

TensorFlow を用いた “Giant Steps” のフレーズの学習と生成

菊地英斗 伊藤彰教 伊藤謙一郎

概要：本研究では、TensorFlow を用いて、転調を含んだジャズ楽曲である “Giant Steps” のフレーズの学習及び旋律の生成を行った。具体的には、“Giant Steps” のフレーズを基に学習用データを作成し、Recurrent Neural Network による学習と旋律の生成を行った。また、生成された旋律の分析を行うことで、生成された旋律において転調がどのように行われているかを検証した。

1. はじめに

コンピュータを用いた音楽生成の分野では、機械学習によって音楽を生成する手法も数多く研究されている。特に、今日における音楽生成の研究では、Convolutional Neural Network (CNN) を用いて音楽の学習と生成を行う Midi Net[1]など、深層学習を用いたものも数多く存在する。また、時系列データを扱うために、Recurrent Neural Network (RNN) や、それを発展させた Long short-term memory (LSTM) を学習に用いた研究も行われている。バッハの聖歌を深層学習によって学習することで、バッハの音楽的様式を模した音楽の生成を行なっている Deep Bach[2]では、LSTM が用いられている。

人間であれば、和声進行をはじめとする音楽理論的背景知識がない状態であっても、数千から数万回同じフレーズを聴取・演奏を行えば、旋律線のみから類似のバリエーションを生み出せる事例はそう珍しいことではない。幼少や思春期から音楽に集中的に取り組む人が、アドリブ演奏などをそれらしく演奏できるという例からもそれがうかがえる。本研究では機械学習において同様の状況を想定し、音階・和音など特段の予備知識を与えることなくメロディーを学習させ、そのネットワークを用いて類似のフレーズを生成することで、どのようなパフォーマンスが得られるかを検証する。特にこれらの課題について人間でも困難である複雑な転調を含むメロディーを事例として検証を行う。学習の事例として、ジャズの楽曲である “Giant Steps” のフレーズを基に学習用データを作成し、TensorFlow を用いて RNN による旋律の学習と生成を行なった。また、楽曲分析を行うことで、“Giant Steps” における和音進行と転調の構造を明らかにしたうえで、生成された旋律との比較を行うことで、生成された旋律における転調の様子を検証した。

2. 実行環境

2.1 TensorFlow と Magenta

本研究では、Google によって開発され、オープンソースで公開されているソフトウェアライブラリである

TensorFlow と、それを用いて機械学習を芸術の分野へ応用する研究プロジェクトである Magenta[3]によって提供されている、音楽の旋律の学習と生成のためのスクリプトを使用した。TensorFlow は 1.10.0 のバージョンを使用し、Magenta は 0.3.12 のバージョンを使用した。

2.2 データセット

学習に用いるデータについては、John Coltrane によって演奏された、アドリブソロを含む “Giant Steps” のフレーズのトランスクリプションを行った。そして、280 小節分の Midi データを作成した。そして、Magenta に含まれるスクリプトを用いることで、tfrecord 形式のデータセットを作成した。

2.3 学習と生成

学習には、Magenta に含まれるメロディ生成のための RNN モデルである melody_rnn を使用した。また、Batch Size は 64 とし、それぞれ 64 ユニットの 2 層の RNN を使用し、学習回数については、学習にかかる時間や、学習に用いるコンピュータのスペックなどを考慮し、20000 回とした。以上の条件でフレーズの学習を実行し、16 小節の旋律を生成した。

3. 事前楽曲分析

3.1 “Giant Steps” の分析

ネットワークが生成した楽曲のパフォーマンスを検証するために、音楽理論を基盤とした人間による事前の楽曲分析を実施する。この結果とネットワークが生成したメロディーを比較することで、音楽的な適合の度合いを検討するためである。まず、“Giant Steps” における和音進行と、その転調の構造を明らかにするために、“Giant Steps” における和音進行の分析に取り組んだ。

