

コンピュータオセロによる自然な手を選択する 棋力調整手法の提案

A Method for Selecting Natural Moves for Adjusting Performance of Computer Othello AI

高木騰也*

藤井叙人†

片寄晴弘‡

概要

ゲーム熟達や学習への動機付けとして、プレイにおける楽しさの確保は必要不可欠である。自分と同程度の強さの AI と対戦することで対戦する楽しさが増加することが分かっており、先行研究では棋譜の学習によって当該プレイヤーと同程度のレベルとする棋力調整システムが実現されているが、プレイヤーの棋譜を用意しないとならないという課題が残る。また、プレイ中に棋力調整を実施するシステムも提案されているが、手加減の際にプレイヤーに不自然さを感じさせてしまう課題がある。よって本稿ではプレイ中に棋力調整を実施しつつ、不自然さをより感じにくくする手法を提案し、システムデザインとシステムパラメータの設定手法について述べる。提案システムの有効性は、1) レベルの違う対局相手に対してもおよそ 50% の勝率となるよう対戦のコントロールができるか、2) 対局相手に手加減の不自然さを感じさせないかの二点で推測される。ここでは 1) の視点における手法の有効性を示した後に、2) について既存の棋力調整手法と提案手法を比較する被験者実験を行うことで有効性を示す。実験の結果、十分棋力調整できていることが確認できた。また、被験者実験の自由記述のアンケートから不自然な手を抑制できていることが示唆される結果を得ることができた。

Abstract

In order to motivate players to master and learn games, it is essential to ensure the enjoyment of playing. It is known that playing against an AI with the same level of strength as the player increases the enjoyment of playing. However, this system requires the player's game record to be prepared. Therefore, in this paper, we propose a system that adjusts the player's strength during play, but makes the player feel less unnatural, and describe the system design and the method for setting the system parameters. The effectiveness of the proposed system can be estimated from two points of view: 1) whether the system can control the game so that the winning rate is approximately 50% even against opponents of different levels, and 2) whether the system does not make the opponent feel unnatural. In this section, after showing the effectiveness of the method from the viewpoint of 1), we will show the effectiveness of 2) by conducting a subject experiment to compare the proposed method with existing methods for adjusting the game strength. As a result of the experiment, it was confirmed that the players were able to adjust their playing strength sufficiently. In addition, we obtained the results suggesting that the unnatural moves were suppressed from the questionnaires written freely by the test subjects.

1 はじめに

ゲーム AI の学術的研究において“強さ”という分野では人間のプロを上回るほどの十分な領域に達していると言える。コンピュータが人間に勝つことは難しいと言われていた囲碁において囲碁プログラム AlphaGO が 2016 年にトッププロ棋士に勝利したニュースは記憶に新しい。2017 年には、AlphaGo Zero という過去の棋譜を一切用いずに自己対戦のみから学習をする囲碁プログラムが、AlphaGo 以上の強さに至っており、AI は人

間積み上げた知識を使わずともプロに勝てるという領域に達している。ゲーム AI の学術的研究における興味は、「強い」コンピュータ AI(以下、AI)の構成から、「楽しさ」を提供する AI の構成へと遷移しつつある。

アマチュアプレイヤーにとって十分すぎるほどに強くなっている AI はゲーム熟達の支援や対戦ゲーム BOT として応用されることが期待される。しかし、これら学習への応用を考えるとプレイヤーがプレイしたいと思えるような動機付けが必要になり、そのためには、プレイすること自体の楽しさが要求される。伊藤の研究 [1] では人間らしいミスを犯す AI のフレームワークのデザインの構築を行い、その中でプレイヤー自身のレベルに近い AI との対戦は、対戦する面白さが増加することを示し、

* 関西学院大学

† 福知山公立大学

‡ 関西学院大学

ゲーム AI に”強さ”という方向以外の新しい評価基準をもたらした。また池田らの研究 [2] ではプレイヤーを楽しませるためにはプレイヤーと同程度の棋力に調整すること、またその際に明らかに手を抜くなど不自然な手を選択しないことが挙げられた。プレイヤーを楽しませるためにはプレイヤーに悟られずに負けることが必要になり、悟られずに負けるには不自然な手を選択しないことが必要になる。

