

科学研究費助成事業 研究成果報告書

令和 6 年 6 月 19 日現在

機関番号：26402

研究種目：基盤研究(C)（一般）

研究期間：2020～2023

課題番号：20K12124

研究課題名（和文）不完全情報・確率的・多人数ゲームにおけるAlphaGo手法の評価と改良

研究課題名（英文）Evaluation and Improvement of AlphaGo Method for Imperfect-Information, Stochastic, and Multiplayer Games

研究代表者

松崎 公紀 (Matsuzaki, Kiminori)

高知工科大学・情報学群・教授

研究者番号：30401243

交付決定額（研究期間全体）：（直接経費） 3,300,000円

研究成果の概要（和文）：本研究ではまず、AlphaGo手法で用いられる探索アルゴリズム（PUCT）について、評価関数の違いが全体の性能に及ぼす影響について「オセロ」を題材にして調査した。次に、確率的一人ゲーム「2048」、不完全情報ゲーム「ガイスター」、非対称二人ゲーム「対戦型2048」、多人数不完全情報ゲーム「DouDizhu」を題材にして、それぞれ深層強化学習によるコンピュータプレイヤー開発を行った。とくに「2048」に対しては、AlphaGo手法におけるPolicy関数の必要性の観点から、より軽量の学習手法の適用の可能性について研究した。

研究成果の学術的意義や社会的意義

AlphaGo手法（後継であるAlphaGo Zero, AlphaZero, MuZeroを含む）は、チェス・将棋・囲碁のような二人完全情報ゲームにおいて人間を超える強さのプレイヤーを実現した。本研究は、不完全情報ゲームや確率的ゲームといったより困難なゲームに対してAlphaGo手法（または一般に深層強化学習）を適用する上で遭遇する問題点をいくつか明らかにした。特に、確率的ゲーム「2048」における深層強化学習において、確率的要素が学習を悪化させることを明らかにし、その対応方法につながる課題の発見に至った。

研究成果の概要（英文）：In this study, we first investigated the impact of different evaluation functions on overall performance of the PUCT search in the AlphaGo method, using the game "Othello". Next, we developed computer players using deep reinforcement learning for various games: a stochastic single-player game "2048", an imperfect information game "Geister", an asymmetric two-player game "Two-player 2048", and a multiplayer imperfect information game "DouDizhu". In particular, for "2048", we explored the possibility of applying a more lightweight learning method from the perspective of the necessity of the Policy function in the AlphaGo method.

研究分野：ゲーム情報学

キーワード：AlphaGo 深層強化学習 モンテカルロ木探索 確率的ゲーム 不完全情報ゲーム 非対称二人ゲーム

様式 C - 19、F - 19 - 1 (共通)

1. 研究開始当初の背景

Deepmind 社による AlphaGo が囲碁のトップ棋士を破ったニュースは、ゲーム情報学の研究者だけでなく社会に広く衝撃を与えた。その後、人間の知識を必要とせず強化学習のみを用いて強いプレイヤーを開発する AlphaGo Zero や AlphaZero (本申請書では、これらを含めて AlphaGo 手法と呼ぶ) が提案され、囲碁・将棋・チェスなどの二人零和確定完全情報ゲームにおいて非常に強いプレイヤーの作成に成功した。

AlphaGo 手法の要点は、

1. ニューラルネットワークによるポリシー(方策)とバリュー(評価値)の2つの評価関数
2. 選択的に深くゲーム木を探索するモンテカルロ木探索
3. 大規模な計算リソースを利用した評価関数の強化学習

の3点に集約される。AlphaGo 手法は、原理的には二人零和確定完全情報ゲームに限ったものではなく、これまで解決が困難と考えられてきたさまざまな問題への適用が可能である。しかしながら、二人零和確定完全情報ゲームと現実の問題との間には大きなギャップが存在する。そのため、実際に AlphaGo 手法が現実のさまざまな問題に有効であるかは未知数であった。

