

音楽ゲームのプレイヤー AIにおける 人間らしく振る舞う強化学習手法の提案

坂本 洸¹ 橋本 剛¹

概要: 現在, ゲーム AI は多くのゲームジャンルで人間の強さを超えており, 近年は人間らしいゲーム AI の研究が注目されている. しかし, 音楽ゲームのようなアクションに対する評価や最高スコアまでの過程が明確なゲームはあまり題材として扱われていない. 本研究では, 音楽ゲームを題材に人間らしい振る舞いを行うゲーム AI を作成することを目標とする. 強化学習に生物学的制約を導入することによって, 人間らしい失敗を表現し, 振る舞いやスコアの変化を検証した. 主観評価実験の結果, 生物学的制約を導入した AI の一部は実際の人間よりも人間らしいと評価され, 極端な失敗を増加させることよりも, 動きの微妙なズレを再現することは人間らしく評価されることがわかった.

Proposal of reinforcement learning method that behaves like a human being in music game player AI

Abstract: Recent game AI has exceeded the strength of human beings in many game genres, and in recent years, study on human-like game AI has attracted attention. However, human-like game AI with a clear process for getting the perfect score, like music games, has not yet been studied. The purpose of this study is to create an AI that behaves like a human in a music game. Reinforcement learning expresses human-like mistakes by imposing biological constraints. As a result of subsequent subjective evaluation experiments, AI with some biological constraints was evaluated as more human-like than human, and AI with small mistakes was evaluated as more human than AI with large mistakes.

1. はじめに

近年, コンピュータゲームやボードゲームの AI に関する研究が盛んに行われている. ボードゲームでは, 特に局面数が大きく AI の開発が難しいとされていた囲碁において, AlphaGo がプロ棋士を破ったことが大きく話題となった [1]. コンピュータゲームでは, Atari 2600 の古典的ゲーム 57 タイトルを強化学習手法の 1 つである R2D2 (Recurrent Replay Distributed DQN) を用いてゲーム画面のみを用いて学習させたところ, 57 タイトル中の 52 タイトルのゲームで人間を上回るスコアを記録した [2]. このように, ゲーム分野において AI は多くのゲームタイトルで人間を超える強さを獲得している.

一方で, 強さを追求したゲーム AI では, 人間がするとは思えないアクションや一手を出すなど, 動きが機械的であることが課題として挙げられる [3]. 例えば, 将棋やオセロ

などで AI が制御する敵側プレイヤーが完封試合を行ったり, 対戦型シューティングゲームで人には全く敵わないような動きでプレイヤーを圧倒してしまうと, プレイヤーはゲームを楽しめなくなるという可能性がある. この課題を解決するために, 人間らしい振る舞いを行うゲーム AI や, プレイヤーを楽しませるゲーム AI に関する研究が行われている. これまでのコンピュータゲームでは, 俗に COM や CPU と呼ばれる対戦相手や味方の強さや振る舞いを開発側の膨大なコーディングによって実現してきた. 機械学習によって自動獲得された AI による人間らしい振る舞いや, 自在な強さの調整が可能になれば, これらの開発の大幅なコストダウンに繋がる. それだけでなく, プレイヤーに対する「手加減」や「焦り」などのより人間らしい振る舞いの実現によってプレイヤーを楽しませたり, プレイヤーと同程度の実力を持つようなライバルになれる AI の実現が期待できる. 人間らしい AI を仮想的なプレイヤーとして設定することで, ゲーム開発におけるデバッグや難易度調整に活用することもできると考えられる.

¹ 松江工業高等専門学校
National Institute of Technology, Matsue College

これまでに人間らしい振る舞いを行うゲーム AI の研究として、ボードゲーム [4] やアクションゲーム [5], シューティングゲーム [3] など、行動に多くの選択肢があるゲームを題材とした研究は行われているが、ボタンをタイミングよく押す音楽ゲームのような、アクションに対する評価や最高スコアに至るまでの過程が明確なゲームはあまり題材として扱われていない。ゲーム AI で人間らしさを表現する手法として生物学的制約を導入する事例 [6] があるが、それらではゲーム内部のパラメータなどを参照する必要があり、他のゲームへの応用が難しく、汎用性が低い点が課題である。また、人間はキャラクタの正確な位置座標などを正確に取得できるとは限らない。本研究では、音楽ゲームを題材に、汎用性を高めるために人間が取得できる情報と同等の情報のみで強化学習に生物学的制約を導入し、人間らしい振る舞いを自動獲得するゲームプレイヤ AI を作成し、得られた振る舞いの比較や人間らしさの評価を行う。

