

Title	着手確率を用いた人間の勝率予測の改善
Author(s)	小川, 竜欣; 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: 166-172
Issue Date	2024-11-15
Type	Conference Paper
Text version	publisher
URL	http://hdl.handle.net/10119/20017
Rights	社団法人 情報処理学会, 小川 竜欣, 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, 166-172. 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	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

着手確率を用いた人間の勝率予測の改善

小川 竜欣^{1,a)} シュエ ジュウ シュエン^{1,b)} 池田 心^{1,c)}

概要：近年、強いゲーム AI を指導や観戦支援に役立てる研究や、指導・観戦用のゲーム AI を作る研究が進められている。指導・観戦の観点から考えると、強いゲーム AI にとっての最善手だけではなく、人間プレイヤーがどのような手を指しそうなのか、その場合の人間プレイヤーにとっての勝率がいくらなのかを知ることができれば有益である。強いゲーム AI は主に AI 自身の勝利を目標に学習しているため、人間の指し手・予測勝率を正確に予測できるとは限らない。人間の指し手を予測する研究は盛んに行われている一方で、人間の勝率を予測する研究には残された課題が多い。本研究では将棋を対象に、既存の AlphaZero のような強化学習手法や教師あり学習手法の人間の勝率予測性能を調べる。そのうえで、着手確率分布を組み合わせることで、人間にとって良さそうに見える手の勝率を高め、思いつきにくい手の勝率を低めに重みづけて勝率を予測する手法を提案する。実験の結果、提案モデルは既存モデルに比べて、より高精度で勝率を予測できることが分かった。

キーワード：勝率予測, ヒューマン AI インタラクション, 将棋

Improving Accuracy of Win Rate Prediction for Human Players Using Selection Probability of Moves

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

Abstract: In recent years, research has been progressing in utilizing strong game AIs for teaching human players or supporting spectators, as well as in developing game AIs specifically for teaching or spectating purposes. From the perspective of teaching or spectating, it would be beneficial not only to know the optimal moves of strong game AIs but also to understand what moves human players are likely to make and what their win rates would be in such cases. Since strong game AIs are mainly trained with the objective of maximizing their own win rates, they may not necessarily be able to accurately predict human moves or the win rates for human players. While much research has been conducted on predicting human moves, there remain many challenges in predicting human win rates. In this study, we target Shogi and investigate the performance of existing reinforcement learning methods, like AlphaZero, as well as supervised learning methods in predicting human win rates. Furthermore, we propose a method that combines move probability distributions to weight the win rates, assigning higher weights to moves that seem good from human perspectives and lower weights to moves that are harder to come up with. Experimental results show that the proposed model can predict win rates with higher accuracy compared to existing models.

Keywords: Win Rate Prediction, Human-AI Interaction, Shogi

1. はじめに

将棋や囲碁、チェスといった二人完全情報ゲームにおいて、人工知能 (AI) は人間のトッププレイヤーを超える棋力

を有している。そこで、強いゲーム AI を指導や観戦支援に役立てる研究や、指導・観戦用のゲーム AI を作る研究が進められている。指導・観戦の観点から考えると、強いゲーム AI にとっての最善手だけではなく、人間プレイヤーがどのような手を指しそうなのか、その場合の人間プレイヤーにとっての勝率がいくらなのかを知ることができれば有益である。強いゲーム AI は AI 自身の勝利を目標に学習しているため、人間の指し手・予測勝率を正確に予測できる

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) hsuehch@jaist.ac.jp
c) kokolo@jaist.ac.jp

とは限らない。例えば、人間プレイヤーにとっては思いつくことが難しい着手の後、勝率が 100% に近くなるような局面があるような状況を考える。この場合強いゲーム AI にとっての勝率は 100% 近いが、人間プレイヤーにとってはその局面にたどり着くことが難しいため、勝率は 100% に近くないかもしれない。

近年、人間プレイヤーの指し手を予測する研究は盛んに行われている一方で [1][2][3][4]、人間プレイヤーの勝率を予測する研究には残された課題が多い。本研究では将棋を対象に、既存の AlphaZero のような強化学習手法 [5] や、教師あり学習手法 [1] はどの程度の精度で人間の勝率予測を行えるかについて調べる。そのうえで、勝率予測に着手確率分布を組み合わせて、着手ごとの人間の指しやすさを考慮する勝率予測手法を提案する。

2. 関連研究

AlphaZero[5] は人間の棋譜を使用せず、自己対局の結果から強いゲーム AI を作る強化学習手法である。AlphaZero は棋譜の勝敗、局面ごとの探索結果から、勝率、着手確率分布を学習する。