棋力調整手法として上田らの研究 [3] ではプレイヤーの棋譜を深層学習に使用することで棋力を調整する手法を提案した。結果として棋力を調整することができ、不自然な手もほとんど感じない結果になったが、プレイヤーの棋譜を用意しなければならないという課題が残った。仲道ら [4] の研究では棋譜を必要とせず AI が優勢であれば、わざと弱い手を選択させ、劣勢であれば強い手を選択させることで棋力を調整する手法を提案した。プレイヤーと同程度の強さに調整できることは確認できたが、プレイヤーが悪手と感じてしまうような不自然な手が確認された。プレイヤーを楽しませるためには棋力調整を行いながら、不自然な手を選択しないことが望ましい。

よって本研究では、棋譜を用いず棋力を調節しながら不自然な手を抑制する動的棋力調整手法を提案する。提案手法で用いるパラメータを実験で決定することで、棋力調整できているかを確認する。のちに既存の動的棋力調整手法と比較して自然な手を選択できているかについての評価実験を行う。

2 関連研究

2.1 手の不自然さについて

手の不自然さについては様々な種類があり、手を抜かれたと感じるような戦略に関する手や、AI の思考時間が長い、思考時間が常に一定による不自然さなど戦略以外の手が存在する。池田らの研究 [2] ではプレイヤーを楽しませるために必要な要素技術を列挙した。その中の一つとして、戦略に関する不自然な手として「形が悪い手」、「流れにそぐわない手」、「明らかに損する手」、「高度すぎる手」があると述べられている。これらの手はあからさまな手加減をされているとプレイヤーに感じさせ、楽しさが減少してしまう要因となる。

2.2 棋力調整手法

現在、棋力を調整させる AI の研究は盛んに行われおり、特にオセロ、チェス、将棋のようなボードゲームでの研究が盛んである [5][6]。調整手法は様々であるが、大きくプレイヤーの棋譜を必要とするものと、しないもので分けることができる。

2.2.1 棋譜を使用する棋力調整手法

上田らの研究 [3] ではプレイヤーに minmax 法によって深さ 5 まで探索する AI と対戦させ、プレイヤーの棋譜を

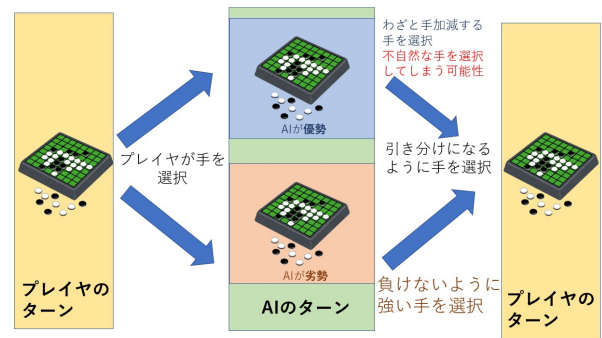


図1 動的棋力調整手法

取得した。その棋譜を用いてプレイヤーと同程度の棋力である AI を深層学習により用意した。その後、プレイヤーに深層学習させた AI と戦わせどの程度の強さと感じたか、不自然な手は感じたかというアンケートを行い、プレイヤーに対して棋力調整ができていることが確認でき、不自然な手も 65 % の人が全く感じないという結果となった。しかし、常にプレイヤーの棋譜が前もって取得できるわけではなく、棋譜がない状態も考えられる。そのような場合、プレイしながらプレイヤーの棋力を測定し、動的に棋力調整を行う必要がある。

2.2.2 棋譜なしの動的棋力調整手法

仲道ら [4] の研究では図 1 のように盤面ごとに算出される評価値に基づき AI 側が優勢ならわざと手加減する手を選択し、劣勢なら勝つための強い手を選択することで棋譜を使わない棋力調整手法を提案した。

結果として十分に AI に手加減できていることが示すことができた。しかし、プレイヤーに比べ十分強い AI にわざと弱い手を選択させることで、「手の見逃し」や「明らかに損する手」など不自然な手を感じる意見が挙げられたことが課題として残った。動的調整手法はプレイヤーと同程度の棋力調整はできたものの、プレイヤーから見て不自然と感じる手があるというのが課題として残っている。よって本研究ではその課題を緩和する棋力調整手法を提案する。次章ではその提案手法のアプローチを説明したうえで、明らかにすべき項目を列挙し、行う実験について述べる。

