

コンピュータ将棋における平均損失を用いた棋力調整手法の提案

林俊佐[†] 伊藤毅志[†]

概要: 本研究では、棋力と高い相関のある平均損失を人為的に模倣することによって将棋 AI の棋力を調整する手法を提案する。先行研究から、将棋 AI によって形勢を表す評価値のプレイヤーの指し手によってどれだけ下がったかを示す損失の平均とそのプレイヤーの棋力との間に強い相関があることが示されている。本研究では、その性質を利用し、人為的に平均損失を模倣することにより将棋 AI の棋力の調整を試みた。このシステムを人間プレイヤーに使用させて、従来の棋力調整方法のプログラムと比較する主観評価をさせたところ、提案手法では従来のものに比べて棋力の高いプレイヤーで不自然さをあまり感じないという好結果を得ることができた。

キーワード: コンピュータ将棋、平均損失、棋力調整

Strength Adjustment Method Using Mean Loss in Computer Shogi

JUIN-ZUO LIN[†] TAKESHI ITO[†]

Abstract: In this study, we propose a method of artificially imitating mean loss that has a high correlation with players' strength to adjust strength of computer shogi. Previous studies have shown that there is a strong correlation between the player's strength and the mean loss, which indicates how much the evaluation value of computer shogi has decreased due to player's moves. We tried to adjust the strength of computer shogi by artificially imitating the mean loss using this nature. Humans played this system and the conventional system and performed subjective evaluations. The proposed method showed that the expert players felt less unnaturalness than the conventional system.

Keywords: Computer Shogi, Mean Loss, Strength Adjustment

1. はじめに

1.1 研究の背景

従来、コンピュータチェス、将棋、囲碁などのゲーム人工知能 (AI) の研究者はプロ棋士を越える強さの実現を求めてきた。チェスにおいて 1997 年 IBM が開発した Deep Blue が当時の世界チャンピオンであるカスパロフ氏に勝ち、将棋において 2017 年に Ponanza が当時の名人である佐藤天彦氏に勝ち、最も複雑と言われていた囲碁においても 2017 年に AlphaGo が互先で当時世界ランク 1 位の柯潔氏に勝利した。その後、AlphaGo の汎用的なバージョン AlphaZero が発表され、コンピュータチェス、将棋、囲碁のそれぞれのトッププログラムを破った。このようにコンピュータチェス、将棋、囲碁などの二人零和有限確定完全情報ゲームにおいて、すでに十分に強いコンピュータプログラムが作られるようになってきている。

近年では、十分に強くなったコンピュータプログラムをどのように利用するかという研究が盛んに行われている。近年の強すぎる AI は、アマチュアから見ると全く勝てない強さであり、対局相手として適さない。そこで、コンピュータの強さを調整することで、より良い対戦相手とする試みが行われている。対戦相手としての AI は、自分より弱く調整してしまうと勝ちすぎ、自分より強く調整してしまうとあまり勝てなくなってしまうため、対戦の楽しさを損ねてしまう。自分から見ると勝敗が五分になる AI が対局相手とし

て好敵手となりうる。しかし、たとえ自分と同じレベルの強さの AI だとしても、対戦中に不自然な弱さを感じてしまうと、対局して楽しいものでなくなってしまう。不自然さを感じないような棋力調整方法が求められている。

1.2 本研究の目的

本研究では、馬場の研究の結果を利用して[1]、将棋の棋力と強い相関のある「平均損失」に着目し、調整目標である棋力に合った平均損失に近い損失を与える手を指すようにすることで棋力を調整する手法について考える。具体的には、馬場の手法を用いて棋力を調整したいプレイヤーの客観的な棋力を測定し、その棋力に相当する同程度の棋力を持つ AI の構築を目指す。客観的な棋力とは、対戦型競技によく使われているイロレーティング (Elo Rating) を用いて表すものとする。二人のプレイヤーが同程度の棋力とは、二人が対局すると勝率が 50%±10% くらいであり、イロレーティングで表すと二人のレーティング差が±70 以内であるとする。その実現方法として平均損失に着目し、同程度の棋力とは、同じぐらいの平均損失を伴うプレイをすることと考え、棋力に応じた平均損失に近い損失を与える手を選択するシステムの実現を目指し、評価実験によりその有効性を示す。

2. 関連研究

2.1 棋力調整手法

2.1.1 評価値を用いた動的な棋力調整手法

評価値とは、コンピュータ将棋においてよく使われてい

る局面の形勢を表した数値である。評価値の詳しい定義は将棋 AI によって異なるが、一般に評価値が 0 の局面は先後互角な局面であり、高い評価値を得たプレイヤーほど優勢であることを示している。なお、ある候補手の評価値とは、その候補手を指した後の局面の評価値を指す。

仲道らの研究では、将棋 AI の評価値を用いて対局相手に応じて動的に将棋 AI の棋力を調整する手法を提案した。そして、この手法を実装したシステムを接待将棋と名付けた[2]。