McIlroy-Young らは、AlphaZero のようなゲーム AI や、弱いものを得るために学習途中で取り出したゲーム AI の指し手は人間の感覚と異なると指摘した [1]。McIlroy-Young らはさらに、AlphaZero と同様のネットワーク構造を持ちながら、人間の棋譜から局面と着手・勝敗結果を教師あり学習する Maia を提案した。Maia は学習データの棋力帯ごとにモデルを分け、各モデルは 1200 万枚ずつの棋譜から棋力帯の特徴を学習した。

小川らは、Maia と比べて 1% 程度の棋譜を使用する場合、棋力帯ごとに学習データを分けない、Maia-S-All モデルのほうが人間の指し手を正確に予測できることを指摘した [4]。また、教師あり学習している Maia-S-All モデルでは人間でもあまり指さない悪手を指してしまうことを指摘した。そこで、Maia-S-All モデルと AlphaZero のように強化学習を行っている DLShogi の着手確率分布を適切に混ぜ合わせることで、模倣したい人間の強さに合わせて指し手をより正確に予測する手法を提案した。小川らの実験設定のもとでは、Blend モデルは Maia-S-All に対して棋力帯ごとに 0.003 から 0.020 ほど一致率（人間の指し手をどの程度当てられるかの指標）が高い。同様に、Blend モデルは DLShogi に対して棋力帯ごとに 0.050 から 0.100 ほど一致率が高い。

AlphaZero の出力する予測勝率から人間同士の対局に適した予測勝率に変換する研究として、平岡の研究が挙げられる [6]。平岡は 0 や 1 に近くなることが多い AlphaZero の予測勝率に対して、温度スケールリングを用いて予測勝率が 0.5 に近くなるようキャリブレーションを行い、Ownership のエントロピーを用いることでキャリブレーションの度合

いを決定する手法を提案した。

3. 提案手法

筆者らは、観戦・検討における予測勝率が人間の感覚と異なっているとき、多くの状況で着手確率分布も人間の感覚と異なっていることに着目した。そこで、人間の勝率を予測する際に、人間の着手をある程度予測できている着手確率分布も利用することで、より正確な勝率予測を行えるのではないかと予想した。本研究では、着手確率分布 π と勝率予測 V を組み合わせる最もシンプルな方法として、内積モデルを提案する。内積を取る理由は、人間が着手確率分布通りに指すと仮定した場合、遷移後の局面の予測勝率を適切に反映できるからである。

内積モデル $V_{ip}(\pi, V)$ では、式 (1) のように、root 局面の合法手 $move$ ごとに、着手確率 π_{move} と局面遷移後の予測勝率 V_{move} を推論し、内積を取る。

$$V_{ip}(\pi, V) = \sum_{move} (\pi_{move} \times V_{move}) \quad (1)$$

提案手法の利点の一つとして、計算の速さが挙げられる。この手法は、root 局面を入力にした着手確率分布と、root 局面から一手遷移した後の局面を入力にした予測勝率があれば計算できる。これらの入力をまとめたミニバッチは、個人用の一般的な GPU メモリに載るサイズであり、一度の出力によって得ることが可能である。

4. 実験

4.1 共通の実験設定

着手確率分布と予測勝率を出力するモデル、データについては、小川らの研究 [4] で使用されたものを参考にした。主な相違点として、小川らの研究で使用された教師あり学習モデルは 50 手目以降の局面のみを学習しているのに対し、本研究で使用する教師あり学習モデルは初手から学習を行っている。モデルは、(1) 教師あり学習を行った Maia-S-All、(2) AlphaZero のような強化学習を行った DLShogi、(3) Maia-S-All と DLShogi の着手確率分布を混ぜ合わせた Blend モデルを使用した。ただし、Maia-S-All は着手確率分布と予測勝率両方を使用し、DLShogi は予測勝率のみ、Blend モデルは着手確率分布のみを使用した。

Maia-S-All の学習データは将棋クエストの棋譜を使った。約 72 万棋譜を棋力帯ごとに 6 等分し、プレイヤーの棋力が最も低いグループから順に「グループ 1」、「グループ 2」として、最も棋力が高いグループを「グループ 6」と呼ぶ。各グループの棋譜の 90% を学習データ、5% を検証データ、5% をテストデータとした。テストデータについては、1 棋譜ごとに 40 手目以降の 1 局面のみを使用した。その理由は、棋譜が同じであれば勝敗結果が同じであり、何度も評価を行うとバイアスのかかった結果になる可能性があるか

