

人間の感覚と異なる囲碁 AI 特有の手の抽出

馬志偉¹ 伊藤毅志¹

概要 : AlphaGo の登場以降, 囲碁 AI は人間を十分に凌駕するレベルになっている. 多くのプロ棋士が囲碁 AI を用いて訓練するようになってきているが, 人間の感覚と囲碁 AI の思考には大きな違いがあることも指摘されており, それが囲碁 AI を用いた学習の妨げになっている. 本論文では, プロ棋士の棋譜を対象に, 最新の囲碁 AI から見ると良い手であるが, 人間的に見ると打ちにくい手を効率よく見つける手法について提案した.

キーワード : 囲碁 AI, 人間の感覚

The Extraction of Go AI-Specific Moves that Differ From Human Sences

ZHIWEI MA^{†1} TAKESHI ITO^{†1}

Abstract: Since the appearance of AlphaGo, the Go AI has reached a level far beyond that of humans. However, it has been pointed out that there is a big difference between the human senses and the moves of the Go AI, which hinders learning with the Go AI. In this paper, we propose a method for efficiently finding the good moves from the viewpoint of the Go AI but difficult to select them from the human senses, from professional Go players' game records.

Keywords: Go AI, Human Sence

1. はじめに

人工知能領域においては, これまで数え切れないほどのゲーム AI に関する研究開発がおこなわれてきた. 将棋, 囲碁などのボードゲームで人間のプロ棋士を超えるレベルに達することは一つの大きな目標であった.

囲碁の分野では, 2006 年頃から, モンテカルロ木探索の出現によって囲碁 AI は強くなり始め, 2008 年には, MoGo が US Go Congress でプロ棋士の金明完八段に 9 子局で勝利を収めた. その後も, 順調にハンデを減らし, 2015 年の第 3 回電聖戦では趙治勲九段を相手に, DoI Baram が 4 子で勝利し, CrazyStone が 3 子で敗れた. 2016 年に深層学習を用いた AlphaGo が登場すると, いきなり互先で人間のプロ棋士を破るレベルに到達し, 2017 年には世界トップの柯潔九段との対戦で 3 連勝し, 人間を凌駕するレベルになったことを示した[1][2].

囲碁 AI は, その後プロ棋士を中心に研究パートナーとして使われるようになり, AI の開発の目的の一つとして, 人間の学習支援としての側面も重要な意味を持つようになってきている. しかし, 人間の思考と囲碁 AI の思考には大きな違いがあることも指摘されており [3], それが囲碁 AI を用いた学習の妨げになっている.

そこで, 人間の感覚的には打ちにくく, 囲碁 AI から見ると良い手を抽出することを考える. もし, そのような手

を効率よく抽出することができれば, 囲碁 AI 特有の手だけを効率よく学習できるのではないかと考えられる.

本論文では, プロ棋士の棋譜を対象に, 最新の囲碁 AI から見ると良い手であるが, 人間的に見ると打ちにくい手を効率よく見つける手法について考察する.

2. 関連研究

AlphaGo Zero の論文を見ると, AlphaGo と AlphaGo Zero のプログラムの特徴を比較している [4].

図 1 と図 2 は, AlphaGo Zero の論文からの引用したものである. 教師あり学習である AlphaGo は赤線, 教師なし学習である AlphaGo Zero は青線である.

図 1 は AlphaGo と AlphaGo Zero のレーティングの比較で, 学習回数が増えるにしたがって, 教師なし学習のほうが教師あり学習を上回るレーティングになっていることが示されている.

一方, 図 2 は教師あり学習と教師なし学習のプロ棋士の手の Policy Network による手の予測確率を表したもので, 教師あり学習の方が, 教師なし学習よりもプロ棋士の手に対して高い予測率であることを示している.

¹ 電気通信大学
The University of Electro-Communications

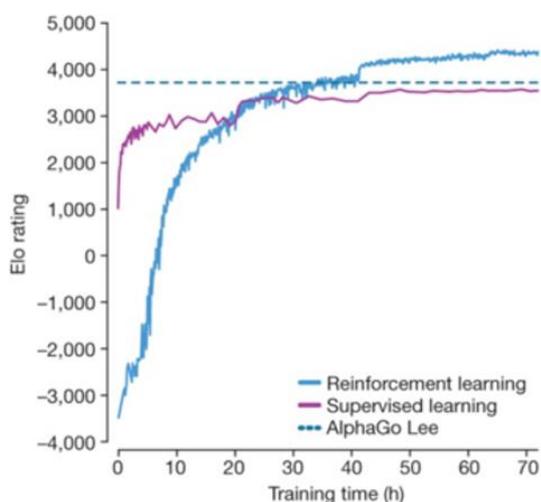


図 1 レーティングの比較

Figure 1 Comparison of ratings.