3 アプローチ・実装

3.1 アプローチ

アプリケーションなどでの実装を想定したとき、棋力調整を行う際に、対局前にプレイヤーの棋譜を用意するのは難しいため、仲道らの研究のように動的に棋力を調整する手法が望ましい。仲道らの動的調整手法を採用する場合、AI の選択する手を常に勝率が 50 % になるように選択させることから、AI が優位の状況ではわざと AI に手加減をさせる必要がある。その手加減させる手の選択がプレイヤーにとって不自然な手と感じる要因になっていると考える。

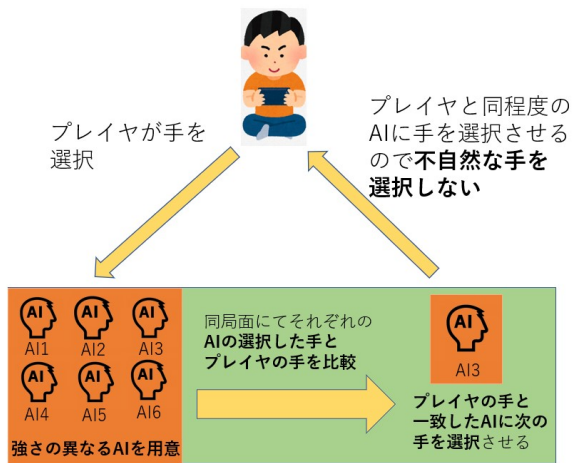


図2 提案アプローチ

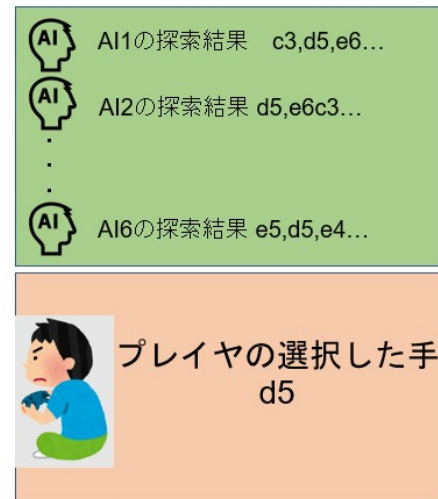


図3 プレイヤーの棋力測定手法

AIに手加減させないようにするためには上田らの研究のようにプレイヤーと同程度の強さのAIを用意する必要があり、同程度の強さのAIの学習にはプレイヤーの棋譜が必要になる。よって本研究ではプレイヤーの棋譜を用いずに同程度の強さのAIと対局させる手法として、図2のように事前に複数の強さの異なるAIを用意しておき、プレイヤーの強さに合わせて使用するAIを決定する手法を提案する。これにより上田らの研究のようにプレイヤーと同等のAIに手を選択させるため、棋力調整を行いながら、不自然ではない手を動的に選択させることができると考える。

3.1.1 プレイヤーの手の強さの評価方法

上田らの研究のように同等の強さのAIに手を選択させるためには、プレイヤーの選択した手がどの程度で、複数用意したAIのどれと同等なのかを判別させる必要がある。

本研究では同じ局面でプレイヤーとAIの選択した手が一致するかどうかで同程度の強さであるか、否かを判別させる。具体的にはプレイヤーのターンにそれぞれのAIにAIならどの手を選択するかを探索させる。そのAIが出した手と実際にプレイヤーが選択した手が一致した深さを用いて、次の手を探索させる。例を図3に示し、図2と共に説明する。

AIを1から6までの6つ用意したとき、プレイヤーのターンにそれぞれのAIに探索させる。図3に示しているのがAIの探索結果例である。プレイヤーの選択した手がどのAIの探索結果と一致するかを判断し、一致したAIに次の手を探索させる。例でいうと、次の手はAI2に探索させる。

5つのAIの第1候補の手とプレイヤーの手がどれも一致しない場合、第2候補、第3候補まで一致するかを判別させる。この手法では動的調整手法で課題の要因と考えられるAIに手加減させることはさせず、元々の棋力がプレイヤーと同程度のAIに探索させることで不自然な手を選択させないようにできると考える。本

