

残りの駒数を考慮したブラフ戦略を行うガイスターAI

Geister AI that Performs a Bluff Strategy Considering the Number of Remaining Pieces

佐藤 光希¹ 穴田 一¹

Kouki Sato Hajime Anada

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トッププレイヤー並みの実力を持つ AI が実現されていない。これは対戦相手の情報が見えず、相手の状態や状況の予測が難しいことが原因だと考えられる。AI が不完全情報ゲームでトッププレイヤー並みの実力を持つためには、人間のように駆け引きを行う能力が必要である。

我々のこれまでの研究では、不完全情報ゲームのガイスターを用いて駆け引きを行う上で重要である相手の駒色とゲーム上での騙し行動である駒色を偽る動きをした駒について時系列データから推定する AI を構築した[1]。その結果、推定精度を向上させることができたが、まだ初心者相手に勝ち越すことはできなかった。これは AI が合理的な行動しかとらず、対戦相手に AI の持つ情報の推測がされやすいことが原因だと考えた。そこで我々は、一見不合理な行動に見えるブラフ戦略をガイスターの勝利条件を逆にした状態で学習を行うことで獲得した[2]。しかし、その戦略を有効的に活用することができていない。そこで本研究では、獲得したブラフ戦略をゲームの進行状況に合わせて活用する方法を提案し、その有効性を検証した。

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 (矢印は出口を表している) 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分からないように盤面手前中央の 2×4 のマスに自由に配置する。図 1 に初期配置の例を示す。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 マス動かさなければならない。この時、相手の駒が存在するマスに自分の駒を動かした場合、相手の駒を取り、その際にその駒の色を知る。

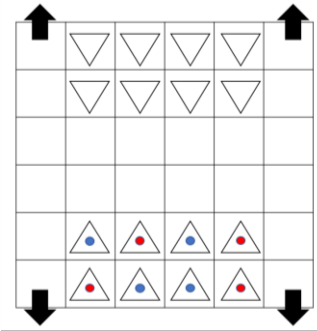


図 1 ガイスターの盤面

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。

- 自分の青駒を脱出マスから脱出させる
- 相手の青駒を全て取る
- 自分の赤駒を全て取らせる

ガイスターで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 のアルゴリズムを用いて、相手の駒色に分かる完全情報状態でガイスター AI を作成した[3]。AlphaZero のアルゴリズム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情報が分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、各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、騙しの動きをした駒の判別が可能と考えたからである。騙しの動きをしている駒の判別を行うた

¹ 東京都市大学大学院 総合理工学研究科
Graduate School of Integrative Science and Engineering,
Tokyo City University Graduate School

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。この結果、駒色を偽る動きをする駒に対しても推定することができた。

しかし、このエージェントでも初心者相手に勝つ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。そこで我々は、ガイスターの勝利条件を逆にした状態で学習を行うことでブラフ戦略を獲得した [2]。具体的には、以下の勝利条件で学習を進めた。

- 自分の赤駒を脱出マスから脱出させる
- 相手の赤駒を全て取る
- 自分の青駒を全て取らせる

この獲得したブラフ戦略を我々の研究で作成した駒色の推定を行うガイスターAI に 25%の確率で選択する形式で導入した。ブラフ戦略を行うと選択した場合でも AI はその戦略の中で合理的な行動を取るため、結果的に効果的なブラフ行動になると考えた。この結果、合理的な行動しか取らなかったこれまでの AI は出口マスに青駒を進めていくが、赤駒を進めていく戦略も選択するようになった。しかし、駒の取得状況やゲームの進行具合などを考慮していなかったため、駒色の推測はしにくくなったが、初心者相手に勝ち越すことはできないという結果が得られた。

4. 提案手法

本研究では、ガイスターの勝利条件を逆にした状態で学習を行うことで獲得したブラフ戦略を駒の取得状況やゲームの進行具合を考慮して行う手法を提案する。具体的には以下の式を用いてブラフ戦略を行う確率を決定する。

$$\frac{\log_{10}(xy^z)}{a} \quad (1)$$

ここで、 x は現在のターン数、 y と a はブラフ戦略を行う確率を調整する定数、 z は自分が取られた駒数である。これは、ゲームのターン数が進むにつれてブラフ戦略を行う確率が高くなる式を表している。また、実際のガイスターでは長期戦になる試合が少ないため、取られた駒数に応じてブラフ戦略を行う確率が高くなるようにしている。

5. 結果

提案手法で表した式のパラメータは予備実験を繰り返した結果 $y=1.2$, $a=6$ とした。このパラメータを用いて表される分布を図 2 に示す。また実際に行った試合からブラフ戦略を行った特徴的な場面を図 3 に示す。図 3 は 30 ターン目と 31 ターン目の盤面である。

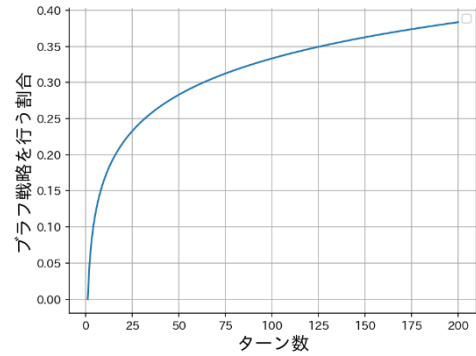


図 2 提案手法の分布グラフ

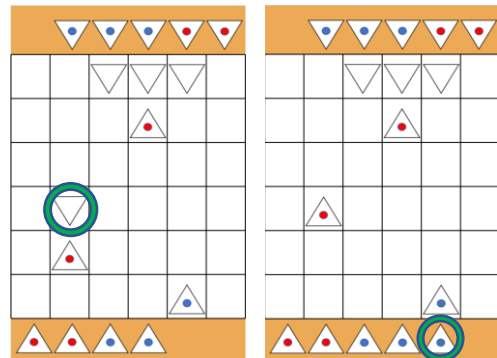


図 3 ブラフ戦略を行った場面

図 3 は緑丸で囲まれた相手の駒が取られる位置に接敵してきた場面である。AI が合理的な行動をしている場合、この駒は赤駒であるはずだが、この駒は青駒であった。この際のブラフ戦略を行う確率は 31.5%であり、ブラフ戦略を行っていたことが確認できた。

6. おわりに

ブラフ戦略を駒の取得状況やゲームの進行具合を考慮して導入することにより、ゲーム終盤により駆け引きを行うことや相手に駒色の推測を困難にさせることができた。しかし、AI 自身でブラフ戦略を行うタイミングを判断することができておらず、ブラフ戦略を行うことにより、自分が不利になる行動を取ってしまう場面が確認できた。今後は、AI 自身でブラフ戦略を行うタイミングを判断する方法を検討する。

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] 佐藤光希, 穴田一, ブラフ戦略を行うガイスターAI, 情報処理学会第 85 回全国大会講演論文集 (2023)
- [3] 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)