

## 自動作曲システム *Orpheus* とその著作権性

中野拓帆<sup>†1</sup> 深山 覚<sup>†1</sup> 嵯峨山 茂樹<sup>†1</sup>

本稿では、まず、自動作曲システム *Orpheus* について紹介する。次に、*Orpheus* のような自動生成手法に関連する、著作権および、著作隣接権等に関するいくつかの問題提起と考察を行う。*Orpheus* は、任意の日本語テキストの持つ韻律に基づき、歌唱曲を自動作曲するシステムであり、<http://hil.t.u-tokyo.ac.jp/~orpheus/> で試用できる。歌唱曲は歌詞との関連性が求められ、特に高低アクセントを持つ日本語では、歌詞を朗読する際の韻律と旋律が一致することが重要とされる。旋律設計を、旋律を音高間を遷移する経路とし、韻律の上下動の制限の下での探索問題と捉えた。文学作品や自作の詩、ニュースなど、あらゆる日本語テキストをそのまま歌詞として旋律を生成し、歌唱曲として出力できる。手軽な作曲のツール、音楽の専門知識を持たない人のための作曲補助ツールとして有用であろう。このような自動生成手法は、著作権問題を回避できる可能性がある他、著作権に関する新しい問題を引き起こさう。

### Automatic Composition System *Orpheus* and Copyright-related Issues

TAKUHO NAKANO,<sup>†1</sup> SATORU FUKAYAMA<sup>†1</sup>  
and SHIGEKI SAGAYAMA<sup>†1</sup>

In this paper, first, we describe *Orpheus*, an automatic music composition system, second, we will discuss copyright-related issues in such systems for automatic generation of intellectual products. *Orpheus* is a system to automatically generate songs from Japanese lyrics considering the prosody of the lyrics. A prototype is available at <http://hil.t.u-tokyo.ac.jp/~orpheus/>. Generally in songs, the lyrics and melodies should be closely related. Especially in Japanese, which is a pitch-accented language, consistent prosody of lyrics is considered to be important across in speaking and in singing. The algorithm is designed by considering composition as an optimal-path search problem under constraints given by the prosody of the lyrics. It can compose a song from any Japanese text (e.g., literature, one's own poems, news, etc.). It is expected to be convenient as an automatic tool or supplementary aid for music composition by amateurs without expert knowledge for music composition. Such automatic generation systems may make it possible to avoid copyright problems and also

bring new copyright-related issues.

#### 1. はじめに

コンピュータに、人間のように歌の作曲をさせることは可能だろうか。可能であれば、誕生日に詩を作って作曲して贈り物にしたり、学校で脚本を作って作曲してミュージカルを作ったり、記憶したいことを歌にして記憶したり、歌詞を募集して作曲は自動で行ったり、著作権の問題のないオリジナル曲を好きなどころで使えたり、さまざまな利用が考えられる。

以下、本稿では著者ら自身の文献 1) に基づいて技術解説し、その後に著作権に関する議論を行う。

コンピュータを用いた作曲は、現在まで様々な手法によってなされてきた。これらは、新たな音楽の創造という「芸術」的な目的を持つものと、人間が、ある形式や理論に基づいて行う作曲の自動化という「工学」的的目的を持つものに分類できる。

前者は、「アルゴリズム作曲」と呼ばれることが多い。これらは、新たな作曲手法(アルゴリズム)の提案により、既存の曲とはまったく異なる音楽の創造を目指している。例えば、初めてコンピュータを用いた作曲を行った「イリアック組曲」、確率論に基づく作曲プログラムを用いて作曲された“Eonta”(I. Xenakis, 1964) などがある。コンピュータのない時代にも、モーツァルト (1756–1791) はいくつかの旋律をさいころを振った組み合わせで決定し、「音楽のさいころ遊び」という曲を作っている。

一方、工学的目的の研究では、既存の音楽理論や楽曲に基づき、人間の作曲過程を自動化する。例えば、遺伝的アルゴリズムを用いたジャズのアドリブソロの自動生成<sup>2)</sup>、「対位法」に基づいた楽曲生成、「和声法」を用いた自動和声づけシステム、コラール生成システムなどが提案されている。本稿は工学の立場から歌唱曲作曲を捉えるので、議論する自動作曲システムの狙いは、名曲やオリジナリティのある曲の作曲ではなく、工学的に捉えることが可能な音楽の要素により、人間が聴いて不自然でない曲、音楽理論的に妥当な曲の作曲にある。

