

ニューラルネットワークによる大貧民の カード提出モデル構築

内田 純平¹ 穴田 一¹

概要：大貧民とは、日本で行われているカードゲームの一種である。様々なローカルルールがあることで知られているが、ゲームの開始時に配布されたカードを順に出していき、カードを早く無くすことを目指すことは共通である。複数人でプレイするゲームで、相手の手札が公開されていないという点から、多人数不完全情報ゲームに分類される。この様なカードゲームで常に上位を目指してプレイを行うと悪い結果を招くことがある。

そこで本研究では、自分の手札と場の状態から目指すべき順位を動的に判断し、それに伴うカード提出を行うモデル構築を目的とする

キーワード：大貧民, 大富豪, エージェント, ニューラルネットワーク

Construction of a Card Submission Model by Neural Network

JUMPEI UCHIDA¹ HAJIME ANADA¹

Abstract: DAIHINMIN is a kind of card game played in Japan. It is known that there are various local rules employed while playing the game. In addition, it is a game played by multiple people, and it is classified as a multiplayer incomplete information game because opponents' hands are not open. While playing this game, it is not good to always play and aim at the top position. Therefore, this study aims to estimate the final position of a submitter using his/her hand and the state of the field.

1. はじめに

ゲーム AI の研究が精力的に行われており、将棋や囲碁は人間のトッププロに迫る強さになりつつある。これらのゲームは完全情報ゲームであり、プレイヤーは互いに全ての情報を手にいれることができる。一方、一部の情報が得られないゲームは不完全情報ゲームと呼ばれる。不完全情報ゲームの研究も行われており、人狼知能[1]等でゲームログのデータ解析などが行われている。このような不完全情報ゲームの一つとして、トランプゲームの大貧民がある。大貧民

をコンピュータにプレイさせる大会が UEC コンピュータ大貧民大会 (UEC-da) として、2006 年より開催されている[2]。

また、アルゴリズム[3][4]の提案も行われている。

大貧民 (大富豪) とは、日本で行われているカードゲームの一種であり、様々なローカルルールがあることで知られているが、ゲームの開始時に配布されたカードを順に出していき、カードを早く無くすことを目指すことは共通である。

このようなゲームの性質のため、各プレイヤーはゲーム開始時に配布されたカードを上手く出していくことが求められる。ゲーム開始時に配布されるカードは無作為に配布されるため、手札の強さによっては最適なプレイを行ったとしても上位を目指すことは難しい。特に、大貧民では前

¹ 東京都市大学大学院総合理工学研究科情報専攻

回の順位に伴うカード交換を行い、順位が高いプレイヤーは強い手札になり、順位が低いプレイヤーは弱い手札からゲームを始めなければいけないため、順位の高いプレイヤーが順位を大きく上げようとする悪い結果を招く可能性が高い。

そこで本研究では、コンピュータ大貧民におけるプレイヤーの手札や場のカード、既に出たカードなど様々な状況を基に目指すことができる順位を推定し、推定結果に基づいた提出を行うことができるカード提出モデルの構築を目的とする。

2. UEC 標準ルール

コンピュータ大貧民は、カードゲームの 1 種である大貧民を計算機上で行うゲームである。電気通信大学において大会が毎年開催されており、統一ルールや各種プログラム、開発環境が公開されている。本研究でもこの枠組みを利用する。UEC コンピュータ大貧民は、場と進行を管理するサーバーと、ゲームをプレイする 5 つのクライアントからなる。大貧民には非常に数多くの地域特有のルールが存在しているが、UEC コンピュータ大貧民大会では次のようなルールが採用されている。

(1) カードの枚数

ゲームに使用するカードは各マークのエースからキングまでの 52 枚とジョーカー 1 枚を加えた計 53 枚である。

(2) ゲーム開始と順番の回り方

各プレイヤーの手札が決まるとゲームが始まる。ゲームは、ダイヤの 3 を持っているプレイヤーから始まり、その次の席順のプレイヤーが提出を行う。この時、必ずしもダイヤの 3 を出す必要はない。なお、この席順は乱数で決定され、何ゲームかその席順でゲームを行った後、再度席替えが行われる。

