

Title	人間プレイヤーの着手を誘導する囲碁AI に向けて
Author(s)	シュエ, ジュウシュエン; 窪木, 響大; 池田, 心
Citation	情報処理学会第57回GI研究発表会, 2026-GI-57(8): 1-8
Issue Date	2026-02-23
Type	Conference Paper
Text version	author
URL	https://hdl.handle.net/10119/20618
Rights	社団法人 情報処理学会, シュエ ジュウシュエン, 窪木 響大, 池田 心, 情報処理学会第57回GI研究発表会, 2026-3, 2026. 正誤表追加版. 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.
Description	情報処理学会第57回GI研究発表会, 早稲田大学, 2026年3月2日-3日

人間プレイヤーの着手を誘導する囲碁 AI に向けて

シュエ ジュウシュエン^{1,a)} 窪木 響大^{1,b)} 池田 心^{1,c)}

概要: 多くのゲームにおいて、AI プレイヤーは人間のトッププロを強さの面で凌駕している。近年は、単なる強化や汎用性向上とは別に、人間を楽しませたり教えたりすることを目的とした AI も注目されている。先行研究では、手加減の度合いと自然さのバランスを取る役割分担により、好ゲームを演出する接待基 AI が提案された。本稿では、この手法をもとに、相手プレイヤーの着手を“誘導”する手法を提案する。例えば、相手（生徒役）の着手における期待目数損失を計算し、その値が小さくなる着手を自分（先生役）が選びやすくすることで、「相手が自然に良い手を打てる基」への展開誘導が可能になると考える。逆にすれば、「相手が失敗しやすい歯ごたえのある基」への誘導も可能になると考える。本稿では、この期待目数損失に基づく誘導を一例として実装・評価する。相手着手の選択確率推定には KataGo の HumanSL ネットワークを用い、誘導対象として囲碁プログラム Pachi を用いた。誘導の方針と強度を制御するパラメータを 0.6（相手が自然に良い手を打てる）、0.8, 1.0（誘導なし）、1.2, 1.4（相手が失敗しやすい）と設定した。その結果、相手の平均目数損失は 1.805, 2.014, 2.310, 2.907, 3.160 と単調増加し、誘導効果が確認できた。

1. はじめに

近年、多くのゲームにおいて、AI プレイヤーは人間のトッププロを凌駕する強さを示している。その代表例として、囲碁における AlphaGo, StarCraft II における AlphaStar, 麻雀における Suphx が挙げられる。このような背景のもと、単に勝つことを目的とする AI だけでなく、人間プレイヤーを楽しませたり、学習を支援したりすることを目的とした研究も注目を集めている。

そのような人間中心の AI 研究の一例として、人間の実力や状態に応じてゲーム体験を調整する研究がある。例えば、動的難易度調整では、プレイ中に観測される行動・成績などのパフォーマンス指標に基づき、敵の強さや出現頻度、課題の難しさなどを変化させることで、過度に簡単でも困難でもない体験の実現が目指されている [1], [2]。また、対戦型ゲームにおいては、AI プレイヤーが意図的に手加減を行うことで、人間にとって楽しめる対局を実現しようとする研究もある [3], [4], [5]。しかし、多くの場合、AI 側の振る舞い設計に主眼が置かれており、「相手（手加減される側）がどのような着手をしやすくなるか」という観点から対局の展開そのものを積極的に設計する枠組みは、十

分に検討されていない。

仮に、対戦相手の棋力や弱点、棋風や好み、さらには「学びたい」「楽しみたい」といった対戦目的がある程度分かっているとすると、それに応じて対局の展開を誘導できれば、上達支援や娯楽性向上に有用であると考えられる。例えば、プレイヤーの強みが発揮されやすい展開、課題が露呈しやすい展開、あるいは戦いが好きなプレイヤーには乱戦となる展開へと導くことができれば、上達支援やモチベーション向上にもつながる。ここで重要なのは、例えば戦いの多い展開に誘導したいとき、たとえ AI 側の着手だけを見たときに「好戦的」「攻撃的」な手を打ったとしても、もし人間プレイヤー側に戦いを避けられてしまえば、展開の誘導にはならないという点である。したがって、AI 側は自分の手だけを考えず、“人間プレイヤーが戦いの手で応じたくるような”手を選ばなければならない。

本稿では、この考え方を囲碁の指導対局に適用し、相手プレイヤーの着手を誘導することで対局の展開を制御する手法を提案する。さらに、その一例として、相手の期待目数損失に基づく誘導を実装し、囲碁プログラムを用いた実験によりその有効性を検証する。

2. 関連研究

本稿では、「どの手が人間にとって打ちたい手、あるいは目につきやすい手で、どの手がそうでないのか」を推測する機能を用い、指導対局の文脈において対局の展開を誘導する。このような人間の着手傾向を推定する機能は、人

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) s2460002@jaist.ac.jp

c) kokolo@jaist.ac.jp

間の着手を模倣する研究と関連が強いため、まず 2.1 節でそれを紹介する。また、指導対局における好ゲームの設計という観点では、従来の強さ調整や手加減研究の系譜に位置づけられるため、2.2 節ではそれらを概観する。最後に 2.3 節では、ゲームにおける人間プレイヤーの行動や選択を誘導する試みについて整理する。

2.1 人間アマチュアの模倣

人間アマチュアの着手傾向を再現することを目的とした研究として、棋譜データに基づいて人間の着手分布を予測するモデルが提案されている。チェス分野では、McIlroy-Young らにより、アマチュア棋譜を用いて異なる棋力帯ごとに学習した着手予測モデル群 Maia が提案されている [6]。この研究では、学習した棋力帯に近い対局を対象とすると、そうでない場合と比べて人間の着手をより良く再現できることが示されている。その後、棋力情報を入力として利用する Maia-2 が提案され、幅広い棋力帯の着手を単一モデルで予測可能であることが報告されている [7]。

