

Title	局面の難しさを考慮したレーティング推定
Author(s)	西川, 就斗
Citation	
Issue Date	2025-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/19809
Rights	
Description	Supervisor: 池田 心, 先端科学技術研究科, 修士 (情報科学)

修士論文

局面の難しさを考慮したレーティング推定

西川 就斗

主指導教員 池田 心

北陸先端科学技術大学院大学
先端科学技術専攻
(情報科学)

令和7年3月

Abstract

In recent years, artificial intelligence (AI) technology has been extensively studied across various fields. In the field of games, powerful AI opponents such as dlshogi for shogi and AZdb for chess have achieved strength surpassing that of top human professionals.

Meanwhile, research is progressing not only on creating strong AI opponents but also on AI that entertains humans. For example, some studies focus on AI players that behave like humans to reduce unnatural behavior and create good opponents or cooperative teammates. Other research aims to estimate player skills and adjust enemy strength accordingly or to generate levels with appropriate difficulty. Among these studies on entertaining AI, we focus on rating estimation, a system that assesses player skill.

A rating system is a method for quantifying player skill numerically and is widely used for various purposes, such as player match-making, in-game handicap settings, and analyzing individual player growth. Therefore, it is important to compute player ratings as accurately as possible. Conventional rating systems, such as the Elo rating, update ratings based on the wins or losses of the matches. As a result, determining an appropriate rating often requires dozens or even more matches.

To estimate ratings with fewer matches, several methods have been proposed that utilize *loss*, a metric representing the magnitude of mistakes made in each move. A smaller loss indicates a stronger player, while a larger loss indicates a weaker player. This approach enables the distinction between crushing and close losses, allowing for a more accurate rating estimation that reflects match content even with a small number of games. Baba et al. further improved the accuracy by having players compete against a shogi AI designed to create close game situations. Additionally, they excluded certain positions according to the number of moves and remaining time, reducing the impact of noise in loss values that do not accurately reflect player skill.

In this work, we further focus on the fact that even players of the same skill level may show different loss values depending on the difficulty of the given positions. A player who faced a larger number of difficult positions by chance may show relatively high loss values, potentially leading to an underestimated rating compared to their actual skill level.

Therefore, in this study, we conducted rating estimation considering position difficulty, using Go as the target game. *Difficulty* is often vaguely defined, but in this research, we adopt *expected loss*, which represents how much error a player is expected to make on average in a given position. The expected loss is calculated as the weighted sum of the loss values of possible moves, where the weights are

their predicted selection probabilities. We use this value as a measure of position difficulty. A position with an expected loss close to zero indicates an easy position where mistakes are unlikely or not critical. Conversely, a higher expected loss indicates a difficult position where mistakes are more likely and have more severe consequences.

Using this expected loss, we adjust the loss of the actual move based on the position difficulty. If an actual move’s loss is moderate in a position with a high expected loss, it is interpreted as avoiding big mistakes in a difficult position, and the loss is adjusted downward. Conversely, if an actual move’s loss is moderate in a position with a low expected loss, it is interpreted as making an unnecessary mistake in an easy position, and the loss is adjusted upward.

First, we analyzed the desirable distribution properties of loss values when using loss-based rating estimation. An ideal loss distribution should accurately reflect player skill with minimal variance, regardless of how the game progresses. Additionally, it is preferable that the distributions of different skill levels (i.e., weaker and stronger players) are well-separated. To evaluate whether these conditions are met, we employed three metrics: stochastic dominance, overlap ratio between distributions, and performance in a trinary classification task.

Next, to validate the appropriateness of the evaluation metrics, we reexamined the effectiveness of the method proposed by Kuboki et al., which demonstrated that using the average loss over the first 150 moves, rather than the entire game, led to more reliable estimations. The results confirmed the effectiveness of this method across all three proposed metrics, thereby also justifying the validity of the metrics themselves.

Finally, we conducted an experiment to estimate player ranks (approximately equivalent to ratings) based on ten game records per player, ranging from 10 kyu to 9 dan. The estimation accuracy was evaluated by measuring the deviation from the players’ actual ranks. The results showed that using raw loss values resulted in a root mean squared error (RMSE) of 3.975[rank], while applying the method of Kuboki et al. reduced the RMSE to 1.935[rank]. Our proposed method, which incorporates adjustments based on expected loss, further reduced the RMSE to 1.511[rank].

We expect the findings of this work to be applicable not only to rating estimation in competitive games such as Go but also to various other domains. For instance, they may be utilized in AI-driven personalized teaching that fits students’ skill levels, as well as in the generation of interesting puzzles.

概要

近年、人工知能技術（AI 技術）は様々な業界で幅広く研究され続けている。ゲーム分野では、将棋の dlshogi やチェスの AZdb などの強い相手 AI が人間のトッププロを超えるほどの強さを獲得している。

一方で、強い相手 AI を作るだけでなく、人間を楽しませる AI に関する研究が進められている。例えば、人間らしく振舞う AI プレイヤを用いて、不自然な挙動を抑制し、良いライバルや良い仲間を演出する研究がある。あるいは、プレイヤの実力を推定しながらちょうどよい強さに敵を調整したり、ちょうどよい難易度のステージを生成するなどの研究もある。このような、楽しませる AI の研究のなかでも、本研究は、プレイヤの実力を推定してくれるシステムである、レーティング推定に着目した。

レーティングシステムとは、プレイヤの実力を数値で表そうとするものであり、プレイヤーのマッチングやゲーム内のハンデ設定、各プレイヤの成長過程を分析するなど幅広い目的で用いられている指標である。そのため、プレイヤの実力をできるだけ正確に算出する事は重要である。Elo レーティングに代表される通常のレーティングシステムは勝敗によってレーティングの更新を行うため、適切なレーティングが判明するまでに数十またはそれ以上の試合数を要する。

レーティングをより少ない試合数で推定するために、1 手ごとのミスの大きさを表す指標である loss を用いる手法がいくつか提案されている。loss が小さいほど強く、大きいほど弱いプレイヤであるということになる。これにより、大敗と惜敗を区別することが可能になり、少ない試合数でも試合内容を反映して精度の高いレーティング推定が可能になった。馬場らはさらに、プレイヤに接戦の局面にするような将棋 AI と対局させ、分析対象となる局面を手数や残り時間により除外することで、力量を反映しないような loss のノイズの影響を抑制し、精度を高めることに成功した。

本研究ではさらに、同じ力量のプレイヤであっても、与えられる局面の難しさによって見かけの loss は変わるはずだという点に着目した。たまたま難しい局面を多く与えられたプレイヤは、相対的に高い loss を示してしまうことにより、力量よりも低いレーティングであると推定されてしまうかもしれない。

そのため、本研究では、囲碁を対象ゲームとして、局面の難しさを考慮したレーティング推定を行った。難しさとは一般的に曖昧に定義されることが多いが、本研究では、ある局面において対象プレイヤが「平均的にどのくらいのミスをする」と予想されるか」という値である“期待 loss”を用いることにした。これは、各着手のミスの大きさを表す loss と、各着手の予想選択確率を重みづけ加算することで得られる期待値であり、各局面の難しさを表す値として扱う。期待 loss が 0 に近いほど、ミスをしにくい、またはミスが致命的でない簡単な局面を表す。期待 loss が大きくなるほど、ミスをしやすく、かつミスが致命的になりやすい困難な局面であることを表す。

この期待 loss を用いて、期待 loss が大きい局面で実着手の loss が中程度ならば、

難しい局面で大きなけがをしなかったと解釈し、loss が低くなるように補正する。また、期待 loss が小さい局面で実着手の loss が中程度ならば、簡単な局面であるべきでない損をしたと解釈し、loss が高くなるように補正する。

まず、loss によるレーティング推定を行う場合に、loss がどのような分布になるのか望ましいかを考察した。試合展開によらずに力量を反映した、ばらつきの少ない分布が理想である。そのうえで、区別したい分布間（弱い人と強い人の loss の分布）が離れていることが望ましい。そこで、それを満たすかを判断する指標である、優越率、分布同士の重なり面積割合を用いた評価、および、3 クラス分類での評価を行うことにした。

次いで、指標が適切なものかを判断するために、窪木らによる先行研究で有効性の示された「一局通しての loss の平均値ではなく、150 手までの平均値を取る」という手法の有効性を再確認した。提案した3つの指標すべてで、この手法が有効であることが分かり、指標そのものも正当化された。