(3) カードの配布方法

初回はランダムにカード配布の順番を決定し、カードを配布する。大貧民、大富豪などの序列が決定されている 2 回目以降は、大富豪を起点として席順にカードを配布する。

(4) カードの強さ

普通は 3 が一番弱く、数字が大きくなるほど強いカードとなる。また、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。革命（後述）が発生するとカードの強弱は逆転し、3 が最も強いカードとなる。

(5) カードの出し方

場に何も無い状態（場が空である状態）であれば、好きなカードを出すことができる。場にカードが出ている場合は、場のカードよりも強いカードを出すことができる。なお、場に出ているカードの出方に従う必要があり、例えば場に 3 枚のペアが出ている時は 3 枚のペアしか出すことができない。

(6) 上がり方

地域特有ルールでは、2 や 8、ジョーカーで上がることが禁止されていることがあるが、UEC 標準ルールではどのカードでも上がるることができる。

(7) しぼり

同じマークが連続して場に出された場合はしぼりが発生する。しぼりが発生すると場にでているカードのマークと同じマークのカードしか提出できなくなる。

(8) ペア

同じ数字のカードを複数枚揃えた役をペアという。この時マークは問わない。ペアを提出できる条件は以下のようになっている。

- 場に何も無い
- 場札がペアであり、以下の 2 つの条件を満たす
 - 提出するペアの数字が場に出ているカードよりも強いこと
 - 場に出ている枚数と同じ枚数によって構成されるペアであること

(9) 階段（連番）

同じマークでつながった数字のカードを 3 枚以上揃えた役を階段（連番）という。例えば、ハートの 9, 10, ジャックを持っている場合、階段として提出できる。階段として出せる枚数に制限はない。階段を提出できる条件は以下のようになっている。

- 場に何も無い
- 場札が階段であり、以下の 2 つの条件を満たす
 - 提出する階段を構成するカードのうち最小の数字が、場に出ている階段を構成するカードのうち最大の数字よりも強いこと
 - 場に出ている階段の枚数と同じ枚数によって構成される階段であること

(10) 革命

カードの強弱が逆転した状態を革命という。この時、3が最強のカードとなり、2が最弱のカードとなる。革命は、以下の条件のいずれかを満たしたときに発生する。カードを出したプレイヤーが革命を起こすかどうかを選択することはできない。

- 4枚以上のカードがペアとして同時に出された時
- 5枚以上のカードが階段として同時に出された時

この状態は、場が流れても継続し、そのゲームが終了するまで継続する。また、革命状態下で上記の条件を満たす提出がされた場合、革命は解除される。

(11) ジョーカー

ジョーカーは単独で使用する場合、常に最強のカードとして扱われる。革命が起きていない状態であれば2よりも強いものとし、革命が発生しているときは3よりも強いカードとして使うことができる。また、ジョーカーはペアや階段を構成する際、どんなカードの代わりとしても使うことができる。例えば、9のカードを2枚もっているとき、ジョーカーを加えて9の3枚組として場に提出できる。このようにあるカードの代わりとして提出した場合、ジョーカーの強さはそのカードと同等のものとなる。また、マークも指定でき、場札のマークと同じものを指定することで、しぼり状態を起こせる。

(12) スペードの3

スペードの3は、基本的に単なる3のカードである。ただし、場にジョーカーが1枚のとき、スペードの3を提出することができる。この場合ジョーカーが出される以前にしぼりが発生していたか否かは関係ない。なお、単独のジョーカーに対してスペードの3が出された場合、無条件で場は流れる。

