

音楽生成行為のモデルとしての自動作曲

古川 聖 藤井晴行

東京芸術大学先端芸術表現学科 東京工業大学大学院建築学専攻

作曲家とインタラクションをしながら学習する自動作曲システムについて述べる。音楽の多義性を扱える作曲システムを制作することがねらいのひとつである。提案システムは音楽の事例を分析して得られた音楽の構成自体の生成文法に基づくのではなく、モチーフを変形する操作の列の構造に着目する点が従来のシステムと異なる。システムは制作途上にある。制作の初期段階として行っている実験からこれまでに得られた知見を示す。

Machine Composition as a Representation of a Model of a Course of Actions in Music Generation

Kiyoshi FURUKAWA & Haruyuki FUJII

Tokyo National University of Fine Arts and Music & Tokyo Institute of Technology

This paper describes the framework of an interactive system for music composition. One of the main target is to make an machine composition system that can deal with the ambiguity in the structure of music. The proposed system differs from the existing system in the sense that the former generate a tune through a course of modifications of a basic motif while the latter relies on a generative grammar constructed by analyzing existing tunes. The proposal system is under development. The findings acquired through the pilot experiments performed as the first stage of the system development.

1. はじめに

筆者らはデザインするという行為に着目し、音楽や空間デザインなど、異種の芸術表現形態におけるデザイン行為間に共通する抽象構造を探究している。作曲は広い意味でのデザインの一クラスであると考えている。研究の最終目的は種々の芸術表現におけるデザイン行為の構造を明らかにし、デザイン行為に関わる知識を埋め込んだデザイン・システムを構築することである。このデザイン・システムは芸術表現を自動的に生成するものではなく、むしろ、作曲家や空間デザイナーなどのパートナーとなりうる、インタラクティブなものである。

本稿では、上記の目的のために制作中の自動作曲システムについて報告する。

自動作曲・編曲システムは次の2種類に大別される[Katayose 2000]。ひとつはシステム開発者が恣意的に与えた作曲に関する規則に基づいて作曲を行うシステム、ひとつは音楽作品から抽出した作曲に関するスタイルや規則に基づいて作曲を行うシステムである。前者の代表例である Music-In-A-Box は入力されたコード進行と選択されたスタイル(音楽ジャンル)からデータベース内のルールに基づいて伴奏を生成するシステムである。後者の代表例のひとつである EMI は事例から抽出したルールを用いて組み立てた楽曲の構成に、同じく事例から抽出したモチーフを埋め込んで音楽を生成するシステムである。これらはシステムとしての実現の興味

から開発されている。筆者らはシステム実現の興味に加え、作曲プロセスの構造の解明を視野に入れている。予め分析された楽曲の構成を表す規則をシステムに与え、その規則を用いて楽曲を生成するのではなく、モチーフを変形する操作の構造に注目して楽曲を生成する点がEMIとは異なる。

2. システムの概要

2.1 前提

提案するシステムが前提とすることを述べる。

先に触れた作曲の不可逆性、すなわち、楽曲の事例の分析プロセスを逆方向に適用したとしても作曲プロセスにはならないという性質は、作曲技法の組み合わせや適用順序に関する作曲家の知（暗黙知と形式知）が楽曲を分析することによっては獲得されないことによると筆者らは考えている。例えば、ある音を強調するためにシンコペーションを採用するといった技法は事例となる楽曲の音程、リズム、音の強弱を分離して抽出した生成文法では表現できない。筆者らはこのことを踏まえ、作曲に関わる知は作曲技法の適用順序に埋め込まれているとの仮定に基づき、音程、リズム、音の強弱を切り離さずになされる実際の作曲技法に即した変形操作の列をモチーフに施すことによって作曲を行うシステムを制作し、システムとのインタラクションによる作曲行為の蓄積から作曲家が持つ知をシステム上に再構成しようと試みている。

2.2 システムの構成

提案する作曲システムはモチーフとモチーフに一連の変形操作を施して得られるフレーズとから多角的な楽曲の楽譜（MIDI データ）を生成するシステムである。モチーフは、作曲者による入力、ノートの遷移確率に基づく生成、ランダム発生など、いくつかの方法で与えることができる。ノートの遷移確率は楽曲の事例から抽出したものをを用いる。単位となるモチーフ変形操作は作曲技法に対応する。モチーフ変形操作の列は、作曲者による入力、変形操作の適用順序のパターンの組み合わせ、モチーフ変形操作列の変形など、いくつかの方法で作成される。適用順序のパターンは作曲者とシステムとのインタラクションを通して機械学習される。

