

多人数不完全情報ゲームにおける 状況に応じて手加減を行うゲームエージェントの構築

内田 純平[†] 穴田 一[†]

東京都市大学[†]

じた手加減を可能にした。

1. はじめに

近年, 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, 麻雀やカードゲームなどプレイヤーごとに得られる情報が部分的なゲームは不完全情報ゲームと呼ばれる. Long ら[1] は不完全情報ゲームを大きく2種類に分けている. 1つはゲームが進み徐々に情報が明らかになるトリック型ゲームである. 2つ目は, ポーカー型ゲームと呼ばれ, 勝利決定までのプロセスでは明らかな情報開示のないゲームである.

本研究では, 多人数不完全情報ゲームであり, トリック型ゲームの一種である大貧民を取り扱う. 大貧民において人間のプレイヤーがカードを提出したときの既出カード, 場のカード, 各プレイヤーの手札数などの状況データから提出カードが勝負手かどうかを判断し, その行動に応じて協力をして手加減を行えるエージェントの作成を目的としている.

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[2]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環境の評価値をニューラルネットワークで動的に推定するモデル, 環境に関する部分的な情報と提出カードから提出カードの評価値をニューラルネットワークで動的に推定するモデルを構築した. 図1にそれぞれのモデルを示す.

エージェントは人間のプレイヤーの提出カードの順位に応じた評価値と環境の順位に応じた評価値を比較することで人間のプレイヤーの提出が勝負手かどうかを判定することで状況に応

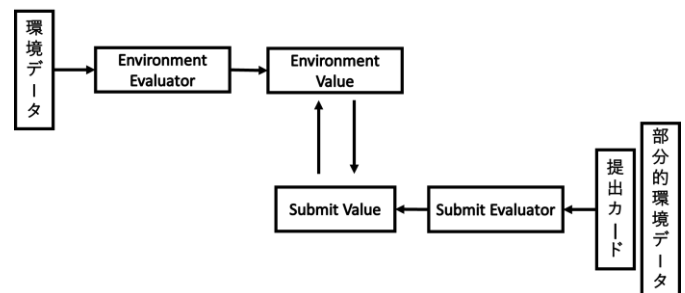


図1 評価値推定モデル

図1のモデル上の Environment Evaluator, Submit Evaluator では, それぞれ本研究で定めた評価値の推定を行っている.

3.1. 評価値

本研究では, 学習データに対してその時取れた順位に応じた評価値を設定し, それを教師データとして学習している. 順位に応じた評価値とは, 大富豪と富豪を 1.0, 平民を 0.5, 貧民と大貧民を 0.0 としている. これをそれぞれのモデルの出力とする.

3.2. 入力データ

本研究で提案するモデルでは, UEC コンピュータ大貧民大会出場エージェントである Blauwerege n, 5 体による 1000 回の対戦ログデータを使用した. ログデータ上から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, 場のカード, 既に出た札, 場のしぼり, 場の役, 各プレイヤーの手札枚数, 各プレイヤーの座席, 各プレイヤーの前試合の順位, 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

Construction of a Game Agent Which can Hold Back Depending on the Situations for a Multiplayer Incomplete Information Game

Jumpei Uchida[†], Hajime Anada[†], [†]Tokyo City University Graduate School, Department of Information Science, Graduate School of Science and Engineering

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 が入る。

Submit Evaluator の入力データとして用いる部分的環境データとは、既に出たカード、各プレイヤーの手札枚数、各プレイヤーの前試合の順位、革命かどうかの情報である。

3.3. Environment Evaluator

Environment Evaluator の役割は環境データを入力として、評価値を推定することである。Environment Evaluator における Neural Network の学習の設定を以下の表 1 に示す。

表 1 Environment Evaluator のパラメーター

入力次元数	130
出力次元数	1
中間層数	1
中間ユニット数	200
活性化関数	Swish 関数
出力次元の活性化関数	Sigmoid 関数
損失関数	Mean Squared Error
学習方法	SGD
学習率	0.001
epoch	2500

3.4. Submit Evaluator

Submit Evaluator の役割は一部の環境データと提出カードデータを入力として、評価値を推定することである。Submit Evaluator における Neural Network の学習の設定を以下の表 2 に示す。

表 2 Submit Evaluator のパラメーター

入力次元数	117
出力次元数	1
中間層数	1
中間ユニット数	200
活性化関数	Swish 関数
出力次元の活性化関数	Sigmoid 関数
損失関数	Mean Squared Error
学習方法	SGD
学習率	0.001
epoch	2500

4. 結果

4.1. Environment Evaluator

図 2 に各 epoch におけるテストデータと学習データに対する Environment Evaluator の Error rate の推移を示す。縦軸は、Error rate、横軸は epoch、青線は学習データに対する推移、オレンジ線はテストデータに対する推移を表す。

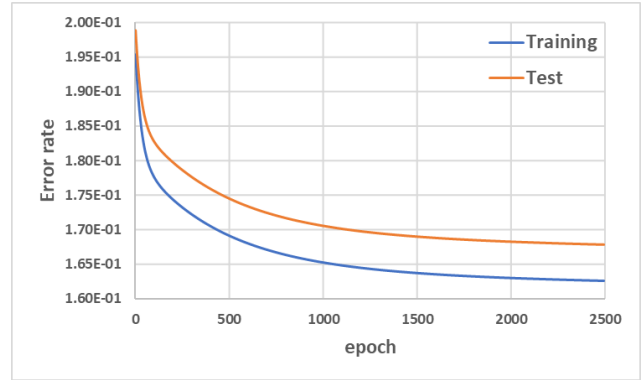


図 2 Evaluator A の正解率推移

4.2. Submit Evaluator

図 3 に各 epoch におけるテストデータと学習データに対する Environment Evaluator の Error rate の推移を示す。縦軸は、Error rate、横軸は epoch、青線は学習データに対する推移、オレンジ線はテストデータに対する推移を表す。

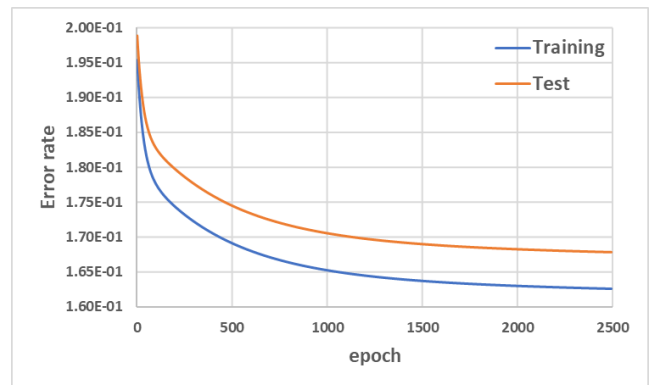


図 3 Evaluator B の正解率推移

結果の詳細と考察は発表で述べる

参考文献

- [1] Long, J., Sturtevant, N.R., Buro, M. and Furtak, T.: Understanding the Success of Perfect Information Monte Carlo Sampling in Game Tree Search, Proceeding of the 24th. AAAI conf., AAAI, pp134 – 140 (2010).
- [2] UECda 運営委員会 : UEC コンピュータ大賞民大会, <http://uecda.nishino-lab.jp/> (参照 2019)