

進捗報告

1 2 つのモデルを用いて交互に学習

2 つモデルを用意し 1 つのモデルを対戦相手にして、もう 1 つのモデルを追加で学習できるように実装した。図 1 に今週行った実験の概略を示す。

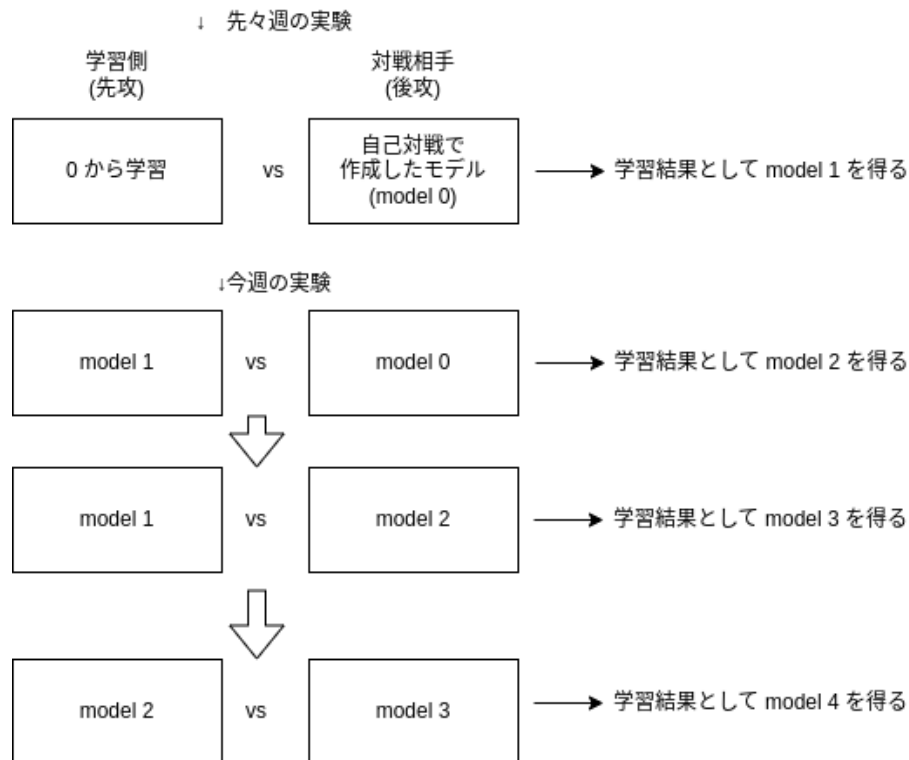


図 1: 今週の実験イメージ図

双方ともデッキは同じにして、1000000 ステップ学習させた。学習後、10000 回同条件で対戦を行い学習側の勝率を記録した。図 2, 3, 4 にそれぞれの対戦の学習中の報酬の推移を示している。

特に model 2 vs model 1, model 3 vs model 2 の学習の場合、学習後の検証段階における勝率はそれぞれ 0.6702, 0.6709 となり、着実にモデルは強くなっていっていることが考えられる。また、図 3, 4 を見る限り学習済みモデルの重みの読み込みが上手く行えておらず 0 からの学習と等しい推移の仕方をしていると考えられる。実装を見直してみる。

2 ChatGPT によるカードプール作成

ChatGPT にカードプールを作成してもらった実験をした。前回、与えたプロンプトが雑なものだったためカードゲームの内容を詳しくプロンプトに記述した。実際に画面見せて説明します。

