

進捗報告

1 今週やったこと

- 学習済みエージェント同士の 1 つ抜きの結果
- バランス調整実践

2 学習済みエージェント同士の 1 種類ずつ抜いた時の結果

表 1 に結果を示す.

表 1: カードを 1 種類ずつ除いたときの学習済みエージェント同士の 10000 回のゲーム実行における先攻の勝率 (太字は表内の最大値と最小値)

後攻 先攻	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
0	0.7596	0.5051	0.5122	0.5104	0.5021	0.5311	0.4947	0.5283	0.4861	0.5448	0.5676	0.5159	0.5306	0.5481	0.5106
1	0.8537	0.6677	0.6655	0.6576	0.6626	0.6788	0.6496	0.6651	0.6456	0.7019	0.6997	0.6676	0.6826	0.6757	0.6607
2	0.8583	0.6637	0.6673	0.6611	0.6664	0.6774	0.6511	0.6785	0.6370	0.7017	0.7097	0.6645	0.6855	0.6889	0.6621
3	0.8620	0.6663	0.6663	0.6713	0.6669	0.6776	0.6522	0.6782	0.6422	0.7041	0.7129	0.6790	0.6889	0.6944	0.6746
4	0.8690	0.6773	0.6742	0.6757	0.6783	0.6880	0.6569	0.6767	0.6488	0.7153	0.7133	0.6766	0.6937	0.6941	0.6710
5	0.8570	0.6508	0.6558	0.6590	0.6596	0.6613	0.6410	0.6562	0.6249	0.6843	0.6881	0.6597	0.6713	0.6749	0.6549
6	0.8745	0.6898	0.6743	0.6839	0.6784	0.6969	0.6629	0.6830	0.6605	0.7133	0.7246	0.6965	0.6977	0.6981	0.6857
7	0.8442	0.6494	0.6515	0.6373	0.6371	0.6520	0.6277	0.6566	0.6231	0.6755	0.6763	0.6531	0.6572	0.6640	0.6462
8	0.8836	0.7041	0.6896	0.6863	0.6960	0.7081	0.6855	0.7018	0.6688	0.7285	0.7358	0.7022	0.7078	0.7119	0.7030
9	0.8327	0.6403	0.6363	0.6357	0.6371	0.6493	0.6093	0.6422	0.5987	0.6554	0.6706	0.6317	0.6522	0.6456	0.6364
10	0.8392	0.6255	0.6219	0.6179	0.6195	0.6393	0.6100	0.6320	0.6019	0.6616	0.6734	0.6437	0.6487	0.6495	0.6219
11	0.8621	0.6725	0.6688	0.6587	0.6669	0.6754	0.6480	0.6604	0.6306	0.6982	0.7019	0.6603	0.6785	0.6734	0.6570
12	0.8598	0.6553	0.6475	0.6518	0.6594	0.6622	0.6282	0.6538	0.6253	0.6988	0.6972	0.6659	0.6658	0.6676	0.6500
13	0.8261	0.6381	0.6388	0.6340	0.6277	0.6466	0.6195	0.6545	0.6047	0.6688	0.6736	0.6442	0.6545	0.6452	0.6364
14	0.8469	0.6617	0.6608	0.6585	0.6604	0.6768	0.6463	0.6668	0.6254	0.6950	0.6971	0.6680	0.6711	0.6732	0.6698

アグロ同士だと 14 が最も弱いカードと判断できたが, 学習済みエージェント同士では表 1 から, ID 8 のカードが最も弱いと判断できる. これはエージェントが構築した戦略におけるカードの優先度まで学習しているため ID 14 のカードはあまり対戦に登場しなかったためと考えられる. 自作したアグロの AI によるシミュレーションと比較して, デッキにおける最適な戦略下において最適なカード選んだ場合の意味のあるシミュレーションができていると考えられる.

3 バランス調整実践

表 2 に示すこれまでの実験で用いてきたデッキはアグロが最適だと分かったため, 表 3 のようにアグロ向けのデッキとして環境に追加することを考える. 調整するパラメータは, 15 種類のカードの HP, 攻撃力, コストの $15 \times 3 = 45$ 個のパラメータである. まず, HearthStone のデッキ間の勝率を 50 % に近づけるように GA を

用いていた関連研究 [1] 通り GA と NSGA-II を用いて最適化する。GA における適応度, NSGA-II における目的関数の 1 つは, 表 3 の r を用いて

$$f = \sum_{i=0}^4 \sqrt{(0.50 - r_i)^2} \quad (1)$$

と計算する。表 4 に適応度を最小にするように GA を適用した結果と, (1) 式の勝率とパラメータの総変更量の 2 つを最適化するように NSGA-II を用いた結果を示す。また, 図 1 に NSGA-II で得られたパレートフロントを示す。また, 表 5, 6, 7 に何も調整してない時, GA で最適化した時, NSGA-II で最適化した時の環境を示す。

表 4 から, 関連研究で示されたように多目的最適化手法である NSGA-II を用いることでパラメータの総変更量が減少されることが本環境においても数値実験で確かめられた。しかし, 変更が合ったカードの枚数は 15 枚中 13 枚となっている。関連研究の文脈では, アプデなどの際に一度にたくさん調整が入るとユーザーにとって好ましくないため NSGA-II を用いてパラメータの総変更量を減らすといったアプローチだったが, 本環境では GA と比べて変更を免れたカードは 2 枚しか無い。ここで, 変更するカードを限定することで「変更があったカード枚数」を直接限定しながら GA を試してみた。変更するカードの限定には, 表 1 を用いて, 以下の手順で変更するカードの優先順位を決定する。