最後に、10 級から 9 段までのプレイヤーについて、プレイヤーごと 10 局の棋譜から段級位（≡レーティング）を推定し、本来の段級位からどれだけ離れているかを評価する実験を行った。その結果、loss をそのまま使う場合は二乗平均誤差が 3.975[段・級]、窪木らの手法を使った場合は 1.935[段・級] となるところ、期待 loss による補正を行った提案手法では、1.511[段・級] まで誤差を抑えることができた。

本研究の成果は、囲碁などの競技におけるレーティング推定のみならず、AI による生徒の力量に合わせた適切な指導や、パズルゲームの面白い問題生成にも活用可能と考えている。

目次

第1章	はじめに	1
第2章	レーティング推定とその課題	3
2.1	Elo レーティング	3
2.2	先行研究	4
第3章	局面の難しさとは	6
第4章	目的とアプローチ	8
第5章	実験設定	10
5.1	対象ゲームと棋譜データ	10
5.2	ばらつき抑制のための手法の評価指標	11
5.2.1	優越率	12
5.2.2	分布の重なり面積	12
5.2.3	クラス分類	12
5.3	手数を変えることによる評価指標の改善の確認	13
第6章	期待 loss を用いた補正	16
6.1	loss の補正式	16
6.2	実験	17
6.2.1	loss 補正式の3つの指標による評価	17
6.2.2	loss 補正式を用いた少数棋譜でのレーティング推定	19
第7章	おわりに	21

目 次

4.1	難しい局面で大けがをしなかった場合の例	9
4.2	簡単な局面ですべきでない損をした場合の例	9
5.1	重なり部分が多い2分布	11
5.2	分散が小さい2分布	11
5.3	分散, 平均の差がともに小さい2分布	11
5.4	半分重なった分布	12
5.5	分離した分布	12
5.6	1試合の loss 平均のヒストグラム	13

表 目 次

5.1	優越率による性能比較	14
5.2	分布の重なり面積による性能比較	14
5.3	平均 loss のみで 1d,4k,8k のどれなのかを推定するクラス分類問題の 正答率	15
6.1	優越率による補正式の効果の比較 (1 試合平均)	18
6.2	分布の重なり面積による補正式の効果の比較 (1 試合平均)	18
6.3	クラス分類による補正式の効果の比較	18
6.4	3-fold 交差検証の結果	19
6.5	テストデータでの性能評価の結果	20

第1章 はじめに

近年、人工知能技術（AI 技術）は様々な業界で幅広く研究、利用され続けている。OpenAI の ChatGPT や Google の Gemini をはじめ、医療への応用、自動運転の実用化など、高度な AI が我々の生活に浸透しつつある [1][2]。

ゲーム分野は、ルールが明確に定義できる、結果を明確に示せる、対戦データが収集しやすいという特徴から、AI の良い研究対象として用いられてきた。将棋の dlshogi¹やチェスの AZdb[3] などの強い相手 AI は、人間のトッププロの実力を超えるほどの強さを獲得している。

また、ゲームの研究は、単に強さを追求するだけにとどまらず、多岐にわたるテーマで進められている。人間らしく振舞う AI プレイヤ [4] は、対人間プレイヤとのゲームが上達するために有効であり、パズルゲームの練習問題 [5] や RPG のダンジョン [6] などのコンテンツを自動生成する AI は、ゲームの多様性を生む。さらに、ゲーム内のキャラクターや道具のステータスを自動調整する AI [7] や、プレイヤの実力を推定してくれるシステム [8] は、正しいバランス調整をすることでプレイヤの不満を減らすことができる。これらの研究のように、強い相手 AI を作ることから、人間を楽しませる AI に関する研究が進められており、ゲーム内の様々な場面で使われている。本研究では、人間を楽しませる AI の研究のうち、プレイヤの実力を推定してくれるシステムである、レーティング推定に着目した。

ここで、レーティングとは、プレイヤーの実力を数値で示す指標である。互角のプレイヤは同じレーティングを持つように、強いプレイヤほどレーティング値が大きくなるようにレーティング値を定義する。ゲーム内において、レーティングはプレイヤーのマッチング、ゲーム時のハンデ設定、自身の実力を示す指標として用いられる。また、プレイヤのみならず、ゲーム外のファンにとっても、勝敗予測、トッププレイヤの勢力図の把握、各プレイヤの成長過程の分析など、レーティングはさまざまな目的で活用できる有益な指標である。したがって、レーティングをできるだけ少ない試合数で、できるだけ正確に算出することは極めて重要である。

レーティング推定する方法の中で、現状最も多く使われているのが Elo レーティングである [9]。Elo レーティングは将棋や格闘ゲーム、サッカーなどの対戦ゲームで用いられているシステムであり、対戦相手の実力と試合結果により、相対評価で実力を表すために使われている。しかし、試合結果のみで実力を推定してい

¹dlshogi [https://github.com/TadaoYamaoka/DeepLearningShogi]（最終アクセス：2025/02/02）

る観点から、実力を適切に表したレーティングになるまでの必要試合数が多くなることが問題点である。

試合結果以外の情報を用いて、より精度の高い推定を行おうとする取り組みもある。バックギャモンでは、プレイヤーの強さを表す指標として Performance Rating (PR) という、コンピュータの計算した最善手に比べて、プレイヤーの着手がどの程度損する手なのかという値 (エラー) を用いている [10]。馬場らも、将棋における棋力推定のために、プレイヤーの着手がコンピュータの最善手に比べてどの程度損したのかという値 (loss) を用いている。大差の試合では loss が大きく出過ぎる傾向があるため、馬場らは接戦の局面を多く作る工夫を行い、4局程度でもある程度の精度で推定ができると主張している [11]。

上記で述べたように、既存のレーティング推定手法の多くでは、局面の進捗や難しさによらずにミスの大きさを計算、平均して推定に用いてきた。しかし、多くの人がミスをする局面でミスをするのと、ほとんどの人が最善手を選べる局面でミスをするのでは、罪の重さは違うはずである。そこで本研究では、先行研究にある1手ごとのミスの大きさに加えて、“局面の難しさ”という要素を考慮することによってレーティング推定する方法の提案を行う。局面の難しさを考慮することで、同じ負け試合であっても、難しい局面が多かったためにミスが多くなり敗北したのか、あるいは、簡単な局面が多かったにもかかわらずミスを多くしてしまったために敗北したのかを区別することが可能となる。また、局面の難しさを推定することは、レーティング推定以外にも応用することができる。例えば、ゲームの一局面を抜き出して問題とした詰将棋やなぞぶよなどの問題作成時に、局面の難しさを推定できれば、適切な難易度の問題を作成することが可能になる。また、人間プレイヤーを指導するAIは、局面の難しさを推定した結果を用いれば、プレイヤーの実力帯に合わせて指導することができる。このように、局面の難しさの推定が可能になれば、ゲームの様々な場面に応用できる。

そこで本研究では、局面の難しさを、ある局面において対象プレイヤーが「平均的にどのくらいのミスをする」と予想されるか」という値と定義した。定義した値を用いて、1手ごとのミスの大きさ (loss) の値を補正することで、レーティング推定を行った。その結果、loss をそのまま使う場合は二乗平均誤差が 3.975 [段・級]、窪木ら [8] の手法を使った場合 1.935 [段・級] になるところ、提案手法では、1.511 [段・級] まで誤差を抑えることができた。

本論文の構成は以下の通りである。2章は、現状のレーティング推定についてをより詳細に説明し、その問題点を提起する。3章は、局面の難しさとは何を指すのかを説明する。4章は、本研究の目的とそのためのアプローチについて述べる。5章は、実験に用いるデータの説明と、それを用いた予備実験を行う。6章は、局面の難しさを表すような期待 loss という指標を用いて、実際にレーティング推定を行う。7章は、本研究をまとめたうえで、今後の展望を記す。

第2章 レーティング推定とその課題

レーティングシステムとは、プレイヤーの実力を数値で示そうとする指標であり、システムによって計算された値がレーティングである。レーティング推定するシステムについて、いくつかの指標が提案されている。レーティングはゲーム内のプレイヤーやゲーム外のファンにとって、幅広い目的で用いられている指標である。ハースストーンでは、プレイヤーに見えるレーティングである外面的レーティングと、プレイヤーから隠されている内部レーティングを用いている [12]。プレイヤー自身の実力を示すもの、一定期間ごとのプレイヤーのプレイ度合いを表すものとして外面的レーティング、プレイヤーのマッチングとして内部レーティングを使っている。また、日本将棋連盟はプレイヤー間の段位差を用いて、ハンディキャップの度合いを公式に発表している [13]。これを用いることで、上級者と対等に真剣勝負をすることが可能になる。さらに、将棋において、これまでの成績と今後の相手から、誰がどのくらいの確率でタイトルを挑戦しそうかを推測することができる [14]。したがって、レーティングをできるだけ少ない試合数で、できるだけ正確に算出することは重要である。

