

予測と深層学習による特定プレイヤーの模倣手法

青木雄亮^{1,a)} 田村仁²

1. はじめに

コンピュータゲームにおいて、人間の対戦相手としての自動プレイヤーはゲーム AI の歴史の最初期から存在し、以降、人工知能分野の研究テーマとして継続的に行われている。その目的の 1 つとして、人間にとって適正なレベルの AI を開発するというものがある。

また、個性的な AI を生み出す目的で、ある個人の個性を学習によって得るという研究も行われている[1]。

人工知能を用いて人間の行動を模倣することを目指した研究は様々な分野に見られる。本研究では、人工知能の研究を行う非営利企業 OpenAI から提供されている強化学習用プラットフォーム OpenAI Gym[4]を用いて、ゲーム AI の学習を行う。具体的には、個人のプレイ特徴を模倣するゲーム AI を、ゲームの局面の予測を組み合わせで作成することを目指す。このゲーム AI の研究によって、近年注目されている e スポーツ等に活用できることが考えられる。

2. 関連研究

「スーパーマリオブラザーズ」シリーズを対象に遺伝的アルゴリズムを用いて特定プレイヤーの模倣 AI を作成する研究を行った例がある[1]。また、Multiplayer Online Battle Arena(MOBA)というジャンルの 1 つである、League Of Legends を対象にプレイヤーの行動の模倣を基に AI キャラクターの行動ルールを生成した研究がある[5]。この研究では Self-Organizing Incremental Neural Network(SOINN)という手法を使用している。この 2 つのようにゲームにおいて模倣をテーマにした研究は存在するが、未来の局面を予測することに焦点を当てたものはない。そのため、本研究では未来の予測に焦点を当て、模倣精度の向上を図る。

3. 提案手法

本研究では、人間である特定プレイヤーのゲームプレイを模倣する AI の作成を主目的としているが、その手法として、前述した OpenAI Gym を使用し、対象のゲームとして、「ソニック・ザ・ヘッジホッグ」を採用した。実装には後述の参考文

献にある書籍[3]を参考にした。このゲームは、操作キャラクターの動きに応じて敵の動きがある程度変化する。これにより、ゲーム局面に一定の変化が発生し、一種のランダム性を有する。これは同じ横スクロールアクションゲームである「スーパーマリオブラザーズ」やスクロール型のシューティングゲームのように動きが固定されているものとは性質が異なる。よって、敵キャラの動きに対応したプレイが求められる。この時、その瞬間の画面情報だけでのプレイも可能だが未来の局面の予測をすることが出来ればプレイを有利に進めることが出来る。

また、人間の模倣をするためには、人間が明らかに未来の予測をしたうえでゲームをプレイしていることを考慮しなくてはならない。よって、人間の特徴を模倣するには未来の予測を AI にさせるという手法が有効だと考えられるため、本研究では模倣に予測を組み合わせることを考えた。

現時点で、未来の局面を予測したうえでの模倣は行われていない。未来予測を「ソニック・ザ・ヘッジホッグ」にて行った例は存在する[2]。しかし模倣の観点では行われていない。よって、模倣に未来予測を組み合わせることでプレイヤーの模倣精度向上を目指す。具体的には、模倣の手法として BC(BehaviorCloning)を使用し、未来予測の手法として GAN を用いることで数フレーム先の局面の画像を生成することを予定している。使用する GAN の手法として、Pix2Pix の使用を予定している。具体的には、ゲーム内で、現在のフレーム画像と数フレーム先の画像をペアで学習させ、数フレーム先の画像を生成させる。これらの手法を組み合わせで実験し、模倣した AI の性能評価を行う。予定としては、AI がプレイした結果、模倣した人間の獲得したスコア、クリアにかかった時間、獲得したアイテム等の、プレイヤーの特徴をどの程度の精度で類似しているかを大まかな評価軸とする。

また、現時点では GAN での画像生成を行う段階にはまだないが、BC を実装し学習を行った段階である。

4. 実験

ここでは「ソニック・ザ・ヘッジホッグ」を対象のゲームとして模倣学習させ、学習の手法は Behavior Cloning(BC)を使用した。まず人間のプレイのキー操作情報と画像をプレイデータとして保存し、それを用いて BC を用いて学習させた。この実験から、人間と、それを模倣した AI がどの程度類似したプレイを行うのかを確かめた。また、今回の実験では 100 プレイ分のプレイデータを用いて学習させ、評価軸は、人間と AI の獲得したスコアとした。

