

令和 2 年 7 月 9 日現在

機関番号：82646

研究種目：基盤研究(C) (一般)

研究期間：2017～2019

課題番号：17K00327

研究課題名(和文) 経験強化型学習と深層学習を組み合わせた新たな機械学習手法の構築に関する研究

研究課題名(英文) Research on new machine learning method combining Exploitation-oriented Learning and Deep Learning

研究代表者

宮崎 和光 (MIYAZAKI, Kazuteru)

独立行政法人大学改革支援・学位授与機構・研究開発部・准教授

研究者番号：20282866

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)：経験強化型学習XoLと深層学習を組み合わせた新たな機械学習手法としてLADQNやDPNと呼ばれる手法を提案した。特に、DPNは、Atari2600ゲーム環境において、一定の条件下で、代表的な深層強化学習手法であるDQNの1/10以下の試行錯誤回数で学習できることを示した。

さらに、深層学習と融合させたXoL手法を、自動車運転者の眠気検知、ツイートデータに基づく病気の症状判定などへ応用し、有効性を示した。これにより、従来の深層強化学習では困難であったリアルタイム性が要求される領域への試行錯誤に基づく学習の適用可能性を広げることに貢献した。

研究成果の学術的意義や社会的意義

強化学習などの試行錯誤に基づく学習は、膨大なデータの中から有効な制御則や戦略を獲得するのに適した接近法である。しかし、一般に、学習には膨大な試行錯誤回数を要するという問題がある。特に近年は、深層学習と融合した深層強化学習の登場により、今まで以上に、試行錯誤回数の削減が重要となっていた。

この問題に対し、本研究課題では、試行錯誤回数の大幅な削減を実現する手法の提案を行った。この成果は、ロボット制御などの、今まで困難であったリアルタイム性が重視される領域への深層強化学習の適用可能性を高めることにつながり、人工知能技術の応用範囲をこれまで以上に広げるものであると考える。

研究成果の概要(英文)：We have proposed new machine learning methods such as LADQN and DPN that have been combined exploitation-oriented learning XoL with deep learning. In particular, DPN can be learned by 1/10th number of trials and errors than that of DQN, a typical deep reinforcement learning method, in the Atari2600 game environment under certain conditions. In addition, we have demonstrated the effectiveness of the XoL method combined with deep learning by applying it to the detection of drowsiness in car drivers and the identification of disease symptoms based on pseudo-tweets. We believe that it has contributed to expanding the applicability of the method based on trial-and-error searches to domains that require real-time performance, which has been difficult in conventional deep reinforcement learning.

研究分野：機械学習

キーワード：強化学習 経験強化型学習 深層学習 深層強化学習 知能ロボット

様式 C - 19、F - 19 - 1、Z - 19 (共通)

1. 研究開始当初の背景

近年、深層学習が注目を集めている。中でも、Deep Q-Network (DQN) や AlphaGo など、強化学習との融合による成果も記憶に新しい。例えば DQN では、Atari2600 のゲームを題材に、代表的な強化学習器である Q-learning (QL) によるプレイ戦略の学習が行われ、一部では人間のエキスパートを凌駕するなど、良好な結果が得られている。また、AlphaGo は、2016 年 3 月に、世界戦での優勝経験を多く持つプロ棋士と対局し、4 勝 1 敗と勝ち越している。

このような深層学習と強化学習を融合させた深層強化学習の成果は、深層学習のパターン認識器としての優れた性能を、強化学習の文脈に導入したという意味で非常に重要なものである。しかし、戦略の学習に強化学習を用いていることからわかる通り、一般に、学習には膨大な数の試行錯誤が必要となる。深層学習および強化学習、それぞれが、多くの試行錯誤を要求するため、例えば、Atari2600 の場合、DQN が学習に要する試行錯誤回数は、通常、数百万単位に達する。

それに対し、研究代表者らは、得られた経験を強く強化することで試行錯誤回数の削減を指向する接近法として、経験強化型学習 Exploitation-oriented Learning (XoL) を提唱している。さらに近年では、DQN に対し、XoL の一手法である Profit Sharing (PS) を融合させたプロトタイプ的な手法である DQNwithPS を提案している。

DQNwithPS は、Atari2600 ゲーム環境中の Pong (ピンポン) と呼ばれるゲームにおいて、DQN よりも少ない試行錯誤回数で学習できることが確認されている。しかしそれ以外のゲームにおける有効性の検証は十分ではない。また、強化学習器として、QL と PS が併用されており、完全なる XoL 手法とはなっていない。そのため、試行錯誤回数の削減効果も限定的なものとなっていることが予想される。そこで、DQNwithPS から QL を除外し、深層学習と XoL との完全なる融合を果たすことができれば、より効果的な試行錯誤回数の削減が実現されると考え、本研究課題を着想するに至った。

