

Title	プログラムの役割分担によって好ゲームを演出する接待基AIとその課題
Author(s)	シュエ, ジュウシュエン; 池田, 心
Citation	情報処理学会第51回GI研究発表会, 2024-3, 2024-GI-51(19): 1-8
Issue Date	2024-03-09
Type	Conference Paper
Text version	publisher
URL	http://hdl.handle.net/10119/18827
Rights	<p>社団法人 情報処理学会, シュエ ジュウシュエン, 池田心, 情報処理学会第51回GI研究発表会, 2024-3, 2024. 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」に従うこと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.</p>
Description	情報処理学会第51回GI研究発表会, 2024-3, 国立情報学研究所

プログラムの役割分担によって好ゲームを演出する 接待碁 AI とその課題

シュエ ジュウシュエン*^{1,a)} 池田 心*^{1,b)}

概要: いまや多くのゲームでコンピュータプレイヤは人間のトッププレイヤよりも強く、上達のためにも利用されている。しかし多くのプレイヤにはそれらは強すぎて、普通に対戦することが楽しみにはつながらない場合も多い。先行研究では、初中級者を対象とし、役割の異なる2つのプログラムを組み合わせることで、好ゲームを演出する手法が提案された。1つ目は人間のトッププレイヤよりも強く、着手の良さを評価するものである。2つ目は着手の自然さ（又は人間らしさ）を評価するものである。囲碁に適用した際には、1つ目に KataGo, 2つ目に人間棋譜を学習したニューラルネットワークが用いられた。実験では好ゲームの演出が概ねできていることが示されたが、初級者の目線でも不自然に感じるような手を打ってしまうなど課題がいくつか残されている。本稿ではそれらの課題をまとめ、可能な改善案を述べる。

1. はじめに

人工知能 (AI) は、ゲームを含む多くの分野で大きな進歩を遂げてきた。長い間、ゲーム分野では主に、強いコンピュータプログラムを開発する研究が行われていた。深層学習技術や探索アルゴリズムの進歩により、AI プレイヤはさまざまなゲームで人間のトッププレイヤを上回っている。AlphaZero はその代表であり、専門家による知識なしに、自己対局によってチェス、将棋、囲碁において人間をはるかに超える強さを獲得した。

このような強いプログラムは、人間プレイヤに新しい打ち方を勉強する機会を与えるが、人によっては、それらとの対戦によって学んだり楽しんだりできない可能性もある。例えば、プログラムが強すぎて負けてばかりいると、人間プレイヤは挫折感を覚え、ゲームを続ける動機を失うかもしれない。また、コンピュータプログラムは時に人間プレイヤに理解や真似ができない手を打つため、最も優れたお手本とは言えないかもしれない。

強いプログラムが人間の指導者のように上手に教えることができるようになれば、多くの人にとって恩恵がある。人間の指導者が生徒に教える方法には、良い例を示したり、悪手を指摘したり、指導対局を行うなど、いくつかの方法

がある。指導対局の文化は将棋やシャンチーにも存在するが、以下は囲碁における指導碁を説明する。図1の黒枠は、人間の指導者が指導碁を行う際に一般的に考慮する内容を示している。指導碁はさらに、好ゲームを演出する段階と、対局を振り返る段階の2つに分けることができるが、本研究では、前者に焦点を当てる。

好ゲームについて、人間の指導者は暗黙的または明示的に以下の要素を達成することを目指していると考えた [1], [2].

- きわどい勝負をする。ハンデ（置き石）の有無に関わらず、勝敗を制御することは重要である。(a1) 指導者が勝つこともあれば、わざと生徒に勝たせることもある。(a2) また、多くの試合では、指導者は大差での勝ち負けを避けようとする。
- 相手の出来不出来を勝敗に反映させる。(a1) に関連して指導者は意図的に生徒を勝たせることがある。生徒に良い手、悪い手を印象深く記憶させるために、指導者は最終的な勝敗と生徒の手の良さ悪しを関連付けようとすることがある。例えば、対局中に生徒が良い手をたくさん打った場合、指導者は一種のご褒美として生徒に勝たせるかもしれない。
- 筋の良い手を打つ。良い勝負をするには、適切なハンデでない場合、わざと損をする必要がある。しかし、生徒が指導者の着手から良い石の形や手筋を学ぼうとすることを考え、指導者は損をするにしても着手の筋や形はできるだけ良く保とうとする。
- 自然な（もしくは人間らしい）手を打つ。人間らしい手を打つことは、人間の指導者にとっては問題ではな

*Equal Contribution

¹ 北陸先端科学技術大学院大学
Japan Advanced Institute of Science and Technology, Nomi,
Ishikawa 923-1211, Japan