仲道らは各局面において各候補手 M の評価値を式 1 の計算式に変換することによって、常に評価値が 0 に最も近い局面となる候補手を選ぶ AI を実現した。

$$V_0(M) = \begin{cases} -V(M), & V(M) \geq 0 \\ V(M), & V(M) < 0 \end{cases} \quad \text{式 1}$$

(ただし、 $V(M)$ は候補手 M の評価値を表す。)

また、接待将棋を評価するための実験も行われた。この棋力調整手法を導入した AI とそうでない AI による着手を対局相手である人間に主観評価させたところ、棋力の低いプレイヤーはあまり不自然さを感じないという結果が得られた。一方、棋力の高いプレイヤーほど指し手に不自然さを感じる傾向が見られた。その理由として、強いプレイヤーほどミスに気づきやすく、自分のミスに気づいたときにそのミスに対してミスで返す指し手に違和感を覚えることが原因ではないかと指摘されている。

接待将棋は、どんな相手に対しても良い勝負になる対局を指せるという利点もあるが、好手に対しては好手で応じ、悪手に対しては悪手で応じるため、AI の棋力が一定していない。そのため、せっかく良い手を積み重ねてもなかなか勝てなかったり、逆に悪手に対して付け込んでくる指し手を選べなかったりする。自身の指し手の良し悪しが勝敗に反映しにくいいため、何が良かったのか悪かったのかを認識することも難しいという問題点もある。

2.1.2 人間のレベルに適應する多様な AI の生成手法

上田らの研究では、オセロ AI において遺伝的アルゴリズムにより相手の棋力に適應する多様な AI の生成手法を提案した[3]。

上田らは、プレイヤーを AI と対局させ、テスター AI を用いてその棋譜を評価し、AI の最善手と実際の着手の評価値の差の平均 μ を計算した。次に、着手の評価値の差の平均が μ に近づくようなエージェント AI を用意した。さらに、特徴量パラメータの違う AI を複数用意し、エージェント AI と対局させ、勝率を算出した。勝率と特徴量パラメータで適應関数を定義し、遺伝的アルゴリズムにより AI の特徴量パラメータを最適化した。この手法で生成した複数の AI を人間に評価させたところ、多様性に関して、中級者には AI の違いが区別可能であるが、初級者には区別しづらかった。これは用いた特徴量パラメータが高度なものであったためと考えられる。

この生成手法で生成した AI には、着手の不自然さをまったく感じられないという評価が多く、勝率も 5 分になる結果が大半となった。しかし、この生成手法は遺伝的アルゴリズムを用いるため、AI 同士に大量の対局をさせる必要がある。上田らの実験を実現するには、1 サイクルあたり 2400 試合が必要であり、15 サイクルには 36000 試合が必要である。本研究で用いる 10 分切れ負け将棋でこの 15 サイクル分を実現しようとする、一局平均 15 分と換算しても 1 年以上かかる計算となり、あまり実用的な手法であるとは言えない。

2.2 棋力推定手法

提案手法のベースとなる棋力推定手法に関する先行研究を紹介する。

馬場らは、平均損失という指標を用いて、少ない棋譜から棋力を推定する手法を提案した[1]。平均損失とは、指した後と指す前の AI による評価値の差の平均であり、この指標が実際のレーティングと非常に高い相関があることを示した。そして、特に互角の局面においてこの性質が強く現れることを利用して、接待将棋と対戦させることで、たった 4 局でかなり正確な棋力を推定できることを示した。本研究では、評価実験の参加者の客観的な棋力を推定するために、この手法を利用する。

3. 棋力調整手法の設計方針

平均損失が棋力との間に強い相関があるという馬場らの先行研究の結果を踏まえ、特定の棋力のプレイヤーを模倣する手段として、そのプレイヤーの棋力の平均損失に近づける指し手を実現することで棋力の調整を試みる。馬場の研究では、序盤と終盤の局面を排除して棋力推定を行っていた。これは、序盤は定跡局面が多く、指し手と棋力との相関が低いこと、終盤は詰むかどうかという局面が多く評価値が極端になりやすく、これも棋力との相関が低くなることを考慮したものである。

したがって、方針としては序中盤では「平均損失の模倣」を行い、終盤では「即詰み局面のミスの模倣」を行うことで、棋力の模倣を実現することにした。

なお、本研究で用いる将棋 AI は技巧 ver2.0.2[4]とする。この AI を用いたのは、十分にアマチュアトップクラスを超える棋力を持つ AI であることと、馬場らの先行研究で用いられた AI であるためである。

4. 予備実験

模倣しようとするレーティングに応じ、そのレーティングの平均損失やミスを再現するためには、レーティングと平均損失との関係式及びレーティングとミス率との関係式を調べる必要がある。この予備実験では、これらの具体的な値を求めるための実験を行う。

