

研究終了報告書

「AI 作曲におけるルールベース手法の機械学習への統合」

研究期間: 2021 年 10 月～2024 年 3 月

研究者: 田中 翼

1. 研究のねらい

作曲家は AI 作曲の可能性に対して否定的な考えを持っていることが少なくない。その理由としては、現状の機械学習による生成曲は、既存の楽曲の不完全な模倣になってしまい、新奇性がない、作りたい曲にならない、品質の保証がないなどが挙げられる。既存楽曲の調査を行った感触としては、むしろ作曲家は、ルールや目的関数を自分でデザインして楽曲を生成する手法を好む傾向が強いと思われる。その代表的な例としては、制約プログラミングや遺伝的アルゴリズムによる楽曲生成が挙げられる。これらの手法においては、制約や適応度のデザインを作曲家自身が行えるため、新奇性のある楽曲スタイルを定義したり、望ましくない響きを回避したり、品質の保証を与えることが可能となる。他方で、こうしたルールベースの手法によって音楽を生成するためには、楽曲の細部にわたってプログラムをしなければならず、大きな労力がかかってしまう問題がある。この点においては、データドリブンな手法による自動生成への期待があるのも事実である。

本研究では、このような観点から、作曲家がデザインしたルールと既存楽曲からの機械学習を融合することで、ルールベースの手法とデータドリブンな手法の良い点を統合するような新たな生成手法を開発することを目的とする。それによって、作曲家が独自のこだわりを持つ点に関してはルールとして意図を反映させつつも、それ以外の部分は機械学習に任せることで、新奇性や意図の導入や品質のコントロールを行えるようにし、作曲家が使いやすくなるようなシステムの実現の基礎となる方法を生み出すことが狙いである。

2. 研究成果

(1) 概要

音楽生成におけるルールベース手法と機械学習手法の統合のための手法として、GAN (Generative Adversarial Network, 敵対的生成ネットワーク) を一般化した新しい深層学習手法を開発した。この手法は、通常の GAN のようにデータドリブンな模倣のための手法としてのみならず、明示的なルールの充足のための手法として用いることを可能にするものである。この手法を複数の旋律ルールを充足する旋律の生成のタスクへ適用した。

GAN の通常の識別器のアウトプットは、本物か偽物かを判別するスカラー値であるが、提案手法では、これをベクトルへと拡張し、各次元が各ルールの充足度を判定するように一般化する。生成器は識別器の出力がそれぞれのルールに対する所望の充足度を持つ旋律を生成するように学習する。その際の学習データとしては、各ルールのみを充足する人工的データセットを作成して用いる。そこから学習することによって、すべてのルールを充足した旋律を生成することを目的とする。つまり、全てのルールを同時に満たす解が学習データに必ずしも存在しなかったとしても、全てのルールを満たす解を生成するという外挿問題を解くことが目標となる。

本手法による実験の結果、単純なメロディ生成のケースにおいて、4つの旋律的なルール

(制約)をすべて満たす旋律を生成することができた。これにより、GAN を明示的なルールの充足のための手法として用いることができ、ルールベースの手法と機械学習の統合の手法の成功例を提示することができた。

(2) 詳細

音楽生成におけるルールベース手法と機械学習手法の統合を実現するため、ACT-X 期間中、モンテカルロ木ベースの Sequence GAN、ニューラルネットによる適応度関数を用いた遺伝的アルゴリズム、Conditional VAE といった手法に基づく音楽生成実験を試みたものの、当初はなかなか精度が上がらなかったが、試行錯誤の結果、GAN (Generative Adversarial Network, 敵対的生成ネットワーク) を一般化した新しい深層学習手法を開発し、単純なメロディ生成のケースにおいて、旋律的なルールとしての複数の制約の充足を精度良く実現することができた。

この手法は、複数のルールを同時に満たす解を、各ルールのみを充足するデータ(人工的データ)から学習することによって実現するものである。つまり、全てのルールを同時に満たす解が学習データに存在しないにも関わらず、全てのルールを満たす解を生成するという外挿問題を解くことを目的とする。

GAN の通常の識別器のアウトプットは、本物か偽物かを判別するスカラー値であるが、提案手法では、これをベクトルへと拡張し、各次元が各ルールの充足度を判定するように一般化する。生成器は識別器の出力がそれぞれのルールに対する所望の充足度を持つ旋律を生成するように学習する。その際重要なポイントは、識別器への入力の本物か偽物かを示す真のデータラベルの代わりに、各ルールの充足度合いの真の値を用いる点である。この値は、識別器に入力される旋律(学習データおよび生成器によって生成されたもの)のルール充足度をプログラムでチェックすることで算出可能である。

