

2-5-1 フレーズ練習記録の一人称研究を例にした ジャズ即興演奏生成システムのための検討

A Discussion about Jazz Improvisation System based on Studies with First-Person's View

安藤大地 *1

Daichi Ando

*1 首都大学東京システムデザイン学部

Faculty of System Design, Tokyo Metropolitan University

In this paper, the author shows an example of jazz improvisation acquisition process record based on the author's own first-person view research. The example is about practice and cognition of 2-5-1 harmony progresses and phrases. The record indicates that musician's cognition of internalize process can be clearly described in such a way as to rule. Then the author discusses possibilities of that it is effective to adopt musician's knowledge into the art creative machine learning system.

1. はじめに

本稿では、ジャズ即興演奏生成システムの構築を目的とし、著者自身の一人称研究的記録の中の2-5-1フレーズの部分を例に事例と問題点を示し、それを元にジャズ即興演奏生成システムに人間と共通の音楽的知識を取り込むことの有効性について議論する。

著者は、人間と音楽知識を共有する作曲システム・即興演奏システムの実現のために、主に遺伝的プログラミング (Genetic Programming, GP) と対話型進化論的計算 (Interactive Evolutionary Computation) へ人間の音楽知識への取り込みを進めている [安藤 05, 安藤 12, 安藤 17, 安藤 18a]。

近年のトレンドである DNN を用いた音楽生成では Google Magenta の例があるが、実際に Google Magenta をクラシック音楽の作曲家が用いた例では、少なくともクラシック音楽に適用した場合には十分な表現力がないという意見がある [大久保 17]。また、Cope は LISP のプログラムが、作曲家が作曲を行う過程のような音楽の高レベルでの表現に有効であると述べている [Cope 05]。OpenMusic という最も成功した作曲支援ソフトウェアも LISP である。

LISP が音楽家にとってベストな音楽生成記述言語であるかの議論は難しいが、少なくとも DNN 的表現による音楽生成よりは音楽家にとって理解しやすい・記述しやすいということが出来る。Cope は、この理由は音楽的知識を音楽的知識のまま入れ込むことができるからであるとしている。

よく知られている一番最初の機械学習的なジャズ即興演奏 (ジャズアドリブソロ) 生成システムは遺伝的アルゴリズムを用いた GenJam [Biles 94] である。GenJam の染色体コーディングに音楽的知識はほぼ入っていない。Google Magenta の DNN 表現は LSTM と同様の表現ノードにより極短期的な時系列的な表現が可能になっているという意味では GenJam に比べ発展しているが、その時系列表現が音楽としては極短期出来であること、音楽表現に特化したノードではないことから、音楽表現力が低いのではないかと著者は考えている。

さらに DNN 表現を用いた機械学習では、データ収集も含め学習に非常に大きなコストがかかることを考えれば、DNN 表現ではなく、「人間の音楽家が定めた (音楽的に破綻せず、そ

れになりのもので生成できることが担保されている、さらに人間の音楽家の個性が反映されている) ルールを元にし、その組み合わせを学習する」といった方針の方が、より良い結果を生むのではないかと、著者は考えた。この「人間の音楽家が定めたルールの組み合わせ」は前述の通り LISP で表現することが容易であり、多くの音楽家が理解し、自ら記述することができる。

また「音楽的知識」だけでは音楽の生成は不可能もしくは不十分なのではないかという議論もある。例えば即興演奏において、五線譜上に表現できる音楽知識ではなく、楽器と身体の関係性で感覚で演奏が行われていることは、Sudnow [Sudnow 79] をはじめとしてジャズピアノのトレーニング等では一般的に知られている。

本稿ではジャズのアドリブソロの単旋律の生成にターゲットを絞り、既存の研究の例を挙げ、さらに著者自身が一人称的に記録してきたジャズアドリブソロ生成の練習メモに対する考察を行う。

2. ジャズ即興演奏の「音楽的知識」と「音楽以外の知識」

Sudnow は、自身のジャズピアノの習得過程の記録 [Sudnow 79] の中で、音名等の音楽情報ではなく鍵盤に対する指の動きという身体性のみで記述を行った。鍵盤楽器は五線譜に一番マッピングしやすい楽器であるが、Sudnow の五線譜なしで指がどのように動くかのみでの即興演奏の習得過程の記述により、鍵盤楽器と指の身体性が少なくとも一部のジャズピアニストには五線譜よりも重要であることを示している。

