



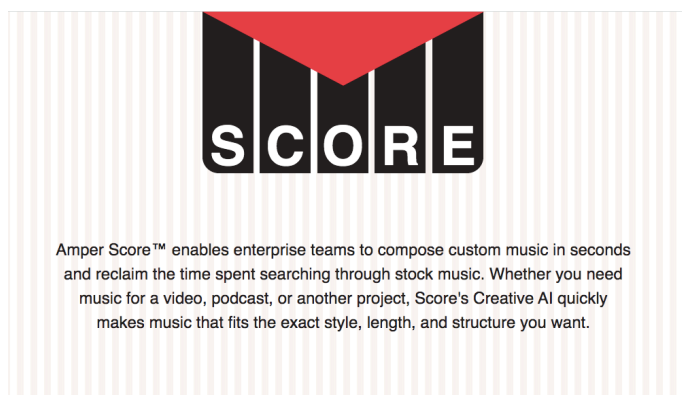
ディープラーニングを用いた 自動作曲システムの研究

事業構想学研究科情報デザイン領域

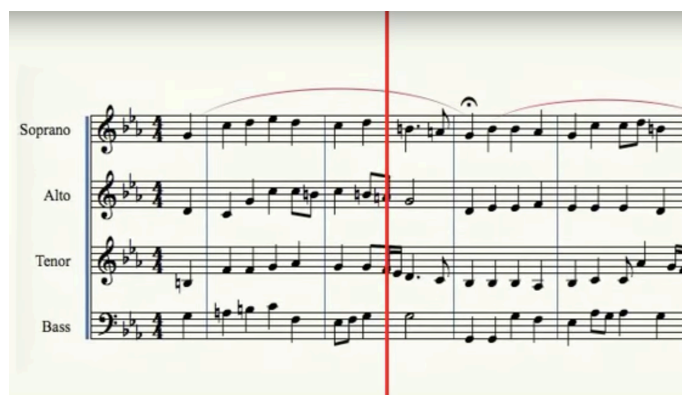
21852007 西野 有

研究目的

- ▶ コンピュータを使った自動作曲は古くからの問題領域である
- ▶ ディープラーニングを用いた自動作曲のシステムは既に存在している



Amper Score



DeepBach

研究目的

▶ 既存のシステムの問題点

- ▶ 学習元となった楽曲に似通ってしまう
- ▶ 主旋律のない楽曲が生成される



研究目的

- ▶ 本研究ではこの2つの問題を解決するための手法を提案する
- ▶ Magentaによる手法
 - ▶ 特定の要素に絞った楽曲を学習させモデル生成を行った

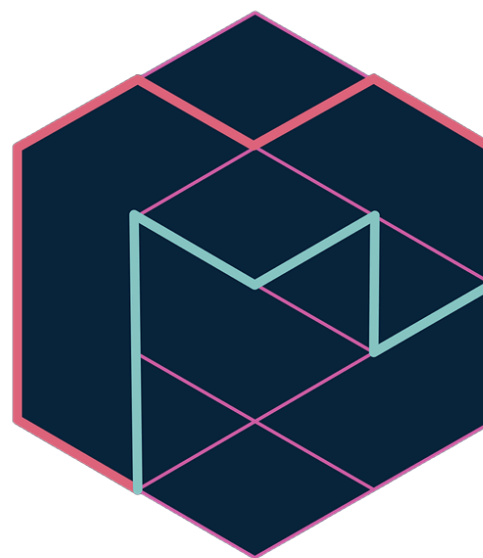
これに対しプロトタイプを作成し、評価を行った

▶ 既存研究

既存研究

Google Magenta

- ▶ ディープラーニングプロジェクトチーム「Google Brain」が行なっている機械学習を使って、音楽や芸術作品を生み出すことが目的のプロジェクト
- ▶ アーティストにも使いやすいプラットフォームを目指す
- ▶ 開発は GitHub にて行われており、ソースコードは公開されている

