Vol.2012-GI-28 No.6
2012/7/13

大貧民において不完全情報性がモンテカルロ法によるプレイヤに与える影響の調査

地曳 隆将[1,a)]　松崎 公紀[2,b)]

概要：近年，ゲームの研究において，乱数によるシミュレーション（プレイアウト）を複数回行うことにより近似解を算出するモンテカルロ法を，プレイヤプログラムに適用する研究が注目されている．大貧民は不完全情報ゲームであり相手プレイヤの手札を知ることはできない．この不完全情報性がプレイアウトの精度を下げる要因のひとつとなる．そこで本研究では次の一手問題を用いた実験を行い，モンテカルロ法およびモンテカルロ木探索の性質と，手札情報の差によって各手役の評価値がどのように変化するか調査した．その結果，モンテカルロ法によるプレイヤは手札公開枚数が多くなるにつれてより正しい評価値を与え，さらに序盤と終盤で最善手の推定に有用な情報が異なることを確認した．

キーワード：大貧民，不完全情報ゲーム，モンテカルロ法，モンテカルロ木探索

Study on Effect of Imperfect Information Nature on Monte Carlo Daihinmin Players

Zibiki Takamasa[1,a)]　Matsuzaki Kiminori[2,b)]

Abstract: In recent game studies, Monte-Carlo-based algorithms, in which we compute an approximate solution by numerous simulations (called playouts), have been widely studied. Daihinmin is an imperfect-information game and in Daihinmin we cannot see the other players' hands, which makes the playouts more imprecise. In this paper, we examine properties of simple Monte-Carlo algorithm and Monte-Carlo tree search by using next-move problems, especially in terms of the information of the other players' hands. We have confirmed the following two: the Monte-Carlo algorithm gives more correct values as the number of unknown cards decreases; the information that is useful for finding the best move differs from the opening stage to the closing stage.

Keywords: Daihinmin, Imperfect information game, Monte-Carlo Method, Monte-Carlo Tree Search

1. はじめに

近年，ゲームの研究において，乱数によるシミュレーションを複数回行うことにより近似解を算出するモンテカルロ法の適用が注目されている．特に，モンテカルロ法を木探索と組み合わせた手法であるモンテカルロ木探索 [4] に関する研究が多く行われている．たとえば囲碁においては，モンテカルロ木探索をプレイヤに適用することでプレイヤの棋力がアマチュア初段レベルまで向上した [14]．将棋においては，モンテカルロ木探索だけでは強力なプレイヤを作り出すことは難しいが，従来の手法と組み合わせて部分的にモンテカルロ木探索を用いることでプレイヤの強化が可能であることが示唆されている [9]．またこの他のゲームでも，オセロやパックマンなどでモンテカルロ木探索を用いることでプレイヤプログラムが強化されることが

[1] 高知工科大学大学院工学研究科基盤工学専攻
Graduate School of Engineering, Kochi University of Technology
[2] 高知工科大学情報学群
School of Information, Kochi University of Technology
[a)] 165064w@gs.kochi-tech.ac.jp
[b)] matsuzaki.kiminori@kochi-tech.ac.jp

明らかになっている [3], [6].

ゲームにおいてモンテカルロ法およびモンテカルロ木探索では，乱数を用いて終局までプレイを行うことを繰り返す．この乱数による終局までのプレイはプレイアウトと呼ばれる．モンテカルロ法やモンテカルロ木探索によるプレイヤを強くするためには，プレイアウトの回数を多くするだけでなく，プレイアウトで得られる値の精度を高めることが必要である．

本研究では，多人数・不完全情報ゲームである，トランプゲームの大貧民を研究の対象とする．大貧民においても，モンテカルロ法を適用することで強いプレイヤを得ることができることが示されている．実際，UEC コンピュータ大貧民大会 [2] の第 4 回，第 5 回，第 6 回において，モンテカルロ法を適用したプレイヤが優勝している [5], [11], [12]. しかし，これらのプレイヤでは，（大会における計算時間のルールなどにより）プレイアウト回数が十分に多くとられていたとは言い難い．また，大貧民では相手プレイヤの持つ手札は分からないので，プレイアウトを行う前に相手プレイヤの持つ手札を仮想的に生成する必要がある．ここで生じる相手プレイヤの手札の差は，プレイアウトの精度を下げる要因のひとつとなる．

大貧民において相手プレイヤの手札を推定する手法に関して，これまでにも研究が行われている．須藤ら [12]，および，西野ら [8] は，相手プレイヤがそれまでに出した手札情報を用いて相手プレイヤの手札を推定する手法を提案している．これらの手法により相手プレイヤの手札を推定して生成することで，より強いモンテカルロ法によるプレイヤが得られることが報告されている．一方で，西野らは，相手プレイヤの手札を推定することの重要性について，手札の集合を同値類に分類することによる考察により，大貧民においては手札推定があまり重要でないと述べている [7].

