

プレイヤーの強さに動的に合わせるオセロ AI

Othello AI That Dynamically Adapts to the Strength of the Opponent Player

塚本 賢

Ken Tsukamoto

法政大学情報科学部コンピュータ科学科

E-mail: ken.tsukamoto.7s@stu.hosei.ac.jp

Abstract

In the realization of the game proficiency support system, AI that plays weakly and naturally according to the strength of the opponent player during the game has an important role. The AI of board games such as Shogi has acquired the strength to beat the professional players. However, most of such AI does not dynamically match the strength of opponent players. To tackle this problem, Nakamichi et al. proposed AI that modified the evaluation function of Shogi to weakly and naturally play it. A major problem with their method is that positive evaluation values are simply converted to negative values; it always uses the same threshold of zero for converting evaluation values. This paper proposes a method for improving Nakamichi et al.'s evaluation function and applies it to Othello AI. Unlike Nakamichi et al.'s method that always uses the same conversion, the proposed method converts evaluation values by using a threshold that it dynamically changes according to the game state. This paper presents the result of experiments on the plays of the proposed method against other AI and human players.

1. はじめに

人工知能の発展とともに、ゲームの AI に関する研究が盛んに行われている。ゲーム AI の歴史を語るうえで欠かせないのは、チェス・将棋・囲碁などのボードゲームの AI である。1996 年には、IBM が開発したチェス AI の Deep Blue が当時チェスの世界チャンピオンであったガルリ・カスパロフに 1 勝し、翌年には 2 勝した。2010 年に行われた情報処理学会 50 周年記念対局では、将棋 AI である、あから 2010 が清水市代女流王将に勝利した。2016 年には、Google DeepMind が構築した囲碁 AI の Alpha Go が囲碁の世界チャンピオンであったイ・セドルに勝利した。チェス・将棋・囲碁に限らず、バックギャモンなどの他のボードゲームでも AI が勝利する事例が起こっている。このように、ボードゲームの AI はプロに勝利することができるほどの強さを誇っている。

ゲームの熟達支援システムの実現で、対局中にプレイヤーの技能に合わせて不自然でなく弱く指す AI が重要な役割を持っている。仲道ら [1] はプレイヤーの技能に合わせて

不自然でなく弱く指す将棋 AI を開発した。仲道らの手法は、局面評価値が最も 0 に近づく手を選択するために、局面における評価値が正の値の時は負の値に変換し、評価値が負の値の時はそのまま用いる。

本研究では、相手のプレイヤーの強さに合わせて弱く指すオセロ AI の開発を目標とする。そのために仲道らの将棋 AI の評価関数をオセロ AI に適用する。しかし、予備実験の結果、仲道らの将棋 AI の評価関数はオセロ AI でうまく働かないことが分かった。このため、評価関数を改良する。具体的には、まず、仲道らの手法における評価値 0 を閾値と見なし、評価値が閾値を超えたときに、評価値が高いほど低い値に変換しているを見なす。仲道らの手法では閾値が 0 に固定されていたのに対して、本研究では局面に応じて動的に閾値を変更する。具体的には、隅にある石が自身のものであるか相手のものであるかに応じて動的に閾値を変更する。この評価関数を加えた提案 AI を評価するために、複数の AI との対戦実験と人間との対戦実験を行った。実験の結果、複数の AI に対して、また人間に対しても強さを動的に合わせられることが分かった。

2. 関連研究

強いボードゲームの AI を実現する研究が多くなされている。Google DeepMind [2] は、ディープラーニングと機械学習を用いた手法で囲碁の局面を評価する新しい手法を確立し、その手法と従来のモンテカルロ木探索の手法を組み合わせた Alpha Go を構築した。関ら [3] は、ゲーム・非ゲームを問わず様々な応用が模索されているモンテカルロ木探索において、チェスや将棋などの一部のゲームでは良い AI が作成できていない問題点に着目している。そのうえで、従来法の発展した将棋において、従来法との比較を通したモンテカルロ木探索の長所短所の解明を行った。

池田 [4] は、強さという意味では十分な域に達しつつある囲碁・将棋 AI の基本的枠組みを解説した。そして、「楽しませるための AI」を取り巻く現状をまとめ、具体的に必要となる要素技術と、そのためのアプローチを紹介した。

3. 仲道らの手法

仲道ら [1] は、相手のプレイヤーの技能に合わせて、不自然でなく動的に弱く指す将棋 AI を開発した。候補手集合

の中で指したときの局面評価値が最も 0 に近づく手を選択するために、局面における評価値が正の値の時は負の値に変換し、評価値が負の値の時はそのまま用いるように評価関数を再定義した。具体的には、探索の結果得られた現局面 M における候補手ごとの評価値 $V(M)$ を次の関数 $V_0(M)$ によって変換している。

