

まとめ: 実験と結果

- Alpha Zeroアルゴリズムの再現性を確認
- 学習毎の自己対戦のグラフから敵対的に学習できていることを確認
- 学習したモデルの経験的評価として、全モデル同士の総当たり戦とMinMaxアルゴリズムとの対局
- この先も学習を続けることでさらに強くなることが示唆された
- 学習序盤では負の方向にバイアスの効いた盤面評価を行っていることの視覚化