

コンピュータ将棋における高精度な深層学習モデル

芝 世式^{†1}

概要: 本研究ではコンピュータ将棋において過去に類のない高精度な深層学習モデルの学習に成功した。本モデルの全般評価を行うと共に過去の研究で困難とされていた終盤の探索無しでの詰み性能を確認する。

キーワード: コンピュータ将棋, 深層学習

An Accurate Deep Learning Model on Computer Shogi

SEIJI SHIBA^{†1}

Abstract: In this research, we succeeded in training an accurate deep learning model for computer shogi. In addition to evaluation of this model, we confirm the performance without searching in Tsume Shogi, which was considered difficult in past research.

Keywords: Computer Shogi, Deep Learning

1. はじめに

ゲームアルゴリズムや探索問題の研究題材としてチェスや囲碁、将棋などの題材が長年取り組まれており、近年は人間の最上位者を遥かに上回る実力を身につけてきたことが明らかになっている。今なお、新しい多くの試みが常に導入されコンピュータ同士の対戦においてその強さというものは年々向上の一途を辿っている。とりわけ AlphaZero[1]の知名度により深層学習モデルを用いた PV-MCTS モデルが多くของเกมに対応できるとの汎用性が示されている。しかしながら、当プロセスの成功にはモンテカルロ木探索[2]と深層学習と強化学習といった3つの要素が含まれるにも関わらず強化学習プロセスにのみ主眼が置かれており、深層学習モデルの作成プロセスやモデル・アルゴリズム等の選定過程は明らかにされていない。特に近年では深層学習モデルの精度向上が AI 強化に寄与する割合が多いものと認識されているにも関わらず十分な研究が行われているとは言えない。

本研究では第 32 回世界コンピュータ将棋選手権（以下 WCSC32）準優勝および電竜戦 TSEC3 で3位となった二番絞りの深層学習モデルを中心に探索無しでの評価を行う。

2. 二番絞りモデルの作成

二番絞りは電竜戦と共に生まれた。電竜戦は 2020 年の世界コンピュータ将棋選手権が中止となって同日オンライン大会が開催されたのをきっかけに、オンラインなら手軽に主催できるのではないかと比較的気軽に始まった。二番絞りは電竜戦第一回予行演習においてテストプログラムの

一つとして作成した。同年5月の選手権に向けて準備された強化学習モデルの副産物である数十億局面の教師データを深層学習モデルに学習させるのが基本コンセプトで絞リカスを絞る行為をなぞって二番絞りと命名された。AlphaZero に倣い 20 ブロック 256 チャンネルの ResNet 構造を用いた。

その後、我々の教師データより良いと思われる AobaZero の教師データが当時公開されており、その 1474 万棋譜全て追加学習させた上、約 1 億局面分の AlphaZero 型自己強化学習を計 2 回行っている。これが 2020 年の電竜戦モデルである。決勝では停電の影響で予選時の設備が利用できなかったが、予選では当時の dlshogi を凌駕する性能を見せており製作者の期待を上回った。[4]

WCSC31 の開催時期では NVIDIA の新しい GPU である A100 がクラウド提供されており探索性能の拡大が期待された。Remi ら一部の開発者には 20 ブロックのニューラルネットワークでは小さ過ぎて A100 の性能を生かすことができないと言われるに至った。我々は 40 ブロックへの移行を試み、20 ブロックモデルで生成した教師データ約 10 億局面を引き継ぐことで実際に高性能なモデルの学習に成功した。加えて 40 ブロックにおいて約 4 億局面分の AlphaZero 型自己強化学習を計 2 回行っている。このモデルは秒 11 局面程度の探索速度で floodgate レート 3300 付近を記録した。[5]

また、40 ブロックと大きなモデルを用いることにより A100 を 16 基搭載されたマシンにおいても探索速度はほぼ線形に推移することを確認している。

^{†1} 岡山県立大学 情報工学部
Okayama Prefectural University
shiba@cse.oka-pu.ac.jp

WCSC32 においては WCSC31 モデルをベースに強化学習を試行錯誤したが大幅な棋力向上を計測することが出来なかった。元々純粋な強化学習であることを放棄し、手段を選ばずに強いモデルを生成することを目指す基本方針であったため 2021 年末に出版された「強い将棋ソフトの創りかた」[8] において配布された教師データを流用することを試みた。結果的には棋力計測後これを持って WCSC32 モデルとした。

