

二人零和有限不確定完全情報ゲームに対する木を用いた 解析手法に関する研究

有我一希^{†1} 小泉康一^{†2} 大槻正伸^{†2}

概要: 本研究では途中の盤面がプレイヤーの意思によらず変化する可能性があるようなボードゲームのうち、二人零和有限不確定完全情報ゲームにおける解析法として、ゲームの木を用いた手法がどれくらい効果的なのかを調査する。そのためにまず適切なプレイ長さのゲームを自作し、そのゲームに対して対人戦用のアルゴリズムを開発する。そして、複数人と複数回対戦し、勝率が5割を超えるアルゴリズムを開発することを研究の1つの目的とする。今回開発した木のカットアルゴリズムを用いると、三手先読みを行うことで実行速度を損なわずに対人戦において5割を超える勝率で勝利した。よって、二人零和有限不確定完全情報ゲームに対して木を用いた解析はある程度有効であることが確認できた。さらに、五手先読みのアルゴリズムやモンテカルロ法を用いたアルゴリズムを開発し、開発したアルゴリズム同士で対戦させた。その結果、最も強いものは五手先読みのアルゴリズムであったが、実用的な実行速度で最も強いものは三手先読みのアルゴリズムであった。それらの検証から、今回作成した探索カットアルゴリズムは改良の余地があるといえる。

キーワード: ゲームの木, コンピュータアルゴリズム, 二人零和有限確定完全情報ゲーム

Study on Analytical Methods Using Game Tree for a Two-person Zero-sum Uncertainty Complete Information Game

ITSUKI ARIGA^{†1} KOICHI KOIZUMI^{†2} MASANOBU OHTSUKI^{†2}

Abstract: In this study, as an analysis method for a two-person zero-sum *uncertainty* complete information game, we investigate whether the method using the game tree is effective or not. For this purpose, we make such a game, and develop the algorithm for this game against the human. One of the purposes of the study is to develop a more than fifty percent winning algorithm by playing multiple times with multiple people. Using the cut algorithm of the game tree, our algorithm could win over 50% against some people. Therefore, the analysis using the tree for a two-person zero-sum uncertainty complete information game is effective. In addition, we developed algorithms prefetching deeper layer and Monte Carlo methods, and made developed algorithms fight each other. As a result, the strongest one was prefetching five depth algorithm, but the one with the most practical execution speed was the three depth algorithm. Since these verifications, our search-cut algorithm can be better.

Keywords: Game tree, Computer algorithm, Two-person zero-sum certainty complete information game

1. はじめに

近年、チェスや囲碁のような二人零和有限確定完全情報ゲームにおけるAI開発が著しく発展し、人間では勝つことができなくなるほど強くなっている。そのため、全ての二人零和有限確定完全情報ゲームにおいてAIが人間より強くなるのは、時間の問題だと考えられる。

そこで本研究では二人零和有限「確定」完全情報ゲームの確定の部分の不確定にした二人零和有限「不確定」完全情報ゲームを自作し、そのゲームにおける強いアルゴリズムの開発を行うことにする。

研究の具体的な内容としては、まず盤面は公開されているが、その進行に乱数を用いるボードゲームを自作し、そのゲームに対して対人戦を目的としたアルゴリズムを開発

する。最終的に複数人と複数回対戦し、勝率が5割を超えるアルゴリズムを開発することを研究の1つの目的とする。また、開発したアルゴリズムと、別の方式を用いて開発したアルゴリズム等と戦わせることにより、この手のゲームにおけるゲーム木を用いた手法の有用性や開発したアルゴリズムの強さについて考察する。

2. ゲームの概要

今回自作したゲームは、まず図1のような中央に十字の壁を設けた7×7の盤面上に、白黒それぞれ5つの駒が配置され、これらを2人のプレイヤーが1つずつ交互に動かして、自分の駒を、後述する勝利条件の形に先に配置した方が勝利するというものである。白(黒)プレイヤーは白(黒)の駒のみを動かせる。初期盤面は対称な駒配置にな

†1 福島工業高等専門学校 電気工学科
Electrical Engineering,
National Institute of Technology, Fukushima College
†2 福島工業高等専門学校 電気電子システム工学科
Electrical and Electronic System Engineering,
National Institute of Technology, Fukushima College