本章では、レーティング推定において、著名なレーティングシステムである、Elo レーティングについての説明とその課題点、レーティング推定に関する先行研究の紹介とその課題点を述べる。

2.1 Elo レーティング

Elo レーティングの仕組みと特徴について、Web での記述 [9] を参考にして説明する。Elo レーティングは数学的に裏付けのある最も著名なレーティングシステムであり、さまざまな改良も試みられている。当初はチェスの実力を表すために用いられたものであり、そこから、将棋や囲碁のオンライン対局や対戦型オンラインゲームのマッチング、サッカーの FIFA ランキングなどにも元祖または改良版が採用されている。

Elo レーティングは、対戦前の相互のレーティング値に基づき期待勝率を計算し、期待勝率と実際の対戦結果との差異に基づいてレーティングを更新する。プレイヤー A とプレイヤー B の対戦における期待勝率は、式 (2.1) で表される。

$$W_{BA} = \frac{1}{10^{(R_A - R_B)/400} + 1} \quad (2.1)$$

ここで、 R_A はプレイヤー A の Elo レーティング値、 R_B はプレイヤー B のレーティング値を示す。また、プレイヤー A のレーティングの更新式は、式 (2.2) で表される。

$$R_{A_{new}} = R_{A_{old}} \pm K \times W_{BA} \quad (2.2)$$

ここで、 $R_{A_{old}}$ は対戦前の Elo レーティング値、 $R_{A_{new}}$ は対戦後の Elo レーティング値、 W_{BA} は B が A に勝つ期待勝率、 K は自由定数であり、勝った場合に符号が+、負けた場合に符号が-となる。 K に使われる値として 16 や 32 が多い。このように、勝ったプレイヤーが増えたレーティングの値だけ、負けたプレイヤーのレーティングが減るような仕組みになっている。

Elo レーティングは試合結果と試合前の両プレイヤーのレーティングが分かれば、レーティング値を更新することができるというのが最大の特徴である。しかし、試合結果のみで推定しているため、新規プレイヤーの初期レーティングの設定が難しく、その結果、新規プレイヤーと対戦相手の実力差がレーティング差以上にあった場合には、レートの変動幅が大きくなることや、新規プレイヤー自身の適性な実力を表すレーティング値になるまでに試合数が多くなることが問題点である。さらに、相対的な評価であるため、平均的な数値の水準が過剰に上昇（いわゆるインフレ）または低下（いわゆるデフレ）してしまうことがあり、最新のレーティングと昔のレーティングでは大きく異なるため比較が困難であることが課題として挙げられる。

2.2 先行研究

レーティングを推定する方法として、試合結果以外の情報をもとに、より精度の高い結果を追求した先行研究がある。

森田らは、大貧民におけるレーティング推定方法として、大貧民は、プレイヤーの強さだけでなく、初期手札の良さが対戦結果に大きく影響されることから、Elo レーティングに加え、初期手札の不均衡性を考慮したレーティングアルゴリズムを提案した [15]。提案されたレーティングアルゴリズムは、2 人仮想ゲームによる検証において、初期盤面勝率が極端だったときに、実際のレートより 200 低いレーティング結果を示したのに対し、提案手法では正しいレーティング結果を示すことができている。しかし、初期手札を考慮しているとはいえ、試合中の行動は評価せず、試合結果を参照しているため、Elo レーティングの問題点であるように、適正な実力を表すようなレーティングになるまでに時間がかかってしまう。

山下らは、勝敗の結果と棋譜の内容から将棋の歴代名人の強さを推測した [16]。将棋倶楽部 24 の棋譜データを用いて、平均悪手、平均好手、好手率、一致率、複雑さの各指標とレーティングデータの関係性を調べることで、平均悪手をもっとも棋力を推測できることを示した。本論文での平均悪手とは、強い将棋 AI である

Bonanza と違う手を指してかつ評価が落ちた場合の平均したもの。複雑さとは、探索途中で変化する評価値の差を合計したものと定義している。平均悪手を用いることで、20 局程度で棋力を推定できると述べている。しかし、平均悪手は悪手が多かった試合と少なかった試合とで、ばらつきが大きいことが難点である。

馬場らは、将棋における棋力推定方法として、接戦の局面が棋力推定に有効に働くことを発見した [11]。手法として、接戦の局面にするような将棋 AI との対局をすること、上手下手関係なく定跡の手になりやすい序盤の 40 手を分析局面から除外すること、焦りから指し手が悪くなることを考慮して持ち時間残り 1 分を下回った手を分析局面から除外することなどの工夫をすることで、平均損失 (loss) を用いてレーティングを推定し、4 局程度で高い推定精度の結果を得ることができた。しかし、接戦の局面を作り出すような将棋 AI との対局をすることが条件となっているため、実際の大会や対局などの棋譜からプレイヤーのレーティングを推定することは難しい。

窪木らは、囲碁における棋力推定手法として、損失（ゲーム AI の最善手とプレイヤーの着手の評価値の差分）と prior（人間の棋譜を学習させたネットワークの選択確率）の 2 つの指標で棋力推定を行った [8]。損失は手数によって平均的な損失が異なることを用いて、損失に対して手数のフィルタと補正をすることで、平均二乗誤差が 1.182[段・級] となった。また、prior は算術平均を取ると非常に低い手の影響が抑制されてしまうため、手数フィルタと幾何平均をとることで、平均二乗誤差が 0.955[段・級] となった。しかし、局面の影響を考慮しているが、1 局の平均を取っているため、1 手ごとの細かな情報が落ちてしまう。1 手ごとのミス大きさを表す loss を使う手法は、レーティング推定に良い指標であるが、loss はノイズを受けやすい指標であり、ノイズ要因の 1 つとして、局面の難しさがあると考えられる。そのため、本研究では、各局面における 1 手ごとのミスの大きさ（損失）である loss を用いること、また、それぞれの“局面の難しさ”を考慮することで、精度が高いかつ少ない試合数でプレイヤーの実力を推定する手法を提案する。

第3章 局面の難しさとは

一般に“難しさ”もしくは何かの事象が“難しい”ということを述べる際、どういった意味で用いられているかが曖昧に使われることが多い。例えば、「課題を提出することが難しい」という文を考えた際に、

- 話者にとって、提出期限が短すぎるため、難しい
- 話者からすると、課題の内容が複雑で理解できないため、難しい
- 話者の理解力で、課題を解くことができないため、難しい
- 話者の環境では、課題を提出することができないため、難しい

など、たくさん難しい原因を考えることができるが、一般的な表現として、これらの原因の1つ、もしくは複数の意味を込めて、単に“難しい”と表現することが多い。

さらに、原因についてだけでなく、難しさの程度や、どういう結果を招きそうなのかについても、先の文「課題を提出することが難しい」だけでは明確に定義できていない。例えば、以下のように難しさや結果などが異なる場合でも、すべて「課題を提出することが難しい」という表現になりえてしまう。

- 提出そのものはできるが、合格点は取れないだろう場合（意味合いが違う）。
- 何らかの理由で、全く可能性はない（表現の問題）。
- 普通に生活しては間に合わないが、徹夜すれば出せるだろう場合（時間を要する）。
- 普通に予習復習するだけでは提出できないが、専門書を読むなど特別な努力をすれば出せるだろう場合（努力や能力を要する）。
- 郵送しなければいけないが、時間内に届く確率はかなり低い（運を要する）。

そのため、“局面の難しさ”を考える際に、どのような意味で難しいを定義するのは非常に重要である。例えば、「プロが最善手を選べる確率が低い局面」と「想定するプレイヤーが最善手を選べる確率が低い局面」は違う。また、「最善手を選べる確率」と「満足手を選べる確率（勝敗に直結するようなとても悪い手を選んで

しまいやすいか)」も違う。以降では、先行研究で各ゲームにおいて“難しさ”はどのような意味で定義されているか、また、“難しさ”をどのように推定するのかを説明する。

相良らは、上海ゲームにおいて、人間らしい忘却をソルバに取り入れることで、人間にとって記憶負担の少ない問題を生成することを目指した [17]。そのために、まず、強いソルバ、弱いソルバの2つの強さの異なるソルバを作成し、クリア率を測ることで、問題を分類し、強いソルバのクリア率が高く、弱いソルバのクリア率が低い問題を良い問題とした。つまり、「技量面での難しさ」を重視し「運を要する難しさ」を排除し、技量さえあれば運に左右されずに勝てる問題を良いとしている。さらに、良いと判断したもののうち、忘却ソルバのクリア率が高い問題を工夫が報われる問題とした。すなわち、「技量面での難しさ」の中から「捨て牌を覚える難しさ」を排除し、覚える苦しさよりも考える楽しさをもたらし問題を良い問題としている。被験者実験で評価した結果、良い問題のうち、忘却ソルバのクリア率が高い問題が、クリア率が低い問題よりも面白い問題である可能性を示した。