2. 音楽ゲーム

本研究では、ゲームの進行に譜面を用いる音楽ゲームを対象とする。譜面を用いる代表的な音楽ゲームとして「太鼓の達人*1」や「beatmania*2」などが存在する。これらの音楽ゲームは、共通する基本的な構成要素として「ノート」「レーン」「判定ライン」「入力デバイス」が存在する(図1)。また、これらが画面に表示されていないものでもそのように捉えることができる場合が多い。ノートは音楽に合わせてアクションを行うタイミングを示した音符のようなものである。音楽ゲームではアクションの種類が複数あることが多く、例えば太鼓の達人では「太鼓の面を叩く」「太鼓の縁を叩く」「連打する」などのアクションが存在する。そして、このノーツの集合がリズムをどのような流れでとるかを決め、音楽用語と同じように「譜面」と呼ぶ。これは横スクロールアクションゲームなどでの「ステージ」に該当する要素である。ノートは画面の外から出現して、レーンに沿って判定ラインに向かって流れていく。そしてノートが判定ラインに到達したタイミングでボタンを押すなど、対応する入力デバイスを操作する。そのタイミングが音楽と合っていればスコアが増えるなどの良い判定を得る。また、多くの音楽ゲームでは複数の判定が存在する。決められたタイミングに合っているほどスコアが増えるなどの良い判定を得て、タイミングがずれるほど増えるスコアが少ない、もしくは増えないなどの悪い判定を得る。全てのノートに対してもっとも良い判定を得るとパーフェクト(最高スコア)となる。パーフェクトや高スコアを目指すなど、いかに良い評価を得るかを目標にプレイすることは音楽ゲームの

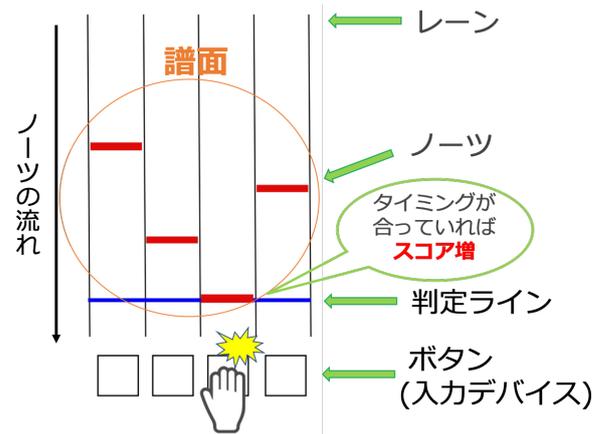


図 1 譜面を用いた音楽ゲームの一般的な構成

Fig. 1 General composition of music game using chart

楽しみ方の大きなひとつである。

マリオのようなアクションゲームにおいても、最短でゲームをクリアした際の残り時間やその最短経路自体を最高スコアとして捉えることは可能であるが、多くの行動の選択肢があるため、その過程を求めることは難しく、最高スコアが明瞭であるとは言えない。一方で音楽ゲームにおいては全て完璧なタイミングでボタンを押せば最高スコアとなるためその過程は非常に明瞭であると言える。

3. 先行研究

人間らしい振る舞いを行うゲーム AI や、その振る舞いを自動獲得するゲーム AI の研究がいくつか行われている。

仲道らは、これまで強くすることを目的とすることが多かった将棋において、ユーザの棋力に合わせる接待将棋 AI を作成し、指し手の人間らしさを検証した [4]。

藤井は戦略型ビデオ TCG(トレーディングカードゲーム)と Infinite Mario Bros(スーパーマリオを模したゲーム)を対象に、人間らしい振る舞いを自動獲得するゲーム AI の実装と人間らしい振る舞いの評価基準に関する検討を行っている [6]。TCG については、最適行動を選択する機構を設計することでプレイヤに戦略レベルで適応するゲーム AI を実現している。Infinite Mario Bros については、人間がゲームをプレイする際に発生する生物学的制約を強化学習に導入することで、人間プレイヤに人間らしいと解釈される振る舞いを表出するゲーム AI を実現している。生物学的制約とは、人間が生得的に持っている制約や欲求であり、人間がゲームをプレイする際にも生じているものである [5]。例えば、生物学的制約を導入したマリオは落ちるとゲームオーバーになってしまう穴がある地点で、穴の直前でパフォーマンスを重視したジャンプをするのではなくある程度離れた位置から余裕を持った安全なジャンプをするようになった。また、ステージを進む際にも単に一直線に進むのではなく、緩急のある行動を取るようになり、人間らしいと評価を得ていた。これらは、「ゆらぎ」や「遅れ」な

