

令和元年5月31日現在

機関番号：12601

研究種目：基盤研究(B) (一般)

研究期間：2016～2018

課題番号：16H02927

研究課題名(和文) 思考ゲームの並列分散探索と機械学習基盤

研究課題名(英文) Machine learning and distributed game-tree search in games

研究代表者

金子 知適 (KANEKO, TOMOYUKI)

東京大学・大学院情報学環・学際情報学府・准教授

研究者番号：00345068

交付決定額(研究期間全体)：(直接経費) 4,100,000円

研究成果の概要(和文)：アルファ碁が人工知能全体の進歩として社会に衝撃を与えたように、ゲームを題材として人工知能研究はめざましく進歩している。この研究課題では、ゲームAIの賢さを実現する機械学習(特に強化学習)とゲーム木探索という基礎技術を研究した。さらに、並列分散計算を活用した高速化にも取り組んだ。現在の強化学習の枠組みではAIが賢さを身につけるために膨大な計算時間が必要で、社会で広く使うことが難しいためである。

研究成果の学術的意義や社会的意義

研究計画の通りに機械学習とゲーム木探索とその並列実行について研究した。また採択後に大きく発展したAlphaGoに代表するGPU計算と強化学習についても、最新技術を研究に反映させた。成果の一つであるUniformity regularizationという学習方式については、tensorflow及びchainer上で実装し、囲碁、将棋、チェスなど代表的なゲームを題材に性能を示した。成果全体を総合して15件の論文を公表した。

研究成果の概要(英文)：After the success of AlphaGo, many researchers are focusing on games to develop technologies in artificial intelligence. In this research project, we have developed new techniques in machine learning (reinforcement learning) and game tree search to improve the performance of AI agents on games. Our research involves parallel and/or distributed computing to speed-up learning, because the learning requires tremendous amount of computing resources in the existing techniques in reinforcement learning.

研究分野：ゲーム情報学

キーワード：ゲームプログラミング

1. 研究開始当初の背景

思考ゲームに関する研究テーマは、実現したい目標と、研究の応用先であるさまざまなゲーム、それに基盤技術という3種類の側面から整理することができる。これまでの研究の主流は、強い思考プログラムをつくるという目標である。その成果の一つとして、将棋において、コンピュータの強さが人間のトップレベルに追いつくにいたった。その際に、中心的な役割を果たした基盤技術は高性能のゲーム木探索であり、それは機械学習と並列分散計算の両方に支えられている。さらに後で説明するように、将棋や囲碁のプログラムの強さを急速に伸ばした機械学習技術では、機械学習プロセス自身がゲーム木探索を内部に含み、より大規模な機械学習のためには並列分散計算が必要というように、複数の技術は互いに深く結びついている。

2. 研究の目的

本応募課題では、思考ゲームを題材に並列分散探索と機械学習の隣接部分について理論と実装の両面から研究を進め、汎用性のあるフレームワークを構築する。現在では広い研究テーマで大規模な機械学習が必要とされている。思考ゲームでも将棋の評価関数、コンピュータ囲碁の方策獲得から、プレイスタイルや局面解説など挑戦的なテーマにまで及ぶ。そのような大規模な学習を現実的な時間で行うために、並列分散計算を活用して、実用的な技術を開発する。

3. 研究の方法

理論と実装の両面から研究を進め、囲碁、将棋、チェスなどの代表的なゲームを題材に実際の性能を評価した。研究計画の採択後に大きく発展したAlphaGoに代表するGPU計算と強化学習についても、計画に組み込んで研究を行った。

4. 研究成果

当初の研究計画の通りに機械学習とゲーム木探索とその並列実行について、新しい手法を開発した。主要な成果の一つはUniformity regularizationというあらたな学習方式で、tensorflow及びchainer上で実装した提案手法の並列・分散環境での学習結果を、囲碁(雑誌論文3)、将棋とチェス(雑誌論文1,2)で示している。全ての研究成果は査読を経て、チェスと将棋(雑誌論文1,2,4,5,9,13,15)、囲碁(雑誌論文3)、強化学習(雑誌論文6,7)、探索(雑誌論文8,10,11,12,14)、プレイスタイル(雑誌論文13,15)などの論文として公表されている。

5. 主な発表論文等

[雑誌論文](計 15 件)

