

対局状況をより正確に表現するための盤面評価値

小川 竜欣^{1,a)} 池田 心^{1,b)}

概要: ゲーム AI は将棋, 囲碁, チェスなどのボードゲームで人間よりも強くなったが, 人間を楽しませるという点から見ると, 残された課題は多い. トッププレイヤーに勝てるようなゲーム AI をそのまま対局に使用すると人間には理解できない高度な手を示され, 探索に手を加えると人間では考えられないミスが発生してしまう. また, 最近ではプロ棋士の対局で局面の評価値が示されることも多いが, 人間プレイヤーの実感または実際と乖離した評価値が示されることもある. 本稿では, 盤面から勝率を予測する教師あり学習を行う際に, 棋力も入力に含めることで, より人間らしい局面の評価を目指す. また, 推定した勝率の性能を確かめるため, 指し継ぎによる評価を行った. 指し継ぎには, 人間的な評価項目を採用している技巧 2 を用いた. 探索の深さを制限した弱い AI による指し継ぎの勝敗は, 我々のモデルの予測勝率のほうが, 強い AI の予測勝率よりも近かった. また, 同一局面で入力レートが変わる場合に予測勝率が大きく異なる局面やほとんど等しい局面をサンプリングして, 局面の解釈を行った. その結果, このサンプリング方法で逆転が起りやすい局面や起りにくい局面をある程度抽出することができた.

キーワード: コンピュータ将棋

Value Function for a more accurate representation of the game situation

OGAWA TATSUYOSHI^{1,a)} IKEDA KOKOLO^{1,b)}

Abstract: Game AI has become stronger than humans in board games such as Shogi, Go, and chess. However, in terms of entertaining humans, there are still many issues to be solved. If a game AI that can beat top players is used as is, it will show advanced moves that humans cannot understand. If we weaken the ability to search for the game AI, it will make mistakes that humans cannot imagine. Recently, evaluation values of professional players' games are shown. However, the evaluation values sometimes deviate from the human player's feelings or reality. In this paper, we aim for a more human-like evaluation of the game by including the player's ability as an input in the supervised learning process to predict the winning percentage from the position. To verify the performance of the estimated probability of winning, we evaluate the probability by weak Shogi AI simulating. For the evaluation, we used Gikou 2, which uses human-like evaluation items. The predicted win rate of our model was closer to the predicted win rate of the strong AI than that of the weak AI with limited search depth. We also sampled positions in which the predicted winning percentage differed greatly or were almost equal when the input rate changed in the same positions and then interpreted the positions. As a result, we were able to extract the positions where comebacks are likely to occur and those where they are unlikely to occur to some extent using this sampling way.

Keywords: Computer Shogi

1. はじめに

近年, ゲーム AI の発展は目覚ましく, 将棋や囲碁, チェスなどの二人完全情報ゲームにおいては, トッププレイヤー

を打ち負かすほどの力を手に入れている. 最近では強いゲーム AI だけではなく, 人間を楽しませるゲーム AI についても研究が行われている. [1] では, 楽しませる AI の要素技術として, プレイヤを理解したり, 不自然な着手を抑制したりすることが重要だと述べられている. また, [2] では, 不自然な着手と棋力には大きな関わりがあるものとされている. [2] ではさらに, ゲーム AI との対局は人間との対局に比べてつまらないという問題点が指摘されてい

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) ogawa.tatsuyoshi@jaist.ac.jp

b) kokolo@jaist.ac.jp

る。例えば将棋であれば、将棋 AI と調整を行わずに戦うと、棋力が高すぎて対局相手として適しておらず、探索を調整すると人間とは異なった弱さが現れ不自然に感じられる。このように、対局相手として将棋 AI を評価すると、残された課題は多い。

