

AI による非対称型対戦ゲームの提案 ～将棋への適用検討～

岡村 涼平[†]尾崎 敦夫[‡]大阪工業大学 情報科学部 コンピュータ科学科[†] (現, 情報知能学科[‡])

1. はじめに

対称型対戦ゲームとは一対一の対戦ゲームにおいて開始時点での戦力差が無いものを指し、逆に、非対称型対戦ゲームは戦力差を有するものを指す。本研究は、対称型対戦ゲームの開始状態を調整し、非対称型対戦ゲームに落とし込むことで対称型対戦ゲームを学習した結果とは異なる戦略を得ることを目的としている。

本稿では、非対称型対戦ゲームの一つである、将棋の駒落ち戦を具体例に検討を行う。

2. 提案手法

2.1 基本概念

一般的な将棋 AI のアルゴリズムでは学習を行う自分と、学習を行うときに自分と対局する相手が全く同じ状態であるため、自己学習という形で学習する[1]。しかし、手合割と呼ばれる駒落ちのハンディキャップを含む対局では自分と相手が同じ状態ではない。両者のパワーバランスは等しくないため、同等に学習を行うと手合割を持つ先手の評価が行えず良好な結果が得られないことから別々に学習を行う必要がある。そこで、勝率に合わせて学習速度を合わせるアルゴリズムが有効となる。これにより、双方の勝率を維持したままより深い学習が実現できると考える。このように、影響力の高い駒や低い駒を予め盤上から除去することで、従来の将棋 AI とは異なる学習結果が得られ、新たな定跡を発見することが期待できる。

2.2 アルゴリズム

通常手合割を含む対局では駒の数が少ない方を先手としており、本稿でも手合割を持つ側を先手とする。図 1 は提案手法でのおおまかな処理フローを示したものである。まず任意の回数対戦し、勝率が低いサイドを学習側とする。同時に、その対戦を通して、初期の学習データを生成する。次に、この学習データおよび AI により、再度任意の回数、対戦して、その勝率を計算すると共に、学習データと AI を更新する。勝率が 50%未満であれば対戦するところからやり直し、50%～60%の間であれば AI と学習データを更新してからやり直す。60%以上あればサイドを変更し、AI と学習データを更新した後に対戦を再開する。また、サイドを変更した回数が所定回数を超えた時に終了する。なお、学習データは深層強化学習により得られたニューラルネットワークの係数などの数値データであり、以後このデータを利用することで対人やデータの解析が可能となる。

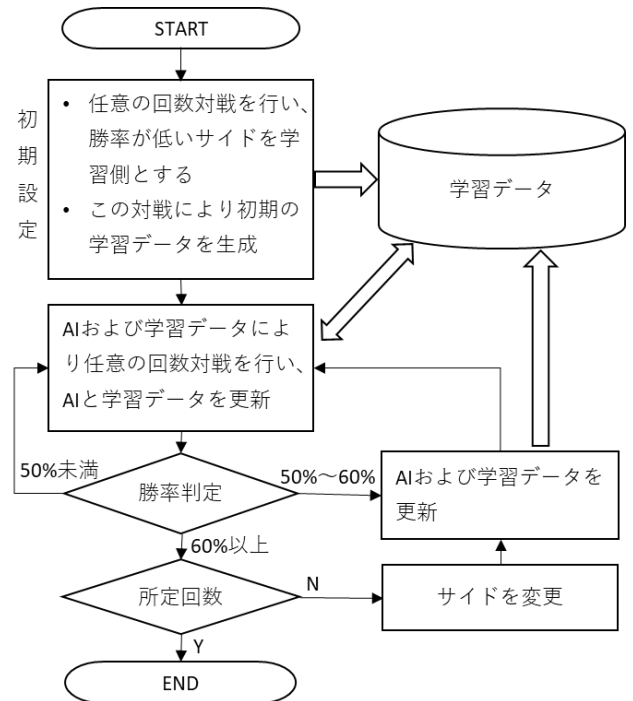


図 1 提案手法(手合割向け将棋 AI の処理フロー)

3. 今後の課題

AI は定跡という形を盤面、棋譜から読み取ることができないため、その定跡がどのような局面で使えるかについての評価方法が今後の課題である。これは、将棋以外への適用についても同様の課題を持つ。また、最終的には手合割を持つ先手が勝てなくなるとされる。このため、できるだけ深く、また勝率を保った状態で学習を止めるアルゴリズムの検討も今後の課題となる。

4. まとめ

対戦ゲームでは AI 適用により、新たな定跡の発見が期待されている[2]。本稿では、将棋を対象に、その発見に有望な手合割向け将棋 AI を提案した。将棋 AI の発展は人間が指す将棋の発展にもつながるため、今後も方式検討やその評価を実施していく予定である。

参考文献

- [1] 山岡忠夫, “将棋 AI で学ぶディープラーニング,” 山口正樹(編), (社)マイナビ出版, 東京, 2018.
 [2] Ritsuko Kawai, “AI が変えた現代将棋の常識と定跡: 北尾まどか氏×森川幸人氏 対談,” モリカトロン AI ラボ, https://morikatron.ai/2019/06/madoka_kitao_interview/, 参照 Jan.5,2019.