1. Shanchuan Wan and Tomoyuki Kaneko. Heterogeneous Multi-Task Learning of Evaluation Functions for Chess and Shogi. ICONIP 2018, pp. 347--358 (DOI 10.1007/978-3-030-04182-3_31, 査読有)
2. Shanchuan Wan and Tomoyuki Kaneko. Building Evaluation Functions for Chess and Shogi with Uniformity Regularization Networks. IEEE Conference on Computational Intelligence and Games 2018. doi: 10.1109/CIG.2018.8490455, 査読有
3. Yusaku Mandai and Tomoyuki Kaneko. An Alternative Multitask Training for Evaluation Functions in the Game of Go. IEEE Technologies and Applications of Artificial Intelligence, Taiwan, 2018, pp. 132--135 (DOI 10.1109/TAAI.2018.00037), 査読有
4. Taichi Nakayashiki and Tomoyuki Kaneko. Learning of Evaluation Functions via Self-Play Enhanced by Checkmate Search. IEEE Technologies and Applications of Artificial Intelligence, Taiwan, 2018, pp. 126--131 (DOI 10.1109/TAAI.2018.00036), 査読有
5. Hanhua Zhu and Tomoyuki Kaneko. Comparison of Loss Functions for Training of Deep Neural Networks in Shogi. IEEE Technologies and Applications of Artificial Intelligence, Taiwan, 2018, pp. 18--23 (DOI 10.1109/TAAI.2018.00014), 査読有
6. Hyunwoo Oh and Tomoyuki Kaneko. Deep Recurrent Q-Network with Truncated History. IEEE Technologies and Applications of Artificial Intelligence, Taiwan, 2018, pp. 34--39 (DOI 10.1109/TAAI.2018.00017), 査読有
7. Tianhe Wang and Tomoyuki Kaneko. Application of Deep Reinforcement Learning in Werewolf Game Agents. IEEE Technologies and Applications of Artificial Intelligence, Taiwan, 2018, pp. 28--33 (DOI 10.1109/TAAI.2018.00016), 査読有
8. Takahisa Imagawa and Tomoyuki Kaneko. Estimating the maximum expected value through upper confidence bound of likelihood. Technologies and Applications of Artificial Intelligence, Taiwan, 2017, pp. 92--95 (DOI:

- 10.1109/TAAI.2017.19)., 査読有
9. Shanchuan Wan and Tomoyuki Kaneko. Imitation Learning for Playing Shogi Based on Generative Adversarial Networks Technologies and Applications of Artificial Intelligence, Taiwan, 2017, pp. 202-207 (DOI: 10.1109/TAAI.2017.19), 査読有
 10. Y. Mandai and T. Kaneko: Improved LinUCT and Its Evaluation on Incremental Random-Feature Tree, IEEE Conference on Computational Intelligence and Games 2016, pp. 1-8 (DOI: 10.1109/CIG.2016.7860440), 査読有
 11. T. Imagawa and T. Kaneko: Monte Carlo Tree Search with Robust Exploration, LNCS, Computers and Games 2016. pp. 34-46. (doi: 10.1007/978-3-319-50935-8_4), 査読有
 12. 万代悠作, 金子知適: 確率的勾配降下法を用いた LinUCT のスケーラビリティの改善, 情報処理学会論文誌. 57(11), 2328 -- 2336, 2016, 査読有
 13. 大森 翔太郎, 金子 知適: 将棋における棋譜から棋風を学習するための研究, 情報処理学会論文誌. 57(11), 2328 -- 2336, 2016, 査読有
 14. Y. Mandai and T. Kaneko: LinUCB Applied to Monte Carlo Tree Search, Theoretical Computer Science. Volume 644, 6 September 2016, Pages 114-126. doi: 10.1016/j.tcs.2016.06.035, 査読有
 15. S. Omori and T. Kaneko: Learning of Evaluation Functions to Realize Playing Styles in Shogi, LNCS, PRICAI, 2016. 367-379. DOI: 10.1007/978-3-319-42911-3_31, 査読有

[学会発表](計 26 件)