3. 棋力計測

WCSC32 前に作成された深層学習モデルは複数バージョン存在しており、各バージョン間の対戦勝率を持って有力候補とした。その後、数個の有力候補を floodgate でレーティング計測することにより代表するモデルを最終決定した。

選手権前には十分な検証期間がなかったため事後検証となったが RTX3060 搭載ノート PC で floodgate レート 4000 を大幅に超えるなど以前のモデルより大幅に強化されていることが確認された。

また、一手の探索も無く入力局面をニューラルネットでは評価した方策出力のみで対戦した場合も floodgate でレーティング計測が可能となった。以前はほぼ勝つことができなかったため有用と思える数値評価できなかったレベルであるので、これは飛躍的な進歩といえる。具体的にはレート 2300 程度であり、これは 15 年前のトップレート付近、言い換えるとアマチュア高段者クラスの棋力となる。また、Yss1000k 相手に 3 割強の勝率を得た。Yss1000k は 2015 年の WCSC で決勝入りしている YSS を探索ノード制限したものであると製作者のコメントを得ている。探索が全く行われていないことを考えると人智の外の大局観とも言える。

この探索無しでの対戦でもっとも大きなパートを占めるのが詰みの場面である。以前の研究では将棋の深層学習モデルでは詰みが扱えず詰将棋で有用性を示せないなどや、対戦時の最終盤で詰みを読めずに逆転をされるなど弱点とされていた。今回のモデルにおいては革新的な結果が得られているので一例を示す。

図 1 は floodgate 上の対局で 13 手詰め局面から最短 13 手で実際に詰めた局面である。敵玉の上部を完全に制圧した形であるため勝ちを認識するのは容易であるが、最短で詰めることが出来たのは快挙である。[9]

次に、図 2 も注目すべき詰み局面である。後手番で 5 手詰めであるが、初手が 9 五金とただで捨てる手から始まる。捨て駒をせずとも完勝の局面であるが、この金捨てで即詰みにするのが最も適切である。探索なくこの即詰みを行うことができるのは過去にない精度の深層学習モデルと考える。

以上のように探索無しの二番絞りが詰み局面まで滞りなく指し切る棋譜は多く生成された。しかしながら、全ての詰み局面を読み切ったわけではなく詰みや必至の局面で無

駄な手を指すようなことも見られた。勝敗に影響が無い場合特に問題が無いとも言える。また、今回は探索が無い場合何手詰めという概念すら AI 側には無いことも特徴である。つまり、このモデルのままさらに長手数の詰みを生成する可能性もあると言える。

