

進化的計算を用いたモンテカルロガイスターのプレイアウトの改善

1245128 栢川純平 【ゲーム情報学研究室】

Improving Playout of Monte Carlo Geister Using Evolutionary Computation

1245128 Jumpei Tochikawa 【Game Informatics Lab.】

1 はじめに

ガイスターは、相手の駒の色がわからない二人不完全情報ゲームである。ガイスターにはモンテカルロベースのプレイヤーがあり、プレイアウト(シミュレーション)でランダム方策を行っている。プレイアウト中はランダム方策よりも人間の知識を用いたルールベース方策の方が強いとされている。しかし、知識を入れるには人間がゲームに習熟している必要があり、ゲームの習熟度によらず作成できると望ましい。Ms.Pac-Man では遺伝的プログラミング (Genetic Programming, 以降では GP) を用いてルールベース方策を作成する研究 [1] があり、平均スコアが向上したという結果が得られている。不完全情報ゲームでのプレイアウトでは、ルート局面で局面の完全情報化を行っているため、相手の状態が正しいとは限らず、適切なルールが不明である。そのため、適切なルールの特徴を調べる必要がある。

本研究では、モンテカルロガイスターで進化的計算でルールベースのプレイアウト方策を作成し性能の改善を目指し、性能の違いによるルールの特徴について調査する。また、GP で用いる適応度についても調査を行う。

2 提案内容

本研究では、ガイスターにおいて進化的計算を用いてプレイアウト中の方策を作成することでモンテカルロベースプレイヤーの性能向上を目指す。プレイアウト方策は、事前にガイスターのルールを基に関数を定義し、それを用いて GP によって作成する。

関数の分類は、「必ず使用する関数」、「色情報を区別しない関数」、「色情報を区別する関数」の3つである。不完全情報ゲームはプレイアウト時に現在の局面を可能な状態から仮定し完全情報状態にする。そのため、関数で色情報の区別で影響すると考えた。

また、GP では一致率と平均悪手を適応度に使用し、教師データは完全情報 (P) と不完全情報 (IP) の場合を使用する。作成した方策とランダム方策のモンテカルロ木探索プレイヤーで対戦により性能を評価する。

ルールベース方策はプレイアウトのランダム性を失うことにより弱くなるため、方策への確率の導入などがされている。よって、作成した方策を確率的に指手選択さ

せるために、遺伝的アルゴリズム (Genetic Algorithm, 以降では, GA) を用いてパラメータ調整を行う。GA 内で使用する適応度は GP で使用した不完全情報の一致率と平均悪手とする。対戦は、パラメータを調整する前のものとの対戦を行う。

2.1 GP の適応度

本研究は、強くすることが目的であるため、適応度は強さと関連があるものである必要がある。また、適応度は計算回数が多いため、計算コストが低い方が良い。本研究では、強いプレイヤーの指手との一致率と平均悪手の2種類を適応度とする。チェスで GA により探索パラメータを調整した研究 [2] では、強い人間の指手との一致率を適応度として採用した。一致率は、この研究で成功していることから使用した。

平均悪手は、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。山下らの研究 [3] では、平均悪手でレーティング推定ができることを明らかにした。よって、平均悪手が強さと関連していることから適応度とする。

3 実験

本実験では、GP によりプレイアウト方策を作成する。GP は色情報なし (N)・色情報あり (C) とすべて使用する (A) で行う。N は「必ず使用する関数 + 色情報を区別しない関数」、C は「必ず使用する関数 + 色情報を区別する関数」、A は「すべての関数を使用」とする。

また、作成した方策の指手選択に確率を導入し GA によってパラメータの調整を行う。方策は勝率が高いものを使用した。

GP, GA は共通して個体数 100, 世代数 50 で行い、変異率 0.2, 交叉率 0.5 とし, 5 回行った。教師は, GAT2020 のガイスター AI 大会で優勝した Naotti-2020 を用いた。

対戦実験は, 1000 試合行い, 勝率を調べた。対戦には, 30, 50 世代の方策を用いる。初期局面は重複なしで 500 局面用意し, 先手後手入れ替えで行った。対戦相手は, GP ではランダム方策のモンテカルロ木探索プレイヤー, GA では調整する前の方策とした。

実験で使用するプレイヤーのプレイアウト回数は 10,000 回とする。また, ルート局面で実現可能な相手駒色の配置をランダムで決定化し, 完全情報化を行っている。こ

れは、1 プレイアウトごとに行っている。

4 実験結果

図1はGPの一致率の推移を表している。グラフは、横軸が世代数、縦軸が一致率を表している。グラフの青が色情報を使わない場合、緑が色情報を使った場合、赤が全ての関数を使用した場合、実線が完全情報、破線が不完全情報のときのグラフとなっている。平均悪手の場合でも、世代が進むにつれ適応度が上昇していた。

