

進捗報告

1 今週やったこと

- 学習済みモデルを先攻後攻両方に配置して対戦させる機能実装
- アグロとコントロールで 50 % 近い勝率を記録するように対戦相手 AI 改良
- 数値実験

2 学習済みモデルを先攻後攻両方に配置して対戦させる機能の実装

今までは、学習済みモデル vs 予め作成しておいた対戦相手しか検証ができなかったが、学習済みモデルを先攻後攻両方に配置して対戦を実行できるようにした。

3 アグロとコントロールで勝率調整

先週は、アグロとコントロールの対戦相手を作成し DQN の対戦相手として適用したが、そもそもお互いの戦略同士の強さが大きく偏っていた。そのため、今週はなるべく 50 % に近づくように戦略を調整した。また、デッキが同じのままで調整することが難しかったためデッキを戦略ごとに分けてデッキ、戦略間の勝率を調整した。表 1 に調整した結果を示す。また、表 2, 3 に各戦略におけるデッキを示す。

表 1: 10000 回のゲーム実行における先攻の勝率

後攻 先攻	アグロ	コントロール
アグロ	0.5255	0.5424
コントロール	0.5121	0.5053

4 数値実験

4.1 対戦相手を変更して DQN を適用

3 節で示した敵 AI を対戦相手として先攻において、DQN を用いて後攻側の行動を学習した。学習後、そのまま 10000 回ゲームを実行し勝率を計算した。また、表 5 に示したデッキを学習側のプレイヤーに持たせている。勝率に有意な差が見られなかったため、10000 回の対戦における学習済みエージェントの行動を計測し分析した。表 6 に結果を示す。全て同じようになった。対戦相手の戦略に応じてではなく学習するデッキにおける戦略の最適解を学習しているのではないかと考えられる。

表 2: アグロのときのデッキ

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

4.2 カードを 1 種類ずつデッキから抜いた場合の勝率計算

デッキ内のカードパワーを測る指標として、これまで Win Rate when Play (WRP) といった指標を試してきたが、これらの指標ではエージェントの対戦記録から計算しようとする構築された戦略がアグロによっているためカードのコストに結果が影響されてしまう問題があった。そこで、森先生から先週アドバイスを頂いたように、同戦略同デッキでカードを 1 種類ずつデッキから抜いた場合の勝率を計算してみた。デッキは表 5 を用いて、表 7 に作成したアグロの行動ルーチンを先攻後攻両方に配置しカードを 1 種類ずつ抜いた場合の 10000 回ゲームを実行した際の勝率を示す。行に注目すると、先攻がカード ID 0 のカードをデッキから除いた時に勝率が大きく減少していることが分かる。また、カード ID 1 のカードをデッキから除いた場合に勝率が大きく増加している。カードパワーを示す指標としてこの表は妥当性があることが分かった。

表 3: コントロールのときのデッキ

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

表 4: 学習後 10000 回のゲーム実行における学習側の勝率 (学習側は後攻)

学習時の対戦相手 勝率計算時の対戦相手	ランダムに戦略が変化する	コントロール	アグロ
コントロール	0.7163	0.7214	0.7131
アグロ	0.7110	0.7108	0.7230

表 5: 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 6: 3 種類の学習済みエージェントの行動

ランダムに戦略が変化		コントロール		アグロ	
行動説明	総数	行動説明	総数	行動説明	総数
ターンエンド	49030	ターンエンド	49463	ターンエンド	49016
手札 1 を盤面に出す	42920	手札 1 を盤面に出す	43340	手札 1 を盤面に出す	42950
盤面 1 で相手プレイヤーに攻撃	39378	盤面 1 で相手プレイヤーに攻撃	39791	盤面 1 で相手プレイヤーに攻撃	39319
盤面 2 で相手プレイヤーに攻撃	20684	盤面 2 で相手プレイヤーに攻撃	20762	盤面 2 で相手プレイヤーに攻撃	20914
手札 2 を盤面に出す	13673	手札 2 を盤面に出す	13569	手札 2 を盤面に出す	13337

表 7: カードを 1 種類ずつ除いたときの 10000 回のゲーム実行における先攻の勝率

後攻 先攻	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
0	0.7857	0.4998	0.4957	0.4779	0.4671	0.5059	0.4914	0.5153	0.4845	0.5105	0.5099	0.4953	0.5128	0.5288	0.4335
1	0.8955	0.6762	0.6798	0.6688	0.6613	0.6963	0.6789	0.6913	0.6807	0.6925	0.6914	0.6855	0.6998	0.6974	0.6345
2	0.8936	0.6763	0.6865	0.6667	0.6652	0.6961	0.6826	0.6850	0.6697	0.6839	0.6864	0.6938	0.6903	0.6958	0.6473
3	0.9088	0.6989	0.6939	0.6844	0.6700	0.7119	0.6991	0.6967	0.6932	0.7130	0.6996	0.7061	0.7116	0.7135	0.6599
4	0.9049	0.7043	0.7073	0.6816	0.6834	0.7261	0.7109	0.709	0.7008	0.7193	0.7173	0.7154	0.7163	0.7269	0.6729
5	0.8773	0.6547	0.6635	0.6511	0.6219	0.6721	0.6586	0.6594	0.6451	0.6780	0.6711	0.6682	0.6732	0.6748	0.6211
6	0.8929	0.6799	0.6845	0.6701	0.6673	0.7083	0.6795	0.6865	0.6802	0.6988	0.6844	0.7040	0.6925	0.7013	0.6397
7	0.8698	0.6539	0.6565	0.6439	0.6317	0.6659	0.6464	0.6602	0.6459	0.6660	0.6637	0.6626	0.6711	0.6750	0.6175
8	0.9028	0.6904	0.6897	0.6741	0.6772	0.7063	0.7011	0.6949	0.6793	0.7102	0.7046	0.6954	0.7156	0.7127	0.6621
9	0.8762	0.6568	0.6625	0.6549	0.6325	0.6677	0.6630	0.6633	0.6539	0.6782	0.6676	0.6618	0.6792	0.6737	0.6215
10	0.8846	0.6711	0.6716	0.6443	0.6454	0.6883	0.6706	0.6724	0.6609	0.6807	0.6740	0.6744	0.6904	0.6886	0.6242
11	0.8853	0.6671	0.6820	0.6580	0.6484	0.6960	0.6718	0.6843	0.6759	0.6840	0.6797	0.6812	0.6913	0.6862	0.6253
12	0.8738	0.6525	0.6594	0.6407	0.6337	0.6794	0.6553	0.6651	0.6576	0.6708	0.6695	0.6689	0.6729	0.6801	0.6234
13	0.8714	0.6428	0.6570	0.6338	0.6262	0.6751	0.6508	0.6673	0.6456	0.6666	0.6616	0.6673	0.6753	0.6710	0.6153
14	0.9296	0.7308	0.7203	0.7199	0.6943	0.7450	0.7287	0.7391	0.7261	0.7400	0.7381	0.7334	0.7456	0.7422	0.6889