

Title	囲碁における人間の着手を模倣するにはKataGoのHumanSLだけで十分か？
Author(s)	シュエ, ジュウシュエン; 池田, 心
Citation	ゲームプログラミングワークショップ2024論文集, 2024: 80-87
Issue Date	2024-11-15
Type	Conference Paper
Text version	publisher
URL	http://hdl.handle.net/10119/20012
Rights	社団法人 情報処理学会, シュエ, ジュウシュエン, 池田 心, ゲームプログラミングワークショップ2024論文集, 2024, 80-87.ここに掲載した著作物の利用に関する注意 本著作物の著作権は情報処理学会に帰属します。本著作物は著作権者である情報処理学会の許可のもとに掲載するものです。ご利用に当たっては「著作権法」ならびに「情報処理学会倫理綱領」に従うをお願いいたします。Notice for the use of this material: The copyright of this material is retained by the Information Processing Society of Japan (IPSJ). This material is published on this web site with the agreement of the author (s) and the IPSJ. Please be complied with Copyright Law of Japan and the Code of Ethics of the IPSJ if any users wish to reproduce, make derivative work, distribute or make available to the public any part or whole thereof. All Rights Reserved, Copyright (C) Information Processing Society of Japan.
Description	第29回ゲームプログラミングワークショップ (GPW-24), 箱根セミナーハウス, 2024年11月15日-17日

囲碁における人間の着手を模倣するには KataGoのHumanSLだけで十分か？

シュエ ジュウシュエン^{1,a)} 池田 心^{1,b)}

概要：人間プレイヤーのゲームプレイを模倣することは、人間プレイヤーを楽しませたり、教えたりする際の重要な要素の1つである。本稿では、囲碁を対象とし、教師あり学習（Supervised Learning, SL）に基づき、大量の人間プレイヤーの棋譜を用いて学習されたKataGoのHumanSLモデルについて調べる。具体的には、(1) 人間らしいゲームプレイを実現するために、教師あり学習モデルが予測した着手の選択確率分布をどのように利用すべきかを調査し、(2) HumanSLの人間らしさを評価し、(3) HumanSLの人間らしさをさらに改善する方法を提案する。(1)については、選択確率最大の着手を選ぶ方法や選択確率分布通りにサンプリングする方法がよく使われているが、我々の結果では、中間的なバージョンであるsoftmaxのほうが、望ましい棋力を達成できることが示される。(2)については、着手一致率、尤度、直前手との距離といった指標を用いる。(3)については、直前手との距離による補正を選択確率を入れる。その結果、HumanSLの人間らしさが改善されていることを確認した。

キーワード：囲碁, KataGo HumanSL, 人間らしさ, 着手の模倣

Is KataGo HumanSL All You Need for Imitating Human Players' Moves in Go?

CHU-HSUAN HSUEH^{1,a)} KOKOLO IKEDA^{1,b)}

Abstract: Imitating human players' gameplay is one of the important factors when entertaining or teaching human players. In this paper, we target the game of Go and investigate KataGo's HumanSL model, which was trained using plenty of human players' game records based on supervised learning (SL). More specifically, we (1) investigate how the probability distributions over legal moves predicted by SL models should be used to achieve human-like gameplay, (2) evaluate the human-likeness of HumanSL, and (3) propose a method to further improve the human-likeness of HumanSL. Regarding (1), researchers often select the move with the highest selection probability or sample according to the probability distribution, while we find that an intermediate version, softmax, can better achieve the desired playing strength. Regarding (2), we employ the metrics of move-matching accuracy, likelihood, and distance to the previous move. Regarding (3), we adjust the selection probability of a move based on its distance to the previous move. We confirm that the human-likeness of HumanSL is improved.

Keywords: Go, KataGo HumanSL, Human-likeness, Imitation of moves

1. はじめに

人間プレイヤーの挙動を再現することは、人間プレイヤーを楽しませたり教えたりするために重要である。このような

AIプレイヤーは例えば練習用の対戦相手やコンテンツ生成時のテストプレイヤーとして使える。将来的に再現したい挙動としては、戦術や好みなどの個性、着手以外の思考時間や投了のタイミングなどもあるが、本稿では「強さごとの平均的なプレイヤーらしい着手をさせること」を目指す。

人間らしい着手を獲得する有力な方法の1つとして、人間のプレイ記録を学習するものがある。ボードゲームの例でいうと、人間プレイヤーの棋譜を用いて教師あり学習を行

¹ 北陸先端科学技術大学院大学
Japan Advanced Institute of Science and Technology (JAIST), Nomi, Ishikawa 923-1292 Japan

^{a)} hsuehch@jaist.ac.jp

^{b)} kokolo@jaist.ac.jp

う方法がよく使われる。少し前までは強い AI プレイヤを作るために上級者の棋譜を学習する研究が多かった [1], [2] が、近年では初中級者の棋譜を学習する研究もますます注目されている [3], [4]。

囲碁において KataGo という強い AI プレイヤが存在する。自己対戦により学習したニューラルネットワークとは別に、2024 年 7 月に発表された v15 版 [5] では、大量の人間プレイヤの棋譜を学習したニューラルネットワーク（以下 HumanSL で略す）が公開された。しかし、HumanSL が予測した着手の確率分布を用いてどのくらい人間らしくプレイできるかは知られておらず、また対戦相手にしたい場合にそれをどう使うのかについても合意はない。