対局以外については、対局の解説や棋譜の解析において将棋 AI に人間らしさが求められる場合がある。対局を観戦しているときは、状況を正確に判断できたほうが対局の臨場感が味わえ、より楽しむことができる。しかし、プロ同士の対局だとアマチュアには理解が難しいため何かしらの手助けが必要になる。ここで活躍するのが人間の「解説」や「聞き手」であるが、解説に将棋 AI が用いられたり、人間がおらず将棋 AI の評価値のみが参考にされたりする場合がある。しかし、将棋 AI を用いた解説では、「その手は人間には指せない、人間ならばこう指すところ」といった指摘がよく見受けられる。このような指摘は、既存の将棋 AI が最善のやりとりを探索した結果の評価値と候補手を表示しており、人間にとって高度すぎる手順を示しているときに見られる。この場合、示されている評価値は探索結果ありきのものなので、実際の対局の状況を正確に、つまり人間プレイヤーの実感や実際の対局結果を予測できるように表わしていない可能性がある。

この課題はプロ棋士の対局に限ったことではなく、アマチュアが将棋を勉強する際にも問題になる。アマチュアが将棋を指す場合、プロ棋士よりも時間が短く、切れ負けや最初から秒読みの将棋が多い。短時間で納得のいく局面まで読み切ることは難しいので、制限時間に従って読みを入れることになる。その場合だと、最善手を指し続ければ必ず勝つが一手でも間違えれば即負けにつながる、という評価値が高い局面よりも、必ず勝てるわけではないが何度かミスをして勝てるような、少し評価値が低い局面を選ぶほうが、結果的に勝率が高くなることが多い。そのため、対局者の強さに応じた「実感・実際の結果に近い」評価値を求めることの必要性は高い。

本研究の目標は、対局者同士の平均棋力を考慮することで、状況をより正確に表現できる評価値を設計することである。

2. 関連研究

どうぶつしょうぎは既に完全解析されていて後手必勝であることが判明している [3] が、逆転の余地を取り入れることで、完全解析の結果以上に人間の考え方に有用な評価関数を設計した研究がある [4]。この研究では、残り 1 手で勝てる局面と残り 10 手で勝てる局面を比較すると、残り 1 手で勝てる局面のほうが指し手を間違える可能性が少なく勝ちやすいとしており、劣勢である先手番での逆転勝ちや、優勢である後手番での安全勝ちを狙うことに優れた評価関数を得ている。

[5] では、チェスを題材にして、人間の次の一手予測や、人間が次に大きなミスをするかという予測について高い精度を得ている。この研究では、棋譜を対局者のレートによって分割しており、レート 1100 からレート 1900 までの棋譜を 9 分割して、それぞれ別のモデルで学習が行われている。本研究ではレートによってモデルを分けず、盤面とともにレートを数値として入力に使用する。このモデルの利点は、学習する際にレートを連続値として扱えるため、比較的必要なデータ数が少量で良いこと、また、実際に使用する際に使用者のレートにより近づけられることが挙げられる。

3. 提案手法と考えられる用途

既存の評価関数は、互いに最善手を指し続けた場合、どのような結果になるか、という探索に基づいて局面の評価を求めている。これは、すでに述べたように、一般的な人間プレイヤーの実感や、人間プレイヤーが実際に指した場合の結果と異なる場合がある。

我々は、盤面と棋力から、そのプレイヤーたちが指し続けた場合の勝率を求めたい。しかし、ある盤面から、複数の棋力のプレイヤーが複数回指し続けた結果を得ることは難しいので、盤面と棋力を入力としてその結果、つまり勝敗を出力とするような教師あり学習で近似できると期待している。例えば、「相手玉に難解な詰みがあるが、白玉は明らかな必至である」という局面においては、棋力が低ければ勝率が低く、棋力が高ければ勝率が高いと判定されたい。もしこれが達成された場合は、次のような用途が考えられる。

- 一度ミスすれば形勢が極端に悪化するのか、それとも何度かミスが許されるのかという、局面の安定度ののようなものが定義・利用できる。
- プロ棋士の対局において、予想される結果に近い、実際の形勢や安定度を示すことができる。
- 初中級者の対局相手 AI として、すぐには気づかない好手がある局面を渡すなどの自然な手加減や、調整なしでは生じにくい攻め合いへの誘導ができる。
- 初中級者の指導 AI として、無理のないアドバイス、実質的に勝率を上げるためのアドバイスが可能になる。