っている。

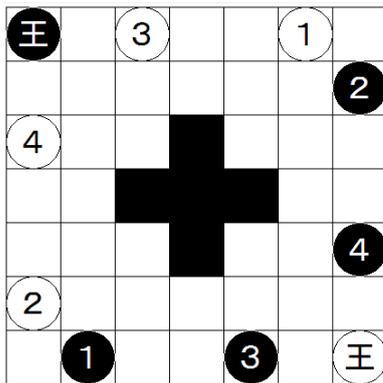


図1 自作ゲームの盤面

Figure 1 The board of our game.

手番となったプレイヤーは、動かしたい自分の色の駒を選択し移動したい方向を指定した後6面サイコロを振り、出目の数だけその方向に移動する。ただし移動により他の駒や盤面の壁に相当する部分へぶつかってしまう場合は、それらの直前で止まることとする。この移動方法のため、ゲームの進行にはランダム性が出てくることとなる。動きの一例として、図1の盤面において黒のプレイヤーが4の駒を左上に動かそうとしたときの様子を図2に示す。ここで、サイコロを振り5が出たとする。このとき4マス以上進もうとした時に白の3の駒にぶつかることとなり、直前の3マス目で止まる。なお、矢印と灰色の球体はそれぞれ移動した方向と4の駒が元いた位置を示している。したがってこの例の場合、1または2が出た場合はその分だけ移動するが、3、4、5、6の目が出た場合はどれが出たとしても3マス分のみ移動する。

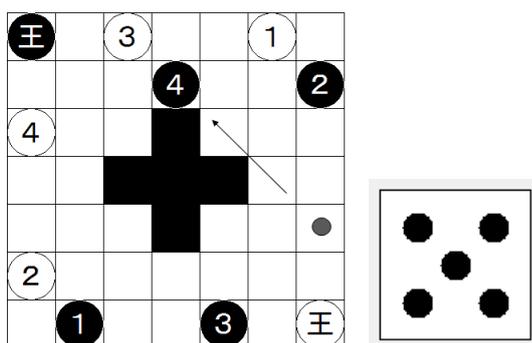


図2 移動後の盤面とサイコロの出目

Figure 2 The board after moving the piece "4".

本ゲームの勝利条件は、「王を含んだ3個以上の自分の駒を縦や横、斜めのいずれかの直線状に隙間なく並べる形を作る」こととした。

また、先手有利をなくすため、先手一手目は比較的強い手である王の斜め移動を禁止した。

このゲームにおける可能な盤面の数は 1.56×10^{13} である。このような可能な盤面の数は、将棋では 10^{60} 以上 10^{70} 未満であると厳密に求められている [1]。厳密ではないが、1つの上界として盤面の数が求められているゲームとしては、オセロ 10^{60} 未満、チェス 10^{120} 未満、囲碁 10^{360} 未満と言われているため、本ゲームは比較的簡単なものとなっている [2]。

3. アルゴリズムの概要

最初に作成した次に指すべき一手を探索するアルゴリズムでは、三手先までの全ての手を考えて、考えた盤面各々について盤面評価をした。

二人零和有限確定完全情報ゲームをゲームの木による解析を行う場合は、数手先を読んで次の手を選ぶときに用いる基準として盤面評価を行うことが多い。ここではミニマックス法 [3] を用いることを一例として取り上げる。ミニマックス法は、自分は自分にとって最良の手を選び、相手は自分にとって最悪の手を選ぶという考え方である。そのため、ゲームの木を用いたミニマックス法では、自分の盤面からのびる手を選ぶ場合には最良の手を選び、相手の盤面からのびる手を選ぶ場合には最悪の手を選ぶことで盤面評価値をゲームの木の根に返している。今回の不確定ゲームではミニマックス法を単純に導入することができない。その理由を示すために図3に二人零和有限確定完全情報ゲームの木と図4に今回の不確定ゲームの簡単なゲームの木を示す。図3にある楕円（節点）はそれぞれの盤面を示していて、楕円からのびる辺（枝）は選べる手を示している。図4は図3に加え、プレイヤーが手を選んだ後にランダムに手が分岐する黒点を加えられている。