“Giant Steps” では、基本的には以下の図 1 で示した和音進行を繰り返すことで楽曲が進行する。

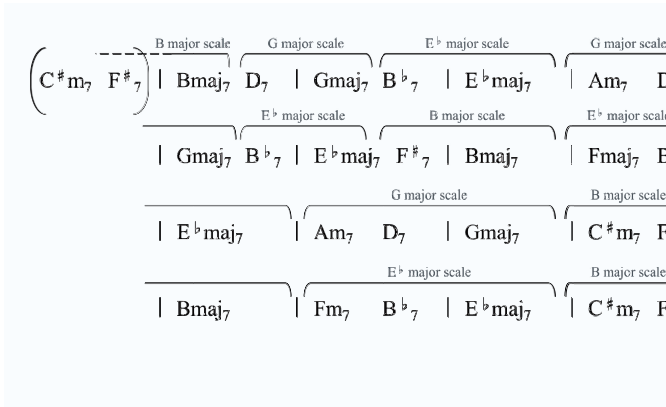


図 1 “Giant Steps” における和音進行

楽曲は Bmaj7 から始まるが、和音進行が一周すると、最後の C#m7 から F#7 の進行と括弧で示した冒頭の和音進行が重なり、曲を通して和音進行が繰り返す仕組みである。C#m7 から、ドミナントである F#7 を經由し、Bmaj7 へと進行する。この部分は B major scale 上の II—V—I であると解釈できる。次にくる属七和音の D7 は、B major scale 上には現れないが、その次に Gmaj7 へ解決することでドミナントモーションとなっている。そのため、この部分では G major scale への転調を行なっていると考えられる。続く Bb7 から Ebmaj7 への進行も同様にドミナントモーションであるため、Eb major scale への転調であると考えられる。その後に見える Am7 は Eb major scale では現れない和音である。しかし、その後の D7 から Gmaj7 への進行と合わせると、G major scale 上の II—V—I と解釈できる。そのため、この部分では Eb major scale から G major scale への転調をしている。その後の Bb7 から Ebmaj7 への進行は、Eb major scale 上の V—I のドミナントモーションによる Eb major scale への転調として考えられ、F#7 から Bmaj7 も同様に、B major scale 上の V—I のドミナントモーションによる B major scale への転調として考えられる。Fm7 は後に続く Bb7 から Ebmaj7 への進行と合わせて、Eb major scale 上における II—V—I の進行によって Eb major scale への転調を行なっている。その後に見える Am7 も同様に、後の D7 から Gmaj7 への進行と合わせて、G major scale 上における II—V—I の進行であるため、G major scale への転調を行なっている。同様にして、C#m7—F#7—Bmaj7 の進行と Fm7—Bb7—Ebmaj7 の進行についても、それぞれ B major scale と Eb major scale 上における II—V—I の進行であり、転調を行なっている。

このことから、“Giant Steps” では、B major scale、G major scale 及び Eb major scale の三つの調性について、非常に短い間隔で転調を行なっていることがわかる。また、用いられている三つの調性はそれぞれ長三度の音程で離れており、五度圏上に示すと以下の図 2 のようになる。

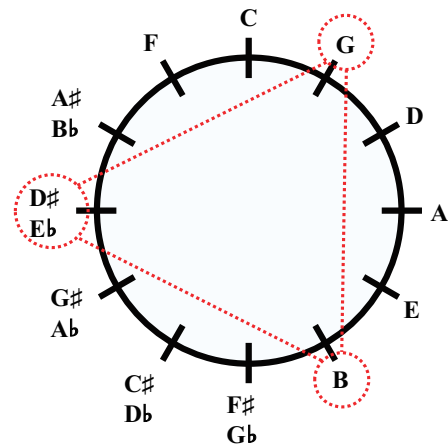


図 2 五度圏上に用いられている調性を示した図。

用いられている調性には、それぞれ長三度ずつ離れているという規則性が見られるが、それらが現れる順番については、規則性は見られない。また、転調の方法についても、V—I と II—V—I の二種類があるが、これらが現れる順番に規則性は見られない。

3.2 生成結果の考察

次に、“Giant Step” のフレーズの学習を実行し、旋律の生成を行なった結果について考察した。

生成された 16 小節の旋律について、前節で行なった楽曲分析結果を基に、和音進行のまとまりごとに分割した。生成を実行し、和音進行のまとまりごとに分割した結果を以下の図 3 に示す。