本研究では、このレートあり予測勝率によって、人間同士の対局状況をより正確に表わすことを目標とする。

4. 実験

4.1 実験内容

実験では、将棋倶楽部 24 万局集 [6] を使用した。9 割の棋譜を訓練データとし、1 割の棋譜をテストデータとして用いた。その際、不利な局面でも最終的に高レート側が勝っている、といったレート差による結果の偏りを防ぐため、対局者のレートの差が 50 以内の棋譜を抽出した。

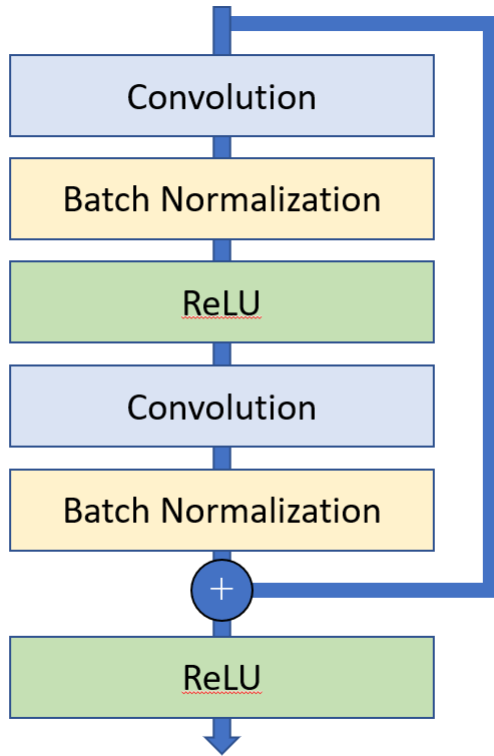


図 1 ResNet の構成

また、最序盤と入玉含みの局面を除外するため、手数が 50 手目から 199 手目までの局面を使用した。最序盤を除外した理由は、似た局面が多く学習に悪影響を及ぼす可能性があることや、勝率に大きく影響を与えない局面が多いと思われるためである。入玉含みの局面を除外した理由は、将棋倶楽部 24 万局集の時点では入玉宣言法が実装されておらず、適切な評価を行えないためである。その結果、訓練データは約 686 万局面、テストデータは約 76 万局面が得られた。

レートなし条件では、局面のみを入力として、レートあり条件では、局面と対局者二人のレートを平均したものを入力とした。入力に使用したレートは平均が 0、分散が 1 になるように標準化を行っている。レートなし条件、レートあり条件ともに局面の勝率と着手確率分布の予測するマルチタスク学習を行っている。

ネットワークの構造や学習オプションについては、山岡による書籍 [7] を参考にした。入力の形式は駒ごとの配置の配列、手番ごとの占有している座標の配列、手番ごとの持ち駒の配列によって構成されているビットボードから入力特徴を作成した。入力特徴は駒の種類ごとにチャンネルを分け、各チャンネルは駒の座標を表す二値画像とした。持ち駒は種類ごとに最大枚数分のチャンネルを割り当てた。中間層のフィルター枚数は 256 として、1 層目は畳み込み層、2 層目からは図 1 のような Residual Network (以下 ResNet) を 5 ブロック重ねる形を採用した。出力の形式に

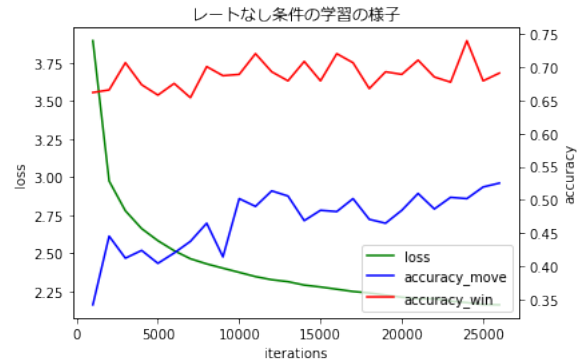


図 2 レートなし条件での学習の様子

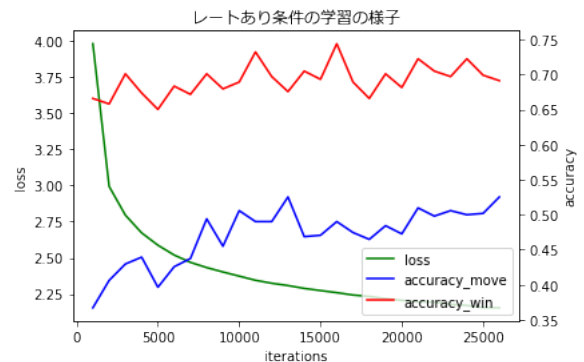


図 3 レートあり条件での学習の様子

については、勝率予測では、勝敗を二値分類の問題として捉え、着手確率分布予測では、移動方向と移動先の座標の組み合わせにより、2187 のラベルを分類する多クラス分類の問題として捉えた。オプティマイザは Adam を採用した。