著者は、ジャズ即興演奏生成に機械学習を適用するにあたり、断続的に人間の演奏家への聞き取り調査とそれを元にした生成ルールの構築 [安藤 15] を行なった。

聞き取り調査では、アマチュアベーシスト初学者を対象として、どのようにアドリブを生成しているかのインタビューを元にベースライン生成のルールベースのモデル化を行ったが、人間の生成モデルは、抽象化された五線譜上のような音楽的情報のルールの組み合わせで表現できるわけではなく、身体性と楽器構造が生成されるフレーズに大きな影響を与えていることがわかった。具体的には、初学者であればあるほど、ベースラインの構成音を「音」「音名」「階名」「コードの中の度数」などの音楽的情報としてではなく、「指のポジション」として把握し

ており、リードシートからのコードネームの視覚的入力音楽情報ではなく身体動作に結び付けられていたことがわかった。

Sudnow の記述と著者のインタビュー結果から、鍵盤楽器・弦楽器など「楽器中の特定の箇所・ポイント」に音高が割り当てられ、手指がそのポイントを移動することによって音高を選択する楽器では、少なくとも初学者の間は音楽情報よりも身体性が重要であることが示唆されている。

これらの事例から、Google Magenta のような音楽的情報要素のみから構成される汎用的なモデルを用いた音楽表現には限界がある、もしくは学習コストが非常に大きくなってしまい、実用レベルで人工知能を用いた音楽生成を行うためには、ジャンル・楽器ごとの音楽知識や、楽器ごとに身体性のみならず楽器特有の音楽認知を適切に取り込んだ知識が必要である、と著者は考えた。

3. 2-5-1 フレーズ即興演奏練習の一人称的記録と考察

そこで著者は、自身を被験者としてジャズサックスの即興演奏を習得する過程を一人称研究的に記録することを行ってきた[安藤 18b]。前述のベース、ギターの例や Sudnow のピアノの身体性が強く反映する例とは異なり、手の位置が基本的に動かない管楽器の場合、音名でフレーズを認知する必要があり、弦楽器や鍵盤楽器とはその習得過程における認知の仕方が異なることが予想される。

被験者でもある著者は、音楽大学の声楽学科の在籍経験を持ちクラシック演奏のフレーズングの基礎能力を有している。また作曲を行っておりジャズの理論を理解するのに苦労することはないという特徴を持っている。

この記録の過程では、ジャズサックス演奏者に指導を仰ぎ練習法の指導を受けるだけでなく、自身で考えた練習法を試して、「どのような練習をしている段階で、どのような獲得感が得られたか」を詳細に文章と五線譜に表記したフレーズにより記録した。

本稿では、そのなかの 2-5-1 フレーズに関する記録を抜き出てまとめて提示し、議論を行う。

2-5-1、または 1-2-5-1 とは「トニック (1) - サブドミナント (2) - ドミナント (5) - トニック (1)」という、「安定 - 緊張」を行き来することにより「解決」感覚を表現する典型的な和声進行の典型例の一つである。ハ長調 C-Major スケールでは「CM7 - Dm7 - G7 - CM7」という進行になる。クラシックやジャズなど多くの西洋音楽はこの繰り返しによって成り立っている。その最初のトニック (1) を外したものが一般的に 2-5-1 と呼ばれる。

ジャズではこの 2-5-1 進行を用いた部分転調がドッペルドミナント等により頻繁に行われ、曲全体の調・キーからこの部分のコードを切り離す必要があるため、この部分のみの進行を内在化する練習を行う必要があるとされる。

3.1 和声感覚の習得過程の記録と考察

3.1.1 指示された練習

まず、コードトーンのアレンジで 2-5-1 の音高の音名を把握する練習法を指示された。図 1 に、練習フレーズの C-major の例を示す。実際の練習では、このような五線譜上に記譜された音符を見てではなく、コードネームを見ながら根音からの音程を度数を頭に思い浮かべながら音名に変換して行なった。(図 1 の例の 3 小節目の前半ならば、練習時に頭に浮かべるのは「Dm7, 7531」であり、「Dm7: 7→ド, 5→ラ, 3→ファ, 1→レ」と逐次認識していく)