これらに対し、本稿では、歌詞に基づく歌唱曲作曲をモデル化し、これに基づき任意の日本語テキストを歌詞とする歌唱曲を自動作曲する手法について解説し、著者らが開発した自

---

<sup>†1</sup> 東京大学  
The University of Tokyo

動作曲システム *Orpheus* を紹介する。続いて、このような自動生成システムによって生じる、著作権に関する新しい問題について、提起、整理、考察を行う。例えば、自動生成された楽曲の権利は誰のものだろうか。

## 2. 歌唱曲旋律への制約条件

### 2.1 音楽の要素

音楽は一般的に、和声、リズム、旋律の3要素から成り立つとされる。本稿で扱う歌唱曲には通常、歌に伴奏が付くので、歌唱曲の構成要素を、和声、リズム、旋律、およびこれを支える伴奏の4要素とする。これら4つの要素は互いに関連があり、旋律とリズム、旋律と和声、和声(と旋律)と伴奏はそれぞれ密接に関連するが、本稿では旋律を生成する以前の和声進行とリズムと伴奏音型は独立に設計可能と仮定する。

この仮定は音楽理論でも暗黙のうちに認められているようである。和声学では、基本的にリズムを扱わない、すなわちリズムの任意性を認めている。また、和声進行と、その分散和音などの伴奏音型への展開とに関して、相互の制約の議論は基本的にされない。

### 2.2 歌詞からの制約：音符数

まず、旋律には音符はいくつ必要になるだろうか。日本語はモーラ言語—俳句の五・七・五として数えるのがモーラ—である。モーラ (mora) とは音韻論の用語で、一定の時間的長さをもった音の分節単位とされる。モーラ言語ではない英語には、「おおさか」と「おさか」の区別がなく、/osaka/と綴る。「来た」/kita/、「聞いた」/kiita/、「切った」/kiQta/の区別も難しく、物理的に真似はできても、別の語であるという意識は持てない。似た概念である音節はどの言語にもあるが、音素の配列の問題であり時間の単位ではない。

### 2.3 リズム設計：リズム木仮説

原則的に、各モーラには少なくとも1つずつ音符が必要である。では、各音符の音価は任意かというところ、そうではない。音価とは、二分音符、付点四分音符といった楽譜上の音符の長さで、その連鎖はリズムを形成する。リズムは曲想を強く表す要素である一方、自由度も高い。特に、日本語の場合は、強弱アクセントを持たず、各モーラがほぼ等時的に発話されるので、リズムの制約は強くない。強弱アクセントを持つ英語では、アクセント位置と強拍位置の一致がよくみられる。

歌の1番と2番の同じ場所で、音符数が違うためにリズムが異なっても同じ曲想を感じることがある。この2つのリズムは音符数によらずに曲想に与える同一の特徴を持っていると考えられる。本稿では、この音符数によらない特徴を「リズムパターン」、同じリズムパ

タンを持つ音符数の異なるリズムの集合を「リズム木」と呼ぶ。そして、一定のリズムパターンの下で、あるリズムが異なる音符数に展開される過程を、以下のように仮定する。

まず、あるリズムパターンから、そのリズムパターンを最もよく実現するリズムが生成される。このリズムを、「標準リズム」と呼ぶ。また、一定のリズムパターンの下で、あるリズムが異なる音符数に展開できる構造として「リズム木構造」を仮定する。リズム木構造に従い、図1のように標準リズムから音価の分割、統合が行われ、一定のリズムパターンを持つ異なる音符数のリズムが形成される。これら異なる音符数のリズムの集合が、リズム木となる。

以上の仮説に確率モデルを導入することで、歌唱曲作曲におけるリズム設計は、曲想を基に確率的に生成されたリズムパターンからリズム木が作られ、歌詞により決まる音符数に合わせ、リズム木から使用するリズムが決定される、とモデル化できる。リズムが決定することで、既に決定している和声と合わせ、旋律音が属する和音が決定する。

### 2.4 歌詞からの制約：抑揚

音符を並べれば旋律になる。そこに言語的制約はないだろうか。音韻論で韻律とは、あるテキストを発話するときの音声の、音の高さ(ピッチ)、強さ、長さ、区切り、息継ぎなど、音色以外の幅広い特徴を指す。しかし、これらの特徴のうち韻律的特徴を強く伝えるものは、主にピッチの変化である。よって、本稿では単に韻律と言った場合、テキストのアクセント句、アクセント型情報を指すものとする。