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

この評価関数を用いて、ミニマックス法に基づくゲーム木探索を行った。

提案手法を評価するために、AI との対戦、初心者と経験者のプレイヤーを対象とした対戦、インターネット対局場での対戦による 3 つの実験を行った。探索の深さ 5 の提案 AI と深さ 1, 2, 3, 4 の既存 AI をそれぞれ 500 回対戦させた結果、深さ 1, 2, 3 の既存 AI に対しては 5 割に近い勝率を記録し、それ以外の深さがあまり変わらない既存 AI で勝率は 2 割近くであった。また、人間に対しての実験では、将棋のアマ級位・段位格の技能を持つ経験者 5 人は、深さが 1 の既存 AI と深さ 5 の提案 AI に対して勝率が 9 割、深さが 5 の既存 AI に対しては勝率が 2 割ほどであった。将棋の経験の無い初心者は、深さが 1 と 5 の既存 AI に対して勝率が 0 割だったのに対し、深さが 5 の提案 AI に対しては勝率が 4 割ほどであった。

4. 提案手法

本研究は仲道らの将棋 AI の評価関数を改良した上でオセロ AI に適用する。オセロは、隅にある石がゲームの形勢に大きく影響し、このため、評価値の計算で 4 つの隅にある石を重視することが多い。この性質のために、仲道らの手法で、評価値を変更する閾値が 0 に固定されていると、隅にある石に応じた形勢の変化に対応できないと考えられる。

そこで本研究では、局面に応じて動的に閾値を変更する改良を行う。具体的には、基本の閾値を 5 とした上で、隅にある自身の石の個数に応じて閾値を 15 ずつ下げ、さらに、隅にある相手の石の個数に応じて閾値を 15 ずつ上げるように動的に閾値を変更する。

すなわち、現局面 M の評価値 $V(M)$ を変換する関数 $V_0(M)$ を次の通り定義する。

$$V_0(M) = \theta - |V(M) - \theta|$$

ただし、 θ は閾値であり、以下の通り定められる。

$$\theta = 5 - 15c_m + 15c_o$$

ここで、 c_m は M で隅にある自身の石の個数であり、 c_o は隅にある相手の石の個数である。

5. 実装

ミニマックス法の改良である $\alpha\beta$ 法を用いたオセロ AI を Java で実装した。基本の評価関数は、葉の局面と葉でない局面とで分かれている。葉の局面では、自身の石の個数と相手の石の個数の差を用いている。一方、葉でない局面では、以下の 4 通りの方法で評価値を計算する 4 つの AI を作成した。

- 自身の石の個数を用いる (以下、石数 AI)。

- 自身の石の個数を基本とし、さらに隅にある自身の石の個数に応じて 8 ずつ評価値を上げ、隅にある相手の石の個数に応じて 8 ずつ評価値を下げる (以下、石数・隅 AI)。
- 自身の石を置くことのできる選択枝の数を用いる (以下、選択枝数 AI)。
- 自身の石を置くことのできる選択枝の数を基本とし、さらに隅にある自身の石の個数に応じて 8 ずつ評価値を上げ、隅にある相手の石の個数に応じて 8 ずつ評価値を下げる (以下、選択枝数・隅 AI)。

さらに、選択枝数・隅 AI の評価関数に前節の提案手法を適用した AI (以下、提案 AI) を作成した。また、比較対象として、同じ評価関数に仲道らの手法を適用した AI (以下、仲道 AI) を作成した。

人間のプレイヤーとの対戦のためにオセロのユーザインタフェースを用意した (図 1)。人間のプレイヤーは盤上のマス目をクリックして石を置くことができる。

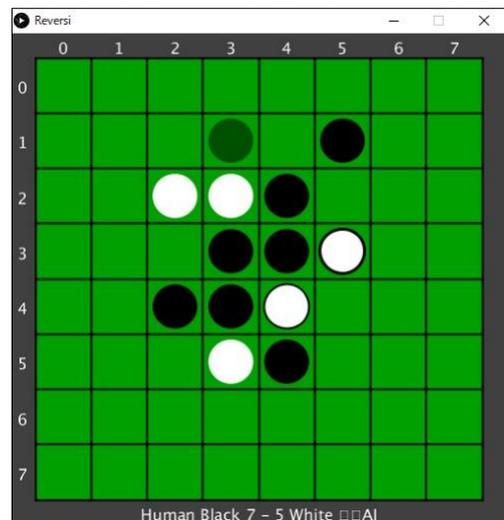


図 1 オセロのユーザインタフェース

6. 実験

本研究の手法を用いて、相手のプレイヤーの強さに動的に合わせられるかを調べる実験を 2 つ行う。1 つ目は、前節に述べた 4 つのオセロ AI と対戦させる実験である。2 つ目は、人間のプレイヤーと対戦させる実験である。