a) hsuehch@jaist.ac.jp

b) kokolo@jaist.ac.jp

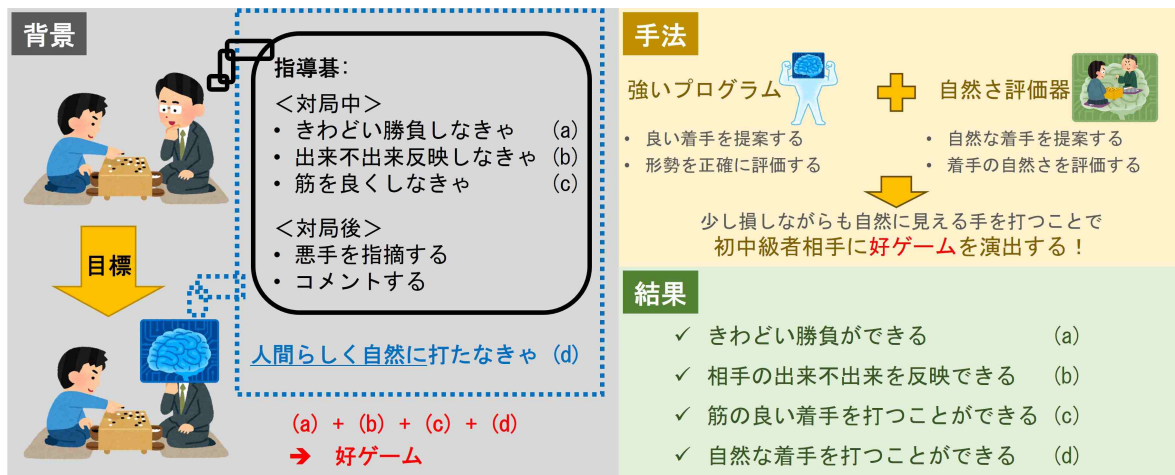


図 1: 本研究の全体像：人間の指導者（実線の黒枠）を強いプログラム（点線の青枠）が模倣する。

いが、コンピュータプログラムにとっては簡単ではない。例えば、人間特に初中級者は局所的な戦いを終えてから別の戦いに移る傾向があるが、AlphaZero ベースの強いプログラムは、こういった傾向が弱い。

本研究の目的は、図 1 の点線の枠にあるように、強いプログラムを用いて初中級者と好ゲームを演出できるコンピュータプログラムを作成することである。具体的には、kgs8k から kgs2d^{*1} 程度のプレイヤーを対象とする。

1つの要素を達成するプログラムの作成は容易でも、4つを同時に考慮することは難しい課題である。例えば (a1) を実現するためには、勝率が 50% に最も近い手を選択する単純な方法がある。しかしこの方法は、人間が大きなミスをしたときに、優勢をなくすために同じくらい大きな悪手を選択してしまうため、相手の手の良し悪しは最終的な勝敗にはほとんど反映されず、(b) が達成できない。また、こういった手法で選ばれた手は、形や筋が悪い、または人間らしくない可能性があり、(c)(d) が達成できない。

そこで、我々は役割の異なる 2つのプログラムを組み合わせることで、好ゲームを演出する手法を提案した [1], [2]. 1つ目は人間のトッププレイヤーよりも強く、着手の良さを評価するプログラムである。2つ目は着手の自然さ（又は人間らしさ）を評価するプログラムである。囲碁においての実現では、1つ目の役割に KataGo [3], 2つ目の役割に人間棋譜を学習したニューラルネットワーク [4] を用いた。実験では好ゲームの演出が概ねできていることを示したが、課題もいくつか見つかった。本稿では先行研究の手法と主な結果を紹介し、見つかった課題を述べる。

2. 関連研究

強いプログラムを意図的に弱める手法はいくつかある。Sephton らはモンテカルロ木探索 (MCTS) における着手

選択法として softmax 方策に基づく着手選択を行った： n_i は着手 i の訪問回数、 z は棋力を制御するパラメータとし、 $(n_i)^z / \sum_j (n_j)^z$ の確率で着手 i を選択する [5]. この方法の問題点は、訪問回数が 0 でない限り、非常に悪い着手が選択される可能性があることである。これを解決するため、Liu らは訪問回数の閾値を導入した [6]. MCTS で訪問回数が最も多い着手が n_{max} 回訪問された場合に、訪問回数 $n_{max} \times R_{th}$ 未満の着手は候補から除外された。

人間らしさの実現のために、人間プレイヤーの棋譜を活用した手法もある。例えば、Nakamichi らは将棋アマチュアの棋譜を学習した評価関数を用い、弱いながらも人間らしいプログラムを作成した [7]. Rosemarin らはいくつかのパラメータを導入した MCTS の変形を提案し、異なるレート帯のチェスアマチュアの棋譜を用いてパラメータを調整し、それぞれの強さに対応するプログラムを作成した [8].

人間プレイヤーの棋譜の活用として、アマチュアの着手を予測するニューラルネットワークを学習するアプローチも用いられる [9]. この研究ではチェスの棋譜を、[1100, 1199], [1200, 1299] などといったレート帯に分け、レート帯ごとにニューラルネットワークを学習し、それがそのレート帯におけるアマチュアの着手を予測するのに適していることを示した。また、Minimax ベースや AlphaZero ベースのプログラムの着手が人間プレイヤーとはかなり異なっていることも示した。予測精度をさらに向上させるために、Jacob らはニューラルネットワークを MCTS に組み込むことが効果的であり、同時にプログラムの棋力も向上させることを示した [10]. Ogawa らは AlphaZero の方策と、人間プレイヤーの棋譜を学習した方策をブレンドすることで予測精度を向上できることを示した [11].

3. 好ゲームを演出するための役割分担

1章で述べたような初中級者との好ゲームを演出するために、我々は役割の異なるプログラムを組み合わせること

*1 KGS は囲碁サーバの 1 つであり、人間とプログラム両方に棋力に対応するランクを付与する。https://www.gokgs.com

を提案し、囲碁に適用した [1], [2].

着手の良さを正確に評価するためには AlphaZero ベースの KataGo [3], 自然さの評価には人間の高級者の棋譜を学習したニューラルネットワーク [4]*2 を用いた. 本稿以降では, このニューラルネットワークによる方策 (着手の確率分布) を $Prior_{human}$ と表記する.

KataGo の着手は強いが, 人間にとって理解しにくい手も多く, 一致率も低い [9]. また, 人間の初中級者が自然と思うような手を「全くありえない手」と判断することも多い. 一方 $Prior_{human}$ は人間の感覚と合っていることが多いが, 教師あり学習の性質上, 場合によっては局面や着手を正確に評価できない. したがって, 異なる役割を異なるプログラムに分けることは合理的である*3.