本研究の目的は主に 2 つある。(1) 着手確率分布の使い方を確立する。(2) KataGo の HumanSL が予測した着手の確率分布を用いてどのくらい人間らしくプレイできるかを調べ、その結果をもとに人間らしさを向上させる方法を提案する。2 つ目の目的に関連して、人間らしさを評価する指標も提案する。

着手確率分布の使い方について、既存研究では「選択確率最大の着手を選ぶ」という方法や、「選択確率通りに着手を選ぶ」という方法がよく見られる [2], [3], [5]。しかし、これらの方法により選ばれる着手は人間らしく見えるかについては明らかになっていない。

人間らしさを評価する方法について、既存研究では一致率（つまり選択確率が最も高い着手と、棋譜で実際に打たれた着手が同じである割合）がよく使われる [2], [3]。しかし、例えば着手が 2 つあり、モデル A での選択確率が 90%・10%、モデル B での選択確率が 51%・49%という場合に、この 2 つのモデルの優劣を区別できないという課題がある。そのため、より多くの面から人間らしさを評価すべきだと考える。

2. 関連研究

McIlroy-Young らは Maia という、チェスの人間初中級者の棋譜を学習したモデルを提案した [3]。初中級者の棋譜を棋力によって分け、各棋力帯に 1 つのモデルを教師あり学習で作った。各棋力帯の人間プレイヤの着手を予測するためには、該当棋力帯あるいはその付近の棋譜を学習したモデルとの一致率が最も高いと報告された。着手確率分布の使い方については明示されないが、評価指標に一致率が用いられたことと、公開されたコードでの説明^{*1} から、「選択確率最大の着手を選ぶ」という方法だったと推測する。

小川らは Maia のような教師あり学習のモデルと、AlphaZero のような強化学習のモデルをブレンドするという手法を提案し、人間プレイヤの着手との一致率を向上させたことを示した [4]。この 2 つのモデルをブレンドする動機と

して、それぞれのモデルが、もう一方のモデルでうまく予測できない状況を得意していることが挙げられている。具体的には、Maia のような教師あり学習のモデルには読みが必要な局面が苦手であることにに対して、AlphaZero のような強化学習のモデルにはうまく予測できるという点と、AlphaZero のようなモデルには人間らしくない着手が多いことにに対して、Maia のようなモデルには人間プレイヤの直感に比較的近い予測ができるという点である。また、初中級者らしさを評価するもう 1 つの指標として、小川らは尤度（人間プレイヤの着手の選択確率の幾何平均）を用いた。具体的には、モデルが予測した、局面 s で着手 a を打つ確率を $P(s, a)$ とし、テスト棋譜にある局面と人間プレイヤがプレイした着手の対の集合 $\{(s_1, a_1), (s_2, a_2), \dots, (s_N, a_N)\}$ での尤度は、 $(\prod_{i=1}^N P(s_i, a_i))^{(1/N)}$ である。尤度を用いるメリットを 3.2 節で説明する。

KataGo の HumanSL は、複数のモデルを作る Maia と異なり、1 つのモデルでさまざまな棋力の着手を予測する [5]。黒番と白番の棋力や日付などの情報を、現局面の情報とともに入力として渡し、モデルが次の一手の確率分布と現局面の勝率などを予測する。入力に直前数手の局面を使わない点と、シチュウや呼吸点などの囲碁の専門知識による情報を使う点も Maia と異なる。KataGo が提供する API では、アマの棋力を preaz_20k から preaz_9d, rank_20k から rank_9k というふうに設定できる。preaz と rank の差は、AlphaZero 流の布石が流行になった前か後である。また、プロ棋士と院生の場合には、proyear_1800 から proyear_2023 というふうに年を設定することができる。着手確率分布の使い方については、公開されたコードの説明^{*2} では「選択確率通りに着手を選ぶ」という方法が推奨される。

3. 予備調査と人間らしさの評価指標

本章では囲碁のアマプレイヤを対象にし、KataGo の HumanSL が予測した着手は、どのくらい指定した段級位らしいかを、4 個の指標で分析する。分析に用いた棋譜は、野狐囲碁というサイトから取り出された対局で、ネットで公開されたものである [6]。18k から 9d の段級位において、同じ段級位同士の対局を 1,000 局ずつサンプリングした。棋譜の殆どは AlphaZero 流の布石が流行になった前のものであるため、KataGo の HumanSL の棋力の設定に preaz を対象とした。比較対象として、(1) KataGo の自己対戦により学習したネットワーク (kata1) と、(2) Leela Zero で公開された、人間のアマ高段者の棋譜を学習したネットワーク (lzh) を含めた。

^{*1} <https://github.com/CSSLab/maia-chess/blob/master/README.md?plain=1#L60>

^{*2} https://github.com/lightvector/KataGo/blob/master/docs/Analysis_Engine.md?plain=1#L435

