

## 進捗報告

### 1 今週やったこと

- 学習ステップ数を増やした DQN の結果
- DQN の実験の結果を踏まえた改善
- 再実験
- コスト, HP , 特殊効果追加した ver の環境の作成

### 2 学習ステップを増やした DQN の実験結果

先週, DQN で 200000 ステップ (約 5000 エピソード) 学習した結果先手が勝率 5 割を切るという結果となった. 一方, 新たに実装したモンテカルロ探索では 1000000 エピソード学習し先手で 0.8011 という勝率を残した. この結果を受け, DQN のステップ数を増やして学習が行われるか実験を行った. パラメータは表 1 に示す.

表 1: DQN のパラメータ

方策	-greedy
	0.1
全結合層の活性化関数	ReLU
全結合層の次元	64
最適化アルゴリズム	Adam
学習率	1e-2
Experience Replay のメモリ量	1000000

3 ???

4 ???

参考文献