囲碁分野においても、同様の考え方に基づくモデルとして KataGo の HumanSL ネットワークが公開されている [8]。HumanSL は、盤面に加えて対局者の棋力などのメタ情報を入力とし、人間の着手確率分布を出力するモデルであり、アマチュアからプロまでを対象として学習されている。本研究では、人間アマチュアの着手傾向を推定する手段として、この HumanSL を用いる。

2.2 対戦 AI における強さ調整

対戦型 AI プレイヤーにおける強さ調整や手加減は、人間にとって楽しめる対局や練習相手として適切な強さを実現するうえで重要な研究テーマであり、長年にわたり取り組まれてきた。モンテカルロ木探索系プログラムに対しては、行動選択方策を操作することで強さや対戦体験を調整する試みが報告されている。Sephton らは確率的な着手選択を導入した対戦相手生成を検討しており [9]、Liu らは低品質な着手を抑制する仕組みを加えた強さ調整とその評価手法を整理している [3]。

また、Nakamichi らは、将棋プログラムの評価関数を調整することで、段階的な棋力弱化を実現する手法を提案している [10]。Fujita は学習済み AlphaZero 系モデルをもとに、人間の練習相手として適切な強さを実現するための複数の調整戦略を比較している [4]。

一方で、単なる勝率や棋力の調整にとどまらず、対局の質や自然さに着目した研究もある。我々は中級者との対局を対象に、囲碁プログラムの役割分担を導入し、好ゲームを演出する接待碁 AI を提案した [5]。具体的には、着手の良さを評価する役割と、人間らしさを評価する役割を分離し、それらを組み合わせて着手を選択することで、自然な手加減対局を演出している。

2.3 ゲームにおける人間プレイヤーの誘導

本節では、「人間プレイヤーの行動傾向を考慮し、特定の行動や展開が生じやすくなるように設計すること」を、広義の人間プレイヤーの誘導として捉える。例えばダンジョン探索型の RPG において、見える範囲に宝箱を置き、その宝箱に到達すると別の宝箱が視界に入り、結果的にプレイヤーの行動を特定の場所にほぼ明示的に誘導するなどはしばしば行われる。Fujihira らの研究は、迷路の交差点における人間の行動傾向をモデル化してテストプレイヤーとして用い、迷いやすい（間違っただけに誘導されやすい）迷路、迷いにくい迷路を作り分けることに成功している [11]。これは隠れた誘導の例と言える。一方、本研究では、環境や盤面を事前に生成するのではなく、対戦中に人間プレイヤーの着手に応答しながら AI が着手を選択するという相互作用の中で、展開を誘導する点が異なる。

対戦ゲームの分野では、相手の行動や戦略をモデル化し、それに基づいて自身の行動を適応させる研究が行われてきた [12], [13]。これらの研究は、相手が誤りやすい選択を予測してそれを利用するという意味では、我々が論じる「人間プレイヤーの誘導」と共通している。しかし、これらの研究は勝利の最大化を主目的としており、相手の行動分布そのものを学習や体験の観点から設計することは想定していない。本研究では、勝敗ではなく、指導対局における展開や学習機会の設計を目的とする点が異なる。

3. 本研究における誘導の枠組み

本章では、本研究で扱う誘導の概念を説明し、その実現に必要な要件を整理する。この概念は、我々の発表済論文 [14] において囲碁の指導対局を想定して予備的に述べたものであるが、本稿ではより一般化した形で再構成する。

3.1 対象とするゲームおよび誘導の目的

誘導の概念を説明する前に、本研究が対象とするゲームの範囲を明確にする。本研究では、完全情報交互ゲームを対象とする。この条件を満たすゲームであれば、相手の着手および対局展開の誘導が可能であると考えられる。説明の簡潔さのため、本稿ではさらに「2 人ゼロサム確定離散」ゲームに限定する^{*1}。

本研究は、勝つことを目的とした研究ではなく、対戦相手に「楽しめた」「成長した」など、何らかのゲーム体験を与えることを目的としている。そのため、対戦相手として人間プレイヤーを想定し、その実力や弱点、プレイスタイル、好みなどがある程度把握できている状況を仮定する。

^{*1} プレイヤー人数については、誘導側を含めて 3 人以上でも原理上は誘導可能であるが、相手が複数になるため議論が複雑になる。ゼロサム性は誘導の本質には関係しないが、説明上扱いやすいため採用する。また、確率性や連続性は原理的には問題にならないが、本研究では先読み計算を前提とするため、行動空間が発散しない設定を想定する。

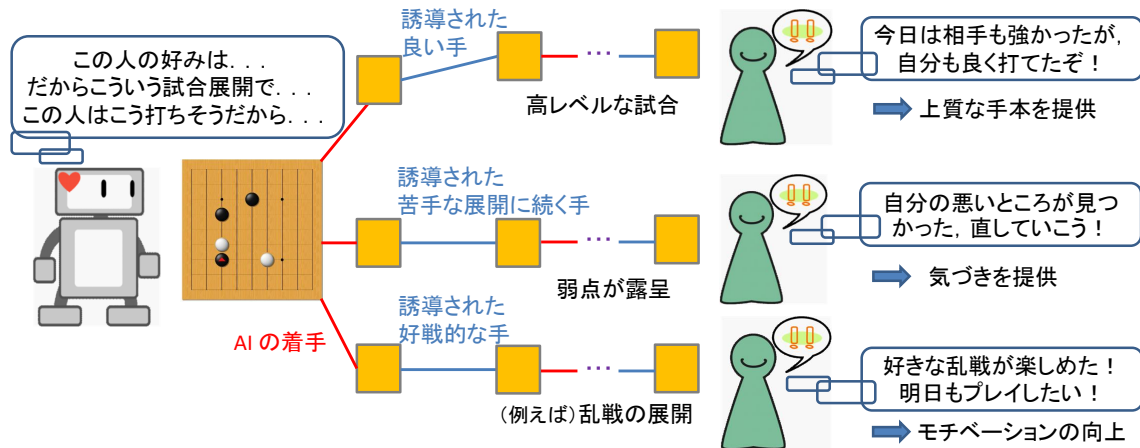


図 1: 誘導の達成イメージ

誘導の達成イメージを図 1 に示す。例えば、相手の強みが発揮されやすい展開へ誘導できれば、高水準の対局を実現し、良質な手本を提供できる。逆に、相手の課題が露呈しやすい展開を作ることで、気づきや上達の方向性を示せる可能性がある。また、戦いを好むプレイヤーに対して乱戦の展開を誘導できれば、対局の満足度やモチベーションの向上につながる。さらに、商業的な観点では、「ゲームから離れるような体験を避けさせる誘導」や「課金アイテムを使いたくなる状況への誘導」といった応用も想定できる。