1. 表内の最大値, 最小値を見て, 最も強いカード, 最も弱いカードを確かめる
2. その 2 枚において対角線上 (先攻後攻そのカードを抜いた時の勝率) を見て, 何も抜いてない時の勝率と比較し, 差の絶対値が大きいカードを変更する優先度が最高のカードとする
3. 先攻で そのカードを抜いた状態で後攻で何も抜いてない場合の勝率を計算する。
4. 先攻が優先度最高のカードを抜いた場合の行において, 3 で計算した勝率と差の絶対値を取る。
5. 値が大きいほど変更する優先度を高く設定する。

今回の実験においては, お互いに何もカードを除いてない時の勝率は 0.66532 となったため, 2 で ID 0 のカードが優先度最高のカードとする。3 において, 先攻が ID 0 のカードを除いて後攻が何も除いてない時の勝率は 0.5348 となった。このように決定した優先順位はカード ID で表すと,

$0 > 8 > 6 > 10 > 4 > 1 > 3 > 14 > 2 > 11 > 13 > 9 > 7 > 12 > 5$ となった。この順番に沿って ID 0 のカードのみ変更という条件から徐々に GA による解空間を次元を増やしていったデータをとっている。表 8 に現時点での結果を示す。

そろそろ時間的にまずい, 多分木曜くらいに計算が終わるので金曜にデータ取り終わったら相談させてください。この実験のデータが使いなさそうだったら 表 1 までで卒論書きます。

参考文献

- [1] Fernando de Mesentier Silva, Rodrigo Canaan, Scott Lee, Matthew C. Fontaine, Julian Togelius, and Amy K. Hoover. Evolving the hearthstone meta. *CoRR*, Vol. abs/1907.01623, , 2019.

表 2: 実験で用いた強すぎるカードと弱すぎるカードを含んだデッキ (ID 0 が強い, ID 14 が弱い)

ID	攻撃力	HP	コスト	特殊効果	枚数
0	4	4	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	取得	2
8	1	3	2	取得	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	1	1	5	治癒	2

表 3: バランス調整する環境

後攻 先攻	新追加デッキ (アグロ)	アグロ	コントロール
新追加デッキ (アグロ)	r_0	r_1	r_2
アグロ	r_3	0.5255	0.5424
コントロール	r_4	0.5121	0.5053

表 4: GA と NSGA-II の結果

評価指標 アルゴリズム	パラメータの総変更量	勝率に関する適応度	変更が合ったカード枚数
GA	80	0.16080	15
NSGA-II	50	0.26060	13

表 5: 何も調整していない時の先攻の勝率

後攻 先攻	新追加デッキ (アグロ)	アグロ	コントロール
新追加デッキ (アグロ)	0.6724	0.8005	0.8859
アグロ	0.3383	0.5255	0.5424
コントロール	0.2774	0.5121	0.5053

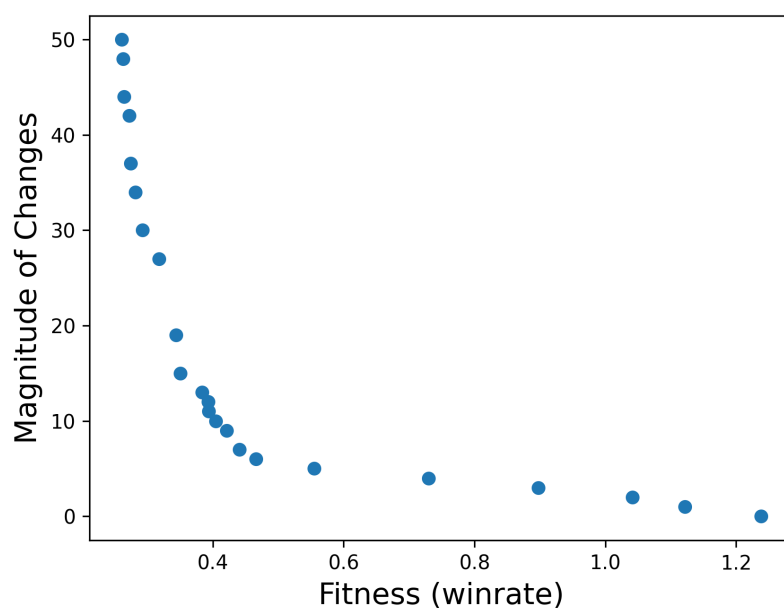


図 1: NSGA-II で得たパレートフロント

表 6: GA で調整した時の先攻の勝率

後攻 先攻	新追加デッキ (アグロ)	アグロ	コントロール
新追加デッキ (アグロ)	0.5362	0.5311	0.5678
アグロ	0.5244	0.5255	0.5424
コントロール	0.4992	0.5121	0.5053

表 7: NSGA-II で調整した時の先攻の勝率

後攻 先攻	新追加デッキ (アグロ)	アグロ	コントロール
新追加デッキ (アグロ)	0.5814	0.5218	0.5662
アグロ	0.5613	0.5255	0.5424
コントロール	0.5211	0.5121	0.5053

表 8: 現時点で得ているデータ

変更するカード ID	パラメータの総変更量	勝率に関する適応度
0	10	0.7330
0, 8	15	0.5288
0, 8, 6	15	0.4893
0, 8, 6, 10	23	0.4559
0, 8, 6, 10, 4	31	0.4471