4.1 レーティングと平均損失の関係

先行研究で定義された平均損失は、序盤を除いた互角な局面のみの損失の平均なので、序盤や互角ではない局面などの局面に直接に利用することは適切ではない。そのため、局面の状況によって局面を序盤、互角、互角ではない局面などのいくつかの状況に分け、各棋力層のそれぞれの状況における平均損失を集計し、レーティングと平均損失との関係式を状況ごとに回帰分析を行って算出する。

4.1.1 定義

本研究では、局面を下のように 5 つの状況に分けることにする。

- 定跡：初手から評価値が 0 で続く局面。
- 序盤：42 手目以前の局面。
- 互角：42 手目以降、かつ評価値の絶対値が 200 以下の局面。
- 偏り：42 手目以降、かつ評価値の絶対値が 200 以上、かつ即詰み手順が見つからない局面。
- 即詰み：どちらかの即詰み手順が見つかった局面。

また、指し手による損失を以下の式 2 のように定義する。

$$\text{指し手 } P_n \text{ による損失} = V(P_n) - V(P_{n-1}) \quad \text{式 2}$$

(ただし、 P_n は n 手目指した後の局面、 $V(P_n)$ は分析 AI による n 手目指した後の局面の評価値)

4.1.2 分析対象とする棋譜

将棋クエストの 10 分切れ負け将棋の棋譜を分析対象とする。ただし、下に挙げる条件に該当する棋譜は除外する。

- 対局者のどちらかが将棋クエスト内の AI である。
- 分析 AI が入玉宣言すべきであると分析した局面がある。

これにより、分析対象となった棋譜は 54959 局となった。

4.1.3 分析 AI 及び分析条件

本研究でベースとなる技巧 ver.2.0.2 を用い、以下の条件で棋譜を分析する。

- 探索深さを 16 と設定する。
- 定跡ファイルは技巧 ver.2.0.2 に付属のものを使う。

4.1.4 実験手順

下の手順で、各状況におけるレーティングと平均損失との回帰直線を引く。

- (1) 分析 AI を用い、対象とする棋譜を分析する。
- (2) 分析した棋譜を、対局者のレーティングにより 50 ごとに区切ったレーティング層に分ける。
- (3) レーティングにより分けた棋譜から序盤/互角/偏りの局面を抽出して、指し手による損失を集計し、平均損失を算出する。
- (4) 抽出した局面の数を重みとして、重み付き最小二乗法により、算出された平均損失とレーティングとの回帰直線を引く。