3.2 誘導に必要な技術要件

本研究の誘導を実現するために必要な要件を次に示す。

(1) **着手の人間らしさ評価器**：最低限 1 つの評価器が必要であり、特に指導対局の文脈では以下の 2 種類に分けることが望ましい。

- **着手の相手らしさ評価器**：相手が打ちそうかどうかを評価する。個人専用モデルが困難な場合は、棋力や棋風などの属性ごとの集団モデルで代替できる。例えば、3 級プレイヤーを想定するなら、3 級の着手分布を予測するモデルを用いる。
- **着手の自然さ・形良さの評価器**：主に誘導側の手の自然さを評価する。人間相手との対局では、着手が自然に見えるかどうか体験品質に影響するためである。また、指導対局では、手本として機能することが望ましく、相手より強い視点からの評価が必要となる。例えば、3 級の相手に対しては、低段者程度のモデルを評価器として用いれば、自然でもあり形も手本になるような手を候補にすることができる。少し強いプレイヤーとの対局が学習効果を高めるといふ指摘もある [15]。

(2) **好ましき関数**：誘導側の着手は、次の 2 種類の関数を組み合わせて決定する。

- **誘導側手番の各着手に対する、プレイ目的ごとの好ましき関数**：通常対局では勝率や minimax 評価値を用いる。指導対局では、形勢損失と自然さを組み合わせ

た指標 [5] が利用できる。

- **相手手番の各局面に対する、誘導目的ごとの好ましき関数**：例えば、相手の強みを発揮させたい場合や課題を露呈させたい場合には、相手の期待形勢損失を指標として用いることができる。この値が小さい局面ほど相手は良い手を打ちやすく、大きい局面ほど失敗しやすいと解釈できる。別の誘導目的として、乱戦への誘導を考えるなら、戦いの発生度を評価する関数が必要になる。囲碁を例にとると、中央志向や実利志向といった棋風に合わせた誘導には、中央・辺隅それぞれで形成される見込み地の評価関数が必要になる。

図 2 を用い、「着手の相手らしさ評価器」が不可欠である理由を説明する。これは我々の予備実験で実際に観測した現象である。誘導側が好戦的な手を選択しても、人間プレイヤーがそれに応じなければ、戦いの展開にはならない場合がある。逆に、平和な着手を選んだつもりでも、相手にとって戦いの手が選びやすければ乱戦になることがある。この事象は、誘導側の手だけを評価し、相手が選びやすい行動を考慮していなかったことに起因する。したがって、人間プレイヤーの選択確率を明示的にモデル化することが、誘導の実現には不可欠である。以上より、人間プレイヤーの着手傾向をモデル化し、それを誘導設計に組み込む枠組みが必要である。次章では、これらの要件を満たす具体的な実装方法を述べる。

4. 囲碁指導対局における誘導実装の一例

本章では、3 章で述べた誘導概念に基づく具体的な実装例を示す。対象として囲碁の指導対局を考え、相手プレイヤーの強みを発揮させる場合、および課題を露呈させる場合の誘導を想定した。説明の簡潔さのため、誘導対象は棋力 3 級 (3k) の人間プレイヤーとしたが、本手法は他の棋力や、棋風・好みといった属性を対象とする場合にも、そのモデルさえあれば適用可能である。

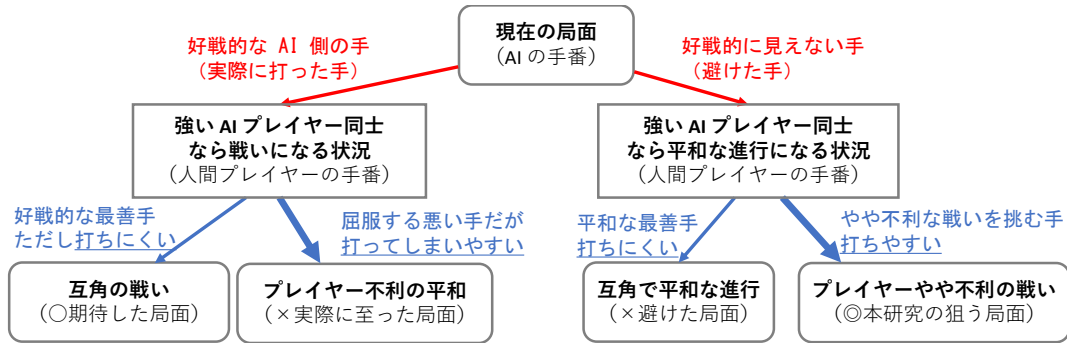


図 2: 好戦的な AI 側の手が、結果的に戦いを導かなかった例

4.1 着手の人間らしさ評価器の実装

「着手の相手らしさ評価器」として、KataGo HumanSL モデルの rank_3k 設定を用いた。また、「着手の自然さ・形良さの評価器」として、同モデルの rank_1d 設定を用いた。これは、初段 (1d) の着手が 3 級にとってやや高度である一方、理解可能であり、指導対局における手本として機能すると考えたためである。ただし、モデル作成にあたりデータ数が潤沢でないゲームで実装する際には、この 2 つの機能分担をそれほど意識しなくてもよいかもしれない。

4.2 好ましさ関数の実装

4.2.1 誘導側着手の好ましさ評価

誘導側手番における各着手の好ましさ評価には、我々が提案した、形勢損失と着手の自然さを組み合わせた指標 [5] を用いた。理論上は全合法手を候補とすればよいが、囲碁では合法手数が多く、探索コストが過大となるため、本稿では以下のいずれにも含まれない着手を候補から除外した。