2. 研究の目的

研究代表者の大きな研究目標は、AlphaGo 手法が、どれだけ広い範囲の問題に対して、どれだけ有効に適用できるかを明らかにすることである。本研究では、これまで研究されてきた二人零和確定完全情報ゲームから離れて、現実的な問題が持ちうる性質を持ったゲームを対象に研究を行う。具体的には、研究代表者がこれまでの研究から知見を得ている、「不完全情報」「確率的」「長手数」の特徴を有するゲームをいくつか選択する。そのようなゲームに対し、AlphaGo 手法を支える3つの技術、ニューラルネットワーク、モンテカルロ木探索、強化学習、を適用・評価する。それにより、AlphaGo 手法を現実問題に適用する際に遭遇しうる課題を発見し、その解決策を探ることが本研究の目的である。

研究代表者のグループでは、これまでに上記の特徴を有するゲームのうち「大貧民」と「2048」に対して積極的に研究に取り組んできた。特に「2048」は、「確率的」「長手数」という、現実の問題を考える上で重要な2つの性質を持つゲームである。研究代表者は、「2048」について複数の研究成果を得ており、本研究の遂行に必要なさまざまな知見を得ている。AlphaGo 手法の性質や問題に適用する際の課題についての研究は世界的に増えているが、複雑度の高い一般の問題ではなく、プレイヤーの評価・解析が比較的容易なゲームを対象として取り組むことは、本研究の重要な独自性である。

3. 研究の方法

本研究は、以下に示す3つの研究課題から構成される。

- (研究課題1) AlphaGo におけるモンテカルロ木探索 (PUCT アルゴリズム) の性質評価
- (研究課題2) AlphaGo 要素技術を用いた強いコンピュータプレイヤーの開発とその応用
- (研究課題3) 確率的ゲームに対して AlphaGo 手法を適用する際の問題点の解明

研究課題1: AlphaGo におけるモンテカルロ木探索 (PUCT アルゴリズム) の性質評価

AlphaGo 手法では、モンテカルロ木探索アルゴリズムとして、(優れた) 評価関数を利用する PUCT が用いられている。AlphaGo 手法では、ニューラルネットワークによる優れた評価関数を用いることで PUCT が高い性能を発揮したとされるが、PUCT における評価関数の影響について解明した研究はなかった。本研究では、公開された非常に強いプレイヤーを用いてほぼ正確に各手を評価できるオセロを対象として、ニューラルネットワーク評価関数を用いた場合と線形な評価関数を用いた場合とで実験により PUCT アルゴリズムの性質評価を行った。

研究課題2: AlphaGo 要素技術を用いた強いコンピュータプレイヤーの開発とその応用

AlphaGo の要素技術である「ニューラルネットワーク」「モンテカルロ木探索」「強化学習」の1つ以上を用いて、様々なゲームに対する強いコンピュータプレイヤーの開発を行う。本研究課題では、以下のゲームを主な対象とした。

- 確率的一人ゲーム「2048」: ニューラルネットワーク, モンテカルロ木探索, 強化学習
- 不完全情報ゲーム「ガイスター」: ニューラルネットワーク, 強化学習
- 非対称二人ゲーム「対戦型 2048」: ニューラルネットワーク, 強化学習
- 多人数不完全情報ゲーム「DouDizhu」: ニューラルネットワーク, 強化学習

「ガイスター」と「DouDizhu」においては、ニューラルネットワークおよび強化学習によって作られる強いプレイヤーを応用することで、それぞれ別の目標を達成することができることも示した。

研究課題 3：確率的ゲームに対して AlphaGo 手法を適用する際の問題点の解明

本来の AlphaGo 手法は非常に大きな計算リソースを要するため、より軽量な手法である Monte-Carlo Softmax Search 学習を用いて、確率的ゲームへの適用の際の問題点を解明した。Monte-Carlo Softmax Search 学習は、(ニューラルネットワーク)、モンテカルロ木探索、強化学習を併用する手法であるが、単純にアルゴリズムを適用しても学習が進まない問題が発生する。その原因が強化学習で用いるデータに含まれるバイアスであることを明らかにした。