4.1.5 結果

得られた回帰直線の方程式は式 3 から式 5 となる。

$$Los_{\frac{5}{2}}^0(R) = 0.104534 \times R - 268.995 \quad \text{式 3}$$

$$Los_{\frac{5}{4}}^0(R) = 0.234493 \times R - 620.877 \quad \text{式 4}$$

$$Los_{\frac{5}{4}}^0(R) = 0.290031 \times R - 925.708 \quad \text{式 5}$$

ただし、 R はレーティングである。

序盤の局面における各レーティング層の局面の数、平均損失及び回帰直線は、図 1 に示される。

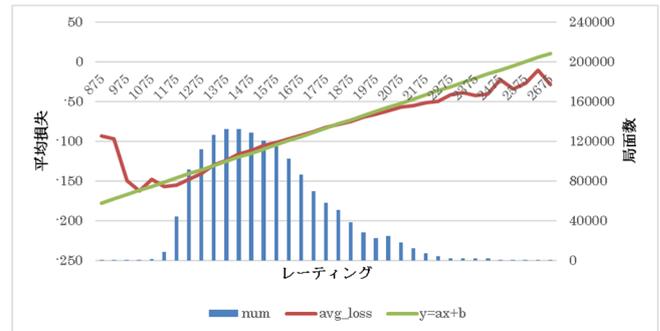


図 1 序盤における局面の数、平均損失及び回帰直線

互角の局面における各レーティング層の局面の数、平均損失及び回帰直線は、図 2 に示される。

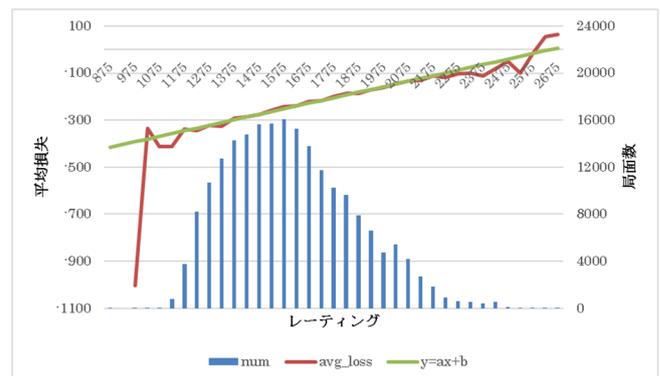


図 2 互角における局面の数、平均損失及び回帰直線

偏りの局面における各レーティング層の局面の数、平均損失及び回帰直線は、図 3 に示される。

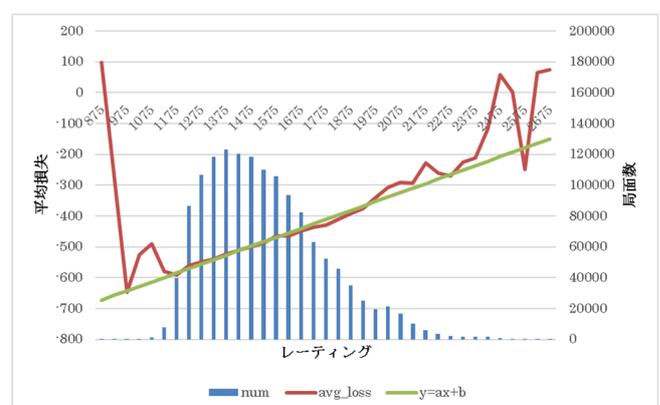


図 3 偏りにおける局面の数、平均損失及び回帰直線

4.1.6 考察

結果いずれでも、概ね正の相関が見える。特に抽出した局面数が多いレーティング層では、レーティングが上昇するにつれて安定して平均損失も増加している傾向が見られる。

図 1 と図 2 の平均損失の値に注目すると、対数関係で近似できるように見える。また、重み付き最小二乗法により回帰分析で対数関数 $a \times \log b \times R + c + d$ により近似すると、結果は図 4 に示されるように、誤差が大幅に（26 百万から 5 百万に）減少した。したがって、レーティングと平均損失の関係は、対数曲線の方がより良い近似ができることが示唆された。

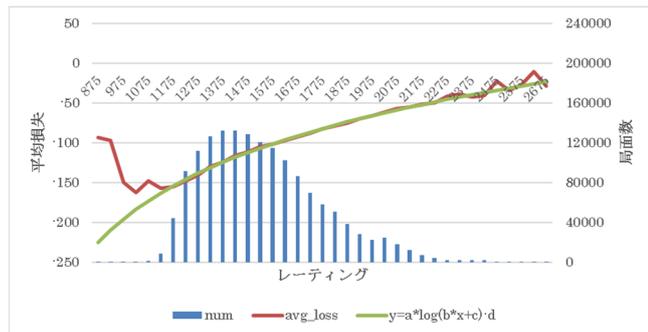


図 4 序盤における局面の数、平均損失及び近似対数曲線

4.2 レーティングと即読み局面におけるミス率

一般の将棋 AI は、即読み手順が見つかると、必ず最短で相手を詰ます手を選ぶ。しかし、人間の場合、棋力によって即読み手順を見つけられる能力には差があることが予想される。そのため、将棋 AI の棋力を調整するには、人間プレイヤーらしく気力に応じて適度に詰みを見逃す必要があると考える。

そのため、各レーティング層の、即読み局面において最短詰まさないかたり詰みを見逃したりする現象（以降、「即読み局面におけるミス」と呼ぶ）が起こる回数を集計し、レーティングと起こる確率との関係を回帰分析により導出する。

4.2.1 定義

本研究では、即読み局面におけるミス率を下のように定義する。

即読み局面におけるミス率 = (即読み局面でその最短手順を選ばなかった数) ÷ (即読み局面の総数)

また、即読み局面におけるミスとなる指し手を下のよう

- 1) ミス 1: 指した後、まだ評価値が正の値である指し手。
- 2) ミス 2: 指した後、評価値が負の値となるが相手の即読み局面とならない指し手。
- 3) ミス 3: 指した後、相手の即読み局面となる指し手。

4.2.2 分析対象及び条件

分析対象とする棋譜、分析 AI 及び分析条件は、4.1 と同じとする。

4.2.3 実験手順

下の手順で、各ミスのレーティングと即読み局面におけるミス率との関係式を導出する。

- (1) 分析 AI を用い、分析対象とする棋譜を分析する。
- (2) 分析した棋譜を、対局者のレーティングにより 50 ごとに区切ったレーティング層に分ける。
- (3) レーティングにより分けた棋譜から即読み局面を抽出し、さらに最短の即読み手順の長さに分ける。
- (4) 最短の即読み手順の長さにより分けた即読み局面から、ミス 1 / ミス 2 / ミス 3 の起こる回数を集計し、即読み局面におけるミス率を算出する。
- (5) 抽出した局面の数を重みとして、重み付き最小二乗法により、最短の即読み手順の長さとして即読み局面におけるミス率との回帰直線を引く。
- (6) 抽出した局面の数を重みとして、重み付き最小二乗法により、前の手順で得られた回帰直線群の係数とレーティングとの回帰直線を引く。

