

## 二番絞り

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気軽に始まった。ドワンゴ社主催の電王トーナメントが開催されていた時期は春に選手権、秋に電王トーナメントと年二回のイベントペースであったのだがこれを踏襲した形で秋にオンラインの電竜戦を開催するべく2020年オンライン大会終了直後から準備された。その後6,7,9,10月の4回の予行を経て11月の本大会へと繋がった。

二番絞りは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生まれた。同年5月開催予定であった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 `dlshogi` に学習させるのが基本コンセプトで絞りカスを絞る行為をなぞって二番絞りと命名された。学習手続きに対する理解も浅く全くと言っていいほど学習が進んでいなかつたため、3勝5敗1分で18チーム中16位という出来であった。

第二回予行演習には一月ほどの学習期間があったため上記教師データを存分に学習することができた。当時将棋の深層学習モデルで最も成功したものが `AobaZero` であることから `ResNet` の20ブロックを採用した。これは5ブロック、10ブロック、20ブロック、30ブロック、40ブロックとテストをして最も強かつたものを選定した結果である。将棋のモデルでは理想的なサイズがこの辺りであるとこの時点では断定した。その後予行では22チーム中10位という好成績を修めることができた。本予行では `AobaZero` がオンライン大会優勝の水匠に勝ち、また21回戦まで不具合なく戦い抜いたことが注目された。一方、二番絞りは水匠に対して千日手の引き分けをもぎ取っている。

第三回予行演習までにはまた一月近くあったこと、加えて NVIDIA の V100 が4枚搭載されたマシンを借りられたことから追加の強化学習実験ができた。追加したものはネットで公開されている `AobaZero` の教師データである。

選手権用に我々が用意した教師データはそれ以前の選手権で用いたエンジンの深さ 12 程度の探索局面データであるためこの指し手はそれほど強いものではないが、`AobaZero` の教師データは `floodgate` にてレート 3000 程度の探索と同程度の教師データである。前述の不具合などが気になるが実験的に導入したものである。この時点での目論見は `AobaZero` と同レベルの深層学習モデルができれば、探索エンジンの実行速度は `dlshogi` の方がやや速いので勝ち目があるかどうかといったところであった。

本実験では3つのモデルを作成した。全く初期状態から `AobaZero` 教師データを 1474 万棋譜全て学習させたもの、前述の第二回予行のモデルに `AobaZero` 教師データを 1474 万棋譜全て学習させたもの、前述の第二回予行のモデルに `AobaZero` 教師データを 1000 万番から 1474 万番棋譜まで学習させたものと3つ制作したが、最も累積教師データの多い二番目のモデルが対戦時に一番強かったのでこれを採用した。予行では不戦不具合が続いたために好成績は残すことができなかったが、`floodgate` に接続することでレート 3900 程度を記録することができた。一部のコンピュータ将棋関係者からは感嘆の声を頂いたと共に手ごたえを感じた。

第四回予行演習にはまた十分な期間があったため追加の学習を考えることになった。しかしながら、手持ちの教師データは使い尽くしたためにこの段階で強化学習へとシフトすることになった。前述のモデル

で自己対局を行い、局面データを作成した。予行1週間前に教師データ生成を止め学習プロセスへと移行した。生成できた学習局面は1億局面程度である。第四回予行演習ではA級リーグとB級リーグを独立で行う運営のテストであったためB級入りを志願し結果B級7位となった。当然強豪ぞろいの本戦でA級入りは不可能との見込みである。

最後に電竜戦本戦に向けた準備として上記強化学習をもう一段階行った。準備できた教師データの局面数は1億2千万局面である。第四回予行時にも同様のテストをしていたのだが教師データを1エポック学習したものと2エポック学習したものと比較していた。前回予行時は1エポックを採用したが本戦では対戦結果から2エポックのものを採用した。この行為は単に2択の紅白戦である。

この段階において学習時に不思議な現象が起きていることを確認している。指し手の一致率が7割近くなっているのである。生成した教師データが偏っているか不足しているのに2エポック回したためかと思われたが対戦して勝った方を本戦に投入することにした。当時は極度な過学習との認識であった。

結果電竜戦初日に予選3位となり本戦A級入りとなった。とりわけ、予選の最終戦においてはV100の4枚を用いる二番絞りに対してA100の8枚を用いるdlshogi本家との対戦となり、同エンジンで4倍～8倍のマシンパワー差から敗北を覚悟していたのであるが中盤から逆転して勝利を修めることになった。電竜戦予選ベストバウトと言って良い。完全に予想外の出来である。

しかしながら、2021年5月開催の第31回世界コンピュータ将棋選手権においては決勝進出8チームのうち7チームが $\alpha\beta$ 探索を用いており純粋なMCTS型エンジンはPALのみであった。そのPALは決勝で2位となり、優勝は $\alpha\beta$ 探索型のelmoであった。これにより探索アルゴリズムの優位は簡単には語り難い状況にある。二番絞りはこの年上記の複数サイズの深層学習モデルを並行開発しており、最終的に最も大きな40ブロックモデルを投入したが予選落ちとなっている。

そういう経緯で、本年の第32回世界コンピュータ将棋選手権には前年モデルに幾つかの追加学習を施したもの投入している。手元の計測では大きな向上を確認していなかったため予選一位突破は想定外であった。前年モデルも計算機リソース次第で上位が狙えたのかかもしれないが今となっては確認不能である。選手権レベルの対戦や検討にはコストが大きいため検証は行っていないが、悪手と言えるものが一手のみであり控えめに評価しても静的な意味では最高精度の評価関数であったことは間違いないだろう。

#### 参考：

芝、「将棋のPV-MCTSに向けた深層学習モデルの最適化」、第45回ゲーム情報学研究会

芝、「探索アルゴリズムに適した時間利用に関する研究」、第46回ゲーム情報学研究会

第32回世界コンピュータ将棋選手権、<https://bleu48.hatenablog.com/entry/2022/05/06/145915>