表1はランダム方策との対戦結果の一部である。「一致率 | IP | A」は、「すべての関数を使用」した場合で一致率を適応度とし教師データが不完全情報での結果を表している。平均・最大勝率は、5回の試行での平均・最大勝率である。対戦の結果、一致率を適応度とした場合は全てで平均勝率は50%を越え、最大勝率は82.5%となるものもあった。平均勝率は、色情報ありが最も高くなり、完全情報の時は70%を越えた。一方、平均悪手を適応度とした場合は不完全情報での色情報なしが平均勝率50.3%で最も高かった。最大勝率の中には有意的に高くなるものもあったが、一致率が適応度の場合に比べて勝率が全体的に低くなっていた。また、世代数による勝率の変化は見られなかった。

GAを行ったところ、世代が進むごとに適応度は上昇していたが、GPの時ほど大きく変化しなかった。対戦を行った結果、一致率を適応度とした場合の30世代で平均勝率50.84%となり、ほとんどの個体で性能の向上は見られなかった。平均悪手を適応度とした場合は平均勝率が約20%と大きく性能が下がった。

5 考察とまとめ

本研究では、GPを用いてガイスターにおけるルールベースのプレイアウト方策を作成し、性能の改善を目指した。実験の結果から一致率を適応度とした場合で性能の良いものが作成できていることが確認できた。

作成された方策を調べたところ、「脱出口に向かう指手(赤駒のみを除く)」が行われることが多い方策のプ

表1 方策プレイヤーの勝率(%) : 対ランダム方策

世代			30		50	
勝率			平均	最大	平均	最大
一致率	IP	A	61.92	82.5	61.32	79.45
一致率	P	C	73.14	79.6	72.21	80.45
平均悪手	IP	N	50.32	57	50.3	55
平均悪手	IP	C	42.38	50.2	42.6	52.2

レイヤーが勝率が高くなっていた。「相手駒を取る」、「自駒を自陣脱出口に向かわせる」を使った方策のプレイヤーは、強いものもあったが弱いものもあった。「相手駒を取る」を含む方策は、相手駒色が正確とは限らないことがプレイアウト結果に影響し、弱くなったと考えられる。また、「相手青駒から逃げる」、「脱出口から遠ざかる」が中心の方策を使用したプレイヤーは勝率が低くなっていた。勝利条件に直接関係があるルールを含む方策ほどプレイヤーが強くなると考えられる。

本研究の実装上、指すことができない場合はランダムに指すことになっている。GAを行ったところ性能が向上しなかったことからGPの時点で十分ランダム性を持っていたと考えられる。また、局面の仮定時にもランダムに駒色の配置を行っており、この部分でもランダム性を持たせられていると考えられる。

適応度が強さに繋がったかを調べるために勝率と適応度との相関係数を調べた。その結果、不完全情報での一致率に正の相関があることがわかった。そのため、現段階では不完全情報での一致率を適応度として用いるの方が良いと考えられる。不完全情報での平均悪手においても弱い正の相関が見られたが、勝率が高くなかったことから適していないと言える。しかし、Naotti-2020の評価関数は粒度が荒かったため、平均悪手の計算に向けた評価関数を利用することで改善する可能性がある。

参考文献

- [1] Atif M Alhejali and Simon M Lucas. Using genetic programming to evolve heuristics for a monte carlo tree search ms pac-man agent. In *2013 IEEE Conference on Computational Intelligence in Games (CIG)*, pp. 1-8. IEEE, 2013.
- [2] Omid E David, H Jaap van den Herik, Moshe Koppel, and Nathan S Netanyahu. Genetic algorithms for evolving computer chess programs. *IEEE transactions on evolutionary computation*, Vol. 18, No. 5, pp. 779-789, 2013.
- [3] 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*, 第2014巻, pp. 9-16, oct 2014.

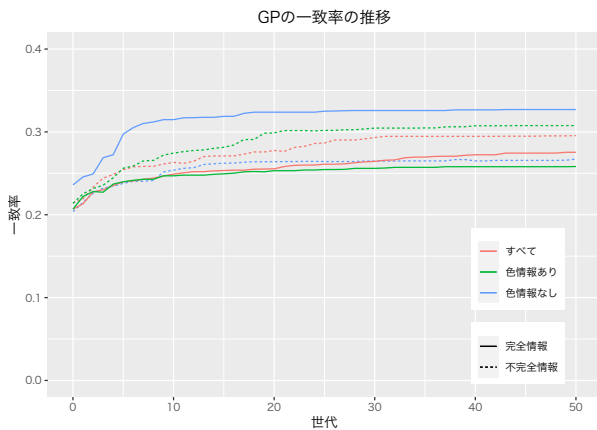


図1 一致率の推移