

ガイスターにおいてプレイ履歴から駒色推定を行うアルゴリズムの研究

1235065 寺村 舞童華 【高度プログラミング研究室】

Study of a Play History Based
Hidden Information Estimation Algorithm in Geister

1235065 TERAMURA, Madoka 【High-Level Programming Lab.】

1 はじめに

ガイスターとは、アレックス・ランドルフが作成し1982年に発売された2人用ボードゲームである[1]。ガイスターは、相手の駒色が分からない状態でゲームが行われる二人零和確定不完全情報ゲームである。一般的に不完全情報ゲームにおいて隠された情報の推定を行うことは重要とされている。ガイスターでの色推定を行う研究として、モンテカルロ法を用いた研究[2][3]があり、色推定の有効性について研究されている[4]。また、機械学習を用いた研究[5]もある。

本研究では、畳み込みニューラルネットワーク(以下CNNと呼ぶ)を用いて相手の駒色を推定する方法を提案し、その有効性を実験により評価する。CNNの学習と評価には、モンテカルロ法ベースのプレイヤーのプレイ履歴を用いた。

2 本研究で用いたCNNの構成

図1は、本研究で用いるCNNの構成である。図1における n は入力で与える情報の数であり、 $Ch(k)$ はフィルタの数を表す。与えられた盤面の縦×横(6×6)と、相手駒の位置、自分の青駒の位置、自分の赤駒の位置、推定する相手駒の位置の6×6×4の形でエンコードを行う。また、相手駒の位置の駒色情報を与えた場合は6×6×5の形としてエンコードを行う。残り駒数の情報は、畳み込み層で得られる出力に対して、全結合層の手前で追加し一次元化を行った後に全結合層に入力する。また学習に用いたモデルは、フィルタサイズ2×2、畳み込み層3層、全結合層2層、ゼロパディングなしのものである。以下このモデルを学習モデルと呼ぶ。

3 評価に駒色情報を与えない場合

実際のプレイでは相手の情報は分からないため、学習時に情報がない場合について実験を行った。

総パラメータ数が30万程度になるようにパラメータを調整し、50万局面をバッチサイズ1000ごとに20エポック学習させ、学習モデルに対して評価を行った。

実験結果を表1に示す。実験結果は、90%以上の性能が出た。

実験結果から、盤面情報から学習が行われていること

表1 学習モデルの一致率

学習	評価	一致率
駒数情報なし	駒数情報なし	0.90455
駒数情報あり	駒数情報あり	0.92515

が確認できたと考察する。

4 評価に駒色情報を与えた場合

実際のプレイでは相手の隠された情報が分かることで試合を有利に進めることができる。そのため、既存のプレイから事前知識として相手の情報が分かっている状況を考え、学習時に駒色情報をどの程度与えると結果が変化するか実験を行った。

学習モデルに駒色情報をそれぞれ0%、25%、50%、75%、100%と与える。学習したモデルに対して結果がどのように変化するか評価する。

実験結果は、図2から残り駒数と駒色情報を与えたモデルの性能が高く、駒色情報のみのモデルも性能が高いことがわかる。また、情報の割合で見ると駒色情報のみでの学習では25%と50%の性能がほぼ変わらず75%のとき性能が高いが100%のときでは性能は良くない。

実験結果から、追加情報について比較すると残り駒数の情報を持つモデルの方が情報を多く持っており、学習したプレイヤーの性能も高くなっていると言える。また、与える情報の割合についてモデルの性能は、情報が多いほど性能も高くなり良くなっているように思われるが、100%のところでは性能が良くない。これは、学習時に情報を与え過ぎて学習すると与えた情報に依存してしまい、情報がなくなるとたちまち性能が落ちるといえる。

5 学習に嘘の情報を与えた場合

通常プレイヤーの相手の駒色の推定が100%正確であるとは必ずしも言えないため、学習時に事前情報として嘘の情報が含むときと含まないときで結果がどのように変化するか評価するために実験を行った。

残り駒数と駒色の情報を25%与えたものに、嘘の情報を全く含まないものと10%と30%含めたもので学習し、学習したモデルに対して評価時に嘘の情報を含めな

い場合と含めた場合について評価する実験を行った。

実験結果を図3に示す。グラフの凡例は実験結果から、一番性能が良くなると考えられる学習や評価に嘘の情報を含めない場合を除き、学習に嘘の情報を含めた場合のものの性能が良いことが分かった。また、嘘の情報を含めずに学習させたものは評価において嘘の情報を含めると性能が下がった。

実験結果から、学習において嘘の情報を含めずに学習すると、評価において嘘の情報に騙されてしまい情報の量が増えるごとにつれて性能が下がるものと考えられる。これは嘘の情報を含める割合を増やせば増やすほど性能は急に低下すると言える。その一方で、嘘の情報を含めて学習させた場合では評価において嘘の情報を含めない場合と含めた場合どちらも情報の量が増えるごとに性能は上がっており、評価に嘘の情報が含まれていない場合の方がより性能が高い。これは学習時に嘘の情報を学習することで駒色が嘘の場合と正しい場合の特徴を学習し、評価時に嘘の情報が混ざる同じ状況下で正しく推定ができたのではないかと考える。

