

ガイスターゲームAIの強化学習のための ニューラルネットワークを用いた形勢評価手法

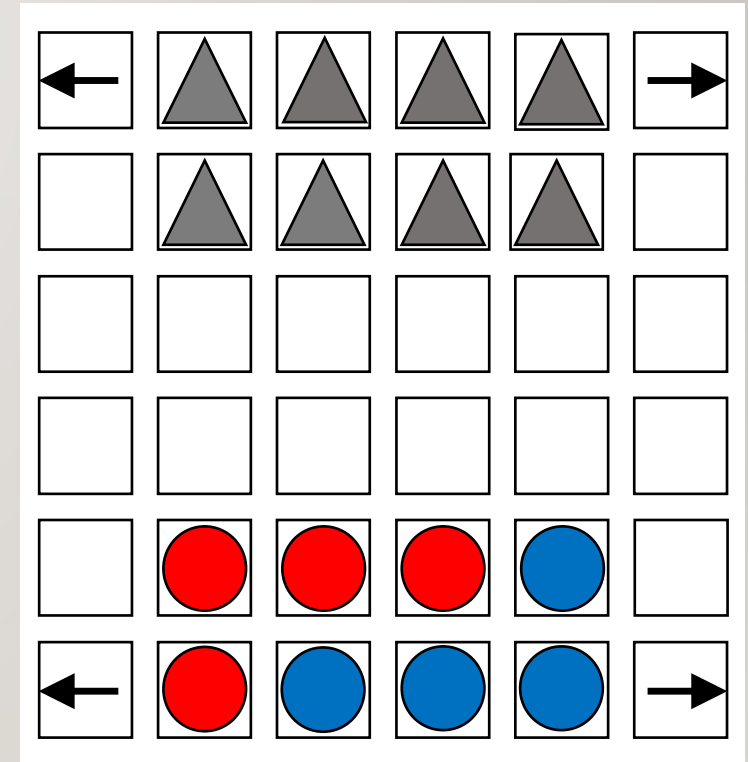
立命館大学 理工学部 電子情報工学科 5回生
動的再構成システム研究室
赤沼 慶
指導教員 泉 知論 教授

ゲームAIとガイスター大会

- 近年、人工知能と呼ばれる分野が注目されているが、中でも特に注目されているのがゲームAI
- ゲームは完全情報ゲームと不完全情報ゲームに分かれる
- 完全情報ゲームは古くから多く研究されており、AIが人間に勝つものも多い
- 本研究では不完全情報ゲームであるガイスターのゲームAIを開発する
- 情報処理学会GAT2021ガイスターAI大会に参加する

ガイスター

- ・プレイヤーは自陣に赤駒と青駒を4つずつ配置
敵駒の色は捕獲するまで分からない
- ・各プレイヤーが交互に自分の駒を上下左右に1マスずつ動かしていく

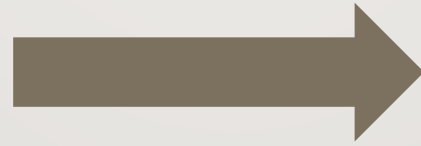


- 勝利条件
- ①敵の青駒を全て捕獲する
 - ②自分の赤駒を全て捕獲させる
 - ③自分の青駒を敵陣側の ← → から脱出させる

AIアルゴリズムの分類と先行研究

学習的アプローチ

- ・大量のプレイデータから学習



佐藤AI [1]

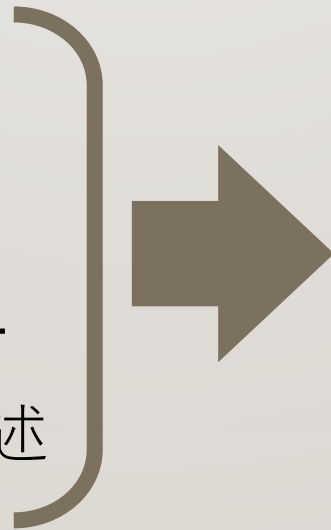
- ・ニューラルネットワークを行動価値関数として用いたもの

探索的アプローチ

- ・盤面状態を先読み

ルールベースアプローチ

- ・開発者の知識をルールベースで記述



園田AI [2]

- ・モンテカルロ法と属性推定を用いたもの

中島AI [3]

- ・Mini-Max法をガイスターゲームに適応させたもの

[1] 佐藤佑史, "ガイスターにおける自己対戦による行動価値関数の学習", 電気通信大学 情報・通信工学 専攻修士論文, Jan. 2016.

[2] 園田夕莉, "モンテカルロ法に基づく駒属性推定によるガイスターゲームAI", 立命館大学 理工学部 電子情報工学科 学士論文, Mar. 2018.

[3] 中島拓弥, "ガイスターにおける不完全情報ゲーム先読み手法の検討", 立命館大学 理工学部 電子情報工学科 学士論文, Mar. 2019.

開発の方針

- 学習的アプローチを選択
- ニューラルネットワーク(NN)を用いた状態評価関数でガイスターの盤面を評価
- 状態評価関数は、盤面情報をそのまま入力とすることで、特徴抽出から形勢評価までの学習を狙う

提案する評価関数

- 4層NN
- ノード数 $144 \times 256 \times 256 \times 1$
- 入力144変数、出力1変数
- 活性化関数はすべてシグモイド関数
- 下の式で教師値を設定、
誤差逆伝達法でNNの重みを更新する。
$$Q(s_t) \leftarrow (1 - \alpha)Q(s_t) + \alpha\{r + \max Q(s_{t+1})\}$$

行動の決定法

- ① 1ターンで可能な行動を全てリストアップ
- ② 行動後の盤面状態を状態評価関数で評価
- ③ 評価値の高い行動をとる

NNの学習

既存の4つゲームAIを対戦相手に
先攻1万回後攻1万回のテストプレイを行った。

実装環境

- ・ C言語 ・ GCC(ver9.2.0) ・ Windows 10 Home
- ・ Intel(R) Core i7-7700 ・ 8GB 2400Hz

実験

テストプレイと同じ対戦相手に各**200**戦ずつ対戦した

	提案AI勝利	提案AI敗北
園田AI [2]	25	175
中島AI [3]	56	144
貪欲脱出手法AI	3	197
貪欲捕獲手法AI	46	154

[2] 園田夕莉, “モンテカルロ法に基づく駒属性推定によるガイスターゲームAI”, 立命館大学 理工学部 電子情報工学科 学士論文, Mar. 2018.

[3] 中島拓弥, “ガイスターにおける不完全情報ゲーム先読み手法の検討”, 立命館大学 理工学部 電子情報工学科 学士論文, Mar. 2019.

まとめ

- ・ ガイスターに**NN**を用いた形勢評価関数を提案・実装
- ・ **8万回**の学習の後に、既存のゲーム**AI**と対戦
全てのゲーム**AI**に対し勝率は**30%**未満だった
- ・ **GAT**ガイスターゲーム**AI**大会に向けて
学習手法の変更や、**NN**の規模拡大などの改善を行い、
優勝を目指す