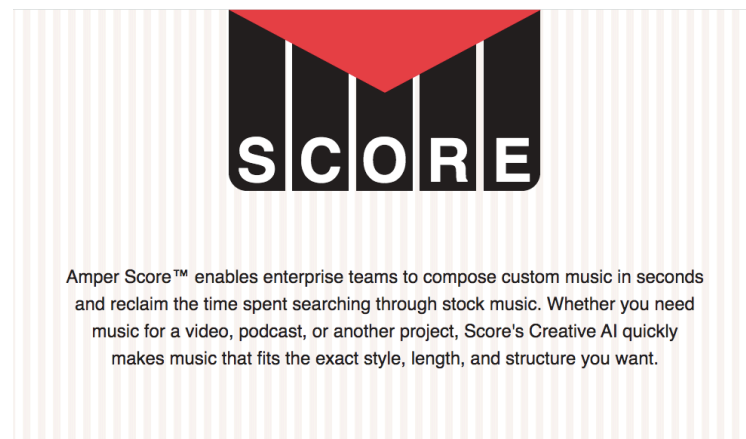


magenta

既存研究

▶ Amper Score

- ▶ 人工知能を用いて作曲ができるサービス
- ▶ 曲のジャンル、雰囲気、長さを指定するだけで作曲可能
- ▶ 楽器やテンポを指定することも可能
- ▶ 作成した曲はダウンロード可能



Amper Score



コンピュータ上の ▶ 音楽表現方法

コンピュータ上の音楽表現方法

- ▶ コンピュータで音楽扱う上で、どのように音を表現するのが重要になる
- ▶ 本研究で重要となる形式と、時間の扱い方について説明する



コンピュータ上の音楽表現方法

▶ MIDI

- ▶ MIDIでは音の高さ (pitch), 強さ (velocity), 長さ (duration)などが情報として、各音符が扱われる
- ▶ 「2, 96, Note_on_c, 0, 60, 90」のように現される
- ▶ 五線譜でも表すことが可能
- ▶ 使い勝手の良い表現方法である一方、微分音などの表現できない音も存在する

```
2, 96, Note_on_c, 0, 60, 90
2, 192, Note_off_c, 0, 60, 0
2, 192, Note_on_c, 0, 62, 90
2, 288, Note_off_c, 0, 62, 0
2, 288, Note_on_c, 0, 64, 90
2, 384, Note_off_c, 0, 64, 0
2, 384, Note_on_c, 0, 65, 90
2, 480, Note_off_c, 0, 65, 0
2, 480, Note_on_c, 0, 62, 90
2, 576, Note_off_c, 0, 62, 0
```



MIDI

コンピュータ上の音楽表現方法

▶ 時間の扱い方

- ▶ 音それ自体の他に、時間をどのように表現するのも重要である
- ▶ 大きく3つの方法がある
 - ▶ 1. 曲全体をグローバルに処理
 - ▶ 2. 時間を等分して処理
 - ▶ 3. 音符ごとに処理



Magentaによる ▶ 自動作曲システム

Magentaによる自動作曲システム

- ▶ MIDIをそのまま学習できるのがふさわしい
- ▶ Magentaをベースに研究を行った
- ▶ 学習データにはMIDIデータセットから著作権上の問題のないクラシック音楽を使用

Magentaによる自動作曲システム

Attention RNN

- ▶ RNN (Recurrent Neural Network) とは、ニューラルネットワークを拡張して時系列データを扱えるようにしたもの
 - ▶ 文章など連続的な情報の利用が有効な場合使用される
- ▶ Attention RNNとは毎回のステップでAttention (注意) すべきタイミングの情報を抽出して学習することができる

Magentaによる自動作曲システム

Attention RNN

$$u_i^t = v^T \tanh(W_1' h_i + W_2' c_t)$$

$$a_i^t = \text{SoftMax}(u_i^t)$$

$$h_t' = \sum_{i=t-n}^{t-1} a_i^t h_i$$

| | |
|--------------|---|
| v | 学習可能なパラメータを持つベクトル |
| W_1', W_2' | 学習可能なパラメータを持つ行列 |
| h_i | 前 n ステップ($h_{t-n}^t, \dots, h_{t-1}^t$)からのRNN出力 |
| c_t | 現在のステップのRNNセルの状態 |
| u_i^t | n ステップごとに1つの値をもつ長さ n のベクトル |
| a_i^t | アテンションマスクと呼ばれるベクトル |
| h_t' | RNNの出力の直近の n ステップ分に a_i^t を乗算し、合計したもの |

Magentaによる自動作曲システム

Attention RNN

- ▶ 次のようなRNNの出力が得られたとする
 - ▶ Step 1: [1.0, 0.0, 0.0, 1.0]
 - ▶ Step 2: [0.0, 1.0, 0.0, 1.0]
 - ▶ Step 3: [0.0, 0.0, 0.5, 0.0]
- ▶ アテンションマスクは次のように計算されたとする
 - ▶ $a_t^i = [0.7, 0.1, 0.2]$



