

相手の駒色予測を考慮したブラフ戦略を行うガイスターAI

佐藤 光希[†] 穴田 一[†]東京都市大学大学院 総合理工学研究科[†]

1. はじめに

近年、完全情報ゲームにおいて、AIは人間のトッププレイヤーを上回る実力を獲得している。一方、不完全情報ゲームでは初心者にさえ勝てない。また本研究で題材とするガイスターでも、ルールが単純なボードゲームにも関わらず初心者相手に勝ち越せるAIが存在しない。これは対戦相手が行ったブラフを見抜くことができないことやAIが合理的な行動しか取らず、駒色が読まれやすいということが原因と考えられる。

我々は、これまでに対戦相手が行ったブラフを見抜くための研究[1]を行った。この研究では、相手の駒色を、ゲーム上での騙し行動である駒色を偽る行動を考慮し、LSTM(Long Short-Term Memory)を用いて推定するAIを構築した。その結果、推定精度は向上し、正解率は約77%となった。それでも初心者相手に勝ち越すことができなかったのはAIが合理的な行動しかとらず、対戦相手に駒色が読まれやすいことが原因だと考えられる。そこで本研究では、我々が提案したLSTMを用いた駒色推定を自身の駒に行い、その推定結果を相手からの見え方とすることで、駒色を偽る行動をするべきか判断する方法を提案し、その有効性を検証する。

2. ガイスター

ガイスターの盤面は図1のように6×6で、盤面の四隅にはそれぞれ脱出マス(矢印は出口を表す)が存在する。各プレイヤーはそれぞれ赤駒と青駒を4つずつ、計8つ所持している。ゲーム開始時、これらの駒を相手に色が見えないように盤面手前中央の2×4のマスを自由に配置する。自分の手番では自分の駒のどれか1つを自分の駒が存在しない上下左右どれかのマスに1マス動かす。この時、相手の駒が存在するマスに自分の駒を動かして相手の駒を取ることができ、その際にその駒の色を知る。各プレイヤーは3つの勝利条件のうち、どれか1つを満たすことでゲームに勝利する。

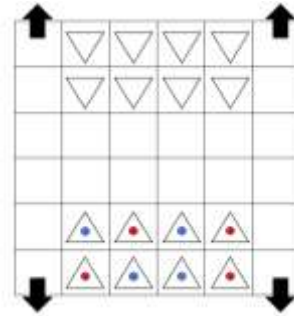


図1 ガイスターの盤面

- ・自分の青駒1つが敵側の脱出マスから脱出する
- ・相手の青駒を全て取る
- ・自分の赤駒を全て取らせる

勝利条件を満たすためには、相手の駒色を正しく推定すること、駆け引きによって相手に自分の駒色を間違えて認識させることが重要である。

3. 既存研究

ガイスターにはトッププレイヤーの棋譜が存在しないため、教師あり学習は不可能である。そこで木村らは強化学習の手法の一つであるAlphaZeroの強化学習サイクルを用いて、相手の駒色がわかる完全情報状態でガイスターAIを作成した[2]。AlphaZeroの強化学習サイクルを用いることで棋譜が無くても学習を行うことができる。しかし、実際のガイスターは相手の駒色がわからない不完全情報状態であるため、実際の試合では用いることができない。そこで我々はLSTM(Long Short-Term Memory)を用いて駒色の推定を行い、これをAlphaZeroで作成したガイスターAIに組み込んだ[1]。これにより、推定した駒色が正しいと仮定することによって、エージェントが完全情報ガイスターとしてゲームを進めることができる。LSTMを用いる理由は、相手の駒の動きの時系列データを使うことで、騙しの行動を考慮した駒色の判別が可能と考えたからである。騙しの行動をしている駒色の判別を行うために、推定内容は赤駒、青駒、青駒のふりをした赤駒、赤駒のふりをした青駒の4つとした。この際の正解率は約52.2%となった。また、この推定結果を騙しの行動が正しかったかどうかは考慮せず色のみにもとめた場

Geister AI with a Bluffing Strategy Considering the Color Pieces Predictions by the Opponent