- rank_1d による選択確率上位 25 手 (自然な着手が含まれる可能性が高い)
- KataGo 自己対戦により学習された強いネットワークの選択確率上位 5 手 (最善手が含まれる可能性が高い)
- パス

これらの各着手について、KataGo に N_{sim} 回訪問の探索を行い、得られた統計量をもとに着手 i の自然な手加減度 s_i を評価する (計算方法は付録 A.1 に示す)。

4.2.2 相手期待目数損失の計算および誘導補正

相手手番における誘導目的を評価するため、相手の着手を探索する。各誘導側着手 i に対し*2、相手の候補着手集合 m_i^{opp} を以下 3 つの集合の和集合に限定した。これも、探索コストを抑えるための妥協案である。

- rank_3k による選択確率上位 4 手 (相手が打ちやすい)
- KataGo 自己対戦により学習された強いネットワークの選択確率上位 1 手 (勝つために有力な着手である可能性が高い)
- パス

*2 実際には、さらにコストを抑えるために、ベース好ましさ s_i が上位の着手に限るなどの実装も有望である。

各着手について、KataGo に N_{sim} 回訪問の探索を行い、得られた統計量を用いて「相手の期待目数損失」を計算する。まず、相手着手 j の rank_3k 選択確率 p_j^{3k} を、集合 m_i^{opp} 内で正規化する: $p_j^{3k'} = p_j^{3k} / (\sum_{x \in m_i^{opp}} p_x^{3k})$ 。誘導側着手 i に対する相手の期待目数損失 l_i^{opp} を

$$l_i^{opp} = \sum_j (p_j^{3k'} \times (\max_{k \in m_i^{opp}} (a_k) - a_j)) \quad (1)$$

と定義した。この値が大きいくほど、相手が誤りやすい展開であると解釈できる。正規化を行わずに期待値を計算すると、探索していない着手の損失を 0 とみなすことになり、損失を過小評価する恐れがあるため、本手法では正規化を行った。ただし、この定義にはいくつかの課題が残されている: (1) m_i^{opp} に最善手が含まれていない場合、「期待損失」とは言いにくいこと、(2) m_i^{opp} 外の着手の損失は実際にはこの値よりも大きく、それらを無視することで損失を過小評価している可能性があることである。

最後に、相手の期待目数損失 l_i^{opp} と自然な手加減度 s_i を組み合わせ、「誘導入り自然な手加減度」 g_i を

$$g_i = s_i \times (c_g)^{l_i^{opp}} \quad (2)$$

と定義した。ここで、 $c_g > 0$ は誘導の強度および方向性を制御するパラメータである。 $c_g \in (0, 1)$ の場合、相手の期待目数損失が大きいくほど g_i は小さくなり、相手が誤りやすい展開を避ける方向に誘導が働く。一方、 $c_g > 1$ の場合、相手の期待目数損失が大きいくほど g_i は大きくなり、相手が誤りやすい展開を積極的に選択することになる。

5. 囲碁プログラムを用いた対戦実験

本章では、囲碁指導対局における誘導手法の挙動を、対戦実験により評価する。本研究は人間プレイヤーの誘導を目的としているため、本来は人間プレイヤーとの対局による評価が望ましい。しかし、パラメータの調整や、統計的な傾向の確認には多数の対局が必要であり、対人戦のみで十分な試行数を確保することは難しい。

そこで本稿では、人間プレイヤーの代替として、囲碁プログラム Pachi を用いた対戦実験を行う。具体的には、棋力

表 1: 誘導パラメータ c_g に対する、誘導側の統計結果

c_g	誘導側の			
	誘導手 頻度	勝率	目数損失 平均	rank_1d 幾何平均
0.6	9.18%	39.5±6.8%	1.729	0.344
0.8	4.62%	44.5±6.9%	1.928	0.355
1.0	0%	57.5±6.9%	2.166	0.351
1.2	4.85%	57.0±6.9%	2.666	0.323
1.4	8.47%	44.5±6.9%	2.851	0.288

が kgs3k となる設定の Pachi を誘導対象とする^{*3}。Pachi はモンテカルロ木探索を基盤とし、シミュレーションに人間知識を用いることから、その着手は一定程度人間らしい性質を持つと考える。なお、本実験は人間プレイヤーを完全に再現することを目的とするものではなく、誘導パラメータの影響を相対的に比較するための評価である。

5.1 実験設定

各候補手および相手着手の探索におけるシミュレーション回数は、いずれも $N_{sim} = 20$ とした。この設定では、1 手の選択あたり約 3,000–4,000 回のシミュレーションが行われる^{*4}。「自然な手加減度」の計算に用いるパラメータ $\alpha, \beta, \tau, \epsilon$ の値は、それぞれ 25.0, 5.0, 2.0, 0.0005 とした。 α および τ については、先行研究 [5] の設定値を援用している。一方、本研究では「着手の自然さ・形良さの評価器」として KataGo HumanSL の rank_1d 設定を用いており、先行研究とは評価器が異なるため、手加減量に影響する β の値は予備実験により調整した。

誘導パラメータ c_g については、0.6, 0.8, 1.0, 1.2, 1.4 の 5 通りを設定した。 $c_g = 0.6$ および 0.8 は相手を間違えさせない方向への誘導、 $c_g = 1.0$ は誘導なし、 $c_g = 1.2$ および 1.4 は相手を間違えさせる方向への誘導を意図している。各 c_g 設定について、Pachi_{kgs3k} と先後交代で 200 局ずつ対戦を行った。対局条件は 19 路盤、日本ルール、コミ 6.5 目とした。勝敗は対局終了後に独立した KataGo により最終局面のリード目数から判定した（ただし、Pachi_{kgs3k} が投了した場合はその結果を用いた）。

5.2 実験結果

表 1 に、誘導パラメータ c_g に対する、誘導補正により着手が変化した頻度「誘導手頻度」を示す。 c_g が 1.0 から離れるほど、誘導手頻度が高くなっており、誘導補正が実際に着手選択へ反映されていることが分かる。相手役着手