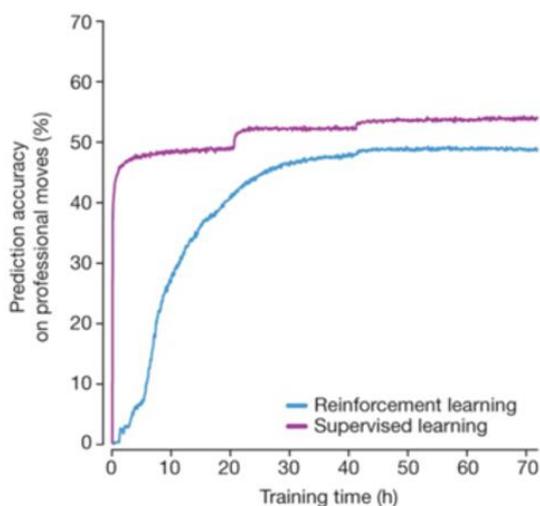


図 2 プロ棋士の着手予測確率の比較

Figure 2 Comparison of prediction accuracy on professional moves.

これらの結果をまとめると、教師あり学習では、プロ棋士の手の予測の精度は高いが、教師なし学習の方がレーティングとしては高くなると言える。

本研究では、AlphaGo と AlphaGo Zero を模倣した Leela と Leela Zero を用いる[5][6]。上記の結果から、Leela (教師あり学習) と Leela Zero (教師なし学習) も同様の傾向があることが予想される。

3. 使用する囲碁 AI

3.1 Leela と Leela Zero

本研究では、人間による知識やプロ棋士の対局棋譜を教師データとして教師あり学習を行っている Leela と人間の知識を使わずにモンテカルロ木探索と強化学習を組み合わせ

せて教師なし学習を行う Leela Zero の二つの囲碁 AI を利用する。

これらのプログラムは、ベルギーのプログラマーである Gian-Carlo Pascutto が開発したもので、AlphaGo, AlphaGo Zero の手法を真似たものとされている。Leela は 2017 年に深層学習を用いた AlphaGo 型のプログラムで、人間のトッププレイヤーの棋譜を教師データとした Policy Network と深層強化学習を行う Value Network を備えたプログラムである。一方、Leela Zero は囲碁の基本的なルールだけが与えられ、AlphaGo Zero の手法を模倣して、モンテカルロ木探索を用いた自己対戦の結果から深層強化学習によって学習したプログラムである。

3.2 AlphaGo Zero 論文との比較

強化学習の初期の段階では、Leela Zero のアルゴリズムとプログラムが正しく機能するかどうかの確認のため、検証を高速化するために、AlphaGo Zero 論文に記載されているいくつかのパラメータを調整していた。

ニューラルネットワークの構造に関しては、両者とも 256 種類のフィルターで 40 ブロックの CNN (畳み込みニューラルネットワーク) が使われている。

強化学習のための自己対局の局数について、両者とも 50 万局の結果が使われている (Leela Zero では直近の 50 万局)。また、AlphaGo Zero が論文で自己訓練や正式な対局の際には、一手ごとに 1600 回のシミュレーションが行われた。Leela Zero では当初は AlphaGo Zero と同様に 1600 回のシミュレーションだったがのちに、3200 回のシミュレーションに変更された。

作者は元の AlphaGo Zero 論文に欠陥があることを発見し、論文の最初の入力は 17 チャネルだったが、白番が基盤の端を認識しやすくするために、Leela Zero で 18 に修正されている。

4. 人間の感覚とは異なる好手の抽出手法

4.1 抽出方針

Leela の Policy Network を用いるとプロ棋士の着手をかなり正確に予測できることは図 1 から推察できる。そのように考えると、Policy Network において選択確率の低い手は人間にとって直感的に打ちにくい手になるのではないかと考える。

プロ棋士の棋譜の中から、Leela Zero の勝率計算を行って、大きく勝率が下がっている局面を調べ、その局面における Leela Zero の最善手を抽出し、その手が Leela の Policy Network で確率の低い手であれば、人間の感覚とは異なる好手になっているのではないかと考え、この手法で手の抽出を試みることにする。

