

▼ 第13回UEC杯コンピュータ囲碁大会参加プログラム

「nlp2」 アピール文
井 建志

代表者 若

はじめに第11回UEC杯では引き分けを引き起こし、運営の皆様に御迷惑をおかけしてことをお詫び申し上げます。

「nlp2」の基本的な構成はpolicy headとvalue headを持つネットワークによるPolicy-Value モンテカルロ木探索という、Silverらが2018年に発表したAlphaZero (文献1)と同様である。ただし自力で強化学習は実施せず、Leela Zeroの公開学習用データを用いている (文献2)。

またネットワーク構成についても、基本はresidual network 20ブロック×2層の後にpolicy headとvalue headを加えるという一般的なもので、一部はKataGo (文献3)を参考にglobal poolingを加えている。ネットワークへの入力、手番の石、手番でない側の石、手番、盤端の4チャンネルに簡略化している。

従来のnlpで使用していたrollout policyは廃止し、ソースコードも文献4のpythonによる実装を参考に代表者がc++で新たに作成したため (連を実装していないなど、完全な移植ではない)、従来バージョンとはほとんど連続性がないことから、今回のプログラムは「nlp2」と改名した。

最近1年間、いくつかの手法を試みたものの、残念ながら実用化には至らなかったため、今回は上記の基本構成による実装の動作を確認し、今後、とくに学習手法の改善を図る際のベースとしたい。

注1 代表者はAIの研究者ではなく、個人の趣味として囲碁プログラムを開発している。

注2 「nlp」は「no life prince」の略で、convolutional neural networkを初めて用いた際、あたかも人間が囲碁を打っているような着手をすることに感動して名付けたものであり、「natural language processing」の略ではない。

Short English summary

The nlp2 ("no life prince", not "natural language processing") program has a network with policy and value heads (20-block residual network). The policy-value Monte Carlo tree search is applied to determine moves. The network is trained by using the training data for Leela Zero.

参考文献 (References)

1. Silver D, Schrittwieser J, Simonyan K, et al. Mastering the game of Go without human knowledge. Nature. 2017; 550: 354-359.

2. <https://leela.online-go.com/training/> (最終閲覧 2022.1.30) .
3. Wu DJ. Accelerating Self-Play Learning in Go. arXiv:1902.10565.
4. Pumperla M, Ferguson K 著、山岡忠夫 訳. 囲碁ディープラーニングプログラミング. マイナビ出版 (東京)、2019.