

# 「アストラ将棋」アピール文書

令和5年5月

恒岡 正年

## 概要：

アストラ将棋 は、山岡氏著「強い将棋の造りかた」という本の内容に沿って dlshogi をベースに作成されたコンピュータ将棋エンジンです。

自前のモデル作成を行い、探索パラメータを本番で使用するハードと持ち時間に合わせて調整しました。

UctSearch.cpp と usi.cpp を改造しましたが、それほど大きな改造ではありません。

学習時に使う data\_loader.py や特に train.py は大きく改造しています。

定跡を作成しましたが考時間の短縮と後手番での特定の局面を避けるのが目的で規模は小さいです。ベンチマークの相手には主に水匠 5 を使用しました。

ベンチマークの開始局面には、たややん氏・山岡氏の公開されたもの、自前で用意した物等幅広く使用しました。ベンチマークは棋譜生成を兼ねており、より多くの学習局面を生成するためです。

## 開発動機：

以前に電通大で開催された囲碁セミナー(松原氏,山下氏が講師をされました)に参加して囲碁 AI の開発を始めたのですが、アルファ碁の登場で桁違いの物量を生かした開発が主流になり撤退。

「強い将棋の造りかた」の本に触発されて将棋 AI の開発に参入。

ディープラーニング(DL)の特に学習部に興味があり、DL の勉強を兼ねて自前の model を作ろうと思いました。そのため、model の学習に最も注力しています。

## 独自に工夫した点：

独自構造の model の模索・作成、学習方法(マクロバッチ実装、他)、生成モデルのチューニング、探索パラメータの調整方法等。

## 開発過程：

2022年1月から「強い将棋の造りかた」(\*1)に沿って pytorch の勉強を兼ねて勉強開始。

2022年5月頃まで色々な構造の model の生成と共に、python-dlshogi2 を改造して floodgate へ投入。GPU に GTX1080ti を使って R3400~R3500 位だったと記憶しています。

2022年5月以降「強い将棋の造りかた」の第7章「GCT 電竜を超える強い将棋 AI を創る」へ移行。

2023年2月後半まで、色々な model の実験・学習に注力、及び棋譜生成を続ける。

マクロバッチを実装し、疑似的に GPU のメモリ量がなくなったのと同じような結果が得られる様にした。(Batch Normalization 部分の挙動が異なるので全く同じにはならない)

最終的に 110 番目のネットワークの model を今回使用。学習途中で中断した物が殆どで、収束するまで学習させたのは 20 個程度。使用する model はこの時点で決定。

2023年4月末まで、探索パラメータの調整(2カ月)及び定跡生成(手作業で2週間ほど)

ネットワーク構造等に関しましては、既提出済のアピール文書(4月29日版)をご参照ください。

## 実験結果：

有意差を理論的に示せるほどの比較実験を行う計算資源が無かったのと、新しい事を多く試したかったので主観的な判断で変更項目の採用・不採用を決めました。

これだけでは面白くないので、パラメータの調整の手法について書きます。

6656po(RTX3090 で 0.5 秒程度)でベンチを取って3つの有力なパラメータセットが得られたとします。3つのベンチの勝率の差は1~2%程度の差でした。これらのパラメータセットを15s+F1.2s(探索量が2倍以上多い)で評価すると勝率の差が8%程度出るケースを確認しています。パラメータセットによって探索時間を大きくした時の伸びしろが異なるという結果でした。

そのためパラメータの調整を行った2カ月間は、6656po (optuna を利用、パラメータ3~4個単位で調整) => 15s+F1.2s(手作業のベンチで調整) => 20s+F5.0s(同左) と3段階で探索量を増やしながらか評価を行い、パラメータセットも最良の物だけではなく2nd,3rd...bestの物も候補として評価し、段階的に候補を絞っていきました。

パラメータセットを2つに絞ったあと、さらにSoftmax\_temperature と C\_base, C\_base\_root(設定の粒度が小さい)の調整を再度行い、最終的に1つを選びました。(もう一つは予備へ)

この2カ月の探索パラメータの調整でR100程度強くできましたと思います。(2月末時点で使用していた手動調整したパラメータの出来が不十分だったとも言えます。)

3月中頃にfloodgateに流した結果では、13056po(RTX3090で1秒弱の探索時間)の探索量でR3800程度ありました。本番のハード(RTX4090)と900s+F5sの持ち時間及びponder有りの条件では32倍強の探索量と推定。探索量が2倍になるとレーティングがR+80増加すると仮定するとR4200(~4300)程度になったと予想しました。(TopクラスはR4800~R5000overでしょうか?)

## 追試可能か：

同一のネットワーク構造を用意し「強い将棋の造りかた」の内容に沿って学習し、パラメータの調整を行えば、同様のモデルが得られると思います。

ただ、短時間で学習する手法(RTX3090で3週間弱)とか、学習後のモデルの調整、optunaを使ったパラメータの効率的な調整は経験に依存する部分も多いと思います。

## 最後に：

dlshogiの成果を利用した上での参加ではありましたが、1台のPCで参加し決勝戦に出場できたのは幸運だったと思います。今回のモデルは定跡を抜けてから80手程度までは上位ソフトと互角に近い状態を維持できましたが、90手目前後以降で読みの精度が低いことを確認しました。これは、KL情報量が多い範囲の局面を意図的に多く学習させ終盤に近い局面を減らした事、学習時に重複局面を除去しなかった事(序盤の局面が増える)によると推定しています。

モデルの出来は満足できる物ではありませんでしたが、2月末以降残り時間を探索パラメータの調整に費やしたのは結果的に良かったと思います。

今回、モデルの学習方法や探索パラメータの調整方法の知見が得られました。反省点を生かした次のモデルではすでに終盤力の向上を確認できております。(序盤の精度が低下している可能性はある。)

## 謝辞：

将棋AIに関わるきっかけとなった本を出版された山岡様、加納様には深く御礼申し上げます。特に加納様にはSNS上で多くのアドバイスをいただきました。

また、WCSC開催に御尽力されている関係者・スポンサー様にお礼を申し上げます。

## 参考文献：

\*1: 山岡忠夫, 加納邦彦. 強い将棋ソフトの創りかた Python で実装するディープラーニング将棋 AI. マイナビ出版, 2021.