

縮小盤オセロにおける AlphaZero と MuZero の比較分析

1240321 佐々木 遥矢 【ゲーム情報学研究室】

1 はじめに

ゲームの AI プレイヤの作成を行う際に用いられる強化学習アルゴリズムは多くの種類がある。有名なアルゴリズムには DeepMind 社が開発した AlphaZero があり、チェス・囲碁・将棋において高い性能を発揮している。このアルゴリズムは学習を行う際にある盤面から一手進めた際の次の盤面を再現するために盤面の状態遷移ルールを知っておく必要がある。そのため状態遷移ルールが完全情報ゲームと比較して複雑になる傾向があるビデオゲームなどには適用しづらい。発展形の MuZero[1] は状態遷移ルール自体を学習するため、同じアルゴリズムで様々なゲームに適用することができる。しかし、状態遷移ルールを用意することができるゲームにおいて性能が AlphaZero に大きく劣っている場合はそのゲームで採用する必要がないと考えられる。

ゲームの性質などに起因するアルゴリズムの性能差について知見を得るために、今回は各アルゴリズムで作成した AI プレイヤの性能比較を行う。

2 関連研究

2.1 MuZero

MuZero は、DeepMind 社が 2020 年に発表した AlphaZero の発展形であり、盤面の状態遷移ルール自体を学習することでマルコフ決定過程の環境に対応した強化学習アルゴリズムである。マルコフ決定過程とは現在の状態から行動を起こした際に次の状態が決まることである。MuZero では学習した環境のモデルとモンテカルロ木探索を組み合わせることによって Atari ゲームで当時最先端の評価を達成している。また、チェス・囲碁・将棋においても AlphaZero の性能に匹敵する評価を達成している。

2.2 Elo レーティングによる比較

Ti-Rong Wu らの研究 [2] では Elo レーティングによる AlphaZero と MuZero の比較を行った際に 9×9 囲碁では MuZero が AlphaZero よりも良く、8×8 オセロでは AlphaZero が良いことが示されている。これは 8×8 オセロのようなターンごとに駒が頻繁に反転するような盤面の変化量が多いゲームでは状態遷移を学習することが困難であるからだと述べられている。オセロは状態遷移ルールを用意することが比較的容易なゲームであるため、ルールを学習させるよりも遷移を人間が用意したほうが正確で性能が向上すると考えられる。

3 提案手法

前述の方法での比較では勝敗による比較のため、対戦相手の強さによる誤差が生じる。また、局面ごとの指し

手の正確さを評価できない。そこで本研究では全局面で勝敗がわかっている完全解析されたゲームを題材とする。ゲームの完全解析結果と学習した AI プレイヤの指し手の比較を行うことによって局面ごとに正しい判断ができていないかを評価する。本研究では完全解析されたゲームとして縮小盤オセロを用いて AI プレイヤの作成を行う。

4 実験

本研究では分散強化学習フレームワークである SimpleDistributedRL¹ を使用して AlphaZero と MuZero の学習を行う。対象とするゲームは 4×4 と 6×6 のオセロである。二つのゲームは完全解析されており、サイズの違いによる差も比較できるためこの二つのゲームをそれぞれのアルゴリズムで学習したプレイヤを作成する。作成したプレイヤが任意の盤面に対してどの指し手を選択するかを集計し、完全解析結果と比較して平均悪手率を求める。また通常の 8×8 サイズのオセロにおいても AI プレイヤを作成し弱解決結果 [3] との比較を行う予定である。

5 まとめ

本研究では、完全解析されている縮小盤オセロを用いて学習した AI プレイヤの指し手を完全解析結果と比較することによって平均悪手率を求める。今後は、実験を受けてゲームの性質とアルゴリズムの性能差について調査・考察する。

参考文献

- [1] Schrittwieser, Julian and Antonoglou, Ioannis and Hubert, Thomas and Simonyan, Karen and Sifre, Laurent and Schmitt, Simon and Guez, Arthur and Lockhart, Edward and Hassabis, Demis and Graepel, Thore and Lillicrap, Timothy and Silver, David, “Mastering Atari, Go, chess and shogi by planning with a learned model“, Nature, 2020, 7839, p.604–609
- [2] Ti-Rong Wu, Hung Guei, Po-Wei Huang, Pei-Chiun Peng, Ting Han Wei, Chung-Chin Shih, Yun-Jui Tsai, “MiniZero: Comparative Analysis of AlphaZero and MuZero on Go, Othello, and Atari Games“, <https://arxiv.org/pdf/2310.11305.pdf>
- [3] Hiroki Takizawa, “Othello is Solved“, <https://arxiv.org/pdf/2310.19387.pdf>

¹https://github.com/pocokhc/simple_distributed_rl