そこで本研究では，大貧民に対するモンテカルロ法とモンテカルロ木探索によるプレイヤについて，その性質を調べるための実験と考察を行った．具体的には，プレイアウト回数に対する評価値の収束について，および，相手プレイヤの手札の情報があることのプレイアウトへの影響について調査した．これを実現するにあたり，大貧民における「次の一手」問題を選定し，それに対して実験を行う手法をとった．

本論文の貢献は，大きく次の 3 点である．

- モンテカルロ法によるプレイヤの性質を評価するため，大貧民のプレイの中から「次の一手」問題を 30 盤面選定した．これらの問題は，他の研究でも利用できるよう公開している．
- モンテカルロ法とモンテカルロ木探索において，その評価値の計算の収束について調査した．
- 大貧民の不完全情報性がモンテカルロ法に与える影響について，実験により評価した．

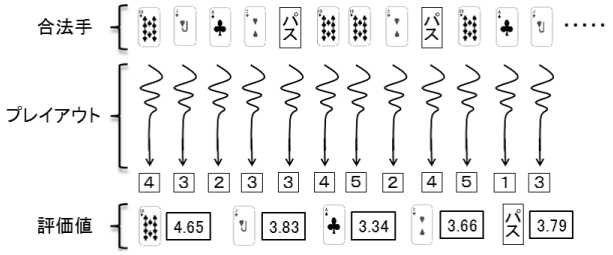

図 1 モンテカルロ法の動作
Fig. 1 Motion of Monte-Carlo Method.

本論文の構成を以下に示す．第 2 章では，モンテカルロ法とモンテカルロ木探索のアルゴリズムについて説明する．第 3 章では，本実験に用いる大貧民のルールおよび 2 種類の大貧民プレイヤの動作について説明する．第 4 章では，本実験に用いる次の一手問題について説明する．第 5 章では，モンテカルロ法による大貧民プレイヤの収束に関して調査する．第 6 章では，手札情報の差が大貧民プレイヤに与える影響を調査する．そして，第 7 章で本論文をまとめる．

2. モンテカルロ法とモンテカルロ木探索

2.1 モンテカルロ法

モンテカルロ法とは，乱数によるシミュレーションを複数回行うことで近似解を算出するアルゴリズムのことである．モンテカルロ法の特長のひとつに，解析的に解を算出しにくい問題を含む広範囲の問題に対して適用可能であることが挙げられる．一般に，モンテカルロ法の精度はシミュレーション回数の平方根に比例する [10].

以下では，ゲームに適用したモンテカルロ法について述べる．ゲームに適用したモンテカルロ法の動作を**図 1** に示す．ゲームに適用した場合のシミュレーションは，ある盤面のある合法手から始め，乱数によって終局までプレイすることに対応する．この乱数による終局までのプレイをプレイアウトと呼ぶ．プレイアウトの結果のそれぞれに点数をつけ，その点数の平均を求めることで評価値を計算することができる．

限られた回数のプレイアウトでより良い結果を得るためには，有望そうな手に対するプレイアウトを多くすることが有効である．どの手をプレイアウトすると最も高い期待値を得られるかは多腕バンディット問題と呼ばれ，これを効率良く解くアルゴリズムが複数提案されている．そのようなもののひとつである UCB1（Upper Confidence Bound）[1] では，手役 j の現在の評価値 $\overline{X_j}$，全体のプレイアウト回数 n，手役 j に対するプレイアウト回数 n_j，およびある定数 c に対して，次の式で示される UCB1 値

$$\overline{X_j} + c\sqrt{\frac{2\log n}{n_j}} \tag{1}$$

が最も大きいものに対してプレイアウトする．この定数 c

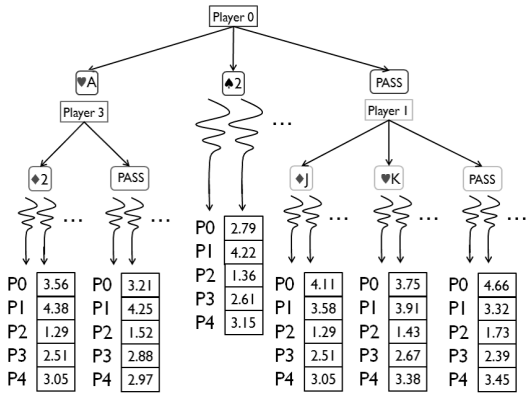

図 2 モンテカルロ木探索の動作
Fig. 2 Motion of Monte-Carlo Tree Search.

は，バランスパラメータと呼ばれる．

モンテカルロ法では，プレイアウト回数だけでなくプレイアウトの精度が結果に大きく影響する．すなわち，乱数によるプレイが実際で行われるプレイと大きく異なる場合には正しい結果に収束しない．さらに，探索する木の深さが 2 以上の場合には，プレイアウト回数をどんなに増やしても正しい結果に収束しない可能性がある [14].

2.2 モンテカルロ木探索

モンテカルロ木探索 [14] は，モンテカルロ法とゲーム木探索手法を組み合わせたアルゴリズムである．基本的な考え方は，プレイアウトによってある盤面のある合法手に対する評価値を計算し（モンテカルロ法），そこで得られた有望そうな手についてその先の探索をより深くする（ゲーム木探索手法）というものである．モンテカルロ法とは異なり，モンテカルロ木探索ではプレイアウト回数を十分に増やすと最善手が得られる [14].

