

IJCAI'97 ワークショップ「AI と音楽における課題 – 評価」 の開催報告

平田 圭二

NTT 基礎研究所

hirata@nefertiti.brl.ntt.co.jp

1997年8月に「AI と音楽における課題 – 評価」というテーマでIJCAI'97 併設ワークショップが開かれた。本稿では、そのワークショップ開催の経緯や背景、10件あった発表の内容などを報告する。これまでの音楽情報処理という研究分野は、先駆的な研究の集合という状態に思える。そこで、研究成果の評価法や研究に対する共通認識を確立することが急務であると考え、本ワークショップを企画し議論を行った。インフォーマルなワークショップという場でこのメタな問題を集中的に議論することができ、今後もこの問題を常に意識しつつ研究を進めていく必要があると感じた。

A Report on IJCAI'97 Workshop on Issues in AI and Music – Evaluation and Assessment

Keiji Hirata

NTT Basic Research Laboratories

In August of 1997, the IJCAI'97 workshop on Issues in AI and Music – Evaluation and Assessment was held. This article reports the background and preparation for the workshop, the contents of 10 presentations and so on. It is observed that at present, the research area of music information processing is a collection of individual pioneering work. Thus, we think that it is an urgent necessity to develop a framework of research evaluation and proper assessment of research results from a common viewpoint. We could intensively discuss how to achieve such a situation at the workshop with an informal atmosphere. It is hoped that we have to proceed our research with paying continuous attention to this meta-problem.

1 はじめに

1997年8月23日から29日にかけて、名古屋コングレスセンターにおいてIJCAI'97 (International Joint Conference on Artificial Intelligence 1997) という国際会議が開催された。この会議は、人工知能関連の国際会議としては、AAAI (American Association for Artificial Intelligence) と並んで最も注目度の高い会議と認識されており、取り扱う話題も理論から応用まで人工知能全般に広がっている。IJCAI と AAAI は交互に隔年で開かれて来ており、IJCAI'97 は今回で15回目を数え、総参加者は約1400名であった。そのIJCAI'97の前半23日から25日にかけて29の併設ワークショップが開かれた。本稿で報告する「AI と音楽における課題 – 評

価」のワークショップはその1つであり23日に開催された。

本ワークショップでの発表や議論の内容は、表題にもあるように、音楽情報処理研究における成果の評価法、共通な例題の設定等に関するものであった。筆者の知る限りでは、音楽情報処理研究のワークショップにおいて、世界的に見てもこのようなメタなトピックが取り上げられた機会はこれまで無かったように思う。本ワークショップで発表された方々や参加された方々は、このトピックに対して何らかの問題意識を持っている方々であり、それだけに、当日はこの問題を集中的に議論でき大変有意義だったと思う。しかし一方では、筆者はもっと多くの方々にこの問題に関心を持ってもらいたかったので、内外より10件の発

表と 16 名の参加者という数字に多少の物足りなさも感じている¹。そこで、本稿執筆の機会を利用し本ワークショップで提起した問題の重要性を述べたいと思う。

2 ワークショップ開催の経緯

まず、これまでの AI と音楽に関連した国内外のワークショップや会議において、どのようなトピックが取り上げられて来たかを概観する。

国内の流れ: まず本ワークショップのテーマに関連の深い日本国内の流れとしては、1996 年 2 月の情報処理学会全国大会にて「コンピュータサイエンスとしての音楽情報処理」というテーマでパネル討論会が開かれ、研究の評価と共通例題の設定が議論された [5]。そもそも、音楽情報処理が研究あるいは学問分野として成立するためには客観的な評価や成果の普及が必要であり、そのためには研究者は何をすべきかという内容のパネル討論会であった。パネリストとポジションペーパーのタイトル及びコメントータを表 1 にまとめる。本パネル討論会の様子は文献 [3] に報告されている。大まかな結論として、必要性は認めるが、共通ベンチマークを構築するほど研究コミュニティに知見が蓄積されているようには思えないという意見が大勢を占めた。