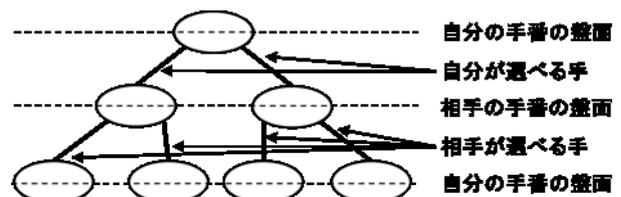


図3 二人零和有限確定完全情報ゲームの木

Figure 3 The game tree for two people zero sum definitive complete information game.

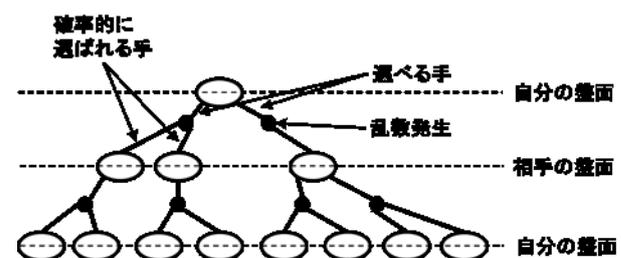


図4 本研究のゲームの木

Figure 4 The game tree for our game.

図4は図3と違って、プレイヤーの選択によらずサイコロの出目によって手が分岐する黒点がある。この存在が数手先の盤面評価と、単純なミニマックス法の導入を困難にする一因である。

そこで考案した次の手探索アルゴリズムでは、黒い点からのびる枝の先の盤面評価値を確率を考慮して足し合わせることににより、あたかも一本の枝と考え、その後ミニマックス法と同様の手法をとるようにした。しかしながらこのゲームにおいては、プレイヤーが同じ手を選んだとしても例えば出目が1ならば強いが、6だと非常に弱くなる、ということも多い。そのため、枝を一本にまとめたからと言って、従来通りの方法をそのまま用いては強いアルゴリズムにならない可能性が高い。

よって最終的なアルゴリズムでは盤面評価の算出方法を改良した。ある盤面から二手先全ての先読みを行ったと仮定したゲームの木を用いて、改良したアルゴリズムの動きの概略図を図5に示す。

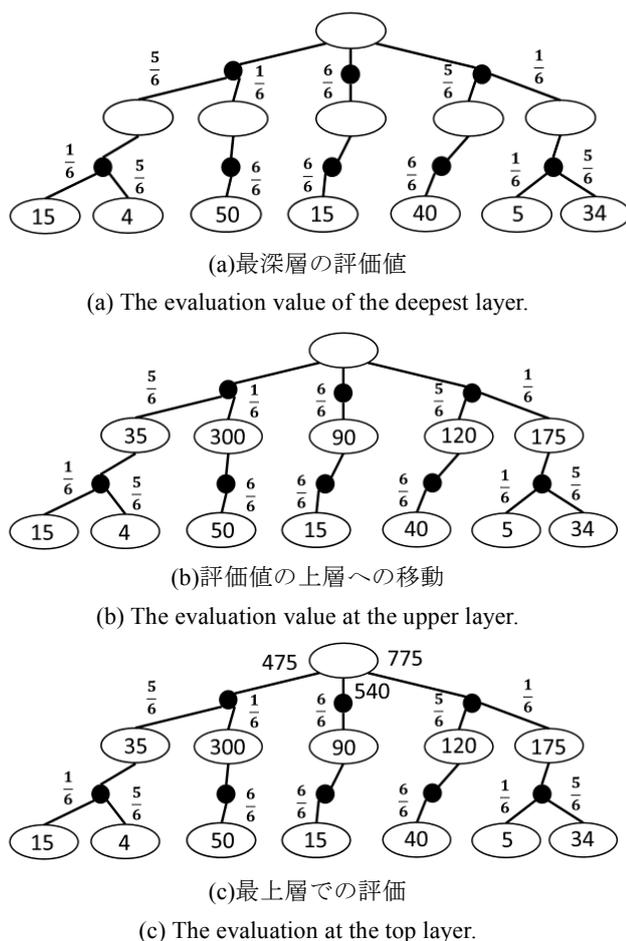


図5 改良したアルゴリズムの動き

Figure 5 Our evaluation of the improved algorithm.

