

## 不完全情報ゲーム UNO のモンテカルロ法によるプレイヤーの構築

福島 健介† 鈴木 徹也†

芝浦工業大学 システム理工学部 電子情報システム学科†

## 1. はじめに

不完全情報ゲームの研究は様々な種類のゲームについて行われている。その中でも、多人数不完全情報ゲームについては麻雀、ポーカー、大貧民などが研究の中心になっている。

しかし、他にも多くの異なる特徴を持った多人数不完全情報ゲームが広く遊ばれている。例として、カードゲーム UNO は手札の増減、特殊なカードの存在などの異なる特徴をもっている。

本研究の目的は、多人数不完全情報ゲーム UNO に対して特徴に適した戦略を見つけることである。そのため平均順位、平均点数、上がり率を上げるプレイヤーと妨害プレイヤー、これらの複合プレイヤーを提案する。

## 2. 不完全情報ゲーム UNO

UNO は 1971 年に Merle Robbins が考案した不完全情報ゲームのカードゲームである。

初めに配られた 7 枚のカードをすべてなくしたプレイヤーがそのラウンド(R)の勝者となり他のプレイヤーから得点を得る。これを 5 R 繰り返し最終的な順位を決める。

UNO は 76 枚の数字カード (各色 0 は 1 枚, 1-9 は 2 枚) と特殊なカードとしてリバース、スキップ、ドロー 2 の 3 種が各色 2 枚, ワイルド, ワイルド・ドロー 4 の 2 種が各 4 枚の合計 108 枚を 1 セットとしてゲームに使用する。

UNO には多くのローカルルールが存在し、公式ルールとは異なるルールでプレイされる場合が多い。

本研究は以下のルールで実験を行う。一般に遊ばれているルールに近づけるために日本 UNO 協会の競技ルールから以下のルールを採用する。

- ドロー 2 の効果とワイルド・ドロー 4 の効果の回避と先送り。

また、本研究では以下のルールを採用しない。

- 手札が 1 枚になる際に「UNO」と宣言する。

- 出せるカードが存在する時でも、もっていないふりをして山札からカードを引くことができる。
- ワイルド・ドロー 4 が使われた場合にチャレンジを行うことができる。

## 3. 関連研究

松岡ら [1] は UNO に対して、平均した失点が少ないプレイヤーが強いプレイヤーだと考え平均失点の少なさを報酬としてモンテカルロ法を適用した。その結果ランダムプレイヤー、ヒューリスティックプレイヤーに対して優位であることを示した。

三石ら [2] は大貧民に対してモンテカルロ法の報酬に対戦相手の得点期待値を組み込むことにより、特定の相手の得点に狙った影響を与えられることを示した。

松岡ら [1] の研究では 1 つの戦略しか試していないが、UNO は他の戦略も考えられる。三石ら [2] の研究は大貧民に対して行っているが、UNO にはドロー 2 などの特殊なカードが存在するため妨害の戦略がより有効であると予想される。また零和ゲームであるため相手の得点を下げること自身で自身の得点を上げることができる。

## 4. 提案手法

3 節で述べた関連研究の問題点と UNO での効果の予想とにもとづき、4 つのプレイヤーを提案する。

- $P_1$  平均点数を報酬としたプレイヤー
- $P_2$  平均順位を報酬としたプレイヤー
- $P_3$  上がる頻度を報酬としたプレイヤー
- $P_4$  特定の相手の得点の低さを報酬とした妨害プレイヤー

また、妨害の対象を対戦中に変更することで順位を上げることができると可能性がある。そのため 1 つの複合プレイヤーを提案する。

- $P_{X-4}$  順位が 1 位ときは  $P_X (X=1, 2, 3)$ , 2 位以下のときは 1 位を妨害の対象とした  $P_4$  となるプレイヤー

この中で  $P_1$  は松岡ら [1] の戦略に近いプレイヤーである。

Construction of a Monte-Carlo Based Player for the Incomplete Information Game UNO

† Fukushima Kensuke, Shibaura Institute of Technology

† Suzuki Tetsuya, Shibaura Institute of Technology

## 5. 実験

提案手法の評価のために3つの実験を行った。

モンテカルロ法のプレイアウト回数は予備実験により 2500 回に決定した。作成したプレイヤーの評価のためにヒューリスティックプレイヤー  $P_H$  を作成した。また、理論上ランダムプレイヤー同士の対戦では平均点数は 0 点であり、平均順位は 2.5 位となる。