石井らは、ガイスターにおいて、将棋における詰将棋のような、詰めガイスター問題を提案した [5]。また、問題を効率よく生成するためのアルゴリズムで、どれほどの時間で生成できるかを検証した。さらに、詰めガイスター問題の面白さや難しさの特徴を、被験者実験で確認し、教師あり学習を用いて、局面の特徴量から面白さや難しさの推定を行った。ここでは、被験者に「感じた難しさ」を問題ごとに評価してもらっている。すなわち、さまざまな意味での難しさを分解せずに、被験者ごとの主観による難しさをそのまま難しさの定義として用いている。一方で、局面の特徴量から難しさの推定を行った部分では、“初期状態で調べる必要がある局面の量”が推定に高い重要度を持っていることを示した。このように、ゲーム探索木がどのような特徴（広さ、深さなど）を持っているかによって難しさを定義または推測する方法は、チェス [18]、数独 [19] などさまざまなゲームで用いられている。

第4章 目的とアプローチ

本研究は、レーティング推定において、できるだけ少ない試合数で、正確にプレイヤーの実力を測ることを目指している。従来の Elo レーティングシステムでは、勝敗のみに着目しているため、僅差と大差は同じとみなし、相手の出来不出来にも依存する可能性があった。そのため、先行研究などは、1手ごとのミスの大きさを表す指標である loss を用いてより高精度にレーティングを推定することを試み、一定の成果を挙げた。しかし、同じミスの大きさだったとしても、全員がミスをするような場合と、ほとんどのプレイヤーがしないミスをする場合を同じとみなしてしまい、局面の難しさに影響されてしまうことが課題であると考えた。そのため、loss に、各局面の難しさを考慮したレーティング推定方法を提案する。

局面の難しさには3章で述べた通りさまざまな解釈や定義がありうるが、本研究では、その局面で当該プレイヤーが「平均的にどのくらいのミスをするか」という値、“期待 loss”を使うことにする。期待 loss は、各着手のミスの大きさを表す loss と、各着手の予想選択確率を用いて計算する。

loss とは、各着手のミスの大きさを表し、対象のゲームによって、その評価値の基準は異なる。将棋では、局面の評価値を点数で表し、loss は最善手を指した場合の点数に比べ、実際の着手がどれほどの点数を損しているかを表す。囲碁では、予測目数差を用いて、loss は最善手を打った場合と比べ、実際の着手がどれほどの目数を損しているかを表す。一方着手の予想選択確率は、その局面の合法手ごとに与えられた「そのプレイヤー、または同じ棋力帯のプレイヤーがその手を選ぶだろう確率」のことである。人間プレイヤーの棋力帯または年代ごとの選択確率は、通常、そのプレイヤーの棋譜を教師あり学習したもので近似することができる。有名なものとして、チェスの Maia[20] や、囲碁の Katago にその機能がある。

期待 loss は、先で述べた loss と着手確率を期待値のように、掛け合わせることで、各局面の難しさを表す値として扱う。具体的な式は式(4.1)で表される。この指標は、0に近いほど、見やすい局面、打ちやすい局面、簡単な局面を表し、大きくなるほど、見にくい局面、打ちにくい局面、困難な局面といえるようなものである。例えば、一見自然で良い手に見える手が大悪手であるような局面はこの値は高くなる。一方、最善手を選ぶことは難しくとも、それを外しても大きな損失にならないような局面、あるいは逆に、大悪手はあってもまず選ばないような局面では、それほど高くない。

$$E[\text{loss}] = \sum_{a \in A_N} \text{loss}_a \times \text{Prior}_a \quad (4.1)$$

この期待 loss を用いて以下のような補正を行う．

- (1) 期待 loss が大きい局面なのに、実着手の loss が中程度であれば、難しい局面で大けがをしなかったと解釈し、loss は低くなるように補正する．
- (2) 期待 loss が小さい局面なのに、実着手の loss が中程度であれば、簡単な局面でするべきでない損をしたと解釈し、loss は高くなるように補正する．

図 4.1 は、(1) の例である．この局面は一見 A に打ちたくなってしまう (2k の選択確率=19.7%) が、実際には loss=8.6 の大悪手である．このような局面では、期待 loss も 2.77 と大きくなる．ここで、あるプレイヤーが X に打ったとすると、その手は loss=0.79 の、最善とは言えない手であるが、A ほどは悪くない．この局面でそれほど悪くない手を選べたことから、loss は 0.79 のままではなく、例えば 0.51 にして、レーティング推定に用いる平均 loss に過剰な影響を与えないようにしたほうがよいだろう．

図 4.2 は、(2) の例である．この局面の選択確率上位 3 手 (A,B,C) は大きくても loss が 0.16 である．このような局面では、期待 loss も 0.96 と小さくなる．しかし、あるプレイヤーが X に打ったとすると、その手は loss=5.56 の大悪手である．この局面で悪手を打ったことから、loss は 5.56 から、例えば 5.91 として、レーティング推定の平均 loss に影響を与えるべきだろう．

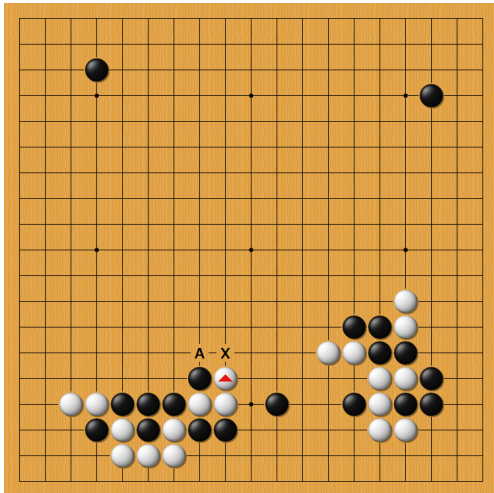


図 4.1: 難しい局面で大けがをしなかった場合の例

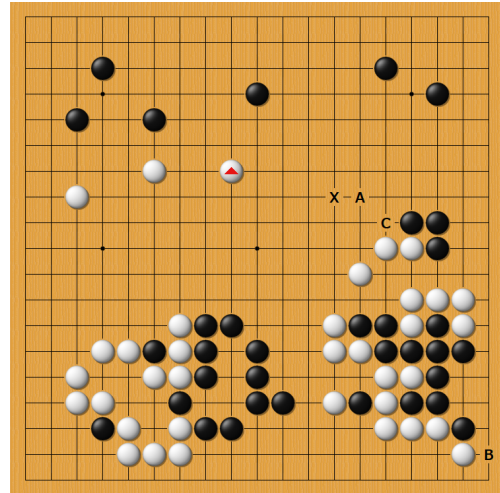


図 4.2: 簡単な局面でするべきでない損をした場合の例

第5章 実験設定

本研究では、期待 loss を用いた loss 補正により、一人のプレイヤーの試合ごとの loss のばらつきを抑え、それによって精度の高いレーティング推定を行うことを目指す。本章では、loss のばらつきがどんな工夫でどの程度抑えられるかを調べるための実験について述べる。5.1 節では、本研究で用いる対象ゲームと棋譜データについて説明する。5.2 節では、ばらつきを抑制する効果を評価するための複数の指標を説明する。5.3 節では、この指標の利用例として、loss 平均をとる区間を変化させた場合に指標がどう変わるかを調べた予備実験の結果を紹介する。

5.1 対象ゲームと棋譜データ

本研究では、対象ゲームに囲碁を選択した。囲碁は、2 人ゲームであり、プレイヤーが交互に黒石と白石を置くゲームである。盤面の中で自分の色で囲まれた陣地を確保しあい、陣地（目数）の大きくするのが目的である。先手の黒番は有利とされているため、ハンディキャップとして、後手の白番は 6.5 目与えられ、目数の多い方が勝利となる。