研究では異なる強さのAIとして、MCTSの探索する深さの上限を変更することでAIの強さを変更する。探索には以下の式、UCT (UCB1 applied to trees)[7]を用いる。

$$\bar{X}_i = c\sqrt{\frac{2\log n}{n_i}} \quad (1)$$

3.2 不自然な手の定義

不自然な手とはプレイヤーの熟達度によって大きく変わってしまう可能性がある。熟達者では様々なプレイヤーモデルを持っており、プレイヤーごとにそのモデルは変わってくるため、不自然な手を定義するのは難しい。よって本研究では熟達者ではない初級者と中級者を対象に不自然な手について評価する。また、本研究では戦略に関する不自然さに注目するため、池田ら[2]の「流れにそぐわない手」、「明らかに損する手」、「形が悪い手」、「高度すぎる手」を不自然な手として採用する。

3.3 実装環境

本研究ではAlphaGOのように複数のPCを使用するのではなく、単一のPCを使用し現実的な探索時間でオセロをプレイできることを目標とする。本研究では以下の条件でアプローチ手法を実装、動作させた。

- ノートパソコン
Razer (Intel Core i7, メモリ 16GB, GPU NVIDIA GeForce RTX2060)

オセロを選択した理由としては、競技人口が多く、ルールが単純であるため、初級者、中級者が感じる不自然な手が分かりやすいと考えたためである。また、本研究ではノートPCを用いて個人でも実装、対局ができるように複数のGPUを必要とせず、手を決定するための探索時間が長すぎないことを目標とする。

3.4 プレイヤのミスについて

本研究のアプローチ手法での考えられる課題として、プレイヤーのミスがあげられる。プレイヤーが誤って手を選択してしまうようなプレイヤーの棋力とは大幅に異なる手を選択されてしまう可能性がある。その手を AI が弱い手や、強い手と判断し、それと同等の手を選択した場合、プレイヤーは本来の実力より劣っている AI の手に対して不自然さを感じてしまう可能性がある。よって本研究では以下の式に従って選択する。

$$T_s = \frac{\sum_{i=0}^S T_i}{S} \quad (2)$$

この式は S は何手目か、 T はプレイヤーと同程度の強さと判断した探索の深さである。ゲーム開始から測定したプレイヤーの棋力の平均をとり、その値に近い強さの AI に手を探索させる。例を挙げるとプレイヤーが強さレベル 5, 5, 5 のような強さの手を選択して次の手に強さレベル 1 の手を選択した時、上記の式から強さレベル 4 の AI に探索させるようにする。これにより、急激な棋力の増減を防ぎ、プレイヤーに不自然な手と感じさせないようにする。

3.5 実験項目

本研究での実験の目的は、大きく二つあり、提案手法が棋力調整できているのかどうかを確認することと、提案手法が既存の動的棋力調整手法と比較して、不自然な手を選択していないかを評価することである。

手の自然さを評価する前に、本研究では棋力調整ができていないということが前提であるため、棋力調整ができていないかの実験が必要になる。更に棋力調整を行うためには、どの程度の強さ(深さ)の AI をいくつ用意するのか、どの程度の強さのプレイヤーなら適用できるのかを明らかにする必要がある。よって本研究では以下の 1 から 3 の実験にて使用する AI の強さの範囲、提案手法を適用することのできるプレイヤーの強さの範囲、使用する AI の数のパラメータを決定し、提案手法が棋力調整ができているのかを確認、その後 4 の実験にて手の不自然さについての評価実験を行う。