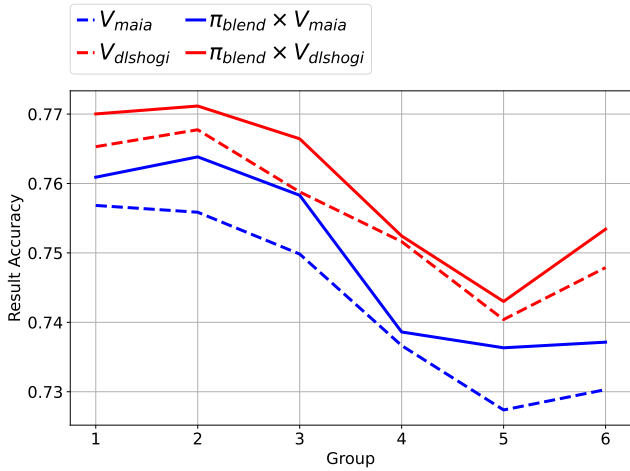


図 1 棋力帯ごとの既存モデルと Blend モデルを使った内積モデルの勝敗一致率。勝敗一致率が高いほうが勝率予測性能が良い指標である。

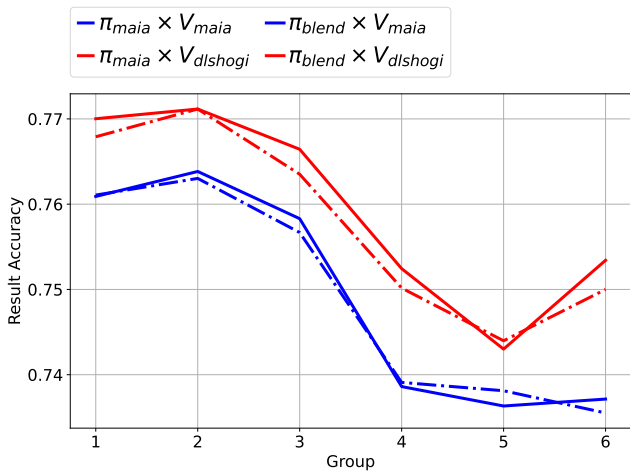


図 2 棋力帯ごとの Maia を使った内積モデルと Blend モデルを使った内積モデルの勝敗一致率。

らである。また、40 手目未満の局面は序盤から中盤の初めに該当し、勝率予測の評価に適していないと考えた。

本研究では勝率予測の評価指標として、勝敗一致率、交差エントロピー、期待校正誤差の 3 つを用いる。次節以降では、各評価指標にもとづく実験結果を検討していく。

4.2 勝敗一致率

勝敗一致率 (Result Accuracy) とは、手番側の予測勝率 $\hat{V} > 0.5$ の場合に最終的に勝利しているか、 $\hat{V} < 0.5$ の場合に最終的に敗北しているかの一致率である。特に観戦の場合、観測できる勝敗は 1 通りだけであるため、終盤の予測勝率と実際の結果が食い違っていると納得しづらく、重要な指標の一つだと考える。

図 1 は棋力帯ごとの既存モデルと Blend モデルを使った内積モデルの勝敗一致率を示したものである。線の種類はモデルの種類を表しており、破線は既存モデル、実線は Blend モデルを着手確率分布として使った内積モデルを表

している。線の色は使っている予測勝率の種類を表しており、青色は Maia-S-All の予測勝率、赤色は DLShogi の予測勝率を使用していることを表している。以降の図でも線の種類・色は同様のことを示す。

既存モデルの予測勝率 (破線) と Blend モデルを使用した内積モデル (実線) を比較する。Maia-S-All の予測勝率を使用したもの (青色) 同士では 0.2 ポイントから 0.8 ポイント、DLShogi の予測勝率を使用したもの (赤色) 同士では 0.1 ポイントから 0.8 ポイントほど内積モデルのほうが勝敗一致率が高い。

既存モデルである Maia-S-All の予測勝率 (青の破線) と DLShogi の予測勝率 (赤の破線) を比較すると、DLShogi のほうが 0.8 から 1.8 ポイントほど勝敗一致率について正確に予測できていることが分かる。着手予測においては、Maia-S-All のほうが DLShogi よりも人間の着手を正確に予測できるため、勝敗予測において DLShogi のほうが Maia-S-All よりも人間の勝敗予測を正確にできるということは意外な結果である。考えられる原因の一つとして、DLShogi と比較して Maia-S-All は学習した棋譜の量が少なく、十分に汎化するにはデータが足りない可能性が挙げられる。