また、棋譜データを解析するために、Katago¹を用いた。Katago は強力な囲碁 AI の 1 つであり、デフォルトのネットワークは自己対戦のみで学習を行っている。loss を計算する際に、Katago で探索された最善手より何目損した手を打ったかで判断した。さらに、大量の人間の棋譜を学習させたネットワークが公開されており、段級位を設定することで、該当する着手の選択確率を手に入れることができ、これを本研究では、着手確率として用いる。本実験では、10k（10 級）, 8k, ..., 2k, 1d, ..., 7d, 9d（9 段）についての着手の選択確率を取得した。

本研究では、棋譜データとして、インターネット囲碁対局サービスである野狐²から、以下の 3 つの条件を全て満たすものを抽出した。

- 対局者どちらか一方の段級位が 10k, 8k, ..., 7d, 9d の対局
- 対局者間の段級位差が 1 以下である対局
- 100 手以上続いた対局

¹Katago [<https://github.com/lightvector/KataGo>]（最終アクセス：2025/02/02）

²野狐 [<https://github.com/featurecat/go-dataset>]（最終アクセス：2025/02/02）

5.2 ばらつき抑制のための手法の評価指標

着手の loss を用いてレーティング推定をする場合、各プレイヤーの1試合あたりの平均的な loss を基準に判断するのが普通である。しかし、どんなプレイヤーであっても試合ごとにさまざまな要因で出来不出来が生じるため、推定に用いる試合数が少ないほど、過大評価、過小評価が生じやすい。

図 5.1 は、あるプレイヤー A (赤) と、あるプレイヤー B (青) の、1 試合あたりの平均 loss を確率分布で示したものである。A は B よりも強いプレイヤーであり、全体的には loss は小さい。しかし、たまたま A が難しい局面を多く与えられることで高い loss (右側) を出してしまい、逆に B がたまたま簡単な局面を多く与えられることで低い loss (左側) を出した場合は、loss だけ見れば B のほうが強く見えるということになってしまう。

ここで、もしうまく補正ができれば、A と B 本来の力量が見分けやすくなるかもしれない。図 5.2 はそのような例であり、平均値の差は同じまま、それぞれ分散を小さくすることができている。この場合、分布 A と B からランダムに loss をサンプルしたとしても、B のほうが小さくなる (強く見える) ことはほとんどない。

“ばらつきを抑制”という言葉は、分散を減らせばよいというふうに聞こえるかもしれない。しかし、図 5.3 に示すように、分散を減らすことで、2 つの分布の平均の差も減ってしまえば、誤判定を導く恐れを減らすことはできない。

本節では、ばらつき抑制のための手法がどれほどの効果をもたらすか、実際のレーティング推定を行わずに、“2 分布間にはっきり差があるか”を数値化することで評価するためのいくつかの指標を提案する。

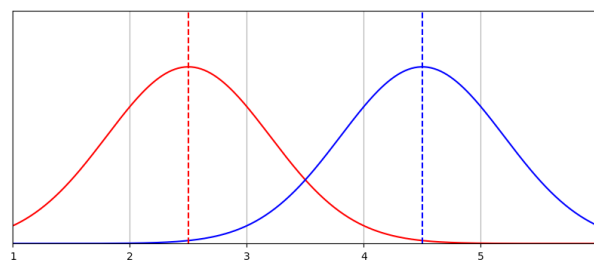


図 5.1: 重なり部分が多い 2 分布

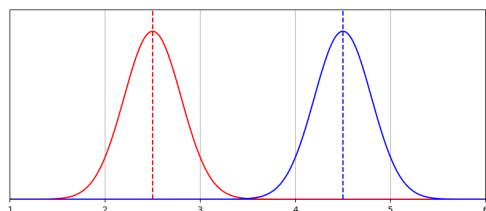


図 5.2: 分散が小さい 2 分布

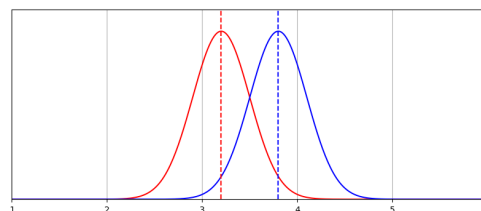


図 5.3: 分散、平均の差がともに小さい 2 分布

5.2.1 優越率

優越率とは、異なる実力帯の2プレイヤーのデータを用意し、ランダムに1データサンプルした時に正しい順序関係となる確率を表す。式で示すと、以下のような式(5.1)となる。

$$f(A, B) = (x_A > x_B \mid x_A \in A, x_B \in B) \quad (5.1)$$

※ A, B は異なる実力帯のデータ群で、 A の方が実力が高いとする。

例えば、2プレイヤーのデータ群が同じだった場合の指標は0.5、図5.4のように半分重なっている分布の場合は0.875、図5.5のように重なった部分がない分布の場合は1となる。本指標は、分布が疎であったとしても、重なり部分が少なくすることを判断する事が可能であり、1に近い方が良い補正であると評価できる。

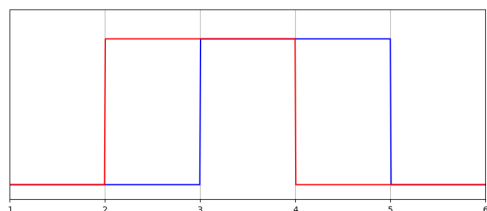


図 5.4: 半分重なった分布

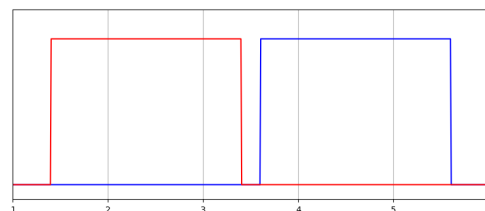


図 5.5: 分離した分布

5.2.2 分布の重なり面積

本指標は、異なる実力帯の2プレイヤーのデータについて、2分布をヒストグラムで表した時の重なっている面積の割合を表す。

例えば、2プレイヤーのデータ群が同じだった場合の指標は1、図5.4のように半分重なっている分布の場合は0.5、図5.5のように重なった部分がない分布の場合は0となる。重なり部分が少ない方が2分布間に差があることを示せるため、0に近い方が良い補正であると評価できる。

5.2.3 クラス分類

本指標は、1d,4k,8kの棋譜の平均lossのみを入力として、1d,4k,8kのいずれかを出力とした3クラス分類を行う。本研究では、手法として、線形SVM（サポートベクターマシン）を用いて、10-fold交差検証している。指標の値は平均正答率を用いる。

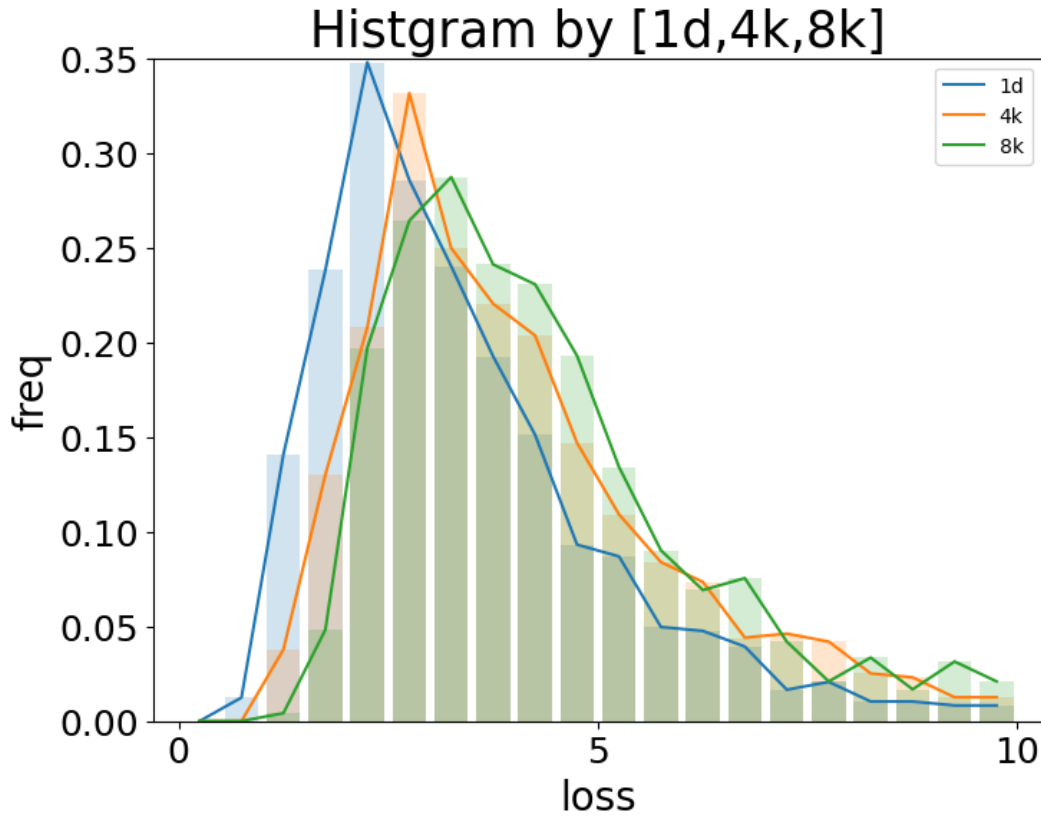


図 5.6: 1 試合の loss 平均のヒストグラム

5.3 手数を変えることによる評価指標の改善の確認

本節では、前節で示したばらつきを抑制する指標の利用例として、1 試合の loss 平均をとる区間を変化させたときに、指標がどのように変わるかを調べた予備実験の方法と結果を紹介する。