*1 BANDAI NAMCO Amusement Inc. 太鼓の達人シリーズ,2001-2019. Video Game.

*2 Konami Amusement Co., Ltd. beatmania シリーズ,1998-2019. Video Game.

どの生物学的制約によって、より安全な行動をしなければ報酬が得られなくなったからだと考えられる。

また、冬木らは音楽ゲームにおける、難しい地点では正確度を低く、簡単な地点では高くなるといった「人間らしさ」を表現したプレイ結果のシミュレーションを行なっている [7]。これは重回帰分析を用いて、人間らしさを表現する要素を数値的に加えることでプレイヤーの熟練度を基に得点を再現している。

4. 研究方針

音楽ゲームのような最高スコアに至るまでの過程が明確なゲームで人間らしさを表現するための手段は限られており、失敗の行動によって表現する必要がある。加えて、ただ単に最終的な得点が低くなるような失敗をするのではなく、難所で発生するミスやゲーム終盤の緊張によるミスなど、「人間らしいミス」を再現することが重要であると予想する。このようなミスは、人間の身体的、精神的なものも多く、これらの振る舞いを獲得するためには音楽ゲームにおいても生物学的制約の導入が有効であると考えられる。しかし、既存の人間らしい振る舞いを行うゲーム AI で生物学的制約など人間らしさを再現する処理を施す際には、自機や敵機の位置座標やステージに存在するオブジェクトの情報など、対象としているゲーム固有のパラメータを参照する必要がある。そのため対象としたゲームに多くの人間らしさを再現する処理が依存してしまい、他のゲームへの応用が難しく汎用性が低い点が課題である。また、実際に人間がゲームをプレイする際も、そのような情報やゲーム内部のパラメータを正確に全て把握できるとは限らない。そこで、ゲーム画面のみの学習で成果を上げている深層強化学習で人間らしさを導入することができれば、ゲーム固有の情報を必要としない汎用的な人間らしい振る舞いを行う AI の実現に繋がると考えられる。

本研究では、深層強化学習に対して生物学的制約の導入を行い音楽ゲームのプレイヤー AI を作成し、得られた振る舞いの人間らしさと、音楽ゲームにおける人間らしい振る舞いについて検証する。

5. 新たな生物学的制約と処理方法の提案

生物学的制約は、藤井が定義したものを参考にして、内部のパラメータなどを用いない処理に変更して「揺らぎ」「認識の遅れ」「行動の遅れ」「緊張」の4種類を定義した。既存処理では自機の座標などのパラメータに対して変更を与えていたものを、入力となる画像とキー操作に対して与え、操作の連続成功時に緊張する状態を「緊張」として新たに定義した。以下でそれぞれの生物学的制約について説明する。

● A : 揺らぎ

人間は、ゲーム画面を見る際に自機や敵機の座標や位置を正確に認識することは難しい。また常に同じ箇所

に視線が向いているとは限らず、視線のズレが生じる。[6] では自機の位置座標や局面情報にガウスノイズを付与することで再現していた処理を、入力するゲーム画面の画像全体の範囲をずらす処理に変更して揺らぎを再現する。

● B : 認識の遅れ

人間は、ゲーム画面を見て局面を認識してから行動に移るまでに遅延が生じる。[6] では自機の位置座標や局面情報を数百ミリ秒過去の情報にすることで再現していた処理を、入力するゲーム画面の画像自体を数フレーム過去のものにすることで認識の遅れを再現する。

● C : 行動の遅れ

人間は、前述のような認識から行動までの遅れだけでなく、実際に行動を開始してから指や手がキーボードなどのオブジェクトを操作するまでの間にも遅延が生じると考えられる。そこで、AI がキー操作を行う際に遅延処理を挿入することで行動の遅れを再現する。