全体的な傾向として、棋力が高いグループほど勝敗一致率は低い傾向にある。これは、棋力が高いほど局面が複雑になり、勝敗を予測するのが難しくなるためだと考えた。

図 2 は、着手確率分布として Maia-S-All を使った内積モデルと、Blend モデルを使った内積モデルの勝敗一致率を示したものである。一点鎖線は Blend モデルを着手確率分布として使った内積モデルを表している。棋力が高いグループほど勝敗一致率が低い傾向や、Maia-S-All の予測勝率を使ったモデルよりも DLShogi の予測勝率を使ったモデルのほうが勝敗一致率が高い、という傾向は変わらなかった。Blend モデルを使った内積モデル (実線) のほうが Maia-S-All を使った内積モデル (一点鎖線) よりも勝敗一致率が高いことが多いが、有意な差であるとは言えない。これは、Blend モデルのほうが着手予測の性能が高いが、Maia-S-All の着手と比べて勝敗が変わるほどの着手があまり存在しないためだと考えた。

4.3 交差エントロピー

交差エントロピーとは、手番側の予測勝率 $\hat{V}$ に対して、以下の式の平均値をとったものであり、分類問題の損失関数としてよく使用される。

$$\text{Cross Entropy} = \begin{cases} -\log \hat{V} & (\text{手番側勝ち}) \\ -\log(1 - \hat{V}) & (\text{手番側負け}) \end{cases}$$

実際の勝率を p とすると、 $\hat{V} = p$ のときのみ、交差エントロピーは最小値を取るという性質がある。4.2 節で扱った勝敗一致率は、最終的に勝った対局であれば、圧勝の局面

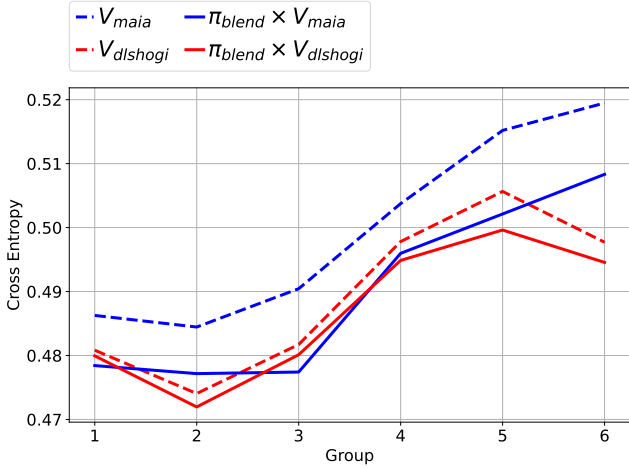


図 3 棋力帯ごとの既存モデルと Blend モデルを使った内積モデルの交差エントロピー（キャリブレーション後）。交差エントロピーは低いほうが勝率予測性能が良い指標である。

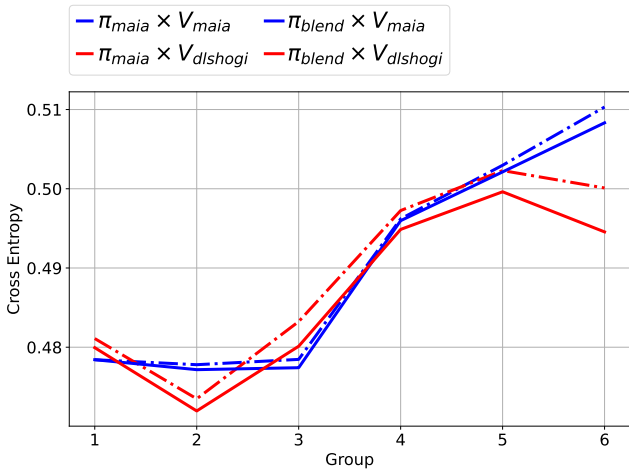


図 4 棋力帯ごとの Maia を使った内積モデルと Blend モデルを使った内積モデルの交差エントロピー（キャリブレーション後）。

で勝率 51%と予測しても正解となる。しかし実用的には、人間がプレイしたときの勝率と予測勝率が近いことが望ましい。交差エントロピーは、圧勝の局面で勝率 51%と予測するよりも、勝率 99%と予測するほうがより良い評価になる指標である。