使用した棋譜データは 5.1 節に記しているもののうち、1d, 4k, 8k のものである。4 段級位離れていれば実力にはかなりの差があり、これらの loss 平均が実際に区別できるのかを確認する。40 局以上対局したプレイヤーの中から、段級位ごとに 20 人分のデータを収集した。Katago で解析した棋譜データの各局について、各プレイヤーの loss の算術平均を求める。この算術平均した値を 1 試合平均 loss とする。また、算術平均する範囲を 100 手まで、150 手まで、全手数対象としている。平均範囲を変えることによって、レーティング推定性能を向上させることができることは窪木らによって示されている [8]。本節では、それを追試し、さらに別のパラメータを試す意図、性能向上が提案した指標からも予想できることを確認する意図で、この予備実験を行う。

図 5.6 に 1d,4k,8k についての、1 試合の loss 平均のヒストグラムを示す。1d が loss が最も小さく（ミスが少ないまたは軽度で）、8k が最も大きい（ミスが多いま

たは重度) ことが分かるが, 重なる部分も多い, つまり試合ごとに出来不出来がかなりあることが分かる.

優越率を求めた結果を表 5.1 に示す. データは 1 試合平均 loss を用いている. 1d-4k の場合, 全手数を含めて平均を取ると優越率は 0.615 に過ぎないが, これを 150 手, また 100 手までに限定すれば, 優越率は 0.639 および 0.665 まで向上する. つまり, 4k の人が 1d の人よりも低い loss になるようなことが減る, 平均 loss に実力が反映されやすくなるということである. 1d-8k や 4k-8k の場合も同様である.

分布の重なり面積を求めた結果を表 5.2 に示す. データは, 1 試合平均 loss, 5 試合平均 loss, 10 試合平均 loss を用いている. ここでは, ヒストグラムのビン幅を 0.1 にしている. まず 1 試合平均の場合を見ると, 4 段級位差があってもかなりの重なりがあることが分かる. これが, 5 試合平均, 10 試合平均になると, 試合ごとの出来不出来が打ち消し合い, そのランクの人たち本来の実力に見合った平均 loss が出やすくなり, 重なりは小さくなる. また, 全手数で loss を平均した場合よりも, 150 手および 100 手までに限定したほうが, 概ね重なりは小さくなり, これは優越率と同じ傾向である.

1d, 4k, 8k の棋譜だけを使い, また, 平均 loss だけを入力とした場合の 3 値分類問題の結果を表 5.3 に示す. 入力のパターンは, 1 試合平均 loss, 5 試合平均 loss, 10 試合平均 loss を用いており, 表内の数字は平均正解率を示している. これまでの傾向と同様, 平均する試合数が多いほど概ね正答率は向上する. また, 試合展開に loss のスケールが左右されやすい後半部分を削除するほど, 正答率は高くなる. とはいえ, 10 試合を見ても 1d, 4k, 8k のどれなのかが 7 割程度の精度でしか当てられないというのはやや不満である.

表 5.1: 優越率による性能比較

	1d-4k			1d-8k			4k-8k		
	-100手	-150手	全手数	-100手	-150手	全手数	-100手	-150手	全手数
	0.665	0.639	0.615	0.773	0.719	0.670	0.626	0.585	0.556

表 5.2: 分布の重なり面積による性能比較

	1d-4k			1d-8k			4k-8k		
	-100手	-150手	全手数	-100手	-150手	全手数	-100手	-150手	全手数
1試合平均	0.737	0.764	0.774	0.576	0.660	0.688	0.780	0.825	0.799
5試合平均	0.539	0.544	0.609	0.351	0.392	0.501	0.678	0.729	0.714
10試合平均	0.450	0.490	0.520	0.250	0.330	0.450	0.490	0.640	0.650

表 5.3: 平均 loss のみで 1d,4k,8k のどれなのかを推定するクラス分類問題の正答率

平均正解率	-100手	-150手	全手数
1試合平均	45.76	40.33	37.48
5試合平均	57.33	54.46	46.73
10試合平均	68.67	58.33	46.33

第6章 期待lossを用いた補正

本章では、提案手法である期待lossを用いて補正を行う。6.1節では、実際にlossをどうやって補正するかのを提示する。6.2節では、6.1節で説明したlossの補正式を用いて、レーティング推定を試みた結果を示す。

6.1 lossの補正式

ある局面の期待lossである $E(\text{loss})$ と、最善手から実際の着手で損した目数を表すlossを用いて、lossの補正を行った。式(6.1)、式(6.2)、式(6.3)、式(6.4)の4通りの補正方法を提案する。基本的には、 $E(\text{loss})$ が高い場合つまり難しい局面ではlossを低めに補正し、 $E(\text{loss})$ が低い場合つまり簡単な局面ではlossを高めに補正することを目的としている。

$$\text{loss}_{\text{new}} = \frac{\text{loss}}{E(\text{loss}) + \alpha} \quad (6.1)$$

$$\text{loss}_{\text{new}} = \frac{\text{loss} + \alpha}{E(\text{loss}) + \alpha} \quad (6.2)$$

$$\text{loss}_{\text{new}} = \text{loss} + \gamma \cdot \frac{\text{loss} - E(\text{loss})}{E(\text{loss}) + \alpha} \quad (6.3)$$

$$\text{loss}_{\text{new}} = \text{loss} + \gamma \cdot \frac{\text{loss} - E(\text{loss})}{\text{loss} + E(\text{loss})} \quad (6.4)$$

式(6.1)は、lossと $E(\text{loss})$ の比率で表すことで、lossと期待lossの大小関係で良い手、悪い手の判断が可能になると考えた。ただし、 $E(\text{loss})$ が0に近いときに値が限りなく大きくなることを防止するために、補正項 $\alpha > 0$ を足している。しかし、lossが0のとき（プレイヤーの着手が最善手のとき）には、 $E(\text{loss})$ の値に関わらず、0となるため、「難しい局面なのに正着が打てた」場合と「簡単な局面で正着が打てた」を同等に評価してしまう。式(6.2)は、式(6.1)の則していない部分を反映するために、分子の項にも補正項 α を足したものである。

式(6.3)は、lossに対して、lossと期待lossの差に基づいた比率を反映して補正した。そのため、もともとのlossのスケールはそのままに、補正することが可能となる。しかし、パラメータ γ の値によっては、新しい指標(loss_{new})の値が負になる可能性がある。式(6.4)は、式(6.3)のパラメータ α をlossに変えることによっ

て、補正幅が $-\gamma$ から γ の間になり、扱いやすくなるが、式 (6.1) と同様に、loss が 0 のとき（プレイヤーの着手が最善手のとき）には、 $E(\text{loss})$ の値に関わらず、 $-\gamma$ となる。

6.2 実験

本節では、前節の loss 補正式がどれほどばらつきを抑制するか、5.2 節で示した指標をもとに評価した結果、また、実際にレーティング補正を行った結果を説明する。

6.2.1 loss 補正式の 3 つの指標による評価

本項では loss 補正式を用いて、5.2 節の指標を用いた評価を行った結果を説明する。使用した棋譜データは 5.1 節で説明している。loss 補正式の計算方法について、5.3 節で説明した 1 試合の loss 平均を計算する範囲が 3 通り、loss 補正式が 4 通り、（ただし α は 0.5, 1.0, 1.5, 2.0 の 4 通り、 γ は 1, 5, 10 の 3 通りのパラメータを含めた補正式が 23 通り）の全 $3 \times 23 = 69$ 通りを試した。また、期待 loss を求めるときに使う prior を 1d とした。

