

# 第 32 回世界コンピュータ将棋選手権 dlshogi with HEROZ 詳細アピール文章

山岡忠夫  
加納邦彦  
山口祐  
大森悠平  
2021/5/7

## 1 独自に工夫した点

ディープラーニング系の将棋 AI は、大局観に優れており、序中盤の形勢判断が従来型将棋 AI と比べて正確であるという特徴がある。一方、終盤の読みが重要になる局面では、従来型将棋 AI の方が正確な場合がある。

dlshogi は、終盤の課題に対処するために、独自の工夫を行っている。

具体的には、「MCTS の葉ノードでの短手数詰み探索」、「ルート局面で df-pn による長手数詰み探索」、「勝敗が確定したノードのゲーム木への伝播」、「PV 上の局面に対する長手数詰み探索」、「強化学習時に初期局面集を使用して局面の多様性を確保する」、「強化学習時に df-pn により詰み探索を行い詰みを報酬とする」という工夫を行っている。

これらのいくつかは、現在、dlshogi 以外のディープラーニング系の将棋 AI には取り入れられているが、dlshogi 以前にこれらを導入しているディープラーニング系の将棋 AI はなかった。

## 2 開発動機

2016 年 3 月に行われた AlphaGo とイ・セドル九段の対局で、AlphaGo が従来と異なるディープラーニングという手法で勝利したことに衝撃を受けた。囲碁の盤面を画像として入力して指し手を予測するという仕組みを理解したいと考えて、論文を読み AlphaGo のクローンの実装を行った。プロの棋譜を学習して対局するところまで実装できたが、囲碁 AI は、すでにアメリカ、中国の巨大企業も取り組んでいたため、個人で囲碁 AI を開発するモチベーションは続かなかった。将棋 AI では、まだディープラーニングの手法が実験レベルでしか試されていなかったため、自分が開発する意義があると考えて取り組むことにした。

## 3 開発過程

2017 年からはじめは教師あり学習でのモデル学習から始め、PV-MCTS の実装をしたことで、GPSFish に勝てることのできたため、ディープラーニングの手法が将棋 AI でも有効であることに確信が持てた。その後 AlphaZero が発表され、将棋でもトップレベルの強さにできることが報告された。AlphaZero は、膨大な計算リソースで実現されていたため、工夫を行うことで個人レベルの計算リソースでも強くできることを目標にして改良を続けた。

今大会では、HEROZ の将棋 AI 開発者のチームで参加した。社内の計算リソースを使って大規

模な強化学習を行った。チームメンバと共に、戦型・定跡といった大会に向けた作戦を準備した。

## 4 実験結果

2021 年秋の電竜戦では ResNet 15 ブロック 224 フィルタのモデルを使用したが、今大会では ResNet 20 ブロック 256 フィルタのモデルを学習した。floodgate の 2017 年から 2018 年 6 月までの棋譜を元に、レーティング双方 3500 以上、評価値 3000 まで、千日手、最大手数を除外という条件で作成したテストデータに対する精度は、以下の通りとなった。15 ブロックのモデルの精度も比較のために記載する。

	方策正解率	価値正解率
15 ブロック 224 フィルタ	52.74%	76.36%
20 ブロック 256 フィルタ	53.76%	76.80%

15 ブロックと 20 ブロックを 1GPU・3 スレッド、持ち時間 400 秒 1 手 2 秒加算で対局した際の結果は、以下の通り(ordo の結果)。持ち時間を 4 倍にした水匠 5 32 スレッドも加えたリーグ戦で計測している。

# PLAYER	:	RATING	ERROR	POINTS	PLAYED	(%)	CFS (%)	W	D	L	D (%)
1 20block	:	67.8	34.2	111.5	176	63	100	100	23	53	13
2 sui sho5-32thx4	:	-27.6	33.9	75.0	169	44	66	62	26	81	15
3 15block	:	-40.2	34.6	72.5	173	42	---	60	25	88	14

大会では、社内の複数台のマシン(A100 を 8 機搭載のサーバ 9 台)を活用するため、Multi Ponder を実装した。MCTS では有望な候補手を任意の N 手取得できることから、上位 8 手を別のマシンに割り当てる構成とした。親ノードでは相手局面からの探索を行うことで、有望な候補手を取得できるようにし、また ponderhit しなかった場合は、その探索結果を利用した。Multi Ponder の効果は測定はできていないが、大会ではほとんどの局面で ponderhit しており、定跡を抜けた後も 0 秒で指す局面が続いていた。

## 5 追試可能か

dlshogi の探索部、学習部のソースコードは GitHub<sup>1</sup>で公開している。学習方法については、ブログ記事<sup>2</sup>や書籍「強い将棋ソフトの創りかた<sup>3</sup>」で解説している。dlshogi の学習に使用しているデータも書籍に付属している。それらを元に、追記を行うことは可能である。

<sup>1</sup> <https://github.com/TadaoYamaoka/DeepLearningShogi>

<sup>2</sup> <https://tadaoyamaoka.hatenablog.com/>

<sup>3</sup> [https://honto.jp/netstore/pd-book\\_31311287.html](https://honto.jp/netstore/pd-book_31311287.html)