

ミニ 2048 における ニューラルネットワークプレイヤーの 学習方法に関する研究

1250304 金子 智弥 【高度プログラミング研究室】

1 はじめに

確率的一人ゲーム「2048」のコンピュータ・プレイヤーについては多くの研究がなされており、人間プレイヤーよりもずっと高いレベルで競われている。近年、さまざまなゲームでニューラルネットワークを用いたプレイヤーが大きな成功を収めているが「2048」の研究においてはこの限りではない。本研究では、ニューラルネットワークのパラメータ数を増やす以外の方法で性能向上が可能かを検討する。完全解析されているミニ 2048 [1] を対象とし、複数のニューラルネットワークモデルを作成する。それらのモデルをネットワークへの入力方法を変更する新手法を用いて学習を行い、既存手法との違いをパーフェクトプレイヤーを用いて評価する [2]。

2 ネットワーク入力

ニューラルネットワークプレイヤーの学習にあたり、ネットワークに入力する入力表現を変更することが有効であると考え、以下の3種類の入力表現を用いた。

通常入力 2048 で一般的に用いられる二値テンソル表現 (図 1)

相対入力 最大のタイルを index 0 に置き、各タイルを最大からの相対位置に配置したもの (図 2)

通常 + 相対入力 図 1 と図 2 を足し合わせたもの。ただし、相対入力の空きタイルは表現しない。

3 実験・実験結果

ネットワークモデルについては、先行研究 [3] で用いられていた deep と resnet を用いた。また、既存の学習改良手法である restart [4] と対称性の利用を組み込んだ学習を行った。ネットワークモデル 2 種 × 入力表現 3 種 × 学習方法 2 つの有無 (4 通り) で計 24 通り、各 24 時間の学習を行った。学習後に評価として行った 1000 プレイで平均得点の高かった上位 3 通りの結果を表 1 に示す。

4 まとめ

本稿では、ミニ 2048 においてニューラルネットワークプレイヤーの学習方法を検討した。その結果、「最大タイルからの相対的大きさ」を付け加えた学習はニューラルネットワークの強さ向上に寄与しなかった。また、restart の付与は学習を高速化できる利点があるが、学習時間を十分に取れば restart を付与しない学習と同等の平均得点を達成することが示された。

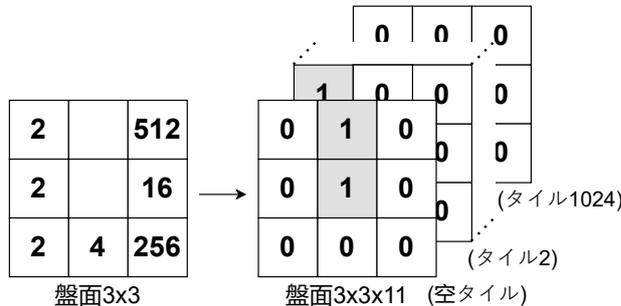


図 1: 入力: 通常入力

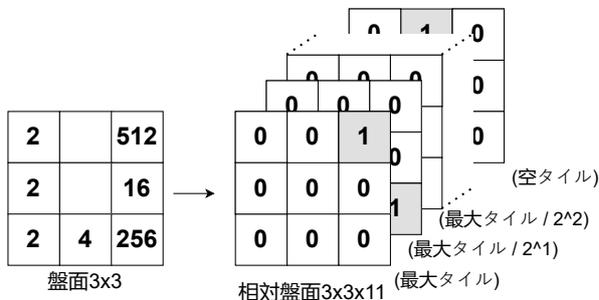


図 2: 入力: 相対入力

表 1: 上位 3 件 24 時間学習結果

プレイヤー	平均得点	中央値
deep	4448.88	5094
deep (通常+相対入力)	4446.70	5064
deep (restart)	4445.22	5144

参考文献

- [1] 山下修平, 金子知適, 中屋敷太一. 3×3 盤面の 2048 の完全解析と強化学習の研究. 第 27 回ゲームプログラミングワークショップ (GPW-22), pp. 1-8, 2022.
- [2] 金子智弥, 松崎公紀. ミニ 2048 における ニューラルネットワークプレイヤーの解析. 第 66 回プログラミング・シンポジウム予稿集, pp. 121-128, 2025.
- [3] 勝田塔子. ミニ 2048 におけるニューラルネットワークプレイヤーの分析と n タプルネットワークプレイヤーとの性能比較. 高知工科大学 修士学位論文, 2024.
- [4] K. Matsuzaki. Developing 2048 player with backward temporal coherence learning and restart. In *Proceedings of Fifteenth International Conference on Advances in Computer Games (ACG2017)*, pp. 176-187, 2017.