4.2.4 結果

得られた回帰直線の方程式は式 6 から式 14 となる。

$$M_1 s_1^{50}(N, R) = a^{50}(R) \times N + b^{50}(R) \quad \text{式 6}$$

$$M_2 s_2^{50}(N, R) = c^{50}(R) \times N + d^{50}(R) \quad \text{式 7}$$

$$M_3 s_3^{50}(N, R) = e^{50}(R) \times N + f^{50}(R) \quad \text{式 8}$$

ただし、 N は即読み手順の長さ、 R はレーティング、

$$a^{50}(R) = -1.5 \times 10^{-6} \times R + 0.016937 \quad \text{式 9}$$

$$b^{50}(R) = -0.00022 \times R + 0.676102 \quad \text{式 10}$$

$$c^{50}(R) = 8.44 \times 10^{-7} \times R - 0.0002 \quad \text{式 11}$$

$$d^{50}(R) = -2.5 \times 10^{-5} \times R + 0.049648 \quad \text{式 12}$$

$$e^{50}(R) = -4.2 \times 10^{-7} \times R + 0.001454 \quad \text{式 13}$$

$$f^{50}(R) = -1.2 \times 10^{-5} \times R + 0.023294 \quad \text{式 14}$$

5. 提案手法

提案手法は以下のエラー! 参照元が見つかりません。に示される通り、AI の n 手目の手番の思考において、現局面が即読み局面であるかどうかにより、方針が「即読み局面のミスの模倣」と「平均損失の模倣」の 2 つの部分に分けられる。

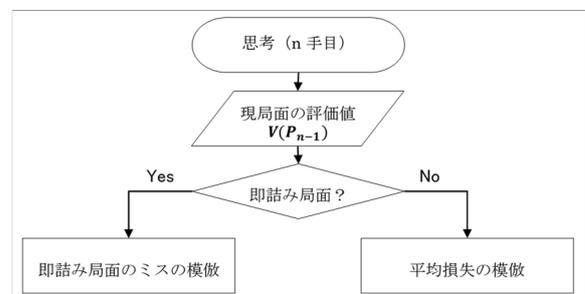


図 5 提案手法の全体のフローチャート

平均損失の模倣を実現するために、評価値変換関数 (5.1 にて後述)、目標値 T の設定 (5.2 にて後述)、誤差保存変数 (5.3 にて後述) の処理を提案する。

即詰み局面のミスの模倣を実現するために、評価値変換関数を片方のみ(後述)とし、目標値 T の設定に確率変数(後述)を導入する。

5.1 評価値変換関数

仲道らの接待将棋は、探索の際に評価値を式 1 の計算式で変換することにより、候補手の元の評価値が0に近いほど、変換した新たな評価値が高くなり、ゲーム AI が常に評価値が最も0に近い一手を指すように実現した。仲道らの式 1 下記の式 15 と同値関係である。

$$V_0(M) = -|V(M)| \quad \text{式 15}$$

式 15 の概念を拡張し、式右边に目標値 T をたし、下記の式 16 になる。

$$V_T(M) = -|V(M) - T| + T \quad \text{式 16}$$

式 16 の目標値 T が0ならば、式 16 は仲道らの提案手法の式 1 となる。したがって、式 16 の変換式により、元の評価値が目標値 T に近いほど、変換した新たな評価値が大きくなる。このような評価値を変換するための計算式を評価値変換関数と呼ぶ。

探索の際に評価値を評価値変換関数で変換するにより、ゲーム AI が評価値の最も目標値 T に近い一手を指すようになることが期待できる。

5.2 目標値 T の設定

評価値変換関数の目標値 T を常に0とすれば、仲道らの接待将棋となる。本研究では、同じ平均損失を持つプレイヤーが同じ棋力を持つように見えると仮定し、同じ平均損失となるような手を指すことで、AI が特定棋力を持つように見えるようにする。そのため、評価値変換関数の目標値 T を基本的に損失が模倣目標の棋力の平均損失に近づくように設定する。 n 手目の目標値 $T(n)$ を数式で表すと、下の式 17 となる。

$$T(n) = V(P_{n-1}) + \text{Los}(\xi R) \quad \text{式 17}$$

ただし、 P_{n-1} は $n-1$ 手目の手を指した後の局面、 $V(P_{n-1})$ は $n-1$ 手目の手を指した後の局面の評価値、 $\text{Los}(\xi R)$ はレーティング R のプレイヤーの平均損失であり、具体的な計算式は 4.1 で得られた回帰直線式 3~式 5 を用いる。