多くの自動作曲システムは既存の楽曲（事例）を分析して得られた楽曲の構成規則を、楽曲を生成するための規則に応用するものであると筆者らは認識している。これらのシステムの背後には楽曲の分析プロセスと生成プロセスは可逆であるとの前提があると推量される。筆者らはこの前提を懐疑的に捉え、生成プロセスの一部には分析プロセスとは不可逆なプロセスがあると仮定している。例えば、音楽の多義性は複雑な構造や慎重に定義された生成文法を用いて記述可能であるかもしれないが、多義的な楽曲は分析によって得られる構造や生成文法をトップダウン的に適用して作曲されるのではない。ひとつひとつは単純であるともいえる基本的かつ局所的な操作（作曲技法）が組み合わせられることによって、複雑かつ多義的な楽曲が生み出される。

2.3 形式言語としての楽曲

システムはモチーフや楽曲を構造化された数値によって記述する。この意味で、モチーフに作曲技法を施しながら進行する作曲プロセスは構造化された数値という形式言語を処理する過程と同等である。楽曲を構造化された数値によって記述することのメリットは五線譜上に記されている情報の関係性を明示的に扱えることである。

ひとつの音 (note) を $s = \langle l, p, v, d, a \rangle$ なる 5 つ組によって特徴づけ、記述する。ここで、 l, p, v, d, a

は音 s が発音される時間的な位置, 音 s の高さ(pitch), 音 s の強さ(velocity), 音 s の長さ(duration), 音 s のアーティキュレーション(articulation)を, それぞれ, 示す. 楽器も音を特徴づける重要な要素であるが, 本稿では楽曲の抽象的な構造に特に注目するため, 楽器による音の特徴付けを行わない. ひとつひとつの音の記述を音楽詞(music word)とよぶ. また, すべての音楽詞の集合 Σ_M を下式によって定義し, 音楽アルファベット(music alphabet)とよぶ. 式中の $\mathbf{L}, \mathbf{P}, \mathbf{V}, \mathbf{D}, \mathbf{A}$ はすべての l の集合, すべての p の集合, すべての v の集合, すべての d の集合, すべての a の集合を, それぞれ, 示す.

$$\Sigma_M = \{ \langle l, p, v, d, a \rangle \mid l \in \mathbf{L}, p \in \mathbf{P}, v \in \mathbf{V}, d \in \mathbf{D}, a \in \mathbf{A} \} \quad (1)$$

楽譜は表記システム (notation system) の規則に従って音楽詞を配置するものである. 音楽詞を配置することによってつくられる記号の集まりを音楽詞配置 (formation) とよぶ. 音楽詞の配置のされ方は, 採用される表記システムの表記規則に従う以外, 特に制約を受けないものとする. すべての音楽詞配置の集合を Σ_M^* によって示す. Σ_M^* には任意の音楽詞をランダムに配置したものも含まれる. すなわち, 音楽詞配置によって表されたものは必ずしも音楽として成立するわけではない. 音楽詞配置のうち音楽として成立しうるものを音楽記述 (music description) とよぶ. 音楽記述の集合を \mathbf{S}_M によって示す. \mathbf{S}_M は音楽の可能な形式の描写のすべてからなる集合であり, \mathbf{S}_M に属さない音楽詞配置は演奏不可能であるとか楽曲として認識されがたいなど, 音楽としての意味を持たない記述である. \mathbf{S}_M の要素のうち, 芸術的であるとみなせる音楽作品(artistry)の音楽記述の集合を \mathbf{S}_{Ma} とする. \mathbf{S}_M の輪郭と \mathbf{S}_{Ma} の輪郭は必ずしも明確ではないが, Σ_M^* , \mathbf{S}_M , \mathbf{S}_{Ma} の間には次式の関係が成り立つ (図 1).