この役割分担を踏まえ, 「自然な手加減度」という着手の評価関数と, 「sequential halving」という探索を実装している [2]. それぞれを 3.1 節と 3.2 節で要約する.

3.1 自然な手加減度

着手の「自然な手加減度」を計算するために, 着手 i の訪問回数 n_i とリード目数 a_i , $Prior_{human}$ p'_i が得られるとする. a_i と p'_i はそれぞれ強いプログラム役と自然さ評価役が評価する. まずは何目程度の手加減を行うかという「目標損失」 l^{target} を計算する. $\alpha > 0$ を手加減度を調整する形勢の範囲, $\beta > 0$ を損失の程度, a_{max} を $\max_{i \in \{j | n_j > 10\}} a_i$ (つまりある程度訪問した着手の中で最も優勢の着手のリード) とし, 図 2a で示している通り, 優勢時に大きな損を, 劣勢時には小さい損を目指す.

得られた目標損失 l^{target} をもとに各着手 i の「手加減度」 q_i を図 2b のように定義する (τ は定数). 目標損失が小さい場合に損失の許容範囲が小さく (橙色の曲線, 山型が比較的狭い), 目標損失が大きい場合に損失の許容範囲が大きい (青色の曲線) という設計になっている.

最後に, 「自然な手加減度」 s_i を $(p'_i + \epsilon) \times q_i$ で計算する. $\epsilon = 10^{-3}$ は自然さ評価が 0 にならないようにするための小さい数である. 式から分かる通り, 自然さ (p'_i) と手加減の程度 (q_i) のバランスを取ろうとする設計になっている. なお, 着手の筋の良さを一定程度保つため, 損失が 20 目を越えた場合に s_i を $-\infty$ とする.

3.2 Sequential Halving による探索

前節で各着手に「自然な手加減度」を与える方法を説明したが, KataGo の通常の探索ではいわゆる「勝つための手」が重点的に探索されてしまい, 「手加減のための少し悪い手」が候補から外れたり, 損失の推定精度が低くなって

しまうという課題がある. そこで, sequential halving [12] という手法を参考に, 勝つための手に探索を偏らせず, 手加減のための手を十分に訪問して評価する探索を提案した.

有望な手に探索資源を重点的に割り当てる通常の MCTS と異なり, sequential halving は, 与えられたシミュレーション数をいくつかのラウンドに分け, 各ラウンドでは候補着手を同じシミュレーション数だけ探索する. 図 2c は簡単な例で, 盤面 s に 4 つの着手 $m_1 \sim m_4$ がある. 探索は $R = \log_2 4 = 2$ ラウンドに分けられる. 総シミュレーション数 N_{sim} を 1000 回とすると, 各ラウンドは $N_{sim}/R = 500$ 回, 各着手は 125 および 250 回探索されることになる. 勝率によって下位半分の手を削除し, これを繰り返すことで最後に残った着手 (図 2c では m_4) を打つ. この例では下位半分としているが, 割合 $\lambda \in (0, 1)$ の着手を次のラウンドに残す拡張も行われている [13].

数百の合法手がある囲碁に対し, 我々は, 1 ラウンド目では 26 個前後の着手のみ, 2 ラウンド目ではそのうち 3.1 節で説明した「自然な手加減度」の高い 4 つの着手のみを探索する. 1 ラウンド目の探索対象は以下の条件のいずれかを満たす手である.

- $Prior_{human}$ 上位 25 の着手*4
- $Prior_{KataGo}$ (KataGo 方策ネットワークによる着手の選択確率) 上位 5 の着手
- パス

$Prior_{human}$ は自然な着手, $Prior_{KataGo}$ は良い着手を提案する. この部分でもお互いの弱点を補いあう役割分担ができていく. これらの候補手を M_0 とする.

総シミュレーション数 N_{sim} 回に対して, ラウンド 1 では $N_{sim}/2$ 回で M_0 の各着手を均等に調べ, 自然な手加減度上位 4 つの着手を残す. これらの着手を M_1 とする. 同様にラウンド 2 では $N_{sim}/2$ 回で M_1 の各着手を調べ, 自然な手加減度最上位*5の着手を打つ.

4. 先行研究の主な結果

本章では先行研究の実験と結果を要約し (4.1 節~4.5 節), 自然な手加減がよくできている例を紹介する (4.6 節).

4.1 実験設定

手加減を行う対象として, 先行研究では GNU Go level 10^{*6} (棋力が約 kgs8k) 及び, Pachi*7 の kgs3k と kgs2d の設定という, 棋力が知られている 3 つのプログラムを選んだ. 本稿では Pachi_{kgs3k} の実験だけを述べるが, 他の結果も概ね似た傾向である. 対戦相手に人間プレイヤーではな

*2 このニューラルネットワークは Leela Zero プロジェクト (<https://github.com/leela-zero>) とともに公開された.

*3 これは囲碁だけの話ではない. 例えばチェスでは Stockfish を強いプログラム役, Maia [9] を自然さ評価役とすることができる.

*4 予備調査により, 上位 25 位以内には, 適切であろう手が概ね 1 つ以上含まれていることを確認した. まれに上位 25 位以内に含まれない重要な手もあり, 5.4 節で例示する.

*5 この際, 最大リード a_{max} の計算には M_0 の全着手を見る.