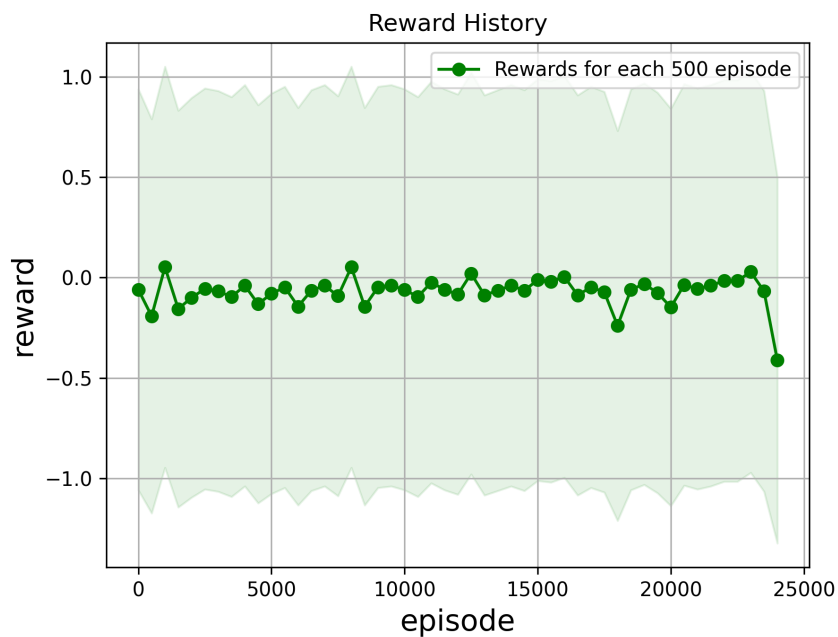


図 2: model 1 vs model 0 の報酬の推移

3 今後の課題

- 自己対戦の実装

自己対戦の部分の実装で詰まり散らかしてるのでなんとかする. 上手く実装ができなさそうな時は逆転オセロニアの研究で用いられていた HandyRL¹ など新しいライブラリの導入を試してみたい. 他の強化学習アルゴリズムの検討もありかもしれない.

- GA を用いたデッキ構築の実装

カードプールから GA を用いてデッキを構築するまでの流れを実装する.

¹<https://github.com/DeNA/HandyRL>

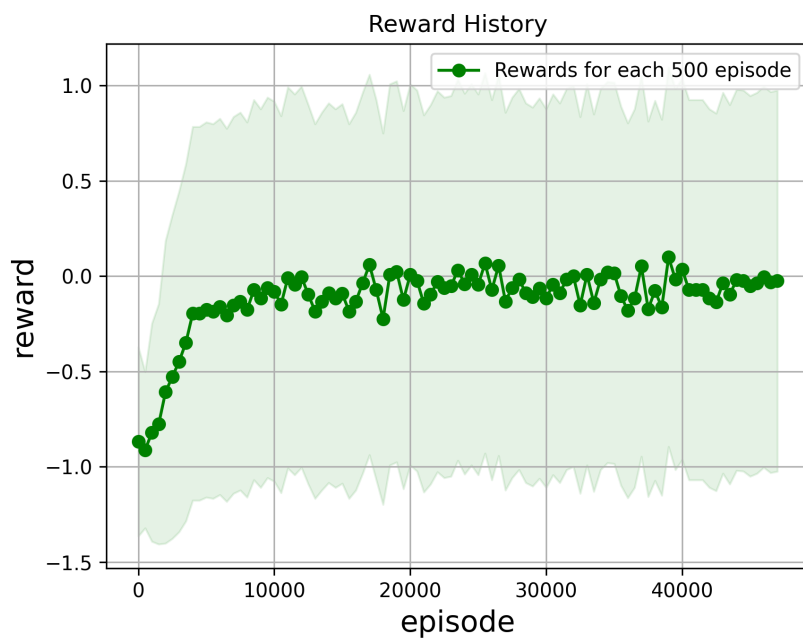


図 3: model 2 vs model 1 の報酬の推移

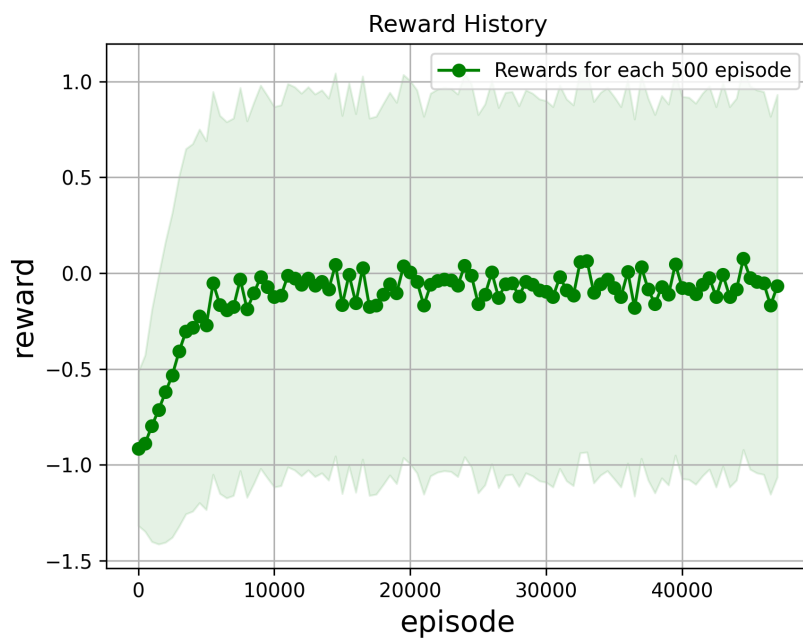


図 4: model 3 vs model 2 の報酬の推移