(13) 8切り

8を含んだ複数枚出しや8を単独で提出した時に場が流れることを8切りという。ただしジョーカーを単独で使用するときに、8として使うことはできない。

(14) パスについて

自分に順番が回ってきた時、カードを出すかパス(何も提出しない)をするかを決める必要がある。もし出せるカードがあっても、出したくないときはパスをすることが可能である。また、場に提出されているカードがペアであるのに、階段を出すなどの

ルール上不正なカード提出を行うとパス扱いになる。

(15) 場が流れる

すべてのプレイヤーがパスをすると、場のカードは取り除かれ、最後にカードの提出を行ったプレイヤーが好きなカードを出す権利を獲得する。これを場が流れるという。

(16) 階級とカード交換

最初に上がったプレイヤーから順に、大富豪、富豪、平民、貧民、大貧民という序列付けを行う。また、序列に応じて次のゲームが始まる際、手札の交換が行われる。この時、大富豪は、大貧民から2枚のカードを貰い、大貧民に2枚渡す。富豪は貧民と1枚交換する。大富豪と富豪の渡すカードの選び方は任意である。逆に、大貧民は2枚、貧民は1枚一番強いカードを献上する。

(17) 得点の獲得

1回のゲームが終わる毎に、各プレイヤーは身分に応じて得点を獲得する。得点は、大富豪は5点、富豪は4点、平民は3点、貧民は2点、大貧民は1点である。

3. 提案手法

本研究では、手札や場のカード、既に出たカードなど環境による情報から目指すことができる順位を動的に推定し、推定結果に基づいたカード提出モデルの構築を行う。図1に本研究におけるカード提出モデルを示す。

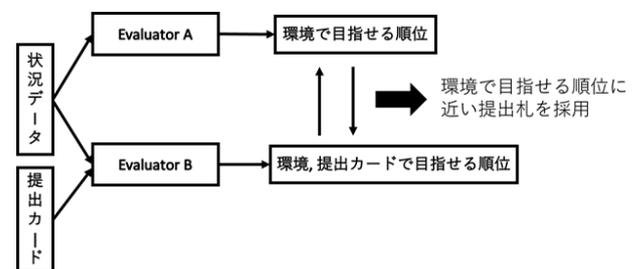


図1 カード提出モデル

図1のモデル上の Evaluator A, Evaluator B では、それぞれ Neural Network による本研究で定めたランクの推定を行っている。

3.1 ランク

本研究で推定するランクは 3 つに分けられたグループのことを指し、大富豪、富豪を富裕層、平民を庶民層、貧民と大貧民を貧困層とした。これを Evaluator A 及び Evaluator B の出力とする。

3.2 入力データ

本研究で提案するカード提出モデルでは、UEC コンピュータ大貧民大会出場エージェントである Blauweregen, jn2019, kou2, GAM, 大会サンプルエージェントの計 5 体のエージェントによる 10000 回の対戦ログデータを使用した。ログデータ上から各エージェントにおける環境データと提出カードデータの 2 つを抽出し利用した。以下でそれぞれについて説明する。

(1) 環境データ

環境データとは、あるプレイヤーが自身のターンで確認できる情報のことを指しており、手札、場のカード、既に出た札、場のしぼり、場の役、それぞれのプレイヤーの手札枚数、革命かどうかの状態のことを指す。

(2) 提出カードデータ

提出カードデータとは、あるプレイヤーが自身のターンで実際に提出したカードのことを指す。なお、最大が 1、最小が 0.5 になるように正規化を行っている。モデルの入力として、提出するカードの位置に 1 が入り、パスを行う場合は入力として全て 0.5 になる。

以下の表 1 に環境データと提出カードデータにおける要素数を示す。

表 1 要素数

手札	53
場のカード	53
既出札	53
場のしぼり	4
場の役	4
各プレイヤーの手札数	5
革命	1
提出カード	53

3.3 Evaluator A

Evaluator A の役割は環境データを入力として、ランクを推定することである。Evaluator A における学習の設定を以下の表 1 に示す。

表 1 Evaluator A のパラメーター

入力次元数	173
出力次元数	3
中間層数	1
中間ユニット数	250
活性化関数	ReLU 関数
出力次元の活性化関数	SoftMax 関数
損失関数	Cross Entropy Error
学習方法	SGD
学習率	0.01
epoch	1000