5.3 誤差保存変数

5.1 と 5.2 の提案を組み合わせれば、AI が一手ごとに目標とした平均損失と最も近い損失を与える候補手を指すように実現できると期待できる。しかし、将棋において平均合法手は 80 手程度であると言われ、本研究で用いられる技巧 ver.2.0.2 では評価値の値域は $-32000 \sim 32000$ であり、候補手の数は評価値と比べると明らかに離散的である。そのため、一手ごとに選んだ候補手の与える損失は目標とした平均損失と最も近いとは言え、全ての指し手の平均損失は目標とした平均損失との間に誤差が生じる。そこで、生じた誤差を保存しておいて、次の手番で目標値 T を調整すれば、この誤差を減少させることができるのではないかと考える。誤差保存変数を式 17 の右边に追加すると、式 18 を得る。

$$T(n) = V(P_{n-1}) + \text{Los}(\xi R) + \text{Diff}(n-2) \quad \text{式 18}$$

ただし、 $\text{Diff}(n-2)$ は $n-2$ 手目の評価値と $n-2$ 手目の目標値との差である。数式で表すと、式 19 となる。

$$\text{Diff}(n-2) = T(n-2) - V(P_{n-2}) \quad \text{式 19}$$

5.4 即詰み局面のミスの模倣

本研究で用いられる技巧 ver.2.0.2 では、即詰み局面においても評価値により指し手を決める。即詰み局面の評価値は見つけた即詰み手順の長さで決める。詰み局面、すなわち合法手のない局面の評価値は最大値であり、即詰み手順の長さが1である局面の評価値は最大値 -1 である。そのため、前述の処理は即詰み局面においても、評価値の視点で処理するという点は適用できる。

しかし、5.1 で述べた評価値変換関数を使うと、選ばれる評価値の最も目標値 T に近い手は評価値が上がる場合もあり、即詰み局面におけるミスという表現とは異なってくる。ここで、即詰み局面における評価値変換関数を片方のみ、すなわち、「目標値 T 以下は変わらず目標値 T 以上は最小値に変換する」とする。数式で表すと、式 20 となる。

$$V_T(M) = \begin{cases} V(M), & V(M) \leq T \\ \text{min}(), & V(M) > T \end{cases} \quad \text{式 20}$$

ただし、 $\text{min}()$ は評価値の最小値の定数関数であり、本研究で用いられる技巧 ver.2.0.2 では -32000 である。

そして、5.2 にて述べた目標値 T の設定方法を使うと、常に損失を与えるため、即詰み手順を常に見逃すようになってしまう。そのため、確率変数を導入し、レーティングに合わせたミス率で目標値 T を変える。数式で表すと、式 21 となる。

$$T(n, p) = \begin{cases} V(P_{n-1}), & p \leq \text{Ms}_1(R) \\ 0, & \text{Ms}_1(R) < p \leq \text{Ms}_2(R) \\ -\text{ChkVal}(), & \text{Ms}_2(R) < p \leq \text{Ms}_3(R) \\ \text{MAX}(), & \text{Ms}_3(R) < p \end{cases} \quad \text{式 21}$$

ただし、 $\text{ChkVal}()$ は即詰み局面であるかどうかの閾値の定数関数で ver.2.0.2 では 30000 、 $\text{MAX}()$ は評価値の最大値の定数関数で ver.2.0.2 では 32000 、 $\text{Ms}_k(R)$ はレーティング R のプレイヤーの k 種目までのミス率の累積であり、具体的な計算式は事前に(4.2に述べた)得られた回帰直線式 6~式 8 をする。

6. 評価実験

提案手法で棋力を調整した AI の人間らしさと自然さを評価するために、実験参加者に提案手法を実装した将棋 AI (以降、提案 AI と呼ぶ) と従来の手法である接待将棋の両方と対局させ、対局後、被験者に評価アンケートを記入させることで比較させる実験を行う。

6.1 実験方法

6.1.1 実験手順

実験は、以下の手順で行う。

- (1) 実験参加者に事前アンケートを回答させる。

- (2) 実験参加者に接待将棋と 10 分切れ負けで対局させる。先後手交代で 4 局行う。
- (3) 実験参加者に接待将棋に対する評価アンケートを回答させる。
- (4) 接待将棋との棋譜を分析用将棋 AI で分析する。
- (5) 実験参加者の互角の局面における平均損失を算出する。
- (6) 算出した平均損失を式 4 の逆関数で算出する。結果を四捨五入し、レーティングを推定する。
- (7) 推定されたレーティングを提案 AI のパラメータとして設定する。
- (8) 実験参加者に提案 AI と 10 分切れ負けで対局させる。先後手交代で 4 局行う。
- (9) 実験参加者に提案 AI に対する評価アンケートを回答させる。

6.1.2 実験参加者

実験参加者の条件は、男女年齢棋力を問わず 10 分切れ負け将棋を指せるプレイヤーとし、公開募集したところ、8 名の応募があった。棋力の内訳は初心者 3 名、級位者 1 名、初段前後 1 名、有段者 3 名であった。