さらに同年 5 月には、情報処理学会音楽情報科学研究会にて「コンピュータサイエンスとしての音楽情報処理 - 計算モデルの構築」というテーマでパネルが開かれた [1]。ここでは、2 月のパネルのテーマを一步進めて、音楽情報処理が独立した 1 つの研究分野あるいは学問分野となるためには、広い意味での計算モデルの構築が必要であるという観点からパネル討論会が企画された。同様にパネリストとポジションペーパーのタイトルを表 2 にまとめる。この研究会は日本音楽知覚認知学会との共催だったので、認知系研究者も交え、音楽のための計算モデルは音楽サイドと工学サイドが協調して研究すべきであり、その協調のために研究者はどのような態度で臨むべきかが主に議論された [4]。

¹参加者が少なかった理由はいろいろ考えられるのだが、IJCAI ワークショップ開催ルールに従う必要があった、というのが効いているように感じる。

海外の流れ: AAAI-88, IJCAI-89, ECAL-90, ECAL-92, IJCAI-95 という著名な AI 系国際会議でも、不定期²ではあるが AI と音楽というテーマで併設のワークショップが開かれている。例えば IJCAI-95 では、“AI Models of Structural Music Understanding” というテーマで Gerhard Widmer がワークショップを主催した [2]。これら一連の併設ワークショップは、AI 系国際会議という状況から、人間の音楽的な振舞い (音楽的な知性) を計算機システムで模擬することが主題となっている。つまり、音楽情報システム構築には (計算機科学的な意味において) どんな知識処理の枠組が適しているのか、音楽はどんな知識構造を持っているのか、またどのように音楽知識を表現すれば良いのか、ということである。実際、発表としては、演奏の表情付け、ビートトラッキング、自動編曲、自動伴奏、音楽解釈などのテーマが多く、音楽ジャンルとしてはクラシックやジャズが多い。

本ワークショップの準備: 本ワークショップの企画、準備は 1996 年 10 月頃から開始された。IJCAI 併設のワークショップを開催するためには、プロポーザルを IJCAI'97 に提出し承認される必要がある。そこで、Roger Dannenberg (Carnegie Mellon University (米国)), Peter Desain (Nijmegen University (オランダ)), 平賀譲 (図書館情報大学), 平田圭二 (NTT), Henk-Jan Honing (University of Amsterdam (オランダ)) が発起人となり、日本在住の平賀、平田を中心にプロポーザル作成が進められた。そして、前述のような国内外の状況を考慮し、現在の AI と音楽の研究コミュニティにおいて最も議論すべきと思われるテーマを選定した。

まず、現時点を個々のパイオニアによる研究がある程度充実してきた時期ととらえ、これまでお互いある程度の距離を置いて研究を進めている状態から脱却すべきではないか、そして今後は、研究コミュニティで何らかの共通認識や共通基盤を持ち、そこで情報交換し、互いの研究を理解し評価し合い、研究成果を体系立てて整理するという作業が必要なのではないかと考えた。他の研究分野においては当たり前の作業のよう

²これらのワークショップは、有志がプロポーザルを提出して採択された時のみ開催できるため。

パネリスト

片寄 晴弘	(財) イメージ情報科学研究所	演奏は誰が上手?
後藤 真孝	早稲田大学大学院 理工学研究科	ビートトラッキングにおけるベンチマーク
堀内 靖雄	千葉大学 工学部	伴奏システムの協調動作の評価
松島 俊明	東邦大学 理学部	楽譜認識は本当に難しいのか
村尾 忠廣	愛知教育大学 音楽教室	音楽認知と音楽情報科学 ~ 音楽情報科学 <から> の <ための> 音楽用語の再検討

コメンテータ

志村 哲	大阪芸術大学 芸術学部
葉 孝之	国立音楽大学 音楽デザイン学科

表 1: 1996 年 2 月情処全大のパネリストとコメンテータ (敬称略五十音順)

井口 征士	大阪大学 基礎工学部	(パネルのタイトルに同じ)
小坂 直敏	NTT 基礎研究所	音響分析・合成における計算モデルについて
竹内 好宏	京都府立 亀岡高校	音楽認知研究から演奏や音楽教育へのモデリング
橋本 周司	早稲田大学 理工学部	感性的メディア技術としての音楽情報処理とその評価
平賀 謙	図書館情報大学	(パネルのタイトルに同じ)

表 2: 1996 年 5 月情処音情研のパネリスト (敬称略五十音順)