平岡の研究 [6] で示されているように、AlphaZero のような強化学習手法で得られたモデルの予測勝率は 0 や 1 に近いことが多く、DLShogi にもこの傾向が見られる。本研究ではより公平な比較を目指し、モデルごとに予測勝率を 0.5 に近くしたり、0 や 1 に近くしたりするキャリブレーションを行う（付録参照）。なお、変換式の性質から、前節で扱った勝敗一致率はその影響を受けない。

図 3 は、棋力帯ごとの既存モデルと Blend モデルを使った内積モデルの交差エントロピーを示したものである。ここで、交差エントロピーは勝敗一致率と異なり、低いほうが勝率予測の性能が良い指標であることに注意する。既存モデルの予測勝率（破線）と Blend モデルを使用した内積

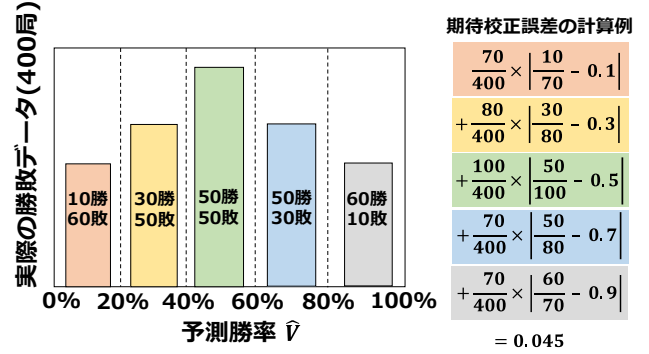


図 5 期待校正誤差の計算例。

モデル（実線）を比較する。Maia-S-All の予測勝率を使用したもの（青色）同士では 0.7 ポイントから 1.3 ポイント、DLShogi の予測勝率を使用したもの（赤色）同士では 0.1 ポイントから 0.6 ポイントほど内積モデルのほうが交差エントロピーが低い。

勝敗一致率とは異なる結果として、Maia-S-All の予測勝率を使用した内積モデル（青の実線）と、DLShogi の予測勝率を使用した内積モデル（赤の実線）では、対象の棋力帯によってどちらがより交差エントロピーが低いかが異なるということが挙げられる。このことから、予測勝率の性能がより低いモデルでも、内積モデルを使うことで、性能が入れ替わる場合があることが分かる。

既存モデルである Maia-S-All の予測勝率（青の破線）と DLShogi の予測勝率（赤の破線）を比較すると、DLShogi のほうが 0.5 から 2.2 ポイントほど交差エントロピーが低いことが分かる。全体の傾向として、棋力が高いグループほど交差エントロピーが高い傾向にあり、これは勝敗一致率の傾向とも合致している。

図 4 は、着手確率分布として Maia-S-All を使った内積モデル（一点鎖線）と、Blend モデルを使った内積モデル（実線）の交差エントロピーを示したものである。基本的には勝敗一致率と似た傾向であり、Blend モデルを使った内積モデルのほうがやや性能が高いことが分かる。

4.4 期待校正誤差 (Expected Calibration Error)

期待校正誤差はキャリブレーションの性能を評価する際によく使われる指標であり、単純な Maia-S-All の予測勝率や DLShogi の予測勝率を人間にとっての勝率により近づけようとする本研究の目的に合致している。この評価指標は、実際の勝敗データを予測勝率の範囲ごとに分割して実際の勝率を集計し、予測勝率との誤差を計算することで求まる。期待校正誤差は例えば、勝率を約 80%と予測する局面を棋譜 100 局から集めてきて、実際に 80 局程度勝っているのか、その予測勝率と実際の勝率について誤差を計算