6.1.3 対局用将棋 AI 及び対局条件

評価実験で対局のために使用する将棋 AI は、以下の条件とする。

用いた将棋 AI：技巧 ver2.0.2

用いた定跡：技巧 ver2.0.2 に付属のもの

定跡の利用：最大 20 手まで

ハッシュメモリサイズ：50

スレッド：4

また、以下の調整手法により、接待将棋と提案 AI を用意する。

接待将棋の調整手法：2.1.1 にて述べた仲道らの調整手法

提案 AI の調整手法：5 にて述べた提案手法

6.1.4 事前アンケート

事前アンケートでは、下の質問を問う。

- あなたの棋力はどのくらいですか？
- あなたはどれくらい将棋を指していますか？
- あなたはどれくらいコンピュータ将棋と指していますか？
- これから、コンピュータ将棋の「人間らしさ」を評価していただきますが、あなたはコンピュータ将棋の「人間らしさ」についてどのようにお考えですか？どこに注目して、何を評価基準に評価されようと思いますか？

6.1.5 評価アンケート

評価アンケートでは、下の質問を問う。

- 先ほど対局したコンピュータ将棋について、「人間らしい」と思いますか？
- 先ほどの対局において、不自然と感じていましたか？

- 先ほどの対局は楽しいと感じていましたか？
- 先ほど対局したコンピュータ将棋について、どれほど強いと思いますか？
- 実際に対戦して、「人間らしさ」を評価していただきましたが、あなたはコンピュータ将棋の「人間らしさ」についてどのようにお考えですか？どこに注目して、何を評価基準に評価されましたか？
- 対戦の中で、特に「人間らしくない」と違和感を覚えた局面や状況があれば、なるべく詳しくお答えください。

6.2 参加者の棋力

実験参加者の平均損失、推定結果及び事前アンケートにて自己申告した棋力は表 2 に示される。

表 2 実験参加者の推定棋力及び自己申告した棋力

被験者	a	b	c	d
平均損失	-1033.21	-533.538	-460.95	-353.754
推定	-1758	372	682	1139
自己申告	初心者	未回答	未回答	2 級
被験者	e	f	g	h
平均損失	-258.651	-140.216	-136.568	-88.4956
推定	1545	2049	2065	2270
自己申告	2 級	二段	2200	2200

6.3 結果

接待将棋と提案 AI に対して、人間らしさ、不自然さ、楽しさ及び強さについての評価は表 1 に示される。表中において、接待将棋を「AI1」提案 AI を「AI2」に、「人間らしくない」から「人間らしい」までの回答を 1 から 4 までの数値に、「不自然と感じなかった」から「不自然と感じた」までの回答を 1 から 4 までの数値に、「楽しいと感じなかった」から「楽しいと感じた」までの回答を 1 から 4 までの数値に、「自分より弱い」から「自分より強い」までの回答を 1 から 5 までの数値に対応している。

表 1 評価アンケートの結果

被験者	人間らしさ			不自然さ			楽しさ			強さ	
	AI1	AI2	差分	AI1	AI2	差分	AI1	AI2	差分	AI1	AI2
a	2	3	1	3	2	-1	4	4	0	3	2
b	1	1	0	4	4	0	2	1	-1	5	2
c	2	4	2	3	2	-1	4	4	0	4	5
d	3	3	0	2	2	0	3	3	0	4	3
e	4	1	-3	3	4	1	4	2	-2	3	5
f	1	2	1	4	2	-2	2	3	1	5	5
g	2	3	1	3	2	-1	4	4	0	3	4
h	2	3	1	3	1	-2	3	3	0	4	5
平均	2.1	2.5	0.4	3.1	2.4	-0.8	3.3	3.0	-0.3	3.9	3.9

6.4 考察

主観評価の結果を t 検定で検証したところ、人間らしさ、楽しさ及び強さには有意差は見られなかったが、不自然さについて p 値は 0.0796 であり、すなわち不自然さについて 92% の信頼水準で有意差が見られた。棋力が高いと言える推

定レーティングの2000を超えた実験参加者は、人間らしさと不自然さの評価において、両者とも接待将棋より提案手法の方が良いという評価が得られた。

7. 自己対戦実験

提案手法の調整の精度を評価するため、提案手法を実装した将棋AIをそれぞれ異なるレーティングに設定し、対戦させて自己対戦実験を行った。

また、評価実験で即詰み局面において不自然と感じた回答が少なくなかった。「即詰み局面のミスの模倣」が棋力をきちんと反映しているかを確認するため、「即詰み局面のミスの模倣」を実装した将棋AIと実装していない将棋AIの2つのAIを比較する。

7.1 実験方法

7.1.1 実験手順

自己対戦実験は、以下の手順で行う。

- (1) 提案手法を実装した将棋AIを二つ用意する。
- (2) 将棋AIにそれぞれ異なる数値をレーティングのパラメータとして入力する。
- (3) 先後交代、10分切れ負けで100局の対局をさせる。
- (4) 勝敗数を集計し、レーティング差を計算する。
- (5) (4)で得られたレーティング差と(2)で入力したレーティングの差との誤差を計算する。

7.1.2 対局用将棋AI及び対局条件

自己対戦実験で使用する将棋AIは、下の条件とする。

将棋AI：技巧 ver2.0.2

定跡：技巧 ver.2.0.2 に付属のもの

スレッド：1

レーティング：1200/1400/1600/1800/2000/2200/2400に設定。ただし、レーティング差が1000以上離れると、表3に示されるように、100局の対局では差が見られなくなるため行わない。