6.1. 実験 1

$\alpha\beta$ 法における探索の深さを 10 に設定した提案 AI と、前節に述べた 4 つの基本的な AI で探索の深さを変えたものを、先手後手を入れ替えて 50 回ずつ、計 100 回対戦させる実験を行った。さらに、比較対象として、選択枝数・隅 AI と石数 AI を対戦させる実験、選択枝数・隅 AI と選択枝数・隅 AI を対戦させる実験、仲道 AI と石数 AI を対戦させる実験も行った。

実験結果を表 1 に示す。この表で勝敗は「自身」の AI から見て勝ったか負けたかを表す。

表 1 AI との対戦実験の結果 (% 値は勝率を示す)

自身	相手	深さ 2	深さ 4	深さ 6	深さ 8	深さ 10
選択 枝数・ 偶 AI	石数 AI	100勝 0敗 (100%)	100勝 0敗 (100%)	100勝 0敗 (100%)	100勝 0敗 (100%)	100勝 0敗 (100%)
	選択 枝数・ 偶 AI	100勝 0敗 (100%)	100勝 0敗 (100%)	92勝 6敗 2分 (94%)	74勝 25敗 1分 (75%)	
仲道 AI	石数 AI	49勝 43敗 8分 (53%)	35勝 52敗 13分 (40%)	35勝 59敗 6分 (37%)	31勝 65敗 4分 (32%)	19勝 76敗 5分 (20%)
提案 AI	石数 AI	77勝 17敗 6分 (82%)	69勝 26敗 5分 (73%)	66勝 27敗 7分 (71%)	60勝 36敗 4分 (63%)	52勝 41敗 7分 (56%)
	石数・ 偶 AI	78勝 20敗 2分 (80%)	69勝 23敗 8分 (75%)	57勝 39敗 4分 (59%)	43勝 54敗 3分 (44%)	43勝 56敗 1分 (43%)
	選択 枝数 AI	56勝 35敗 9分 (62%)	44勝 52敗 4分 (46%)	33勝 58敗 9分 (36%)	17勝 79敗 4分 (18%)	
	選択 枝数・ 偶 AI	55勝 36敗 9分 (60%)	38勝 54敗 8分 (41%)	32勝 59敗 9分 (35%)	14勝 85敗 1分 (14%)	

6.2. 実験 1 の考察

まず、実験 1 のうち石数 AI を相手とした場合について考察する。選択枝数・偶 AI は、探索の深さ 2 から 10 の石数 AI のすべてに対して全勝した。一方、仲道 AI は、選択枝数・偶 AI と同じ深さ、同じ評価関数を用いているが、探索の深さ 2 から 6 までの石数 AI に対して、勝率が 4 割から 5 割程度となっており、強さを調整できている。しかし、探索の深さ 10 に対しては勝率が 2 割であり、十分に調整できていない。石数 AI は 4 つの AI の中で最も弱い評価関数を用いているので、一定以上強い評価関数の AI には、強さを合わられないと考えられる。

提案 AI は、仲道 AI と同様、選択枝数・偶 AI と同じ深さ、同じ評価関数を用いているが、探索の深さ 2 から 6 までの石数 AI に対して勝率が 7 割から 8 割程度であった。選択枝数・偶 AI に比べると勝率が 2 割から 3 割程度低いことから、弱い評価関数の AI に対しても強さを動的に合わせられているが、仲道 AI に比べると十分に強さを調整できているとは言えない。しかし、探索の深さ 8 と 10 の石数 AI に対しては勝率が 5 割から 6 割に近い値であることから、より強い AI に対しては仲道 AI よりも動的に合わせられると考えられる。

次に、提案 AI と石数・偶 AI の対戦結果について考察する。石数・偶 AI の評価関数は、石数を基本としたものに、取っている隅の数を考慮していて、前述の石数 AI よりも強くなっている。提案 AI は深さ 2 と 4 の石数・偶 AI に対して勝率が 7 割から 8 割程度であり、ある程度勝た