2. 研究の目的

本研究課題の目的は、深層学習に XoL の成果を取り入れ、これまで深層強化学習が必要としていた学習に要する試行錯誤回数の大幅削減を実現することにある。例えば、DQN に対する数値目標としては、1/10 以下の試行錯誤回数での学習を掲げる。そして、提案手法を各種の問題に適用し有効性を広くアピールする。

この目的を達成するために、各年度ごとに、以下の各項目の達成を目標に研究を進める。

- (1) 平成 29 年度：DQNwithPS の問題点の解明
- (2) 平成 30 年度：深層学習と XoL の融合
- (3) 最終年度：応用例の拡充

3. 研究の方法

本研究課題では、まず、平成 29 年度に、DQNwithPS の問題点を明らかにする。これまでも DQN に PS を組み合わせた DQNwithPS により、ある程度の試行錯誤回数の削減は実現されている。しかし対象問題や実験の設定によっては、十分な効果が得られない場合がある。これには強化学習が有する報酬と罰の設計問題が深く関係していると考えている。また、これ以外にも様々な要因が、DQNwithPS の性能に影響していると思われる。そこで、DQNwithPS で可能な限り多くの実験を行い、そのような問題点の洗い出しを平成 29 年度において行う。

続く平成 30 年度には、深層学習と XoL を融合させた手法を完成させる。そのために、まず、代表的な XoL 手法である PS の性能向上を図る。具体的には、PS の基本的性能に大きな影響を与える報酬の分配方法を再考した手法の提案などを行う。さらに、平成 29 年度に得られた成果をもとに、DQNwithPS の問題点を解決する手法を完成させる。そして、最終的には、DQNwithPS から QL の機能を除外した新たな手法を提案する。

その後、最終年度である令和元年においては、平成 29 年度および 30 年度に得られた成果を踏まえ、深層学習と融合させた XoL 手法の応用例の拡充を図る。これまで、研究代表者らは、ゲーム問題を中心に提案手法の有効性を検証してきた。また、Keepaway タスク、意識的意思決定システム、および、大学の卒業認定に関する方針を題材としたテキストマイニング研究などへの各種 XoL 手法の応用を図ってきた。これらの成果を踏まえ、最終年度では、リアルタイムな意思決定が要求される「ヒトを中心に据えた応用例」の探求を行う。これは、学習に要する試行錯誤回数の大幅削減を指向する本研究課題にとって特に適した応用先になると考える。

4. 研究成果

2 で述べた研究の目的にしたがって研究成果をまとめると図 1 のようになる。以下では、順次、この図に従って研究成果を述べる。

(1) DQNwithPS の問題点を解明し、それを克服する手法を提案した。

平成 29 年度において、DQNwithPS の問題点として知られる報酬と罰の設計問題に対するひとつの解決方法として、報酬にスケジューリングを導入した手法を提案した。提案手法を、オリジナルの DQNwithPS では学習困難であった Atari2600 ゲーム環境中の Breakout (ブロック崩し) と呼ばれるゲームに適用し、DQN よりも少ない試行錯誤回数で、より高いスコアを獲得できるこ

研究の目的：

学習に要する試行錯誤回数の大幅削減を実現した手法の提案と応用例の拡充

- (1) DQNwithPSの問題点の解明 ⇒ DQNwithPSに報酬のスケジューリングを導入
Learning Acceleration DQN (LADQN)の提案
- (2) 深層学習とXoLの融合 ⇒ Deep P-Network (DPN)の提案
LADQNの改善 (Rainbowとの融合)
PSの改善 (DPSMやSPSの提案)
- (3) 応用例の拡充 ⇒ 自動車運転者の眠気検知
ツイートデータを利用した病気の症状判定

図1：研究の目的と研究成果の対応関係

とを示した。本成果は原著論文として学術雑誌に掲載された。

さらに、Atari2600 ゲーム環境を利用した様々な実験を行い、DQNwithPS が持つ新たな問題を見つけた。具体的には、DQNwithPS では学習手法としてQLとPSを併用しているが、これら2手法による学習が競合し、不適切な学習が行われる可能性があることを示した。この問題を解決するための手法として、Learning Acceleration DQN (LADQN) と呼ばれる手法を、共同研究を行っている大学院生らとともに提案した。LADQN では、QLの学習とPSの学習が競合する場合にはPSの学習を抑えることで学習性能の悪化を抑制している。これにより、DQNwithPS では学習が不安定であった Enduro (カーレース) と呼ばれるゲームにおいて、より安定的にDQNの性能を上回ることが確認できた。

これらの手法には、QLの機能が残されてはいるものの、本研究課題の目的である「DQNの学習に要する試行錯誤回数の大幅削減」に大きく寄与するものである。特に、DQNと同一の条件下で利用できるため、これまでDQNが使われていた問題に対しそのまま適用できる点が特に重要である。単純に手法を置き換えるだけで、試行錯誤回数の削減が期待できるため、利用のしやすさの観点からもインパクトが大きい成果である。