$$\mathbf{S}_{Ma} \subseteq \mathbf{S}_M \subseteq \Sigma_M^* \quad (2)$$

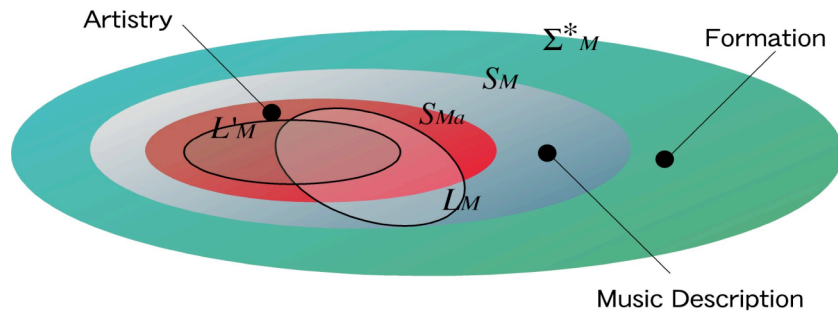


図1 音楽詞配置, 音楽記述, 音楽作品, 言語の関係

筆者らが作曲システムに期待することは \mathbf{S}_{Ma} の部分集合の要素 (図 1 の \mathbf{L}'_M の要素) あるいは \mathbf{S}_{Ma} の要素でもある場合がある \mathbf{S}_M の部分集合の要素 (図 1 の \mathbf{L}_M の要素) を生成することである. 多くの作曲システムは楽曲の構成を生成文法などで明示することによって \mathbf{L}_M や \mathbf{L}'_M を定義する. 筆者らのシステムはモチーフとモチーフの変形操作の適用順序を \mathbf{L}_M または \mathbf{L}'_M を定義するためのパラメータとする. 従って, 適切な変形操作の列を見つけることがシステムを期待通りに働かせるための鍵となる.

2.3 言語処理としての作曲

提案するシステムはユーザーに指定されたモチーフを変形しながら作曲を進める. 作曲の

開始時に、指定されたモチーフは基本モチーフ記憶バッファに格納されるとともに変形モチーフ記憶バッファに格納される。また、楽譜バッファは白紙の状態に初期化される。変形モチーフ記憶バッファに格納されている情報は、作曲プロセスにおいて、特定の操作によって楽譜バッファに書き出される。楽譜バッファの内容は、全ての変形操作の終了時すなわち作曲終了時に、MIDI ファイルとして出力される。現行システムでは、MIDI ファイルに書き出される内容は楽曲の抽象構造を表すものであり、最終的な音楽表現の楽譜ではない。いわゆる鑑賞される音楽として楽譜を自動生成させるためには、修辞や制約などを用いた抽象構造の具体化プロセスをシステムに組み込む必要がある。

システムに与えたモチーフの変形操作は次の通りである。数字は変形操作の識別番号を示す。識別番号は操作の適用順序の指定に用いられる。

- (1) 変形モチーフ記憶バッファの内容を別途指定された度数だけ移調する。同時に、変形モチーフ記憶バッファの内容を楽譜バッファに書き込む際の接続位置を 6 単位後ろに移動する。また、移調度を -2 に設定する。
- (2) 変形モチーフ記憶バッファの内容を別途指定された度数だけ移調する。同時に、変形モチーフ記憶バッファの内容を楽譜バッファに書き込む際の接続位置を 9 単位後ろに移動する。
- (3) 変形モチーフ記憶バッファの内容のアーティキュレーション（スタッカートやレガート）を変更する。同時に、モチーフの拡大率を 0.5 に設定する（1/2 に縮小）。
- (4) 変形モチーフ記憶バッファの内容のアーティキュレーション（スタッカートやレガート）を変更する。同時に、モチーフの拡大率を 2 に設定する（2 倍に拡大）。
- (5) 変形モチーフ記憶バッファの内容を逆行形にする（前後関係を逆にする）。同時に、変形モチーフ記憶バッファの内容を楽譜バッファに書き込む際の接続位置を 5 単位後ろに移動する。
- (6) 変形モチーフ記憶バッファの内容を反転形にする（指定された高さを軸として音の高さを折り返す）。また、移調度を 7 に設定する。
- (7) 変形モチーフ記憶バッファの内容の一部を反転する。同時に、変形モチーフ記憶バッファの内容を楽譜バッファに書き込む際の接続位置を 10 単位後ろに移動する。
- (8) 変形モチーフ記憶バッファの内容を原型（基本モチーフ記憶バッファの内容）に戻す。
- (9) 変形モチーフ記憶バッファの内容を転回形にする（音の高さがある範囲外にあるとき、指定された高さの軸と線対称の高さに変更する）。同時に、変形モチーフ記憶バッファの内容を楽譜バッファに書き込む際の接続位置を 2 単位後ろに移動する。
- (10) 変形モチーフ記憶バッファの内容の一部を切り出して展開する（切り取った部分を移調しつつ、繰り返す）。
- (11) 変形モチーフ記憶バッファの内容の一部を切り出し、上とは違うパターンで展開する。
- (12) 変形モチーフ記憶バッファの内容の前後を入れ替える（モチーフの開始点をシフトする）。
- (13) 何も行わない。
- (14) 変形モチーフ記憶バッファの内容を別途指定された度数移調して楽譜バッファに書き出す。このとき、別途指定された接続位置において不協和音が発生する場合、接続

位置を微調整する。また、拡大率、接続位置、移調度数を初期値（それぞれ、1.0, 6, 2）に戻す。

(15) 同上。

(16) 同上。

3. 作曲実験

3.1 着眼点

システム制作の初期段階として、主として次の2点に注目する実験を遂行中である。(1) 作曲家と提案システムとのインタラクションを通じて音楽記述または音楽作品を生成するためのモチーフ変形操作の適用順序が学習されるか否かを確認する。(2) 学習された場合に、変形操作の適用順序のパターンを抽出し、変形操作を構造化する。