図 1: 練習用として指示された、または著者自身が考えた 2-5-1 のコードトーンアルベジオフレーズ。

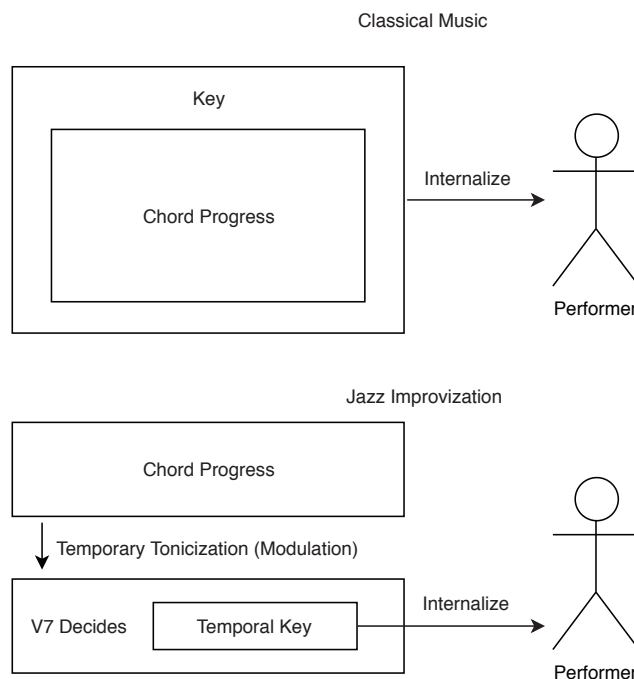


図 2: 著者の中でのクラシックとジャズの調性の把握感覚の違い。クラシックの和声進行は調・キーに内包されているのに対して、ジャズ即興演奏の場合はコード進行の中の属七 (図中「V7」) が部分転調の調・キーを生成しているイメージ。クラシックとジャズでは五線譜上での表現に差がないのに、演奏時の著者の「認識」もしくは「認知」は明らかに異なっている。

3.1.2 和声感覚把握時点での著者の問題点と考察

一般的にアヴェイラブルノートスケールに基づきジャズの即興演奏を行う場合、本来 2 の和音 (図 1 中 Dm7 の部分) のスタート時点で 5-1 (図 1 中の G7 → CM7 の部分) に到るまでの調号を把握している必要があるが、著者は図 1 に示したフレーズの練習を始めてから 2 年以上経っても 5 の和音 (クラシック和声で属七と呼ばれる和音) のアルベジオに到るまで、全体の調号を把握する、つまり調性を感覚的に明確に把握することができなかった。「頭と指ではわかっているのに、聴取感覚がついてこない」とも言える。音楽の原理的には属七のみが調・キーを決定するのはクラシック・ジャズ双方で自然なことであり、五線譜上に表現してしまえばどちらにも差がないのだが、著者は自身の「認識」もしくは「認知」の感覚の違いに注目した。

図 2 に著者のクラシックとジャズの 2-5-1 の和声進行の感じ方の違いを、記述したメモを元に表した模式図で示す。

クラシックの演奏やルールベースの音楽生成システムを考えた場合、調・キーは主旋律から導き出され、演奏を始める前に既知であることが前提であり、コード進行も調の中に内包さ

れている範囲に限定される。このような著者の様な「調がわからない」感覚は発生しないと考えられる。転調の際も「ディアトニック転調」と呼ばれる手法では移動する先の調・キーがわかっていることが一般的である。

それに対して、ジャズの即興演奏の時はコード進行のみが提示されており、調整を決定づけるような主旋律の提示は行われない。この状態では「それまでに演奏された旋律」が調性のガイド足り得ないため、5のコード「属七」が出てくることのみが、その部分の調整が決定する要因となっている。

しかしながらクラシック経験者の著者がこの練習をする時に感じた「属七」が出るまでわからなくなってしまう感覚は、ジャズ特有のごく短短いスパンでの部分転調の極端な多さや、アウト感を演出するために別のコードとも捉えられるテンションノートを中心とした音使いをする場合には、適切な認知なのではないかと考えた。