● D : 緊張

人間は、成功が連続している時に失敗したくないと感じ、緊張が生じることが多々ある。また、緊張した際には、普段ではしないような失敗や自分の意図に反した行動を取ってしまうことがある。そこで、連続成功時にランダムな行動を取る確率を増加させることに加え、前述の行動の遅れを発生させることで緊張を再現する。

6. 実装

深層強化学習の1つである DRQN[8] を用いて既存の音楽ゲームである「osu!mania[9]」のゲームプレイヤー AI を作成し、ゲーム画面の情報のみで学習した。

6.1 osu!mania の仕様

「osu!mania」は PC でプレイすることができる音楽ゲームである。世界中で多くのユーザがプレイしており、ユーザが譜面の作成を行うことが可能である。「osu!mania」は、複数のゲーム形式を持つ「osu!」のゲームモードの1つであり、上から降ってくるノーツが判定ラインに重なるタイミングでキーを押す音楽ゲームである。また、「osu!mania」では操作するキーの数を変更することができ、今回は4つのキーを操作する設定にした。図2に今回用いる「osu!mania」のゲーム画面を示す。

「osu!mania」には6種の判定があり、良い順に「虹300」「黄300」「200」「100」「50」「0(ミス)」である。各判定の数値は、スコアの計算時に用いる。また、それぞれの判定が得られる指定されたタイミングとのズレの許容時間の幅を「判定幅」と呼ぶ。表1に各判定の判定幅を示す。例えば、正確なタイミングから60msずれていた場合は、49.5ms~82.5msの間であるため判定200を得る。ゲームスコアは達成率(%)で表され、以下の式で求められる。



図 2 osu!mania のゲーム画面
Fig. 2 Game screen of osu!mania

表 1 各判定の判定幅

Table 1 Judgment width of each judgment

| 判定 | 判定幅 (± ms) |
|-------|------------|
| 300 虹 | 16.5 |
| 300 黄 | 49.5 |
| 200 | 82.5 |
| 100 | 112.5 |
| 50 | 136.5 |

$$\text{達成率} = \sum_{i=\text{判定 } 0(\text{ミス})}^{\text{判定 } 300 \text{ 虹}} \frac{\text{各判定の数値} * \text{各判定の個数}}{\text{全ての判定の個数} * 300} \quad (1)$$

最高スコアは 100% であり, 全てのノートで 300 以上の判定を得ると 100% となる。

6.2 ネットワーク構成

強化学習のニューラルネットワークには DRQN (Deep Recurrent Q-Network) を用いる。DRQN は深層強化学習の代表的な手法である DQN (Deep Q-Network)[10] の CNN (Convolutional Neural Network) 層と全結合層の間に, 時系列データに対するモデルである LSTM (Long Short-Term Memory) 層を追加した手法である。LSTM 層の追加によって過去の情報が長期的に渡って活用できるようになり, Atari2600 のゲームで DQN よりも高いスコアを記録している。表 2 にネットワーク構成を示す。入力は「osu!mania」のゲーム画面 (800*600) を 1/4 のサイズ (200*112) に縮小した画像データで, 出力はどのキーを押すかを決定する行動評価値ベクトルとした。

6.3 学習の流れ

図 3 に学習の流れを示す。既存のゲームを用いるため, ゲームの操作では実際にキー入力を行い, どの報酬を与えるかは画面の画像から判別する必要がある。そのため, Python の GUI 自動化ライブラリである PyAutoGUI[11]

表 2 DRQN のネットワーク構成
Table 2 DRQN Network Configuration

| 層 | 構成 |
|-----|---------------------------|
| 入力層 | 200*112*3 |
| 中間層 | CNN3 層, LSTM1 層, 全結合層 1 層 |
| 出力層 | 行動評価値ベクトル |



図 3 学習の流れ

Fig. 3 Learning flow

表 3 学習のハイパーパラメータ

Table 3 Learning hyperparameters

| パラメータ | 設定 |
|-------------|----------------------------------|
| exploration | ϵ -greedy 法 (0.4 → 0.1) |
| Bach Size | 64 |
| γ | 0.98 |
| Optimizer | AdaDelta |