モンテカルロ木探索の動作を**図 2** に示す．モンテカルロ木探索では，まず，それまでに得られている評価値を利用して木を探索する．ある葉ノードに到達すると，その葉ノードに対応する盤面から始めてプレイアウトを行う．ここで，ある葉ノードにおけるプレイアウト回数がある閾値に達した際には，そのノードを展開して，その盤面の 1 手先からプレイアウトするようにする．UCB1 などの手法を用いると，有望そうな手にはプレイアウトが多く実行され，そのようなプレイアウトが集中するところはより深く木を探索できることになる．

3. 実験に用いる大貧民プレイヤ

本章では，まず，実験に用いる大貧民のルールについて概説する．その後，本研究で実験に用いる，モンテカルロ法によるプレイヤとモンテカルロ木探索によるプレイヤの，2 つの大貧民プレイヤについて説明する．

3.1 大貧民のルール

本研究では，UEC コンピュータ大貧民大会の標準ルール [2] を用いた．そのうち本研究に大きく影響を与える重要なルールを以下に示す．

あがり時の制約 プレイヤの手札枚数が 0 になった状態をあがりと呼ぶ．あがり時にはどのような手役を提出してもよい．あがり時に出す手役の制約がないことにより，モンテカルロ法やモンテカルロ木探索においてプレイアウトを単純に実装することができる．

8 切り ランクが 8 のカードを含む手役が場に提出されると 8 切りが発生する．8 切りが発生すると場が流れ，8 切りを発生させたプレイヤが次の手役を出せる．ランクが 8 のカードのランク自体はそれほど強くないが，プレイにおいて適切に使用すると強力なカードにもなるため，モンテカルロ法とモンテカルロ木探索での差が生じる可能性がある．

しばり 場に同じスート（複数枚の手役のときには，全てが同じスート）の手役が連続して提出されるとしばりが発生する．しばりが発生すると，場が流れるまで，同じスート（の組み合わせ）の手役しか場に提出できなくなる．このルールにより，特定スーツのカードを持つかどうかに意味があり，相手手札の不完全情報性の一要素となる．

3.2 モンテカルロ法プレイヤ

本研究でのモンテカルロ法によるプレイヤは，すべての合法手に対して同じ指定回数のプレイアウトを実行する．プレイアウトの対象となる合法手には，（全く意味のない，場が新しい場合を除き）パスも含める．

プレイアウトでは，各プレイヤは以下のように動作する．

(1) 合法手のうち，それを出すことであがりとなるような手があれば，それを選択する．
(2) パス以外の合法手が存在する場合には，パス以外の合法手の中から等確率に選択する．
(3) 合法手がパスのみである場合には，パスを選択する．

上記のとおり，プレイアウト中では意図的なパスはしないことになる．

1 回のプレイアウトが終了すると，選択した手役に対して点数を割り当てる．このときに割り当てる点数は，大富豪であれば 5 点，富豪であれば 4 点，平民であれば 3 点，貧民であれば 2 点，大貧民であれば 1 点とした．それまでのプレイアウトで得られた点数の相加平均を，その手役の評価値と呼ぶ．

3.3 モンテカルロ木探索プレイヤ

モンテカルロ木探索によるプレイヤにおいて，プレイアウト中の動作および割り当てる点数はモンテカルロ法によるプレイヤと同じとする．木の探索においてどの子ノード

を選択するかは，式 1 の UCB1 値によって決定した．ここで，バランスパラメータ c は 4 とした．また，葉ノードのプレイアウト回数が 100 回を越えるときにそのノードを展開するものとした．

プレイアウトでは，ルートノードのプレイヤだけでなく全てのプレイヤに点数を割り当てる．葉ノードにおいては，割り当てられた点数の相加平均により，各プレイヤの評価値を求める．その他のノードについては，$\max^n$ アルゴリズムにより，各プレイヤがその子ノードの評価値のうちそのプレイヤにとって最大となるものをローカルに選択するようにした．例えば，図 2 でルートからハートの A を選んだ先のノードでは，その子ノード（いずれも葉ノードである）のうちプレイヤ P3 にとって評価値の高い $[P0 \rightarrow 3.21, P1 \rightarrow 4.25, P2 \rightarrow 1.52, P3 \rightarrow 2.88, P4 \rightarrow 2.97]$ がその評価値となる．

4. 大貧民における次の一手問題の作成

モンテカルロ法やモンテカルロ木探索の性質をより詳しく調べるため，本研究では次の一手問題を作成しそれを用いた．

次の一手問題とは，ある盤面においてプレイヤが次にどのような行動をするのが最良であるかを問う問題のことである．将棋や囲碁の分野では，人間向けのクイズとしてだけでなく，プレイヤの性能評価にも次の一手問題が用いられ成果を挙げている [9], [13]．著者の知る限り，大貧民における次の一手問題は存在しなかったため，コンピュータ大貧民のプレイの中から盤面を選択してそれを次の一手問題とした[*1]．

本研究で作成した次の一手問題は，以下の情報からなる．

- プレイヤに関する情報（次にプレイするプレイヤを `Player0` とする）
  - プレイヤが持つカードの集合
  - プレイヤがそのターンでプレイすることができるかどうか（パスしているか，すでにあがりである場合にはプレイできない）
- 場に関する情報
  - 最後に場に出されたカード（場が新しい場合には空とする）
  - 強さの順番（革命の有無）
  - しばりがある場合にはそのスート

