

## 進捗報告

### 1 今週やったこと

- B3 実験用データ集め, 資料とスライド作成
- カードの特殊効果実装
- 初期手札を変えた際の勝率の変化
- ゲームバランス調整手法の検討

### 2 カードの特殊効果実装

環境に変化を加える時のために, 以下の 2 つの特殊効果を実装した.

- 自身が盤面にいる時は対戦相手は自身しか攻撃ができない
- 盤面に出す時, 自盤面と敵盤面のカードを全て破壊する

デバックは終わっているため入れようと思えばいつでもデッキに入れることができるが, まずは現在の環境で成果を出してからにする.

### 3 初期手札を変えたときの勝率の変化

先週にランダム環境における後攻の初期手札の枚数を 3 → 4 枚に変更しただけで後攻の勝率が 1 割近く向上した. この結果を受けて, デッキのバランスの悪さを表す指標になるのではないかと考えていた. 森先生に初期手札の組をいろいろ変えてやってみようアドバイスを頂いたので実験した. 表 1 にデッキの内容を示す. 変更点は無い.

カードの特殊効果は,

- 盤面に出したら (攻撃力, HP) = (1, 1) のユニット追加で出す. (召喚)
- 盤面に出したら自プレイヤーの HP を 2 回復 (治癒)
- 盤面に出したら敵プレイヤーの HP を 2 削る (攻撃)
- 盤面に出したら自プレイヤーは 1 枚カードをドロー (循環)
- 盤面に出たターンに攻撃できる (速攻)

となっている. なお, 結果として示しやすいようにカードに ID を割り振っている.

先週にカードパワーを図る指標として導入した Win Rate when Play (WRP) が高い順から 5 分割してそれぞれ 3 種類ずつ後攻の初期手札として決定してランダム動詞の対戦環境の後攻の勝率を計算した. 表 2 に結果を示す. 初期ドローによる勝率の変化が大きかったため, 新しく Win Rate when Draw (WRD) の指標を追加して確認してみた. 表 3 に結果を示す.

表 1: デッキの内容

ID	攻撃力	HP	コスト	特殊効果	枚数
0	1	1	0	無し	2
1	2	1	1	無し	2
2	3	2	2	無し	2
3	4	3	3	無し	2
4	5	4	4	無し	2
5	2	2	2	召喚	2
6	2	3	3	召喚	2
7	1	1	1	循環	2
8	1	3	2	循環	2
9	2	1	2	速攻	2
10	3	1	3	速攻	2
11	1	2	2	攻撃	2
12	2	3	3	攻撃	2
13	1	1	1	治癒	2
14	2	1	3	治癒	2

表 2: 初期手札の組合せを変化させたときの勝率

初期手札 カード ID	勝率
4, 3, 6	0.57672
8, 12, 5	0.57734
2, 7, 11	0.51678
1, 14, 10	0.37751
0, 13, 9	0.37728

## 4 考えてみたこと

1. 深層強化学習でバランス調整したいデッキを用いてアグロ, コントロール寄りの戦略を取る対戦相手に対してどちらにも勝てる戦略を作成する.
2. 作成したエージェントの戦略下で, WRP や WRD などを用いてカードパワーを調べる.
3. バフ・ナーフするカードを限定し, 目標となる勝率になるまで GA を用いてカードのパラメータを調整する.

### 4.1 GA でパラメータ調整

カードの HP, 攻撃力, コストのどこをどう調節するかということが定式化できればよかったが, そこで行き詰まっているため, カードの 3 種のパラメータの調整を, 組合せ最適化問題, 離散最適化問題とみなして遺伝的アルゴリズムを用いて調節する方法をとった. なお, GA の実装の都合上デッキの初期値を表 4 のように変更した. また, DQN の適用実験についてプレイヤーの初期手札を 5 枚と変更していたため, ランダム同士の対戦環境でも初期手札の枚数を 5 枚とした.

表 3: カードごとの WRD (降順)

カード ID	WRD
5	0.51898
4	0.51643
8	0.51609
3	0.51453
7	0.51439
6	0.51276
2	0.51272
12	0.51010
13	0.50555
11	0.50461
1	0.50437
0	0.50294
10	0.50016
9	0.50007
14	0.49625

この時, 50000 回ランダム同士で対戦させたところ, 後攻の勝率が 0.32962 となった. まず, 何も制限を掛けずにカードの攻撃力と HP, コストについて遺伝的アルゴリズムを用いてランダム同士の対戦環境において後攻の勝率が 0.50 に近づくように調整した. デッキにおいて ID 0 ~ 14 のカードについてそれぞれ 3 つパラメータがあり, 計 45 個パラメータが存在する. 45 要素の一次元配列として遺伝的アルゴリズムに適用した. 個体の適応度として, 一定回数シミュレーションを行い勝率を計算し

$$fitness = (0.5 - winrate)^2 \quad (1)$$

で計算した. なお, この定義の仕方のため適応度の値が小さい個体ほど優秀な個体とされる. 表 に各種パラメータを示す.