に聞こえるが、音楽的な知性という特殊な研究対象に、このような伝統的アプローチで研究を進めるのは冒険かも知れない。しかし、音楽的な知性を研究する最良の方法論は伝統的アプローチの延長線上にあると仮定すれば、まず我々は伝統的アプローチで研究を進めて行くしか方法がないように思えるのである。

この共通認識や基盤を持つというメタな問題は、個々の研究者の倫理的な態度から科学的な方法論に到るまでの大変広い範囲を含んでいる。それだけに、ワークショップというインフォーマルな場が議論に適していると考えた。論文募集に掲載したワークショップで受付けたトピック一覧を表 3 に示す。さらに、このワークショップがこれまでの研究成果を広い視野のもとで整理し分類し、戦略的な研究プログラムの策定やベンチマークのような標準的なデータ集を構築するようなプロジェクトに発展することを期待して、本ワークショップを開催した。

3 各発表について

プログラムは 10 発表 3 セッションから成り、発表の順序は抽象的なトピックから段々具体的なトピックに流れるように組まれている(表 4)。当日は平賀謙氏が司会を務めた。以下、各発表の

内容を要約する(敬称略)。

(1) Milicevic の主張は、音楽を作り出す知識構造(脳)や音楽を聴く知識構造(脳)はすべて主観的であるという徹底したものである。何故なら、その人が成長してきた経験により構成されたニューロンの結合によってのみ音楽は解釈され演奏されるのであり、それは常に変化し続け、そこに絶対的なものは存在しないからである。だから逆に、そのような脳の機能を模倣して、音楽を理解したり演奏したりする機械を作るとは科学的に非常に興味深いと考える。

(2) 平賀の発表は、AI と音楽の研究領域における評価や共通認識を考える時に、最も基本となる問題は何かを議論し整理した。例えば、評価の方法は知見の蓄積とともに進化し続けるという循環性、音楽固有の主観と客観という対立概念、実世界の現象を抽象化する時の問題点、人間の認知活動としての音楽とそれ自体閉じた存在としての音楽の差異、演奏者と聴取者それぞれの計算モデルの評価の非対称性、妥当な計算モデルとはどうあるべきか等の点が指摘され、これらが研究成果の評価とどのように関係しているのかを議論している。

General issues

- significance and difficulties of evaluation
- merits and demerits of evaluation
- what is to be/can be evaluated
- action program, authorization

Individual practice

- criteria of a "sound" presentation
- assessment of existing work
- comparison of related works

"Subjective/Objective" duality

- distinction and characterization of subjective/objective aspects
- subjective aspects of music
- objective aspects of music

Methodology

- viewpoints of evaluation (methodical, performance, etc.)
- specific topics and forms of evaluation
- enhancing objectivity of evaluation
- formalization/systematizing of evaluation methods
- evaluation methodology in AI, psychology, and other related fields

Common workgrounds

- terminology
- common problem sets, benchmarks
- corpora, data archives, standard formats
- (multimedia and network) support for sharing research results
- forums for discussion and critical reviewing

表 3: 論文募集に掲載されたトピック一覧表

(3) Rousseaux と Pachat が製作しているジャズ即興システムを例にとり、音楽的なシステムを構築する方法論について述べた。Ontology さえきちんと定義すれば、後は記号の世界で知識を記述でき、記号の世界は客観的である。曖昧な部分を含む音楽という対象を扱うシステムを従来の自然科学的な手法だけで評価するのは難しいので、音楽的チューリングテストを提案する。音楽的な知識を可視化する例として、Harmonic space と Tonal analysis システムが紹介された。

(4) 音楽システムが目的とする仕事をどれ位うまく実行しているかを評価するには、まず最初

Philosophy

- (1) Mladen Milicevic (Univ. of South California)
Objectivity in Music: An Impossible Fantasy
- (2) 平賀 謙 (図書館情報大学)
Some Considerations on the Basis of Evaluation in AI and Music
- (3) Francis Rousseaux and François Pachat (Université Paris 6)
Distinction and Characterization of Subjective/Objective Aspects for the Design of Artificial Assistance Systems in "AI and Music" Field
- (4) Lorin Grubb (CMU)
The Importance of Metrics for AI and Music[†]