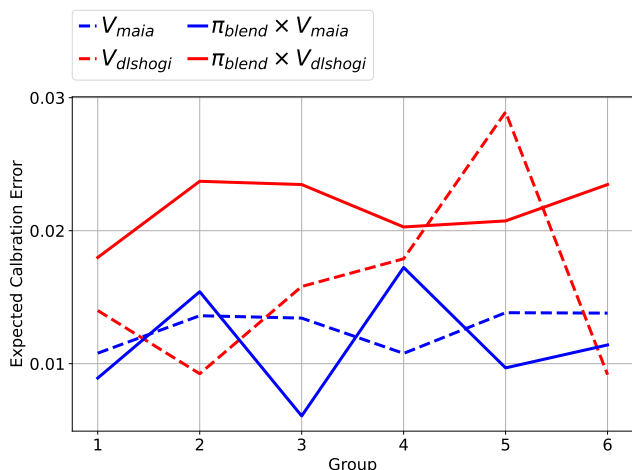


図 6 棋力帯ごとの既存・提案モデルの期待校正誤差（キャリブレーション後）。期待校正誤差は低いほうが勝率予測性能が良い指標である。

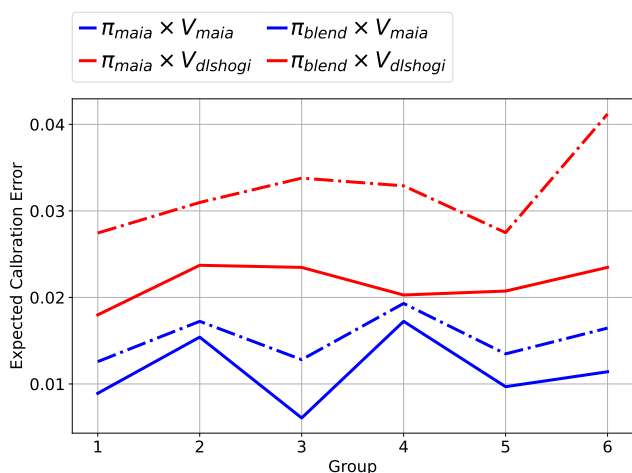


図 7 棋力帯ごとの既存・提案モデルの期待校正誤差（キャリブレーション後）。

する指標である。図 5 は期待校正誤差の計算例である。本研究では、予測勝率を 10% ごとに分割する 10 分割を採用する。

期待校正誤差は小さいほうが望ましいが、0 であっても適切な勝率予測モデルだとは限らない。例えば手番側の予測勝率を常に 0.5 と出力するモデルは期待校正誤差を 0 近くにすることができるが、これは望ましいモデルとは言えない。

図 6 は、棋力帯ごとの既存モデルと Blend モデルを使った内積モデルの期待校正誤差を示したものである。

予測勝率に DLShogi を使っているモデル（赤色）と予測勝率に Maia-S-All を使っているモデル（青色）を比較すると、DLShogi を使っているモデルのほうがばらつきが大きいことが分かる。モデルに使用しているキャリブレーションが交差エントロピー用に調整されているため、期待校正誤差の調整としてあまり適していない可能性がある。

図 7 は、着手確率分布として Maia-S-All を使った内積

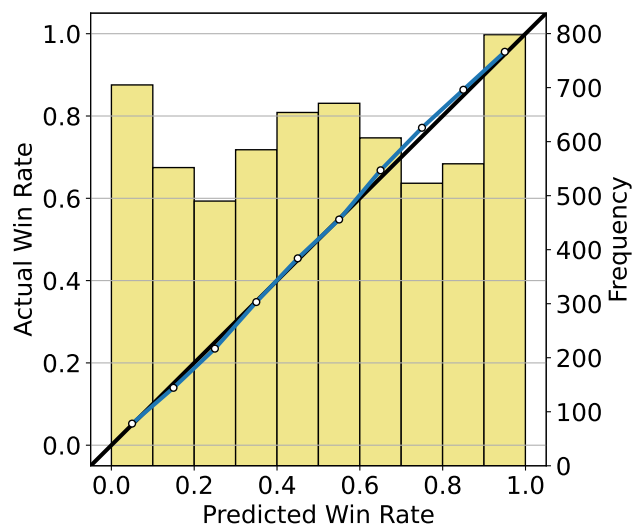


図 8 予測勝率と実際の勝率の関係、および予測勝率のヒストグラム。モデルは Blend モデルを使った内積モデル、対象の棋力帯はグループ 1。

モデル（一点鎖線）と、Blend モデルを使った内積モデル（実線）の期待校正誤差を示したものである。これまでの評価指標とは異なり、着手確率分布に Maia-S-All を使った内積モデルよりも、Blend モデルを使った内積モデルのほうが期待校正誤差が明確に小さいことが図から分かる。

本研究の期待校正誤差がどの程度の値なのか可視化するため、図 8 に予測勝率と実際の勝率の関係、および予測勝率のヒストグラムを示す。モデルは Blend モデルを使った内積モデル、対象の棋力帯はグループ 1 である。黒い直線は $y=x$ となる直線であり、青の折れ線と黒い直線の誤差に対して、ヒストグラムの頻度で重みづけしたものが期待校正誤差（0.01 程度）となる。図から分かる通り、Blend モデルを使用した内積モデルの予測勝率は、実際の勝率とかなり近い。