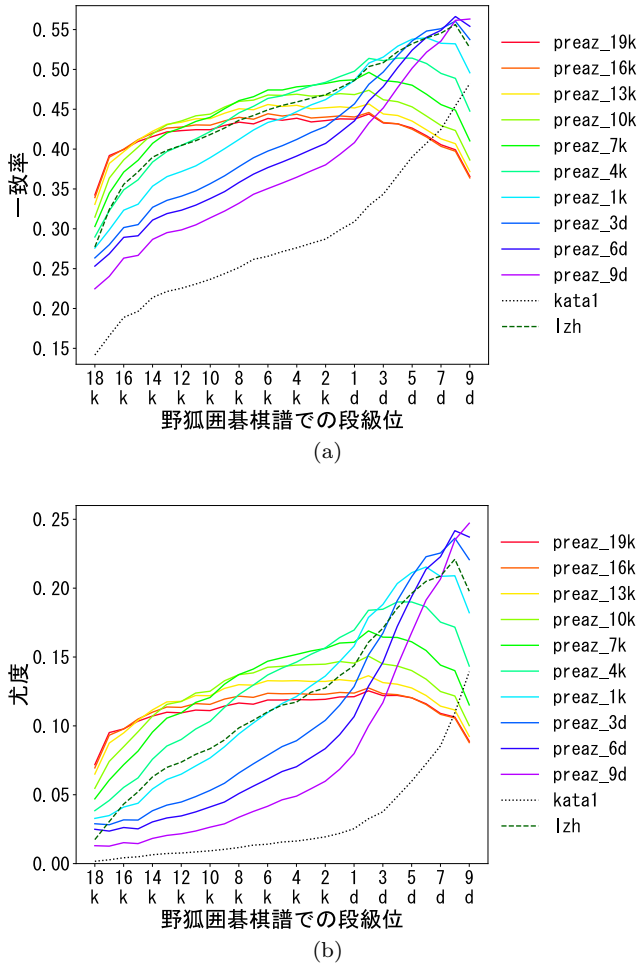


図 1: 実験に用いた野狐囲碁の棋譜において, KataGo の HumanSL preaz_{19k, 16k, ..., 9d} と kata1 と lzh の (a) 一致率と (b) 尤度.

Fig. 1 (a) Move-matching accuracy and (b) likelihood of KataGo's HumanSL preaz_{19k, 16k, ..., 9d}, kata1, and lzh in the game records used in the experiments that were from the Fox Go Dataset.

3.1 一致率

人間らしさの評価にまずはよく使われる一致率（選択確率が最も高い着手と、棋譜で実際に打たれた着手が同じである割合）で調べた。KataGo の HumanSL の preaz_19k から preaz_9d（虹色の曲線）の、野狐囲碁 18k から 9d の棋譜（横軸）においての一致率を図 1a に示す。なお、見やすさのために HumanSL の段級位を間引いている（19k, 16k, ..., 9d）。ある段級位の棋譜において、該当段級位とその付近の設定の一致率が最も高い傾向が見られた。この結果は既存研究 [3], [4] と同様である。参考に、kata1（自己対戦）の一致率を黒い点線で示す。値は大体 preaz より低く、kata1 は比較的人間らしくないと言える。また、lzh（アマ高段者）の一致率を濃い緑色の破線で示す。級位者での値が比較的低く、高段者での値が比較的高い傾向が見られ、予想通りの結果である。

3.2 尤度

2 章で紹介した通り、尤度は人間プレイヤーの着手の、モデルにおける選択確率の幾何平均である。一致率と比べて、尤度を用いるメリットは次の通りである。分かりやすさのため、 s_1 という局面において、人間プレイヤー 10 人中 8 人が着手 a_{11} 、残り 2 人が着手 a_{12} を選んだという例を考えよう。人間プレイヤーの着手の真の確率分布は a_{11} が 80%, a_{12} が 20% である。仮にモデル A と B があり、 $a_{11} \cdot a_{12}$ の予測選択確率はモデル A が 80% · 20%, モデル B が 60% · 40% であるとする。どちらのモデルも a_{11} が選択確率最大の着手であり、一致率が 80% になる。一方で、モデル A と B の尤度はそれぞれ $((0.8^8 \times 0.2^2)^{(1/10)} \approx 0.61$ と $((0.6^8 \times 0.4^2)^{(1/10)} \approx 0.55$ になり、モデル A が優れていることを表せる。また、幾何平均を用いているので、特に、実際に着手された手を低い確率で予測することの悪さをしっかり評価できる。

KataGo の HumanSL の、野狐囲碁 18k から 9d の棋譜（横軸）における尤度を図 1b に示す。ある段級位の棋譜において、該当段級位とその付近の設定の尤度が最も高い傾向は一致率と同様である。なお、曲線の形は一致率と異なる、特に preaz_高段者の場合には違いがはっきりと見える。尤度の結果から、「弱い人間プレイヤーの着手は高段者（のネットワーク）目線では不自然（選択確率が非常に小さい）なものが多い」と言えると考えられる。

3.3 直前手との距離

囲碁の人間プレイヤーは、相手の着手に応じてその近くに打つ傾向があり、弱いプレイヤーほどこの傾向が強いと考える。この傾向を反映する指標として、本稿では「着手と直前手の距離」を用いる。距離については Rémi Coulom 氏の定義 [1] を用いる（以下 Rémi 距離と呼ぶ）。具体的には x 成分差の絶対値 $|\delta_x|$ と y 成分差の絶対値 $|\delta_y|$ 及びそれらの最大値の総和、つまり、 $|\delta_x| + |\delta_y| + \max(|\delta_x|, |\delta_y|)$ である。Rémi 距離はユークリッド距離やマンハッタン距離に比べて、囲碁プレイヤーの感覚に近いと言われている。

野狐囲碁の棋譜において、人間プレイヤーが打った着手とその直前手の Rémi 距離の平均を図 2 の茶色の点で示す。18k の結果を除いたら概ね予想通りの傾向である。同じ棋譜の各局面から KataGo の HumanSL の確率分布通り打った場合の、直前手との Rémi 距離の結果を図 2 の虹色の曲線で示す。HumanSL の結果は人間プレイヤーの結果とだいぶ違うことが分かる。また、確率分布ではなく、選択確率最大の着手においても分析し、値が少し小さくなったことを確認したが、まだ人間プレイヤーの結果と離れている。この結果から、HumanSL の人間らしさに改善の余地が残されていると言えるかもしれない。