広く研究に利用できるようにするため，相手プレイヤの持つカードについても明示的に持つようにしている[*2]．通常の大貧民のプレイの場合には，相手プレイヤのカードについては，カードの枚数のみ分かるようにした上で使われていないカードの集合を見せれば良い．

[*1] 本研究で用いた次の一手問題は，http://ipl.info.kochi-tech.ac.jp/daihinmin_database/ にて公開している．

[*2] この盤面に至るまでの手役の履歴については含んでいないため，相手プレイヤの手札を推定してプレイするようなプレイヤの研究にはこれでは使えない．

```
Order Normal
Lock
Last Group H5 D5
Player0 Yes C3 S7 H7 D7 H8 C8 SJ HQ DQ
Player1 Yes S3 S4 H4 S5 S6 S8 S9 S10 CQ HA D2
Player2 Yes H6 HJ CJ SQ SK HK
Player3 Yes C7 D8 H10 DK CA C2
Player4 Yes C9 DJ CK S2 H2
```

図 3 次の一手問題の例（序盤）
Fig. 3 A Next-move Problem (Opening Stage)

```
Order Normal
Lock
Last
Player0 Yes C6 S7 C7 D9 C2
Player1 Yes CQ CK SA
Player2 Yes D3 H10 HJ CJ
Player3 Yes C4 D4 D5 H6 H7 D8 S9 C9
Player4 No
```

図 4 次の一手問題の例（終盤）
Fig. 4 A Next-move Problem (Closing Stage)

以上の情報からなる次の一手問題をコンピュータ大貧民のプレイの中から 30 個選択して作成した．その内訳は，枚数を基準として序盤を 10 個，中盤を 10 個，終盤を 10 個とした．序盤・中盤・終盤を区別する枚数の基準は以下の通りとした．

序盤 まだ使用されていないカードが 36 枚以上
中盤 まだ使用されていないカードが 21 枚以上 35 枚以下
終盤 まだ使用されていないカードが 20 枚以下

作成した次の一手問題の例を**図 3** と**図 4** に示す．各プレイヤは 1 行で示されるが，その 2 項目目の `Yes` または `No` が，プレイヤがそのターンでプレイすることができるかどうかを表している．

5. 大貧民におけるモンテカルロ法の収束に関する調査

この章では，モンテカルロ法プレイヤとモンテカルロ木探索プレイヤを用いてプレイアウトを行った際の，各手役の評価値の推移を調査した．実験に使用する 30 問の次の一手問題は相手プレイヤの手札をすべてわかっている状態とした．本実験の目的は，モンテカルロ法プレイヤおよびモンテカルロ木探索プレイヤの評価値の収束について調査し，どの程度の回数のプレイアウトを行えばよいか明らかにすることである．

5.1 モンテカルロ法の収束

モンテカルロ法プレイヤの収束について調査するため，各手役に対して 10000 回のプレイアウトを行った．**図 5** は，序盤においてモンテカルロ法プレイヤを用いてプレイアウトを行ったときの各手役の評価値の推移の一例であ

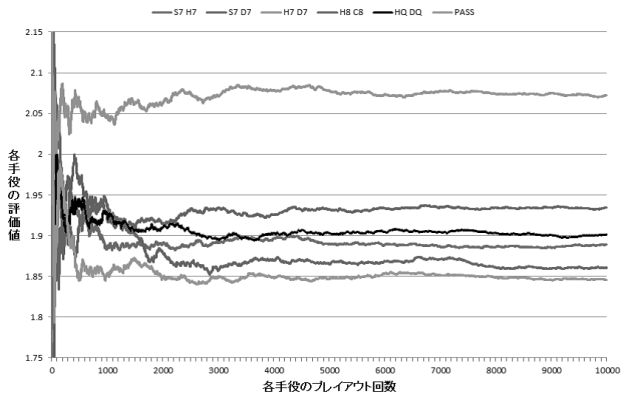

図 5 序盤におけるモンテカルロ法プレイヤの評価値の推移
Fig. 5 Transition of evaluation values of Monte-Carlo Method player in early stage.

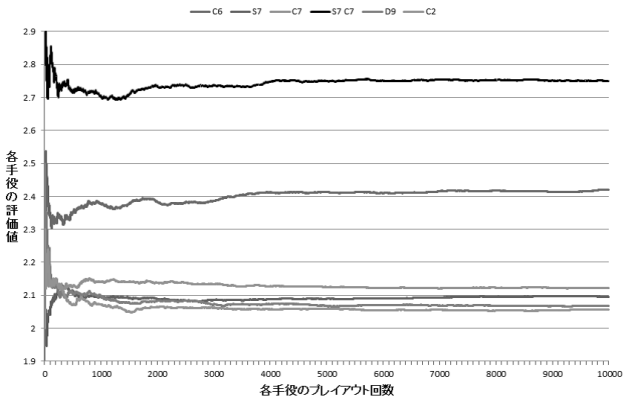

図 6 終盤におけるモンテカルロ法プレイヤの評価値の推移
Fig. 6 Transition of evaluation values of Monte-Carlo Method player in closing stage.