^{*3} <https://github.com/pasky/pachi> の README において、Fixed Strength の設定例として、`pachi -t=5000 --nodcnn` が kgs 3k 相当として示されている。

^{*4} Core Ultra 9 285K と Geforce RTX 5090 を搭載したマシンでの考慮時間は 2.6 秒であった。もう少し古いマシンでは 8.5 秒と、少し人間の相手をするには反応が遅い程度の時間がかかる。実用上では、脚注 *2 で述べた計算量削減の方法は重要である。

表 2: 誘導パラメータ c_g に対する、相手役 (Pachi_{kgs3k}) 着手の誤り傾向に関する統計結果

c_g	相手役 (Pachi _{kgs3k}) 着手の		
	目数損失 平均	≥10 目損失 着手の頻度	KataGo 選択 確率の幾何平均
0.6 (期待損失 ↓ ↓)	1.805	2.33%	0.048
0.8 (期待損失 ↓)	2.014	2.88%	0.046
1.0 (誘導なし)	2.310	3.83%	0.039
1.2 (期待損失 ↑)	2.907	6.05%	0.035
1.4 (期待損失 ↑ ↑)	3.160	7.00%	0.034

における誘導効果については 5.2.1 項で述べ、誘導側から見た結果については 5.2.2 項で述べる。

5.2.1 相手役着手における誘導効果の評価

表 2 に、誘導パラメータ c_g に対する、相手役 (Pachi_{kgs3k}) の着手統計を示す。相手着手について、複数の指標から誘導効果を評価した。

まず、相手着手の目数損失の平均に着目する。 c_g の増加に伴い、目数損失の平均は単調に増加している。これは、誘導パラメータによって、相手が良い手を打ちやすい展開から、誤りやすい展開へと段階的に誘導できていることを示している。次に、10 目以上のやや大きな損失を伴う着手の割合に着目する。この指標においても、 c_g の増加に伴って単調な増加が見られる。具体的には、 $c_g = 0.6$ では約 2.33% 程度であるのに対し、 $c_g = 1.4$ では約 7.00% に達しており、誘導パラメータの設定によって、明確な失着が生じる頻度を段階的に制御できていることが分かる。

目数損失以外の指標として、KataGo 自己対戦ネットワークによる選択確率にも着目する。この選択確率は、着手の有望性を反映する指標と解釈でき、値が高いほど、強い意味で良い着手であると考えられる。相手着手に対する選択確率の幾何平均は、 c_g の増加に伴って単調に減少している。この結果は、 c_g を大きく設定することで、相手が強いネットワークにおいて有望とされる着手を選びにくくなっている可能性を示唆している。以上の結果より、相手プレイヤーの強みを発揮させる場合、および課題を露呈させる場合のいずれにおいても、本稿で提案した誘導手法が有効に機能していることが確認できた。

5.2.2 誘導側における自然な手加減の評価

本項では、誘導側が自然な手加減を行っていたかを評価する。まず、誘導側から見た勝率を表 1 に示す。 c_g の設定によっては、勝率が 50% からやや離れる場合も見られるが、極端な一方的勝利にはなっておらず、一定程度の手加減が行われていると解釈できる。

次に、誘導側の目数損失平均に着目する。この値は c_g の増加に伴って単調に増加している。この傾向については、(1) 相手に誤りやすい展開を作るために、誘導側自身も形勢を損なう着手を選択している可能性、(2) 相手が誤った

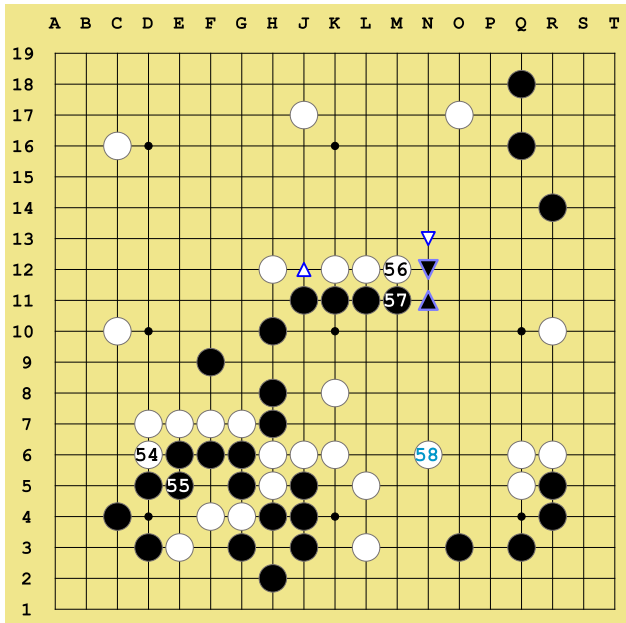


図 3: 誘導の局面例 1 (黒番). 黒三角は誘導手を表し、上向き (▲) は「相手に間違えさせない」手、下向き (▼) は「相手に間違えさせる」手、小さい白三角 (△, ▽) は各誘導手に対して相手が選択すると想定される着手を表す。

着手を選択することで形勢差が拡大し、その結果として誘導側が意図的に手加減している可能性、の 2 通りの解釈を考えた。この点については、さらなる分析が必要である。

最後に、誘導側着手の自然さについて、 $rank_1d$ に基づく選択確率の幾何平均を用いて評価する。 $c_g = 1.2$ および $c_g = 1.4$ では、この値がやや低下しており、相手を誤らせる方向の誘導において、誘導側がやや不自然な着手を選択している可能性を示唆している。以上の結果から、誘導側では、手加減の度合いと着手の自然さを一定程度保ちながら、誘導の強さや方向性に応じた着手選択が行われていることが確認できた。

6. 誘導の局面例

本章では、囲碁における誘導の典型例を 3 つ、想定と違う例を 1 つ紹介する。

局面例 1 図 3 (黒番) は、相手役が白 58 と弱い白石を連絡し、中央黒への攻撃態勢をとった局面である。最善は黒 J12 から白の包囲網の弱点を攻撃する手であるが、黒 N11 ノビ (▲) や黒 N12 ハネ (▼) と右上に連絡しようとするのも自然であり、約 7 目の損ではあるが、手加減する側の打ち方としてはありうる。