4.2 予備調査

まずは、Leela が Leela Zero よりもプロ棋士の手を予測す

ることができるのかを調べるために、それぞれの Policy Network を用いて、2020 年度の第 68 回 NHK 杯テレビ囲碁トーナメントの全棋譜に対して、プロ棋士の着手との一致率を調べた。その結果、図 3 のように、Leela の方が高い一致率を示した。

また、Leela Zero でプレイアウト回数を変えて最善手の変化の様子を調べた[8]。実環境で調べたところ、6400 回以上のプレイアウトで最善手の変化率が減少することがわかった。Leela Zero で最善手を求めるためには、6400 回以上のプレイアウトが望ましいことが示唆された。実際に一試合における棋譜の分析にかかる時間を考慮すると、Leela Zero のプレイアウト数を 10,000 回まで増やすことが限界であるとわかったので、本研究では Leela Zero のプレイアウト回数を 10,000 回とすることにした。

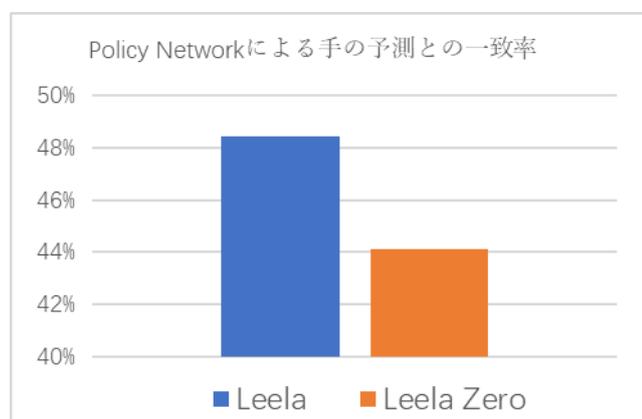


図 3 Leela と Leela Zero の手の一致率

Figure 3 The degree to which Leela and Leela Zero match

4.3 抽出手法

具体的には、以下の手順で人間には打ちにくい囲碁 AI 特有の好手を抽出する。

<手順 1> Leela Zero を用いて、プロ棋士の実践譜のすべての着手を調べ、勝率において一定以上低下する手に着目し、その局面で Leela Zero が選ぶ最善手を抽出する。

<手順 2> Leela Zero が選ぶ最善手の中で、Leela の Policy Network による評価で着手確率が一定の値より低い手を抽出する。

5. 実験

5.1 抽出実験と抽出基準

4.3 の抽出手順で、第 68 回 NHK 杯テレビ囲碁トーナメントの全棋譜 (49 局, 10582 局面) を対象に抽出を試みた[7]。勝率を下げた手として、10%以上勝率を下げた手を調べたところ、640 局面、全体の約 6%程度の手が見つかった。分析対象とするのに適当な数であると考えたので、とりあえず、10%で分析を行うことにした。さらに、それらの局

面を Leela の Policy で AI の示す手を評価させ、図 4 のように選択確率の頻度を調査した。その結果、選択確率 1%未満のものが、全体の 3 割程度を占めていることがわかった。

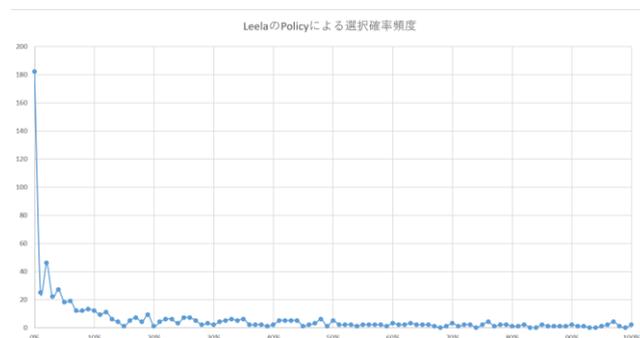


図 4 選択確率の頻度

Figure 4 Frequency of selection probability.