著者の場合のクラシック的な感じ方はトップダウン的な音楽知識に基づくのに対して、ジャズ的な感じ方は「生成的」である、とも言える。

前述の通り五線譜に記述、つまり五線譜上の音楽表現になってしまうクラシックとジャズのアドリブソロは全く同じルールで表現可能なのだが、そのバックグラウンドの和声進行の感じ方には、このように大きく違いがあることがわかった。

3.2 曲中での 2-5-1 フレーズ演奏の習得方法

次に実際の曲での 2-5-1 フレーズの使用についての練習の記録について述べる。

3.2.1 一般的な習得方法

曲中で 2-5-1 を即興演奏する一般的な習得方法としては、

1. 多くの演奏家が残している 2-5-1 フレーズ (リック) をコピーする。
2. そのフレーズを主要キーに移調できるようにする
3. ターゲットとなる曲の進行中に 2-5-1 を探して、コピーしてきたフレーズを当てはめる

という手順が推奨されている。実際に著者もそのような指導を受けた。ここで 1. の「多くの演奏家が残している 2-5-1 フレーズ」とは、ジャズ特有のアウト感を出すための非コードトーン (オルタードスケールやコンビネーションオブディミニッシュスケールに含まれている音高) が多く含まれており、演奏家がフレーズのストックを増やして使うことで自然にアウト感を出せるようになっている。

3.2.2 著者の問題点と解決法

しかしながら著者の場合、「あらかじめ覚えたフレーズを演奏し出すと、コードトーンが頭の中で鳴らなくなってしまう。認識できなくなってしまう」という問題が発生した。この現象はクラシック演奏経験者で即興演奏を習得しようとする人間には頻繁に見られる傾向のようで、音楽大学でピアノ即興演奏を教授している指導者から同じような状況が多く見られることを伝えられたことから、いわゆるフレーズを認識し演奏するというクラシックの練習経験やフレージングの発想が、コードトーン認知の邪魔をしているのではないかと考えた。今後聞き取り調査も含め詳細な研究が必要である。

そこで著者は、コピーをしないでコードトーンと 9 度 13 度の音を使って「生成的に」演奏をする方向に練習切り替えた。いわばこれは「与えられたフレーズをそのまま覚えるのではなく、与えられたフレーズをマルコフ過程的に一度分析してその生成モデルを獲得する」と言える。アウト感を出すためには

チャーリー・パーカーが取り込んだとされるドビュッシーやラヴェルの考え方をそのまま用い [Amy 86], 「より解決感を強く出すための基本の属七からズラす」と捉えることとした結果、アウト感を持つ 2-5-1 フレーズを即興演奏で生成ができるようになった。

フレーズの場合も和声感覚の時と同じく、フレーズを覚えるトップダウン的な知識に基づくのに対して、実際に著者が自然と感じた方法は生成的とも言える手法であると言える。

4. 考察

ここまで即興演奏習得に関する既存研究と、著者の 2-5-1 フレーズの即興演奏習得について一人称的記録をまとめた。

既存研究の調査では、即興演奏には楽器によって身体知が大きく影響している可能性があり、単なる五線譜ベースの音楽知識の表現だけでは即興演奏の表現そのものが難しいのではないかと推測した。

また、即興演奏習得過程の記録のまとめから、クラシック音楽の演奏者としての経験があり、かつ理論的には一般の学習者よりも高いレベルにある著者の場合、1. 和声進行把握がクラシックとジャズで異なっている、2. クラシック音楽演奏経験のフレーズ把握認知が邪魔をしジャズの一般的なリックを習得する練習法では効率が悪い、と推測した。

このように一言にジャズの即興演奏生成モデルと言っても、演奏する人間の音楽的バックグラウンドによりその認知の仕方、感じ方も様々であり、それが演奏者の個性のような演奏の違いを生んでいると言うことができると著者は考えている。個性が多く反映されることが「創作において評価されるクリエイティビティ」だとするするならば、創作に機械学習を用いる時に音楽の良さ・演奏の個性に絶対的な評価ができなくなってしまう、機械学習を行うためのモデル化の際に特定のモデルを「いいモデル」であると決定することは難しくなる。

