

大富豪における人間らしい行動で接待する コンピュータプレイや実現の試み*

東野勇人^{†1} 横山大作^{‡1}

¹ 明治大学理工学部

1 はじめに

近年, 様々なゲームでゲーム AI が導入されており, ゲーム AI にはプレイヤーである人間をより楽しませることが求められている. そのためには, ゲーム AI がプレイヤーである人間の實力に合わせたプレイができることが望ましい. その際に注意すべき点として, 明らかに實力を合わせているという不自然な行動は除外しなければならない. これらを実現することで人間らしい行動をし, なおかつ實力を調整できるゲーム AI の作成を試みる. そこで本研究では多人数ゲームである大富豪をゲーム環境に選択した. 理由として, 多人数ゲームであれば實力を調整することは中間の順位を取ることで簡単に実現できると考えたからである.

2 研究方針

本研究の目的は, 様々なゲームにおいて, ゲーム AI がプレイヤーである人間を楽しませることを可能にすることである. そのためにはプレイヤーの實力を認識し, その實力に拮抗するようにゲーム AI の實力を調整する必要がある. しかしその際に, ゲーム AI が實力を調整しているとプレイヤーに感じさせてしまうことは望ましくない. そこでプレイヤーから見た時に明らかに不自然と捉えられる行動を選択しないようにすることが求められる. 例えば大富豪においては, 合法手が存在するのにもかかわらず連続でパスを選択し続けることはとても不自然である. その他にも, 上がることが確定している手札でわざと上がらずに順番を回すなども不自然な行動である. このような不自然な行動に共通しているのは手加減をしてい

るという点である. 本研究ではその中でもパスを連続して行うという行為に焦点を当て, そのような行為が無いような手加減を行えるコンピュータプレイの構築を試みる.

3 検証用ゲーム環境

3.1 大富豪

大富豪はトランプを用いた多人数不完全情報ゲームである. ゲーム進行の流れとしては, 初めに手札が配られ, その中から 1 枚または複数枚のカードを役として場に提出する. 早く手札を無くした人から順位が高くなっていくゲーム性である. 今回はプレイ人数を 5 名として対戦を行う. また, 現在大富豪には様々なローカルルールが存在するが, 本研究では電気通信大学が主催するコンピュータ大貧民大会 UECda[1] で採用されているルールを参考にして実装した. 以下にそのルールの詳細を示す.

カードの枚数 各マークのエースからキングまでの 52 枚にジョーカー 1 枚を加えた計 53 枚.

カードの強さ 3 が一番弱く, 数字が大きくなるほど強いカードとなる. なお, エースはキングよりも, 2 はエースよりも強いものとする. また, ジョーカーは全てのカードの中で 1 番強い (スペードの 3 を除く).

役 単体役は 1 枚のみで場に出せる役. 複数役は同じ数字のカードを 2 枚以上で場に出せる役. 階段役は同じマークでつながった数字のカードを 3 枚以上揃えた役.

カードの出し方 場が空の時, どんな役も出すことができる. 場に出ている場合, そのカードの役と枚数が同じであり, かつ場のカードよりも強いカードを出すことができる. また, カードを出すことができない場合は, 何もカードを出さないパスを選択する.

* Towards the strength adjustment of the computer player using human-like plays in Daifugo

[†] Higashino Hayato

[‡] Yokoyama Daisaku

革命 4枚以上のカードが複数役として提出された時、カードの強弱が逆転した状態となる。この状態は場が流れても継続し、一回のゲームが終了するまで継続する。また、革命状態で再び革命の条件を満たす提出がされた場合、革命は解除される。

スペ3切り ジョーカーが単体役で場に出された時にスペードの3を場に出すことができる。その後場は流れる。

8切り 数字の8を含む役が場に出された場合、その場は流れる。

縛り 場のカードと同じマークでカードが出されると縛りとなり、場のカードと同じマークの役しか出せなくなる。

都落ち ランクが大富豪のプレイヤーより先に上がったプレイヤーがいるとき、大富豪であったプレイヤーが自動的に大貧民になる。

3.2 中間の順位を狙うプレイヤー

本研究では、原始モンテカルロ法でゲームプレイヤーを構築する。大富豪ではそれぞれのプレイヤーの手札の枚数は単調に減少していくため、乱数におけるシミュレーションの実装がしやすいという点から原始モンテカルロ法の使用を検討した。今回作成するプレイヤーは合法手を選択し、乱数によるシミュレーションを行い、プレイアウトごとに報酬を与える。プレイアウトとはそのゲームを終局までプレイすることである。報酬は終局時の順位によって決定され、大富豪なら-1、富豪なら1、平民なら1、貧民なら1、大貧民なら-1を報酬値として返す。これは中間の2,3,4位をとった際にはプラスの報酬を与え、それ以外の時はマイナスの報酬を与えている。また、不自然な行動を観測した際にも報酬を与える。具体的には合法手を持っているという条件下でパスを2回以上連続して行った際には-1の報酬を与える。

4 実験設定

本研究では様々な方針のプレイヤーの中で2,3,4位を意図的に狙うことができるのか、またその際に不自然な行動が無いように手加減を行えるのかを検証する。不自然な行動に関しては、マイナスの報酬を与えない場合よりも与える場合の方が不自然な行動が減っているのかを検証する。

4.1 その他のプレイヤー設定

中間の順位を狙うプレイヤーの他に4種類のプレイヤーを用意し、合計5種類のエージェントで対戦を行い、その優劣を比較する。完全なランダムプレイヤーやヒューリスティックな戦術を持つプレイヤー、モンテカルロ法で上位を狙うプレイヤーを用意することで、様々な方針のプレイヤーの中で中間の順位を取るという検証を行うことを目指す。4種類のプレイヤーの戦術の内訳は、合法手の中から完全にランダムで手を選択するプレイヤー1種類、合法手の中から一番枚数が多く、ランクが高いものを選択するプレイヤー1種類、合法手の中から一番枚数が多く、ランクが低いものを選択するプレイヤー1種類、モンテカルロ法において報酬値を大富豪なら2、富豪なら1、平民なら0、貧民なら-1、大貧民なら-2に設定した自分の順位を上げることを目指したモンテカルロプレイヤー1種類である。

4.2 対戦環境

ルールに都落ちを採用しているため、初戦以外は1つ前の試合の結果が必要となってくる。そのため連続的にゲームを進行していく。試合回数は1000回としている。

5 まとめ・今後の課題

本研究では、大富豪において人間らしい行動で接待をするコンピュータプレイヤーの実現に向けて、パスを連続して行うといった不自然な行動が無いように手加減をして中間の順位を意図的に取得することが可能かどうかを検証する。本稿の作成時点では、プレイヤーの作成段階であり、今後はプレイヤーの構築を行い、実験を進めていく。

参考文献

- [1] 電気通信大学. UEC コンピュータ大貧民大会 2021, <http://www.tnlab.inf.uec.ac.jp/daihinmin/2021/> (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 (計算理論とアルゴリズムの新潮流)." 数理解析研究所講究録 1894 (2014): 84-88.
- [3] 岡和人, 松崎公紀. "札譜データの学習を用いた大貧民モンテカルロプレイヤーの強化." 第56回プログラミング・シンポジウム予稿集 2015 (2015): 13-24.