通常の GAN の定式化は次のようになっている:

$$\max_D \left\{ E_x [\log(D(x))] + E_z [\log(1 - D(G(z)))] \right\}$$

$$\min_G \left\{ E_z [\log(1 - D(G(z)))] \right\}$$

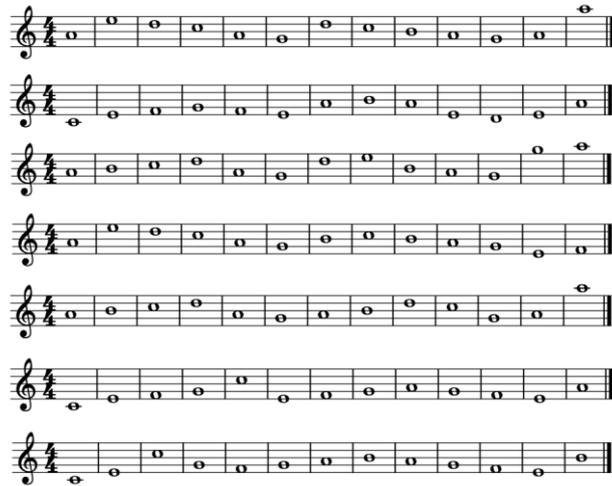
それに対し、提案手法である複数のルールの充足のための一般化は次のように定式化する:

$$\min_D \left\{ E_x \left[\|R_x - D(x)\|^2 \right] + E_z \left[\|R_{G(z)} - D(G(z))\|^2 \right] \right\}$$

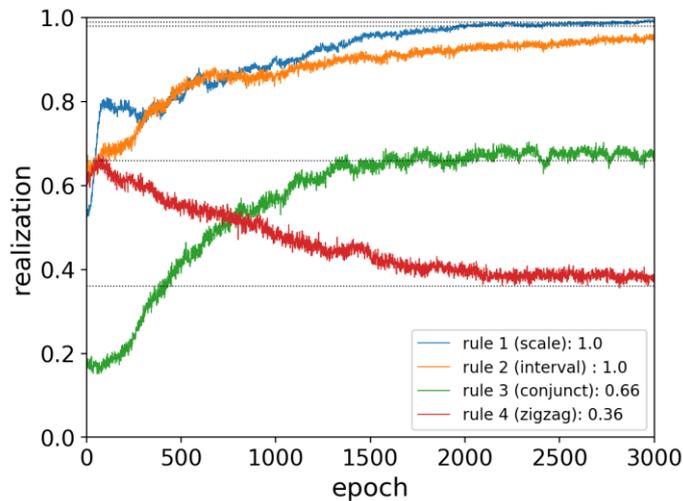
$$\min_G \left\{ E_z \left[\|T - D(G(z))\|^2 \right] \right\}$$

ここで、 R_x は学習データサンプル x のルール充足度のベクトル、 $R_{G(z)}$ は生成器の生成サンプル $G(z)$ のルール充足度のベクトル、 T はターゲットとなる理想的なルール充足度のベクトルである。

具体的な実験結果としては、4つの旋律ルール(音階内の音の使用、特定の旋律的音程の禁止、順次進行の多用、旋律の方向転換の数を小さくする)を設定し、次の図のような、すべてのルールを充足する旋律を生成することができた。



また、その学習プロセスにおいては、次の図のように、各ルールの充足度がターゲットの値へと収束していくことが確認できた。



本手法が成功した要因は、ルールをどのくらい満たしているかを判定する機能を GAN の識別器に担わせることで、ルールを充足する例を真似るだけでなく、ルールを満たさないもの(ターゲットベクトルと異なるルール充足度をもつもの)を排除する学習も行いうることができるためだと考えられる。それに対し、失敗した手法、例えば Conditional VAE の手法においては、基本的にデータサンプルを再構成することのみが学習されており、ルール充足度による条件付けを行っても、条件付けがうまくいかない生成例を排除する学習が不十分であるため、ルールを充足していないサンプルが生成されることが多くなってしまったものと考えられる。

本研究により、GANをデータドリブンな模倣のための手法としてのみならず、明示的なルールの充足のための手法としても用いることが可能であることを示すことができた。そして、ルールベースの手法と機械学習の統合の手法の成功例を作ることができた。ただし、すでに研究発表

が完了した部分において本研究で達成したのは、全音符による単純なリズムでの短い単旋律であり、より複雑なリズム、より長い曲、複数声部への対応などの改良を行うことが実用化のためには必要である。研究期間終了に近くなった時点において、リズムが可変のケースや二声の楽曲の生成についての新たな生成実験に成功し、大きく進展しつつある状況である(論文査読中)。

