

## 中心音を考慮したわらべうたの創発的設計

布袋田由理子\* 竹中毅\*\* 上田完次\*\*

\*東京大学大学院工学系研究科

\*\*東京大学人工物工学研究センター

本研究は、人間の認知に基づいてどのように旋律が発生するのかを明らかにすることを目的とするものである。そのために旋律生成を創発と捉え、人間の認知に基づき設定した評価指標を用いた創発的旋律生成モデルを提案し、それに基づいたわらべうたの旋律生成実験を行う。提案モデルにおいては、エージェントが音の音長・音高を繰り返し決定することで旋律が生成される。本稿では特に、中心音との協和と近接を考慮して行った旋律生成実験について述べる。実験の結果として、各実験条件において異なる傾向を持つ旋律が生成され、旋律生成の条件を操作することにより、異なる印象を持つ旋律を生成できる可能性を示した。

### Emergent Design of Japanese Children's Song Considering the Idea of Central Note

Yuriko Hoteida\* Takeshi Takenaka\*\* Kanji Ueda\*\*

\*School of Engineering, The University of Tokyo

\*\*Research into Artifacts, Center for Engineering, The University of Tokyo (RACE)

In this paper, we aim at understanding the relationship between cognitive features and melody generation processes. For the purpose, we propose a model using reinforcement learning which generates melodies mainly based on sensory consonance with central note and proximity. We conduct computer simulation in which Japanese children's songs are generated. The result shows that generated melodies have discriminative trend according to each conditions. This means we may be able to control the impression of generated melodies by setting proper conditions.

#### 1. はじめに

作曲においては、個々の楽音間の局所的な関係を築きながら、同時に曲全体の構造的秩序を実現することが求められる。また、聴取者は一つひとつの音を個別に聴いているのではなく、時間軸に沿って変化していく秩序を認知しているため、要素である個々の楽音と全体の秩序の関係を動的に設計することは極めて複雑な問題であると言える。

その中でも、旋律の設計問題は非常に古く、かつ現代においても最も難しい問題の一つである。現代の音楽学においては、調性音楽の和声に関する理論は体系化されてきたものの、旋律に関する理論体系は、旋律作法と呼ばれるいくつかの経験的規則や分析理論を除いては殆ど存在しない。しかしながら、調性音楽成立以前から存在する教会音楽や各地の民族音楽においては「旋法性」の存在が指摘されており、学問的体系化は難しいものの、現在に至るまで経験的・暗黙的に受け継がれてきている。

近年、情報技術と複雑系科学の発達に伴い、音楽を複雑系と捉え

た研究が行われており[1][2]、我々は特に音楽の生成を創発現象と捉えることで、人間の基礎的な認知的特性に基づいて音楽における秩序が生成される過程について明らかにしてきた[3][4][5]。本研究では、旋律の設計問題に対してこの視点からアプローチする。

本研究では、特に人間の認知とかわりの深い音楽を対象として考えるため、まず西洋調性音楽以外の音楽から始めることとした。西洋調性音楽では記号論的理論が重視され、必ずしも人間の認知に根ざした音楽が実現されているわけではないためである。そこで注目したのが各地の民族音楽である。特に日本のわらべうたに関しては、他の音楽要素に対する旋律の優位性が指摘されており、旋律の詳細な分類がなされている[6][7]。特に東川[7]は、使用音階と終止音に着目した分類法を提案している。この分類に従って旋律生成の対象を定めることで、計算機上で旋律を扱う際に使用音と終止音の条件を統制することが可能となった。様々な音楽分野を扱うことのできる旋律設計モデルを構築する第一歩として、本研究ではわらべうたの旋律生成を対象とする。

本研究の目的は、認知的特性に基づいた評価指標を導入した強化学習による旋律生成の創発的モデルを用いて、人間の認知に基づいた旋律生成を行うことである。我々はこれまでに、既存曲においては隣り合う音の近接性を多く満たしていること、そのことが既存曲に安定感を与えていることを示した。また生成曲はまとまり・安定感において既存曲に劣っていた[4][5]。

本稿では、近接性の評価に加えて、既存曲にまとまりを持たせている要因としての中心音という概念に着目した。「中心音」というものが音楽において大きな役割を果たしているであろうということは、古くから指摘されてきた[8]が、具体的にそれが何であるのかは明らかになっていない。そこで、感覚的協和[9][10]を考慮することで、音楽構造に一定の秩序を与えられるという我々の結論[3]を基に、中心音との協和を考慮することで、旋律にまとまりを持たせることができるという仮説を立て、評価指標の設定を行った。

