

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、囲碁や将棋などの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。これは対戦相手の所持している手が見えず、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の見えない情報の推測がされやすいことが原因だと考えられる。

そこで本研究では、ガイスターを用いて、一見不合理な行動に見えるブラフ戦略を学習することを目的とする。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によって不合理な戦略を獲得し、その戦略の活用方法の学習について検討する。

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。

各プレイヤーは次の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。

Bluffing Strategies Geister AI

[†] Graduate School of Integrative Science and Engineering,
Tokyo City University Graduate School

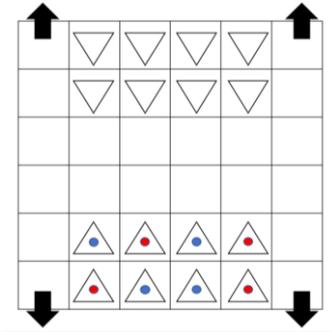


図1 ガイスターの盤面

- 自分の青駒を1つ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、相手の駒色分かる完全情報状態でガイスターAIを作成した[2]。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できる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。そこで岸野らは、AIにブラフ戦略を持たせる手法を提案した[3]。ブラフ戦略を行う駒の選択方法として、2つの手法を行っている。1つ目は決められた確率で自身の駒色に対する認識を逆にする手法である。2つ目は最初からブラフ戦略を取る駒を決めておくという手法である。ブラフ戦略をする場合はその駒を動かす際に青駒ならあえて取られるような位置に動き、赤駒なら取られない位置に動くといった駒色を偽る動きをする。その結果、どちらの手法でも初心者相手に勝ち越すことはできないが、駒色の推測はしにくくなった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ブラフ戦略を獲得する。具体的には、以下の勝利条件で学習を進める。

- 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

例えば、これらの条件で学習を行うことで、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できるようになる。このように獲得した戦略を通常のガイスターに組み込むことで不合理な戦略を取ることができる。

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**b**%の確率で選択する形式で導入する。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られる。

5. 結果

表1に学習したブラフ戦略を行うエージェントの試合結果を示す。対戦相手はルールを学んだばかりの初心者10人と対戦を行った。また勝敗のつき方の結果を表2、表3で示す。

表1から、初心者相手に勝ち越すことはできないが、ブラフ戦略の導入により勝利回数を増やすことができた。

表1. ブラフ戦略を加えたエージェントの対戦結果

	b=0(%)	b=25(%)
勝利	1	4
敗北	9	5
引き分け	0	1

表2. ブラフ戦略を行わない場合の対戦結果詳細

勝ち方		負け方	
青駒を取る	0	青駒を取られる	5
赤駒を取られる	1	赤駒を取る	1
青駒脱出	0	青駒脱出	3

表3. ブラフ戦略を25%で行う場合の対戦結果詳細

勝ち方		負け方	
青駒を取る	0	青駒を取られる	1
赤駒を取られる	4	赤駒を取る	2
青駒脱出	0	青駒脱出	2

表2からブラフ戦略を行わないと青駒を取られる負け方が一番多くなったが、これは駒色の予測がされやすいためだと考えられる。表3からブラフ戦略を行うことで赤駒を全て取られる勝ち方が多くなったが、これはブラフ戦略により駒色の予測が難しくなったことが原因だと考えられる。

6. おわりに

ブラフ戦略を導入することにより、相手の駒色の推測を困難にすることができた。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。

発表では、初心者相手との試合結果だけでなく、LSTMを用いて駒色推定を行うエージェントとの対戦結果や駒色推定の結果を出力し、より詳しくブラフ戦略を導入した有効性について述べる。

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)
- [3] 岸野圭汰, 川上直人, 橋本剛, ガイスターAIにおけるブラフ戦略の提案, 情報処理学会研究報 43, pp. 1-7 (2020)