せることができている。さらに、探索の深さ 6 から 10 までの AI に対しては勝率が 4 割から 6 割に近い値であるので、より強い AI に対して動的に合わせられると考えられる。

次に、提案 AI と選択枝数 AI の対戦結果について考察する。提案 AI と比べ、選択枝数 AI の評価関数は、隅の数を考慮していないものとなっている。提案 AI は深さ 2 から 6 までの選択枝数 AI に対して 4 割から 6 割と勝率を調整できている。選択枝数 AI は前述の石数 AI、石数・偶 AI よりも強くなっているが、提案 AI は負けすぎることなく、強さを動的に合わせられている。しかし、深さ 8 の AI に対しては勝率が 2 割に満たなかった。探索の深さが 2 しか変わらないことと、評価関数にあまり差がないことから、提案 AI は強さを調整する余裕がなく負けすぎたと考えられる。

最後に、提案 AI と選択枝数・偶 AI の対戦結果について考察する。選択枝数・偶 AI は、選択可能な手の数と取っている隅の数を考慮しており、提案 AI と同じ評価関数を用いている。提案 AI は、探索の深さ 2 から 6 までの選択枝数・偶 AI に対して勝率が 4 割から 6 割程度であった。選択枝数・偶 AI に比べると勝率が 4 割から 5 割程度低いことから、同じ評価関数を用いた、探索の深さの異なる AI に対しても、強さを動的に合わせられている。しかし、提案 AI は、探索の深さ 8 の AI に対しては勝率が 2 割に満たなかった。選択枝数・偶 AI は深さ 8 の選択枝数・偶 AI に対して勝率が 7 割程度であるので、深さ 6 に比べて勝率が 2 割程度低くなっており、深さ 6 より明らかに強くなっていることが分かる。このことから、選択枝数 AI の場合と同様に提案 AI と選択枝数・偶 AI が同様の評価関数を用いており、探索の深さが 2 しか変わらないため、提案 AI が強さを調整する余裕がなく負けすぎたと考えられる。

以上の結果から、提案 AI は異なる強さを持つ複数の評価関数を用いた AI に対して強さを調整できていることが分かる。また、探索の深さを変えた場合も、ある程度強さを動的に合わせられていると考える。しかし、強さが近い AI に対しては、強さを調整する余裕がなく合わせられないことが分かる。

6.3. 実験 2

探索の深さが 10 である提案 AI を含む複数の AI と、人間のプレイヤーを先手後手 2 回ずつ対戦させる実験を行った。実験は、まず深さ 4 の選択枝数・偶 AI と対戦させ、その結果に応じて、深さ 2 または深さ 6 の選択枝数・偶 AI と先手後手の 2 回ずつ対戦させる。さらに、深さ 10 の提案 AI と先手後手の 2 回ずつ対戦させる。

実験参加者は年齢 20 歳から 54 歳までの 5 人 (平均年齢 28 歳、男性 4 人、女性 1 人) である。実験参加者のうち参加者 1~4 はオセロの経験が少ししかない人と、友人には負けないほどの強さを持つ人である。一方、参加者 5 はスマートフォンアプリケーション「オセロクエスト」で 5 万人中 1000 位程度のランクである。

実験の結果を表 2 に示す。この表で勝敗は AI から見て勝ったか負けたかを表す。

表2 人間プレイヤーとの対戦実験の結果
(%値は勝率を示す)

参加者	提案 AI	選択手数・偶 AI (深さ 2)	選択手数・偶 AI (深さ 4)
1	2勝 2敗(50%)	1勝 3敗(25%)	3勝 1敗(75%)
2	3勝 1敗(75%)	2勝 2敗(50%)	3勝 1敗(75%)
3	3勝 1敗(75%)	3勝 1敗(75%)	4勝 0敗(100%)
参加者	提案 AI	選択手数・偶 AI (深さ 4)	選択手数・偶 AI (深さ 6)
4	2勝 2敗(50%)	1勝 2敗 1分 (33%)	4勝 0敗(100%)
5	1勝 3敗(25%)	0勝 4敗(25%)	1勝 3敗(25%)

6.4. 実験 2 の考察

参加者と選択手数・偶 AI の対戦結果から、参加者 1 は探索の深さ 2 の選択手数・偶 AI より強く、参加者 2 は深さ 2 と同程度の強さ、参加者 3 は深さ 2 より弱いことが分かる。一方、参加者 4 は深さ 4 と同程度の強さで、参加者 5 は深さ 6 より強いことが分かる。提案 AI は、比較的弱い参加者 1~3 にも勝たせることができています。このことから、弱い人間のプレイヤーに対して強さを動的に合わせられることが分かる。また、提案 AI は、比較的強い参加者 4 と 5 にも勝ちすぎることなく負けすぎることなく対戦できています。このことから、強い人間のプレイヤーに対しても強さを動的に合わせられることが分かる。提案 AI の実験参加者全員に対しての勝率は、11 勝 9 敗で 55% であった。この結果から、総合的にも、異なる強さを持つ人間のプレイヤーに対して強さを動的に合わせられることが分かった。