4.2 実験結果

学習にかかった時間は、GTX-1070 の一般的な PC でおよそ 4 時間だった。このため、学習する棋譜がもう少し増えても問題ないと思われる。図 2 は、レートなし条件での学習の様子、図 3 は、レートあり条件での学習の様子である。実験の結果、勝率予測の正解率は、レートなし条件では 0.7000、レートあり条件で 0.7005、着手予測の正解率は、レートなし条件では 0.4977、レートあり条件では 0.4983 という結果になった。着手予測、勝率予測のどちらもレートあり条件のほうが正解率が高いが、有意な差ではなかった。これらの条件に大きな差が見られないことは意外な結果であった。我々は、棋力によって結果が異なる局面が多くあると考えていたため、レートを入力に入れることで、予測がより正確になると考えていたためである。結果から、この前提が間違っている可能性も示唆されるが、別の可能性もある。局面は対局者が選んだ手の積み重ねであるため、「初心者らしい局面」や「プロらしい局面」といったものが存在する。そのため、レートを明示的に入力に含めずとも局面から棋力の大まかな推定が出来てしまい、推定精度の差がほとんど出なかったという可能性がある。

表 1 手数ごとの予測勝率の正解率の比較

手数	50-79	80-109	110-139	140-169	170-199
局面数	363112	265826	98299	26388	6183
レートなし条件	0.6459	0.7393	0.7658	0.7831	0.7847
レートあり条件	0.6457	0.7395	0.7676	0.7889	0.7962

表 2 平均レートごとの予測勝率の正解率の比較

平均レート	0-499	-999	-1499	-1999	-2499	2500-
局面数	81885	208233	184283	200948	83516	862
レートなし条件	0.6942	0.7135	0.7038	0.6907	0.6857	0.7135
レートあり条件	0.6947	0.7132	0.7060	0.6906	0.6857	0.7135

ここからは勝率予測を分析していく。表 1 は、レートあり条件での予測勝率の正解率と、レートなし条件での予測勝率の正解率を、手数ごとにまとめたものである。結果から、どちらの条件でも手数が進んで終局に近づくほど勝率の予測が正確になることが分かった。また、手数が大きくなるにつれて、レートあり条件のほうが正解率が上がっている。このことから、終盤に近いほど、レートを入力に含めて学習することでより状況を正確に判断できることが分かった。これは、先ほど述べたような、難解かつ勝敗に直結するような手が終盤だと現れやすいからだと考えている。表 2 は、レートあり条件での予測勝率の正解率と、レートなし条件での予測勝率の正解率を、対局者二人の平均レートごとにまとめたものである。平均レートの高低によってレートあり条件とレートなし条件での正解率に差は出なかった。また、どちらの条件でも、レートが高いほど予測がしやすいというわけでもないことが分かった。さらに、レートあり条件では、このレート帯だと結果がおかしくなる、といった学習が苦手なレート帯がないことも判明した。

4.3 評価

4.3.1 指し継ぎによる評価

評価のため、人間的な評価項目を使用している技巧 2 によって指し継ぎを行う。人間によって指し継ぎを行うことが理想的だが、現実的には困難なため、ある程度弱くした将棋 AI によって指し継がせる。Bonanza は、[2] によってレートが算出されており、探索の深さが 4, 5 の Bonanza の強さはそれぞれ、将棋倶楽部 24 のレートで 1740, 1984 に相当する。探索の深さが 1 の技巧 2 は、探索の深さが 4 の Bonanza に 61 勝 38 敗 1 分、探索の深さが 5 の Bonanza に 42 勝 58 敗 0 分であったので、将棋倶楽部 24 のレートで 1850 程度として使用した。