を用いて, 学習した行動評価値に基づいてキー入力を与え, ゲーム画面をキャプチャした画像を入力として与えると同時に, 得点の表示を検出して報酬を与えた。またゲーム画面のキャプチャは 0.1 秒ごとに行い, これを 1 フレームとした。AI が行う行動は, 4 つの操作するキーそれぞれを押すかと, 何も押さないの 5 つから選択するようにした。表 3 に学習のハイパーパラメータを示す。 ϵ -greedy 法で最初はランダム行動選択率を 40% とし, 学習が進むにつれ確率を下げ, 最終的に 10% まで下げるようにした。表 4 に与えた報酬をまとめたものを示す。報酬は良い判定ほど高い報酬を与えた。また, 0(ミス)と空打ちには負の報酬を与えた。そして生物学的制約は, ゲーム画面の入力時に揺らぎと認識の遅れを与え, AI が行動する際に行動の遅れと緊張を与える。

6.4 生物学的制約の導入

定義した生物学的制約を再現する処理を実装した。実際の処理内容を以下に示す。

- A : 揺らぎ

10% の確率で, 5 フレーム (0.5 秒) 間, ゲーム画面の

表 4 与える報酬
Table 4 Reward value to give

| ゲーム内判定 | 報酬値 |
|--------|-----|
| 300 虹 | 200 |
| 300 黄 | 150 |
| 200 | 30 |
| 100 | 15 |
| 50 | 10 |
| 0(ミス) | -50 |
| 空打ち | -1 |

キャプチャする範囲の座標を-20~20ピクセルの間でランダムにずらす。

● B：認識の遅れ

入力するゲーム画面の画像を1フレーム(0.1秒)前のものにする。

● C：行動の遅れ

キー操作時に0.05秒の遅延処理を挿入する。

● D：緊張

10連続で0(ミス)を出さなかったら、ランダム行動選択率を2倍にして、キー操作時に0.05秒の遅延処理を挿入する。

7. 実験1 モデルの作成とスコアの比較

7.1 実験概要

実装した学習環境を用いて、各生物学的制約を導入して強化学習を行うモデルを作成した。モデルは以下の5つを作成した。

- A：揺らぎ
- B：認識の遅れ
- C：行動の遅れ
- D：緊張
- E：生物学的制約の導入無し(以下「無し」とする)

AIにプレイさせる譜面は、1プレイ20秒の楽曲*3で24個のノーツが流れてくる譜面を「osu!mania」の譜面作成機能を用いて作成した。各モデルを、作成した譜面を用いて1000プレイずつ学習させ、生物学的制約の有無や種類による振る舞いやスコアの変化を検証した。

7.2 実験結果

作成した各モデルのゲームスコアの推移(移動平均を20で取ったもの)を図4に、スコアの変化がある程度収束したと考えられる、プレイ回数900~1000回の平均スコアを表5に示す。最終的な平均達成率が「モデルE：無し」では85~90%程度と高スコアであるのに対し、生物学的制約を導入したモデルA~Dでは全てで達成率がモデルEよりも下回る結果となった。このことより、生物学的制約の導入によりミスが増加したことがわかる。また、生物学的制

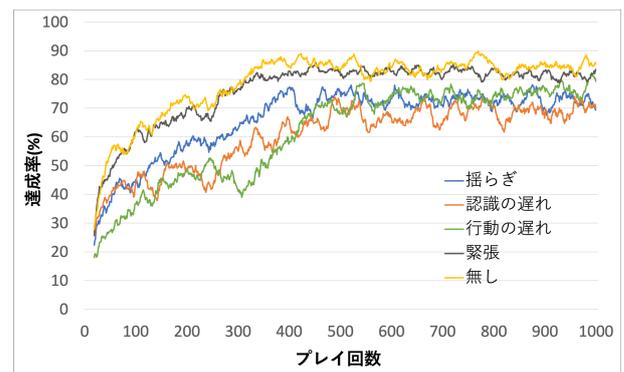


図 4 スコアの推移(移動平均 20)

Fig. 4 Transition of score(moving average 20)

表 5 プレイ回数 900~1000 回の平均スコア

Table 5 Average score of 900-1000 plays

| モデル | A | B | C | D | E |
|---------|-------|-------|-------|-------|-------|
| スコア (%) | 72.70 | 69.58 | 76.83 | 81.36 | 84.91 |

