

令和 6 年 6 月 14 日現在

機関番号：12612

研究種目：基盤研究(C)（一般）

研究期間：2020～2023

課題番号：20K12120

研究課題名（和文）将棋人工知能に関するディープマインド社の大規模実験とその知識獲得過程の検証

研究課題名（英文）DeepMind's large-scale experiment on shogi artificial intelligence and verification of its knowledge acquisition process

研究代表者

保木 邦仁（Hoki, Kunihito）

電気通信大学・大学院情報理工学研究科・准教授

研究者番号：00436081

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）：先行研究の将棋における大規模な深層強化学習の追試を、市場に出回っているハードウェアを用いて行い、学習過程や生成された人工知能の性能を観測した。報告者は追実験を行うため、グラフィックス・プロセッシング・ユニット（GPU）を活用した自己対局生成の計算効率を追求した。NVIDIA社の1個15万円程度のGPUを用いて、日産1万局程度の効率を達成した。そして、市販ハードウェアのみを用いて構築された追実装は、先行研究に準じた性能を獲得した。また、本研究において生成された大量の棋譜を整理する過程から着想を得て、ボードゲームの状態空間を疎な組合せ集合として表現する方法を新規開発するに至った。

研究成果の学術的意義や社会的意義

将棋人工知能と本研究課題で公開するプログラムを比較することにより、現在の日本における将棋人工知能技術と国際標準となったAlphaZeroの技術とを性能・コスト・プレイスタイルなどの面で比較検討することが可能となる。本研究の成果物であるプログラムコードAobaZeroは、GitHubリポジトリ「AobaZero」(<https://github.com/kobanium/aobazero>)にて公開している。インターネットの検索エンジンにて、二つのキーワード「将棋」と「AobaZero」で検索すると、これが多数のWebページで紹介されていることが確認できる。

研究成果の概要（英文）：This research conducted a follow-up test of large-scale deep reinforcement learning for shogi in previous research using hardware on the market, and observed the learning process and the performance of the generated artificial intelligence. In order to conduct follow-up experiments, the presenter pursued the computational efficiency of self-game generation using graphics processing units (GPUs). Using NVIDIA's GPU, which costs about 150,000 yen, we achieved an efficiency of about 10,000 gameplays per day. The shogi player, which was constructed using only commercially available hardware, achieved performance comparable to previous research. In addition, inspired by the process of organizing the large amount of game records generated in this research, we developed a new method to represent the state space of a board game as a sparse set of combinations.

研究分野：人工知能応用

キーワード：AlphaZero 深層学習 強化学習 将棋 グラフィカル・プロセッシング・ユニット

## 様式 C - 19、F - 19 - 1 (共通)

### 1. 研究開始当初の背景

ここ数年のゲーム人工知能の性能向上は目覚ましく、2016年の3月にグーグル・ディープマインド社のAlphaGoが囲碁プレイヤーのイ・セドルを破った出来事は記憶に新しい。このテンソル・プロセッシング・ユニット (TPU) を駆使した深層学習を伴う大規模強化学習の結果は、囲碁のみならず、他のゲーム人工知能領域にも革新的な影響を与えた。

2018年には同社の研究チームが、人工知能プレイヤーが自己対局形式で知識を獲得する強化学習アルゴリズム AlphaZero をサイエンス誌にて発表した。このアルゴリズムは囲碁以外のボードゲームにも適用可能なものであり、同発表では囲碁に加えてチェスと将棋の実験結果も示されている。

2016年のAlphaGoのイベント以来、複数の研究グループが追実験を行ってきた。特に、囲碁とチェスにおいてはコミュニティ・ベースド・プロジェクトの形態で人々により再実装が開発され続けていて、ソースコードがGitHub上で公開されている。

- ・ 囲碁 (Leela Zero) : 多数の若手プロプレイヤーによって既に利用されていて、イ・セドルを破ったグーグル・ディープマインド社の囲碁マシンに準じた性能は獲得しているとみなされている。
- ・ チェス (LCZero) : 最高水準の性能をもつチェスプログラム Stockfish をわずかに勝る性能を獲得している。

これらの再実装は市場に出回っているハードウェア上で動作可能であり、多数の人々が囲碁やチェスでAlphaZeroアルゴリズムの性質や特徴を検証している。

### 2. 研究の目的

本研究の目的は、将棋におけるAlphaZeroアルゴリズムに関して、

1. 市場に出回っているハードウェア上で動作する再実装を公開
2. 生成される人工知能プレイヤーの性能を分析
3. ランダムプレイヤーを開始点とした強化学習法によって将棋知識が獲得される過程を分析することにある。本研究課題を遂行することにより、AlphaZeroの将棋における性能やプレイスタイルが明らかとなる。