図 4 は指し継ぎを行った 1 つ目の局面で、先手番である。先手は二枚馬が手厚いが玉が薄く、後手は玉が固いが駒の効率が悪い局面で、先手がうまい攻めを見つければ押し切れそうだが、あまりにゆっくりした攻めだと後手に 6 四香車と打たれたり飛車を下ろされたりしてしまう。この局面

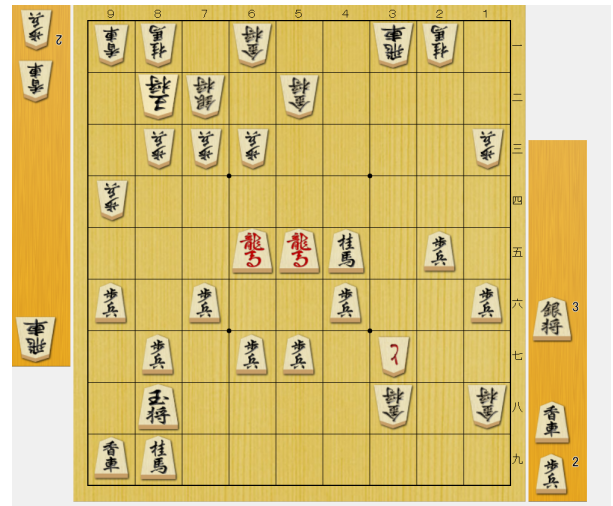


図 4 指し継ぎを行った局面 1

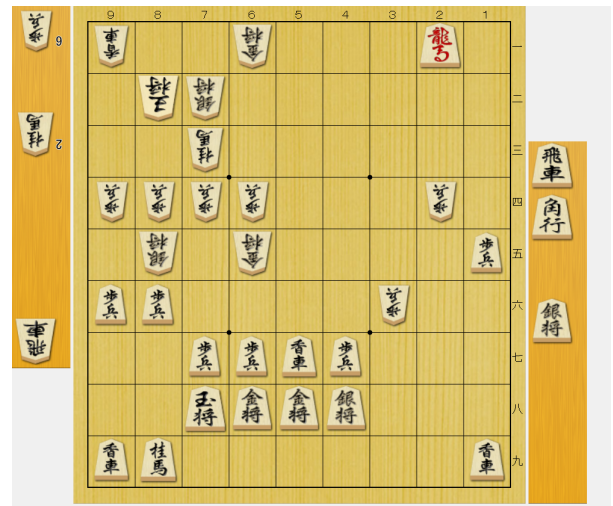


図 5 指し継ぎを行った局面 2

を水匠 4 で 10 億ノード読ませると、先手優勢 (1260) と評価する。評価値を勝率に変換する式は $1 / (1 + \exp(-\text{評価値}/600))$ とされている。この式に 1260 という値を当てはめて勝率に変換すると、先手勝率 0.8909 となる。しかし本研究のモデルに評価させると、対局者のレートが 1850 のときは、先手の勝率が 0.2283 と評価する。ここで、探索の深さが 1 の技巧 2 同士で指し継がせると、先手 35 勝後手 65 勝引き分け 0 と、本研究のモデルの評価に比較的近い結果となった。

図 5 は指し継ぎを行った 2 つ目の局面で、後手番である。先手は大きく駒得していて、玉もそれなりに固いが 8 筋が弱点で、逆に後手は駒損を補うため 8 筋を狙っていく将棋になりそうである。この局面を水匠 4 で 10 億ノード読ませると先手有利 (760) と評価する。これを勝率に変換すると先手勝率 0.7802 となるが、本稿のモデルがこの局面を評価すると、対局者がレート 1850 のときは、先手勝率 0.4314 と出力する。同様に、探索の深さが 1 の技巧 2 同士で指し継がせると、先手 50 勝後手 50 勝引き分け 0 と、本