Methodology

- (5) Peter Desain (Nijmegen University) and Henkjan Honing (University of Amsterdam)
How to Evaluate Generative Models of Expression in Music Performance
- (6) 橋本 周司 (早稲田大学)
On Evaluation and Assessment of Musical Art Technology – Universality and Individuality in Kansei Aspect –
- (7) 村尾忠廣 (愛知教育大学)
Evaluation by Computational Performance Studies on Analyses of Mozart's Theme in A Major Piano Sonata

Experiment

- (8) 平賀 瑠美 (筑波大学)
Objective Evaluation for Computer Generated Musical Performance
- (9) Roger B. Dannenberg (CMU)
Meaningful Evaluation of Style Classification Techniques[†]
- (10) 後藤 真孝, 村岡 洋一 (早稲田大学)
Issues in Evaluating Beat Tracking Systems

†: ビデオテープによる発表

表 4: ワークショップのプログラム

に、用いる評価尺度を決めることが重要である。評価尺度を決めると、比較、精密な分析、改良が容易になる。安定な一貫した結果が得られるように実験の状況を限定すれば、そこでは客観的な実験が可能となる。これらが、共通基盤を醸成する必要条件となるだろう。

(5) 最初からシステム評価を考慮した計算モデルを作るアプローチの紹介. これまで演奏の表情付けの(計算)モデルは多く提案されて来たが, どれも部分的なものであった. それらを SECT (Structural Expression Component Theory) という枠組で組み合わせ, analysis-by-synthesis でパラメータを最適化する. すると, 重なっていた音楽構造を分離でき, 各計算モデルの評価が可能となる. 今後, SECT を実装したシステム DISSET を用いて実験を行うが, 組み合わせの方法, 最適化の方法などの課題を解決する必要がある.

(6) 音楽などのアートにおいては, 特に emotion という非論理的な部分をどう扱うかが問題であり, 橋本は感性情報処理という枠組みで研究を展開している. 感性情報処理が従来の情報処理の枠組みと異なるのは, emotion の共有に由来する reality が評価の対象となる点, 一般性より個人性の方が重要である点などである. 音楽を情動的な面から評価できるのは人間のみであり, 聴衆を模倣するシステムを評価する時は, 複数の視点で行うべきである. もし感性システムのベンチマークが存在しその結果が平均として出力されるとしたら, それは評価としてはあまり意味がないだろう.

(7) モーツァルトのイ長調ピアノソナタを分析する研究は多いが, この曲で分析手法を評価するという仕事はない. 多くの解釈ができるからこの曲が題材として選ばれた. メロディ, リズム, 和声の構造やそれらの複合構造に関して, いろいろな解釈をしても客観的な正しさが証明できないので, 結局その分析法に研究者が納得できるかどうかになってしまう. 以前は楽曲の統計的性質をもとに分析法の比較を試みたが, 楽曲の構造との関係を明らかにできず行き詰まってしまった. 現在村尾らが進めている計算機を利用した実験的手法なら, 構造を定量的に分析できる可能性がある.

(8) 計算機科学の研究室において自動演奏システムを客観的に評価する方法を提案した. 基本的には, 比較すべき演奏を用意して, 演奏パラメータの差異や演奏法の差異を定量的に測定する. ここで演奏法とは曲の構造と演奏パラメータの

関係を意味する. 比較出力の差分は視覚化ツールを利用して表示すると効果的である. この評価法の利点は, 手軽に実行でき定量的な結果を出せるという点, 最初に比較演奏さえ用意すればあとは機械的に評価が行える点であり, 逆に課題曲の選定, 比較演奏の用意, 人間の評価と食い違った時の対処などが問題となろう.

(9) Dannenberg らのスタイル分析器製作の経験を踏まえ, そこで生じたシステム評価の問題を議論した. スタイルはインフォーマルな概念であり正確に定義することは困難である. スタイル分析器があるメロディを聴いた時, 音符数や音域など個々のパラメータの値を正確に出力する必要はなく, 答えとして正しい意図を出力しさえすれば OK と考えた. すると分析器の正しさとは演奏者の意図と分析器の出力との合致と等価になる. 訓練データの作成時, 主観に基づいてメロディの断片にラベルを付けたのでゆらぎが生じてしまった. 過学習の問題は cross validation で避けた. 実際のライブでインタラクティブ演奏を行っている状況では, 音楽的文脈を考慮した訓練データで学習していなかったため演奏者の意図と食い違うこともあった. 訓練データが与えられている環境での分析器の評価は容易だが, いかにして良い訓練データを与えるかが根本的な問題である.