3.4 Evaluator B

Evaluator B の役割は環境データと提出カードデータを入力として、ランクを推定することである。Evaluator B は Evaluator A と同じ環境データを使用している。そこで、環境データと提出カードデータのバランスをとるためパラメーター $\alpha(0 \leq \alpha \leq 1)$ を導入し、パラメーター α と環境データの積をとり環境データの入力を小さくした。また、提出カードの最小値を 0.5 にし、常に提出カードの情報がある状態で学習を行うようにした。この時、Evaluator B における学習の設定を以下の表 2 に示す。

表 2 Evaluator B のパラメーター

入力次元数	226
出力次元数	3
中間層数	1
中間ユニット数	500
活性化関数	ReLU 関数
出力次元の活性化関数	SoftMax 関数
損失関数	Cross Entropy Error
パラメーター : α	0.06
学習方法	SGD
学習率	0.01
epoch	1000

3.5 カード提出方法

本研究で提案したカード提出モデルは、Evaluator A による環境のランク推定を行い、そのランクを目標ランクとする。そして、その状況で提出ができるカードの組み合わせ全てについて Evaluator B でランクの推定を行う。この時、Evaluator B によるランク推定結果が最も Evaluator A の推定結果に近いカードの組み合わせを提出カードとする。

4. 結果

4.1 Evaluator A

図 2 に各 epoch における損失値の推移を示す。縦軸は、損失値、横軸は epoch 数、青色実線は損失値の推移、赤色実線はバリデーションデータに対する損失値の推移を表す。

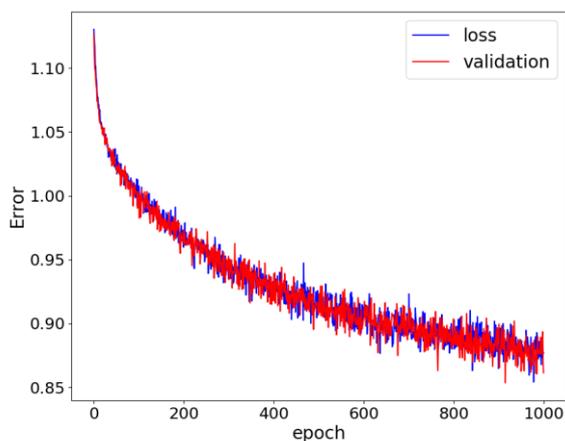


図 2 Goal Setter の損失値推移

図 3 に各 epoch におけるテストデータに対する Evaluator A の推定正解率の推移を示す。縦軸は、推定正解率、横軸は epoch、表す。

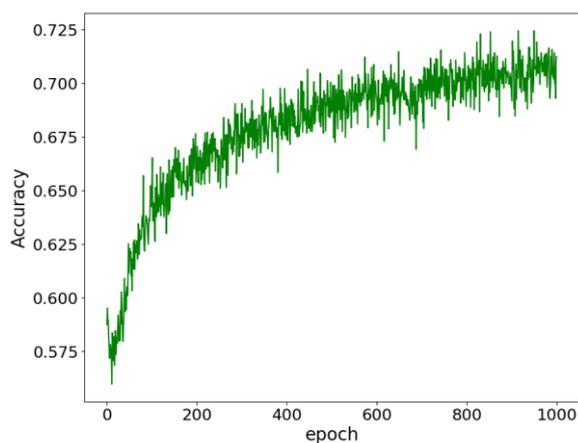


図 3 Evaluator A の正解率推移

Evaluator A では、テストデータに対して約 70%正解していることが、図 3 からわかる。

4.2 Evaluator B

図 4 に各 epoch における損失値の推移を示す。縦軸は、損失値、横軸は epoch 数、青色実線は損失値の推移、赤色実線はバリデーションデータに対する損失値の推移を表す。