4. 研究成果

本研究では、研究方法で述べた各研究課題に対し、以下のような成果を得た。

研究課題 1：AlphaGo におけるモンテカルロ木探索 (PUCT アルゴリズム) の性質評価

複数のニューラルネットワーク構造を設計し、それぞれ Zebra を用いて計算した (ほぼ) 正しい評価値を教師あり学習にて学習させた。そうして得られた複数のニューラルネットワーク評価関数を用いて、PUCT アルゴリズムと古典的な 枝刈り探索とを比較した。その結果、PUCT アルゴリズムにおいてシミュレーション回数を増やすと、評価関数の性能に依らず性能向上するという性質が確認された。この結果は、古典的な評価関数を用いた研究代表者らの既存研究の結果と一致しており、PUCT アルゴリズムの利点の 1 つを明らかにできたことになる。(ゲーム情報学研究会, 2020)

研究課題 2：AlphaGo 要素技術を用いた強いコンピュータプレイヤーの開発とその応用

(1) 確率的一人ゲーム「2048」: 2048 に対して、初めて、畳み込みニューラルネットワークを用いて人間と同等以上 (平均得点 9 万点以上) のプレイヤーを作成することに成功した (Journal of Information Processing, 2019)。そこで得られた知見をさらに拡張することで、ニューラルネットワークと強化学習を組み合わせ平均得点 40 万点のプレイヤーを得た (Journal of Information Processing, 2021)。このプレイヤーは、ニューラルネットワークに基づくプレイヤーの中では、その後 DeepMind の研究成果に抜かれるまで最先端であった。さらに、従来の木探索の代わりにモンテカルロ木探索を用いた場合の結果の考察を行った。これらの成果は、後述の研究課題 3 の基礎である。

(2) 不完全情報ゲーム「ガイスター」: 小さな盤面を持つ「ガイスター」の変種に対し、ニューラルネットワークと強化学習の組合せがどれだけ効果的に働くかを検証した。ここで得られた成果は、不完全情報ゲームを分類・評価するための基礎手法へと発展した (IEEE CoG, 2021)。

(3) 非対称二人ゲーム「対戦型 2048」: 対戦型 2048 でも同様に、ニューラルネットワークと強化学習によるプレイヤー開発を行った。研究代表者らが実施した既存研究よりも強いプレイヤーが開発できただけでなく、非対称の 2 プレイヤーを同時に学習すると学習が安定しないという問題点があることなどが明らかになった (プログラミング・シンポジウム, 2022)。

(4) 多人数不完全情報ゲーム「DouDizhu」: ニューラルネットワークと強化学習を用いて、DouDizhu のカードプレイフェーズの優れたプレイヤー開発を目指して研究を行った。その過程で、そこで得られたプレイヤーを利用することで、DouDizhu のビディングフェーズのプレイヤー開発が容易にできることを示し、先行研究のプレイヤーを超える性能のビディングプレイヤーを開発できた (IEEE CoG, 2023)。

研究課題 3：確率的ゲームに対して AlphaGo 手法を適用する際の問題点の解明

ゲーム 2048 に対して、ニューラルネットワーク、モンテカルロ木探索、強化学習の 3 つの要素技術を適用できることを確認した後、これら要素技術を用いる学習手法である (AlphaGo 手法より軽量な) Monte-Carlo Softmax Search (MCSS) 学習の適用を試みた。その結果、MCSS 学習を単純に適用しても学習が進まない問題が発生することを確認した (ゲームプログラミングワークショップ, 2022)。その問題が発生する理由の 1 つは Double Q 学習の手法を適用することで解決できたが、それだけでは完全には学習の問題を解決できず、確率的ゲームが持つもう 1 つの問題点について考察した (ゲーム情報学研究会, 2024)。

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 2件/うち国際共著 0件/うちオープンアクセス 1件）