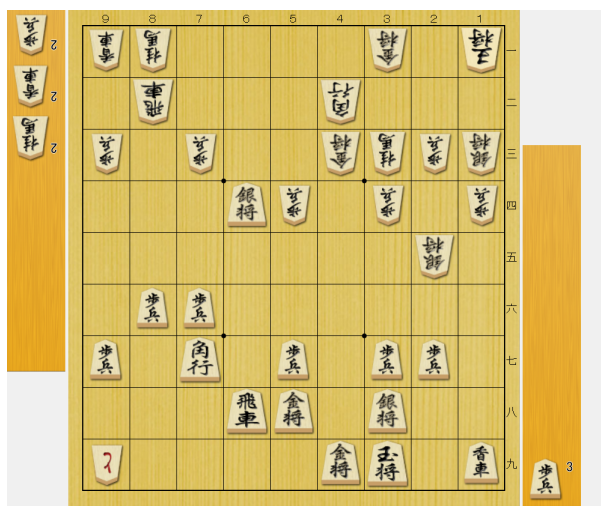


図 6 予測勝率差が最も大きい局面 (60~79 手目)

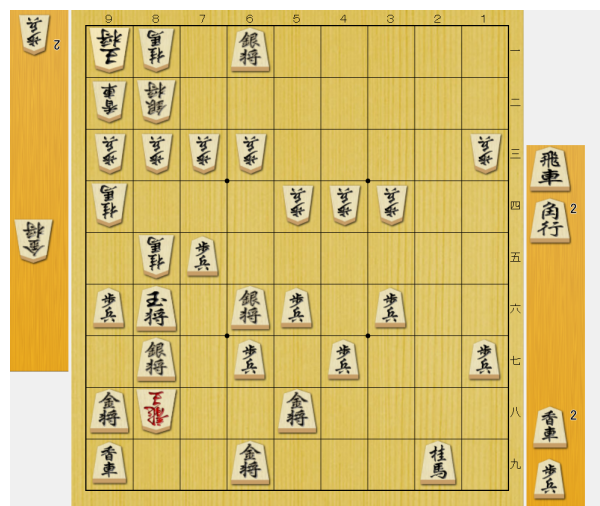


図 8 予測勝率差が最も大きい局面 (80~99 手目)

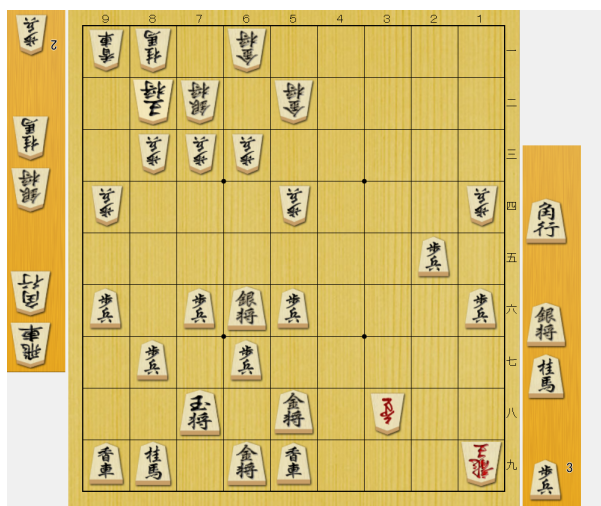


図 7 予測勝率差が最も小さい局面 (60~79 手目)

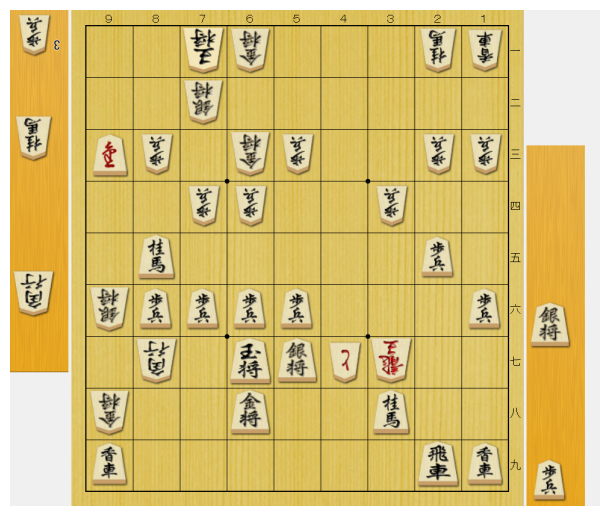


図 9 予測勝率差が最も小さい局面 (80~99 手目)