6 まとめ

本研究では、CNNを用いたアプローチで相手の駒色を推定する実験を行った。モンテカルロ法ベースのプレイヤーのプレイ履歴を学習と評価として使用した。学習モデルを元にして残り駒数と相手の駒色の情報と嘘の情報を加えたモデルを作成し学習と評価実験を行った。実験結果より、学習モデルに駒数情報を加えることで性能を向上させることができ、駒色情報を入れるとさらなる性能の向上が見られた。また、嘘の情報を含めずに学習を行うと嘘の情報に騙されやすくなり、含めて学習すると嘘の情報と正しい情報の特徴を学習し、嘘の情報が混ざる状況かでも正しく推定ができたと思われる。

参考文献

- [1] Geister, <http://www.luding.org/cgi-bin/GameData.py/ENgameid/19686> (2013).
- [2] 三塩武徳, 小谷善行: ゲームの不完全情報推定アル

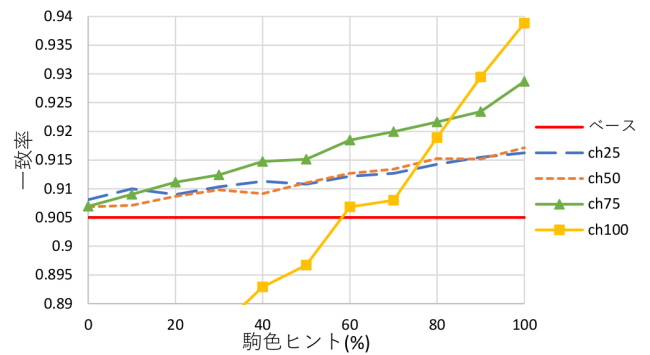


図2 学習に駒色情報を与えた場合の実験結果

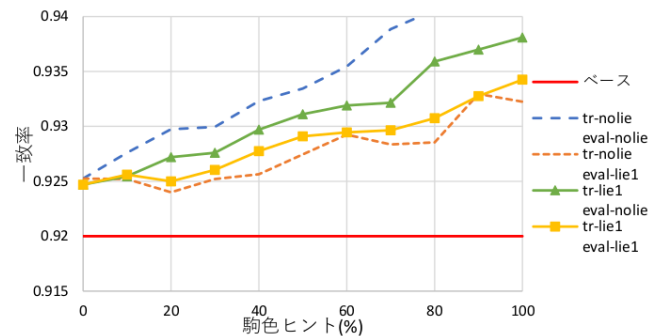


図3 学習に嘘の情報を与えた場合の実験結果

ゴリズム UPP とそのガイスターへの応用, 第31回ゲーム情報学研究会 (2014).

- [3] 鴛淵 隆斗, 佐藤 直之: ガイスターゲームにおけるモンテカルロ法を利用した駒推定およびブラフ手の生成可能性の検証, 第43回ゲーム情報学研究会 (2020).
- [4] 梶川純平, 竹内聖悟: ガイスターの初期盤面における相手駒推定の有効性, 第25回ゲームプログラミングワークショップ (20)
- [5] 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, 第24回ゲームプログラミングワークショップ (2019)

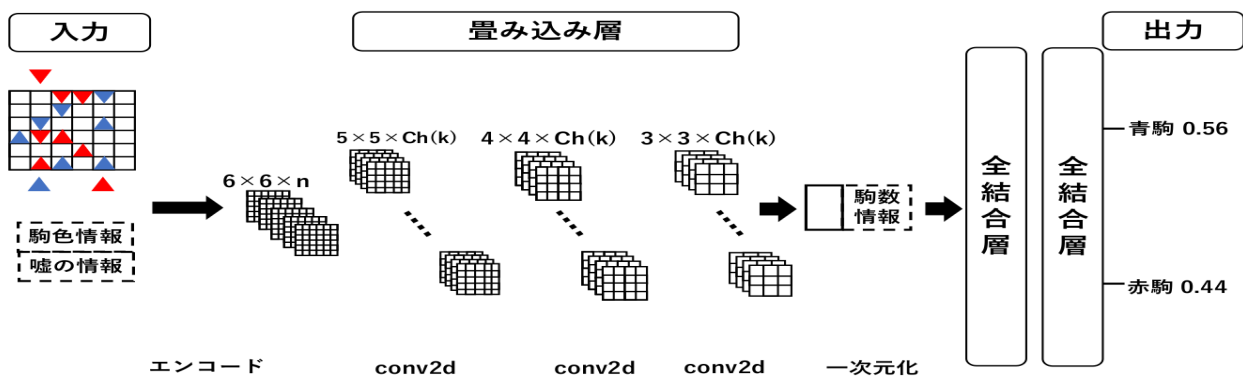


図1 CNNの構成