約によるスコアへの影響は「モデルB：認識の遅れ」が最も大きかった。

図5に、プレイ回数900~1000回の各モデルの判定構成比を示す。横軸は判定の種類、縦軸はその判定が1プレイの中でどれだけ取得されたかの比率を示している。「モデルA：揺らぎ」では最低判定の割合が増加し、「モデルB：認識の遅れ」では最高判定と最低判定の両方が増加した。また、「モデルC：行動の遅れ」では他のモデルと比べて極端に判定200が極端に増加した。

各モデルのプレイ中の様子*4を図6に示す。実際にAIのゲームプレイ映像を見て振る舞いを確認すると、「モデルA：揺らぎ」、「モデルB：認識の遅れ」では特に0(ミス)の増加が目立っていたが、ミスの際にそのタイミングでは押す必要がない他のキーを空打ちするような振る舞いが多く見られた。「モデルC：行動の遅れ」では、0(ミス)までにはならず判定200を得る程度のリズムのズレが増加しており、生物学的制約によって異なる挙動が見られた。「モデルE：無し」では、多くのプレイで全てのノーツを判定300以上で取得しており、ほぼパーフェクトのスコアも確認した。

7.3 生物学的制約を後から導入した場合

追加実験として、生物学的制約が与えられなかったモデルに対して後から生物学的制約を導入した場合にスコアの推移や得られる振る舞いに変化があるのかを調べるため、学習済みの「モデルE：無し」に対して生物学的制約を導入し、同様の楽曲を用いて更に1000回学習させた。後から各生物学的制約を導入した場合と「モデルE：無し」のゲームスコアの推移(移動平均を20で取ったもの)を図7に、プレイ回数900~1000回の平均スコアを表6に示す。既

*3 otosozai.com <http://www.otosozai.com>

*4 松江高専 情報工学科 橋本研究室
<https://youtube.com/channel/UCe2u8IcMBVnDTinE3ypHiIg>

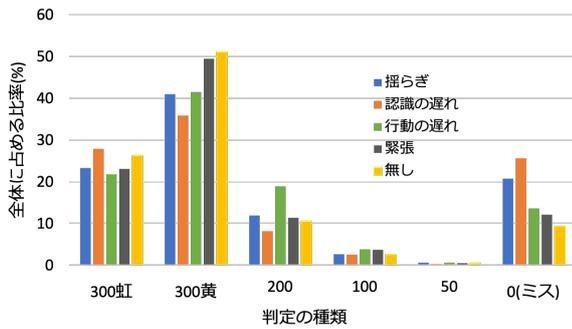


図 5 各モデルの判定構成比

Fig. 5 Composition ratio of each judgment for each biological constraint

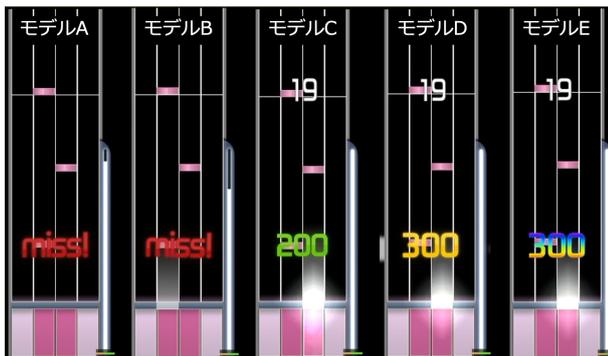


図 6 各モデルのプレイ

Fig. 6 Play of each model

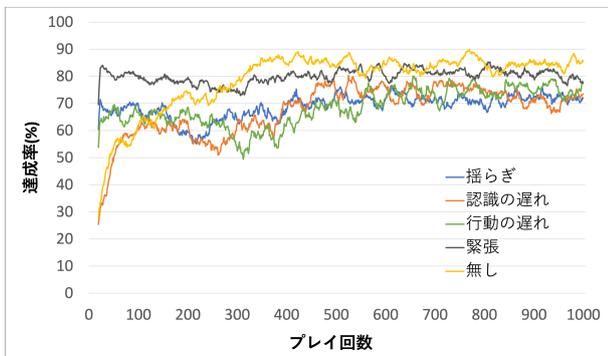


図 7 後から生物学的を導入した場合のスコアの推移 (移動平均 20)
Fig. 7 Transition of score when biological constraints are introduced later (moving average 20)

