

# ゲームの特性に着目して抽出した盤面を用いた 2048 プレイヤの学習

1190341 寺村 舞童華 【高度プログラミング研究室】

## 1 はじめに

ゲーム 2048 とは、2014 年に Gabriele Cirulli によって公開されたパズルゲームである [1]。ゲーム 2048 は、縦横 4×4 の盤面でタイルの移動と結合を繰り返し 2048 のタイルを作成することが目標である。他のパズルゲームより比較的ルールを把握しやすいがゲームの性質上極めるのが難しく世界中で多くの人々を惹きつける魅力がある。

ゲーム 2048 において最も成功したアプローチは N タプルネットワークによるものである。本研究では、ゲームの特性に着目して抽出した盤面を用いて、N タプルネットワークベースで 2048 プレイヤを学習させる。学習方法は松崎が提案した後退 TC 学習 [2] を用いて 10<sup>8</sup> 盤面分を行う。また、学習させたプレイヤに greedy 法と expectimax 法を用いてゲームをプレイさせ、得られたスコアを比較する。

## 2 N タプルネットワークプレイヤについて

N タプルネットワークは、n 個のタプルによって構成されており 1 つのタプルは m 個のマス目の組である。N タプルネットワークプレイヤとは、各タプルで一致するマス目のタイルが一回出現するたび特徴量を抽出し、その特徴量の総和から盤面の評価値を算出し移動方向を選択していくプレイヤである。

## 3 後退 TC 学習について

時刻  $t$  における盤面  $S_t$ 、移動後のスコア  $r_t$  と盤面を  $S'_t$ 、盤面の評価値を  $V(S_t)$  および  $V(S'_t)$  とする。また、時刻  $t$  から 1 時刻前を  $t-1$  とする。

評価値の誤差  $\Delta$  は以下のように定義される。

$$\Delta = r_t + V(S'_t) - V(S_{t-1})$$

誤差  $\Delta$  を小さくするために全ての N 個のタプルについて特徴量  $V_i(S'_{t-1})$  を  $\Delta$  のある部分だけ更新する。また、学習率  $\alpha$  は誤差の累積  $E_i[S]$  と絶対誤差の累積  $A_i[S]$  を用いて、比率  $|E_i[S]|/A_i[S]$  によって決まる。

後退 TC 学習は以上の更新を行いながらゲームの終了盤面から開始盤面まで遡りながら学習を行う方法である。

## 4 実験

### 4.1 実装する N タプルネットワークプレイヤについて

本研究では、6 個のマス目の組を 4 つ用意し N タプルネットワークプレイヤを実装する。

### 4.2 実験で使用する盤面について

本研究では、松崎による N タプルネットワークプレイヤのプログラム [2] を元に初期盤面を変更し実験を行う。

実験で使用する盤面を図 1 と図 2 に示す。図 1 は上部に大きい値のタイルがある盤面を示しており、図 2 は下部に大きい値のタイルがある盤面を示している。

2048	2	2048	1024
256	128	64	32
8	4	2	2
4	2		

図 1 盤面 a

4	2		
8	4	2	2
256	128	64	32
2048	2	2048	1024

図 2 盤面 b

## 4.3 実験結果

梗概提出時点では実験を進めている段階であるため、実験結果については論文本体にに記載する。

## 5 まとめ

本研究では、ゲームの特性に着目して抽出した盤面を使用し、N タプルネットワークベースで 2048 プレイヤを学習させた。また本論文では、学習させたプレイヤに greedy 法と expectimax 法を用いてゲームをプレイさせ、得られたスコアを比較した結果を述べる。

## 参考文献

- [1] Gabriele Cirulli, “2048”, <https://play2048.co>, 2018 年 10 月 17 日閲覧。
- [2] Kiminori Matsuzaki, “Developing 2048 Player with Backward Temporal Coherence Learning and Restart” Proceedings of Fifteenth International Conference on Advances in Computer Games (ACG 2017), pp. 176-187, 2017.