表 1 モンテカルロ法プレイヤの各盤面における最小値と最大値の差の幾何平均
Table 1 Geometric mean of difference between minimum and maximum values in each board of Monte-Carlo Method player.

測定範囲	序盤	中盤	終盤
100〜1000	0.1474	0.1097	0.0864
1001〜2000	0.0351	0.0266	0.0224
2001〜3000	0.0214	0.0132	0.0121
3001〜4000	0.0152	0.0101	0.0098
4001〜6000	0.0150	0.0111	0.0090
6001〜8000	0.0103	0.0071	0.0063
8001〜10000	0.0085	0.0058	0.0049

り，**図 6** は，終盤においてモンテカルロ法プレイヤを用いてプレイアウトを行ったときの各手役の評価値の推移の一例である．図 5 と図 6 を見ると，プレイアウト回数が増加するにつれて評価値の変化の幅が小さくなっていることがわかる．

表 1 は各手役の評価値の測定範囲をプレイアウト回数で区切り，その範囲内の最小値と最大値の差の幾何平均を示した図である．表 1 を見ると，モンテカルロ法プレイヤの評価値はプレイアウト回数が増加するにつれて変化の幅が小さくなっていることがわかる．また，序盤，中盤，終盤となるにつれて評価値の変化の幅が小さくなっていることがわかる．測定範囲が 100〜1000 の範囲での各手役の評価値は，序盤で 0.15，中盤で 0.11，中盤で 0.09 程度の変化の幅がある．それ以降の測定範囲を見ると，徐々に変化の幅が小さくなっているのがわかるが，3001〜4000，4001〜6000 の測定範囲の変化の幅は大差ないことがわかる．3001〜4000，4001〜6000 の測定範囲では評価値の変化の幅が 0.01 程度に収まっているため，評価値が収束しているのではないかと考える．このことから，モンテカルロ法プレイヤが各手役に対して行うプレイアウト回数は 4000 回程度が適当であると考える．

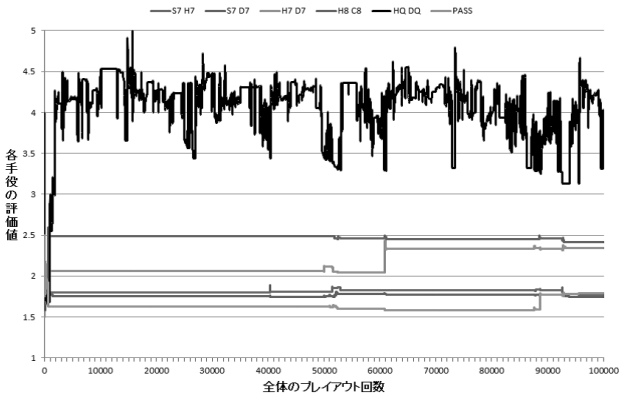

図 7 序盤におけるモンテカルロ木探索プレイヤの評価値の推移の一例
Fig. 7 Transition of evaluation values of Monte-Carlo Tree Search player in early stage.

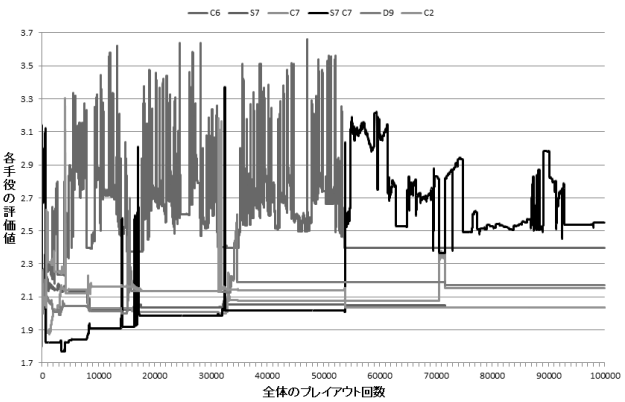

図 8 終盤におけるモンテカルロ木探索プレイヤの評価値の推移の一例
Fig. 8 Transition of evaluation values of Monte-Carlo Tree Search player in closing stage.

5.2 モンテカルロ木探索プレイヤの収束

モンテカルロ木探索プレイヤの収束について調査するため，合法手全体で 100000 回のプレイアウトを行った．**図 7** は，序盤においてモンテカルロ木探索プレイヤを用いてプレイアウトを行ったときの各手役の評価値の推移の一例で

表 2 モンテカルロ木探索プレイヤの各盤面における最善手の最小値と最大値の差の幾何平均

Table 2 Geometric mean of difference between minimum and maximum values of best hand in each board of Monte-Carlo Tree Search player.