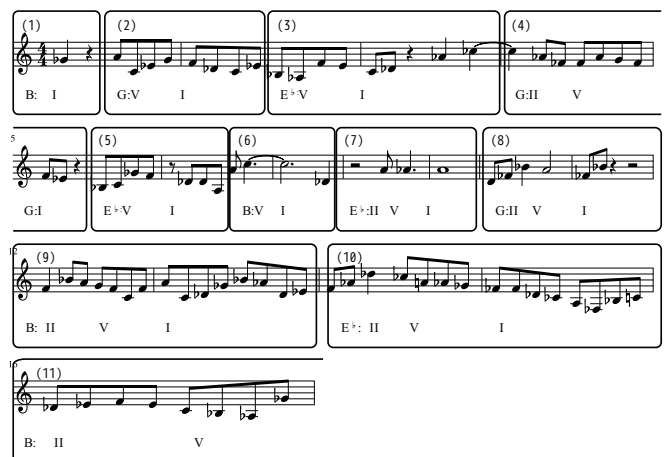


図 3 生成を実行し、和音進行のまとまりに分割した結果

そして、楽曲分析及び聴取による主観評価を行い、(a)「スケール・和音進行に適合しており、かつ自然である」(b)「スケール・和音進行に適合しているが、不自然である」(c)「スケール・和音進行に適合していないが、自然な外し方をしている」(d)「スケール・和音進行に適合しておらず、不自然である」の 4 つに分類を行なった。なお、この分類は、

研究者らによる主観的な感覚に基づくものである。

(1) B : I のブロック

まず、はじめの(1)B:Iのブロックでは、旋律はG♭を鳴らしている。これは、異名同音のF♯がB major scale上のIの和音における第五音である。また、ブロックの中に含まれている音が一音のみであるため、(a)に該当すると判断した。

(2) G : V—I のブロック

次に現れるG:V—Iのブロックでは、初めの4音にA, C, E♭, G♭が現れる。G major scale上におけるVの和音の構成音はD, F♯, A, Cであるため、CとA、そしてG♭の異名同音であるF♯の3音が一致する。また、E♭については、G major scale上には現れないが、同主調であるG minor scale上に現れる音でもあり、♭9thとして和音の構成音と捉えることもできる。一方で、続くF, D♭, C, E♭の旋律は、G major scale上に現れる音はCのみであり、和音の構成音とも捉えがたい。また、旋律全体を通して考えると、どちらかといえば、G major scaleではなく、D♭ harmonic minor scaleとして考えた方が、スケールに含まれる音も多く、より自然である。このことから、(d)と判断した。

(3) E♭ : V—I のブロック

次にくるE♭ : V—Iの部分では、旋律はB♭, A♭, F, E♭, C, D♭, A♭, Bの8音が用いられている。前半部分のB♭, A♭, F, E♭は、それぞれE♭ major scale上におけるV, IV, II, Iの音であり、全てE♭ major scale上に現れる音である。中でも、B♭とFは和音に含まれており、どちらも強拍に現れた後に長2度下行している。旋律の後半の部分であるC, D♭, A♭, C♭では、D♭及びC♭の異名同音であるBの2音がE♭ major scaleに含まれない音である。しかし、Bの音は次のG major scaleのブロックまで延びており、G major scaleにBの音は含まれるため、次のブロックの音を先取りした音として解釈することができ、全体を通して聴いても、あまり不自然さを感じることはなかった。そのため、この部分は(c)に該当すると判断した。

(4) G : II—V—I のブロック

次のブロックであるG : II—V—Iでは、冒頭に前のブロックから保留されたBの音が鳴っている。これは、G major scale上のIIの和音の9thに当たる音であり、和音の構成音としても捉えることができる。しかし、続くA♭, F♭, A♭, G, F♭, F, E♭は、G及びF♭の異名同音であるE以外の音がG major scaleに含まれない音である。スケールに含まれない音が多く用いられており、聴いてみてもあまり自然に聴こえなかったため、(d)と判断した。

(5) E♭ : V—I のブロック

次のE♭ : V—Iのブロックでは、前半部分のB♭, C, G♭, Fのうち、E♭ major scaleに含まれない音はG♭のみである。さらに、G♭の後に半音下行して和音であるFに進むため、不自然さも感じにくい。一方で、後半部分で用いられ

