一意に求まるということは不自然である。

本研究が提案するシステムは、GTTM に以上の拡張を行い実装を試みた。GTTM による自動分析システムの概要は次のとおりである。

入力された楽曲に対しボイスリーディング、第一声部進行を求め、これらに対し GTTM のグルーピング構造の各規則適用をおこなう。グルーピング構造と拍節構造は同時進行で相互に参照しながら分析をおこなう。その結果得られた構造をタイムスパン簡約分析モジュールの入力とする。タイムスパン簡約によって楽曲の構造を木構造として分析をおこなう。各中間出力は XML ファイルとし、本システムの有効性を検証した。

検証の結果、ボイスリーディング、第一声部進行に関してはその有効性を確認できた。ボイスリーディングは、当初予測していた規則適用に関する曖昧性の解決や、GTTM の規則には欠けている音符間のつながりを見いだす規則としての効果以外に、楽曲の並行性を見いだすための手がかりとなる効果まで確認できた。第一声部進行は、ボイスリーディングが成立しない音符間に対する規則適用からその有効性が確認できた。並列計算の試みに関しては、現在出力側としては有効な解候補すべてを求める実装状態であるが、入力する側がこれに対応するように実装していないため、今回有効性の検証は見送った。

システム全体の有効性としては、グルーピング構造、拍節構造の両構造分析は GTTM の文献 [4] と専門家による人手の分析結果との検証をおこなった結果、その十分な有効性が確認できた。タイムスパン簡約に関しては、現システムでは十分な結果が得られないことが確認できた。その理由としては、GTTM の理論には書かれていない和声に関する分析が欠けているため、システムに実装できる規則が少ないことが挙げられる。これに関しては、GTTM 以外の和声分析の理論を拡張する必要があると考えられる。

今後の課題としては、GTTM の理論には欠けている和声分析の理論を現システムに拡

張し，和声の情報を参照するタイムスパン簡約の規則の実装をおこない，木構造の精度を上げる事が挙げられる．また，並列計算に対しては入力側の拡張を行ない，より上位の構造において解候補を絞りこむ実装が挙げられる．

これらの拡張をおこなった結果，十分に有効なタイムスパン木が得られることを確認し，既存の延長的簡約分析のモジュール [16] への入力をおこない，その結果を検証することが考えられる．これにより，GTTM の理論に基づいた楽曲分析の完全自動化がおこなわれることを期待する．

最後に本研究の成果が，これからの GTTM を用いた楽曲の構造分析に関する研究の有力な先行研究となることを期待する．

# 謝辞

本研究を進めるにあたり，日頃から方針，内容についてご助言，ご指導賜りました東条敏教授，に厚くお礼申し上げます．

そして，研究の全般にわたって終始細かなご助言を頂いた平田圭二氏 (NTT コミュニケーション科学基礎研究所) に深く感謝いたします．

また日頃からご協力，ご助言をいただきました音楽知研究会の佐藤健教授 (国立情報学研究所) に心から感謝いたします．

本研究にご理解と多大なご協力を賜った，鳥澤健太郎助教授，永田裕一助手をはじめとする東条研究室，鳥澤研究室の皆様感謝いたします．

## 参考文献

- [1] Alan Ruttenberg, Review and Discussion of A Generative Theory of Tonal Music, <http://alanr.www.media.mit.edu/people/alanr/Jackendoff&LerdahlFinal.html>, 1994.
- [2] 東和信, 森博彦, 小杉信 計算器を用いた旋律の規則に関する研究, 情報処理学会 音楽情報科学研究会 研究報告, No.42, pp.1-6, 2001.
- [3] Curtis Roads, コンピュータ音楽 -歴史・テクノロジー・アート- 東京電機大学出版局, 2001.
- [4] Fred Lerdahl and Ray Jackendoff, A Generative Theory of Tonal Music, The MIT Press, 1983.
- [5] Gerd Castan, Michael Good, and Perry Roland, Extensible Markup Language(XML) for Music Applications: An Introduction COMPUTING IN MUSICOLOGY, No.12, pp.95-102, 1999-2000.
- [6] 平賀譲, 音楽認知研究の諸問題, 情報処理学会 音楽情報科学研究会 研究報告, No.6, pp.15-22, 1994.
- [7] 平田圭二, 青柳龍也, パーピース: 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, No.37, pp.17-23, 2000.
- [8] 平田圭二, 青柳龍也, パーピース: 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 音楽情報科学研究会 研究報告, No.31, pp.7-12, 1999.

- [9] 平田圭二, 青柳龍也, 音楽理論 GTTM に基づく多声音楽の表現手法と基本演算, 情報処理学会論文誌, 2002.
- [10] 池内友次郎, 長谷川良夫, 他和声 “理論と実習 I”, 音楽之友社, 1964.
- [11] 片寄晴弘, 竹内好宏, 演奏解釈の音楽理論とその応用について, 情報処理学会 音楽情報科学研究会 研究報告, No.7, pp.15-22, 1994.
- [12] MAURO BOTELHO, TONAL GROUPING:AN ADDENDUM TO LERDAHL AND JACKENDOFF'S “A GENERATIVE THEORY OF TONAL MUSIC”, International Conference for Music Perception and Cognition(ICMPC), No.3, pp.265-266, 1994.
- [13] Péter Halász, COMPUTER SIMULATION OF TIME-SPAN REDUCTION, International Conference for Music Perception and Cognition(ICMPC), No.3, pp.267-268, 1994.
- [14] 竹内好宏, 音楽の構造解析とその応用, コンピュータと音楽の世界, 共立出版, 1999.
- [15] 東条敏, ヘッドの概念を用いた和声学の文法, 情報処理学会 音楽情報科学研究会 研究報告, No.42, pp.55-60, 2001.
- [16] 上符裕一, コーパスを用いた音楽構造解析, 大阪大学 基礎工学部 修士論文, 1998.
- [17] 鷺坂光一, 標準 MIDI ファイルからのメロディの自動抽出, 情報処理学会 音楽情報科学研究会 研究報告, No.5, pp.7-12, 1994.
- [18] 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.