3. 今後の展開

本研究におけるルールベース手法と機械学習の統合というテーマは、音楽だけの特殊なテーマではなく、より普遍的な問題だと考えられ、広い応用先があると思われる。したがって、将来的には、音楽のみならず、他分野の問題への応用をも見いだしていきたい。一例として、詰将棋の生成への応用を構想している。AI による詰将棋の創作は、通常の一先手を予測する将棋 AI とは質が異なり、数手先の状況に関する制約を厳密に守らなければならないという意味で極めて困難な課題である。そのため、既存の詰将棋の局面をデータとした通常の機械学習で模倣して生成するには困難がある。実際、既存研究においては機械学習の適用例は少なく、むしろルールベースの手法(ここでは遺伝的アルゴリズムにおける適応度関数の手動設計もルール設計と解釈する)が多く用いられている。なぜなら、詰将棋においては、通常の将棋のルールに加えて、王手をかけ続けること、解が一意であること、持ち駒を使いきることなど、多くのルールを同時に満たす必要があり、ルールを充足する解が極めて珍しいという特徴があり、機械学習の適用が困難である。また、一間龍、頭金など、特定の形を満たすというデザイン上の制約を課すことで、作者の創意や教育的配慮をルールで表現することも重要となる。本研究で開発した手法を応用すれば、こうした制約を考慮しながら機械学習に基づく詰将棋生成が可能になるかもしれない、さらに、ルールベース手法のネックである探索時間の問題を回避できるようになるかもしれない。

楽曲生成については、引き続き、より大規模な楽曲の生成を目指して手法の拡張を行っていく予定である。その他、ACT-X の実施期間中に見出した新たなアイデアの検証を行なっていきたい。そのアイデアには、Diffusion モデルを用いた制約に基づく生成手法、遺伝的アルゴリズムと機械学習の融合、音楽におけるルールベース手法とデータ・ドリブン手法の統合に関するベンチマークとなりうる問題としての12音技法に基づく無調音楽の生成である。12音技法に基づく作例の数は限られており、大きなコーパスを作ることは困難である。また、音列が曲ごとに異なるため、通常の機械学習のみでは生成は不可能であろう。しかし、創始者のシェーンベルクの12音技法楽曲などは、リズムの面などで、それ以前の調性音楽との連続性が高いため、調性音楽のコーパスからの学習が部分的には有用だと考えられる。そのため、この問題は、ルールベース手法と機械学習の統合が必要な典型的な問題だと考えられる。

4. 自己評価

研究目的の達成状況に関しては、初期においては、当初の構想通りには研究が進まず、様々な別のアイデアを試しながら、後半になってから部分的な成功に到達し、最終年度によりやく成果発表を行うことができ、その後研究が加速している状況である。その理由としては、一般論として芸術の研究は1人で行うことが多く、今回も研究実施体制は研究代表1人であり、初期にはサポートしてもらえらる高度な人材の確保が困難であったことが挙げられる。音楽の研

究を理解できるプログラマを育てながらの実施であったが、研究期間の後半になり、サポートを受けられる水準に成長し、研究の遂行が楽になってきたことで進展が加速した。芸術の研究においてもチームづくりが重要だと痛感した次第である。

得られた研究成果である GAN を一般化した手法は、AI 作曲に限らず、他の領域にも応用可能な極めて一般的なものであり、大きなインパクトをもたらさうものだと考えている。その点で、自己評価としては極めて満足のいくものだと自負している。ただし、AI 作曲としての実用化までの道のりはまだ遠く、より規模の大きな楽曲(より長い曲や複数声部への対応など)の生成に向けた改良のための研究を継続する必要がある。

5. 主な研究成果リスト

(1) 代表的な論文(原著論文)発表

研究期間累積件数: 1件

1. Tsubasa Tanaka, “Melodic Constraint Satisfaction with a Generative Adversarial Network”, in Proceedings of International Computer Music Conference, pp.209–212, 2023

概要: 敵対的生成ネットワーク (GAN) は通常、データセットを模倣したインスタンスを生成するために用いられる。しかし、本研究では GAN のメカニズムを一般化し、制約充足の方法として使用できるように改変した。このように、トレーニングデータセットが必ずしも満たすとは限らない明示的な制約をニューラルネットワークに与え、充足する能力を獲得できれば、作品のオリジナリティや、生成 AI の制御可能性、信頼性を向上できる可能性がある。本手法の有効性を確かめるため、4つの音楽的制約を設定し、それらを満たすシンプルなメロディーを生成する実験を行った結果、提案手法の GAN は、それぞれの制約についての合成データセットから制約を学習し、すべての制約を同時に満たす解を生成することができた。

(2) 特許出願

研究期間全出願件数: 0件(特許公開前のものは件数にのみ含む)

(3) その他の成果(主要な学会発表、受賞、著作物、プレスリリース等)

International Computer Music Conference 2023 (深圳、中国)において上記の論文に関する口頭発表を行った。