3.2 方法

指定した方法で生成した変形操作の列を用いてシステムに音楽詞配置を生成させる。音楽詞配置に基づいて演奏された音(楽曲として成立しているとは限らない)を作曲家(筆者の一人、古川)が聞き、10点満点で評価する。評価値は0：音楽詞配置，1~4：音楽記述，5~10：音楽作品である。システムは変形操作の列と評価値に基づいて変形操作の遷移の頻度表を更新する。以下に実験の手順を示す。なお、図2に示す旋律を基本モチーフとした。

- (1) 指定した長さの変形操作列をランダムに10本程度発生させる。
- (2) 各変形操作列に基づいて生成される楽曲を作曲家が評価する。
- (3) 評価値が最高の変形操作列を選び、列の一部のみをランダムに変更した変形操作列を10本程度発生させる。
- (4) 各変形操作列に基づいて生成される楽曲を作曲家が評価する。
- (5) 音楽記述を生成した変形操作列から、変形操作の連続のパターンの頻度表を作成する。頻度の積算において、変形操作の評価値を当該変形操作にて出現する変形操作の連続パターンの頻度に乗じる。
- (6) 1~5を繰り返し、変形操作の連続パターンの頻度表を複数生成する。
- (7) 頻度表を適当に組み合わせ、コンテキストが0から10の変形操作の遷移確率行列を作成し、同行列を用いて変形操作列を複数発生させる。
- (8) 各変形操作に基づいて生成される楽曲を評価し、変形操作の連続パターンの頻度表を作成する。
- (9) 1または3または7から繰り返す。



図2 実験に用いた基本モチーフ

3.3 結果と考察

定性的かつ直観的であるが、以下の知見が得られている。

- (1) 手順1~4によって音楽作品を生成するための変形操作の列が絞り込める。ただし、

試行毎に得られる変形操作の連続のパターンの頻度表は異なる。また、試行によって、音楽作品に近い音楽詞配置が生成されたりされなかったりする。すなわち、システムは、この段階において、ランダムに変形操作列を発生させるよりは音楽記述や音楽作品に近い音楽詞配置を生成しうるが、常にそうであるわけではない。

- (2) 手順7によって音楽詞配置を生成する場合、それまでに作成された全頻度表を用いると、ランダムに発生させた変形操作列を用いるよりも音楽記述が生成される確率が高くなる。しかし、生成される音楽詞配置は音楽記述ではあるが音楽作品とは言い難いものとなる。全頻度表を用いる場合は、全ての試行において高い評価値を得た、いわゆる定石的な変形操作の連続が発生される確率が高くなり、音楽記述の生成に成功する確率が高くなるものと推量される。
- (3) 一方、同じく手順7によって音楽詞配置を生成する場合、数種類だけの頻度表を組み合わせて用いた場合、音楽作品であると評価される音楽詞配置が生成される確率が高くなる。適当な頻度表を用いる場合は、定石的な変形操作の連続の発生確率が高くなる傾向にあるものの、それ以外の連続パターンも混入する。すなわち、定石的な変形操作パターンに定石的ではない変化操作パターンを加えると、音楽作品が生成される可能性が上がる傾向がある。
- (4) 手順7を用いる場合、コンテキストが5程度のとき、音楽作品が生成される割合が上がる傾向がある。

4. 今後の課題

システムは制作途上にあり、多くの課題が残されている。主要な課題を以下に記す。

- (1) 変化操作のパターンを変形操作の列の構造として抽出及び表現する方法を検討する必要がある。現行のようにマルコフ過程を用いると定石的な変形操作の連続パターンが高い頻度で発生されるが、そのことが音楽作品ではない音楽記述を頻繁に生成する原因ともなっている。
- (2) 作曲の意図と整合的な音楽記述または音楽作品を生成する方法を検討する必要がある。作曲の意図、使用したモチーフ、変形操作列の関係がその手がかりになると推量される。
- (3) 音楽以外の芸術表現形態との連係の方法を検討する必要がある。現行のシステムは控曲システム単体として独立しうるものである。しかし、筆者らの興味は異種の芸術表現形態に共通するデザイン行為の構造にある。この構造を明らかにすることによって、作曲を視覚的に行うとか空間デザインを聴覚的に行う、すなわち、聴覚と視覚を連係させるデザインが可能になる。

参考文献

[Katayose 2000] 長尾真, 片寄晴弘ほか; 文字と音の情報処理, 岩波講座マルチメディア情報学4, 岩波書店 2004年