黒 N11 ノビ▲はどちらかというと局所的にはおとなしい手であり、白は局所的な戦いよりも広い視野を持って考えやすい。具体的には、弱点を残したまま局所的な形である N12 や O13 に打ってしまう割合はそれほど高くなく ($rank_3k = 0.20, 0.14$. 約 4-5 目損)、最善の白 J12 △ ($rank_3k = 0.21$) や白 G13 など守りに目が行く可能性が

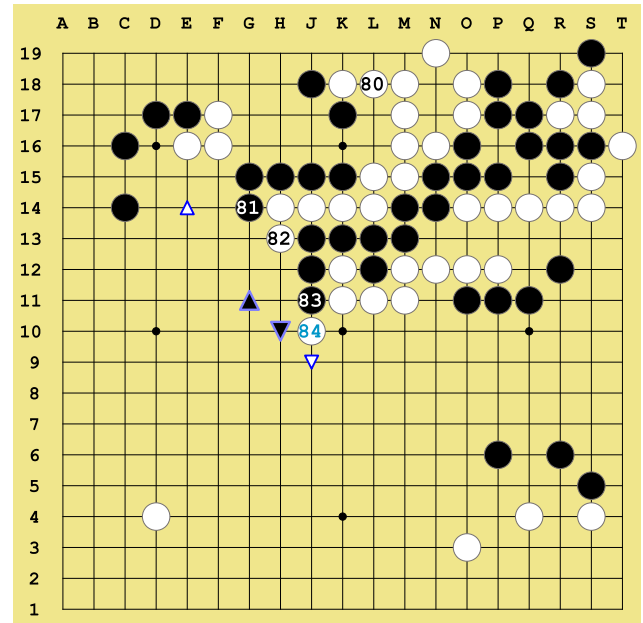


図 4: 誘導の局面例 2 (黒番). 黒三角は誘導手を表し、上向き (▲) は「相手に間違えさせない」手、下向き (▼) は「相手に間違えさせる」手、小さい白三角 (△, ▽) は各誘導手に対して相手が選択すると想定される着手を表す。

高い。一方、黒 N12 ハネ▼は、ノビに比べて白に局所的に働きかけている手であるため、中級者はその付近に目がいきがちである。具体的には、N13 ハネ▽ ($rank_3k = 0.25$, 約 6 目損) あるいは N11 キリ ($rank_3k = 0.32$, 約 8 目損) と打ってしまい、不利な戦いに誘導される可能性が高い。すなわち、相手に間違えさせたいなら N12 ▼、間違えさせたくないなら N11 ▲ が、有力な手になるのである。

局面例 2 図 4 (黒番) は、相手役が白 84 とハネ、まだ活きが確定していない右上の黒の大石を包囲しようとしている局面である。一方で、左上白 3 子にも十分活力が残っており、仮にこの白が逃げ出せば、上辺黒が攻められるような局面である。最善は黒 F13 にコスミ、上辺黒を安全にしながら白 3 子を確実に捕獲する手であるが、これはやや見えにくい。黒 H10 ハネ (▼) や黒 G11 トビ (▲) と頭を出す手も自然であり、約 7 目の損ではあるが、手加減する側の打ち方としてはありうる。

黒 H10 ▼に対しても G11 ▲に対しても、白は左上白 3 子を逃げ出す E14 トビ (△) や D15 ノゾキが最善である。黒 G11 ▲が白石に働きかけていない手であり、白 E14 △が比較的に見えやすい ($rank_3k = 0.23$)。一方で、黒 H10 ハネ▼に対しては、つい白 J9 ▽ ($rank_3k = 0.29$) などと相手をしたくなってしまうため、白 E14 に目が向かう可能性が減ってしまう ($rank_3k = 0.11$)。すなわち、相手に間違えさせたいなら H10 ▼、間違えさせたくないなら G11 ▲が、有力な手になるのである。

局面例 3 図 5 (白番) は、右上に黒がスベった局面である。白の最善手は左上 C17 と三々に入るか C14 とカカる

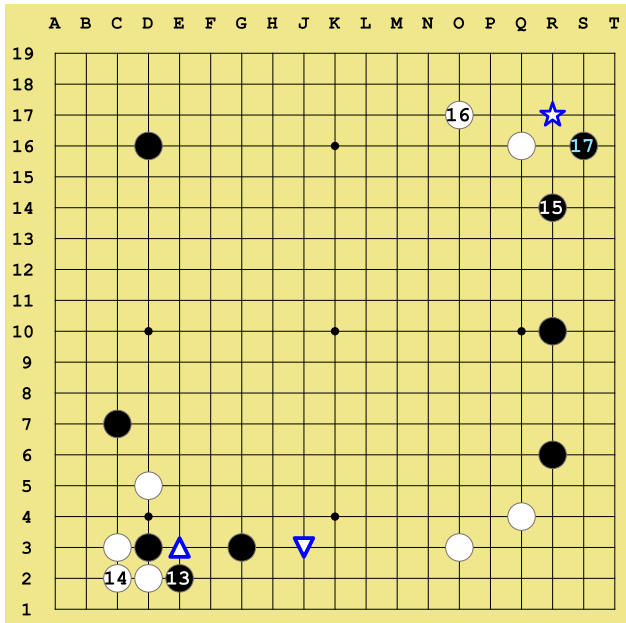


図 5: 誘導の局面例 3 (白番). 白三角は誘導手を表し, 上向き ($\triangle$) は「相手に間違えさせない」手, 下向き (∇) は「相手に間違えさせる」手, 白星 ($\star$) は誘導補正を行わない場合に選択される手加減手を表す.