5. 実験結果

今回の実験によって 100 プレイ分の人間とそれを基に模倣学習した AI のそれぞれのプレイした結果のスコアをグラフに表すことで結果を得ることができた。

以下の図 1 に人間の 100 プレイ分のスコアのグラフを示す。

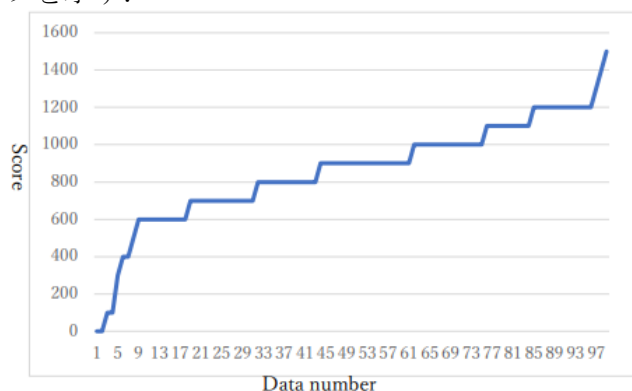


図 1 人間のプレイ結果

次に、100 プレイ分のプレイを模倣学習した AI のプレイのスコアのグラフを以下の図 2 に示す。

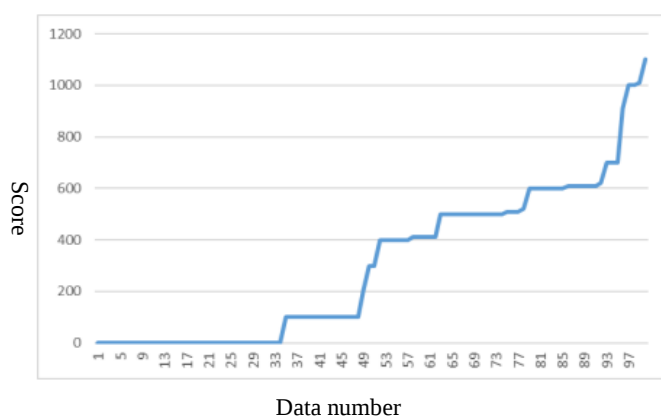


図 2 AI のプレイ結果

6. 考察

実験結果から、スコアのグラフは類似したものにならなかった。この原因として、人間のプレイデータ不足や、AI が獲得できる、人間の特徴に関しての情報が不足していたためだと考えられる。前者に関してはある程度まで数を増やすことは可能だが 1000 プレイや 10000 プレイといった桁のプレイデータを収集することは非現実的である。よって、AI が如何に、ゲームのスコア、クリアまでに費やした時間、獲得したアイテム等の人間のプレイ特徴を獲得できるかという点に焦点を当てるべきだと考えられる。また、BC によってある程度の模倣が可能であると分かったので、今後の研究の可能性を探ることが出来ると考えられる。

7. まとめ

実験結果から、人間とそれを基に学習した AI のスコアのグラフは類似度は十分ではないと考えた。今回の実験での手法では特定の個人の模倣の段階までは進まなかったが、模倣 AI を作成する足掛かりにはなった。

8. 今後の予定

今回は模倣学習を行う段階まで研究は進んだが、数値的に人間の特徴にそれほど類似しない結果になったため、より類似性を高めるための手法の研究を行う。具体的には、人間のゲームプレイをキャプチャした画像から、GAN を使用することで数フレーム先の画像を生成し、AI にゲーム画面の先読みをさせることで個人の特徴の取得が可能であると考えている。今後はそのような方針で研究を進めていく。

参考文献

- 1) 池田 裕太郎 池田 心, アクションゲームにおける特定のプレイヤーの特徴を模倣する AI プレイヤーの作成, 研究報告ゲーム情報学(GI),2020-GI-43,9,pp1-8 2020
- 2) World Models applied to Sonic <https://dylandjian.github.io/world-models/> 2021年7月26日アクセス
- 3) 布留川 英一 OpenAI Gym / Baselines 深層学習・強化学習 人工知能プログラミング実践入門 株式会社ボーナデジタル 2020年2月25日発行
- 4) OpenAI Gym <https://gym.openai.com/> 2021年7月26日アクセス
- 5) 張輝陽 星野准一 プレイヤ行動の模倣に基づく AI キャラクター行動ルールの自動生成, 研究報告ゲーム情報学 (GI) ,2014-GI-31,5,pp1-4 2014