Magentaによる自動作曲システム Attention RNN

- ▶ 次のように計算される
 - ▶ Step 1 (70%): [0.7, 0.0, 0.0, 0.7]
 - ▶ Step 2 (10%): [0.0, 0.1, 0.0, 0.1]
 - ▶ Step 3 (20%): [0.0, 0.0, 0.1, 0.0]
- ▶ これらの値を合計する
 - ▶ $h'_t = [0.7, 0.1, 0.1, 0.8]$



Magentaによる自動作曲システム 楽曲の生成

- ▶ 下記の方法でモデルの学習、楽曲の生成を行った
 - ▶ 1. MIDI→NoteSequences変換
 - ▶ MIDIをTensorFlowの学習形式であるTFRecord に変換する
 - ▶ 2. SequenceExamples 作成
 - ▶ NoteSequencesを学習用データと評価用データに分ける
 - ▶ 3.モデル学習
 - ▶ SequenceExamplesを使用してモデルの学習
 - ▶ 4.楽曲生成
 - ▶ 学習したモデルから楽曲を生成

Magentaによる自動作曲システム 生成楽曲の評価

- ▶ システムを使用して生成される楽曲を具体的に想定し、学習させるデータを変化させて6つのモデルを作成した
- ▶ 生成時の入力音にはすべて子犬のワルツの最初の4音を使用
- ▶ 生成される楽曲は全て単旋律

Magentaによる自動作曲システム 生成楽曲の評価

- ▶ 今回は以下の6つのモデルを作成した
 - ▶ (1) 全て学習モデル
 - ▶ (2) ベートヴェンのみを学習させたモデル
 - ▶ (3) ベートヴェン+バッハモデル
 - ▶ (4) 明るいベートヴェンモデル
 - ▶ (5) ピアノベートヴェンモデル
 - ▶ (6) 全ピアノ学習モデル

Magentaによる自動作曲システム 生成楽曲の評価

- ▶ (1) 全て学習モデル
 - ▶ データセット内のすべての楽曲計4666曲を学習させたモデル
 - ▶ 他のモデルとは異なり、比較的新しい楽曲も学習している
 - ▶ 複雑なリズムをとる現代的なリズムに感じる楽曲が生成された
 - ▶ 似たようなメロディーを少し変化させながら繰り返す部分もあり、音楽らしさはある
 - ▶ 全体的にはどこかぼやけた印象



A musical score example for the entire learning model. It consists of eight staves of music in 4/4 time, marked with a tempo of ♩ = 120. The music is written in a key with two flats (B-flat and E-flat) and features a complex, modern rhythmic style with various note values and rests.

全て学習モデル楽譜例



A musical score example of a repeated section. It shows a single staff of music in 4/4 time, marked with a tempo of ♩ = 120. The music is written in a key with two flats (B-flat and E-flat) and features a complex, modern rhythmic style with various note values and rests.

繰り返し部分

Magentaによる自動作曲システム 生成楽曲の評価

- ▶ (2) ベートヴェンのみを学習させたモデル
 - ▶ データセット内のベートヴェンの楽曲176曲を学習させたモデル
 - ▶ 冒頭部分はベートヴェンの交響曲9番4楽章の展開部を思わせる激しさがある
 - ▶ 全体を通してまるでベートヴェンの交響曲第5番4楽章のような力強い雰囲気を感じさせる楽曲が生成された
 - ▶ 実際のベートヴェンの楽譜と類似している楽譜も生成された



ベートヴェンのみを学習させたモデル



ベートヴェン作曲交響曲第9番より抜粋

Magentaによる自動作曲システム 生成楽曲の評価

- ▶ (3) ベートヴェン+バッハモデル
 - ▶ ベートヴェンとバッハの楽曲からそれぞれ176曲抽出し、学習させたモデル
 - ▶ スタッカートを多用しているのが特徴
 - ▶ ベートヴェンとバッハのそれぞれの雰囲気はあるようにも感じられたが、なんとも言い難い楽曲であった楽曲が生成された
- ▶ (2) よりもベートヴェン感は薄れた

The image shows a musical score for a piece generated by the Beethoven+Bach model. It consists of six staves of music in 4/4 time, with a tempo marking of ♩ = 120. The key signature is two flats (B-flat and E-flat). The score is written in treble clef and features a mix of eighth and sixteenth notes, with some staccato markings. The music is a single melodic line.