参考に、kata1（自己対戦）の直前手との Rémi 距離を黒い点線で示す。値は大体 preaz より大きく、直前手の付

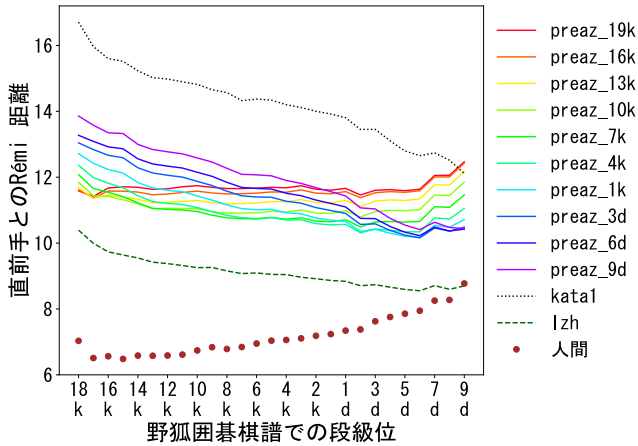


図 2: 実験に用いた野狐囲碁の棋譜において, KataGo の HumanSL $\text{preaz}_{\{19k, 16k, \dots, 9d\}}$ と kata1 と lzh の, 着手確率分布通りに打つ場合の直前手との Rémi 距離, 及び人間プレイヤーの着手と直前手の Rémi 距離.

Fig. 2 The Rémi distance to the previous move of KataGo’s HumanSL $\text{preaz}_{\{19k, 16k, \dots, 9d\}}$, kata1 , and lzh when selecting moves according to the probability distributions and the Rémi distance of human players’ moves to the previous moves in the game records used in the experiments that were from the Fox Go Dataset.

近に打つ傾向が比較的弱いと考える. また, lzh (アマ高段者) の直前手との Rémi 距離を濃い緑色の破線で示す. KataGo の kata1 や preaz より小さい値を得るという結果は興味深い. なお, この現象の原因究明は本稿の主目的ではないため将来の課題として残す.

3.4 棋力

棋力の評価指標について, 本来ならば各設定の HumanSL プレイヤをいろんな人間プレイヤーと対戦させるべきであるが, 十分な数の人間プレイヤーと対局を集めることは容易ではない. そのため, 人間プレイヤーの代わりに, 棋力がよく知られているプログラムを対戦相手とした. 具体的には GNU Go level 10^{*3} (以降 GNU Go, 棋力が約 kgs8k), 及び Pachi^{*4} の kgs3k と kgs2d の設定という 3 つのプログラムを用いた.

KataGo の HumanSL については, 用いたプログラムの棋力 (段級位) に合わせた設定を選び, GNU Go には $\text{preaz}_{.8k}$, $\text{Pachi}_{\text{kgs3k}}$ には $\text{preaz}_{.3k}$, $\text{Pachi}_{\text{kgs2d}}$ には $\text{preaz}_{.2d}$ で対応させた. 各 preaz 設定に「選択確率最大の着手」と「選択確率通り」という 2 つの使い方両方で実験した. $\text{preaz}_{.8k}$ を例でいうと, 「 $\text{preaz}_{.8k}$ の選択確率最大の着手を選ぶプレイヤー対 GNU Go」と「 $\text{preaz}_{.8k}$ の選択確率通りに着手を選ぶプレイヤー対 GNU Go」といった対戦実験を行った. 19 路盤, 日本ルール, コミ 6 目半, 各対戦組み合わせに先後

表 1: KataGo の HumanSL に基づくプレイヤー対 GNU Go・ $\text{Pachi}_{\text{kgs3k}}$ ・ $\text{Pachi}_{\text{kgs2d}}$ の勝率とその 95% 信頼区間

Table 1 Win rates and the 95% confidence intervals for players based on KataGo’s HumanSL playing against GNU Go, $\text{Pachi}_{\text{kgs3k}}$, and $\text{Pachi}_{\text{kgs2d}}$.

対戦相手	GNU Go	$\text{Pachi}_{\text{kgs3k}}$	$\text{Pachi}_{\text{kgs2d}}$
HumanSL	$\text{preaz}_{.8k}$	$\text{preaz}_{.3k}$	$\text{preaz}_{.2d}$
確率最大	$82.50 \pm 5.27\%$	$97.50 \pm 2.16\%$	$96.00 \pm 2.72\%$
確率通り	$28.00 \pm 6.22\%$	$38.50 \pm 6.74\%$	$24.50 \pm 5.96\%$

交代で 200 試合行った.

対戦の勝率とその 95% 信頼区間を表 1 に示す. HumanSL の $\text{preaz}_{.8k}$ と $\text{preaz}_{.3k}$ と $\text{preaz}_{.2d}$ のいずれも, 「選択確率最大」の着手を選ぶ場合に設定した段級位よりはるかに強く, 「選択確率通り」に着手を選ぶ場合に設定した段級位より弱いという傾向が見られた. チェスにおいて, アマプレイヤーの棋譜を学習したモデルについても似た傾向が報告されている [7].