(2) 深層学習とXoLとの融合手法を完成させた。

次に、平成30年度において、DQNwithPSからQLに相当する部分を切り離した手法であるDeep P-Network (DPN) を、共同研究を行っている大学院生らとともに提案した。DPNは、QLに依存せずPSのみで学習を行うことができる初めての深層強化学習手法である。Atari2600 ゲーム環境中のPongで検証したところ、条件次第では、DQNの1/10以下の試行錯誤回数で学習できることが確認された。これは本研究課題の目的である「学習に要する試行錯誤回数の大幅な削減」を実現する重要な成果である。

さらに、平成29年度に提案したLADQNの発展も行った。具体的には、DQNの改善手法を統合したモデルとして知られるRainbowの知見を利用して、LADQNの性能を向上させた。現時点では、DPNはそのままでは罰を扱えないので、報酬と罰が混在する問題に対してはLADQNが有望である。また、既に述べたようにLADQNは、DQNをそのまま置き換えて利用できる手法なので、LADQNの性能改善は、深層強化学習の適用領域の拡大に大きく寄与するものである。

これらとともに、手法としてのPSの拡張も行った。具体的には、PSの本質である報酬の分配方法を決定する割引率に独自の視点に基づく考察を加えたDetour Path Suppression Method (DPSM) と呼ばれる手法や、新たなPSベース手法であるStable Profit Sharing (SPS) と呼ばれる手法を、共同研究を行っていた大学院生らとともに提案した。なお、DPSMについては、原著論文として学術雑誌に掲載された。DPSMやSPSは、純粋にPSの性能改善を図ったものであり、今後、これらの手法をDPNやLADQNに取り入れることで、さらなる性能改善が期待できる。

(3) 深層学習と融合させたXoL手法の応用例の拡充を図った。

最終年度では、応用例を拡充するために、まず、研究代表者らがこれまで行ってきたKeepawayタスク、意識的意思決定システム、および、大学の卒業認定に関する方針を題材としたテキストマイニング研究への各種XoL手法の適用を進めた。これらに加え、新たに、自動車運転者の眠気検知、ならびに、ツイートデータを利用した病気の症状判定といったヒトを中心に据えた応用例の探求を行うとともに、人工知能研究のテストベッドとして知られるAngry Birds AI Competitionへの参戦も行った。

これらのうち、自動車運転者の眠気検知、および、ツイートデータを利用した病気の症状判定に対しては、本研究課題の中心的手法である深層学習と融合させたXoL手法の適用を開始し、予備的実験において有効性を確認した。中でも、自動車運転者の眠気検知は迅速な判断が要求される問題のため、学習に要する試行錯誤回数の大幅削減を指向する本研究課題にとって、特に適した応用先となっている。したがって、最終年度に得られた成果は、本研究課題の目的達成に大きく寄与するものである。

現在、これらの応用例に対する提案手法のさらなる改善を検討中である。例えば、自動車運転者の眠気検知では、ヒューマノイド型ロボット「ナオ」と犬型ロボット「アイボ」を比較した予備的実験を開始しているが、本比較についても、今後さらなる実験を重ね、より広い視点から、提案手法の有効性を主張したいと考えている。

本研究課題では、当初は対象問題として、ゲーム問題を中心に取り上げ、従来手法である DQN の 1/10 以下での学習を実現してきた。それに加え、最終年度では、新たにヒトを中心に据えた応用例の拡充を図った。これは、今後、提案手法の広範囲な領域への応用を強くサポートする成果である。

このように、本研究課題で得られた成果は、今後、XoL に関する研究を推進・発展させる上で、中心的な役割を担うものばかりである。また、すべての研究成果は、国内の学会や国際会議で発表し高い評価を得るとともに、一部の成果は、原著論文として学術雑誌に掲載された。

試行錯誤に基づく学習と深層学習とを融合させた手法は、膨大なデータの中から有効な制御則や戦略を獲得するのに適した接近法である。しかし、これまでは、学習には膨大な試行錯誤回数を要していたため、一般に、リアルタイム性が重視される領域への適用は困難とされていた。

それに対し、本研究課題では、学習に要する試行錯誤回数的大幅削減を実現し、応用例を通じ、各種の提案手法の有効性を示した。この成果は、ロボット制御などの、今まで困難であったリアルタイム性が重視される領域への深層強化学習の適用可能性を高めることにつながり、人工知能技術の応用範囲をこれまで以上に広げるものであると考える。

