

令和 5 年 6 月 7 日現在

機関番号：14303

研究種目：基盤研究(C)（一般）

研究期間：2020～2022

課題番号：20K11988

研究課題名（和文）深層強化学習に基づく組合せ最適化法

研究課題名（英文）Combinatorial Optimizer Based on Deep Reinforcement Learning

研究代表者

飯間 等（IIMA, Hitoshi）

京都工芸繊維大学・情報工学・人間科学系・准教授

研究者番号：70273547

交付決定額（研究期間全体）：（直接経費） 3,400,000円

研究成果の概要（和文）：短時間での最適解の発見が困難であることが知られる組合せ最適化問題に対して、囲碁やテレビゲームに対して提案されている深層強化学習に基づく最適化法を開発した。また、学習の大半を最初に一度だけ実行するだけで、他の問題に対する解を短時間に得る方法の基本的な枠組みを検討した。これらの解法を配送スケジューリング問題に適用する数値実験を行い、それらの性能を明らかにした。

研究成果の学術的意義や社会的意義

深層強化学習が学習問題に対して高い性能を示すことが知られているが、これを組合せ最適化法として構成し直すことにより、従来の方法とは一線を画す解法を開発できた。この解法を既存の最適化法と比較することにより、さらに優れた最適化法を開発することが期待できる。また、優れた最適化法を開発することで、生産、物流、通信ネットワーク、金融、交通、土木、農業などの多くの分野のシステム開発に寄与できる。

研究成果の概要（英文）：We have developed optimizers based on deep reinforcement learning for solving combinatorial optimization problems, which are known to be challenging to find optimal solutions in a short time. These optimizers are classified into two types: Go and video games. We have also studied the basic framework of a method that can find a solution for a new instance in a short time after pre-training for another instance. These methods were applied to a delivery scheduling problem, and their performance was investigated experimentally.

研究分野：ソフトコンピューティング

キーワード：機械学習 深層学習 強化学習 深層強化学習 最適化 組合せ最適化

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、研究者個人に帰属します。

様式 C - 19、F - 19 - 1、Z - 19 (共通)

1. 研究開始当初の背景

よく知られているように、人工知能の中心技術の1つである深層強化学習が、囲碁、ビデオゲームなどにおいて人間を凌駕する極めて高い性能を示している。この事実は、深層強化学習が様々な分野で高い性能を発揮できることを示唆しており、深層強化学習を最適化にも利用することが考えられる。最適化技術はありとあらゆる分野で用いられている重要な技術であるが、多くの最適化問題は難解であり、短時間で最適解を求める方法が存在しない。深層強化学習で最適化問題を解くことが可能となれば、これらの分野の発展に貢献できる。

2. 研究の目的

組合せ最適化問題に対して深層強化学習で決定変数の最適値を学習させ、その学習が組合せ最適化問題を解くことに貢献するかを研究する。

3. 研究の方法

まず、ビデオゲームに対して提案されている深層強化学習に基づいた最適化法を開発する。そもそも強化学習は、学習エージェントが得る報酬の総和が最大となるような、各状態に対する学習エージェントの行動を獲得するための技術である。強化学習を適用するためには、これら状態、行動、報酬をどのように与えるのかを適切に定める必要があるため、これらを検討する。

次に、囲碁に対して提案されている深層強化学習に基づいた最適化法を開発する。囲碁に対する深層強化学習の大きな特徴は、深層強化学習だけでなく、モンテカルロ木探索も併用する点にある。そこでまずモンテカルロ木探索のみを用いて最適化を行う方法を開発する。その後、モンテカルロ木探索を深層強化学習に組み込んだ解法を開発する。

一般に、機械学習法は未知のデータに対しても正解を導く汎化性能を有している。そこで、この能力を備えた深層強化学習に基づく最適化法を検討する。すなわち、この解法では、まず多くの問題例に対する学習を事前に実行しておき、実際に解きたい他の問題例が現れたら、その問題例に対する学習を少しだけ行って解を短時間で得ようとする。

4. 研究成果

(1) ビデオゲームに対して提案されている深層強化学習に基づいた最適化法を開発した。この強化学習では、現在の解候補を状態、新しい解候補を生成するための現在の解候補の変更方法を行動とし、初期解からの目的関数値の改善度を報酬としている。また、この深層強化学習最適化法を配送スケジューリング問題に適用する際の手続きを提案した。分散深層強化学習 Ape-X を用いた数値実験を実施したところ、表1に示されるように、いくつかの問題に対しては最適解を、その他の問題に対しても下界値に近い解を求めることができることを確認した。また、この研究に関連して、配送スケジューリング問題の優れた解を考察するための最適化法や、深層学習や強化学習の基本性能を向上させる方法を検討した。

表1 深層強化学習最適化法で得られた目的関数値 (目的関数値は小さい方が良い)

問題例	1	2	3	4	5
目的関数値	5	5	10.23	8.58	7.0
下界値	5	5	9.33	7.74	6.5

(2) モンテカルロ木探索を用いて最適化を行う方法を開発した。特に、モンテカルロ木探索のシミュレーションにおいて、過去に探索した中で優れた解候補を保存しておき、この保存した解候補の情報を用いて新たな解候補を生成する手続きを導入した。ナップサック問題に適用する数値実験を実施したところ、表2に示されるように、6割以上の問題例で最適解を発見できることを確認した。また、他の最適化法である模擬アニーリングや遺伝的アルゴリズムよりも多くの問題例で最適解を発見できることを確認した。