### 5.1 実験1 最適な報酬の調査

UNO に適したモンテカルロ法の報酬を調査するために、3 つのプレイヤー  $P_1, P_2, P_3$  のそれぞれと 3 つのランダムプレイヤー  $P_R$  との対戦を行った。1R 毎の結果と最終結果の平均を表 1 と表 2 に示す。

表 1 ランダムプレイヤーとの対戦 1R

	$P_H$	$P_1$	$P_2$	$P_3$
平均点数	7.994	15.833	24.130	10.888
平均順位	2.368	2.321	2.153	2.372
上がり率	0.265	0.325	0.336	0.299

表 2 ランダムプレイヤーとの対戦 最終結果

	$P_H$	$P_1$	$P_2$	$P_3$
平均点数	39.970	79.164	120.652	54.442
平均順位	2.276	2.154	1.930	2.250

実験の結果 1R、最終結果ともに  $P_2$  が平均点数、平均順位、上がり率全てにおいて最も高くなっている。これより、 $P_2$  が最も UNO に適したプレイヤーであるといえる。

### 5.2 実験2 妨害の効果の調査

$P_4$  の妨害効果を検証するために  $P_4$  と 3 つの  $P_R$  との対戦を  $P_4$  の妨害対象を変えて 2 通り行った。実験結果を表 3 と表 4 に示す。表中の  $P_{R-n}$  はプレイ順で  $P_4$  の  $n$  個次の  $P_R$  である。また、 $P_{4-n}$  は  $P_{R-n}$  を妨害対象とする  $P_4$  である。

表 3  $P_{R-1}$  を対象とした  $P_{4-1}$  の最終結果

	$P_{4-1}$	$P_{R-1}$	$P_{R-2}$	$P_{R-3}$
平均点数	57.258	-63.414	-1.536	7.692
平均順位	2.230	2.792	2.516	2.458

表 4  $P_{R-2}$  を対象とした  $P_{4-2}$  の最終結果

	$P_{4-2}$	$P_{R-1}$	$P_{R-2}$	$P_{R-3}$
平均点数	52.366	-13.282	-30.056	-9.028
平均順位	2.216	2.598	2.628	2.550

$P_{4-1}, P_{4-2}$  ともに対象としていないプレイヤーよりも対象としたプレイヤーの平均点数、平均順位が低くなっている。これより妨害が実現できているといえる。また、どちらもランダムプレイヤーに対しての優位性を持っているといえる。

$P_{4-2}$  が  $P_{4-1}$  よりも妨害の効果が小さいのは、直接カードで干渉できないためであると考えられる。

### 5.3 実験3 複合プレイヤーの調査

実験 1 の結果から最適と判断した  $P_2$  と  $P_4$  とを組み合わせた  $P_{2-4}$  を評価するために実験を行った。

ランダムプレイヤー 3 つとの対戦とヒューリスティックプレイヤー 3 つとの対戦を  $P_2, P_{4-2}$  それぞれで行った。最終結果の平均を表 5 と表 6 に示す。

表 5 ランダムプレイヤーとの対戦 最終結果

	$P_2$	$P_{2-4}$
平均点数	120.652	59.160
平均順位	1.930	2.244

表 6 ヒューリスティックプレイヤーとの対戦 最終結果

	$P_2$	$P_{2-4}$
平均点数	63.372	30.554
平均順位	2.158	2.344

$P_2, P_{2-4}$  ともにヒューリスティックプレイヤー、ランダムプレイヤー両方に対して優位性を持っているといえる。しかし、 $P_{2-4}$  は  $P_2$  と比較すると平均点数が半分以下になり平均順位も下がっているので明らかに弱い。これは、 $P_{2-4}$  は 2 位以下の場合に自身の順位を考慮しなくなるために、上がる頻度や順位を落としてしまったからだと考えられる。

## 6. おわりに

本研究では、UNO のモンテカルロ法プレイヤーを 5 つの戦略で作成しそれらの効果を調査した。その結果平均順位を上げる戦略が最も適していることを示した。

今後の課題として、相手手札の推定などを組み込むことでプレイアウト精度の改善を行いより効果的なプレイヤーを作れるのではないかと考えている。また、より適した戦略や戦略の組み合わせの調査が課題として挙げられる。

### 参考文献

- [1] 松岡 確, 堀内研, 中山 泰一. 第 76 回全国大会講演論文集, 2014(1), pp. 605-606 (2014-03-11).
- [2] 三石 亮, 大久保 誠也, 若月 光夫, 西野 哲郎. 研究報告ゲーム情報学, (GI), 2019-GI-41(26), 1-6 (2019-03-01).