1U-6

不完全情報ゲームUNOのモンテカルロ法による解法

松岡確† 堀内研‡ 中山泰一†

†電気通信大学大学院情報理工学研究科 ‡電気通信大学情報理工学部

1 はじめに

近年，ゲームの研究においてモンテカルロ法が注目されている．特に多人数不完全情報ゲームの大貧民ではモンテカルロ法を用いたソフトが大会 [4] で活躍している．本研究では多人数不完全情報ゲームの一種であるUNOにおいてモンテカルロ法を用いたプログラムを実装し性能を評価した．その結果，1ゲームごとの失点の平均を減らすことに成功した．また，強いアルゴリズムを実装しているプレイヤに隣接したプレイヤは，上がる頻度が大幅に下がり，また平均の失点数を大きく落とすことが分かった．

2 関連研究

2.1 大貧民のモンテカルロ法

須藤ら [2][3] は多人数不完全情報ゲームの一種「大貧民」で，対戦相手の手をランダムに想定する，モンテカルロ法で実装したプレイヤプログラムを作成し，第4回および第5回UECコンピュータ大貧民大会で優勝した．また，小沼ら [1] は須藤らのプログラムをモンテカルロ法のプレイアウトの最善手の推定にε-Greedy法を用いるなどの改良を行い，同大会の第6回で優勝した．

3 題材と方針

3.1 UNOについて

UNOは1971年にMerle Robbinsが考案した多人数不完全情報ゲームである．

日本版の公式ルールでは誰か一人が手札が0枚になったとき（上がった）に，残ったプレイヤの手札の失点を計算し，あがったプレイヤはその各プレイヤの失点分を自分の加点とする零和ゲームとなっている．

本研究でのルールは日本版のルールを採用するものとする．

3.2 方針

モンテカルロ法を用いて強いプログラムを作成するにあたり，UNOで強いプレイヤを明快に定義するのは難しい．強さの例としては次のような場合が考えられる．

1. ほかのプレイヤよりも平均した失点が低い
2. ほかのプレイヤよりも上がる頻度が高い
3. ほかのプレイヤよりもより効率的なタイミングでスクロールカードが使える
4. トータルで勝っているプレイヤの足を引っ張ることに長ける

UNOというゲームを考える上で負けにくいプログラムであることがトータルで見たとき強いプログラム，と考えられるので今回は1の条件で実装を行った．

3.3 実装

失点の少ないプログラムを作成するにあたり以下の式を定義する．

$$R_n = \sum_{p=1}^{m}(n_p)$$

この式はある局面に対して合法手 R_n を選択したプレイアウトの終端局面における失点の和である．すなわちmはプレイアウトの回数，

$$n_p$$

は．この値が最小となる手を最善手として選ぶことになる．

4 実験と考察

4.1 モンテカルロ法を用いない場合

まず，モンテカルロ法を用いて実験を行う前に失点を減らすことに主眼を置いたヒューリスティックな知

A solution for a game UNO with imcomplete information by Monte Carlo Method

Akira MATSUOKA†, Ken HORIUCHI‡and Yasuichi NAKAYAMA†

†Graduate School of Informatics and Engineering, The University of Electro-Communications

‡Faculty of Informatics and Engineering, The University of Electro-Communications

182-8585, Chofu, Tokyo, Japan

matsuo-a@igo.cs.uec.ac.jp

Copyright ©2014 Information Processing Society of Japan.
All Rights Reserved.

識を与えたプレイヤプログラムを作成し，完全にランダム[1]にプレイするプレイヤプログラム4体と1万回の対戦を行い性能を比較した．結果を下に示す．プレイヤ1が今回実装したプログラムである．

表 1: モンテカルロ法未使用の場合

プレイヤ	平均失点数	平均順位	上がった回数
1	-8.66	2.78	2023
2	6.20	3.09	1864
3	-0.77	2.95	2083
4	1.19	2.96	2018
5	2.03	2.92	2012

このことから以下の知見が得られた．

- プレイヤの強さとゲームで上がる回数にほとんど相関は見られない
- ただし，強いプレイヤが直前にいるプレイヤは失点数，平均順位，上がる回数全てが著しく下がる

4.2 モンテカルロ法を用いた場合

提案手法の元で実装したプログラムを用いてヒューリスティックなプレイヤと同様の条件の下で実験を行った．また，対戦相手のプレイヤにヒューリスティックなプレイヤを用いて同様に行った結果を以下に示す．

表 2: モンテカルロ法対ランダムプレイヤ

プレイヤ	平均失点数	平均順位	上がった回数
1	-8.41	2.76	2049
2	4.39	3.07	1938
3	0.53	2.99	2083
4	0.65	2.96	2042
5	2.85	2.95	1933

表 3: モンテカルロ法対ヒューリスティックプレイヤ

プレイヤ	平均失点数	平均順位	上がった回数
1	-0.52	2.99	2024
2	6.20	3.09	1985
3	-0.77	2.95	2023
4	1.19	2.96	1988
5	2.03	2.92	1980

ヒューリスティックな知識だけのプレイヤのときの実験と大きな差は見られなかった．しかし，モンテカルロ法を用いたプレイヤとヒューリスティックな知識で戦うプレイヤ4人を相手にしても，1体を除き勝ち越したため，強さが僅かに向上していると考えられる．

また，上がった回数で僅かにほかのプレイヤの数を上回ったので目的の結果は得られた，と考えられる．

[1]ただし，カードを出すことができるときは必ず出す

4.3 考察

今回の実験での1万回の試行では1回のゲームごとの誤差が大きなものになっている可能性が考えられる．また，実装したプレイヤから1つ飛んだ先のプレイヤの成績が良くなる傾向があるのは，実装したプレイヤの使用する，スクロールカードの影響で自身がカードをプレイする機会が増えてるものと予想される．

5 まとめと今後の課題

今回は，UNOにおいてモンテカルロ法を用いてプレイヤの強さが向上することを実証できた．また，UNOのゲーム性の特徴として，強いプレイヤに隣接，特に初期の回りで直後にあたるプレイヤは強いプレイの影響を受け，成績が同等の性能のものに比べ下がることが分かった．

今回ゲームの終了時の得点を，報酬の与え方として最善なものであったかどうかは，依然不明である．報酬の与え方をゲームの失点のみに依存したものや上がる回数に着目したものも実装して追加実証して行く．今回は実装の簡単さから原始モンテカルロ法を用いたが，モンテカルロ木探索を行うなどの改善が考えられる．また，プレイアウトにおいて対戦相手のプレイヤが持ち得ぬ手札を渡している，などの問題点もある．そのためプレイアウト自体の条件の改善などを行っていきたい．

参考文献

[1] 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-25, No.3) (2011).

[2] 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 (2009).

[3] 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 (2010).

[4] UECコンピュータ大貧民大会，
http://uecda.nishino-lab.jp/

Copyright ©2014 Information Processing Society of Japan.
All Rights Reserved.