*6 <https://www.gnu.org/software/gnugo/>

*7 <https://github.com/pasky/pachi>

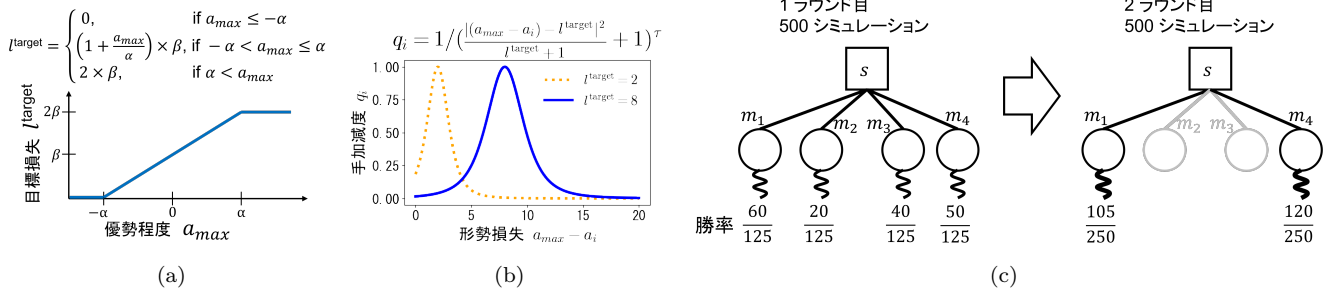


図 2: 先行研究 : (a) 目標形勢損失, (b) 手加減度, (c) Sequential Halving.

くプログラムを用いたのは、分析のために多くの対局を集めるためである。

既存手法 4 つ、(先行研究の) 提案手法 2 つを、それぞれ Pachi_{kgs3k} と対戦させた結果 [1], [2] をまとめる。19 路盤、日本ルール、コミ 6 目半、先後交代で 500 試合行った。

既存手法は下記の 4 つである。

- P1: 相手プログラムと同じもの。この場合、Pachi_{kgs3k}
- P2: Maia [9] のような教師あり学習の Prior
- P3: Softmax 法 [6]
- P4: 単純な方法で形勢が 0 目に最も近い着手を選ぶ

先行研究の提案手法は下記の 2 つである。

- P5: 役割分担の早期版 [1]*8
- P6: 3.1 節の自然な手加減度 + 3.2 節の halving [2]

P3~P6 は N_{sim} が 1,000 の KataGo に基づいたものである*9。また、人間プレイヤー同士の対局も野狐囲碁データセット*10 から集めて分析した。kgs3k におよそ該当する 1d の棋力ランク同士の対局を約 30 局抽出した。

4.2 要素 (a) きわどい勝負の結果

棋力が調整可能な P3, P5, P6 のパラメータを予備実験で決めた。P3 には Liu ら [6] と同様に R_{th} を 0.1 とし、 z で棋力を調整した。P5 と P6 には β で棋力を調整し、共通パラメータの α を 25.0, ϵ を 0.001 にした。P5 の γ を 0.4, P6 の τ を 2.0 にした。

P2 には $\text{Prior}_{\text{human}}$ を使い、選択確率 p_i が最も高い着手 i を打つようにした。しかしこれでは棋力の調整は不可能であり、kgs3k よりは強くなってしまふ。そこで我々は $(p_i)^\rho$ に比例して着手を選ぶという実装も行い、kgs3k に対しては $\rho = 3.0$ とした。この実装を以降 P2 ρ と呼ぶ。

これらの手法の設定及び要素 (a) きわどい勝負の結果を表 1 に示す。要素 (a) をさらに (a1) 勝率は互角か、(a2) 大差での勝ち負けを避けているかという 2 つの指標を分け

*8 KataGo に局面を探索させ、訪問回数が 10 以上の着手の中で、 $(p_i + \epsilon) \times \gamma^{e_i}$ 最大の手を選ぶ。 p_i と ϵ は 3.1 節と同じで、 e_i は着手 i の損失と目標損失の誤差、 γ は e_i の影響程度を決める定数である。

*9 P3 と P5 では内部のコミを 4 目楽観させるという工夫をした。

*10 <https://github.com/featurecat/go-dataset>

て評価した。(a1) について、P2 と P4 以外は勝率が 50% に近いという結果を得た。(a2) について、終局時の地合差が ± 10 目以内の対局の割合という指標を用いた。P4 は形勢 (地合差) が 0 目に最も近い着手を選ぶため、この割合が特に高いが、提案手法 (P5, P6) も大勝大敗をある程度抑制できていると言える。

4.3 要素 (b) 相手の出来不出来の反映の結果

先行研究では、勝敗を予測する分類問題におけるマッシュアップ相関係数 (MCC) を用い、相手 (仮想人間プレイヤー) の出来不出来をうまく反映できるかを評価した。出来が悪いのに勝たせたり、出来が良いのに負けさせた場合には低い値になり、これは指導碁としては好ましくない。「出来の悪さ」は、KataGo が解析した、1 試合中の着手の損失 (最善手と比べて損する目数) の平均値とした。

各手法の Pachi_{kgs3k} に対する最大 MCC を表 1 に示す。P4 は勝率が低いこともあり、相手の出来不出来を殆ど反映できなかった。P1, P2, P2 ρ は自分側も幅のある頻度と規模のミスをするため、相手の出来不出来にあまり関わらない勝敗結果となってしまっていることが分かる。P3, P5, P6 は比較的良好な結果を取めた。

4.4 要素 (c) 着手の筋の良さの結果

着手の筋の良さの評価には定量評価と定性評価両方を用いた。定量評価では 2 つの指標を用いたが、本稿では 15 目以上損する着手を打った割合だけを示す。P2 と P2 ρ の割合が最も高い。これらは探索なしに着手を決定するため、大悪手を打ってしまう場合があるようである。5.4 節で具体例を示す。P6, P3, P5 は強いプログラムの探索結果を用い、悪い手を抑制する機構があるため、割合は低くできている。P1 は弱いプログラム、P4 は相手の大悪手に大悪手で返すプログラムであるため、抑制が十分でない。

定性評価では高段者 5 人にいくつかの手法の筋の良さについて 5 段階評価をしてもらった。定量評価と異なり、P2 の評価が最も高く、大きな損失と筋の悪さは必ずしも一致しないという結果となった*11。次に評価が高いのは

*11 P2 は P5 よりはやや強いという理由の 1 つであろう。

表 1: 手法 P1~P6 の対 Pachi_{kgs3k} の棋力調整用のパラメータ設定及び要素 (a) きわどい勝負と要素 (b) 相手の出来不出来の反映の結果, \pm の後ろに 95% 信頼区間を示す.