図5の葉の中の値はその盤面の評価値を示しており、辺の横に書かれている分数はその手を選ばれる発生確率を示し

ている。値を親に返す際は評価値と発生確率の分子を掛け合わせて足し合わせることで評価する。この値は期待値の6倍になる。内部ノードでも同様にして親の値を計算する。根では次の手を選ぶために期待値を出すことはしないで、各枝から来た値を比較し、最も高い値の枝に対応する手を次の手として選ぶ。

改良したアルゴリズムではミニマックス法と異なり、深読みした葉に対応する盤面の、確率を考慮した全ての平均評価値を返すことにより、どのようにサイコロの目が出ても比較的強くなるようにしている。しかし、この方法で全ての可能な手を見ていくと最大で一千万以上の盤面評価をしなければならず、極めて多くの時間を費やさねばならなくなる。

したがって、今回の改良アルゴリズムでは、例えば三手先を読むとき、三手先すべての盤面評価を始める前にまず一手先のみすべての盤面評価をし、そこで基準となる評価値を超えた手のみさらに三手先まで読むという探索カットを行い、木の探索時間の短縮を図る。

4. 盤面評価

葉の盤面を評価するための盤面評価法では大きく分けて以下の3つの項目について採点する。なお、括弧内の数値はその評価項目でつけられる点数の範囲である。

- ① 自分のリーチの有無 (0~50)
- ② 王と他の駒の距離 (1~12)
- ③ 自駒が動くことで発生する

相手が次の手で勝利するリーチの有無 (-200 or 0)

これらの項目各々に得点を設けてコンピュータが有利な場合は加点、不利な場合は減点している。例えば、白視点で見た時に図1の初期盤面であれば左上部に3と4の駒の弱いリーチがあるため5程度、下に示した図6であれば王と1の強いリーチがあるため50程度の評価値になる。

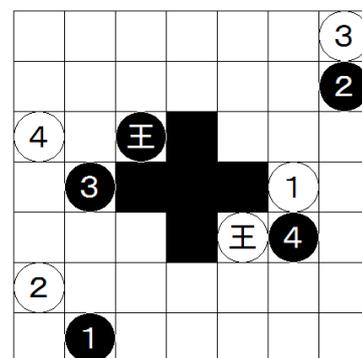


図6 ゲーム終盤の盤面の例

Figure 6 An example of the board of the endgame.

先程説明した盤面評価値は1つの盤面に対して計算される。ゲーム序盤であれば、互いにほぼ同等な条件であるた

め点数は 100 程度が、優位であれば、1000 程度の評価値が付く。プレイヤーが次の手でゲームに勝つ盤面であれば、評価値は -2000 程度になる。

さらに上の 3 つの評価法の他に、現状の盤面で明らかに勝負がつく手がある場合は先読みをせずに動かす手を選ぶという「明白な手」を取り入れている。白側から見た時に明白な手が決まる盤面を図 7 と図 8 に示す。

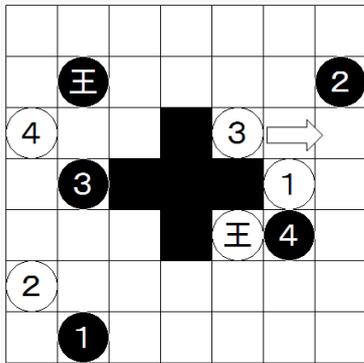


図 7 勝つために明白な手が指される盤面
Figure 7 The obvious move for win.

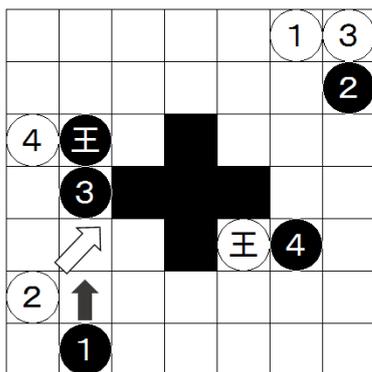


図 8 負けないために明白な手が指される盤面
Figure 8 The obvious move against the opponent.