研究のモデルの評価に比較的近い結果となった。これらの指し継ぎによって、少なくともある局面においては、本研究のモデルが既存のモデルよりも対局の状況をより正確に表現できたといえる。

4.3.2 同一局面での予測勝率の差による評価

我々は、同一局面でも入力レートが変わることによって予測勝率が大きく異なる局面、ほとんど等しい局面について解釈を与えることで本稿のモデルについて一定の評価を下せると考えた。60 手目から 79 手目、80 手目から 99 手目、100 手目から 119 手目より、棋譜が被らないようにそれぞれ 512 局面ランダムサンプリングした。それぞれの手数サンプルについてレートが 500 の場合とレートが 1850 の場合で勝率予測を行い、予測勝率の差が最も大きい局面と最も小さい局面を抽出した。図 6 から図 11 が実際の局面図である。なお、先後にかかわらず、手番が全て手前側に来るように調整した。後に述べる勝率も手前側にとっての勝率である。以下で局面の解説と解釈を行う。

図 6 ではレート 500 が入力された場合は勝率 0.5996、

レート 1850 が入力された場合は勝率 0.3065 と出力される。後手は桂馬と香車を得しており、先手からの攻めをいなせば有利になりそうな局面である。先手からは 4 四歩という攻めが目が付く。5 三金とあたりを避けると同銀成同角 4 三步成がある。後手が困っているようだが、実は 4 四歩には同金と取る手が成立し、同角 4 三香と打って反撃すれば後手が指せる局面である。レート 500 同士の対局であれば先述の攻めが成功するが、レート 1850 同士の対局になると比較的多くの人間が反撃の筋に気づくために予測勝率の差が大きくなるのだと解釈できる。

図 7 ではレート 500 が入力された場合は勝率 0.2242、レート 1850 が入力された場合は勝率 0.2244 と出力される。後手は大体飛車一枚分得しており、玉も固いままである。先手の逆転筋は 7 四桂馬や 4 六角による王手飛車の筋だが、前者は同歩、後者は 4 九竜と自然に対応すれば問題ない。

図 8 ではレート 500 が入力された場合は勝率 0.3496、レート 1850 が入力された場合は勝率 0.6313 と出力される。先手が王手をかけられている局面で、先手が正しく応対す

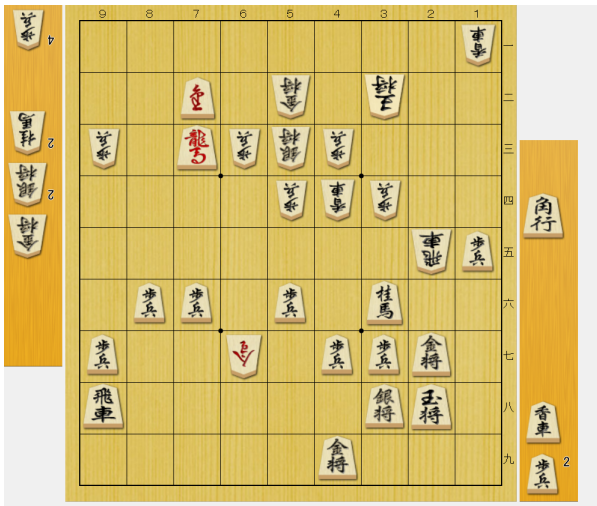


図 10 予測勝率差が最も大きい局面 (100~119 手目)

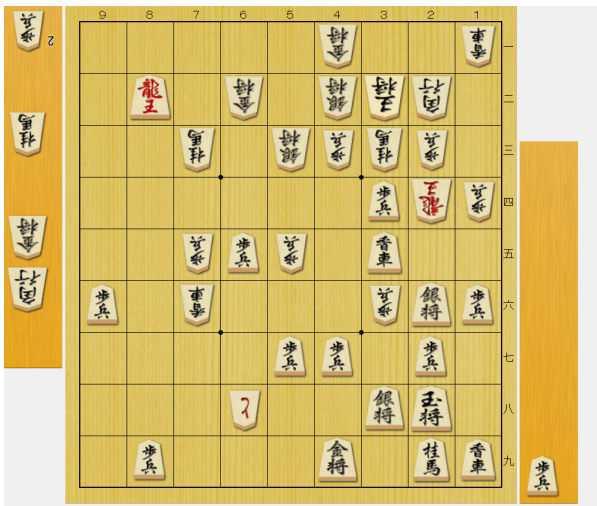


図 11 予測勝率差が最も小さい局面 (100~119 手目)

