

接待 AI に向けて、多様な強さでプレイ傾向が一貫するプレイヤー生成の研究

1210377 三谷 愛 【ゲーム情報学研究室】

1 はじめに

近年では、ゲーム情報学の研究はゲーム AI の強さのみならず一般的な人間プレイヤーが楽しめるよう接待 AI の分野にも進んでいる。接待 AI に必要な要素として対等な力量であり不自然さが無いことが挙げられるが、動的に力量を調整しつつ一貫性がある戦略を保った接待 AI の作成は難しいのが現状である。

対等な力量を持つ接待 AI に関して、杉本ら是对戦中に数ステップごとの勝率を予測し対戦させる AI を変更する手法を提案している [1]。また、上田らは特徴パラメータと勝率による適応度計算を用いた遺伝的アルゴリズム (GA) により同程度の力量でありながら多様な戦略を持つ AI 群の作成を行っており今回扱う問題とは真逆の接待 AI となっている [2]。杉本らの手法と、上田らの適応度計算式を上手く落とし込めば、盤面ごとの状況により動的に強さを切り替えながら戦略に一貫性を持つ不自然さのない接待 AI の作成が期待できる。

本研究では、オセロを題材に前段階である GA を用いたパラメータ最適化による異なる強さを持つ AI プレイヤーの生成法が有効であるかを検討をする。また、AI プレイヤーのプレイ傾向が保持されたままかの調査も行う。

2 関連研究

2.1 遺伝的アルゴリズム (GA)

GA は、生物の進化の過程を模倣した解の探索手法である。ある世代において解の候補となる遺伝子 (パラメータ) を持つ個体群の中から確率により突然変異や淘汰、交叉するものが選ばれ行われる。その中から、適応度が高いものが次世代の個体として残りやすく新たな次世代の個体群を形成する。N 世代に達した時点での適応度が最も高い個体の遺伝子が最適解となる。複数の異なるパラメータを持つ個体の動的な調整を上田らの研究で行われている [1]。

2.2 平均悪手

教師プレイヤーの指し手を最善とし、生徒プレイヤーの指し手がどれだけ悪手であるのか評価をする。平均悪手は、評価で得た値の合計を局面数で割ったものであり強さの指標として利用できる。

3 提案手法

プレイ傾向に特徴を持ったオセロ AI の状態評価関数のパラメータを GA により調整を行う。GA 中の適応度の計算は、異なる力量、似たプレイ傾向のものが高くなるように設定をする。前者は、Zebra との平均悪手の評価値を利用し、後者は最適化前のオセロ AI の状態評価に使用される重みパラメータとの差を用いる。

4 実験手順

オセロの代表的な $\alpha\beta$ 法による状態評価関数に角に置きやすい、辺に置きやすいプレイ傾向を持つような新たな重みパラメータを追加する。プレイ傾向の重みを持たないオセロ AI、重みを持つ AI を MCTS プレイヤーとの対戦をさせ、勝率、角に置いた回数、辺に置いた回数をカウントしてプレイ傾向が現れているかの調査をする。

GA によるパラメータ最適化後に、再度 MCTS プレイヤーと対戦をして GA 前より勝率及びプレイ傾向に変化があるか調査をする。

5 実験結果

表 1、表 2 は特徴を角を重視、辺を重視と設定したときの勝率、及び駒の置いた回数を最適化前と後について示したものである。辺を重視するものは特徴が多少欠けているが勝率が 8% 向上した。逆に、角を重視するものは特徴を維持しながら勝率が 8% 下がり弱くなっている。

表 1 GA 最適化前のプレイ傾向

特徴 AI	角置き	辺置き	勝率
特徴なし	59 回	620 回	10%
角を重視	196 回	1163 回	26%
辺を重視	70 回	1492 回	1%

表 2 GA 最適化後のプレイ傾向

特徴 AI	角置き	辺置き	勝率
角を重視	204 回	1172 回	18%
辺を重視	149 回	1205 回	9%

6 まとめ

本研究では、GA を用いたプレイ傾向に一貫性を持ちながら強さが異なる AI プレイヤーの作成を行った。今後はこの作成法が有効であるかの検討を行う。

参考文献

- [1] 杉本 直樹, 鶴岡 慶雅, “戦略の動的推定による 2 人対戦ゲーム接待 AI の提案”, ゲームプログラミングワークショップ 2018 論文集, p114-119, 2018.
- [2] 上田 陽平, “GA を用いたプレイヤーのレベルに適應する 多様なオセロ AI の開発”, JAIST 学術研究成果リポジトリ, 2012.