こういった傾向が表れた原因を次の通り推測する. 平均的なプレイヤーの着手を学習したモデルの場合, 仮にモデルがよく学習できたと仮定すると, 選択確率最大の着手を選ぶことは, 多数決を行うことと同様の効果がある. 少し極端な例であるが, 例えばある局面において必勝手 A と必敗手 B があり, 10 人中 6 人が A, 4 人が B を選び, モデルも A と B の選択確率をそれぞれ 60% と 40% と予測したとする. 選択確率最大の着手を選ぶという使い方だと常に必勝手 A が選ばれ, 平均的なプレイヤーより強くなってしまふことが分かる. 選択確率通りに着手を選ぶという使い方はこの局面に限って言うと, 平均的なプレイヤーと同じ棋力が得られる. しかし, 一般的にはそもそもモデルがうまく学習できている保証がない. また, 実際の対局ではもっと多くの局面に関わり, 「運が悪くてひどい着手を選んでしまったらもう逆転の余地がない」という状況はめずらしくないと予想する. そのため, 選択確率通りに着手を選んだら平均的なプレイヤーより弱くなってしまふことは合理的であると考え. 設定した段級位通りの棋力でプレイさせるために, 選択確率の使い方について検討する余地があると言える.

4. 直前手との距離による補正

3.3 節で示したように, HumanSL は人間のプレイヤーよりも直前手から離れたところに打つ傾向があることが見られた. 本章では, 直前手との距離による補正を入れることで着手の選択確率を調整する方法を提案し, 調整された確率分布を 3 章の評価指標で評価する.

4.1 提案手法

直前手との距離を縮めるために, 直前手に近い着手の選

^{*3} <https://www.gnu.org/software/gnugo/>

^{*4} <https://github.com/pasky/pachi>

択確率が高く、遠い着手の選択確率が低くなるような重みづけを行う。局面 s において HumanSL が予測した、着手 a の選択確率を $P(s, a)$ とし、直前手を a^- とする場合に、新しい選択確率は

$$P'(s, a) = \frac{W(a^-, a) \times P(s, a)}{\sum_b (W(a^-, b) \times P(s, b))} \quad (1)$$

になる。 d を着手 a と直前手 a^- の Rémi 距離（どちらもパスではない場合）、 m を線形関数の傾きとし、重みを次の通り計算する。

$$W(a^-, a) = \begin{cases} 1, & \text{if } a^- \text{ が } a \text{ がパス} \\ 0.1, & \text{if } d > 14 \\ 0.1 + (14 - d) \times m, & \text{if } 14 \geq d > 4 \\ 0.1 + (14 - 4) \times m, & \text{if } 4 \geq d \end{cases} \quad (2)$$

m が大きければ補正の影響が大きくなり、直前手に近く打つ傾向が強くなると予想する。この重みづけの補正は Shi らの方式 [8] を参考にしているが、ユークリッド距離の代わりに Rémi 距離を用いることと、階段関数の代わりに線形関数 ((2) の 3 行目) にすることで、より自然な補正になると考える。

4.2 結果

4.2.1 直前手との距離

直前手との距離による補正を入れる効果を、まずは直前手との距離で評価した。3 章の実験で用いた野狐囲碁の棋譜のうち、棋力が 16k, 13k, 10k, 7k, 4k, 1k, 3d, 6d, 9d のものを対象とした。人間プレイヤーの着手と直前手の Rémi 距離を、表 2 の 2 列目に示す (図 2 の茶色の点と同じ)。HumanSL については、対象棋譜と同じ段級位の preaz 設定を用いた。距離補正なしの、HumanSL の確率分布通りに打った場合の直前手との Rémi 距離の結果を 4 列目に示す。図 2 に示した通り、人間プレイヤーの統計量と離れている。

距離補正について、(2) の線形関数の傾き m というパラメータを、0.03, 0.05, 0.1, 0.15, 0.2, 0.25 で実験した。新しい確率分布通りに打った場合の直前手との Rémi 距離の結果を表 2 の 5–10 列目に示す。 m を適切に設定すれば、人間プレイヤーの統計量に近いものが得られることを確認した。また、棋力が弱いほど、 m を大きく設定したほうが良いという傾向が見られた。3.3 節で説明した通り、弱いプレイヤーほど、相手の着手に応じてその近くに打つ傾向が強いため、この結果は予想通りである。

4.2.2 一致率と尤度

次に、距離補正を入れることでの一致率と尤度への影響を調べた。各段級位には、人間プレイヤーの直前手との距離に最も近い結果を得た m を選んだ。具体的には、preaz_16k に $m = 0.25$, preaz_13k と preaz_10k と preaz_7k に $m = 0.2$,

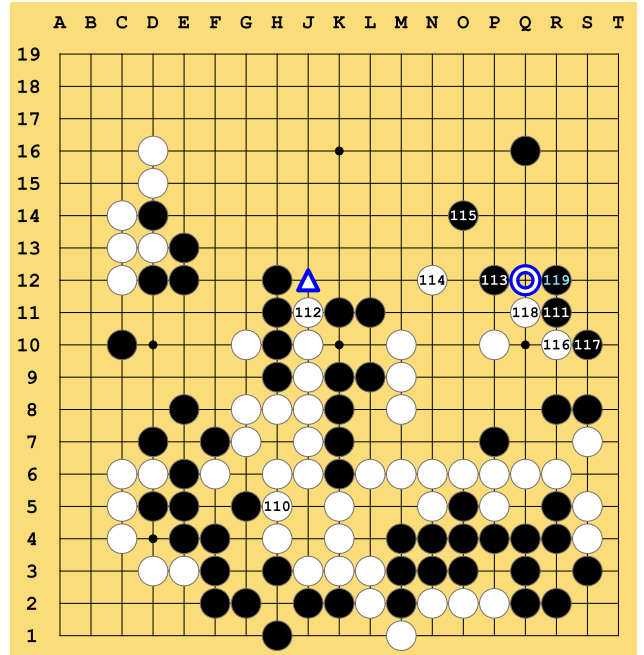


図 3: 直前手との距離による補正を入れることによって選択確率最大の着手が 1k 人間プレイヤーの着手と一致した局面 (マル: 実際の着手, 三角: 補正なしの preaz_1k 選択確率最大の着手)。