図 7 の盤面は 3 の駒を右に動かすことにより、高い確率で勝利できるため先読みを行うことなくこの手を選ぶ。図 8 の盤面では 2 の駒を右上に動かすことで次の黒番で 3 つの駒が並ぶことによる敗北を防ぐことができる。

以上の盤面評価法を取り入れたアルゴリズムを typeA とする。

また、typeA に追加して相手の駒の位置を盤面評価に取り入れた評価法も別に作成した。これを typeB とする。

今回特に明記しない場合には、木の先読みの深さは 3 とする。

typeA

- ・基本的に自分の駒の配置のみを評価する。
- ・明白な手を選ぶ機能が搭載されている。

typeB

- ・常に互いの駒の配置を評価する。
- ・明白な手を選ぶ機能が搭載されている。

type(A+B)

- ・盤面によって typeA か typeB のどちらかの評価を行う。

5. 対戦結果

今回の研究に用いたコンピュータの性能を以下に示す。アルゴリズム同士で戦わせる場合は、1 台のコンピュータ上で同時にプログラムを動作させる。

プロセッサ：Intel®Core™i5-4590 CPU
動作周波数：3.3GHz
メモリ：4GB
RAM 速度：1600MHz

5.1 アルゴリズム対人間

複数人との対戦結果を表 1 に示す。比較のため、各アルゴリズムにおいて先読みの深さ一手読み版と三手先読み版それぞれと対戦していただいた。5 人がそれぞれ 4 回ずつ対戦し、計 20 戦の平均勝率を表している。表中の type(A+B) とは盤面の状況に応じて typeA と typeB の盤面評価のどちらかを行うアルゴリズムである。

表 1 各アルゴリズムの 20 回対戦の勝率 [%]

Table 1 The winning percentages in 20 battles
for each algorithm.[%]

	typeA	typeB	type(A+B)
一手読み	50	50	55
三手読み	60	40	55

5.2 アルゴリズム対アルゴリズム

追加で五手先読み版やモンテカルロ法^[4]を用いたアルゴリズムを作成した。それぞれのアルゴリズムの特徴と各アルゴリズム対三手先読み版 typeA との対戦の様子を説明する。

5.2.1 typeB 対 typeA

typeB の戦略では盤面上の相手の駒の配置を考慮しているため、序盤から相手の動きに合わせた動きになっている。そのため、自分の駒の配置のみを考慮する typeA との対戦では多くの場合で後手に回ることが多く、その結果猛攻を防ぎきれずに負けるような展開も多かった。数千回の対戦では 6 割程度で typeA に負けていた。

5.2.2 type(A+B) 対 typeA

type(A+B)は、負けているときには typeB の戦略をとり、

それ以外の時は typeA の戦略をとっている。このような評価の変え方では、負け始めた時に typeB と同じ様な展開となり、当然後手に回り続けた時には typeA に勝つことは難しかった。5000 回程度の typeA との対戦では typeB と同じように 6 割程度で負けていた。

5.2.3 五手先読み対 typeA

五手先読みのアルゴリズムは三手先読みのアルゴリズムを拡張し先読み深さを変えたものである。そのため、評価関数等はそのままである。また、木のカットは、まず一手先を読んでカットし、次に残った手を三手先まで読んでカットして、最後に残った手を五手先まで読むようにしている。13 戦ほど戦わせた結果としては五手先読みが 10 勝となり、より深く読んでいる五手先読みの方が三手先読みよりも多く勝利していた。ただし、一手を指すまでの時間としては、三手先読みが 10 数秒であるのに対し、五手先読みは 2 時間以上かかっていた。

5.2.4 モンテカルロ法対 typeA

モンテカルロ法を用いたアルゴリズムは、一手目のみ決まった手を指し、その後は互いにランダムに指してゲームが終わらせ、勝利回数を記録しておく。これを可能な全ての手について同回数行い、勝利数が最も多かった手を一手目として選んで指すということをしているアルゴリズムである。三手先読みより 50 倍程度の実行時間がかかる試行回数のモンテカルロ法を用いたアルゴリズムを三手先読みの typeA と戦わせると、モンテカルロ法の勝率は 3% ほどであった。

