学際大規模情報基盤共同利用·共同研究拠点公募型共同研究 平成30年度採択課題

10th Symposium

jh180067-DAJ

金子知適 (東京大学)

大規模な強化学習技術の実証と応用

研究背景と目的

- 強化学習手法の台頭
- DQN, A3C (Atari 2600)
- AlphaGo Zero (囲碁), Alpha Zero (将棋・チェス)

人間の知識なしで人間を超えた

囲碁将棋を主な対象に大規模強化学習の知見を蓄積

- 探索の効率化による学習の高速化
- 計算資源の削減 lacksquare

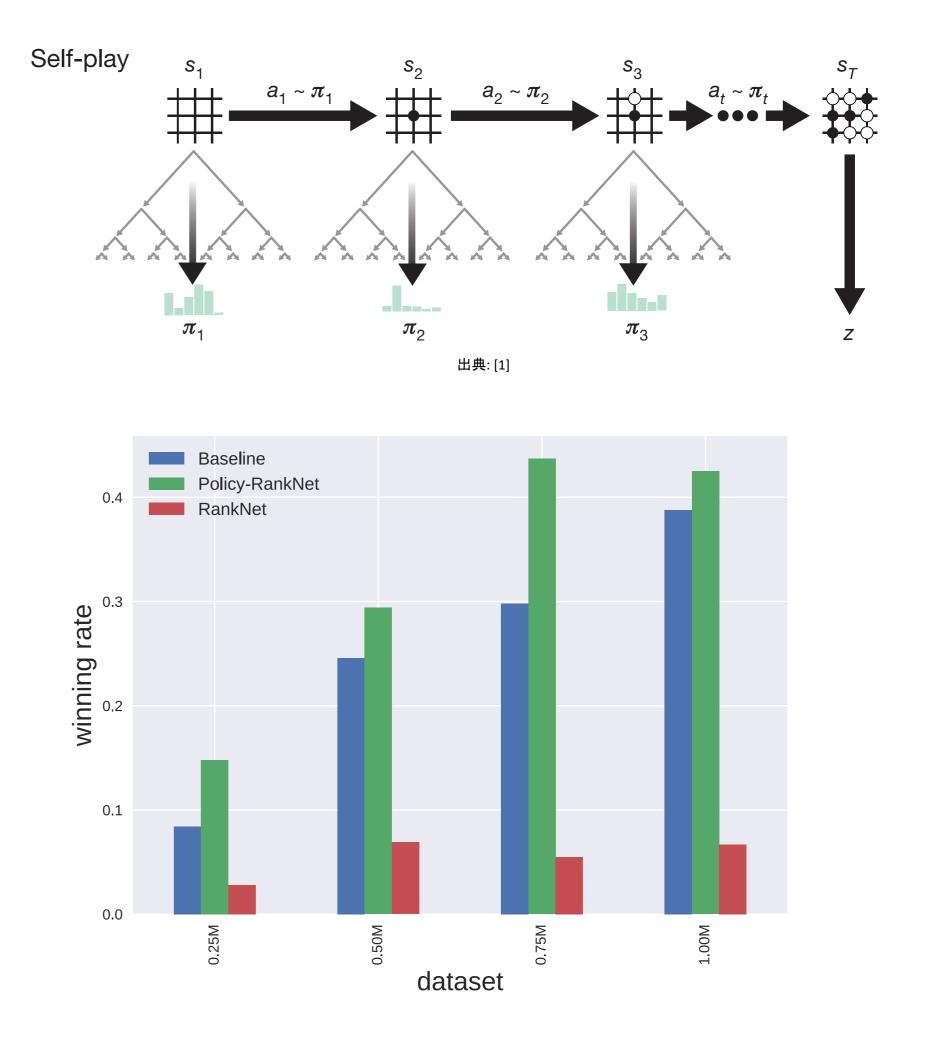
Joint Usage / Research Center for Interdisciplinary Large-scale Information Infrastructures



一方	で、
• =	- •

- 計算資源の肥大化 AlphaGo Zero の再現には 1700 年かかる
- 学習結果の信頼性の検証 なぜこの結果になるのか人間に説明

という課題



学習結果の可視化と説明

「なぜこの局面だとこの手が評価が高いのか」

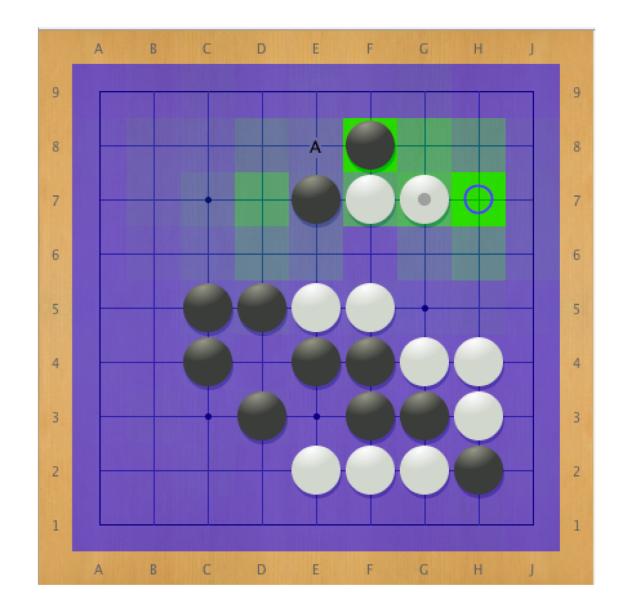
局面対による学習の高速化

- 2つの局面の優劣を学習することにより必要な学習局面 (i.e. 自己対戦回数)を削減
- → Learning to Rank によって価値関数を学習 学習のデータ数を二乗のオーダーで増やせる

既存手法と比較して全体的に性能が向上

学習器の説明

学習したニューラルネットワークが「よく見ている」場所を可視化 Saliency Mapを用いて最も注目している位置をハイライト





- 探索部分の効率化を行い、学習を更に高速化 まだ良くわかっていない局面をより重視するような探索
- 探索を組み合わせた学習器の説明

References

- D. Silver et al, "Mastering the game of Go without human knowledge", In: Nature 24270
- O. David et al, "DeepChess: End-to-End Deep Neural Network for Automatic Learning in Chess", In: ICANN 2016

JHPCN

学際大規模情報基盤共同利用・共同研究拠点 第10回シンポジウム

Japan High Performance Computing and Networking plus Large-scale Data Analyzing and Information Systems

2018年7月12日,13日