Fig. 3 A position (game state) where a 1k human player's move matches the move with the maximum selection probability after the adjustment according to the distance to the previous move (circle: the actually played move, triangle: the move with the maximum selection probability of preaz_1k without the adjustment).

preaz_4k と preaz_1k に $m = 0.15$, preaz_3d に $m = 0.1$, preaz_6d に $m = 0.05$, preaz_9d に $m = 0.03$ といった設定を用いた。距離補正がある場合とない場合の一致率と尤度の結果をそれぞれ表 3 と表 4 に示す。いずれの段級位においても、一致率と尤度両方を向上させたことを確認した。また、棋力が弱いプレイヤーほど改善が大きい傾向が見られた。

直前手との距離による補正を入れることで、一致率と尤度を向上させた原因を実際の局面を用いて説明する。野狐囲碁 1k プレイヤ同士での対局から取り出した 1 つの局面を図 3 に示す。119 手目に黒のプレイヤーが R12 に打ち、120 手目に白のプレイヤーが Q12 と、先の黒石の隣に打った。113 手目から右辺での攻防が行われてきたため、120 手目に Q12 に打つことは級位者にとって流れとして自然であると考えられる。一方、120 手目に Q12 の preaz_1k での選択確率は 0.077 であり、部分的には非常に小さい値であった。 $m = 0.15$ の距離補正を入れたら、Q12 と J12 それぞれの選択確率が 0.413 と 0.241 となり、選択確率最大の着手は Q12 になって人間プレイヤーの着手と一致した。つま

表 2: 対象の野狐囲碁棋譜において、人間プレイヤーの着手と直前手の Rémi 距離、及び HumanSL に距離補正がある場合とない場合の、着手確率分布通りに打つ場合の直前手との Rémi 距離

Table 2 The Rémi distance of human players' moves to the previous moves and the Rémi distance to the previous move of HumanSL when selecting moves according to the probability distributions, with and without the adjustment according to distance, in the target game records from the Fox Go Dataset.

対象棋譜	人間	HumanSL							
		棋力設定	補正なし	距離補正重みの m					
				0.03	0.05	0.1	0.15	0.2	0.25
16k	6.56	preaz_16k	11.58	8.94	8.28	7.42	6.97	6.67	6.46
13k	6.58	preaz_13k	11.22	8.77	8.15	7.34	6.91	6.63	6.43
10k	6.74	preaz_10k	11.04	8.76	8.18	7.41	6.99	6.72	6.53
7k	6.84	preaz_7k	10.74	8.65	8.11	7.39	7.00	6.75	6.57
4k	7.06	preaz_4k	10.70	8.73	8.21	7.52	7.15	6.90	6.72
1k	7.24	preaz_1k	10.70	8.82	8.33	7.66	7.30	7.06	6.88
3d	7.62	preaz_3d	10.58	8.82	8.35	7.71	7.36	7.13	6.96
6d	7.95	preaz_6d	10.22	8.56	8.13	7.54	7.22	7.00	6.84
9d	8.77	preaz_9d	10.44	8.75	8.31	7.71	7.39	7.17	7.01

表 3: 距離補正がある場合とない場合の一致率

Table 3 The move-matching accuracy with and without the adjustment according to distance.

対象棋譜	棋力設定	補正なし	補正あり	改善
16k	preaz_16k	0.400	0.468	+0.068
13k	preaz_13k	0.431	0.494	+0.063
10k	preaz_10k	0.444	0.502	+0.058
7k	preaz_7k	0.466	0.518	+0.052
4k	preaz_4k	0.473	0.522	+0.048
1k	preaz_1k	0.474	0.519	+0.045
3d	preaz_3d	0.497	0.534	+0.037
6d	preaz_6d	0.540	0.565	+0.025
9d	preaz_9d	0.563	0.578	+0.015

表 4: 距離補正がある場合とない場合の尤度

Table 4 The likelihood with and without the adjustment according to distance.

対象棋譜	棋力設定	補正なし	補正あり	改善
16k	preaz_16k	0.098	0.130	$\times 1.333$
13k	preaz_13k	0.117	0.154	$\times 1.308$
10k	preaz_10k	0.125	0.160	$\times 1.277$
7k	preaz_7k	0.141	0.175	$\times 1.238$
4k	preaz_4k	0.146	0.178	$\times 1.220$
1k	preaz_1k	0.147	0.177	$\times 1.208$
3d	preaz_3d	0.167	0.194	$\times 1.161$
6d	preaz_6d	0.214	0.235	$\times 1.098$
9d	preaz_9d	0.247	0.259	$\times 1.047$

り、実際に着手された手の選択確率を 0.077 と過小評価せず、0.413 と高い確率で着手されると予測する確率分布にできた。このような局面は他にもたくさんあったため、直前手との距離が人間の統計量に近くなっただけでなく、

表 5: 直前手との距離による補正を入れた HumanSL に基づくプレイヤー対 GNU Go・Pachi_{kgs3k}・Pachi_{kgs2d} の勝率とその 95%信頼区間

Table 5 Win rates and the 95% confidence intervals for players based on KataGo's HumanSL with the adjustment according to the distance to the previous move playing against GNU Go, Pachi_{kgs3k}, and Pachi_{kgs2d}.