1. 探索深さとレーティングの関係
2. 探索時間の測定
3. BOTAI との対戦
4. 着手の自然さの評価実験

4 実験

前項で述べたように手の自然さについての評価を行う前に棋力調整ができていないのか、棋力調整させる際に使用する AI の深さの範囲、提案手法を適用することがで

異なる深さのレーティング

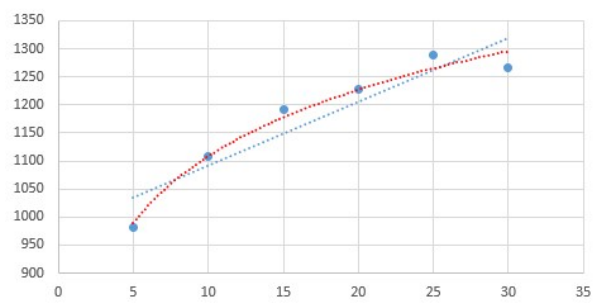


図4 深さ別レーティング

きるプレイヤーの強さの範囲、使用する AI の数を明らかにさせる必要がある。まず探索させる深さの範囲を決定するために、探索深さとレーティングの関係について明らかにする。その後、用いる AI の総数を決定するために、一つの AI の一手当たりの平均探索時間を測定することで、オセロにおいて一局あたりの持ち時間と比較し本研究で用いる AI の総数を決定する。そして決定したパラメータを使用し、実装した提案手法が棋力調整できているのかを BOT と対戦させることで評価する。本研究ではプレイヤーの強さの指標としてオンラインアプリケーション「オセロクエスト」のレーティングを用いる。

4.1 探索深さとレーティングの関係

提案手法で用いる探索させる深さの範囲と適用することのできるプレイヤーの強さの範囲を明らかにするために探索深さとレーティングの関係について明らかにした。実験内容としては MCTS の深さ別に「オセロクエスト」[8] の「5分対局」でレーティング測定を行い、探索の深さが変わることでのレーティングの変化について検証した。ここで MCTS の深さを 5, 10, 15, 20, 25, 30 とし MCTS の制限としては総探索時間が 5 分を超えないように一手当たりの探索最高時間を 8 秒 (300 秒/終局) とした。

図 4 が深さ別にレーティングを測定した結果である。直線近似を青で、曲線近似を赤で表示した。近似直線を見ると右肩上がりであるが、深さ 30 のレーティングは深さ 25 よりも低い値であったこと、近似曲線を見ると深さ 25 以上はレーティングは収束しているように見える。この理由として探索最高時間を 8 秒としているからだと考える。この制限がない場合は深さを深くすることでよりレーティングは上がっていくと推測する。実験結果より、本研究ではプレイヤーのレーティングが約 980 以上 1300 以下のプレイヤーを評価対象とし、使用する深さの範囲を 5 から 25 とする。本研究での評価対象であるレーティング 980 以上 1300 以下というのはおおよそ初級者から中級者にあたる。

4.2 探索時間の測定

前述の実験により提案手法が適用できるプレイヤーの強さと本研究で使用する探索深さの範囲が決定した。次に

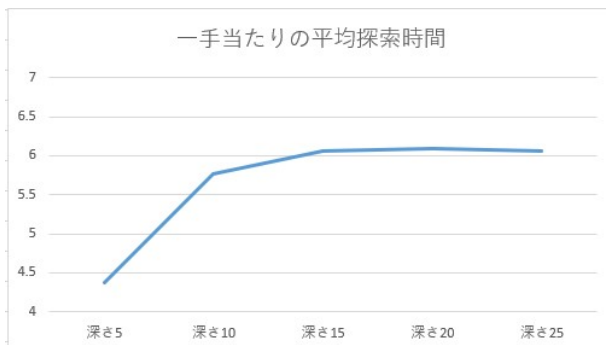


図5 一手当たりの平均探索時間

一手当たりの平均探索時間を測定することで、深さ5から深さ25までのAIの中で使用するAIの数を決定する実験を行う。実験に使用するエージェントはMCTSの深さ5,10,15,20,25で制限として最高探索時間は8秒とする。5つのAIそれぞれを100戦自己対局させ計200局分の一手当たりの探索時間を測定した。

図5は実験結果の深さ別の一手当たりの思考時間を測定結果である。深さ5の時は一手4.3秒で深さ15では一手6.0秒で深さ15以上は探索時間は一定になるという結果になった。これは最高探索時間を8秒と制限したことで深さ15以上では一手当たりの思考時間が一定になったと考察する。

本研究での一局の想定持ち時間としては、アマチュア大会の持ち時間として多く採用されている20分とする。よって一手当たり約40秒の探索時間がある。実験結果より、深さ5と深さ25の平均探索時間は約5.5秒かかるので計7つのAI深さ5,9,12,15,18,21,25を使用する。