## 2. 強化学習を用いた旋律生成

### 2.1 旋律生成モデル

本節では、本研究で提案する、強化学習を用いた旋律生成モデルについて説明する。

1章でも述べたように、本研究では日本のわらべうた[6]を扱う。簡便のため全ての旋律を八均(使用音 C, D, E, G, A)に統一した。また、既存のわらべうたの数の多い陽調レ旋法を対象とした。これにより終止音の音高はD音となり、音長は2音符と設定した。また、拍子は既存曲に多い4分の2拍子とし、旋律の長さは4小節とした。

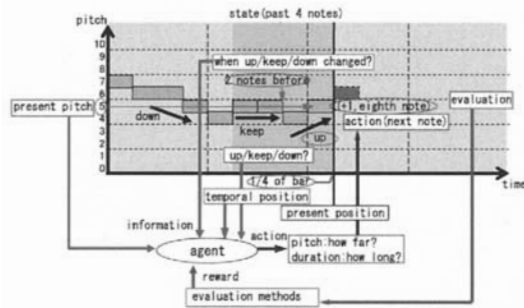


図1. モデルの概念図

提案モデルの概念図を図1に示す。楽音は音高と音長という2側面を持つ。本モデルでは、エージェントが楽音の音高と音長の決定を繰り返すことで旋律が生成される。エージェントの意思決定には強化学習が用いられる。エージェントはまず直近の過去4音の音高と拍節構造を状態として観測し、次の楽音の音高と音長を決定する。決定がされると、周辺の楽音との関係性を基にした評価が行われ、報酬がエージェントに与えられる。エージェントは与えられる報酬を基に、試行錯誤を通してよりよい行動を学習していく。なお、

開始音は既存曲を参考に所与とした(D, 8音符)。強化学習はQ学習、行動選択手法はソフトマックス手法を用いている[11]。

### 2.2 評価指標

旋律が生成される際の評価は、人間の認知に基づいて設計した。それ以外に身体的な制限から発生すると考えられる制約を設定した。身体的制約には、文化的背景による使用音階や拍子の制約や、使用される楽器などによる物理的な制約(音域など)が含まれる。本研究においては対象がわらべうたなので、人間、特に子どもの歌声に基づく旋律への制約を考慮する。

旋律生成に用いた指標を以下に示す。

#### 認知的特性評価

- ・感覚的協和

旋律で使用されている全ての音と、ある1つの基準音(中心音)との不協和度を計算し、楽音1つあたりの平均値を基に評価。不協和度の計算には Sethares の計算式[9]を用いた。

- ・時間的・空間的近接性

時間的近接と空間的近接が同時に実現された場合に正の報酬

#### 身体的制約

- ・拍節構造

小節を超えて伸びる音が現れた場合に負の報酬

- ・出現音高種数

出現した音高の種類が3種以下もしくは6種以上の場合に負の報酬

- ・音域

強化学習のエージェントの環境として設定

C3~C5

- ・音高行動の幅

エージェントの行動選択肢として設定

## 3. 旋律生成の計算機実験

### 3.1 実験1: 異なる中心音条件下の旋律生成

中心音を「仮想的に旋律とともに鳴っている音」と定義する。中心音と旋律中で使用されている音との協和を考えることで旋律にまとまりを持たせ、印象を制御できるのではないかと、更に複数の中心音を設定することで印象の制御がより精緻に行えるのではないかと、この仮説に基づき実験を行った。

ある音と旋律で使われている全ての音との不協和度を音数で平均したもの(旋律中の1音あたりの不協和度)を、その音の平均不協和度とする。平均不協和度の低い音の順位を使用音階の中で指定し、それを満たす旋律が生成された場合にエージェントに正の報酬を与えることで評価を行った。1位のみを指定することから、全ての順位を指定することまで可能である。

### 3.1.1 実験設定

陽類レ旋法[6]の旋律生成実験を行った。

- ・使用音 CDEGA (C3~C5 の 2 オクターブ)
- ・開始音 D4 8 分音符
- ・終止音 D4 2 分音符

協和評価の順位指定について以下のような 3 条件を設定した。

条件 1 (d) : 1 位 D

条件 2 (da) : 1 位 D 2 位 A

条件 3 (dag) : 1 位 D 2 位 A 3 位 G

なお、エージェントの行動選択範囲は3~+3とした。

### 3.1.2 結果と考察

実験 1 における、エージェントの獲得した報酬の推移の一例を図 2 に示す。試行錯誤を通して学習回数 300 前後で局所解に到達し、その後も探索を繰り返し学習回数 8300 前後で収束解を得ているのを見て取れる。