6. 対戦結果の考察

6.1 本ゲームの性質

アルゴリズム対人間やアルゴリズム同士、人間同士の対戦を見て体感した本ゲームの特徴について以下に示す。

まず、攻守の配分について述べれば、ほぼ間違いなく攻めに重点を置く方が強い。なぜなら、最初から相手の動きを予想して駒を動かしたとしても駒の動きにはランダム性があるため、手が進むごとに考慮した動きではない動きになっていくからだ。しかしながら、攻めにばかり重点を置いていても、相手の方がサイコロの出目がよく、負ける可能性も大いに存在するため、全く守りを考えないことは問題がある。そのため、このゲームに対する攻守の配分としては、基本的には攻めについて考え、相手の勝利確定が近づいて来る前兆が見受けられたならば、取り返しがつかなくなる程度に防ぐことが好ましいと考えられる。

次に先手後手に対する優位であるが、差はあまり感じられなかった。今回作成したゲームはまだプレイし尽くされておらず、様々な戦略の登場の可能性があるため断言はできないが、現状行われた対戦を見る限りでは、極めて高い確率でどちらかが勝つと言ったような手は見受けられなかった。さらに、本ゲームにおいてはサイコロの目によって

は先手が 1 手分損をして、後手があたかも先手であるかのように入れ替わることも珍しくない。そのため、現段階では先手と後手の間に有利不利の大きな差はないと考えられる。

6.2 アルゴリズム対人間

今回、アルゴリズム typeA を用いて目的の 1 つである勝率 5 割以上を満たすことができた。この結果から、ゲームの木を解析に利用し、さらに木のカットを用いたアルゴリズムはある程度有効であるといえる。

このアルゴリズムをさらに強化するのであれば、別の評価関数を複数用意してそのうち 1 つを選ぶなどの方法を取り、様々な戦略に対応できるような対策を行えばよいと考えられるが、今回試しに type (A+B) で対戦したところ大した強化にはならなかった。したがって、より良い組み合わせ法を考案する必要があると考える。

6.3 アルゴリズム対アルゴリズム

6.3.1 typeB 対 typeA

対戦結果の考察としては、このゲームにおいては攻めを主軸にして防がなければ負けるようなときのみ守るような戦略が有効であると考えられ、その考えにより近かった typeA に軍配が上がったものと考えられる。しかし、時には早めの段階から相手の駒の動きを考慮した手を指していたがために守りが堅牢となり、逆転して勝った対戦もあった。そのため、どんな相手にも勝つことができるような改良をするのであれば、typeB の評価関数の一部を修正し、導入することも悪いことではないと考えた。

6.3.2 type(A+B) 対 typeA

対戦結果としてはおおむね typeB と同じようになってしまった。しかしながら、これは 6.1 で述べたように合成の方法の悪さが引き起こした結果であり、新たな合成方法を考案する必要があると考えられる。例えば、方向性は異なるが同程度の強さを持つ 3 つの評価があったとして、これらを数手毎に切り替えるような合成方法や、全ての方法で評価して多くの評価で採用された手を選ぶ方法、もしくはそれぞれに選ばれた 3 つの手からランダムに 1 つを選ぶ方法などが考えられる。どの方法で合成するにせよ今回作成したアルゴリズムは似通った盤面評価のみで作成されているため、実際に試すことはできなかった。

6.3.3 五手先読み対 typeA

この結果だけ対戦結果の数が著しく少ない理由は、五手先読みの実行時間にある。五手先読みは一手指すのにおよそ 2 時間かかり、三手先読みの 500 倍程度かかっている残念ながら決して実用的なアルゴリズムだとは言えない。しかし、カットなしであれば、実行時間は理論上三手先読みの 2 万倍となるはずであり、そういった意味でカット自体は効果が極めて大きいと考えられる。

以上のことから、五手先読み、さらに言えば五手以上の先読みを行うアルゴリズムを開発するのであれば、より有

効だと思われる木のカットを考案する必要があり、そのためにはゲーム自体の性質に対してより深い考察が必須であると考えられる。また、対戦データを多く集めることにより深さに対する強さの変化を見て、ある程度の深さで強さが飽和するのであれば、それ以上の先読みは行わない等の処置も必要だと考えられる。