この実験により、上記の7つのAIを用いた場合、一手当たりの探索時間は38.5秒となり、一手40秒以内に抑えることができ、思考時間が長すぎることなく、現実的に対局が可能であることが確認できた。

本研究ではさらに探索時間の短縮を目指し、7つのAIの探索を並列化させる実装を行った。並列処理を行う前は一手当たり平均38.5秒かかっていたのを平均9.6秒まで短縮させることができた。これにより、待ち時間による手の不自然さや、退屈さを軽減させることができると考える。

4.3 BOTAI との対戦実験

前述の二つの実験で求めたパラメータを使用した提案手法が棋力調整を行うことで勝率を50%程度に調整できているかどうかについて実験を行う。深さ5,9,12,15,18,21,25のMCTSを使用し、一手の最大思考時間を8秒とした提案手法を実装し、深さ5,10,15,20,25のMCTSのBOTAIとそれぞれ50局対局させ、勝率を測定したものを図6に示す。

実験結果から深さ5に対しては勝率60%を若干超えてしまっているものの、それ以外では約50%という結果になった。どの強さのAIと対局しても約50%を保って

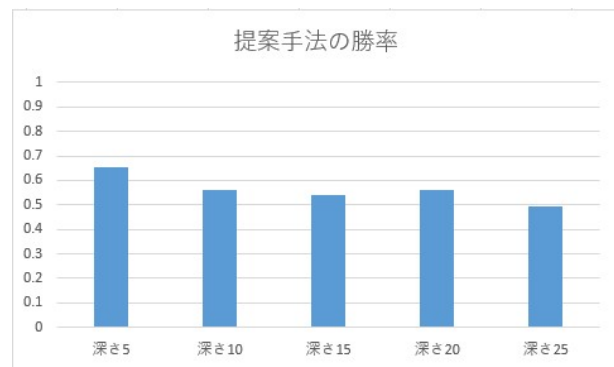


図6 提案手法のBOT との勝率

いることから、レートが980から1300までの範囲では高い精度で棋力調整ができていたことが分かった。

これらの実験結果から、深さ5から深さ25までの7つのAIを用いて提案手法を実装することで、一局あたりの思考時間を20分以内に抑え、深さ5から25までのBOTAIに対してはすべてにおいて勝率を約50%に調整し、精度の高い棋力調整ができていたことが分かった。

4.4 着手の自然さの評価実験

4.1章から4.3章までで一手あたりの思考時間が長すぎることなく、棋力調整するためのパラメータを求め、それをを用いることで提案AIが自分より弱い相手でも棋力調整できるということを確認できた。本章では既存の動的に手加減する手法と比較して、提案AIより弱い相手に対し、自然な弱さをもつ同程度の対戦相手として利用できるかを人間による主観的評価を行う実験により検証する。

4.4.1 実験参加者

実験参加者はオセロのルールは知っているが、特に学習経験のない3人を対象とした。被験者には4.2章で求めた本研究での対応できるレーティングの範囲、980以上1300以内であることを判別するため、オセロクエストを用いて最低50戦以上プレイさせることでレーティングを収束させて提案手法が対応可能か判別した。以下がレーティング結果である。

表1 被験者のレーティング

被験者	1	2	3
レーティング	1149	1188	1242

結果より、3人の被験者は本研究の提案手法で対応可能だと考える。

4.4.2 実験手続き

上記3人を対象に、3章で述べた本研究で定義する手の自然さについて説明を行い、仲道らの手法で手を選択するAI(以後、動的AI)、提案手法のAIと先後1回ずつ、計4回をランダムに対局して一局ごとに勝敗を回答させ、手の強さについて、不自然な手の有無について、対局の楽しさについての5段階のアンケート、対局後の感想、

どのような手が不自然と感じたかについての自由記述のアンケートを行った。対局中のオセロ AI の手の不自然さは着手以外の情報からの良し悪しを評価しないように教示した。実験手続きは以下の通りである。

1. 実験についての説明, 定義する手の自然さについての説明をした。
2. 動的 AI と提案 AI の 2 種類のうち一つと対局させた。
3. 対局後, アンケートに答えさせた。
4. 2 と 3 を 4 回繰り返した。