ベートヴェン+バッハ
モデル楽譜例

Magentaによる自動作曲システム 生成楽曲の評価

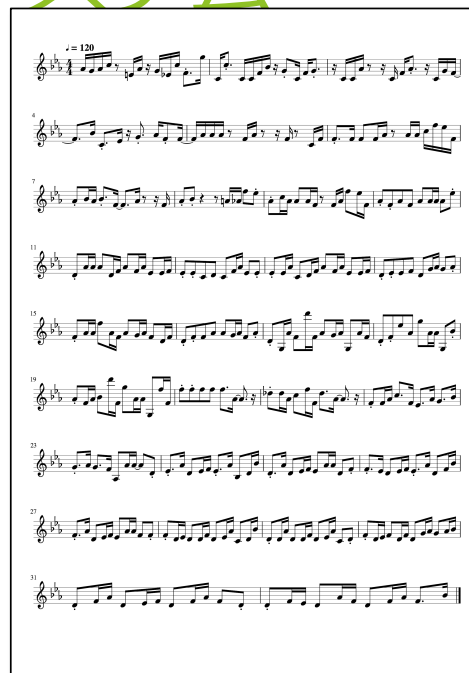
- ▶ (4) 明るいベートヴェンモデル
 - ▶ ベートヴェンの楽曲の内、明るいと感じた楽曲71曲を学習させたモデル
 - ▶ 全く同じフレーズを4回繰り返すなど、聞いていて不自然に聞こえる部分もある
 - ▶ 学習曲数が少ないことが原因か
 - ▶ (2)と比較して若干明るい楽曲が生成された
 - ▶ 「明るさ」焦点を絞る価値はある

A musical score example for the 'Bright Beethoven Model'. It consists of ten staves of music in 4/4 time, marked with a tempo of quarter note = 120. The key signature is two flats (B-flat and E-flat). The music is characterized by repetitive rhythmic patterns and melodic fragments, illustrating the model's tendency to generate repetitive or unnatural-sounding passages.

明るいベートヴェンモデル楽譜例

Magentaによる自動作曲システム 生成楽曲の評価

- ▶ (5) ピアノベートヴェンモデル
 - ▶ ベートヴェンのピアノソナタ45曲を学習させたモデル
 - ▶ 音と音の音程の間隔が大きい、リズムがとても複雑など聞いていて不自然なところが多い
 - ▶ 1つのフレーズを少しずつ変化させながら展開している部分などの評価できる点はある
 - ▶ 学習曲数が少なすぎたことが原因か
 - ▶ モデルとしての評価は低いですが、全く評価できないわけではない



A musical score for a piano piece, likely a sonata, in 4/4 time with a tempo marking of 120. The score is written in treble clef and features a complex, fast-paced melody with many sixteenth and thirty-second notes. The piece is divided into measures, with some measures containing multiple notes, indicating a high level of rhythmic complexity.

ピアノベートヴェンモデル楽譜例



A short musical phrase in treble clef, showing a sequence of notes: G4, A4, B4, C5, B4, A4, G4. The interval between G4 and C5 is a major third, which is highlighted as an example of a large interval.

音の間隔例

Magentaによる自動作曲システム 生成楽曲の評価

- ▶ (6) 全ピアノ学習モデル
 - ▶ すべての作曲家のピアノ曲計1891曲を学習させたモデル
 - ▶ 自然で軽やかなピアノ曲らしい楽曲が生成された
 - ▶ 途中の四分音符を境に異なるメロディーが始まる2部構成のような楽曲
 - ▶ 今回作成したモデルの中で一番良かった
 - ▶ 学習対象をピアノに限定したこと、学習曲数が多かったことが要因か



The image shows a musical score for a piano piece, likely in 4/4 time and a minor key. The score is written on a grand staff with two staves. The tempo is marked as ♩ = 120. The music consists of several measures, with a clear change in melody around the 10th measure, illustrating the two-part structure mentioned in the text. The notation includes various rhythmic values and accidentals.

全ピアノ学習モデル楽譜例

Magentaによる自動作曲システム

Accuracy / Loss

- ▶ 学習する曲数が増えるほどAccuracy値は小さくなり、Loss値は大きくなる
- ▶ (4)(5)のモデルで生成された曲の評価から、Accuracy値高ければ良いという訳ではないことがわかった

| | Accuracy | Loss |
|-----|----------|---------|
| (1) | 0.7537 | 0.8707 |
| (2) | 0.8867 | 0.3633 |
| (3) | 0.7538 | 0.801 |
| (4) | 0.9353 | 0.2049 |
| (5) | 0.9874 | 0.03963 |
| (6) | 0.7789 | 0.7693 |