測定範囲	序盤	中盤	終盤
100〜10000	1.5394	0.7868	0.5031
10001〜20000	0.3550	0.4027	0.3729
20001〜30000	0.4806	0.4373	0.4129
30001〜40000	0.4723	0.3703	0.4429
40001〜50000	0.4412	0.1111	0.3695
50001〜60000	0.3794	0.3164	0.2461
60001〜70000	0.3756	0.6482	0.3521
70001〜80000	0.4322	0.5886	0.4239
80001〜90000	0.4402	0.5886	0.4239
90001〜100000	0.5151	0.5518	0.1515

あり，**図 8** は，終盤においてモンテカルロ木探索プレイヤを用いてプレイアウトを行ったときの各手役の評価値の推移の一例である．図 7 と図 8 を見ると，各手役の評価値が大きく変動している箇所が多数あることがわかる．これは，プレイアウトをしている手役の葉ノードを展開したときに，その 1 つ下のプレイヤがベストな手役を提出したことでおこる現象であると考える．このことから，多人数ゲームにおいてモンテカルロ木探索を行った場合には，木の深さを 1 つ深くするだけで各手役の評価値が大きく変動することが確認される．

表 2 は最善手の評価値の測定範囲をプレイアウト回数で区切り，その範囲内の最小値と最大値の差の幾何平均を示した図である．なお，測定範囲内で評価値が全く変化していないものは無視して測定した．表 2 を見ると，モンテカルロ木探索プレイヤの評価値は，全体のプレイアウト回数が増加しても変化の幅があまり変わらないことがわかる．これは，図 7 や図 8 のように評価値が大きく変動しながら推移するためであると考える．このことから，モンテカルロ木探索プレイヤを用いて正しい評価値を算出するためには，膨大なプレイアウト回数が必要になることがわかる．

6. 不完全情報性が大貧民プレイヤに与える影響の調査

本実験では，30 問の次の一手問題をモンテカルロ法プレイヤとモンテカルロ木探索プレイヤに解かせる実験を行った．また，モンテカルロ法プレイヤの相手プレイヤの手札公開枚数を変更して実験を行った．

6.1 実験方法

各 30 盤面に対して 6 種類のシミュレーションを行い，そのときの各手役の評価値を調査した．6 種類のシミュレーションは以下の通りである．

MCM:全ランダム 相手プレイヤの全てのカードをランダムに割り当て，モンテカルロ法プレイヤでプレイアウトを行う．

MCM:強固定 各相手プレイヤの最強のカードを一枚固定し，その他のカードをランダムに割り当て，モンテカルロ法プレイヤでプレイアウトを行う．

MCM:弱固定 各相手プレイヤの最弱のカードを一枚固定し，その他のカードをランダムに割り当て，モンテカルロ法プレイヤでプレイアウトを行う．

MCM:強・弱固定 各相手プレイヤの最強と最弱のカードを一枚固定し，その他のカードをランダムに割り当て，モンテカルロ法プレイヤでプレイアウトを行う．

MCM:全固定 相手プレイヤの全てのカードを固定し，モンテカルロ法プレイヤでプレイアウトを行う．

MCT:全固定 相手プレイヤの全てのカードを固定し，モンテカルロ木探索プレイヤでプレイアウトを行う．

補足すると，全ランダムとは相手の手札が一枚もわかっていない (通常の大貧民と同じ) 状態であり，全固定とは完全情報ゲームになっている状態である．

本実験では，MCT:全固定は合計 100000 回のプレイアウトを行い，それ以外の 5 つのシミュレーションは各手役に対して 10000 回のプレイアウトを行った．

6.2 実験結果と考察

本実験では，MCT:全固定時の各手役の評価値を「解答の評価値」とし，MCT:全固定時に最も評価値が高い手役をその盤面における「最善手」とした．本実験では，MCT:全固定時の各手役の評価値とそれ以外の 5 つのシミュレーションの各手役の評価値の幾何平均誤差と，最善手の的中数について調査した．

各盤面における最善手の評価値の幾何平均誤差を**図 9** に示し，最善手の的中率を**表 3** に示す．表 3 中のスラッシュの左の数値は各シュミレーション中で最も評価値が高い手役が最善手と一致した数を表し，スラッシュの右の数値は各シミュレーション中で二番目に評価値が高い手役が最善手と一致した数を表している．また，各盤面における各手役の評価値と解答の評価値の幾何平均誤差を**図 10** に示す．各シミュレーションの評価値の詳細は付録に記載する．なお，図 9，図 10 および表 3 中のシミュレーション手法の表記は MCM:を省略している．

図 9 を見ると，序盤でははは全固定や強・弱固定よりも全ランダムの方が評価値の誤差が少ないことがわかる．このことから，序盤においては相手の手札が完全にわからない状態でも最善手の推定は可能であると考える．また，序盤と中盤では弱固定が最も評価値の誤差が小さいため，相手プレイヤの最弱のカードを予想することでプレイヤが強化できることがわかる．終盤では，弱固定が評価値の誤差が最も大きく，全固定が評価値の誤差が最も小さいことがわかる．また，全ランダムと強・弱固定の評価値に大差がな

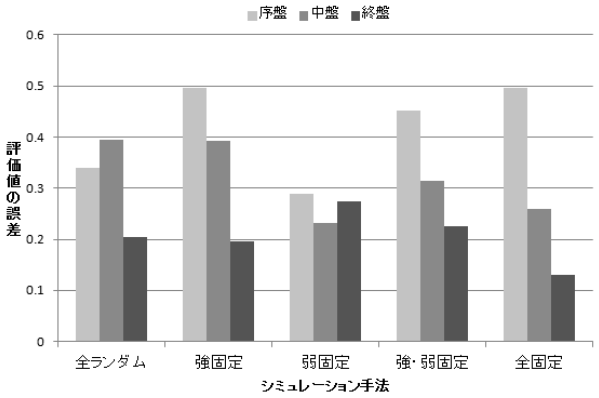

図 9 各盤面における最善手の評価値の幾何平均誤差
Fig. 9 Geometric mean of error evaluation value of best hand in each board.