[†]Kouki Sato, Hajime Anada

Graduate School of Tokyo City University

合の正解率は約 77%となり、最初から赤駒、青駒のみで推定した場合よりも 24%向上させることができた。しかし、このエージェントでも初心者相手に勝ち越すことができなかった。これは AI が合理的な手しか選択しないため、人間が AI の駒色を見破ることが容易であることが原因だと考えられる。

4. 提案手法

本研究では、我々が提案した LSTM を用いた駒色推定を自身の駒に行い、その推定結果を相手からの見え方とすることで駒色を偽る行動をするべきか判断する方法を提案する。

また本研究では、AlphaZero の強化学習サイクルを用いて作成したガイスターAI を使用する。既存研究[1]と同様にエージェントに LSTM を組み込む。推定内容は赤駒、青駒、青駒のふりをした赤駒、赤駒のふりをした青駒の 4 つとする。

自分の手番では、AI が動かすと決定した駒に対して LSTM の推定を行い、推定結果から駒色を偽る行動をするべきか判断する。判断基準については、分類結果が 70%以上の確率で判定していた場合に自身の駒色の認識を入れ替えてから行動することでブラフ戦略を実行する。具体的には、推定結果が赤駒、青駒、青駒のふりをした赤駒、赤駒のふりをした青駒である確率がそれぞれ 90.8%, 1.2%, 3.3%, 4.7%となった場合、正解の駒色は赤駒である場面で、推定結果は赤駒である確率が約 90%であるため、この赤駒は青駒として行動する。

5. 結果

表 1 に本研究で提案エージェントとブラフを行わないエージェントが被験者 10 人と対戦した際の試合結果を示す。エージェントの対戦順は、各エージェントが 1 試合目にくる回数が同じになるようにランダムで順番を決定した。10 人中 7 人は過去の実験にも協力してもらっており、残りの 3 人は初めてガイスターをプレイする。またそれぞれのエージェントの勝敗のつき方を表 2、表 3 で示す。

表 1 からどちらの AI も初心者相手に勝ち越していない。表 2 からブラフ戦略を行わないと駒色の予測がされやすいため、青駒を取られる負け方が一番多い。表 3 からブラフを行うことで赤駒を全て取られる勝ち方が一番多く、青駒を全て取られる負け方が減少した。

表 1. 提案エージェントの対戦結果

	ブラフなし	提案手法
勝利	3	4
敗北	7	6

表 2. ブラフを行わない場合の勝敗のつき方

勝ち方		負け方	
青駒を取る	1	青駒を取られる	5
赤駒を取られる	2	赤駒を取る	1
青駒脱出	0	青駒脱出	1

表 3. 提案 AI の勝敗のつき方

勝ち方		負け方	
青駒を取る	1	青駒を取られる	2
赤駒を取られる	3	赤駒を取る	3
青駒脱出	0	青駒脱出	1

6. おわりに

ブラフを行うことで青駒を全て取られる負け方が減少し、赤駒を全て取られる勝ち方が増えたことから、相手に駒色の推測を立てにくくさせることができたと考える。しかし、勝利回数が著しく増加したわけではない。これは対戦順が原因の一つだと考えられる。ブラフを行う AI から対戦したプレイヤーが 2 試合目であるブラフを行わない AI に対してもブラフを疑ってしまう様子が見られた。合理的な行動を取り、駒色がわかりやすいという問題点があるにも関わらず、赤駒を全て取られる勝ち方をしており、ブラフを行わない AI も勝ちやすい状況になってしまった。

また本研究では、AlphaZero の強化学習サイクルで獲得した AI によって動かす駒を決定してから、その駒にブラフ行動をさせている。しかし、これではブラフ行動の最善手とは言えないのではないかと考える。

発表では、上記の問題点を修正し、対戦人数を増やし、より詳細な結果を述べる。

参考文献

- [1] 佐藤光希, 穴田一, 駒の動きの時系列データを用いて駒色を予測するガイスターAI の構築, 第 21 回情報科学技術フォーラム(FIT2022)講演論文集 (2022)
- [2] 木村勇太, 伊藤毅志 深層強化学習を用いたガイスターAI の構築, 電気通信大学, ゲームプログラミングワークショップ論文集, vol. 2019, pp. 130-135 (2019)