

平均悪手を適応度とした遺伝的アルゴリズムによる探索パラメータの調整

1200375 山崎 蓮太 【ゲーム情報学研究室】

1 はじめに

強いゲーム AI の実現には最新のゲーム木探索の実装が重要な要素になっている。ゲーム木探索の効果を最大化するためには、探索内部のパラメータを適切に調整する必要がある。調整に遺伝的アルゴリズム (GA) を用いる際には、目的関数 (適応度) に強さと相関があるものを設定することが望ましい。単純な方法は対戦を行い、勝率で強さを測る手法であるが、時間がかかりすぎるために実用的ではない。一方で、従来手法は早いが強さとの関連性が薄いものが使われてきた [1]。適応度には強さとの関連が強く、対戦よりも時間がかからないものを利用したい。そこで本調査では、平均悪手を適応度に設定して GA によるパラメータ調整を行う。平均悪手は山下らによりレーティングを推定できることが明らかになっており、対戦よりも時間がかからない [2]。この平均悪手が探索パラメータ調整に有効であるかを調査する。

2 関連研究

2.1 遺伝的アルゴリズム

遺伝的アルゴリズムは、生物の進化の過程を模したアルゴリズムであり、最適化問題によく用いられる。

ある適応度を最大化するように調整する際には、その適応度が適応度関数となる。その適応度が高いほど生き残りやすく、次の世代に子孫を残しやすくなる。次世代の個体は親のペアの遺伝子同士を交叉させ、確率に応じて突然変異を起こし、生成される。交叉や突然変異、遺伝子のデータ構造については様々な手法がある。

世代ごとに N 個体あるとし、 M 世代まで繰り返すとした場合、具体的な手順は以下の通りになる。

1. 個体を N 個生成
2. 個体ごとに適応度を計算
3. 適応度が十分に大きいか、 M 世代に達しているならば終了。終了しない場合、以下の手順で子孫を作成
 - (a) 現在の世代から、親となるペアを適応度に応じて選択
 - (b) ペアの遺伝子を確率 P_c で交叉、確率 P_m で突然変異させる
4. 古い世代を新しい世代で置き換えて 2. へ戻る

2.2 平均悪手

平均悪手とは、プレイヤーが悪手をどれだけ指しているかの指標である。悪手とは、「教師に設定したプレイヤーと別な指し手」かつ「盤面の評価値が下がった場合」の指し手のことである。この悪手と教師の最善手の評価

値の差を足し合わせて、悪手を調べた局面数で割った値が平均悪手になる。この平均悪手で、山下らがレーティングを推定できることを明らかにしている [2]。

3 提案手法

1. で述べたように、本研究では提案手法として、GA による探索パラメータ調整を行う際の適応度を平均悪手に設定することを提案する。

平均悪手を提案した理由は、2.2 で述べたように、強さとの相関があり、対戦よりもレーティングの推測に時間がかからないためである。

4 調査と実験

調査と実験は次の手順で行う。まず、平均悪手が適応度として有用であるかを調べるために、平均悪手を適応度とした探索パラメータの調整を行う。そして調整されたプレイヤーに対して、対戦や適応度、探索パラメータの計測、教師との一致率の比較を行うことで性能評価を行う。

4.1 実験設定

GA については竹内らの実験に従い、10 個体 50 世代として、変異率 0.05、交叉率 0.75 として調整を行った [1]。また GA には Python の遺伝的アルゴリズムライブラリである DEAP を利用する。実験用のオセロプログラムは、探索パラメータの調整に用いるテストプログラムとして、モンテカルロ木探索を実装したプレイヤーを利用した。

調整するパラメータの詳細は卒業論文に記載する。

調整に用いる適応度に設定する平均悪手については、教師としてオセロプログラムである Zebra を採用した。局面数は 500 局面とした。

5 まとめ

本研究では、平均悪手を適応度とした遺伝的アルゴリズムによる探索パラメータの調整を行った。卒業論文では、実験結果から平均悪手が有効であるかを考察する。

参考文献

- [1] 竹内聖悟, 金子知適. 探索パラメータの調整に適した目的関数の調査: モンテカルロ木探索将棋の探索パラメータの調整. ゲームプログラミングワークショップ 2012 論文集, pp. 84-91, 2012.
- [2] 山下宏. 将棋名人のレーティングと棋譜分析. ゲームプログラミングワークショップ 2014 論文集, pp. 9-16, 2014.