

# モンテカルロ法を用いた UNOのプレイヤープログラムの提案

堀内 研<sup>†</sup> 松岡 確<sup>††</sup> 中山 泰一<sup>††</sup>

<sup>†</sup>電気通信大学 情報理工学部 <sup>††</sup>電気通信大学大学院 情報理工学研究科

## 1. はじめに

近年、ゲームのプレイヤーアルゴリズムとしてモンテカルロ法が注目され、コンピュータ囲碁[1]やコンピュータ大貧民[2]などで適用され、成果をあげている。

本研究ではモンテカルロ法を用いた多人数不完全情報ゲームUNOのプレイヤープログラムを提案、実装する。UNOはゲーム中に手札が増加する点や、ゲーム終了時の手札に応じた得点と失点をもとに順位を決定する点など、既存のモンテカルロ法で用いられてきたゲームとは異なる特徴を持つ。本研究ではその特徴に対してモンテカルロ法がどの程度有効であるか検証することを目的とする。

## 2. 設計

モンテカルロ法は各合法手に対して多数回のプレイアウトを行い、その報酬和が最大となる手を選択する。プレイアウトの報酬の与え方として、あがり、順位、手札の残数、得点、失点を、それぞれ提案し実装した。プレイアウトの回数は300回とし、プレイアウト中ではヒューリスティックなアルゴリズムを用いたプレイヤーを使用した。

## 3. 実験・考察

まず、各報酬を実装したプレイヤーの性能を比較するために、ランダムにプレイするプレイヤー4体を対戦相手として1万回の対戦を行った。その結果が図1である。プレイアウトの報酬を変えることでプレイヤーの性能が変化することが確認できた。平均得点において、得点、失点といった得点計算に関わる部分を報酬として与えた方が性能の向上がみられるようである。しかし、順位を報酬としたものに対して、明確に強いという結果は得られなかった。

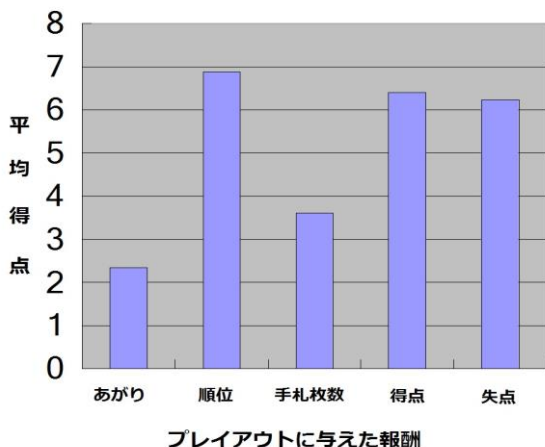


図1. 報酬を変化させた場合の対戦結果

表1. ヒューリスティックなプレイヤーとの対戦結果

プレイヤー	平均得点	平均順位	勝利数
モンテカルロ	-4.39	3.23	2161
ヒューリスティック	3.08	2.89	2161
ヒューリスティック	0.19	2.89	1926
ヒューリスティック	1.40	2.82	1985
ヒューリスティック	0.26	2.83	1894

次に得点を報酬として与えたプレイヤーを用いて、ヒューリスティックなアルゴリズムのプレイヤー4体と1万回の対戦を行った。その結果が表1である。モンテカルロ法を用いたプレイヤーはヒューリスティックなアルゴリズムのプレイヤーに比べて明確に強いとは言えない結果となった。得点を報酬として与えた際に、あがったプレイアウトの報酬について適正に評価されていない可能性や、UNOの合法手がそれほど多くないため、ヒューリスティックなプレイヤーのルーチンが有利に働いた可能性などが考えられる。

## 4. まとめと今後の課題

本研究ではモンテカルロ法を用いたUNOのプレイヤープログラムを提案し、それを実装した。得点や失点をプレイアウトの報酬として与える方式は一定の性能の向上がみられたが、順位を報酬とするプレイヤーと大差が出るほどではなかった。ヒューリスティックなプレイヤーに対して、モンテカルロ法を実装したプレイヤープログラムは及ばなかった。

今後の課題としてモンテカルロ法の精度を上げるための手法を検討する必要がある。得点や失点を報酬とするとき、上がった時の報酬に適当な重みを付けて評価するといった手法や、他プレイヤーの得点や失点を計上するなどといった手法が考えられる。また、今回実装したプレイアウトプログラムは原始的なものであったため、対戦相手の出したカードをもとに手札の推測を行うなどすれば、プレイアウトの精度を上げることができると期待される。モンテカルロ法の精度を高めることで、少なくとも同程度の実力を得ることができると考えている。

## 参考文献

- [1]松原仁 編, 美添一樹, 山下宏 著:コンピュータ囲碁 モンテカルロ法の理論と実践, 共立出版株式会社, 2011年11月10日 第1刷発行。  
[2]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, 情報処理学会研究報告, Vol2011-GI-25 No.3.