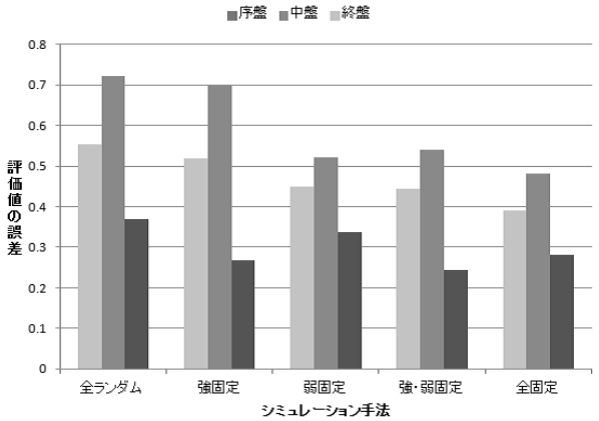

図 10 各盤面における各手役の評価値と解答の評価値の幾何平均誤差
Fig. 10 Geometric mean of error evaluation value of each hand in each board.

いため，終盤では相手プレイヤの手札全体を推定しなければ意味がないことがわかる．

表 3 を見ると，序盤では全固定の的中数が 4 であり，それ以外のシミュレーション手法の的中数が 3 である．また，中盤も序盤と比べて最善手の的中数に大差がないことがわかる．しかし，二番目に評価値が高い手役が最善手と一致した数は増加しているため，中盤のほうが最善手の推定ができていると考える．終盤では弱固定と強・弱固定の的中数が 7 であり，それ以外のシミュレーション手法の的中数は 8 である．このことから，相手の手札が完全にわかったとしてもカード残存枚数が多い盤面では最善手の推定が難しいことがわかり，カード残存枚数が少ない盤面では相手の手札が完全にわからない状態でも最善手の推定がしやすいことがわかる．また，本実験で用いた次の一手問題の合法手数は序盤で 6〜7 手，中盤と終盤で 5〜6 手程度と大差ないが，最善手の的中率には大差がある．そのため，合法手数が多ければ最善手の判定が難しく，少なければ最善手の判定がしやすいとはいえないことがわかる．

表 3 最善手の的中数
Table 3 The number of correct answers of and best hand.

シミュレーション手法	序盤	中盤	終盤
全ランダム	3 / 1	2 / 2	8 / 1
強固定	3 / 1	3 / 2	8 / 1
弱固定	3 / 1	2 / 3	7 / 1
強・弱固定	3 / 0	3 / 3	7 / 2
全固定	4 / 0	4 / 3	8 / 1
合法手数の幾何平均	6.6779	5.2249	5.5667

7. おわりに

本研究では，モンテカルロ法およびモンテカルロ木探索の評価値の収束，および不完全情報性がモンテカルロ法によるプレイヤに与える影響について調査した．

モンテカルロ法によるプレイヤでは，ある 1 手について 3000〜4000 回程度のプレイアウトを行うことで十分に評価値が収束していると考える．これは，これまでコンピュータ大貧民大会で使われいたモンテカルロ法プレイヤのプレイアウト回数よりも多い．一方，モンテカルロ木探索によるプレイヤでは，十分に評価値が収束するプレイアウト回数を判断することができなかった．

単純なモンテカルロ法によるプレイヤは，手札公開枚数が多くなるにつれてより正しい評価値を返す傾向が見られた．一方で，相手プレイヤの手札の全てが分かっている状態であっても，単純なモンテカルロ法による評価値とモンテカルロ木探索による評価値には差があった．さらに，序盤から中盤において相手プレイヤの手札のうち弱いカードの情報がより有用であり，また終盤においては相手プレイヤの手札のうち強いカードの情報がより有用であった．このように，公開枚数だけでなくカードの内容が不完全情報性に影響することが確認できた．

今後の課題は，プレイアウト中の各プレイヤの動作を改善することにより，単純なモンテカルロ法による評価値をモンテカルロ木探索の評価値に近づけ，より精度の高い結果を得ることが挙げられる．また，本研究において予想と異なる結果が得られた点について，さらに追実験が必要であると考える．

参考文献

[1] P. Auer, N. Cesa-Bianchi and P. Fischer. Finite-time Analysis of the Multi-armed Bandit problem. *Machine Learning*, Vol. 47, pp. 235–256 (2002).
[2] 電気通信大学. UEC コンピュータ大貧民大会, http://uecda.nishino-lab.jp/2011/ (2011).
[3] 池畑 望, 伊藤 毅志. Ms. Pac-Man におけるモンテカルロ木探索. **情報処理学会論文誌**, Vol. 52, No. 12, pp. 3817–3827 (2011).
[4] L. Kocsis and C. Szepesvári. Bandit Based Monte-Carlo Planning, *17th European Conference on Machine Learning (ECML 2006)*, Lecture Notes in Computer Science 4212, pp. 282-293 (2006).