そこで、本研究では、「実際にプロ棋士が選択した手よりも 10%以上評価値が高く、Policy Network で 1%未満の手」を、人間の感覚的には打ちにくい好手であるとみなして、抽出するものとする。

5.2 抽出例

ここで、上述の基準で抽出された一つの局面を紹介する。

図 5 は第 68 回 NHK 杯テレビ囲碁トーナメント 3 回戦第 3 局、一力遼 (天元) 対大西竜平 (七段) の第 99 手目 (黒番) である。

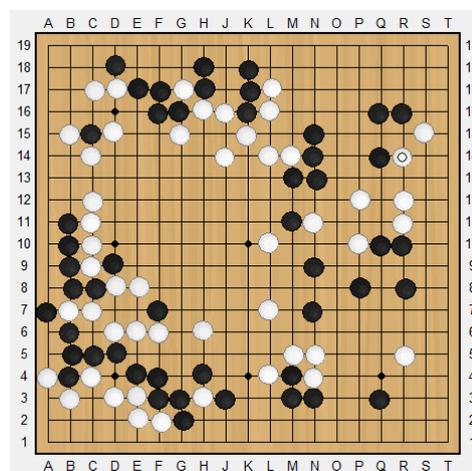


図 5 抽出された局面

Figure 5 Extracted aspects.

図 6 と図 7 のように、この局面におけるプロ棋士の着手は Q11 で、AI の最善手は C2 であった。Leela Zero によれば、プロ棋士の着手の勝率は約 54.1%、AI の着手の勝率は 64.3%であり、プロ棋士の選択した着手は勝率を 10%以上下げている。

一方、Leela を用いた着手選択確率は、Q11 は 21.77%であるのに対して、C2 の選択確率は、0.10%であった。したがって、C2 は人間の感覚的には着手しにくい好手である可能性が高い。

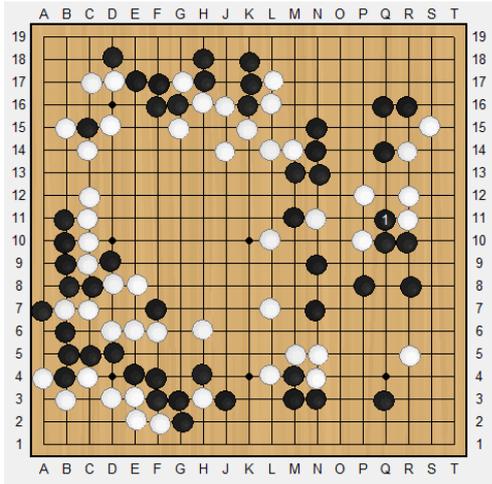


図 6 プロ棋士の着手
Figure 6 Professional move.

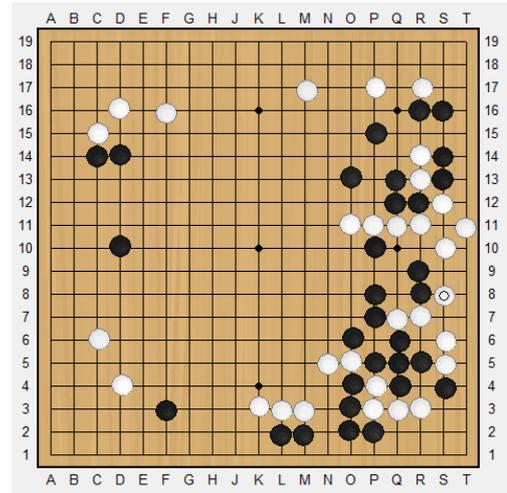


図 8 抽出された局面
Figure 8 Extracted aspects.

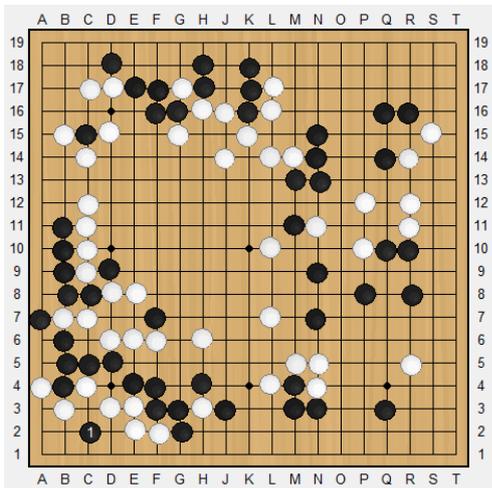


図 7 囲碁 AI の着手
Figure 7 Go AI move.