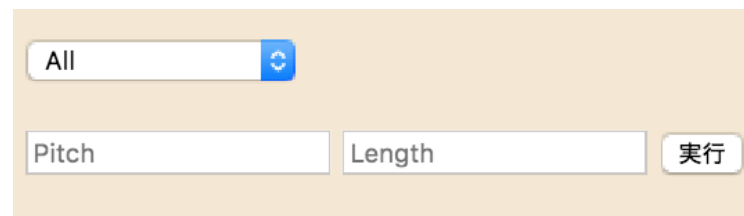
Magentaによる自動作曲システム

Accuracy / Loss

- ▶ (3) (4) から、特定の要素のみの学習と要素の加算を行うことは有効
- ▶ (1) (6) の比較から学習データは何かしらのまとまりを持って学習させた方が良い

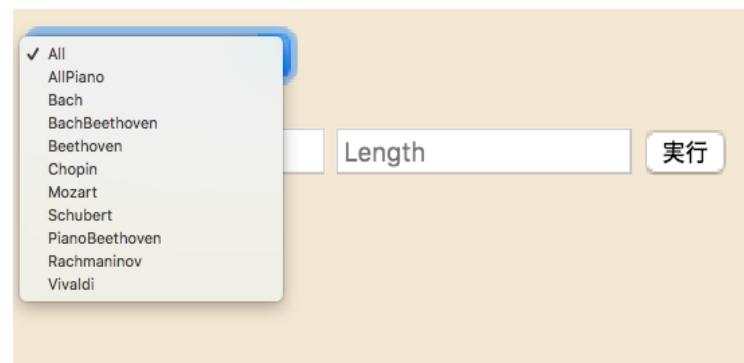
Magentaによる自動作曲システム Webアプリケーション

- ▶ PythonのWebアプリケーションフレームワーク「Flask」を使用
- ▶ 生成したい楽曲のモデル、初めの音の音程、曲の長さを指定して楽曲を生成
- ▶ 音程、長さの入力には多少のMIDI知識が必要
- ▶ UIは今後改善予定



A screenshot of the web application interface. It features a dropdown menu at the top with 'All' selected. Below it are two input fields labeled 'Pitch' and 'Length', followed by a button labeled '実行' (Execute).

Webアプリケーション



A screenshot of the model selection screen. A dropdown menu is open, listing various composer models: All (checked), AllPiano, Bach, BachBeethoven, Beethoven, Chopin, Mozart, Schubert, PianoBeethoven, Rachmaninov, and Vivaldi. To the right of the dropdown are a 'Length' input field and an '実行' (Execute) button.

モデル選択画面

Magentaによる自動作曲システム 評価

- ▶ 作曲したい楽曲を想定した上でモデルを作成し、自動作曲させる手法については十分な成果
 - ▶ 汎用性に欠けている
 - ▶ ユーザに合わせた学習を行うシステムにするか、全く別な手法を取る必要がある
- ▶ Webアプリケーションは自動作曲に必要な機能は揃えてある
 - ▶ MIDIの特性を理解しないと扱いづらい
 - ▶ ユーザーにダウンロード前に視聴してもらい、気に入らなければ再度調整する機能などがあればなお良い

終わりに

- ▶ Magentaによる手法
 - ▶ 汎用性は低く、改善の余地あり
 - ▶ 要素の加算については実現することができた
- ▶ 今後の課題
 - ▶ 汎用性の高い自動作曲システムに仕上げること



参考文献

- ▶ [1] L. A. Hiller and L. Isaacson: Experimental Music: Composition with an Electronic Computer, McGraw-Hill, 1959.
- ▶ [2] Amper Score: AI Music Composition Tools for Content Creators, <https://www.ampermusic.com/>
- ▶ [3] DeepBach: harmonization in the style of Bach generated using deep learning, <https://www.youtube.com/watch?v=QiBM7-5hA6o>
- ▶ [4] Magenta, <https://magenta.tensorflow.org/>

参考文献

- ▶ [5] TensorFlow, <https://www.tensorflow.org/>
- ▶ [6] J.P. Briot, G. Hadjeres, F. Pachet:
Deep Learning Techniques for Music Generation, Springer, 2017.
- ▶ [7] The Largest MIDI Collection on the Internet,
https://www.reddit.com/r/WeAreTheMusicMakers/comments/3ajwe4/the_largest_midi_collection_on_the_internet/
- ▶ [8] Flask | The Pallets Projects, <https://palletsprojects.com/p/flask/>

ご清聴ありがとうございました