優越率を求めた結果を表 6.1 に示す。表の結果は 1 試合平均 loss で、すべてのパターンの中で最良のパラメータのときの値を用いている。1d-4k の場合、loss をそのまま使って 1 試合平均を取ると優越率は 0.665 に過ぎないが、最も効果が高かった期待 loss の式 (6.2) を用いると、優越率は 0.773 まで向上する。1d-8k や 4k-8k の場合も同様であり、特に 1d-8k の式 (6.2) は 0.901 という高い数値になっており、通常の loss より 0.128 も向上した。

分布の重なり面積を求めた結果を表 6.2 に示す。表の結果は 1 試合平均 loss で、すべてのパターンの中で最良のパラメータのときの値を用いている。ここでは、ヒストグラムのビン幅を 0.1 にしている。loss をそのまま使って平均した場合に比べ、最も効果が高かった期待 loss の式を用いると、1d-4k は 0.173、1d-8k は 0.223、4k-8k は 0.161 の向上がみられた。

1d, 4k, 8k の棋譜だけを使い、また、平均 loss だけを入力とした場合の 3 値分類問題の結果を表 6.3 に示す。入力のパターンは、1 試合平均 loss、5 試合平均 loss、10 試合平均 loss を用いており、表内の数字は平均正解率を示している。また、表の結果はすべてのパターンの中で最良のパラメータのときの値を用いている。loss をそのまま使って平均した場合に比べ、最も効果が高かった期待 loss の式を用いると、1 試合平均 loss を入力時は 10.28 ポイント、5 試合平均 loss を入力時は 13.96 ポイント、10 試合平均 loss を入力時は 9.33 ポイントの向上が見られた。

以上のように、どの指標を用いても、loss 補正が補正なしよりも効果的であることが分かった。一方で、5.3 節で述べたように補正なしの場合では「100 手まで」

が最も良い結果であったのに比べ、表 6.1～表 6.3 では必ずしもそうではないことが分かった。表 6.3 の比較で分かるように、試合数が小さい場合は手数を限定してしまうと参考にできる着手が小さくなりすぎるためと考える。表 6.1, 6.2 の結果はどちらも 1 試合平均であるため、全手数のほうがよいこともあったのだろう。

また、式 (6.1)-(6.4) の比較では、式 (6.4) は一段劣るものの、他の補正は一長一短があることが分かった。loss に乗るノイズの原因や影響はさまざまであり、まだこれらよりもよい補正式がありうることを示唆している。

表 6.1: 優越率による補正式の効果の比較 (1 試合平均)

	通常	式(6.1)	式(6.2)	式(6.3)	式(6.4)
1d-4k	0.665 (-100手)	0.772 (全手数, $\alpha=0.5$)	0.773 (全手数, $\alpha=0.5$)	0.757 (-150手, $\alpha=0.5,\gamma=10$)	0.688 (-100手, $\gamma=5$)
1d-8k	0.773 (-100手)	0.900 (全手数, $\alpha=0.5$)	0.901 (全手数, $\alpha=0.5$)	0.881 (-150手, $\alpha=0.5,\gamma=10$)	0.819 (-100手, $\gamma=10$)
4k-8k	0.626 (-100手)	0.701 (-100手, $\alpha=1$)	0.701 (-100手, $\alpha=1.5$)	0.695 (-100手, $\alpha=1,\gamma=10$)	0.661 (-100手, $\gamma=10$)

表 6.2: 分布の重なり面積による補正式の効果の比較 (1 試合平均)

	通常	式(6.1)	式(6.2)	式(6.3)	式(6.4)
1d-4k	0.737 (-100手)	0.573 (全手数, $\alpha=0.5$)	0.572 (全手数, $\alpha=0.5$)	0.564 (全手数, $\alpha=0.5,\gamma=10$)	0.680 (-100手, $\gamma=10$)
1d-8k	0.576 (-100手)	0.357 (全手数, $\alpha=0.5$)	0.353 (全手数, $\alpha=1$)	0.360 (-150手, $\alpha=0.5,\gamma=10$)	0.495 (-150手, $\gamma=10$)
4k-8k	0.780 (-100手)	0.689 (-100手, $\alpha=0.5$)	0.675 (全手数, $\alpha=0.5$)	0.619 (-100手, $\alpha=0.5,\gamma=10$)	0.707 (-100手, $\gamma=10$)

表 6.3: クラス分類による補正式の効果の比較

	通常	式(6.1)	式(6.2)	式(6.3)	式(6.4)
1試合平均	45.76% (-100手)	55.87% (全手数, $\alpha=0.5$)	56.04% (全手数, $\alpha=1$)	53.87% (-150手, $\alpha=0.5,\gamma=10$)	49.86% (-100手, $\gamma=10$)
5試合平均	57.33% (-100手)	70.95% (-150手, $\alpha=0.5$)	71.29% (-150手, $\alpha=0.5$)	69.61% (-100手, $\alpha=0.5,\gamma=10$)	62.35% (-150手, $\gamma=10$)
10試合平均	68.67% (-100手)	78.00% (-100手, $\alpha=0.5$)	76.67% (-150手, $\alpha=0.5$)	76.67% (-100手, $\alpha=1.5,\gamma=5$)	72.67% (-100手, $\gamma=5$)

6.2.2 loss 補正式を用いた少数棋譜でのレーティング推定

本項では、loss をそのまま使う手法（通常手法）、窪木らが提案した手法による補正をした loss を用いる手法（先行研究）と、本研究で提案する、期待 loss を用いて loss を補正した値を用いる手法（提案手法）の 3 つにおいて、少数棋譜でレーティング推定をした結果を示す。

通常手法と先行研究の手法については、5.3 節で説明した 1 試合の loss 平均を計算する範囲を先行研究と揃える目的で、150 手まで、全手数対象の 2 通りにした。さらに、提案手法については、平均を計算する範囲の 2 通りに、loss 補正式が 6.1 にて説明した 4 通り（ただし α は 0.5, 1.0, 1.5, 2.0 の 4 通り、 γ は 1, 5, 10 の 3 通りのパラメータを含めた補正式が 23 通り）、期待 loss を求めるときに使う prior が 10k, 8k, ..., 7d, 9d の 10 通りの全 $2 \times 23 \times 10 = 460$ 通りを試した。推定手法については、棋譜の loss とそのプレイヤーの棋力帯をペアとして教師あり学習を行い、モデルの学習手法として LightGBM を用いている。

学習時に使用した棋譜データは、10k, 8k, ..., 7d, 9d までの 300 局 $\times$ 10 段級位の計 3000 局を使用した。学習データは以下の手順で作成した。

- (1) 段級位ごとにランダムに 10 局取り出し、その平均値をとる。
- (2) 1 を段級位ごとに 500 回繰り返し、
500 個 $\times$ 10 段級位の 5000 個の学習データを作成する。

まず、レーティング推定精度が向上しているかを確かめるために学習データに対して棋力帯ごとに 3-fold 交差検証した結果を表 6.4 に示す。3 回の RMSE の平均値で評価しており、表の結果はすべてのパターンの中で最良のパラメータのときの値を用いている。従来手法の RMSE が 2.832[段・級]、先行研究の手法が 2.342[段・級] だったのに対して、最も効果が高かった期待 loss の式 (6.2) ($\alpha = 1$, prior=1d) は 1.730[段・級] と 1.102[段・級] も誤差を抑えることができた。

表 6.4: 3-fold 交差検証の結果

	通常	窪木らの手法	式(6.1)	式(6.2)	式(6.3)	式(6.4)
手数	-150手	-150手	-150手	-150手	-150手	-150手
パラメタ			$\alpha=0.5$	$\alpha=1$	$\alpha=0.5, \gamma=10$	$\gamma=10$
prior			1d	1d	3d	1d
RMSE	2.832	2.342	1.740	1.730	1.772	2.105

次に、通常手法、先行研究の手法について、最も効果が高かった平均 loss の計算範囲を 150 手までに限定する方法を、提案手法について、最も効果が高かった期待 loss の式 (6.2) ($\alpha = 1$, prior=1d) に対して、テストデータを用いて少数棋譜での棋力推定を行う。テスト時に使用した棋譜データは、10k,8k,...,7d,9d の 10 段級位 20 人に 1 人 50 局の計 10000 局を使用した。評価方法は以下の手順で行う。

- (1) プレイヤーごとにランダムで 10 局取り出したデータに対して、各指標ごとに 1 試合平均 loss を求め、10 局の平均を求めたものをモデルの入力に与えて、棋力推定を行う。
- (2) 1 を 100 試行行い、100 試行分の推定誤差を RMSE で算出する。
- (3) 2 の結果を棋力帯ごとに 20 人分の平均を計算する