対戦相手	GNU Go	Pachi _{kgs3k}	Pachi _{kgs2d}
HumanSL	preaz_8k	preaz_3k	preaz_2d
距離補正	$m = 0.2$	$m = 0.15$	$m = 0.1$
確率最大	76.00 $\pm$ 8.37%	94.00 $\pm$ 4.65%	84.00 $\pm$ 7.19%
確率通り	31.00 $\pm$ 9.06%	37.00 $\pm$ 9.46%	29.00 $\pm$ 8.89%

一致率と尤度の向上にも貢献したと考えている。

4.2.3 棋力

最後に 3.4 節と同様に、preaz_8k と preaz_3k と preaz_2d により作ったプレイヤーをそれぞれ、GNU Go と Pachi_{kgs3k} と Pachi_{kgs2d} と対戦させた。各 preaz 設定に「選択確率最大の着手」と「選択確率通り」という 2 つの使い方両方で実験した。距離補正用のパラメータについて、preaz_8k に $m = 0.2$ 、preaz_3k に $m = 0.15$ 、preaz_2d に $m = 0.1$ といった設定を用いた。各対戦組み合わせに先後交代で 100 試合行った。

対戦の勝率とその 95%信頼区間を表 5 に示す。表 1 と同様に、「選択確率最大」だと強すぎて「選択確率通り」だとちょっと弱いという傾向が見られた。距離補正を入れた場合に、Pachi_{kgs2d} の「選択確率最大」プレイヤーが距離補正なしに比べて少し弱くなったこと以外に、勝率の値は表 1 と大きな差が見られなかった。この結果から、直前手との距離による補正を入れることは棋力への影響が小さいと言えるかもしれない。

表 6: 直前手との距離による補正を入れた HumanSL をもとに, softmax 方式で着手を選ぶプレイヤー対 GNU Go・Pachi_{kgs3k}・Pachi_{kgs2d} の勝率とその 95%信頼区間

Table 6 Win rates and the 95% confidence intervals for players based on KataGo's HumanSL with the adjustment according to the distance to the previous move that select moves by the softmax method playing against GNU Go, Pachi_{kgs3k}, and Pachi_{kgs2d}.

対戦相手	GNU Go	Pachi _{kgs3k}	Pachi _{kgs2d}
HumanSL	preaz_8k	preaz_3k	preaz_2d
距離補正	$m = 0.2$	$m = 0.15$	$m = 0.1$
$\rho = 2.00$	67.00±9.22%	79.00±7.98%	72.00±8.80%
$\rho = 1.50$	57.00±9.70%	78.00±8.12%	71.00±8.89%
$\rho = 1.20$	51.00±9.80%	73.00±8.70%	48.00±9.79%
$\rho = 1.10$	36.00±9.41%	63.00±9.46%	36.00±9.41%
$\rho = 1.05$	36.00±9.41%	49.00±9.80%	36.00±9.41%
$\rho = 1.00$	31.00±9.06%	37.00±9.46%	29.00±8.89%

5. Softmax

表 1 と表 5 の結果では, 「選択確率最大の着手」を選ぶと強すぎ, 「確率通り」に着手を選ぶとちょっと弱いという傾向が見られた. 本章では, この 2 つの着手確率分布の使い方の中間的なものを用い, ちょうどよい棋力に達成することを目指す.

5.1 用いる手法

局面 s において着手 a の選択確率を $P(s, a)$ とする場合に, 「選択確率最大の着手」を選ぶことを, $P'(s, a) \propto (P(s, a))^\rho$ と見なすことができる. 「確率通り」つまり $P'(s, a) \propto (P(s, a))^1$ との中間版を,

$$P'(s, a) \propto (P(s, a))^\rho \text{ subject to } \rho \in (1, \infty) \quad (3)$$

により作成できる. ρ が大きいほど強くなると予想する. 先行研究 [9], [10] ではこの方式を softmax と呼ぶ.

5.2 結果

3.4 節と 4.2.3 項と同様に, preaz_8k と preaz_3k と preaz_2d により作った softmax プレイヤーをそれぞれ, GNU Go と Pachi_{kgs3k} と Pachi_{kgs2d} と対戦させた. 距離補正用のパラメータは 4.2.3 項と同じ設定 (preaz_8k に $m = 0.2$, preaz_3k に $m = 0.15$, preaz_2d に $m = 0.1$) を用いた. 各 preaz 設定ごとに, softmax のパラメータ $\rho \in \{2.00, 1.50, 1.20, 1.10, 1.05\}$ を実験した.

対戦の勝率とその 95%信頼区間を表 6 に示す. ρ を適切に設定すれば, 目標棋力に達成できる, つまり勝率が 50% 程度にできることを確認した. 実験の結果から, preaz_8k に $\rho = 1.20$, preaz_3k に $\rho = 1.05$, preaz_2d に $\rho = 1.20$ が適していることが分かったが, 全体的に各 preaz 設定に

表 7: プロ棋譜での HumanSL とブレンドの一致率

Table 7 The move-matching accuracy of HumanSL and the blend models in professional game records.

	HumanSL	ブレンド	改善
GoGoD 2012	0.555	0.568	+0.013
GoGoD 2022	0.566	0.586	+0.020

どの ρ 設定が良いかについては, 明確なパターンが見つからなかった. このようなパターンが存在するかの解明を将来の課題として残す.