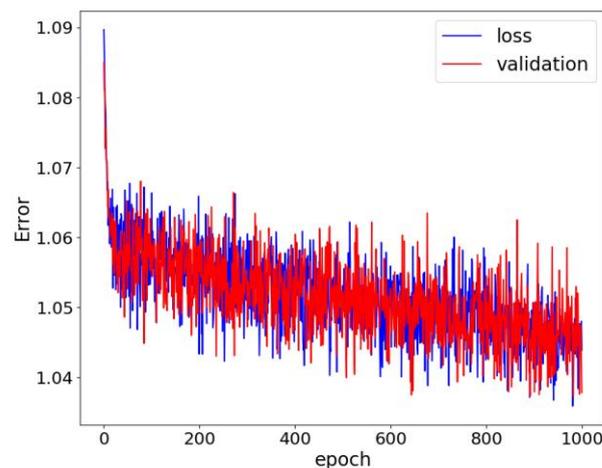


図 4 Evaluator の損失値推移

図 5 に各 epoch におけるテストデータに対する Evaluator B の推定正解率の推移を示す。縦軸は、推定正解率、横軸は epoch 数を表す。

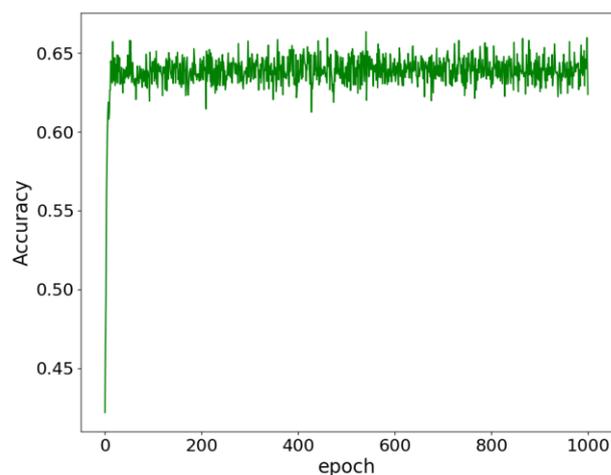


図 5 Evaluator B の正解率推移

Evaluator B では、テストデータに対して約 65%正解していることが、図 5 からわかる。

4.3 ランク推定一致率

環境データの重みを決定するパラメーター α を0.01から0.1に0.01刻み、0.1から1.0に0.1刻みで変動させ、それぞれのパラメーターを用いて学習した Evaluator B のランク推定結果と Evaluator A によるランク推定結果の一致率を調べた。これを標準一致率とした。この時、学習した Evaluator B が環境データに過度に依存した推定を行ってないか確認するために、推定時の提出カードデータの入力を0にしたときの Evaluator B のランク推定結果と Evaluator A のランク推定結果の一致率をデータ欠損一致率として求めた。データ欠損一致率は、Evaluator A と Evaluator B の類似度を表し、標準一致率は、Evaluator A による環境に対するランクの推定結果に Evaluator B がどの程度正しく提出カードデータを推定できたかどうかを表す。標準一致率とデータ欠損一致率を比較することで、Evaluator B の推定が提出カードデータの変更に対応できているどうかを確認している。

図6に、テストデータに対するランク推定一致率のグラフを示す。縦軸は推定一致率、横軸はパラメーター α の値、青の棒グラフは標準一致率、オレンジの棒グラフは、データ欠損一致率を表す。

