

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 Construction of a Card Submission Model by Applying Neural Network

内田 純平[†] 穴田 一[†]
Junpei Uchida Hajime Anada

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、順位の低いプレイヤーが順位を大きく上げようとすると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。

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。

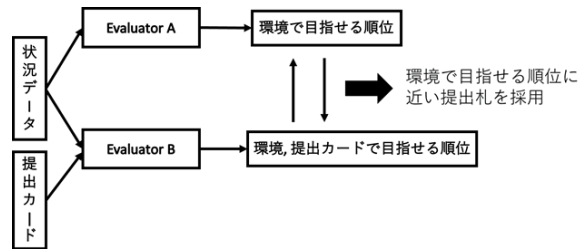


図 1 カード提出モデル

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。

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 Neural Network の学習の設定を以下の表 1 に示す。

[†] 東京都市大学大学院 総合理工学研究科 情報専攻
Tokyo City University, Department of Information Science,
Graduate School of Science and Engineering

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 α ($0 \leq \alpha \leq 1$)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 Neural Network の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

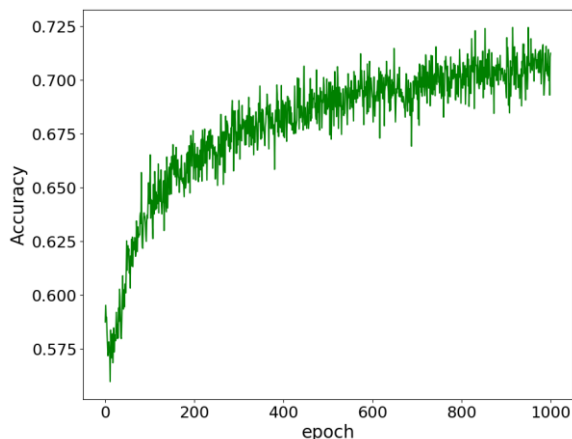


図 2 Evaluator A の正解率推移

4.2 Evaluator B

図 3 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。

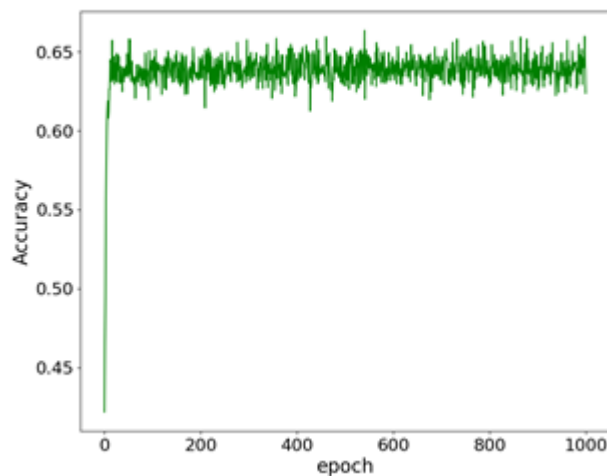


図 3 Evaluator B の正解率推移

結果の詳細と考察は発表で述べる

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・パラランシングを用いた方策学習, 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).