6. プロ棋譜での初歩結果

本章では囲碁のプロ棋士を対象にし, KataGo の HumanSL と自己対戦のネットワークをブレンド [4] する効果を分析する. 分析に用いた棋譜は, GoGoD データセット*⁵ の 2012 年と 2022 年の棋譜から 1,000 局ずつサンプリングしたものである. 2012 年と 2022 年の棋譜それぞれに, KataGo の HumanSL の proyear_2012 と proyear_2022 という設定にした. HumanSL とブレンドの一致率を表 7 に示す. どちらの年の結果でも, 自己対戦ネットワークとブレンドすることで, 一致率が上がったことを確認した. この結果から, HumanSL のプロ棋士の着手模倣においても改善の余地があることを示した.

7. おわりに

本稿では KataGo の HumanSL という, 大量の人間プレイヤーの棋譜を用いて学習したモデルの適切な使い方とその人間らしさについて調べた. HumanSL には, 予測したい段級位を入力で指定することができ, 出力には着手選択確率の分布が含まれる. しかし, 対戦相手にしたい場合にこのような確率分布をどう使うのかについて合意はない. 本稿では「選択確率最大の着手を選ぶ」方法 (つまり $p' \propto p^\infty$) と「選択確率通りに着手を選ぶ方法」 (つまり $p' \propto p^1$), 及びそれらの中間版である softmax (つまり $p' \propto p^\rho$, $\rho \in (1, \infty)$) を調べた. この 3 つの方法を, 棋力が既知のプログラムと対戦させた結果, 「選択確率最大の着手」だと強すぎ, 「選択確率通り」だとやや弱く, softmax の場合はパラメータ ρ を適切に設定すれば, HumanSL に入力した段級位とほぼ同じ棋力にできることを示した.

また, アマの囲碁プレイヤーの棋譜を用い, HumanSL の人間らしさを, 着手一致率, 尤度, 直前手との距離といった指標で評価した. これらの結果をもとに, さらに人間らしさを向上させる方法を提案した. 具体的には直前手との距離による補正を選択確率に入れた. 直前手に近いほど選択確率が大きく, 遠いほど選択確率が小さくなるように調整を行った. これは, 人間プレイヤー (特に弱いプレイヤー) が直前手に応じてその近くに打つ傾向が見られたためであ

*⁵ <https://gogodonline.co.uk/>

る。実験結果では、弱いプレイヤほど人間らしさの向上幅が大きく、一致率は最大 6.8 ポイントの改善が見られた。

将来の課題としていくつかの点が残されている。まず、着手選択確率の分布の使い方について、「確率分布通りに選ぶ」方法でも softmax 方式でも、確率 0.001 のような、不自然かもしれない着手が選ばれる可能性がある。選択確率が非常に小さい着手を除外するフィルタリングを検討したい。また、本稿では距離補正や softmax 用のパラメータをグリッドサーチで手動で決めたが、予測したい段級位が変わったり、新しいモデルが出たりするたびに、パラメータを手動で決めるのは効率が悪いので、何かの関数に落とし込めないかを考えている。関連して、今回の分析ではアマのプレイヤを中心に行ったが、プロの棋譜において、例えば直前手による距離補正が効果を持つかも調べたい。最後に、距離補正と softmax をもとに作ったプレイヤの着手がどのくらい自然に見えるかについて、実際に人間プレイヤに評価してもらうことも考えている。

謝辞 本研究は JSPS 科研費 JP23K17021 と JP23K11381 の助成を受けたものです。

参考文献

- [1] Coulom, R.: Computing Elo ratings of move patterns in the game of Go, *CGW 2007*, pp. 113–124 (2007).
- [2] Silver, D. et al.: Mastering the game of Go with deep neural networks and tree search, *Nature*, Vol. 529, No. 7587, pp. 484–489 (2016).
- [3] McIlroy-Young, R., Sen, S., Kleinberg, J. and Anderson, A.: Aligning superhuman AI with human behavior, *KDD '20*, pp. 1677–1687 (2020).
- [4] Ogawa, T., Hsueh, C.-H. and Ikeda, K.: Improving the human-likeness of game AI's moves by combining multiple prediction models, *ICAART 2023*, pp. 931–939 (2023).
- [5] lightvector: New Human-like Play and Analysis, GitHub (online), available from (<https://github.com/lightvector/KataGo/releases/tag/v1.15.0>) (accessed 2024-10-10).
- [6] featurecat: Fox Go Dataset, GitHub (online), available from (<https://github.com/featurecat/go-dataset>) (accessed 2024-10-10).
- [7] Zhang, E. et al.: Transcendence: Generative Models Can Outperform The Experts That Train Them (2024).
- [8] Shi, Y., Fan, T., Li, W., Hsueh, C.-H. and Ikeda, K.: Position control and production of various strategies for game of Go using deep learning methods, *Journal of Information Science and Engineering*, Vol. 37, No. 3, pp. 553–573 (2021).
- [9] Liu, A.-J., Wu, T.-R., Wu, I.-C., Guei, H. and Wei, T.-H.: Strength adjustment and assessment for MCTS-based programs [research frontier], *IEEE Computational Intelligence Magazine*, Vol. 15, No. 3, pp. 60–73 (2020).
- [10] Hsueh, C.-H. and Ikeda, K.: Improvement of move naturalness for playing good-quality games with middle-level players, *Applied Intelligence*, Vol. 54, No. 2, pp. 1637–1655 (2024).