評価実験において比較のために、逆に人間の感覚的には打ちやすいであろう局面の抽出も試みた。図 8 は第 68 回 NHK 杯テレビ囲碁トーナメント準々決勝第 1 局、山下敬吾（九段）対高尾紳路（九段）の第 65 手目（黒番）である。

図 9 と図 10 のように、この局面におけるプロ棋士の着手は S9 で、AI の最善手は S3 であった。Leela Zero によれば、プロ棋士の着手の勝率は約 36.08%，AI の着手の勝率は 51.60% であり、プロ棋士の選択した着手は勝率を 10% 以上下げている。

一方、Leela を用いた着手選択確率は、S9 は 3.10% であるのに対して、S3 の選択確率は、84.60% であった。したがって、S3 は人間の感覚的には着手しやすい好手である可能性が高い。この局面において、プロ棋士は感覚的にも打ちにくい悪手を打った可能性がある。

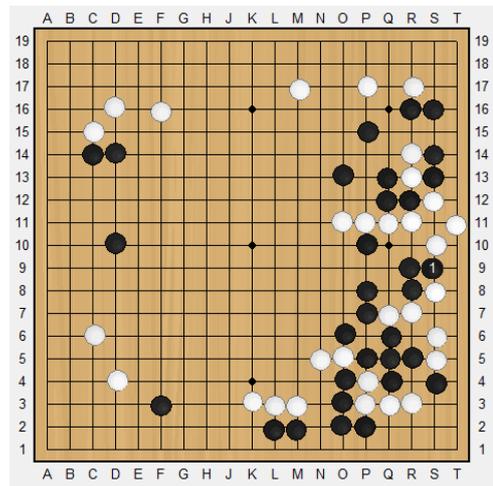


図 9 プロ棋士の着手
Figure 9 Professional move.

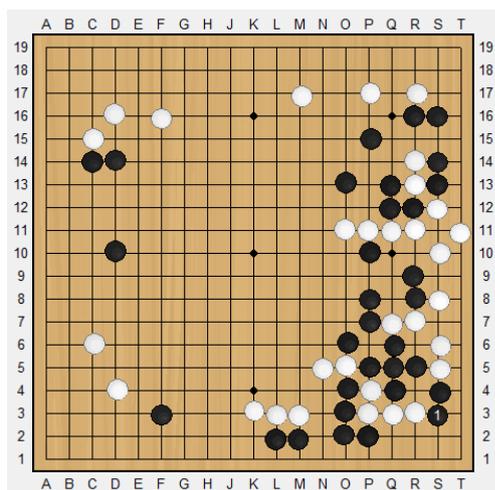


図 10 囲碁 AI の着手
Figure 10 Go AI move.

6. 評価実験

6.1 実験目的

提案手法で抽出した手が本当に人間にとって感覚的に着手しにくい手であるのかを確かめるため、評価実験を行う予定。これによって、本手法の有効性を示していきたい。

6.2 実験対象

囲碁の熟達者、及びプロ棋士のそれぞれ 5 人ずつ 10 名程度を本評価実験の対象とする、

6.3 実験方法

本実験では、被験者に Google Form に用意した囲碁対局の局面問題を回答してさせることと、それによる特定の着手に対しての主観評価を行わせる。

<Step 1>図 11 のように、用意した局面を被験者にランダムに提示して、提示した局面において自分なら直感的にどの手を選ぶのかを回答させる。

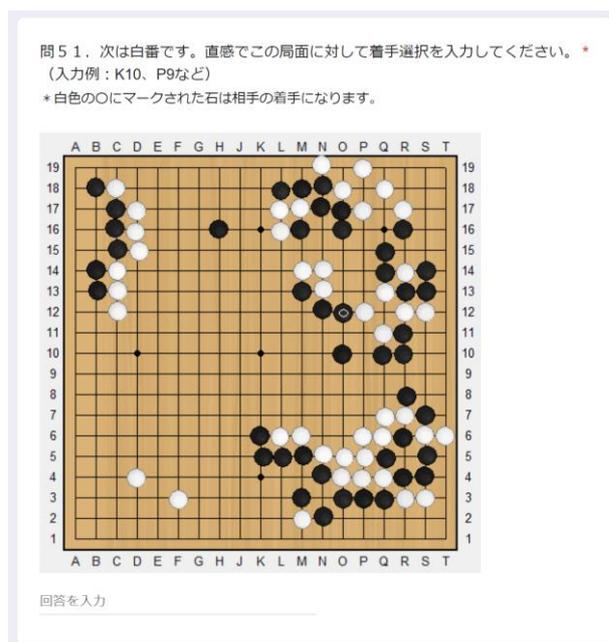


図 11 評価実験で提示する質問の例

Figure 11 An Example of questions to present in the evaluation experiment.