また、著者が行ってきたように、訓練された音楽家は、自身の感じ方の分析やとった解決法などを明確に記述できることがわかった。和声感覚の認知では自らの感じ方を元にクラシックとジャズの和声感覚の違いを図示することができた。曲中での 2-5-1 フレーズの演奏習得では、既存のフレーズを使わずに分析することでその生成モデルを獲得することが自身にとって自然であることを「コードトーンが頭の中で鳴らなくなる」という問題と共に考察することができた。

Google Magenta では、ジャンルを問わず同一のモデルにより学習を行う。しかしこのようにクラシックとジャズ即興演奏、または楽器、演奏家のバックグラウンドにより認識・認知が大きく異なることから、よりジャンルや楽器に特化したモデルを用意する必要があるのではないかと著者は考えている。さらに、このように明確に記述された内容が表す音楽的知識に基づく認知・生成モデルは、Google Magenta のような DNN での音楽表現モデルよりも高レベルでの認識・認知であり、同様に音楽認知科学でのモデルよりも高レベルであると言える。また、Google Magenta のモデルは音楽家からみれば非常に非合理的な表現をされているということもでき、音楽家の認知を取り込む、すなわち人間の音楽家と知識を共有できるモデルを導入すれば、機械学習の効率は大きく上がると著者は予想している。

5. おわりに

本稿では、ジャズ即興演奏習得に音楽知識以外の身体性を考慮する必要性があること、著者自身の一人称研究に基づく

ジャズ即興演奏習得過程記録の例を挙げ、習得過程の認知が人によって大きく異なり、単に五線譜上で表現できる音楽知識だけではジャズ即興演奏を表現し切れない可能性、の2つを示した。

今後は、楽器ごとに聞き取り調査などを進め、人間が把握している演奏生成モデルの実例の様々なパターンを作成し、人間と音楽知識を共有する機械学習のための高レベルなモデルへの応用を目指す。

参考文献

- [Amy 86] Dommel-Diny, A., 笠羽 映子：ドビュッシー、演奏家のための和声分析と演奏解釈 / アミ・ドメル・ディエニー著, シンフォニア (1986)
- [Biles 94] Biles, J.: GenJam: A genetic algorithm for generating jazz solos, in *Proceedings of 1994 International Computer Music Conference*, Aarhus (1994), ICMA
- [Cope 05] Cope, D.: *Computer Models of Musical Creativity*, MIT Press, Cambridge (2005)
- [Sudnow 79] Sudnow, D.: *Ways of the Hand - The Organization of Improvised Conduct*, MIT Press (1979)
- [安藤 05] 安藤 大地, Dahlstedt, P., Nordahl, M. G., 伊庭 斉志：対話型 GP を用いたクラシック音楽のための作曲支援システム, 芸術科学会論文誌, Vol. 4, No. 2, pp. 77–86 (2005)
- [安藤 12] 安藤大地：クラブ系ループ音楽パフォーマンスのためのリアルタイムに音楽を Breeding するシステム, 情報処理学会論文誌, Vol. 53, No. 3, pp. 1030–1040 (2012)
- [安藤 15] 安藤 大地：ジャズ即興演奏における4ビートベースラインの2段階対話型遺伝的プログラミングによる生成, 情報処理学会インタラクション 2015 シンポジウム (2015)
- [安藤 17] 安藤 大地：対話型遺伝的プログラミングを用いた無調性音高配列からの和声進行の生成, 情報処理学会インタラクション 2017 シンポジウム (2017)
- [安藤 18a] 安藤 大地：CPPN 表現と Neuro-Evolution を用いたエレクトロニカのためのノイズ音素材生成, 情報処理学会インタラクション 2018 シンポジウム (2018)
- [安藤 18b] 安藤 大地：クラシック演奏家のジャズ即興演奏習得過程の記録と考察, 日本音楽即興学会第10回大会 (2018)
- [大久保 17] 大久保 雅基：携帯音楽プレイヤーを使用した遠隔演奏同期システムの考案, 先端芸術音楽創作学会会報, 第9巻 (2017)