

ゲーム 2048 におけるモンテカルロ木探索の実装手法に関する研究

1240285 安藤 陸人 【高度プログラミング研究室】

1 はじめに

2048 ゲームは G.Cirulli が 2014 年に公開した不完全情報ゲームである。本稿では、ゲーム 2048 におけるモンテカルロ木探索において、ボルツマン分布を用いた過去の研究との比較として、UCT 探索を用いた研究となっている。

2 モンテカルロ木探索

モンテカルロ探索とは、一つのアクションに対して、選択枝を平均的に複数回実行しそれぞれの選択枝の勝率を計算することによって、勝率の高い選択枝を選ぶ手法である。モンテカルロ探索では一手先の選択枝しか考慮しないため、その先の選択枝に無駄が生じてしまう。モンテカルロ木探索では、選ばれた選択枝の子ノードを展開することによって無駄な手の先読みを減らしていく。

また、本研究で用いる UCT 探索は、モンテカルロ探索での選択枝を平均的に複数回実行することによって、勝率の低い選択枝を実行することに無駄が生じてしまう事象を改善した手法である。

以下が UCT 探索における式である。

$$UCT_a = \frac{w_a}{n_a} + c\sqrt{\frac{\ln N}{n_a}}$$

左項は勝率を表し、 w_a が勝利回数、 n_a が試行回数である。右項はペナルティを表し、 c が定数、 N が累計試行回数である。

累計試行回数 N に対して自然対数を取っているため、序盤の試行では優先的に回数の少ない手が選択され、終盤の試行では勝率に応じて選択される割合が大きくなるように調整されている。

3 実験方法

2048 ゲームでの一手に対するシミュレーション回数を 100 回と 200 回で 100 ゲーム行い、score の平均値の上昇幅を比較する。本研究では、2048 ゲームに勝敗の判定がないため、左項の勝率に対して評価関数を用いた評価値を扱っている。評価関数は、ニューラルネットワーク評価関数 CNN22 を借用する。UCT 探索において、左項の勝率の部分を盤面の評価値である q とする。

なお今回使用する評価関数である CNN22 では、評価値の値が 10^6 桁と膨大である為、本来の UCT 探索の式である勝率の値に近づけるために選択枝の評価値の最大値で割っている。また、もう一つの評価値調整方法として \log_2 で対数変換を行うといった手法も行う。

q	sim	c	score
q/max(q)	100	0.01	339234
		0.1	316948
		1	280769
q/max(q)	200	0.01	346633
		0.1	350726
		1	301428
$\log_2(q)$	100	0.01	336074
		0.1	318961
		1	304466
$\log_2(q)$	200	0.01	351791
		0.1	314377
		1	302130

表 1 sim=100,200 の結果

4 実験結果

c の値が小さくなればなるほどスコアの値が大きくなる傾向にある。シミュレーション回数が 100 回の段階では、評価値の調整方法に結果の大きな差はない。シミュレーション回数が 200 回では、 $q/\max(q)$ の $c = 0.01, 0.1$ と $\log_2(q)$ の $c = 0.01$ の結果が上昇幅が大きい。

5 まとめ

本稿では、2048 ゲームにおける UCT 探索において、最適な実装方法を研究した。完成度の高い評価関数では、試行回数による影響を受けずに、評価値のみを基準とした場合でも良い結果を得ることができる。しかし、一手当たりのシミュレーション回数を増やすことにより、試行回数の少ない選択枝をどれほど考慮するかによって、平均スコアの伸びが期待できそうである。今後の研究では、シミュレーション回数の増加による成長性が望まれる実装方法でシミュレーション回数 800 回でのスコアの比較を行う。

参考文献

- [1] K. Matsuzaki, “Developing value networks for game 2048 with reinforcement learning,” *Journal of Information Processing*, vol. 29, pp.336–346, 2021.
- [2] S. Watanabe and K. Matsuzaki, Enhancement of CNN-based 2048 Player with Monte-Carlo Tree Search, *presented at the 27th Int. Conf. Technol. Appl. Artif. Intell. (TAAI2022)*, Tainan, Taiwan, Dec. 1–3, 2022.