表3 理論上の勝敗数

	1400	1600	1800	2000	2200	2400
1200	24-76	9-91	3-97	1-99	0-100	0-100
1400		24-76	9-91	3-97	1-99	0-100
1600			24-76	9-91	3-97	1-99
1800				24-76	9-91	3-97
2000					24-76	9-91
2200						24-76

また、下の調整手法により、即詰み局面のミスの模倣ありのAIと即詰み局面のミスの模倣なしのAIを用意する。

即詰み局面のミスの模倣ありのAI：5にて述べた提案手法。

即詰み局面のミスの模倣なしのAI：5にて述べた提案手法の中に式21を式22に変えるもの。

$$T(n, p) = \{MAX(), \quad 0 \leq p \leq 1 \quad \text{式22}$$

7.2 結果

即詰み局面のミスの模倣があるAIの推定レーティング差と設定値との誤差の関係を表したものが表5である。即詰み局面のミスの模倣のないAIの推定レーティング差と設定値との誤差の関係を表したものが表6である。後者のほうが誤差平均の少ないことが示された。

表4 推定レーティング差と設定値との誤差

	1200	1400	1600	1800	2000	2200	2400	平均
1200		-77	-69.8	-198	-124			-117.2
1400	-77		-69.1	-112	-122	-1.7		-76.34
1600	-69.8	-69.1		-9.2	-36.8	-151	-124	-76.567
1800	-198	-112	-9.2		-100	-98.7	3.9	-85.667
2000	-124	-122	-36.8	-100		-9.2	-18.3	-68.367
2200		-1.7	-151	-98.7	-9.2		-44.5	-60.94
2400			-124	3.9	-18.3	-44.5		-45.7
平均								-75.6

表5 推定レーティング差と設定値との誤差

	1200	1400	1600	1800	2000	2200	2400	平均
1200		-35.9	-36.8	-176	-1.7			-62.525
1400	-35.9		-77	-53.9	-218	-196		-116.24
1600	-36.8	-77		40.8	49.4	76.1	-196	-23.933
1800	-176	-53.9	40.8		-18.3	1.9	-47.9	-42.183
2000	-1.7	-218	49.4	-18.3		-44.5	-18.3	-41.95
2200		-196	76.1	1.9	-44.5		9.9	-30.54
2400			-196	-47.9	-18.3	9.9		-63.1
平均								-52.356

7.3 考察

表5のレーティングの平均誤差は-75.6であり、本研究の目標とした±70を超えてしまった。また、レーティングと誤差の平均との相関係数は0.87549となり、高い相関があることが言える。レーティングが高くなるほど調整の精度も高くなると示唆された。結果を見るとレーティング2000以上のみに限定すれば、平均誤差は±70以内に収まっていることが示された。

一方、即詰み局面のミスの模倣を除いた実験では、除いていない場合に比べて誤差が小さいことが確認され、即詰み局面のミスの模倣が棋力調整に悪影響を与えている可能性があることも示唆された。

棋力がうまく調整できなかった理由としては、棋力推定に用いた棋譜が10分切れ負けの棋譜であったことも影響している可能性がある。切れ負けの将棋では、持ち時間が短くなると、必ずしも正確に詰ますことが目的でなくなるため、そのような棋譜が悪影響を与えている可能性も考えられる。

8. まとめ

評価実験では、比較的棋力の高いプレイヤーで、人間らしさと不自然さの評価において、両者とも提案手法の評価が従来手法より良い評価が得られた。また、自己対戦実験では、調整の精度と設定棋力との間に高い相関があることが確認され、比較的棋力が高い場合ほど、調整の精度が高くなることがわかった。すなわち、高い棋力のプレイヤーに対して、本手法の有効性が示唆された。

一方、即詰み局面のミスの模倣は、単に出現確率だけを模倣するだけでは棋力の調整が難しいことが示唆された。

謝辞

本研究は JSPS 科研費 18H03347 の助成を受けたものです。

参考文献

- [1] 馬場匠, 伊藤毅志 “少ない棋譜からの将棋プレイヤー棋力推定手法の提案”. 研究報告ゲーム情報学 (GI), 2019, 2019.13: 1-8.
- [2] 仲道隆史; 伊藤毅志. “人を楽しませる接待将棋システム”. In: 人工知能学会全国大会論文集 第 28 回全国大会 (2014). 一般社団法人 人工知能学会, 2014. p. 1E5OS23b5i-1E5OS23b5i
- [3] 上田陽平, 池田心 “遺伝的アルゴリズムによる人間のレベルに適応する多様なオセロ AI の生成”. 研究報告ゲーム情報学 (GI), 2012, 2012.5: 1-8.
- [4] “技巧”, <https://github.com/gikou-official/Gikou>, (参照 2019-04-25).