に学習済みではあるため、図 4 の最初から導入した場合と比べて序盤のスコアは高かったが、最終的な達成率に大きな差は見られなかった。また、200~300 回辺りでは生物学的制約を導入したどのモデルでもスコアが若干低下する傾向が見られた。

8. 実験 2 主観評価実験

8.1 実験概要

実験 1 で作成したモデルの振る舞いの人間らしさを評価するための主観評価実験を行なった。モデル A~E (5 つ)

表 6 後から生物学的を導入した場合のプレイ回数 900~1000 回の平均スコア

Table 6 Average score of 900-1000 plays when biological constraints are introduced later

| モデル | A | B | C | D | E |
|---------|-------|-------|-------|-------|-------|
| スコア (%) | 72.33 | 71.46 | 75.60 | 79.90 | 84.91 |

のゲームプレイ映像と、そこへ人間の音楽ゲーム初心者と上級者のゲームプレイ映像を加えた計 7 種類の映像を 12 名の被験者に見せた。それぞれの映像は、「モデル A : 揺らぎ」は「映像 A : 揺らぎ」のように、対応するアルファベットの番号をつけ、人間の初級者を「映像 F : 初級者」、人間の上級者を「映像 G : 上級者」とした。映像は被験者ごとに 1 種類ずつ、ランダムな順番で見せ、人間の映像もあることは伝えずに見せた。その後、以下の内容について回答してもらった。

- (1) 人間らしさの絶対評価 (5 段階評価)
- (2) 自由記述 (評価の理由、どのような振る舞いが人間らしかったか、または人間らしくなかったか)

8.2 実験結果

各映像の振る舞いの人間らしさ平均評価値と標準偏差を表 7 に示す。平均評価値は 5 段階評価の平均値で、数値が大きいほど人間らしいと評価されたものである。「映像 G : 上級者」は 5 段階評価で最も人間らしいと評価された。また、「映像 F : 初級者」も 5 段階評価で 3 番目に人間らしいと評価され、多くの被験者は実際の人間のプレイを見破っていた。生物学的制約を導入したモデルでは、「映像 C : 行動の遅れ」が 2 番目に高い評価値で、人間の初級者よりも人間らしいと評価された。しかし、それ以外は人間プレイよりも人間らしいと評価されなかった。また、「映像 E : 無し」は 4 番目の評価値となったが、人間らしさの評価が分かれ、唯一評価値 1 と 5 の両方の評価を得て、標準偏差も 7 種類の中で最も大きくなった。

自由記述では、人間らしかった振る舞いとして以下の意見が挙げられた。

- キーを押すタイミングが時々ずれる
- 密度が高い (難しい) 局面でリズムが崩れる
- ゲームが上手い (下手) な人のような動き
- 近くのキーを間違えて押す
- 大きなミスはしないが全体でずれている

また、人間らしくない振る舞いとして以下の意見が挙げられた。

- 上手いところと下手なところの差が大きい
- 全く違うところを叩く (空打ちする)
- 同じキーの連打
- 最高判定が連続する
- 露骨すぎるミス

表 7 人間らしさ平均評価値と標準偏差

Table 7 Humanity average evaluation value and standard deviation

| 映像 | A | B | C | D | E | F | G |
|-------|------|------|------|------|------|------|------|
| 平均評価値 | 2.17 | 2.08 | 3.83 | 3.08 | 3.33 | 3.75 | 4.33 |
| 標準偏差 | 0.80 | 0.86 | 1.14 | 0.76 | 1.18 | 1.09 | 0.75 |

この中でも、「キーを押すタイミングが時々ずれる」「密度が高い(難しい)局面でリズムが崩れる」「上手いところと下手なところの差が大きい」「上手いところと下手なところの差が大きい」は多くの被験者から同じ意見を得た。