現状, オセロクエストにて最低 50 戦以上プレイしてもらうのは被験者にとって大きな負荷であり, 3 人の被験者しか実験可能ではないため, 実験結果では対局後の感想, どのような手が不自然と感じたかの自由記述で得られた回答を実験結果で述べ, 動的 AI と提案 AI での不自然な手の有無, どのような手が不自然と感じられたかについて考察する。

4.4.3 実験結果

不自然な手の記述について動的 AI のアンケートでは「勝てたけど, 角をとらせるような手があった」が 1 件, 「角を取れるタイミングでとらなかった」が 1 件あった。また, 「勝てて楽しかったが, 勝たせてもらった気がした」という記述があり, 盤面の角での不自然さが目立つことが確認できた。一方で提案 AI では角を取れるところでとらないなどの明確な不自然な手の記述はなかった。

楽しさについての記述は動的 AI と提案 AI 両方で「楽しかった」という記述はあったが, 動的 AI では「勝てて楽しかった」という勝敗による楽しさの記述があったことに対し, 提案 AI では「特に中盤から終盤にかけて予想通りの手を打ってくるのが多くあったので AI が強さを合わせてくれる感じがして楽しかった」という記述があった。

4.4.4 検討

実験結果の手の不自然さについて, 動的 AI は盤面の角で不自然な手が確認されたが, 一方で提案 AI では確認されなかった。これは提案 AI は対局で盤面の角周辺で不自然な手を選択させることなく手を選択することができるという示唆される。

また, 楽しさについて, 提案 AI では「特に中盤から終盤にかけて予想通りの手を打ってくるのが多くあったので AI が強さを合わせてくれる感じがして楽しかった」という記述についてこのように自分と同レベルに感じたという記述は動的 AI には見られなかった。動的 AI は形勢を保つために選択する手の強さを変更していくのに対し, 提案 AI はプレイヤーと同等のレベルの手を選択し続けることによってこのような記述があったと考えることができる。

被験者実験については予備調査を実施した段階にある。今後, 最低 10 人程度に被験者数を増やし, 不自然な手について既存手法と比較して抑制できているかどうか, また提案手法にはどのような不自然な手がみられるかについて評価, 考察していきたい。

5 結論

本研究では既存の棋力調整手法の不自然な手を選択されてしまう原因について考察したうえで自然な手を選択する棋力調整手法を提案した。その後, 既存手法と不自然な手について被験者実験を行うことで評価した。また提案手法が棋力調整できているのかという問題に対しては実験結果から提案手法は実験で求めたパラメータを使用することで高い精度の棋力調整ができていることが分かった。被験者実験についてはまだ被験者数は少ないものの, アンケートの自由記述から, 既存手法より, 提案手法の方が自然な手を選択できていることが示唆されるような結果が得られた。今後の展望としては手の自然さについての評価実験の被験者数が少ないので, さらに被験者数を増やすことで有意差の有無の確認, どのような不自然な手を減らすことができたか, 提案手法で確認される不自然な手は何かあるかを確認していく。

参考文献

- [1] 伊藤毅志, ミスを犯す人間らしいゲーム AI の研究, 科学研究費助成事業 研究成果報告書 (2013-2015), <https://kaken.nii.ac.jp/ja/file/KAKENHI-PROJECT-25280130/25280130seika.pdf>
- [2] 池田心, Simon Viennot, モンテカルロ碁における多様な戦略の演出と形勢の制御～接待碁 AI に向けて, 情報処理学会シンポジウム論文集 (2012)
- [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, 第 27 回ゲーム情報学研究会 (2012)
- [4] 仲道隆史, 伊藤毅志, プレイヤの技能に動的に合わせるシステムの提案と評価, 情報処理学会論文誌 57(11), 2426-2435(2016)
- [5] I-Chen Wu, On Strength Adjustment for MCTS-Based Programs, AAI(2019)
- [6] Reid McIlroy-Young, Aligning Superhuman AI with Human Behavior: Chess as a Model System, KDD '20: Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining(2020)
- [7] Levente Kocsis, Csaba Szepesvári, Bandit based Monte-Carlo Planning, European Conference on Machine Learning(ECML), (2006)
- [8] <http://wars.fm/reversi?lang=ja>