<Step 2>図 12 のように、Leela Zero が示した最善手を提示して、それが感覚的に見えにくい手かどうかを 5 段階のリッカート尺度（打ちやすい、比較的打ちやすい、どちらとも言えない、比較的打ちにくい、打ちにくい）で被験者に回答させる。

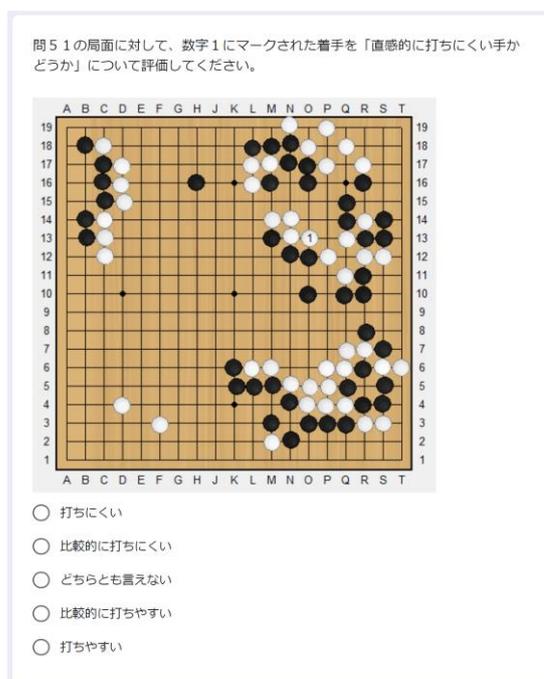


図 12 Leela Zero の手を直感的に評価させる質問例

Figure 12 An example of questions that make Leela Zero's moves intuitive to evaluate.

6.4 実験に使用する問題

着手選択確率とその時の AI の最善手の関係を現した表が表 1 である。

表 1 抽出した局面の構成

Table 1 Structure of extracted aspects.

	着手選択確率	AI の最善手
抽出候補とする局面	1%未満	207
比較のため用意した局面	1~20%	238
	20~40%	80
	40~60%	57
	60~80%	35
	80%以上	23

本評価実験では、比較のため、Leela Zero が示した最善手の中に、Leela の Policy による着手選択確率が提案手法で抽出した手よりも高い手を 20%刻みで、この表から、各 10 局面を用意した。1%未満のもの 1~20%、20~40%、40~60%、60~80%、80%以上の合計 60 局面を用いることにする。

7. おわりに

本論文では、囲碁 AI を用いた学習に対して、本来人間の思考と違って、感覚的には打ちにくく、囲碁 AI 特有の手をより効率よく学習するため、プロ棋士の棋譜を対象に、最新の囲碁 AI から見ると良い手であるが、人間的に見ると打ちにくい手を効率よく見つける手法を提案した。

本発表までに、評価実験を進行させ、囲碁 AI 特有の手を抽出する手法の妥当性を議論していきたい。さらに、抽出された手を用いた学習支援の方法についても考察していきたい。

参考文献

- [1] 松原仁：“コンピュータ囲碁の進歩”、日本ロボット学会誌、Vol.35,no.3,pp.191-194 (2017).
- [2] 伊藤毅志、松原仁：“AlphaGo の技術と対戦”、人工知能学会誌、31(3),pp.441-443 (2016).
- [3] 伊藤毅志、高橋克吉、猪爪歩、加藤英樹、村松正和、松原仁：“人間とコンピュータの思考の違い～囲碁の次の一手問題による考察～”、ゲームプログラミングワークショップ 2012、p.9-16 (2012).
- [4] Silver, D., Schrittwieser, J., Simonyan, K. et al.: “Mastering the game of Go without human knowledge”. Nature 550, 354–359 (2017).
- [5] Leela
<https://sjeng.org/leela.html> (2017-10-24 更新)
- [6] Leela Zero
<https://github.com/leela-zero/leela-zero> (2021-8-25 更新)
- [7] 棋譜ダウンロードサイト
<http://goyuan.aquifer.jp/materials/download/index.html> (2022-6-16 更新)
- [8] 竹内聖悟、プレイアウト数増加に伴うモンテカルロ木探索の振舞い、情報処理学会、ゲームプログラミングワークショップ