れば勝てそうだが、頓死筋があったり、後手玉が絶対に王手がかからない局面であったりするため、逆転が起こりやすい。具体的には、8五玉で詰まないと考えて指すと8六金で頓死したり、7六玉と指した後に9八竜と切って同香に6四金と抑える筋があり、実戦的には少し大変である。ただ、ある程度の棋力があれば受けきれられるために、レート1850が入力された場合は先手有利になるのだと解釈できる。

図9ではレート500が入力された場合は勝率0.0059、レート1850が入力された場合は勝率0.0060と出力される。先手玉には詰めろがかかっているが、後手玉に詰みはない。先手玉は受けても一手一手なうえに、後手が間違えやすくなる局面に誘導することはできないため、入力レートによって予測勝率に差がない局面である。

図10ではレート500が入力された場合は勝率0.7839、レート1850が入力された場合は勝率0.5097と出力される。先手の銀と桂馬と後手の角が交換され、後手のほうが駒得だが玉が薄い。ここでは2六香が最善手で、4四桂は手順

前後で互角になる。具体的には、4四桂同步2六香同飛車と切って2一香と打つ手がある。4四桂を指していなければこのとき2四歩が打て、特に問題ない。このように後手の反撃筋は飛車を切って小駒で攻めをつなげるような難易度の高いものになり、初級者同士であれば先手がそのまま押し切れると考えたため、低レートを入力した場合は先手勝率が高くなり、高レートを入力した場合は予測勝率が互角に近くなるのだと考えた。

図11ではレート500が入力された場合は勝率0.0738、レート1850が入力された場合は勝率0.0738と出力される。後手が大きく駒得しており、先手の攻め駒も足りない。先手の指し手は3三歩成くらいだが、同角同香成同銀4五桂4四角で先手の攻めは切れており、なおかつ先手側はこの一本道を辿るしかないため、レートにかかわらず後手が勝ちやすい局面である。

5. おわりに

本稿では、人間の対局状況をより正確に表わすことを目標に、レートを入力に含めて勝率を推定した。また、指し継ぎや、レート差によって勝率が変化する局面を実際に示して解釈を行い、モデルの評価を行った。

将来的には、うっかりや勘違いなどのミスを再現した探索モデルとともに使用することで、ミスが起こりやすい局面の人的な評価も目指したい。また、現時点では指し継ぎにおいて詰みがある局面で将棋AIは必ず詰ませてしまう。アマチュアや、持ち時間が短いときのプロが詰みを見逃すことはよくあるので、人間らしい終盤モデルの作成を行うことで、本稿のモデルの評価を終盤の局面でも行いたい。

参考文献

- [1] 池田心, 楽しませる囲碁・将棋プログラミング, オペレーションズ・リサーチ: 経営の科学, 58(3), 2013, 167-173
- [2] 仲道隆史, 熟達度に着目した将棋プログラムに対する不自然さに関する研究, 電気通信大学大学院情報理工学研究所博士論文, 2020
- [3] 田中哲朗, 「どうぶつしょうぎ」の完全解析, 研究報告, ゲーム情報学 (GI), 2009(3), 1-8.
- [4] 中屋敷太一, 金子知適, 逆転の余地を考慮した評価関数の設計とどうぶつしょうぎによる評価, The 25th Game Programming Workshop 2020
- [5] Reid, M-Y., Siddhartha, S., Jon K, Ashton A Aligning Superhuman AI with Human Behavior: Chess as a Model System, KDD, 2020
- [6] 久米宏, 将棋倶楽部 24 万局集, 2002, ナイタイ出版
- [7] 山岡忠夫, 将棋 AI で学ぶディープラーニング, 2017, マイナビ出版