

大富豪における特定プレイヤーへの付度を可能にする コンピュータプレイヤーの検討

大室 光† 横山 大作†

明治大学大学院理工学研究科情報科学専攻†

1. はじめに

ゲーム戦略の研究では、プレイヤー自身が勝つことを目的とし、その為にいかに自分に優位な状況を作りだすかを考えることが多い。しかし、現実のゲームでは特定の相手プレイヤーを勝たせたい場面も存在する。その場合、自身ではなくその相手が有利になるような戦略、いわば特定の相手に付度する戦略が必要となる。

本研究では、多人数不完全情報ゲームの1つである大富豪において、原始モンテカルロ探索を用いたプレイヤーを構築し、相手に付度する戦略が実現可能かを検証する。

特定のプレイヤーの順位に基づく報酬を最大化するようなプレイヤーを構築し、それぞれ次のプレイヤー、対面のプレイヤー、前のプレイヤーに対して付度する実験を行ったところ、各プレイヤーに対して付度が実現可能であることを確認した。

2. 検証用のゲーム

大富豪は複数人プレイの不完全情報トランプゲームである。トランプを全てのプレイヤーに均等に配り、手持ちのカードを順番に場に出していき、早く手札を無くすことを競うゲームである。定番のトランプゲームの1つではあるが、「革命」、「スート縛り」などの基本ルール他に「Jバック」などのローカルルールが数多く存在し、プレイする際に事前にどのルールを採用しているかを確認する必要がある。

本研究では、基本ルールとして UEC 標準ルール[1]を採用し、異なる点は以下の通りである。

- プレイヤーの人数は4人。席順は常に固定しており、時計回りに付度プレイヤー、CPUA、CPUB、CPUCの順とする。
- ゲーム開始時の手札交換において、大富豪が大貧民に、富豪が貧民に渡すカードは任意のカードではなく、最も弱いカードから選択するものとする。

3. 検証用プレイヤーの実装

本研究では、原始モンテカルロ探索を採用し、大貧民におけるモンテカルロ法を参考に付度プレイヤーを実装した[2]。

パスを含んだ全ての合法手に対して、以下の手順を100回繰り返す。

- ①合法手を選択
- ②相手プレイヤーの手札を回収し、ランダムに再配布する
- ③全プレイヤーをランダムプレイヤーとして、付度を受けるプレイヤーが上がるまでゲームを続ける
- ④ゲームの結果、付度を受けるプレイヤーが大富豪であれば3、富豪は2、貧民は1、大貧民は0を報酬値として返す

各合法手の報酬値の平均を比較し、最大の合法手を最適手とする。

なお、既に付度を受けるプレイヤーが上がっている場合は、パスを除く合法手の中からランダムに手を選択する。

4. 付度可能性の評価

上記のモンテカルロ法を用いた付度プレイヤーが、次のプレイヤー(CPUA)、対面のプレイヤー(CPUB)、前のプレイヤー(CPUC)に対して付度することが可能かを検証する。

実験は以下の4種類行う。

- AllRandom
全プレイヤーをランダムプレイヤーとする
- SontakuA, B, C
付度プレイヤーがそれぞれ次のプレイヤー(CPUA)、対面のプレイヤー(CPUB)、前のプレイヤー(CPUC)に対し付度する

全てのゲームにおいて、相手プレイヤーは、パスを除く合法手の中からランダムに手を選択するランダムプレイヤーである。

ゲーム終了時に大富豪であれば3点、富豪は2点、貧民は1点、大貧民は0点を得られるとし、プレイヤーが1000試合で得た得点の平均点をそのプレイヤーの評価値とする。

A study of a computer player that enable letting a specific opponent win in Daifugo.

†Hikaru Ohmuro, Daisaku Yokoyama, Meiji University

4.1. AllRandom

全プレイヤーをランダムプレイヤーとしてゲームを行った. 表 1 は各プレイヤーの順位と評価値をまとめたものである.

表 1 各プレイヤーの順位の回数と評価値

	1 位	2 位	3 位	4 位	評価値
付度プレイヤー	281	251	254	214	1.599
CPUA	253	239	241	267	1.478
CPUB	231	261	247	261	1.462
CPUC	235	249	258	258	1.461

この AllRandom の評価値と 1 位 (大富豪) になった回数を参考値とし, 付度を受けたプレイヤーがこの数値をどこまで超えることができるかを評価する.