7. 議論

提案 AI の 4 つのオセロ AI と対戦させる実験から、提案 AI は、自身より弱い評価関数を用いた AI に対しても、自身と同等の強さの評価関数を用いた AI に対しても、強さを調整できた。そして、探索の深さを変えた場合も、ある程度強さを動的に合わせられていると考えられる。また、人間に対する実験では、異なる実力の参加者に対して、強さを動的に合わせられた。しかし、同等の評価関数を用いた AI で探索の深さをほとんど変わらなくしたものを相手にした場合、強さを調整する余裕がなく、動的に合わせられないことが分かった。これらのことから、閾値を動的に変更するという提案手法は、自身とほとんど強さが変わらない場合以外に適用できると考える。

改善点としては、以下の 3 つが考えられる。1 つ目は、より多くの対戦実験を通じて、4 節に述べた評価関数の精度を高めることである。精度を高めるためには、より多くの評価関数を用いた AI と対戦させる必要がある。本研究に用いた評価関数以外にも、囲い具合や、パターンに基づく評価をしたものなどが考えられる。2 つ目は、隅以外で閾値を動的に合わせられるようにすることである。本研究では、隅にある石の個数に応じて動的に閾値を変更していた。これに加えて、隅の 1 つ斜め内側にある石の個数に応じて閾値を動的に合わせることによって、より強さを動的に合わせることができると考えられる。

3 つ目は、提案 AI の深さをより深くすることである。本研究の実験では、自身とほとんど強さが変わらない AI に対しては、強さを動的に合わせる余裕が無かった。提案 AI の深さをより深くすることで、深い探索を行う相手の AI に対して強さを動的に合わせることが考えられる。

8. おわりに

本研究では、相手のプレイヤーの技能に合わせて弱く指すオセロ AI を開発した。そのために仲道らが提案した将棋 AI の評価関数を改良してオセロ AI に適用した。具体的には、評価値を変換する閾値を局面に応じて動的に変更する改良を行った。AI との対戦実験、人間との対戦実験により、提案 AI が相手のプレイヤーの強さに動的に合わせられることを示した。

今後の課題としては、以下の 3 つが考えられる。1 つ目は、本研究で作成した評価関数と異なる、さらに複数の評価関数を用いた AI と対戦させることである。本研究では、評価関数として、その局面での石数、取っている隅の数、その局面での選択可能な手の数などの基準を用いて作成した。しかし、評価関数の基準の設け方はこの限りではない。このため、より多くの評価関数を用いた AI と対戦させることで、提案 AI の精度を高める。2 つ目は、より強いプレイヤーに対しても同様の結果が得られるかを検証することである。本研究の実験の中で最も強い参加者はスマートフォンアプリケーション「オセロクエスト」で 5 万人中 1000 位程度のランクであった。提案 AI の精度を高めた上で、より強い人間のプレイヤーと対戦させたときにどうなるかを検証する。3 つ目は、不自然でなく指すことができているかを検証することである。仲道らの研究では、人間から見て明らかな悪手を指していないかを検証するために、AI が指すたびに良い手、悪い手、そうでない手とラベルを付けさせることにより、評価値上の悪手がプレイヤーに見つけられる割合を調べた。本研究の人間に対する実験では、ほとんどの参加者が初心者のため、AI が打った手が良いか悪いかを判断することができなかった。このため、より強くオセロに精通しているプレイヤーに対して不自然でなく指すことができているかを実験し、提案 AI の精度を検証する必要がある。

文 献

- [1] 仲道隆史, 伊藤毅志, "プレイヤー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, vol. 57, no. 11, pp. 2426-2435, 2016.
- [2] D. Silver et al., "Mastering the game of Go with deep neural networks and tree search," *Nature*, vol. 529, no. 7587, pp. 484-489, 2016.
- [3] 関栄二, 三輪誠, 鶴岡慶雅, 近山隆, "将棋におけるモンテカルロ木探索の特性の解明," ゲームプログラミングワークショップ論文集, pp. 68-75, 2012.
- [4] 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, vol. 58, no. 3, pp. 167-173, 2013.