6.3.4 モンテカルロ法対 typeA

モンテカルロ法は、今回開発した typeA などの盤面評価方式に近い要素があったために比較対象として開発した。近い要素とは全ての手について考慮している点であり、モンテカルロ法ではある程度の回数の対戦結果を考えることにより、統計学的推定のようにその手の強さを推し量っている。モンテカルロ法の方が今回開発したアルゴリズムより優れている点は制度を向上させたときに増加する実行時間の長さにある。例えば、盤面評価方式であれば一手先読みから三手先読み、五手先読み、七手先読みと拡張させていくと、カットを考慮しなければ理論上 4 万倍程度ずつ時間が伸びていく。しかし、モンテカルロ法であれば、実行時間が各手を見る回数を増やした分だけしか増えないのに対し、移動に使われるサイコロの目が出る確率も収束していき精度があがる。さらに、対戦が終盤になり見る手の数が減れば、一手にかかる実行時間も大きく削減される。

しかし、対戦結果を見れば明らかな通りモンテカルロ法のアルゴリズムは極めて弱かった。その理由は、ランダムに手を選び対戦を終わらせたことにあると考えられる。例え 2 万回、3 万回と結果を集めてもその多くが実際の対戦において使われないような手であったならば見る意味がない。そのため、時間が掛かるものの深くまで先を予測することができるようなゲームにおいてモンテカルロ法は明確な盤面評価方式を超えられなかったのではないかと考えられる。

ただし、拮抗する強さを持ついくつかの手の中からランダムに 1 つを選ぶような場合であれば、完全にランダムにするよりはモンテカルロ法の方が多少の有用性があると考えられる。

6.4 アルゴリズムの改良に対する提案

6.3 までに述べたことを統合して考えたアルゴリズムについて述べる。

考えたアルゴリズムの動作は主に複数の方法を用い、それぞれの方法で次の手となる候補を選ぶ。その後、候補となったそれぞれの手に関してモンテカルロ法を用いて次の手を確定させる。このアルゴリズムに用いる複数の方法の 1 つには typeA を強化したものを使うことを考えている。ここでいう typeA の強化とは現在のアルゴリズムにはない typeB 等の要素を加えることを意味している。

7. まとめ

今回の研究を行うにあたって研究に適したゲームを作成

することから始めた。作成したゲームは先手や後手によって有利不利の差があまりなく、プレイ長さや複雑さの面からみると将棋やオセロといった有名なゲームよりも短い手数で終わるものの、戦略によって盤面の進行、展開が変わるには十分な複雑さだと言える。

今回、開発したアルゴリズムにおいてゲームの木を用いて三手先読みを行い、進行上不確定さがあるようなゲームにおいても勝率 5 割を超えることのできるゲームの木探索を行うことができた。目標の対戦結果を出せ、本研究は目的の 1 つは達することができたといえる。しかし、この結果はあくまで本ゲームのルールを覚え対戦ができる段階になったレベルの人との対戦結果であるため、仮にアルゴリズムの行動パターンを解析し、それに適する手を選ぶことを瞬時にできるような極めて強い人間プレイヤーとの場合の対戦は想定していない。よって、いかなる場合においても効果的となるように、次の 1 手を探索する方法を複数個開発して効果的になるように組み合わせることで、さらなる強化をするべきだと考える。

謝辞 対戦実験において協力いただいた福島工業高等専門学校電気工学科 15 人の皆様に謹んで感謝の意を表します。

参考文献

- [1] 篠田正人, “将棋における実現可能局面数について,” 情報処理学会 IPSJ Symposium Series, Vol.2008, No.11, 2008.
- [2] “ゲーム理論,” <http://agent.inf.kyushu-u.ac.jp/~yokoo/lecture/game-theory-5.pdf>, (参照 2017-01-30).
- [3] 佐々木宏夫, “入門ゲーム理論 戦略的思考の科学,” 日本評論社, 2003.
- [4] 小谷善行, “ゲーム計算メカニズム ー将棋・囲碁・オセロ・チェスのプログラムはどう動くー,” コロナ社, 2010.