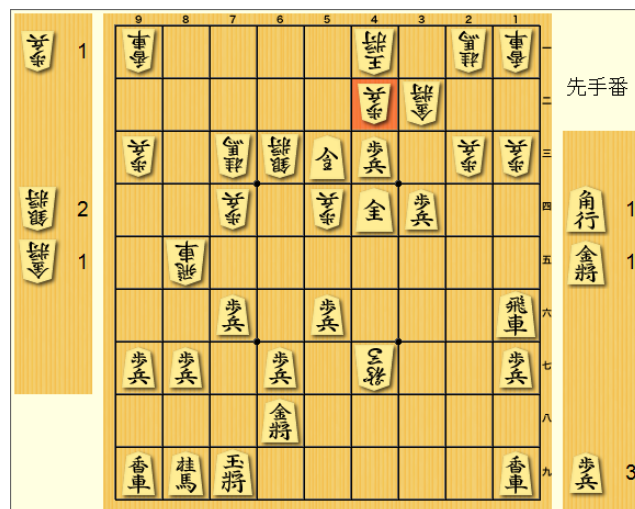


図 1 先手番 13 手詰め局面

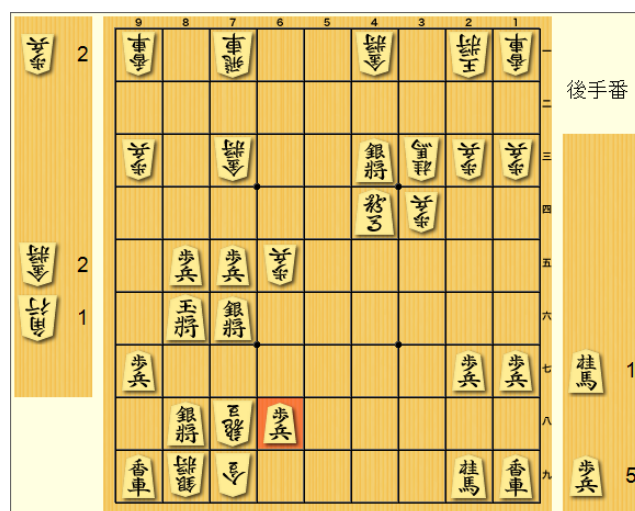


図 2 後手番 5 手詰め局面

4. 詰み局面計測

前節の結果を踏まえ深層学習モデルが詰み局面をとらえられているかどうかの計測を行った。計測には詰み局面集 [11] から同一シードのランダムサンプリングで 1000 局面を選び正確に詰ませられるかどうかで行った。対戦相手には Kristallweizen を用いた。比較対象には AobaZero の AlphaZero 追試最終モデルである w3459 および公開されている dlshogi の最新モデルである第 2 回世界将棋 AI 竜電戦エキシビジョンバージョンを用いた。計測の便宜上、詰み局面から 20 手以上続いた場合は詰めきれなかったとして

扱う。また、無駄合いの定義等により設定手数より短手数のものも僅かながら存在することを確認している。偶数手数で終了しているものは逆に詰まされているものである。

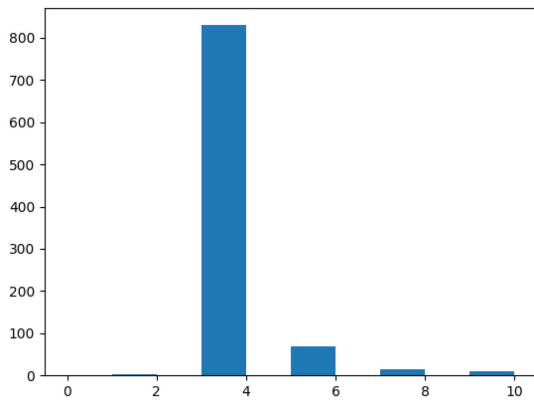


図3 二番絞り WCSC32 三手詰め局面の手数分布

図3に WCSC32 の二番絞りの三手詰め局面からの手数分布を示す。83%の局面において三手で詰ましていることが分かる。20手未満で勝利する割合は95%であった。

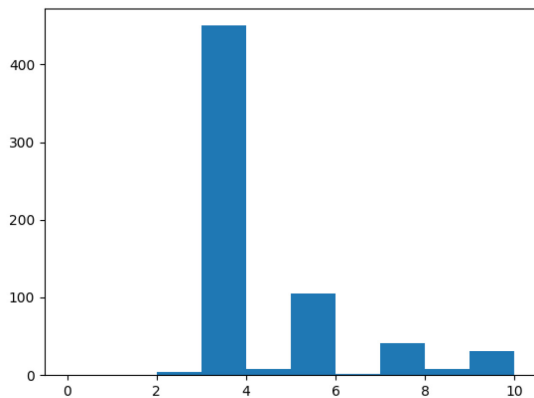


図4 AobaZero 三手詰め局面の手数分布

図4に AobaZero の三手詰め局面からの手数分布を示す。45%の局面において三手で詰ましている。20手未満での勝利割合は71%であった。偶数手で終了する局面も約5%存在した。他のモデルではこのような負け局面は1.2%程度であるため少々最終盤に甘いモデルと言える。