各条件において生成された旋律の一例を図 3 に、収束 10 パターンにおける、各音高と各音長の出現頻度を図 4・5 に示す。

図 4 から、協和の条件が厳しくなるにつれて (条件 1→3) A 音の出現頻度が高くなることからわかる。A 音は協和順位指定 1 位の D 音との不協和度が D 音自身に次いで低い (5 度音程)、A 音には、それ以外の、比較的不協和度の高い音を経由せずに到達できるため、その出現頻度が高くなっていると考えられる。図 5 を見ると、協和の条件が厳しくなるにつれて 8 分音符が少なく、その分 4 分音符と 2 分音符が多くなっている。このことは旋律で使用される音の数が少ないことに対応する。エージェントが音数を少なくすることで不協和度の高い音の使用を避けようとする戦略を採っていることが考えられる。

生成された旋律を実際に聴くと、協和の条件の差による違いははっきりとは感じられなかった。今回の設定とは違う中心音設定に基づいて生成した旋律と比較することが有効と考えられるが、そういった実験設定の下では収束パターンが得られなかった。これは、そもそもわらべうたの 5 音階 (ヨナ抜き音階) が何らかの協和基準により選別されている可能性[3]とも関係があると考えられる。つまり、音階選別の基準と異なる音を中心音として設定した場合、中心音との不協和度の低い音が極端に少なくなり、協和を満たそうとすると音高種数の制約を満たせないために、収束が得られないということが起こっていると思われる。この問題を回避し、様々な設定による実験を通して中心音による印象制御の方法を議論するためには、現在の 5 音階ではなく、7 音階 (西洋調性音楽など) あるいは音律からの旋律生成実験が必要であると考えられる。

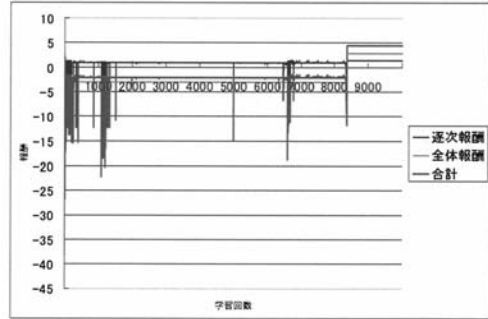


図 2. 実験 1 報酬の推移



図 3. 実験 1 生成された旋律の一例

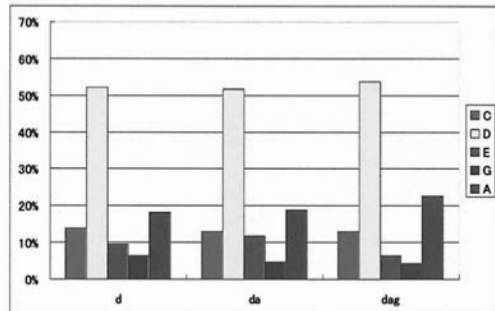


図 4. 実験 1 音高出現頻度

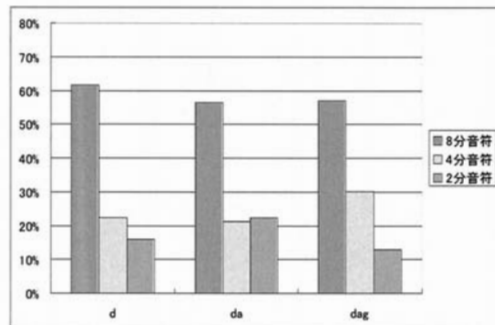


図 5. 実験 1 音長出現頻度

### 3.2 実験2：近接制約下の旋律生成

これまでの計算機実験と聴取実験[5]から、近接が多く実現されるほど聴取者が旋律に安定感を感じている可能性が示された。このことから、より安定感のある旋律を生成するために、近接を制約化した実験を行った。

#### 3.2.1 実験設定

陽類似旋法[6]の旋律生成実験を行った。

- ・使用音 CDEGA (C3~C5の2オクターブ)
- ・開始音 D 8分音符
- ・終止音 D 2分音符

エージェントの行動選択範囲によって以下の3条件を設定した。

- 条件1：エージェントの行動選択範囲 -3~+3
- 条件2：エージェントの行動選択範囲 -2~+2
- 条件3：エージェントの行動選択範囲 -1~+1

なお、協和評価の条件は実験1と同様(d)とした。

#### 4.2.2 結果と考察

実験2においても実験1と同様に学習が進み、収束解を得た。

各条件において生成された旋律の一例を図6に、収束10パターンにおける、各音高と各音長の出現頻度を図7・8に示す。

図6を見ると、エージェントの行動選択範囲が狭くなるにつれて(条件1→3) A音の出現頻度が低くなっていることがわかる。A音は、中心音であるD音との不協和度が低く(5度音程)協和の評価を満たすためには、この音を多用することが有利となる。条件1では、他の不協和度が比較的高い音を経由せずに直接開始音DからA音に遷移することができるため、他の条件に比べてA音の出現頻度が高くなっていると考えられる。条件2では、開始音から上昇するとD音との不協和度の高いE音などを經由しなければならず、これを避けるために、開始音から下降するパターンが多く見られた。これに伴い条件2ではE音の出現頻度が著しく低くなっている。条件3では、音高を移動する際に隣り合う音にしか移動できないため、他の条件に比べて各音高の出現頻度のばらつきが少ない。