[5] 小沼 啓, 西野 哲朗. コンピュータ大貧民に対するモンテカルロ法の適用. **研究報告ゲーム情報学** *(GI)*, Vol. 2011-GI-25, No. 3, pp.1–4 (2010).
[6] 前原 彰太, 橋本 剛, 小林 康幸. 局面評価関数を使う新たな UCT 探索法の提案とオセロによる評価. **研究報告ゲーム情報学** *(GI)*, Vol. 2010-GI-24, No. 5, pp. 1–5 (2010).
[7] 西野 順二, 西野 哲朗. 多人数不完全情報ゲームのモンテカルロ木探索における推定の効果. **研究報告バイオ情報学** *(BIO)*, Vol. 2011-BIO-27, No. 31, pp. 1–4 (2011).
[8] 西野 順二, 西野 哲朗. 大貧民における相手手札推定. **研究報告数理モデル化と問題解決** *(MPS)*, Vol. 2011-MPS-85, No. 9, pp. 1–6 (2011).
[9] 佐藤 佳州, 高橋 大介. モンテカルロ木探索によるコンピュータ将棋. **情報処理学会論文誌**, Vol. 50, No. 11, pp. 2740–2751 (2009).
[10] 島内 剛一, 有澤 誠, 野下 浩平, 浜田 穂積, 伏見 正則. **アルゴリズム辞典**. 共立出版株式会社, pp. 804–806 (1998).
[11] 須藤 郁弥, 篠原 歩. モンテカルロ法を用いたコンピュータ大貧民の思考ルーチン設計. 第 1 回 UEC コンピュータ大貧民シンポジウム (2010).
[12] 須藤 郁弥, 成澤 和志, 篠原 歩. UEC コンピュータ大貧民大会向けクライアント「snowl」の開発. 第 2 回 UEC コンピュータ大貧民シンポジウム (2011).
[13] 高橋 克吉, 伊藤 毅志, 村松 正和, 松原 仁. 次の一手問題を用いた囲碁プレイヤの局面認識についての分析. **情報処理学会論文誌**, Vol. 52, No. 12, pp. 3796–3805 (2011).
[14] 美添 一樹. モンテカルロ木探索 : コンピュータ囲碁に革命を起こした新手法. **情報処理**, Vol. 49, No. 6, pp. 688–693 (2008).

付　　録

A.1 付録：各問題に対する評価値の結果

以下に，第 6 章で述べた実験の結果の詳細を示す.

手法	ランダム	強固定	弱固定	強弱固定	全固定	MCT
序 1	2.9073	2.9251	2.7143	2.7802	2.6624	2.7498
序 2	4.3047	3.8216	3.8398	3.8106	3.6468	4.6200
序 3	2.3412	2.2518	2.3077	2.2941	2.1811	2.4087
序 4	2.4458	2.2313	2.4069	2.2238	2.5566	1.5723
序 5	4.7441	4.7920	4.7335	4.8264	4.7204	4.7366
序 6	3.5864	3.6080	3.5322	3.5659	3.4914	4.0788
序 7	2.8521	2.7602	2.6650	2.6247	2.6248	3.2347
序 8	2.4197	2.3117	2.2465	2.2121	2.1282	5.0000
序 9	2.3388	2.4046	1.8115	1.8753	1.9012	3.7752
序 10	3.6899	3.7031	3.7068	3.6737	3.8498	2.3521
中 1	3.7303	3.6937	3.5718	3.6457	3.7808	3.8331
中 2	2.8734	2.7451	2.4699	2.4797	2.4217	2.4167
中 3	4.7817	4.8952	4.6505	4.8544	4.7727	4.4049
中 4	4.3985	4.3266	4.2631	4.3928	4.5559	5.0000
中 5	1.9721	1.9454	1.8251	1.8170	1.6514	2.0326
中 6	3.2027	3.2550	2.6656	2.7242	2.7135	2.2266
中 7	2.4731	2.4310	2.6405	2.7305	2.7495	3.1496
中 8	3.7925	3.8028	3.5884	3.3831	3.3447	4.0000
中 9	4.5395	4.4872	4.1450	4.2731	4.3494	4.1208
中 10	3.8617	3.8011	3.7612	3.7058	3.5756	3.0000
終 1	3.9456	3.9903	3.9515	3.9963	3.9982	4.0000
終 2	2.5629	2.4937	2.5700	2.7861	2.6108	1.5053
終 3	3.9165	3.8917	3.9361	3.8888	3.9467	4.0000
終 4	3.5819	2.4250	3.2318	2.7277	2.7494	2.5495
終 5	2.8216	2.7280	2.8886	2.7979	3.0763	3.1130
終 6	2.9841	2.7871	3.1594	2.7591	2.8778	2.9775
終 7	3.7254	3.7126	3.4855	3.5266	3.5042	4.0000
終 8	3.8071	3.8462	3.7360	3.6986	3.5618	3.0000
終 9	3.6109	3.4480	3.4605	3.4578	3.1095	2.8360
終 10	2.8857	2.9007	2.8524	2.7039	3.2020	2.9622