日本語はピッチアクセントの言語である。言いかえると、アクセントは語のピッチ変化としてあらわれる。長いテキストを発話するときは、テキストをいくつかのグループに分割し、そのグループにイントネーションをつけて発話する。このグループをアクセント句と呼ぶ。

関東方言(いわゆる標準語)では、一般に  $m$  モーラの単語あるいはアクセント句には  $0$  型から  $m$  型までの  $m+1$  種類のアクセント型があり得る。 $0$  型はピッチの下降のないものをさす。他は、ピッチの下降の直前のモーラ(アクセント核)が何番目であるかにより、 $m$  型と称する。例えば「端」「箸」「橋」はいずれも2モーラの語で、発音は「はし」であるが、アクセント型はそれぞれ  $0$  型、 $1$  型、 $2$  型である。「端」と「橋」の相違は、それぞれに助詞「を」をつけて発音すると、 $\text{ハ}|シヲ$ 、 $\text{ハ}|シヲ$ 、 $\text{ハ}|シ|ヲ$  となることからわかるだろう。

例えば「箸の端が橋から伸びた」というテキストは、 $\text{ハ}|シノ}$   $\text{ハ}|シガ}$   $\text{ハ}|シ|カラ}$   $\text{フ}|ビタ}$  とアクセント句に分割され、イントネーションがつけられる。図2に示すような曲があるとすれば、イントネーションと旋律が一致していることが感じられるだろう。歌唱曲の作曲では、日本語歌詞のイントネーションと旋律が一致することが要求されてきた。特に、一度きりの作曲、一度きりの聴取を想定する場合等では、韻律旋律一致の原則は有効だと思わ

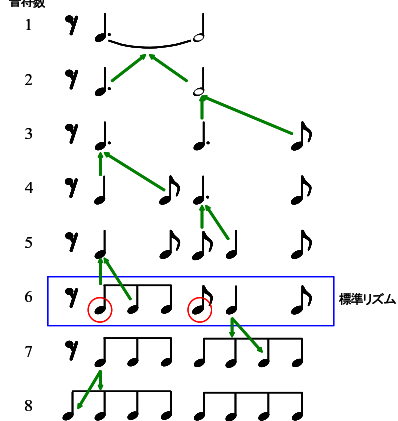


図 1 リズム木構造による標準リズムからのリズム木の形成。標準リズム中で丸で示した位置がリズムパターンを表現する。標準リズムがリズム木構造に従って音符数に応じて展開され、一定のリズムパターンを持つリズム木をなす。

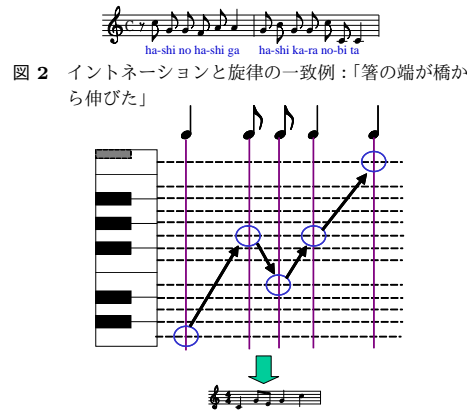


図 2 イントネーションと旋律の一致例：「箸の端が橋から伸びた」

図 3 旋律経路の例。1 オクターブ内で旋律を設計すると、音程候補は 13 音になる。図上部のようにリズムが与えられると、音程候補の選択タイミングが決定する。音程とリズムから形成される格子点上を図中のように遷移する経路は、図下部に示した旋律となる。

れる。本稿ではこの原則を守ることにする。つまり、音高(ピッチ)の上下動と、アクセントによる歌詞発話時のピッチの上下動を一致させる必要が生じる。旋律設計は、韻律の上下動を守りながら、曲想に応じてどの音域を使うか、どの程度の跳躍をするか、などを考慮して進められる。この際、音程だけでなく音価も同時に考えられる。強弱アクセント言語を歌詞とする場合と比べて、歌詞が日本語の場合はリズムの制約は弱い、歌詞のモーラ数により音符数は決まるので、器楽曲に比べればリズム設計の自由度はかなり小さくなる。こうした制約の中で、曲想に合わせて静かなリズム、躍動的なリズム、などを決定する。

