

ponkotsuアピール文書

概要

本プログラムは定跡を用いることなくDL系ソフトに振り飛車を指させることを目的に開発を行っている。

採用手法

本プログラムでは2つのNNモデルを使用して指し手を生成している。一方はPolicy専用のモデルであり、他方はValue専用のモデルであり、いずれも10bのResNetである。

Policy専用のモデルは振り飛車を指すソフトとして有名なHoneyWaffle同士を手元で5分1手5秒加算で対局させた棋譜約2万局分を教師データとして20epoch学習させた。尚、教師データ生成において初期局面集やrandom moveは使用していない。また、重複局面の削除も行っていない。

Value専用のモデルはfloodgateの2015年~2021年の棋譜おおよそ50万局分(数えていないため正確な値ではない)を教師データとして学習させた。こちらは、年単位で分割して4epochずつ学習を行った。ただし2015年と2016年はまとめて学習を行った。各年毎にレーティングの条件指定を変えており、レーティングの条件指定は以下の通りである。下表のレーティング以上のソフト同士の対局を教師データとして用いた。

年	レーティング
2015~2016	1600
2017	2000
2018	2500
2019	2800
2020	3000
2021	3500

尚、こちらも同様に重複局面の削除も行っていない。

どちらのモデルも学習時には精度向上のため次に指す手(Policy)と勝敗(Value)双方を学習するマルチタスク学習を行った。

探索アルゴリズムはMCTSを採用した。尚、探索の並列化は行っていない。

課題

本大会を通じて以下の課題が明らかとなった。

- 評価値が控えめに出る(他のソフトなら1000点を出しそうな局面でも300点と出す)
- 1つの局面に対し2度推論を行うため、NPSが低い
- 捌く手を指さない(振り飛車を指すにも関わらず居飛車のような将棋を指す)
- 疑問手や悪手が多い

これらは2つのモデル間で相異なる傾向にある教師データを用いて学習を行っていることに起因していると考えられる。また、疑問手や悪手が多い原因としてPolicy専用のモデルにおいて単独のソフト、かつ重複局面の削除を行っていない棋譜を教師データとしているため他のソフトなら悪手と判断するような手を学習しているためと考えられる。

今後の展望

今後は公開されているGCTの学習に使用したデータセットやプロ棋士の振り飛車の棋譜を使用して学習を行うことを考えている。また、NPS向上のため探索の並列化を行う予定である。

使用したライブラリ

- [python-dlshogi2](#)
- [HoneyWaffleWCSC28](#)