棋力調整	(a) きわどい勝負		(b) 出来不出来の反映	
	(a1) 勝率	(a2) 最終地合差 ± 10 以内の割合	勝敗予測の MCC	
	50% に近いほどいい	高いほどいい	高いほどいい	
P1 (Pachi _{kgs3k})	-	47.8 \pm 4.4%	20.4 \pm 3.5%	0.240
P2 (Maia-like [9])	-	65.8 \pm 4.2%	18.4 \pm 3.4%	0.165
P2 ^{ρ} (Maia-like [9])	$\rho = 3.0$	52.0 \pm 4.4%	22.0 \pm 3.6%	0.189
P3 (Softmax [6])	$z = -0.5$	47.4 \pm 4.4%	53.4 \pm 4.4%	0.720
P4 (単純な方法)	-	7.4 \pm 2.3%	97.0 \pm 1.5%	0.022
P5 (提案手法早期版 [1])	$\beta = 3.0$	48.4 \pm 4.4%	59.4 \pm 4.3%	0.645
P6 (提案手法 3 章 [2])	$\beta = 3.2$	49.2 \pm 4.4%	56.6 \pm 4.3%	0.700
野狐囲碁 1d	-	-	7.1 \pm 9.5%	-

表 2: 手法 P1~P6 の対 Pachi_{kgs3k} の要素 (c) 着手の筋の良さ要素 (d) 着手の自然さの結果.

	(c) 着手の筋の良さ		(d) 着手の自然さ			
	15 目以上損 の着手の割合	専門家評価 1~5	直前手との距離	p' 算術平均	$p' > 0.9$ 着手を 打たない割合	専門家評価 1~5
	低いほどいい	高いほどいい	人間に近いほどいい	高いほどいい	低いほどいい	高いほどいい
P1 (Pachi _{kgs3k})	1.87%	3.0	3.627	0.381	0.857%	2.9
P2 (Maia-like [9])	3.03%	4.1	3.139	0.571	0%	4.3
P2 ^{ρ} (Maia-like [9])	3.32%	-	3.352	0.551	0.003%	-
P3 (Softmax [6])	0.22%	2.4	5.860	0.253	4.429%	1.8
P4 (単純な方法)	1.37%	-	5.949	0.281	4.531%	-
P5 (提案手法早期版 [1])	0.28%	3.5	3.711	0.468	0.588%	3.4
P6 (提案手法 3 章 [2])	0.21%	-	3.435	0.487	0.152%	-
野狐囲碁 1d	7.35%	3.3	3.426	0.364	0.589%	3.6

P5 で, 総合的には提案手法が概ね良いと言える考える.

4.5 要素 (d) 着手の自然さの結果

着手の自然さについても定量評価と定性評価を行った. 定量評価には「直前手とのユークリッド距離」「 p' 平均」「 $p' > 0.9$ の着手を打たない割合」の 3 つを用いた. 直前手との距離は人間のデータに近いほど自然に見えると考え. 1 章で述べた通り, 人間プレイヤーは局所的な戦いを終えてから別の戦いに移る傾向があり, これは直前手との距離にある程度反映されると考える. P3 と P4 の値は明らかに大きすぎ, あちこちに転戦する傾向であり, 不自然さを感じさせるであろう.

$Prior_{human}$ は人間プレイヤーの棋譜を学習したため, p' の平均が高いほど, 人間プレイヤーから着手が自然に見える予測する. 平均について算術平均と幾何平均両方を用いたが, 本稿では前者の結果だけ述べる. P2 と P2 ^{ρ} は直接 p' によって着手を選ぶため, この値が最も高いのは当然である. 次に高かったのは, p' を自然さ評価として用いる提案手法である. p' を用いたもう 1 つの指標は $p' > 0.9$ の着手を打たない割合である. $p' > 0.9$ の着手は人間プレイヤーが打つ可能性が非常に高いものであり, このような着手が放置されたら人間プレイヤーには不自然に感じるかもしれない.

い. 結果としては前 2 つの指標と同様な傾向を示した.

4.4 節と同様に, 囲碁の高段者 5 人にいくつかの手法の自然さに対して 5 段階評価をしてもらった. 同様に P2 が最も好評であり, 提案手法の P5 は次に高かった.

ここまで 4 つの要素計 9 つの評価指標の結果を示した. 要素 (a)~(d) を総合的に見たら, 提案手法は自然な手加減を行う (好ゲーム演出) には最も優れていると考える.

4.6 自然な手加減の例

本節では自然な手加減がよくできていると考える例を 3 つ紹介する. 図 3a (白番) は P5 (提案手法早期版 [1]) 対 Pachi_{kgs3k} の 1 局の序盤で, 専門家により自然さ 5 点満点と評価されたものである. 50 手目時点で P5 が +2 目程度のリードに留まっており, 手加減できていることが分かる. また, 展開も自然に見え, 不自然な着手は特になかった.

図 3b (黒番) は P6 (提案手法 [2]) 対 GNU Go の 1 局の序盤である. 全体的に自然な展開になっており, 例えば 49 手目 B18 ($p' = 0.10$) や 51 手目 D9 ($p' = 0.22$) は, 自然かつ穏やかに損する手となっている. どちらも $Prior_{KataGo}$ の選択確率はかなり低く (それぞれ 0.0002 と 0.0006), $Prior_{human}$ によって初めて候補に入った手である.