(10) ビート追従システムを評価するための共通の指標を提案し, 現在後藤らが製作しているシステムに適用し例題 40 曲で調べた. その評価プロセスは, オーディオ信号のレベルで人間がビート位置にラベル付けし, それをビート追従システムの出力と比較し, リズムの階層構造に合わせて 2 小節, 1 小節, 2 分音符, 4 分音符, 8 分音符の各レベルでビート位置のズレを測定する. 評価結果は, そのズレから計算された平均, 分散, 最大値とビート追従システムにおいて典型的な 2 種類のエラーを示すフラグ (テンポが 2 倍か 1/2 倍と誤認識されているかどうか, 位相のズレが 0 度か 180 度か) から成る. この指標はまだ完全とは言えないが, ビート追従の認識率や正しくトラッキングするまでの立ち上がり時間などが測定できるようになった.

General Discussion 本ワークショップでは、全発表が終了した後 30 分程の自由討論を行った。各発表者へ今日の発表では何故その曲を例題として選んだのかという質問に対し、主張したいポイントをアピールするのに適していた、複数の解釈が可能であるなどの回答があった。共通の例題曲を設定することに関しては、そういうものは自然発生に任せておくべきもので、意図的には何もする必要がない、あるいは有志がフリーで質の良い入力データを web 等で公開するよう心掛けるだけで十分ではないかという意見があった。

4 おわりに

音楽情報処理という言葉は新しいが、その研究テーマ自体は 15 年ほど前 (1980 年代前半) から認識されていたと思う (もちろんそれより以前の先駆的な研究もある)。しかし、音楽情報処理においてこれまでどんな際立った (工学的な科学的な) 研究成果が得られているのかと問われた時、答えに窮する人は少なくないだろう。その理由は、音楽の持つ本質的な困難さや、研究に対する問題意識が自然科学寄りだからなのかも知れない。あるいは、問題意識の持ち方や研究目標設定自体の困難さのためかも知れないし、芸術的な要素や主観が関係しているため、研究成果の提示が分かりにくいからなのかも知れない。しかし、それならば、これまで持っていた問題意識や方法論をもう一度見直してみる必要があると思う。この意味で、本ワークショップは音楽情報処理において研究成果の評価と共通基盤の醸成という面を見直す作業の一環であったと考えている。

しかし今の時点では、どんな方向に研究を進めるべきなのか、どんな方法論が正しいのかということは全く分からない。むしろ、個々の研究者がこのテーマに対する意識を高く持ち続けることが大切である。今後も機会があれば、このテーマに関して議論する場を企画したいと考えている。

謝辞: 本ワークショップにて発表して下さった方々、議論に加わって頂いた方々に感謝致します。本ワークショップのためのプロポーザルの作成や CFP の作成は local organizer の一人である平

賀譲氏に依る所が大でありました。ここに感謝致します。

参考文献

- [1] 井口他, パネル討論: コンピュータサイエンスとしての音楽情報処理 - 計算モデルの構築, 情報処理学会研究報告 96-MUS-15, pp.33-37, 1996.
- [2] G. Widmer (ed.), Working Notes of IJCAI-95 Workshop on AI and Music, 1995.
- [3] 片寄他, パネル討論会「コンピュータサイエンスとしての音楽情報処理」の報告, 情報処理学会研究報告 96-MUS-15 付録, pp.91-98, 1996.
- [4] SIGMUS, 第 15 回研究発表会質疑記録, 情報処理学会研究報告 96-MUS-16, pp.58-59, 1996.
- [5] 平田, コンピュータサイエンスとしての音楽情報処理, 情報処理学会 第 52 回全国大会シンポジウム・インダストリアルセッション論文集, March, 1996. (<http://www.brl.ntt.jp/people/hirata/Papers/mip-as-cs.dvi.gz>)