表2 最適解が得られた問題例数の割合

問題例数の割合	モンテカルロ木探索	模擬アニーリング	遺伝的アルゴリズム
	0.66	0.14	0.04

(3) 囲碁に対して提案されている深層強化学習に基づいた最適化法を開発した。元にした深層強化学習は基本的には AlphaGoZero のものを用いるが、より優れた学習を行うことができる MuZero の手続きも一部導入している。配送スケジューリング問題に適用する数値実験を実施したところ、最適解または下界値に近い解を求めることができることを確認した。また、(1)で開発した深層強化学習最適化法より優れた解を求められる問題例があることも確認した。

表3 AlphaGoZero を基本とした深層強化学習最適化法で得られた目的関数値
(目的関数値は小さい方が良い)

問題例	1	2	3
(1)の最適化法での目的関数値	5	8.06	20.30
(3)の最適化法での目的関数値	5	8.06	18.74
下界値	5	7.74	14.10

(4) 汎化能力を有する深層強化学習最適化法を検討した。この深層強化学習におけるネットワークには問題データも入力させ、いくつかの異なる問題例に対する学習を事前にまず実施し、実際に解きたい他の問題例が現れたら、その問題例に対する学習を少しだけ行って解を求める。配送スケジューリング問題に適用する数値実験を実施したときの目的関数値の推移を図1に示す。図の横軸は行動回数すなわち解候補変更回数、縦軸は目的関数値であり、小さい方が良い。青線は(1)で開発した深層強化学習最適化法、オレンジ線はここで検討した深層強化学習最適化法である。図より、汎化能力を考慮していない(1)の解法と比較して、速く優れた解を求めることができることが確認できる。

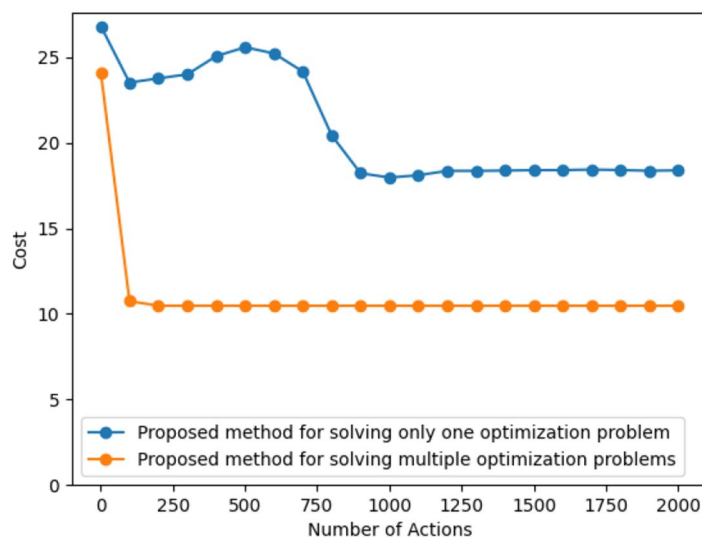


図1 目的関数値の推移

<引用文献>

V. Mnih, K. Kavukcuoglu, D. Silver, A. Rusu, J. Veness, M.G. Bellemare, A. Graves, M. Riedmiller, A.K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis, Human-level control through deep reinforcement learning, *Nature*, Vol.518, pp.529-533, 2015.

D. Horgan, J. Quan, D. Budden, G. Barth-Maron, M. Hessel, H. van Hasselt, and D. Silver, Distributed prioritized experience replay, *Proceedings of International Conference on Learning Representations*, 2018.

D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton, Y. Chen, T. Lillicrap, F. Hui, L. Sifre, G. van den Driessche, T. Graepel, and D. Hassabis, Mastering the game of Go without human knowledge, *Nature*, Vol.550, pp.354-359, 2017.

J. Schrittwieser, I. Antonoglou, T. Hubert, K. Simonyan, L. Sifre, S. Schmitt, A. Guez, E. Lockhart, D. Hassabis, T. Graepel, T. Lillicrap, and D. Silver, Mastering Atari, Go, chess and shogi by planning with a learned model, *Nature*, Vol.588, pp.604-609, 2020.

5. 主な発表論文等

〔雑誌論文〕 計6件（うち査読付論文 6件／うち国際共著 0件／うちオープンアクセス 3件）

1. 著者名 Yohei Hazama, Hitoshi Iima, Yoshiyuki Karuno and Kosuke Mishima	4. 巻 15
2. 論文標題 Genetic algorithm for scheduling of parcel delivery by drones	5. 発行年 2021年
3. 雑誌名 Journal of Advanced Mechanical Design, Systems, and Manufacturing	6. 最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） 10.1299/jamdsm.2021jamdsm0069	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 -

1. 著者名 飯間 等、兵野拓海	4. 巻 140
2. 論文標題 モンテカルロ木探索によるナップサック問題の解法	5. 発行年 2020年
3. 雑誌名 電気学会論文誌C	6. 最初と最後の頁 1141-1146
掲載論文のDOI（デジタルオブジェクト識別子） 10.1541/ieejeiss.140.1141	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3件（うち招待講演 0件／うち国際学会 3件）

1. 発表者名 Hitoshi Iima, Yohei Hazama
2. 発表標題 Genetic Algorithm with Machine Learning to Estimate the Optimal Objective Function Values of Subproblems
3. 学会等名 International Conference on Intelligent Systems, Metaheuristics & Swarm Intelligence（国際学会）
4. 発表年 2022年

1. 発表者名 Hitoshi Iima, Yoshiyuki Nakamura
2. 発表標題 Hyperheuristic Method Based on Deep Reinforcement Learning
3. 学会等名 International Congress on Advanced Applied Informatics（国際学会）
4. 発表年 2022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
---------	---------