### 2.5 その他の制約

これまで述べてきた以外にも、和声構造、非和声音、伴奏、音域と跳躍などの制約が考えられる。全て曲想、旋律に関わるが、特に前者は主に音楽理論による禁則条件などが存在する。詳細は省略するが、和声進行はユーザがライブラリから選択、非和声音にはひとまず経過音、補助音、倚音を認める、とする。伴奏、旋律間には音楽理論に基づく禁則を犯さないようにするという制約が与えられる。音域や跳躍の度数は、和声、リズム、伴奏と同様に曲想に基づき確率的に決定されるものであるが、和声、リズムと異なり、旋律の音程が決定して初めて実現される。音程の出現確率(音域)、音程の遷移確率(跳躍度数の起こりやすさ)の2つの確率により、旋律設計に曲想を表現する音域、跳躍の度数が反映される。

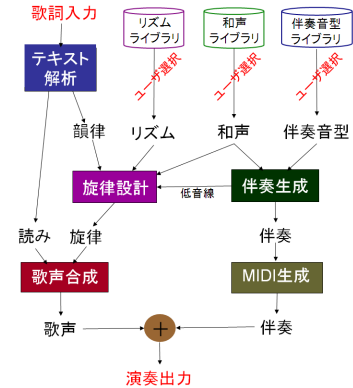


図 4 Orpheus のシステム構成

### 2.6 動的計画法を用いた旋律設計

以上の議論から、曲想を基にユーザが和声、リズム、伴奏音型を独立に選択し、そこから旋律および伴奏を設計するという順序で歌唱曲作曲のプロセスが考えられる。旋律の声域と、旋律のリズムが与えられると、旋律が取りうる音程の候補と、その候補の中から使用する音程を選択するタイミングが決定される。このとき、旋律は図3に示されるように、音程とリズムからなる格子点上をリズムの進行に従って遷移する経路と捉えることができる。

この中から確率最大の経路を探索する問題は、作曲家達が以上のさまざまな制約を守り、歌詞の韻律を最大限活かして作曲を行うならば、最も高い確率で作られる曲はどのような曲か、と喩えられる。確率最大経路探索問題の解を、単純な全探索で求めようとする、 $N$ 個の音符がそれぞれ取りうる音高候補が  $M$  個のとき、 $M^N$  通りの経路について確率の積を比較する必要があるが、動的計画法を用いれば  $MN$  に比例する計算量で解くことができる。

### 3. 自動作曲システム Orpheus

自動作曲システム Orpheus は、以上に述べた作曲の原理に基づき、動作するシステムである。処理の流れを、図4に示す。ユーザは、漢字かな混じりの日本語テキストを入力するとともに、自分が望む曲想から、(1) リズム木、(2) 和声進行、(3) 伴奏音型、(4) 伴奏音色、(5) 合成音声の音色(男/女)、を選択肢から選択する。

システムは、テキスト解析、リズムの決定、伴奏生成、旋律設計、統合出力と音響出力を行い、楽曲生成を行う。詳細は、文献4)、5)を参照されたい。テキスト解析部では、入力された日本語テキストを解析し、韻律情報を得る。さらに、解析されたテキストは、2小節を

単位とする節に分割される。よって、各節の音符数が決定し、これを基にユーザが選択したリズム木から音符数が合致するリズムパターンを抽出する。リズム木の一例は図1に示したが、節の含む音符数に応じたリズムが記述されている。リズム木は、人手で作成したものを用いているが、音楽データからの自動生成も可能であり、今後の研究テーマである<sup>3)</sup>。伴奏は、和声、伴奏音型を用いて生成される。非和声音と和声学については本稿では省略する。

旋律設計において、本システムで最も重要となるのは韻律と旋律の上下動の関係である。リズムの決定から、音程以外の旋律要素が決定され、各音符の遷移の方向が韻律に従って決定される。ピッチ変化が起きる位置では、上行または下降音型、ピッチがあまり変化しない部分やアクセント句の境界部分では、平行音型としている。平行音型では、発話において不自然でないと思われる程度、ピッチの上下動を許し、上下動の方向をアクセント句内の位置により制限した。音域は、中心を定め、中心から離れる程その音の出現確率が段階的に小さくなるような確率を設定した。跳躍音程の制御は、跳躍、順次進行のどちらが多いかをユーザが選択するようにした。これら制約および確率の下で、2.6節で述べたように動的計画法を用いて旋律を設計する。以上により生成された旋律および伴奏は、楽譜出力とMIDI出力され、歌声合成システム<sup>6)</sup>を用いた旋律の歌声と伴奏を合成し自動演奏出力を行う。