1. Hanhua Zhu and Tomoyuki Kaneko. Playing the Flappy Bird with Reinforcement Learning Algorithms. The 23rd Game Programming Workshop, pp. 153-159, 2018
2. Chen Chen and Tomoyuki Kaneko. Counterfactual Regret Minimization for the Board Game Geister. The 23rd Game Programming Workshop, pp. 137-144, 2018
3. Shanchuan Wan and Tomoyuki Kaneko. Pos2Pos: Automatic Position-to-Position Translation in Chess-Like Games. The 23rd Game Programming Workshop, pp. 51-54, 2018
4. ZheJie Hu and Tomoyuki Kaneko. Reinforcement Learning with Effective Exploitation of Experiences on Mini-Games of StarCraft II. The 23rd Game Programming Workshop, pp. 168-174, 2018
5. 万代悠作, 金子知適. 囲碁ニューラルネットワークの判断根拠の可視化. 第23回ゲームプログラミングワークショップ, pp. 9-15, 2018
6. 中屋敷 太一, 金子 知適. 将棋用ニューラルネットワークへの顕著性抽出手法の適用. 第23回ゲームプログラミングワークショップ, pp. 1-8, 2018
7. 金川裕司, 金子 知適. ロードライクゲームによる強化学習ベンチマーク環境 Rogue-Gym の提案. 第23回ゲームプログラミングワークショップ, pp. 120-127, 2018
8. 王 天鶴, 金子 知適. 人狼エージェントにおける深層Qネットワークの応用. 第23回ゲームプログラミングワークショップ, pp. 16-22, 2018
9. Oh Hyunwoo, 金子 知適. LSTM の初期状態の学習による DRQN の改善. 第23回ゲームプログラミングワークショップ, pp. 220-227, 2018
10. 藤村 悠太郎, 金子 知適. 階層を考慮した模倣学習と強化学習の組み合わせ 第23回ゲームプログラミングワークショップ, pp. 145-152, 2018
11. 藤村 悠太郎, 金子 知適. 様々な学習戦略と学習環境における Hybrid Reward Architecture の性能の評価. 人工知能学会全国大会 2D4-02, 2018
12. 今川 孝久, 金子 知適. モンテカルロ木探索における状態価値の推定方法の改善. ゲームプログラミングワークショップ 2017, pp. 34-41
13. 万代 悠作, 金子 知適. 局面の組合せを用いた囲碁評価関数の学習. ゲームプログラミングワークショップ 2017, pp. 8-14
14. OH HYUNWOO, 金子 知適. GVG-AI のための Monte Carlo Tree Search の改善に関する研究. ゲームプログラミングワークショップ 2017, pp. 56-63
15. 嶽 俊太郎, 金子 知適. 強化学習を用いた評価関数の作成手法の信頼性の分析. ゲームプログラミングワークショップ 2017, pp. 250-257
16. 黄 柱皓, 金子 知適. Settlers of Catan におけるマップ自動生成の研究. ゲームプログラミングワークショップ 2017, pp. 195-200
17. Shanchuan Wan and Tomoyuki Kaneko. Style Transfer in Playing Shogi Based on Generative Adversarial Networks. ゲームプログラミングワークショップ 2017, pp. 138-143
18. 渡辺 敬介, 金子 知適. 将棋における勾配ブースティング木を用いた評価関数. ゲームプログラミングワークショップ 2017, pp. 158-162
19. 王 天鶴, 金子 知適. 人狼ゲームエージェントにおける行動選択手法の比較. ゲーム

- プログラミングワークショップ 2017, pp. 177-182
20. 藤村 悠太郎, 金子 知適. Hybrid Reward Architecture を用いたリアルタイムな意思決定の改善. ゲームプログラミングワークショップ 2017, pp. 215-218
 21. Tomoyuki Kaneko, Machine Learning of Evaluation Functions and Playing Styles in Shogi, TCGA, June 4, 2016
 22. ゲーム木に基づく並列探索での下位局面の分担. 横山秀, 金子知適. 1--5, ゲーム・プログラミング ワークショップ 2016.
 23. モンテカルロ木探索における子孫の勝敗確定時のプレイアウト結果の修正. 今川孝久, 金子知適. 13--20, ゲーム・プログラミング ワークショップ 2016. (研究奨励賞)
 24. 畳み込みネットワークによる No-Limit Hold'em の研究. 黄柱皓, 金子知適. 94--99, ゲーム・プログラミング ワークショップ 2016.
 25. 経験的知識と類似局面を用いた証明数探索の効率化. 島野拓也, 金子知適. 129--134, ゲーム・プログラミング ワークショップ 2016.
 26. UCT における Prior knowledge と方策学習を用いた個性の実現. 渡辺順哉, 金子知適. 141--148, ゲーム・プログラミング ワークショップ 2016.

6 . 研究組織

(1)研究分担者

研究分担者氏名 :

ローマ字氏名 :

所属研究機関名 :

部局名 :

職名 :

研究者番号 (8 桁) :

(2)研究協力者

研究協力者氏名 :

ローマ字氏名 :

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。