表 6 に結果として得られたデッキを示す. 変更されたパラメータ量は合計 25 となり原型をとどめていない. これを深層強化学習で最適化するカードを限定させ, マナレシオといったカードゲームにおけるヒューリスティックな知識を用いて変化量を抑えることができれば良いと考えている.

## 4.2 学習中に戦略が変わる相手に対しての DQN

以前までの対戦相手のルーチンは, 相手の盤面に破壊できるカードがあれば攻撃し, 緊急時以外は相手プレイヤーを攻撃してくるものだった. 今回は Algorithm 1 に示すような盤面を積極的に処理するルーチンを作成し, 対戦相手のルーチンを学習のエピソードごとに乱数で変化させて学習した. 先攻で 1000000 ステップ学習し, 10000 回対戦を行った結果, 0.9806 という勝率を得た. また, 表 7 に学習済みモデルで WRP, WRD を計測した結果を示す. 上位 5 つ, 下位 3 つが同じカードとなった.

## 5 今後やること

表 7 から調整対象のカードを選んで, 遺伝的アルゴリズムに適用させる. どのようにカードを選ぶか, また調整したデッキの評価方法を検討する必要がある.

表 4: GA の実装の都合上変更したデッキの内容

ID	攻撃力	HP	コスト	特殊効果	枚数
0	1	2	1	無し	2
1	2	2	2	無し	2
2	3	3	3	無し	2
3	4	3	4	無し	2
4	5	4	5	無し	2
5	2	2	2	召喚	2
6	2	3	3	召喚	2
7	1	1	1	循環	2
8	1	3	2	循環	2
9	2	1	2	速攻	2
10	3	1	3	速攻	2
11	1	2	2	攻撃	2
12	2	3	3	攻撃	2
13	1	1	1	治癒	2
14	2	1	3	治癒	2

表 5: 遺伝的アルゴリズムパラメータ

パラメータ	値
1 世代あたりの個体数	100
世代数	50
交配の確率	0.4
交配の種類	二点交配
突然変異の確率	0.2
選択	3 個体のトーナメント方式
勝率計算の際の回数	500

---

**Algorithm 1** 新たに作成した盤面処理を優先する行動ルーチン

---

```

1: 盤面にカードを 1 枚プレイ
2: for 盤面のカード (プレイ順が古い方から) do
3:   if 敵の盤面に 1 回の攻撃で破壊できるカードがある then
4:     その攻撃対象を選んで攻撃
5:   else
6:     if 敵盤面にカードがある then
7:       敵盤面の最も攻撃力高いカードを攻撃
8:     else
9:       敵プレイヤーを攻撃
10:    end if
11:  end if
12: end for
13: ターンを終了

```

---

表 6: GA を全てのカードに対して適用した場合のデッキの内容

ID	攻撃力	HP	コスト	特殊効果	枚数
0	1	2	1	無し	2
1	2	3	4	無し	2
2	1	3	3	無し	2
3	1	3	4	無し	2
4	1	4	5	無し	2
5	2	4	2	召喚	2
6	1	3	3	召喚	2
7	1	5	2	循環	2
8	1	3	2	循環	2
9	2	1	2	速攻	2
10	3	1	3	速攻	2
11	1	2	3	攻撃	2
12	1	3	3	攻撃	2
13	1	4	1	治癒	2
14	2	1	3	治癒	2

表 7: 学習済みモデルにおける WRP , WRD の計測結果 (降順)

WRP		WRD	
ID	値	ID	値
8	0.97787	8	0.97892
5	0.97720	5	0.97861
7	0.97631	7	0.97703
0	0.97585	0	0.97663
1	0.97418	1	0.97615
13	0.97346	2	0.97563
11	0.97288	11	0.97520
6	0.97287	4	0.97510
2	0.97239	3	0.97507
9	0.97235	6	0.97496
4	0.97076	13	0.97486
3	0.97040	9	0.97434
12	0.97001	12	0.97384
10	0.96775	10	0.97304
14	0.96474	14	0.97178