#### 4. 著作権の問題

さて、Orpheusのような自動生成手法を用いると、文化的な作品を、比較的容易に作成することができる。実はこの事は、著作権に関連する新しい問題を生む可能性がある。例えば、既存の楽曲を和声・リズム・旋律といった要素に分解した後、再び合成を行う際の問題、システムの生成物であることが伏せて発表された作品の問題、作品を大量生成し、著作権を主張できた場合、他の人が同じ曲を生成した場合の問題などが考えられる。

分解と再合成にあたり、著作権はどうなるであろうか。非常に極端な例ではあるが、パッヘルベルのカノンの和声進行はよく「再合成」のように利用されるなど、多くの作曲家が、知ると知らざるとのうちに用いている。和声進行だけを取り出しても、特徴は十分にあり、和声進行には何らかの個性が現われているといえる可能性がある(著作権の発生する可能性)。つまり、他人の曲の和声進行を取り出し、用いてはいけない可能性がある。

次の問題は、どうやって「著作物」と判定するか、という問題と言えるかもしれない。著作権の切れた詩から、自動作曲を行うとしよう。「リズムや和声進行、伴奏をよく選んで」できた曲は、創造的活動によって生み出された、と考えてよさそうである。しかし、たまたま「適当リズムや和声進行、伴奏」からできた曲はどうであろうか、システムを用いた人は創

作したと言えるのか。また、作曲のパラメータとともに公開された曲があったとき、他の人がそのパラメータを用いて(同一の)曲を得た場合、著作権は主張できるのか、公開した人は、手法は公開したものの、作品を自由に使っていいと言った訳ではない。作品が「独自創作」されたものであれば、権利侵害ではないが、どうなるだろうか。

機械による著作であることを伏せて発表した場合はどうであろうか。ある人が、入力日本語テキスト、和声やリズム等の選択(いずれも、自由に使えるものだったとしよう)を、大量の組み合わせについて行い、公開したとしよう。別の人が、その中から「この曲はいい」と思った曲を見つけた場合、権利が付与されるのは、また、「創作」したのは一体どちらか。意味もなく大量公開しただけの前者に、後者が許諾を得ることは必要なのだろうか。

#### 5. おわりに

本稿では、歌詞の韻律を旋律設計に反映した歌唱曲の自動作曲手法と、それを実装したシステムについて述べ、さらに、自動生成手法の利用により生じうる著作権に関連する問題を提起した。Orpheusのような自動作曲システムは、誰でも手軽に作曲ができるという点で、作曲行為の敷居を大きく下げるといってよいだろう。本稿ではいくつかの問題を指摘するとどめたが、一般の自動生成手法の利用に際しどのような問題が生じうるかを調査、議論することは今後の課題と言えよう。

#### 参考文献

- 1) 嵯峨山茂樹, 中妻 啓, 深山 覚, 酒向慎司, 西本卓也: 日本語歌詞からの自動作曲 (<特集> 音楽とOR), オペレーションズ・リサーチ: 経営の科学, Vol.54, No.9, pp. 546-553 (2009).
- 2) Biles, J.A.: Genjam: A Genetic Algorithm for Generating Jazz Solos, *Proceeding of 1994 International Computer Music Conference* (1994).
- 3) 中野拓帆, 深山 覚, 小野順貴, 嵯峨山茂樹: リズム木構造仮説に基づく楽曲解析とリズム木学習, 日本音響学会春季研究発表会講演論文集, to appear (2010).
- 4) 中妻 啓, 酒向慎司, 小野順貴, 嵯峨山茂樹: 歌詞の韻律を用いた自動作曲, 日本音響学会春季研究発表会講演論文集, pp.739-740 (2007).
- 5) 深山 覚, 中妻 啓, 米林裕一郎, 酒向慎司, 西本卓也, 小野順貴, 嵯峨山茂樹: Orpheus: 歌詞の韻律に基づいた自動作曲システム, 情報処理学会研究報告. [音楽情報科学], Vol.2008, No.78, pp.179-184 (2008).
- 6) 酒向慎司, 宮島千代美, 徳田恵一, 北村 正: 隠れマルコフモデルに基づいた歌声合成システム, 情報処理学会論文誌, Vol.45, No.3, pp.719-727 (2004).