図 3c (白番) は同じく P6 対 GNU Go の対局である.

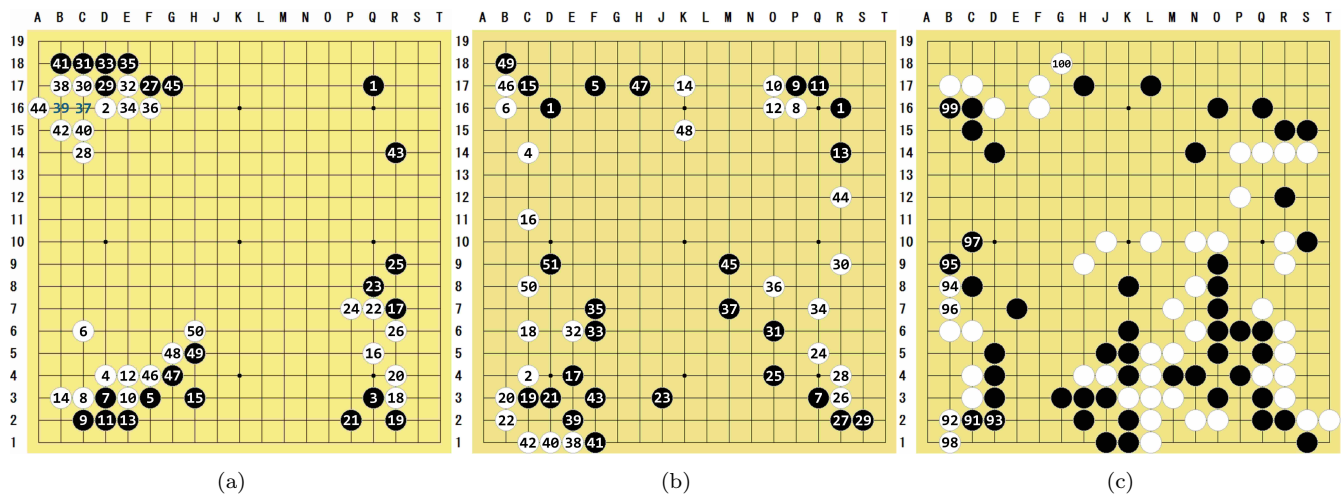


図 3: 自然な手加減がよくできている例.

94~98 手目は相手（黒）を固くしたうえで後手なので計 10 目ほど損する。100 手目 G18 も相手の 99 手目 B16 に反応したような着手で 3 目ほどの損であるが、どれも穏やかで不自然には見えにくい手加減になっている。

5. 残っている課題

4 章で示した通り、先行研究の手法は好ゲームの演出が概ねできている。しかし、棋譜を確認した際に課題もいくつか見つけた。本章ではこれらの課題を、攻撃的な手 (5.1 節)、損しすぎる手 (5.2 節)、小さすぎる手 (5.3 節)、 $Prior_{human}$ の誤りによる手 (5.4 節)、不自然な手 (5.5 節) に大別し、具体例と可能な解決策を述べる。

5.1 攻撃的な手

提案手法は攻撃的な手を打つことがしばしばある。適度な戦いは練習のためにも楽しさのためにも必要であるが、初中級者を難しすぎる戦いに導いたり、畏にはめるような手は抑制すべきと考える。

図 4a (白番, 9 目リード) はその一例である。88 手目白 G12 は自然な着手 ($p' = 0.34$) で、最善手 S8 に比べ適度に (4.2 目) 損する手であるが、初中級者向けの手加減には適切でないかもしれない。大きく包囲された黒 2 子が逃げ出し、損を重ねて好ゲームが台無しになるかもしれないためである。G12 よりも、小さく取る F13 ($p' = 0.13$) や単なる大場である M3 のほうが適切かもしれない。

図 4b (黒番) の 35 手目 Q17 ($p' = 0.56$) は右上の白石を攻める着手で自然であるが、白 34 手目 O18 に応対する意味でも N17 ($p' = 0.14$) で受けるほうが穏やかではある。

なお、攻撃的な手そのものが良くないわけではない。局面は手数進んで、白が 46 とやや薄くトンド局面で、黒 47 手目は P18 ($p' = 0.02$) であった。これは白の目を奪い、明示的に活きを促す手である。白はこの場所を活かれば形勢

互角に近く、「生きてください」という黒 47 は攻撃手でありながら、むしろ穏やかな局面へと導く手でもある。

着手が攻撃的であるかは KataGo の盤面占有率 (ownership $\in [-1, 1]^{19 \times 19}$, 各地点が最終的に相手の地か自分のか) を用いてある程度判別できる。実際我々は、着手後の盤面占有率と着手せずパスした後の盤面占有率の差分をもとに、相手の石をどのくらい弱化したかを分析したことがある [14]。しかし、乱戦を導くような手、初中級者の間違いを誘発するような手の判断には他の方法が必要であろう。

5.2 損しすぎる手

提案手法は手加減を行うとはいえ、初中級者にも分かるほど損しすぎる手を打ってしまうこともあり、これはやる気をそぐ結果につながりうる。例えば図 4c (黒番) の 23 手目 R8 ($p' = 0.08$) はそのような着手である。R8 は KataGo の最善手 S4 ($p' = 0.31$) に比べ 5.7 目損である。目数差以上に、切り違いの要石が取られるのは敗着に近く感じられ、手加減とはいえ適切でない。この局面は S4 の他は R6 ($p' = 0.54$) が適切であろう。

損しすぎの判定には、単に損失の目数だけではなく、対局の段階 (序盤・中盤・終盤) なども考慮する必要がある。例えば同じ 5 目の損失は序盤では大きいですが、大ヨセでは許容できるかもしれない。あるいは大石の戦いの場合に、活きるかどかで 10 目以上の差が付いていることがあってもおかしくない。目標損失式 (図 2a) になんらかの補正を入れることで、この課題に対処できると考える。