9. 考察

「A: 揺らぎ」や「B: 認識の遅れ」を導入したモデルは、ミスは増加したが、ミス際に他のキーを必要以上に空打ちをする振る舞いが人間らしいと評価されなかった。また、「B: 認識の遅れ」では、最高判定とミスの両方が多かったことから、上手い場面と下手な場面差が大きく、局面ごとの振る舞いが極端すぎたことが人間らしく評価されなかった原因であると考えられる。「C: 行動の遅れ」「D: 緊張」は判定 200 の比率が増加し、判定 0(ミス)にはならない程度のリズムのズレが、人間の中級者から上級者程度の人間プレイヤーに近い振る舞いだったと考えられる。「E: 無し」では人間らしさの評価が大きく分かれた。これは、極端に制度の高いプレイを、熟達した上級者の完璧なプレイと捉えた被験者は人間らしいと評価した一方で、機械的な動きと捉えた被験者は人間らしくないと評価したためだと考えられる。そのため、音楽ゲームのようなスコアを追求するようなゲーム性では、完璧なプレイを人間らしいと感じるユーザは一定数存在するが、そう感じないユーザも存在するため人間らしさの表現には有効な手段ではないと考える。「F: 初級者」ではミスは多かったが、空打ちなどが少なかったため人間らしい評価を得たと考えられる。「G: 上級者」は高スコアながら人間らしいと評価された。このことから、最高判定が極端に多くなく、完璧すぎない上手いプレイは人間らしいと評価される可能性が高いことがわかった。これらのことから、極端な失敗を増加させることよりも、微妙なズレなどを再現することが音楽ゲームで人間らしさを表現する上で重要だと考える。

10. まとめと今後の展望

本研究では、音楽ゲームを題材に人間らしい振る舞いを行うゲームプレイヤー AI の作成を行った。生物学的制約の導入により失敗の行動を再現し、一部では人間プレイヤーと同等の人間らしい振る舞いを獲得することができた。しかし、生物学的制約の種類によっては人間らしいとは評価されず、制約の与え方や処理方法について検討し改善する必要がある。また、現在の環境では AI に等速でゲームをプレ

イさせるため、学習に時間を要する点が課題であり、学習を実行する手段についても検討する必要があると考える。

今後は、ゲームの難易度を上げての学習や別の生物学的制約の導入を検討し、より人間らしい振る舞いの獲得を目指す。また、今回は音楽ゲームの画面情報のみを用いたため、より人間のプレイしている際の状態に近づけるには、音を聞きリズムを取ることができるようにする必要があると考える。そのため、音の情報を入力に導入することを検討する。

謝辞 本研究は JSPS 科研費 JP17K00514 の助成を受けたものです。AI の作成に関して、dsanno 様が公開されている [12] コード群に多く助けられました。感謝いたします。実験に協力していただいた被験者の皆様に感謝いたします。

参考文献

- [1] David Silver et al.: Mastering the game of Go with deep neural networks and tree search, Nature529, pp. 484-489, 2016.
- [2] Steven Kapturowski et al.: Recurrent Experience Replay in Distributed Reinforcement Learning, ICLR2019.
- [3] 佐藤直之ら: Influence Map を用いた経路探索による人間らしい弾避けのシューティングゲーム AI プレイヤ, ゲームプログラミングワークショップ 2016 論文集, pp. 57-64, 2016.
- [4] 仲道隆史・伊藤 毅志: 人を楽しませる接待将棋システム, 2014 年度人工知能学会全国大会, 2014.
- [5] 藤井 叙人ら: 生物学的制約の導入によるビデオゲームエージェントの「人間らしい」振る舞いの自動獲得, 情報処理学会論文誌 Vol 55, No. 7, pp. 1655-1664, 2014.
- [6] 藤井 叙人: 人間らしい振る舞いを自動獲得するゲーム AI に関する研究, 関西学院大学博士論文, 2016.
- [7] 冬木 一輝・伊與田 光宏: 音楽ゲームにおける習熟度を想定したプレイ結果のシミュレーション, 2018 年度電子情報通信学会東京支部学生会研究発表会, pp. 121, 2019.
- [8] HAUSKNECHT, Matthew; STONE, Peter. Deep recurrent q-learning for partially observable mdps, arXiv preprint arXiv:1507.06527, 2015.
- [9] peppy: osu!, <https://osu.ppy.sh/home>, accessed 2021-01-29
- [10] MNIH, Volodymyr, et al.: Playing atari with deep reinforcement learning, arXiv preprint arXiv:1312.5602, 2013.
- [11] asweigart: PyAutoGUI, <https://pyautogui.readthedocs.io/en/latest/>, accessed 2021-01-09
- [12] dsanno, Deep Q-Network Implementation using chainer, <https://github.com/dsanno/chainer-dqn>, accessed 2021-01-20