ているD♭, Aの音は、どちらもE♭ major scaleに含まれない音である。しかしながら、全体を通して聴いてみても、あまり不自然さを感じることはなかったため、(c)と判断した。

(6) B : V—I のブロック

続くB : V—Iのブロックでは、Aの音が非常に長く延ばされているが、これはB major scaleには含まれない音である。また、冒頭にあるAの音もB major scaleには含まれない。最後に現れるD♭は、異名同音のC♯がB major scaleには含まれるものの、ブロックのほとんどがスケール・和音進行から外れており、聴いてみても不自然さがある。そのため、(d)に該当すると判断した。

(7) E♭ : II—V—I のブロック

次のE♭ : II—V—Iのブロックでは、冒頭にE♭ major scaleには含まれないAの音が現れる。しかし、その後すぐに半音下行して和音の構成音であるA♭へ進むため、あまり不自然さは感じない。また、その後、再びAの音が現れる。しかし、全音符であることから、和音の音と同化し、#11thの音のように響くため、自然に聴こえた。そのため、このブロックは(c)に該当すると判断した。

(8) G : II—V—I のブロック

続くG : II—V—Iのブロックでは、D, F♭, B♭と上行した後に、Aの音が二分音符で伸ばされている。この中では、B♭の音がG major scaleには含まれない音である。しかし、その後半音下行して和音の構成音であるAの音へ進行するため、不自然さはあまり感じない。一方で、その後のF♭, B♭では、G major scale上のIの和音が鳴っているため、どちらも和音には含まれない音となる。そのため、あまり自然な旋律には聴こえず、(d)と判断した。

(9) B : II—V—I のブロック

次のB : II—V—Iのブロックでは、後半のD♭, G♭, B♭, A♭, D♭, E♭の6音は、それぞれ異名同音のC♯, F♯, A♯, G♯, C♯, D♯として、B major scaleに含まれているが、それ以外の音のほとんどがスケールに含まれておらず、自然な旋律にも聴こえなかった。そのため、(d)であると判断した。

(10) E♭ : II—V—I のブロック

次のE♭ : II—V—Iのブロックでは、ほとんどの音がE♭ major scaleに含まれない音であり、聴いてみても不自然な旋律であった。そのため、(d)に該当すると判断した。

(11) B : II—V のブロック

最後のB : II—Vのブロックでは、異名同音を考慮すると、B major scaleに含まれない音はFとCの2音である。また、B major scale上のIIの和音を9thとして考えると、どちらも次に和音の構成音へ進行している。そのため、不自然さを感じにくく、通して聴いても自然な旋律に聴こえた。スケール・和音進行へ適合する音も多いため、(a)と判断した。

以上をまとめると、(a)と判断したブロックは 2 つあり、(b)と判断したブロックはなかった。また、(c)と判断したブロックは 3 つあり、(d)と判断したブロックは 6 つあった。

4. おわりに

本研究では、TensorFlow を用いた RNN によって “Giant Steps” の学習と生成を行った。そして、“Giant Steps” の楽曲分析を行い和音進行と転調の構造を明らかにしたうえで、生成結果について考察し、生成結果における転調の様子を検証した。その結果、生成された旋律の多くはスケール・和音進行に適合しておらず、中には音楽的に自然な外し方をしているものもあったが、不自然なものが多くあることが確認できた。

今後の研究では、学習の際のパラメータや学習回数の設定の妥当性や、評価や判断の妥当性についても検討が必要である。また、本研究では、研究者らの主観的な感覚に基づいた評価を行ったが、「音楽的に自然である・不自然である」という判断基準をどのように設定するのかについても、今後さらなる検討が必要である。

参考文献

- [1]Yang, L. -C., Chou, S. -Y and Yang, Y. -H.. MidiNet: A Convolutional Generative Adversarial Network for Symbolic-domain Music Generation. 2017, arXiv:1703.10847.
- [2]Hadjeres, G., Pachet, F. and Nielsen, F.. Deep Bach:a Steerable Model. 2017, arXiv:1612.01010.
- [3]“Magenta project”. <https://magenta.tensorflow.org/>, (参照 2018-10-04).