4.2. SontakuA, B, C

付度プレイヤーが各プレイヤー (CPUA, CPUB, CPUC) に対して付度を行うゲームである. 表 2 はゲーム毎に各プレイヤーが大富豪になった回数をまとめたものである.

表 2 ゲーム毎の各プレイヤーの大富豪の回数

	SontakuA	SontakuB	SontakuC
付度プレイヤー	0	0	0
CPUA	588	386	279
CPUB	208	386	340
CPUC	204	228	381

4.3. 対戦実験結果

図 1 は各ゲームにおけるプレイヤー毎の評価値をまとめたものである. 付度を受けたプレイヤーの評価値は上昇しており, 他のプレイヤーと比べても高くなっていることが確認できる.

また, 付度を受けたプレイヤーの前の手番のプレイヤーの評価値もわずかながら上昇する傾向があることがわかる.

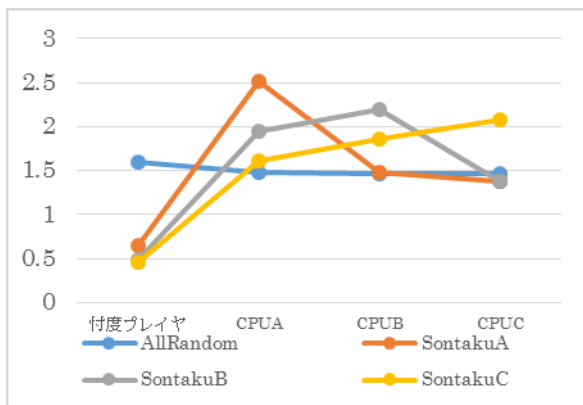


図 1 各ゲームのプレイヤー毎の評価値

図 2 は各プレイヤーが付度を受けた際的评价値の上昇率である. 付度を受けるプレイヤーの手番が付度プレイヤーの手番の直後であるほど得点の上昇率が高く, CPUA と CPUC では 27%ほどの差がある.

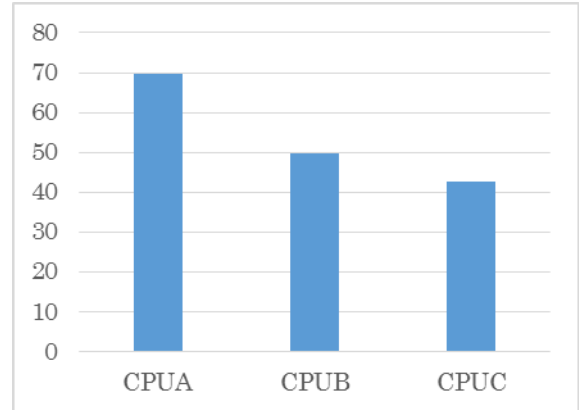


図 2 付度を受けたプレイヤーの評価値の上昇率

5. まとめ

本研究では, モンテカルロ法を用いたプレイヤーを構築し, 相手に付度する戦略が実現可能かを検証した.

実験から, 次のプレイヤー, 対面のプレイヤー, 前のプレイヤーそれぞれに対して付度するプレイヤーの構築に成功した. また, 付度を受けるプレイヤーの手番が付度プレイヤーの手番の直後であるほど, 付度をしやすいことがわかった.

今後の課題として以下のことが考えられる.

- ・自身の手番から遠いプレイヤーへの付度の精度の上げ方の模索
- ・相手プレイヤーがある程度の戦略を持つ場合の付度法
- ・別の手法による付度プレイヤーの作成法の模索

また, 今回のモンテカルロ法を用いた手法が, 麻雀や類など他の不完全情報ゲームにも適用可能なのを確かめたい.

参考文献

- [1] 電気通信大学, “UECd-2021 コンピュータ大貧民大会” (<http://www.tnlab.inf.uec.ac.jp/daihinmin/2021/>)
- [2] 漆畑雅士, “多人数不完全情報ゲームに対する局面評価値を用いたモンテカルロ法”, 数理解析研究所講究録 第 1894 巻 2014 年 pp. 84-88