少数棋譜でのレーティング推定を行った結果を表 6.5 に示す。従来手法の全体が 3.975[段・級]、先行研究の手法が 1.935[段・級] だったのに対して、提案手法は 1.511[段・級] と良い結果が得られた。特に、10k や 8k などの低段級位の人や、9d や 7d の高段級位の人に対しては、より正確な結果を得ることができた。

表 6.5: テストデータでの性能評価の結果

棋力帯	通常	窪木らの手法	提案手法 式(6.2)
10k	6.903	2.315	1.069
8k	5.094	1.605	1.267
6k	3.092	2.134	2.495
4k	2.139	2.215	2.534
2k	2.454	3.024	2.365
1d	3.526	2.712	1.827
3d	4.637	1.894	1.253
5d	5.192	1.808	1.091
7d	4.865	1.346	1.048
9d	1.937	0.289	0.130
全体	3.975	1.935	1.511

第7章 おわりに

人間プレイヤーの強さはしばしばレーティングという指標で表される。レーティングは通常、多くの試合を重ねて求められるものであるが、指導などの面で、もっと少ない試合数でおおよそそのレーティングを推定したい場合もある。この目的のために、“評価対象のプレイヤーの着手が最善手とくらべてどのくらいの損をしているのか”である loss という値を用いてレーティングを推定する手法が研究されている。

しかし、間違えやすく“難しい”局面が多い対局の loss と、平易な局面が多い対局の loss を同等に扱うことは、誤推定の原因となる。本研究では、loss をそのまま用いるのではなくて、局面の難しさに応じて補正を行うことで、運不運をできるだけ排して精度の高いレーティング推定を行うというアプローチをとる。

まず、局面の“難しさ”にはどういったものがあるかを検討したうえで、その中で補正目的に最も適していそうな定義を選択した。具体的には、ある局面における各着手の loss と、各着手をプレイヤーが選ぶ“着手確率”の内積を取り、これを“期待 loss”と呼んで局面の難しさとして扱うことにした。

次に、レーティング推定で loss を用いるときに、どういった分布になることが望ましいかを考察した。loss のばらつきが一番大きな要素であり、分布間が離れていることが望ましいと考え、それを満たすかを判断する指標である、優越率、分布同士の重なり面積割合、3クラス分類での評価を行うことを提案した。指標が適切なものかを判断するために、先行研究 [8] でその効果が示されている、「一局全体から loss 平均値を計算するのではなく、150 手までに限定して計算する」というノイズ抑制のアイデアについて、追試を行った。この結果、3つの指標が loss のばらつきを判断する指標として正しいことを確認できた。

最後に、期待 loss を用いて loss の補正を行う提案手法の評価を行った。補正法としては4つの式を提案した。優越率などの3つの指標の比較だけでなく、レーティング推定の精度の比較も行った。補正式4つは、全ての指標において補正しない場合よりも優れた性能を示したが、その中でも、実際の loss と期待 loss の比率を重視する補正式が最も良い結果を示した。

今後の課題として、パラメータのチューニングを詳細に行うこと、他の補正式の提案、複数手法を混ぜた結果の提示などが挙げられる。今回パラメータを α, γ の2つに対して、4通りと3通りで推定した。しかし、さらに小さい値や大きい値なども候補に挙げられる。実際、式 (6.1) や式 (6.2) で用いた α は、調べた範囲で最小の $\alpha = 0.5$ が、式 (6.3) や式 (6.4) で用いた γ は、調べた範囲で最大の $\gamma = 10$

が良い傾向になっているため、下限や上限を調べる必要がある。補正式については、今回は大きく2つに分かれているが、他の手法もあるはずであるため、他の手法を提案するとともに、本手法と比較したい。また、本研究では、lossに着目して、priorを用いた期待lossで補正することでレーティング推定を行った。しかし、今回用いたKatagoには、予測勝率やownershipと呼ばれるどちらの手番のプレイヤーの陣地になるかを示した指標などたくさんある。そのため、他指標を含めた推定結果などはさらに良い結果が出ると期待できる。

さらに、本手法の局面の難しさを表した期待lossの指標は、レーティング推定以外にも応用できると期待する。ゲームにおいては、プレイヤーに適したパズル問題やゲームの部分問題を作成することが可能となるだろう。さらに、ゲーム外においても、人の意思決定を推測することや教育における問題作成の助けになることを期待する。

参考文献

- [1] OpenAI, “GPT-4 technical report”, arXiv pp.2303-08774(2023)
- [2] Gemini Team Google, “Gemini: A Family of Highly Capable Multimodal Models” arXiv pp.2312-11805(2024)
- [3] D Silver, Thomas Hubert, et.al, “Mastering the game of Go without human knowledge” Nature, Vol.550, No.7676, pp.354-359(2017)
- [4] Chu-Hsuan Hsueh and Kokolo Ikeda, “Improvement of Move Naturalness for Playing Good-Quality Games with Middle-Level Players”, in Applied Intelligence, Vol.54, No.2, pp.1637-1655(2024)
doi: 10.1007/s10489-023-05210-2
- [5] Chu-Hsuan Hsueh, Takefumi Ishii, Tsuyoshi Hashimoto, and Kokolo Ikeda, “Proposal and Generation of Endgame Puzzles for an Imperfect Information Game Geister”, Entertainment Computing, Vol.52(2025)
- [6] Nam Sanggyu, Hsueh Chu-Hsuan and Ikeda Kokolo, Generation of Game Stages with Quality and Diversity by Reinforcement Learning in Turn-based RPG”, in IEEE Transactions on Games, Vol. 14, No. 3, pp. 488-501(2021)
doi: 10.1109/TG.2021.3113313.
- [7] 眞鍋和子, 三宅陽一郎, “遺伝的アルゴリズムを用いたプレイヤーエージェントによるデジタルゲームのバランス調整”, JSAI 社会における AI 研究会, Vol.2017-SIG-SAI-30, pp.1-6(2017)
- [8] 窪木響大, シュエジュウシュエン, 池田心, “ゲームの展開に過度に影響されないプレイヤ強さの推定方法”, ゲームプログラミングワークショップ 2024 論文集, pp.117-124(2024)
- [9] イロレーティング [<https://ja.wikipedia.org/wiki/イロレーティング>] (最終アクセス：2025/02/02)
- [10] BMAB[<https://bgmastersab.com/about>] (最終アクセス：2025/02/02)

- [11] 馬場匠, 伊藤毅志, “少ない棋譜からの将棋プレイや棋力推定手法の研究”, 情報処理学会論文誌, Vol.61, No.6, pp.1190-1199(2020)
- [12] 開発者の洞察:「ハースストーン:バトルグラウンド」レーティングシステムのアップデート [<https://hearthstone.blizzard.com/ja-jp/news/23523064>] (最終アクセス: 2025/02/02)
- [13] 初心者必見! 上位者と対等に真剣勝負を楽しむ方法? 「駒落ち戦」の魅力とは [https://www.shogi.or.jp/column/2017/05/post_127.html] (最終アクセス: 2025/02/02)
- [14] 第10期叡王戦(シミュレーション) [<http://kishibetsu.com/kisenR/eiou10R.html>] (最終アクセス: 2025/02/02)
- [15] 森田茂彦, 松崎公紀, “大貧民における初期手札の不均等性を考慮したレーティングアルゴリズムの提案”, 研究報告ゲーム情報学 (GI), Vol.2014, No.14, pp.1-5(2014)
- [16] 山下宏, “将棋名人のレーティングと棋譜分析”, ゲームプログラミングワークショップ 2014 論文集, pp.9-16(2014)
- [17] 相良駿昇, シュエジュウシュエン, 池田心, “上海ゲームの工夫が報われる問題生成のための人間の記憶選択と忘却のモデル”, ゲームプログラミングワークショップ 2023 論文集, pp.48-55(2023)
- [18] Matej Guid and Ivan Bratko, “Computer analysis of world chess champions”, ICGA journal, Vol. 29, No.2, pp.65–73(2006)
- [19] 中野龍一, 飛田克哉, 亀谷由隆, “難易度を考慮した遺伝的アルゴリズムによるナンバークロスワードパズルの生成とその評価”, 第 83 回全国大会講演論文集, Vo.2021, No.1, pp.737-738(2021)
- [20] Reid McIlroy-Young, Siddhartha Sen, Jon Kleinberg and Ashton Anderson, “Aligning superhuman ai with human behavior: Chess as a model system”, Proceedings of the 26th ACM SIGKDD, pp.1677–1687(2020)