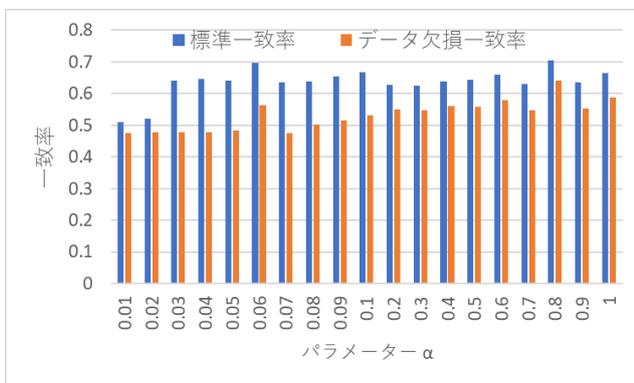


図6 ランク推定一致率

図7に図6における標準一致率とデータ欠損一致の差を表したものを示す。縦軸は一致率の差、横軸はパラメーター α の値を表す。

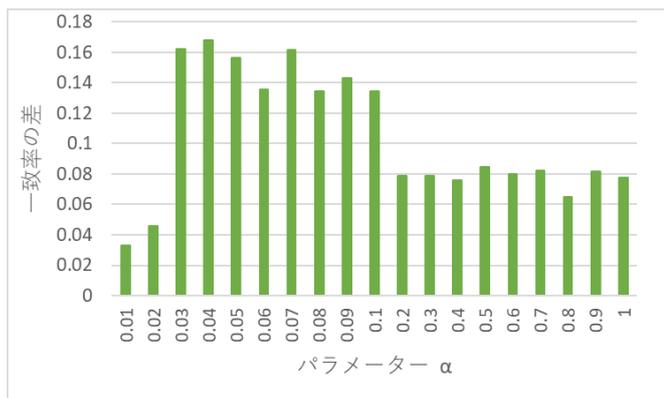


図7 推定一致率の差

図6から標準一致率が最も大きいのは α が0.8の時であるとわかる。しかし、図7から、標準一致率とデータ欠損一致率の差がほとんどないことがわかる。これは、Evaluator A と Evaluator B の類似度が大きく、提出カードデータの変更に伴う Evaluator B のランク推定結果の変化が少ないことを表している。しかし、標準一致率が2番目に大きい α が0.06の時、標準一致率とデータ欠損一致率の差は α が0.8の時よりも大きいことがわかる。これは、提出カードデータの情報が欠如したことによる影響を受けているためである。よって、 α を0.06にすることで環境データに対するランク推定を行う Evaluator A と提出カードデータに対してランク推定を行う Evaluator B の2つのモデルによって動的に目指すランクを決定し、それに伴う提出を行うことができると考えられる。

5. 今後の展望

本研究で提案したカード提出モデルは、状況に応じて動的にランク推定を行い、目指すランクを決定し、そのランクに近いカードの提出を行うモデルである。コンピュータによる面白い大貧民ゲームを設計できる。大貧民ゲームを人間がコンピュータと対戦する状況の時、本モデルの特徴を利用し、エージェントの強さをゲームの途中で適応的に変化させることで、勝てそうで勝てない、負けそうだけど勝てるといった状況を作り出すことができる。作成したエージェントを利用することで、コンピュータが弱すぎてつまらない、強すぎてつまらないという状況をなくし、プレイヤーの満足度が高い大貧民ゲームを作ることができる。そこで、本研究で提案したカード提出モデルを組み込むことで、付度を行う大貧民エージェントの開発を行うことを考えている。

参考文献

- [1] 人狼知能プロジェクト Artificial Intelligence based Werewolf, <http://www.aiwolf.org/> (2020).
- [2] UECda 運営委員会：UEC コンピュータ大貧民大会, <http://uecda.nishino-lab.jp/> (参照 2019).
- [3] 飯田伸也, 藤田 悟：大貧民におけるシミュレーション・バランスを用いた方策学習, 第77回全国大会講演論文集, 人工知能と認知科学, pp. 93-95 (2015).
- [4] 大渡勝己, 田中哲朗：方策勾配を用いた教師有り学習によるコンピュータ大貧民の方策関数の学習とモンテカルロシミュレーションへの利用, 情報処理学会研究報告, 2016-GI-35, No. 10, pp. 1-8 (2016).