生成された旋律を実際に聴くと、制約の幅が小さくなるにつれて旋律の違和感が少なくなるように感じた。詳細な検討には聴取実験が必要であるが、近接を制約化することで安定感のある旋律が生成されたと言える。



図6. 実験2 生成された旋律の一例

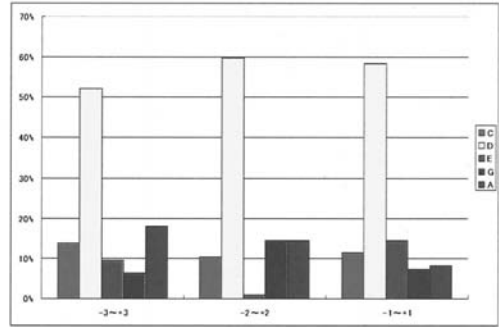


図7. 実験2 音高出現頻度

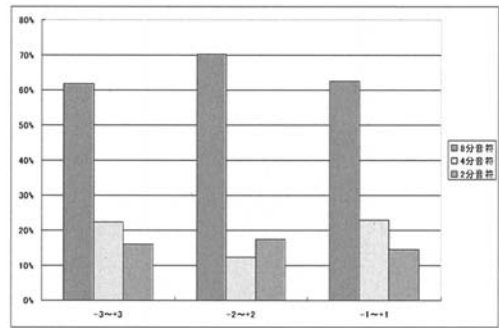


図8. 実験2 音長出現頻度

### 4. おわりに

本論文では、旋律生成を創発と捉えることで、人間の認知に基づいて旋律を生成するモデルを提案した。また、提案モデルを用いて行った二つの旋律生成実験について述べた。中心音の設定に注目して行った実験では、条件毎に異なる傾向を持つ旋律が生成された。不協和度の低い順位を下位まで指定すると、指定音との不協和度の低いA音の出現頻度が高くなった。近接を制約として設定した実験においては、制約の幅を小さくすることで、より安定感のある旋律が生成された。

今後の展望としては、中心音と旋律の関係のより深い議論のための、西洋音階や各種音律など、より多くの音を用いた旋律生成実験が挙げられる。そこで生成された旋律を用いた聴取実験を通して、中心音の設定と旋律の持つ印象の関係性を明らかにでき、旋律設計の方法論の確立につながると期待される。

### 参考文献

- [1] 大竹孝昌, 徳丸正孝, 村中徳明, 今西茂 (2003). カオスニューラルネットワークによる作曲支援システム:音高と音長分離によるメロディーのコーディング, 電子情報通信学会技術研究報告 ET, 教育工学, Vol.102, No.697, pp19-24.
- [2] 江村哲二 (2000). 一般設計学の芸術分野への応用としての楽曲創作過程のモデリング, 日本機械学会論文集C編, Vol.66, No.651,

pp3805-3810.

[3] 竹中毅, 鈴木晋太郎, 布袋田由理子, 上田完次(2006). マルチ・エージェント学習モデルを用いた音楽的構造の創発, 情報処理学会研究報告 2006-MUS-65(1).

[4] Y. Hoteida, S. Suzuki, T. Takenaka, K. Ueda (2006). Agent-based melody generation model according to cognitive and bodily features: Toward composition of Japanese traditional pentatonic music, Proceedings of 9th International Conference on Music Perception and Cognition (ICMPC9), 235-238.

[5] 布袋田由理子, 竹中毅, 上田完次(2007).中心音を考慮したわらべうたの旋律生成モデル, 日本音楽知覚認知学会 2007 年度春季研究発表会資料 pp.7-12.

[6] 小泉文夫(1958). 日本伝統音楽の研究 1 (民謡研究の方法と音階の基本構造), 音楽之友社.

[7] 東川清一(1990). 日本の音階を探る, 音楽之友社.

[8] Costere, E. (小宮徳文 訳) (1980). 和声の変貌—音高組織の論理, 音楽之友社.

[9] W.A.Sethares (1993). Local Consonance and the relationship between timbre and scale, Journal of the Acoustical Society of America, Vol.94, No.3, pp.1218-1228.

[10] 小畑郁男(2001). 楽器の音色を視野に入れた音高構成の研究—感覚的協和理論の音楽への応用—, 九州芸術工科大学博士学位論文.

[11] Sutton, R.S. and Barto, A.G. (2000). Reinforcement Learning, 森北出版.