5.3 小さすぎる手

前節の損しすぎる手と似ているが、提案手法は小さすぎる手を打ってしまうこともある。例えば図 4d (白番) の 172 手目 D15 ($p' = 0.24$) はそのような着手である。KataGo の最善手 G14 ($p' = 0.03$) に比べて損失の程度は

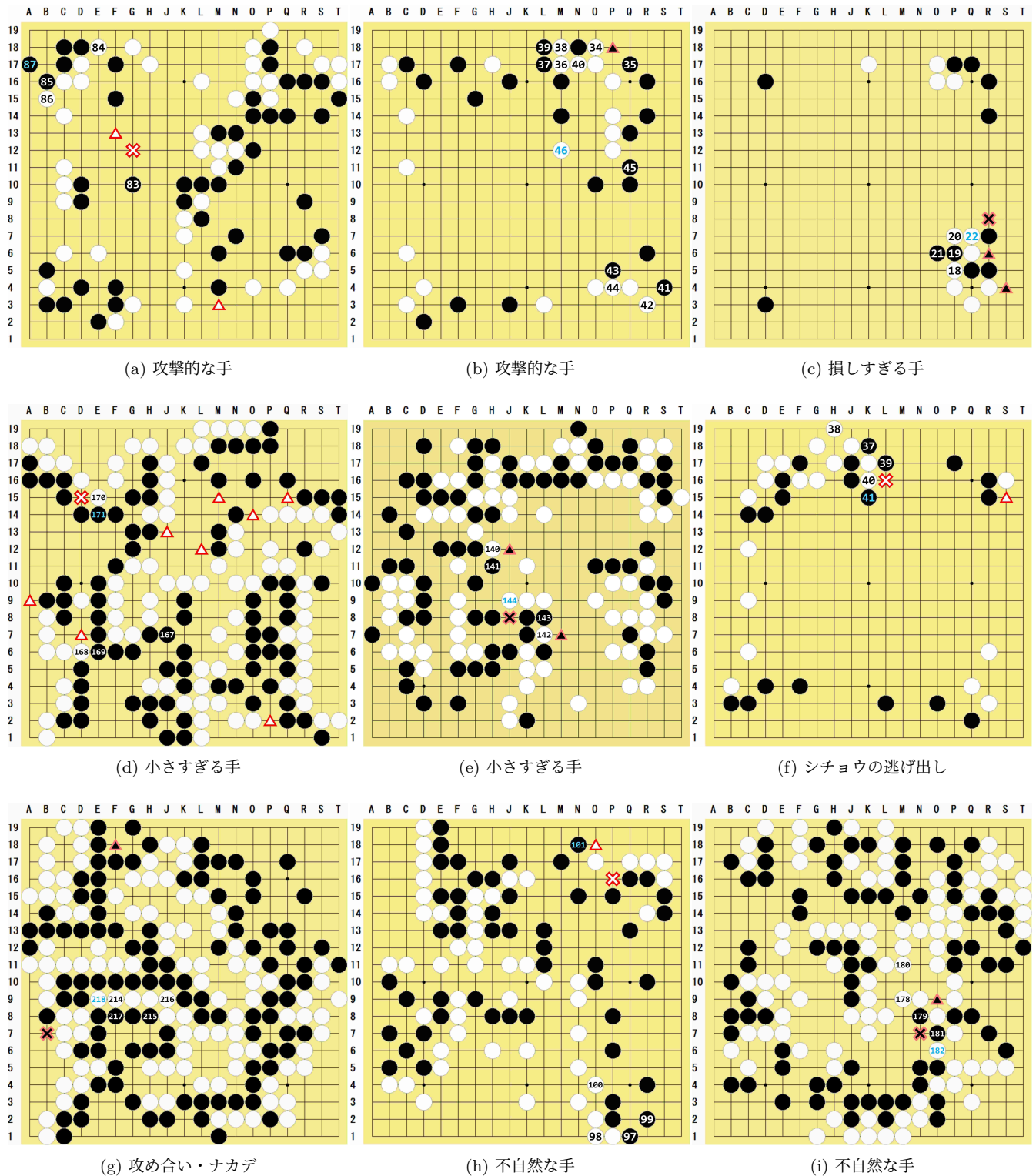


図 4: 課題を説明するための局面例 (ハツは課題となっている着手).

4 目程度と適切かもしれないが、これはパスと同等、いわゆるゼロ目の手である。この局面において 1 目以上の価値がある手はいくつか残され (三角の位置など)、D15 のような小さすぎる手を打つことは、手本になるという意味からも適切でない。

このようなパスとあまり変わらない小さすぎる手は、パ

スの場合のリード目数と比較すれば判別できる。実際こういった指標が用いられ、終盤での無駄手がどの程度あるかを分析した研究がある [15]。

図 4e (黒番) の 145 手目 J8 ($p' = 0.10$, +31.1 目優勢) はもう 1 つの例である。この局面では左の白石を全滅させており、ここまでの展開が悪かったかもしれない。KataGo

の最善手 J12 ($p' = 0.26$, +44.0 目優勢) に比べて J8 は 13 目ほど損し、拮抗状態に近づこうとしているが、もう数十目勝ちの局面なので、負けることを諦めて勝つ方が良いかもしれない。なお、J8 はパスに比べれば 5.7 目の価値はある手であり、図 4d ほどは解決策が単純でない。最善手の価値（この場合 19 目ほど）との比率や、形勢を考慮するなどが必要だと考える。