か,あるいは H4 のカタも有力であるが,いずれも中級者には見えにくい.むしろ,白 R17 に受けたり ($\star$, $\text{rank_3k} = 0.27$),左下の黒弱石を攻める E3 ($\triangle$, $\text{rank_3k} = 0.23$),J3 (∇ , $\text{rank_3k} = 0.09$) などが自然に見えるであろう.これらの手はいずれも KataGo 上で 1 目ほど (勝率では 8% くらい) 損な手であるが,手加減としては自然に見えて ($\text{rank_1d} = 0.22, 0.17, 0.19$) 有力である.

これらの手はそれぞれ性格が異なる.白 R17 $\star$ は黒に手を渡しており,黒としては全体を見渡して左下の弱点を守る E3, D4, K3 などの準最善手に目が行きやすい (期待目数損失 2.26)^{*5}.いわば,「自由に打たせる手加減」である.白 E3 $\triangle$ は,一度白 14 とついだ後なので黒 D3 の石は軽い.黒としては F3, K3 などと軽く打つ手が見えやすい (期待目数損失 1.88)^{*6}.これは R17 $\star$ に比べても,「相手が対応を間違えにくい」タイプの手加減である.

一方で,白 J3 ∇ は「黒さん攻めますよ,どうですか」と,課題局面を与えている.中級者にとって自然なのは D4 ($\text{rank_3k} = 0.59$) や E3 ($\text{rank_3k} = 0.24$) であるが,これらは軽い石を重く動く手であり,3.8 目ほど損である.最善は黒 E6 カケであるが,これは中級者には見えにくい ($\text{rank_3k} = 0.0008$).まだしも準最善の黒 G5 (1 目損, $\text{rank_3k} = 0.02$, $\text{rank_1d} = 0.08$) は軽く動く分かりやすい手である.これらから J3 ∇ の期待目数損失は 4.40 と他の 2 手に比べて高く,「間違わせて指導する」ような,い

^{*5} 黒 C14 と K3 がほぼ最善, E3 や D4 は 2 目半程度損

^{*6} 黒 C14, N16, K3 がほぼ最善, F3 が 1 目損,級位者に見えやすい F2 が 5 目損

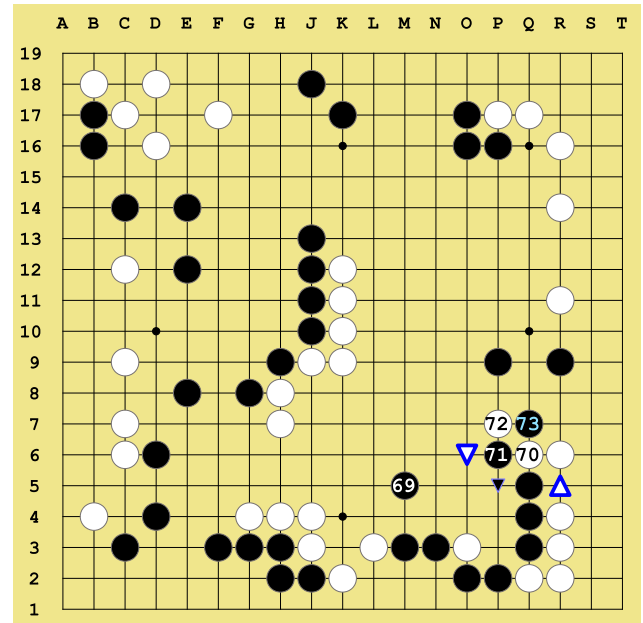


図 6: 誘導の局面例 4 (白番). 白三角は誘導手を表し, 上向き ($\triangle$) は「相手に間違えさせない」手, 下向き (∇) は「相手に間違えさせる」手, 小さい黒下向き三角 ($\blacktriangledown$) は白の ∇ に対して相手が選択すると想定される着手を表す.

わば教育用手加減になっている.

ここまでの 3 つの例は全て,補正なしの場合に比べ,補正を行うことで「間違えさせる」あるいは「間違えさせない」ような手が選べるようになった例である.

局面例 4 図 6 (白番) は,少し想定とは違う「間違えさせ方」をしようとした例である.相手役が黒 73 と切った局面であり,最善は白 R7 とアテることで間接的に R5 の傷を守る手である.補正なしの手加減手法では白 R5 $\triangle$ と単純に傷を守り,黒 P5 の傷を強調する (約 5 目損).この場合,黒には O6 ノビで傷を守る手と, O7 カカエで白一子を取る (右下 3 子は見捨てる) 手があり,どちらも有力でありかつ目につきやすい.計算された期待目数損失は 0.69 目であり,「黒は間違わないだろう」と判断している.

ここに「間違えさせる」補正を行うと,白 O6 アテ ∇ というあまり見ない手が選ばれてしまう.白 O6 アテ ∇ に対して一見自然な ($\text{rank_3k} = 0.91$) P5 ツギ $\blacktriangledown$ が,最善手である R5 デに比べて 7 目も損であり,期待目数損失が 6.57 目,つまり「 ∇ を選べば黒を間違えさせる」と判断するためである.このように,わざと隙のある手を打ったうえで「正しく咎めることができないことを期待する」のは,指導の一環としてありえないことはないが,多用すると「上手が嘘手ばかり打っている」ことになってしまうため,注意が必要かもしれない.