表1 詰み対局割合

	WCSC32	WCSC31	2020	DL	Aoba
三手	83.0	81.8	74.7	41.4	45.0
20未満	95.0	93.0	91.1	74.7	71.0

以上のような手続きで詰ませられた割合を表1に示す。

最も三手詰め局面を三手で詰まさせられないのは dlshogi となったが、これは恐らく ResNet15 ブロックと最も小さなニューラルネットワークで表現力に劣るためと思われる。対局時のレーティングの高さから比べると意外であるがニューラルネットワークの大きさが詰み局面では大きく影響すると推察される。

左から3つが三世代の二番絞りであるが新しいモデルが優秀であることが分かる。特に40ブロックのニューラルネットワークを採用した左の二つは秀逸である。しかしながら、2020年モデルは20ブロックの ResNet と AobaZero と同じサイズでありながら圧倒的な詰み性能を誇る。終盤力の差が定量化されたと考えてよいだろうか。

最後に WCSC32 の二番絞りの長手数の詰み性能を表2に示す。三手詰めは既出であるが、五手詰め局面も6割近く最短で詰めていることが分かる。七手詰めから急激に最短で詰められる割合が落ちるが、十一手詰め局面でも16%程度は詰めている。20手未満の勝利条件であれば大きく割合が増えるため恐らく勝勢時に甘い手が出て手数が伸びる傾向にあることが推察される。

表2 二番絞り WCSC32 の長手数詰み対局割合

	三手	五手	七手	九手	十一手
最短	83.0	59.4	31.6	21.4	16.5
20未満	95.0	83.3	67.1	53.8	40.5

5. おわりに

二番絞りプロジェクトという将棋の深層学習モデルを生成するプロジェクトを進めて2年強になる。今年生成されたモデルは従来言われていた詰み局面での弱点を大幅に克服し最大13手詰めまで実戦で確認した。また、捨て駒の手筋も理解していると考えられる棋譜も生成した。

また、ニューラルネットワークの大きさが詰み局面の特徴量を捉えるために必要ではないかと示唆される複数の結果を得た。

学習プロセスだけで一か月単位と非常に計算コストがかかるテーマであるが、今後も可能な限り比較実験を交えて深層学習モデルの差異を明らかにしていきたいと考えている。

参考文献

- [1] David Silver et al. “Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm”, <https://arxiv.org/abs/1712.01815>
- [2] Rémi Coulom. “Efficient selectivity and backup operators in monte-carlo tree search”. In International conference on computers and games, pp. 72–83. Springer, 2006.
- [3] dlshogi, <https://github.com/TadaoYamaoka/DeepLearningShogi>
- [4] 芝, “将棋の PV-MCTS に向けた深層学習モデルの最適化”, 2021-GI-45

- [5] 芝, “探索アルゴリズムに適した時間利用に関する研究”, 2021-GI-46
- [6] 千葉ら, “強化学習を用いて人間の直感を再現した詰将棋解答 AI の作成”, GPW2018
- [7] 芝, “コンピュータ将棋の新棋戦「電竜戦」について”, 2022-GI-GI48
- [8] 山岡, 加納, “強い将棋ソフトの創りかた”, マイナビ出版
- [9] nibanshibori_1p vs. coduck_pi2_600MHz_1c (2022-07-22 19:30), http://wdoor.c.u-tokyo.ac.jp/shogi/view/2022/07/22/wdoor+floodgate-300-10F+nibanshibori_1p+coduck_pi2_600MHz_1c+20220722193005.csa
- [10] coduck_pi2_600MHz_1c vs. nibanshibori_1p (2022-07-23 17:30), http://wdoor.c.u-tokyo.ac.jp/shogi/view/2022/07/23/wdoor+floodgate-300-10F+coduck_pi2_600MHz_1c+nibanshibori_1p+20220723173007.csa
- [11] やねうら王公式からクリスマスプレゼントに詰将棋 500 万問を謹呈, <https://yaneuraou.yaneu.com/2020/12/25/christmas-present/>
- [12] AobaZero, <http://www.yss-aya.com/aobazero/index.html>