5.4 $Prior_{human}$ の誤りによる悪手

$Prior_{human}$ の値は人間の感覚にかなり近いが、教師あり学習という方式上、いくつかの誤りもある。誤りの 1 種類目はシチョウの逃げ出しである。図 4f (白番) の 40, 42 手目は損な逃げ出しであり、初中級者でも少し読めば分かる手である。42 手目で KataGo の最善手は S15 であるが、 $Prior_{human}$ p' の値は 0.004 と低い。実戦の L16 は $p' = 0.96$ と非常に高い値になってしまっており、かつ 5.1 目の損であるため、提案手法からすると「自然に数目手加減できる、良い候補」ということになってしまっている。

誤りの 2 種類目は攻め合いやナカデなどの見落としである [16]。図 4g (黒番) 219 手目はそういう見落としである。提案手法は B5 ($p' = 0.03$) に打ったが、現状は白 218 に切られたところであり、F18 に打って大石を生きるのが自然であろう。F18 の $Prior_{KataGo}$ は 0.43 と高いため、探索の候補には入っているのであるが、 $Prior_{human}$ は 0.002 に過ぎず、つまり提案手法にとって「手加減もできず、不自然である手」となってしまっている。

このような誤りを解決あるいは抑制するには、 $Prior_{KataGo}$ とブレンドすること [11] が有効かもしれない。 $Prior_{KataGo}$ の確率を p とし、ブレンドというのは $p^\omega \times p'^{(1-\omega)}$ で新しい確率分布を作ることである。図 4f 図 4g とも打つべきの着手の p が一定程度ある（それぞれ 0.78 と 0.43）ため、 $Prior_{human}$ の過小評価をある程度抑制できる。

5.5 不自然な手

提案手法は、手加減をするために、自然に見える良い手を打たずに不自然な手を打ってしまうことがある。例えば図 4h (白番) の 102 手目 P16 ($p' = 0.05$, 9.8 目損) はそのような着手である。KataGo の最善手 O18 ($p' = 0.87$) は、右上白への攻撃を防ぐ着手であり、直前手 N18 の応手としても自然であるが、大差の碁を無理に互角に近づけようとして悪い手を打ってしまっている。

もう 1 つの例として、図 4i (黒番) の 183 手目 N7 ($p' = 0.09$, 4.7 目損) がある。これも細かいミスをわざとしたありがちな手であり、当然 O9 に抜くのが自然である。

このような不自然な手の原因は複合的で、解決のための手段もこれまで述べてきたものを組み合わせる必要があると考える。人間が不自然と感じる理由はさまざまであり、それを $Prior_{human}$ だけで判断するのは難しいであろう。

6. おわりに

本稿ではまず、初中級者相手に好ゲームを演出するための、役割の異なるプログラムを組み合わせる先行研究を要約した。囲碁において、強いプログラム KataGo に局面と着手の正確な評価を、人間棋譜を学習したニューラルネットワークに着手の自然さを評価させることで、好ゲームの演出（自然な手加減）が一定程度達成された。

しかし、この手法にもまだ課題は多い。本稿ではそれらの課題を具体例を用いて説明し、可能な解決策も提案した。今後の展望として、これらの解決策を実現し、KGS や幽玄の間などの囲碁サーバに投入することで人間プレイヤーからの評価を受けたい。

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

参考文献

- [1] Hsueh, C.-H. and Ikeda, K.: Playing good-quality games with weak players by combining programs with different roles, *IEEE CoG 2022* (2022).
- [2] Hsueh, C.-H. and Ikeda, K.: Improvement of move naturalness for playing good-quality games with middle-level players, *Applied Intell.*, Vol. pre-press (2024).
- [3] Wu, D. J.: Accelerating self-play learning in Go, *AAAI-20, RLG Workshop* (2020).
- [4] Leela Zero: A network trained from human games, GitHub (online), available from (https://sjeng.org/zero/best_v1.txt.zip) (accessed 2024-02-09).
- [5] Sephton, N., Cowling, P. I. and Slaven, N. H.: An experimental study of action selection mechanisms to create an entertaining opponent, *IEEE CIG 2015* (2015).
- [6] Liu, A.-J. et al.: Strength adjustment and assessment for MCTS-based programs [research frontier], *IEEE Comput. Intell. Mag.*, Vol. 15, pp. 60–73 (2020).
- [7] Nakamichi, T. and Ito, T.: Adjusting the evaluation function for weakening the competency level of a computer shogi program, *ICGA J.*, Vol. 40, pp. 15–31 (2018).
- [8] Rosemarin, H. and Rosenfeld, A.: Playing chess at a human desired level and style, *HAI'19* (2019).
- [9] McIlroy-Young, R. et al.: Aligning superhuman AI with human behavior, *KDD'20* (2020).
- [10] Jacob, A. P. et al.: Modeling strong and human-like gameplay with kl-regularized search, *ICML 2022* (2022).
- [11] Ogawa, T., Hsueh, C.-H. and Ikeda, K.: Improving the human-likeness of game AI's moves by combining multiple prediction models, *ICAART 2023* (2023).
- [12] Karnin, Z., Koren, T. and Somekh, O.: Almost optimal exploration in multi-armed bandits, *ICML 2013* (2013).
- [13] Fabiano, N. and Cazenave, T.: Sequential halving using scores, *ACG 2021* (2021).
- [14] シュエジュウシュエン, 池田 心: 囲碁布石問題における不正解手の自動生成, *GPW 2023* (2023).
- [15] Kuboki, K., Hsueh, C.-H. and Ikeda, K.: Factor analyses on positive and negative evaluations of games against Go programs, *TAAI 2023* (2023).
- [16] 池田 心, ビエノシモン: 行動評価関数を用いたモンテカルロ木探索の重点化と見落としの抑制, 情報処理学会論文誌, Vol. 55, pp. 2377–238 (2014).