先述したように、状況が判断できていなくとも、勝率を 0.5 近くだと予測すれば期待校正誤差は 0 に近くなる。つまり同じ期待校正誤差であれば、予測勝率が 0 や 1 に近い極端な値を取るほど局面への理解がより深いと言える。そのため、勝率予測の正確さを評価したい場合は、期待校正誤差と、勝敗一致率・交差エントロピーのバランスを取れるような手法が望ましい。

5. おわりに

本研究では、人間の勝率予測の精度を測るため、既存モデルと提案モデルについて勝敗一致率・交差エントロピー・期待校正誤差の 3 つの指標で評価を行った。既存モデルとして、AlphaZero のような強化学習を行っている DLShogi の予測勝率や、教師あり学習を行っている Maia-S-All の予測勝率を用いた。提案モデルでは、これらの予測勝率に着手確率分布を組み合わせる内積モデルを提案した。

現在は1局面先の遷移のみを用いた内積モデルを提案しているが、実現確率探索 [7] のように、着手確率分布をもとにより深い探索を行う手法も有効だと考える。また、Maia-S-All の予測勝率と DLShogi の予測勝率を混ぜ合わせる手法も効果があるかもしれない。

謝辞

本研究は、JST 次世代研究者挑戦的研究プログラム JP-MJSP2102 の支援を受けたものです。また、本研究では、Mindwalk 株式会社様から「将棋クエスト」の棋譜提供を受けています。深く感謝申し上げます。

参考文献

- [1] McIlroy-Young, R., Sen, S., Kleinberg, J. and Anderson, A.: Aligning superhuman AI with human behavior: Chess as a model system, in *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 1677–1687 (2020).
- [2] McIlroy-Young, R., et al.: Learning models of individual behavior in chess, in *SIGKDD*, pp. 1253–1263 (2022).
- [3] Jacob, A. P., Wu, D. J., Farina, G., Lerer, A., Hu, H., Bakhtin, A., Andreas, J. and Brown, N.: Modeling strong and human-like gameplay with KL-regularized search, in *International Conference on Machine Learning*, pp. 9695–9728PMLR (2022).
- [4] Ogawa, T., et al.: More Human-Like Gameplay by Blending Policies from Supervised and Reinforcement Learning, *IEEE ToG* (2024).
- [5] Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., Lanctot, M., Sifre, L., Kumaran, D., Graepel, T., Lillicrap, T., Simonyan, K. and Hassabis, D.: A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play, *Science*, Vol. 362, No. 6419, pp. 1140–1144 (2018).
- [6] 平岡和幸：囲碁対局における推定勝率のキャリブレーション：KataGo の勝率 8 割は人間なら何割か, Technical Report 4, 亜細亜大学経営学部データサイエンス学科 (2024).
- [7] Tsuruoka, Y., Yokoyama, D. and Chikayama, T.: Game-tree search algorithm based on realization probability, *Icga Journal*, Vol. 25, No. 3, pp. 145–152 (2002).

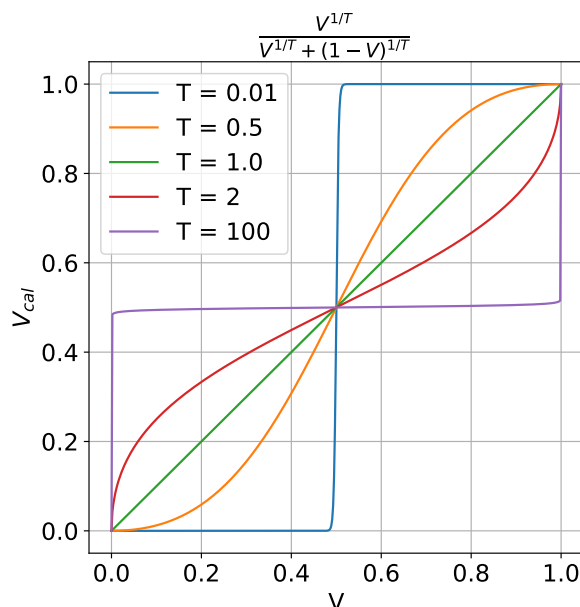


図9 式(2)による温度ごとの勝率キャリブレーションの例。

付録．DLShogi の予測勝率のキャリブレーション

平岡の研究で示されているように、AlphaZero のような強化学習から得られたモデル [6] は、0 や 1 に近い極端な勝率予測を行う傾向がある。図 10 は、グループ 1 の検証データに対しての、DLShogi の予測勝率のヒストグラムである。多くの局面に対して 0.0 から 0.1, または 0.9 から 1.0 の勝率を予測している。