1. 著者名 竹内 聖悟、栃川 純平、松崎 公紀	4. 巻 63
2. 論文標題 不完全情報ゲーム「ガイスター」における相手駒色推定の有効性評価	5. 発行年 2022年
3. 雑誌名 情報処理学会論文誌	6. 最初と最後の頁 787 ~ 795
掲載論文のDOI（デジタルオブジェクト識別子） 10.20729/00217475	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Matsuzaki Kiminori	4. 巻 29
2. 論文標題 Developing Value Networks for Game 2048 with Reinforcement Learning	5. 発行年 2021年
3. 雑誌名 Journal of Information Processing	6. 最初と最後の頁 336 ~ 346
掲載論文のDOI（デジタルオブジェクト識別子） 10.2197/ipsjip.29.336	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計13件（うち招待講演 0件/うち国際学会 4件）

1. 発表者名 小田 駿斗、松崎公紀
2. 発表標題 対戦型2048におけるニューラルネットワークプレイヤーとNタブルネットワークプレイヤーの性能比較
3. 学会等名 情報処理学会第49回ゲーム情報学研究会
4. 発表年 2023年

1. 発表者名 Shota Watanabe, Kiminori Matsuzaki
2. 発表標題 Enhancement of CNN-based 2048 Player with Monte-Carlo Tree Search
3. 学会等名 The 27th International Conference on Technologies and Applications of Artificial Intelligence (TAAI 2022) (国際学会)
4. 発表年 2022年

1. 発表者名 Runa Kubota, Lucien Troillet, Kiminori Matsuzaki
2. 発表標題 Three Player Otrio will be Strongly Solved
3. 学会等名 The 27th International Conference on Technologies and Applications of Artificial Intelligence (TAAI 2022) (国際学会)
4. 発表年 2022年

1. 発表者名 渡邊翔太, 松崎公紀
2. 発表標題 2048へのMC Softmax探索の適用
3. 学会等名 情報処理学会ゲームプログラミングワークショップ2022
4. 発表年 2022年

1. 発表者名 小田駿斗, 松崎公紀
2. 発表標題 対戦型2048 におけるニューラルネットワークプレイヤーの 探索による強化
3. 学会等名 情報処理学会ゲームプログラミングワークショップ2022
4. 発表年 2022年

1. 発表者名 Weikai Wang, Kiminori Matsuzaki
2. 発表標題 Improving DNN-based 2048 Players with Global Embedding
3. 学会等名 IEEE International Conference on Games (国際学会)
4. 発表年 2022年

1. 発表者名 荒武佑磨, 松崎公紀, 竹内聖悟
2. 発表標題 探索結果の評価値に基づく手選択の工夫による相手に合わせた将棋AI
3. 学会等名 情報処理学会第48回ゲーム情報学研究会
4. 発表年 2022年

1. 発表者名 中山友里歌, 植田桂広, 鴨川翔太, 松崎公紀
2. 発表標題 コンピュータ大貧民における手札推定の有効性についての再評価と考察
3. 学会等名 情報処理学会第48回ゲーム情報学研究会
4. 発表年 2022年

1. 発表者名 Lucien Troillet, Kiminori Matsuzaki
2. 発表標題 Analysing simplified Geister using DREAM
3. 学会等名 3rd IEEE Conference on Games (国際学会)
4. 発表年 2021年

1. 発表者名 小田 駿斗, 松崎 公紀
2. 発表標題 攻撃側が置くタイルの数を選択できる対戦型2048に対するニューラルネットワークプレイヤーの学習
3. 学会等名 第63回プログラミング・シンポジウム
4. 発表年 2022年

1. 発表者名 Lucien Troillet, Kiminori Matsuzaki
2. 発表標題 An Experimental Evaluation of PUCT Algorithm with Convolutional Neural Network Evaluation Functions
3. 学会等名 情報処理学会第44回ゲーム情報学研究会
4. 発表年 2020年

1. 発表者名 横山 智洋, 松崎 公紀
2. 発表標題 ニューラルネットワークと強化学習による対戦型2048プレイヤーの作成
3. 学会等名 情報処理学会第62回プログラミングシンポジウム
4. 発表年 2021年

1. 発表者名 寺村舞童華, 松崎公紀
2. 発表標題 ニューラルネットワークを用いたガイスターの相手駒色推定とその拡張
3. 学会等名 情報処理学会第45回ゲーム情報学研究会
4. 発表年 2021年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------	---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8 . 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------