7. おわりに

本稿では,対戦ゲームにおいて,勝敗の最適化ではなく,局面や対戦相手の着手をある方向に誘導する枠組みを提案

した。これにより、将来的には例えば「楽しめた」「成長した」と感じられるようなゲーム体験を実現することを狙う。人間プレイヤーを対戦相手として想定し、その実力、弱点、プレイスタイルや好み、さらには対戦目的がある程度把握できている状況のもとで、対局中の相互作用を通じて、特定の行動や展開が生じやすくなるように誘導する点に特徴がある。この枠組みは、指導対局における上達支援や娯楽性向上にとどまらず、多様な対戦目的に応じた体験設計へと拡張可能である。

提案した誘導概念の具体例として、囲碁の指導対局を対象に、人間プレイヤーの着手傾向を評価するモデルを用いた誘導実装を示した。相手の期待目数損失を指標とし、誘導パラメータによって対局展開を制御する手法を実装した結果、囲碁プログラム Pachi を相手とした対戦実験において、相手が良い手を打ちやすい展開から、誤りやすい展開までを段階的に誘導できることを確認した。

今後の課題として、まず誘導補正における探索対象の選択方法が挙げられる。本稿では、各誘導側着手に対して相手着手の一部を探索する実装としたが、実用化を考えると、誘導側にとって有望な着手に限定して誘導評価を行うなど、探索対象を段階的に絞り込む工夫が必要である。また、本稿では相手の期待目数損失を誘導指標として用いたが、「片方が中央志向／辺隅志向になりやすくする」「戦いが起きやすく／起きにくくする」といった誘導や、対局目的に応じた多様な誘導関数の設計にも取り組みたい。さらに、人間プレイヤーとの対局実験を通じて、誘導が学習効果やモチベーションに与える影響を評価することも、今後の重要な課題である。

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] Hunnicke, R. and Chapman, V.: AI for dynamic difficulty adjustment in games, *AAAI-04 Workshop on Challenges in Game Artificial Intelligence* (2004).
- [2] Souza, C. H. R. et al.: Extending a MAPE-K loop-based framework for Dynamic Difficulty Adjustment in single-player games, *Entertain. Comput.*, Vol. 52 (2025).
- [3] Liu, A.-J. et al.: Strength adjustment and assessment for MCTS-based programs [research frontier], *IEEE Comput. Intell. Mag.*, Vol. 15, No. 3, pp. 60-73 (2020).
- [4] Fujita, K.: AlphaDDA: strategies for adjusting the playing strength of a fully trained AlphaZero system to a suitable human training partner, *PeerJ Comput. Sci.*, Vol. 8 (2022).
- [5] Hsueh, C.-H. and Ikeda, K.: Improvement of move naturalness for playing good-quality games with middle-level players, *Appl. Intell.*, Vol. 54, No. 2 (2024).
- [6] McIlroy-Young, R. et al.: Aligning superhuman AI with human behavior: Chess as a model system, *KDD '20* (2020).
- [7] Tang, Z. et al.: Maia-2: A unified model for human-AI

- alignment in chess, *NeurIPS 2024* (2024).
- [8] lightvector: New Human-like Play and Analysis, <https://github.com/lightvector/KataGo/releases/tag/v1.15.0>.
- [9] Sephton, N., Cowling, P. I. and Slaven, N. H.: An experimental study of action selection mechanisms to create an entertaining opponent, *CIG 2015* (2015).
- [10] Nakamichi, T. and Ito, T.: Adjusting the evaluation function for weakening the competency level of a computer shogi program, *ICGA J.*, Vol. 40, No. 1 (2018).
- [11] Fujihira, K., Hsueh, C.-H. and Ikeda, K.: Procedural maze generation considering difficulty from human players' perspectives, *ACG 2021* (2021).
- [12] Billings, D. et al.: Opponent modeling in poker, *AAAI '98* (1998).
- [13] Liu, M. et al.: Safe opponent-exploitation subgame refinement, *NeurIPS 2022* (2022).
- [14] 早下雅弘, 池田心, シュエジュウシュエン: 囲碁 AI が好ゲームを演出するための要因分析, 技術報告, 北陸先端科学技術大学院大学, GI-49 (2023).
- [15] Yen, S.-J., Chen, Y.-L. and Lin, H.-I.: Scaffolding learning for the novice players of Go, *ICITL 2019* (2019).

付 録

A.1 自然な手加減度の計算方法

本付録では、我々が提案した「自然な手加減度」の計算方法 [5] を示す。局面の形勢評価には KataGo が出力するリード目数 (scoreLead) を用いる。着手 i のリード目数を a_i , rank1d による選択確率を p_i^{1d} とする。

まず、目標とする手加減量を表す「目標損失」 l^{target} を定義する。 $\alpha > 0$ を形勢範囲を調整する定数、 $\beta > 0$ を損失量の基準値とし、十分に探索された着手の中で最も優勢なもののリード目数を $a_{\max} = \max_{i \in \{j | n_j > 10\}} a_i$ とすると、 l^{target} は次式で与えられ、優勢局面では比較的大きな手加減を行い、劣勢局面では過度な損失を避ける。

$$l^{\text{target}} = \begin{cases} 0, & \text{if } a_{\max} \leq -\alpha \\ (1 + \frac{a_{\max}}{\alpha}) \times \beta, & \text{if } -\alpha < a_{\max} \leq \alpha \\ 2 \times \beta, & \text{if } \alpha < a_{\max} \end{cases} \quad (\text{A.1})$$

次に、各着手 i の手加減度 q_i を定義する。着手 i の実際の損失量と目標損失との差を $e_i = |(a_{\max} - a_i) - l^{\text{target}}|$ とし、誤差の影響を制御する定数 $\tau > 0$ を用い、

$$q_i = \frac{1}{(\frac{(e_i)^2}{l^{\text{target}} + 1} + 1)^\tau} \quad (\text{A.2})$$

とする。 $l^{\text{target}} + 1$ により、目標損失の大きさによって誤差の意味合いが変わることを考慮する。

最後に、自然な手加減度 s_i を次式で定義する。

$$s_i = (p_i^{1d} + \epsilon) \times q_i \quad (\text{A.3})$$

ϵ は自然さ評価が 0 となることを防ぐための微小定数である。ただし、着手の筋を一定程度保つため、損失が 20 目を超える着手については $s_i = -\infty$ とする。

正誤表

表 1（誘導側の統計結果）の誤り

誘導手頻度の算出において、分母に相手手番の着手も含めて計算していた。本来は誘導側手番のみを分母とすべきであり、真の誘導手頻度は表 1 に記載した値のおよそ 2 倍となる。なお、他の統計量および結論には影響しない。誘導手頻度の修正値は以下の通りである：

- 0.6：9.18% → 18.32%
- 0.8：4.62% → 9.24%
- 1.0：0%
- 1.2：4.85% → 9.68%
- 1.4：8.47% → 16.94%