### 3. 研究の方法

応募者はこれまで追試の準備を行ってきた。この活動を通して、計算効率のさらなる追求が必須であることが分かっている。再実験で最も時間を要する計算は、グラフィックス・プロセッシング・ユニット (GPU) を活用した自己対局の生成である。応募者がもつ現在の実装と、常時利用可能なハードウェア (NVIDIA GeForce GTX 1080 Ti 32枚程度) では、自己対局の生成スピードは日産1万局程度であることに対して、AlphaZeroの将棋における強化学習実験を一度繰り返すために要する自己対局数は1千万程度である。彼らの実験を一度繰り返すだけでも数年の月日を要すうえに、AlphaZeroが確率論的アルゴリズムであることを考えるとこれを一度繰り返すだけでは十分な検証を行えそうにない。

日産16万局以上の自己対局生成能力を獲得し、1年間で追試を4~5回程度繰り返すことを可能にするため、本計画では実験の規模をより大きくする。

・ 将棋のAlphaZeroアルゴリズムの実行に特化させたコード開発 (計算のボトルネックになる部分ではCaffeやTensorFlowなどの汎用深層学習ライブラリおよびPythonなどのスクリプト言語を使用しない)。現在実装途中であるが、C++言語やOpenCLを用いた現状の独自実装は汎用ライブラリを用いた実装よりも5~6倍速くAlphaZeroアルゴリズムを実行した。最終的には8倍以上速くなると見込まれる。

・ GPUが担当する計算手順のバッチ処理 (各種GPUそれぞれの計算ユニットすべてを最大限活用するために、複数の処理を一度に行う。)

・ NVIDIA社のハードウェアに特化したコード開発 (NVIDIA社が同社ハードウェア用に高度にチューニングした数値計算ライブラリを使用する。)

である。これらの方針の有効性は囲碁やチェスの再実装において既に確認されている。

高速化達成後に、計算機実験を本格的に始動させる。また、2018年以降に公開された及びされるであろう最新の将棋人工知能と対局実験を行い、AlphaZeroの強化学習アルゴリズムが生成する将棋人工知能の性能を計測する。そして、既知の将棋の手筋や好形が出現したり消滅したりする様子を観測する。研究分担者はこれまで、ボードゲームのプレイヤーの学習過程を研究してきた。このような方法で将棋人工知能の認知過程を分析する。

### 4. 研究成果

本研究では、この先行研究の将棋における大規模実験の追試を市場に出回っているハードウェアを用いて行い、強化学習過程や生成された人工知能の性能を観測した。報告者は追実験を行うため、グラフィックス・プロセッシング・ユニット (GPU) を活用した自己対局生成の計算効率を追求した。NVIDIA社の1個15万円程度のGPU (GeForce RTX 2080) を用いて、日産1万局

程度の効率を達成した。そして、市販ハードウェアのみを用いて構築された追実装は、AlphaZero の将棋の実験に準じた性能を獲得した (山下宏、保木邦仁、小林祐樹、AobaZero の高速化と現在の状況、コンピュータ将棋協会誌、vol. 32, pp. 72-86, 2021)。

本研究の成果物であるプログラムコードは、GitHub リポジトリ「AobaZero」 (<https://github.com/kobanium/aobazero>) にて公開している。インターネットの検索エンジンにて、二つのキーワード「将棋」と「AobaZero」で検索すると、これが多数の Web ページで紹介されていることが確認できる。

また、本研究において生成された大量の棋譜を整理する過程から着想を得て、ボードゲームの状態空間を疎な組合せ集合として表現する方法を新規開発するに至った (山本敦也、保木邦仁、NOCCA × NOCCA の強解決、情報処理学会論文誌、64, 1678-1688, 2023)。

5. 主な発表論文等

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

1. 著者名 山本 敦也、保木 邦仁	4. 巻 64
2. 論文標題 NOCCA × NOCCAの強解決	5. 発行年 2023年
3. 雑誌名 情報処理学会論文誌	6. 最初と最後の頁 1678 ~ 1688
掲載論文のDOI（デジタルオブジェクト識別子） 10.20729/00231448	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 -

〔学会発表〕 計0件

〔図書〕 計0件

〔産業財産権〕

〔その他〕

- GitHub リポジトリ「AobaZero」, <a href="https://github.com/kobanium/aobazero">https://github.com/kobanium/aobazero</a> , 最終アクセス2024 - 山下宏、保木邦仁、小林祐樹、AobaZero の高速化と現在の状況、コンピュータ将棋協会誌、vol. 32, pp. 72-86, 2021
--

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
研究 分 担 者	伊藤 毅志  (Ito Takeshi)  (40262373)	電気通信大学・大学院情報理工学研究所・教授          (12612)	

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
---------	---------