この問題に対応するため、本研究では平岡が使用した以下の式によって予測勝率のキャリブレーションを行う。

$$V_{cal} = \frac{V^{1/T}}{V^{1/T} + (1-V)^{1/T}}. \quad (2)$$

図 9 に式 (2) による温度ごとの勝率キャリブレーションの例を示す。式 (2) は点 (0, 0), (0.5, 0.5), (1, 1) を通る単調増加関数であり、温度パラメータ $T \in [0, \infty)$ が小さくなるとほとんどの値が 0 や 1 に変換され、逆に温度パラメータ T が大きくなるとほとんどの値が 0.5 に変換される。

本研究では検証データを用いてキャリブレーションの温度パラメータ T を決める。ただし、期待校正誤差についてパラメータ調整を行うとほとんど 0.5 のみ出力されるような高すぎる温度パラメータになってしまう。また、勝敗一致率はキャリブレーションの影響を受けないため、交差エントロピーについて調整したパラメータ T を全体で使用する。温度パラメータの調整として、0.5 から 3.0 の範囲を 0.01 刻みで全探索し、検証データに対して最も性能が良いパラメータを採用する。

図 11 はグループ 1 の検証データに対して温度パラメータ T を調整した DLShogi の交差エントロピーである。DL-

温度 T	V_{maia}	$V_{dlshogi}$	$\pi_{maia} \times V_{maia}$	$\pi_{maia} \times V_{dlshogi}$	$\pi_{blend} \times V_{maia}$	$\pi_{blend} \times V_{dlshogi}$
グループ 1	1.03	2.61	0.98	2.43	0.98	2.46
グループ 2	0.98	2.45	0.93	2.30	0.93	2.31
グループ 3	0.87	2.21	0.85	2.11	0.85	2.12
グループ 4	0.89	2.24	0.85	2.10	0.85	2.13
グループ 5	0.88	2.23	0.88	2.08	0.88	2.11
グループ 6	0.80	2.04	0.78	1.89	0.78	1.93

表 1 グループごと、モデルごとの調整後温度パラメータ

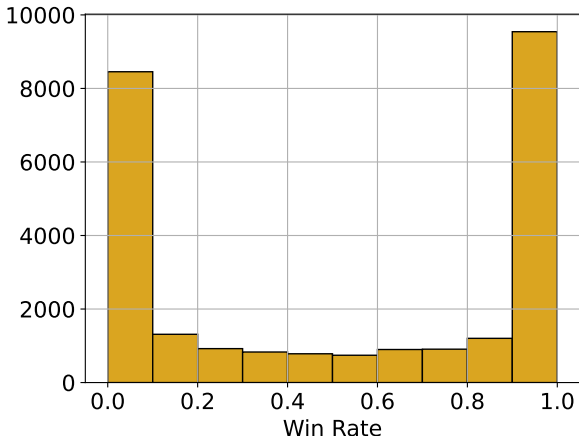


図 10 DLShogi の予測勝率のヒストグラム (グループ 1).

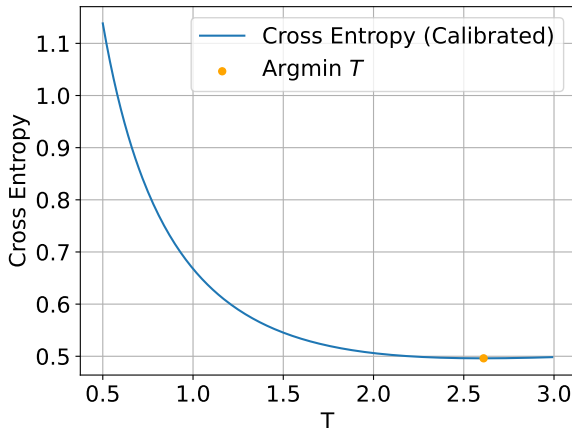


図 11 DLShogi の交差エントロピーのキャリブレーション例 (グループ 1).

Shogi のそのままの予測勝率，つまり温度パラメータが 1 のときは交差エントロピーが 0.7 ほどであるのに対し，最適な温度パラメータ 2.61 に対しては交差エントロピーが 0.5 付近に下がっている．他のモデル，他のグループに対してもこのように $T = 0.5$ から $T = 3.0$ の間に頂点がある下に凸な関数の形になっていた．最終的なグループごと，モデルごとの調整後温度パラメータを表 1 に示す．