5. 主な発表論文等

〔雑誌論文〕 計7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 6件）

1. 著者名 宮崎 和光、井田 正明	4. 巻 139
2. 論文標題 Character-level CNNを用いたディプロマ・ポリシーとカリキュラム・ポリシーの整合性判定システムの構築	5. 発行年 2019年
3. 雑誌名 電気学会論文誌C	6. 最初と最後の頁 1119 ~ 1127
掲載論文のDOI（デジタルオブジェクト識別子） https://doi.org/10.1541/ieejeiss.139.1119	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 SHIRAIISHI Daisuke, MIYAZAKI Kazuteru, KOBAYASHI Hiroaki	4. 巻 12
2. 論文標題 Proposal and Evaluation of Detour Path Suppression Method in PS Reinforcement Learning	5. 発行年 2019年
3. 雑誌名 SICE Journal of Control, Measurement, and System Integration	6. 最初と最後の頁 190 ~ 198
掲載論文のDOI（デジタルオブジェクト識別子） https://doi.org/10.9746/jcmsi.12.190	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 Miyazaki Kazuteru, Ida Masaaki	4. 巻 102
2. 論文標題 Construction of consistency judgment system of diploma policy and curriculum policy using character level CNN（雑誌論文「Character-level CNNを用いたディプロマ・ポリシーとカリキュラム・ポリシーの整合性判定システムの構築」の翻訳版）	5. 発行年 2020年
3. 雑誌名 Electronics and Communications in Japan	6. 最初と最後の頁 30 ~ 39
掲載論文のDOI（デジタルオブジェクト識別子） https://doi.org/10.1002/ecj.12223	査読の有無 無
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 KODAMA Naoki, MIYAZAKI Kazuteru, KOBAYASHI Hiroaki	4. 巻 11
2. 論文標題 Proposal and Evaluation of Reward Sharing Method Based on Safety Level	5. 発行年 2018年
3. 雑誌名 SICE Journal of Control, Measurement, and System Integration	6. 最初と最後の頁 207 ~ 213
掲載論文のDOI（デジタルオブジェクト識別子） https://doi.org/10.9746/jcmsi.11.207	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 MIYAZAKI Kazuteru	4. 巻 123
2. 論文標題 Proposal of a Deep Q-network with Profit Sharing	5. 発行年 2018年
3. 雑誌名 Procedia Computer Science	6. 最初と最後の頁 302 ~ 307
掲載論文のDOI (デジタルオブジェクト識別子) https://doi.org/10.1016/j.procs.2018.01.047	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 Miyazaki Kazuteru	4. 巻 21
2. 論文標題 Exploitation-Oriented Learning with Deep Learning - Introducing Profit Sharing to a Deep Q-Network -	5. 発行年 2017年
3. 雑誌名 Journal of Advanced Computational Intelligence and Intelligent Informatics	6. 最初と最後の頁 849 ~ 855
掲載論文のDOI (デジタルオブジェクト識別子) 10.20965/jaciii.2017.p0849	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 Miyazaki Kazuteru, Furukawa Koudai, Kobayashi Hiroaki	4. 巻 21
2. 論文標題 Proposal of PSwitchEFP and its Evaluation in Multi-Agent Reinforcement Learning	5. 発行年 2017年
3. 雑誌名 Journal of Advanced Computational Intelligence and Intelligent Informatics	6. 最初と最後の頁 930 ~ 938
掲載論文のDOI (デジタルオブジェクト識別子) doi: 10.20965/jaciii.2017.p0930	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33件 (うち招待講演 1件 / うち国際学会 8件)

1. 発表者名 Kazuteru Miyazaki
2. 発表標題 Classification of Medical Data using Character-level CNN
3. 学会等名 The 3rd International Conference on Information Science and System (ICISS 2020)
4. 発表年 2020年

1. 発表者名 Naoki Kodama, Taku Harada, Kazuteru Miyazaki
2. 発表標題 Deep Reinforcement Learning with Dual Targeting Algorithm
3. 学会等名 2019 International Joint Conference on Neural Networks (IJCNN)
4. 発表年 2019年

1. 発表者名 宮崎和光
2. 発表標題 深層強化学習を利用したドライバーの眠気防止システムに関する一考察
3. 学会等名 第47回 知能システムシンポジウム
4. 発表年 2020年

1. 発表者名 小玉直樹、原田拓、宮崎和光
2. 発表標題 Profit Sharingによる方策の直接強化手法の提案
3. 学会等名 第47回 知能システムシンポジウム
4. 発表年 2020年

1. 発表者名 宮崎和光
2. 発表標題 意識的意思決定システムへの深層強化学習の適用可能性に関する一考察
3. 学会等名 計測自動制御学会 システム・情報部門 学術講演会2019
4. 発表年 2019年

1. 発表者名 宮崎和光
2. 発表標題 経験強化型学習によるAngry Birds AI Competitionへの挑戦
3. 学会等名 計測自動制御学会 システム・情報部門 学術講演会2019
4. 発表年 2019年

1. 発表者名 宮崎和光、高橋望、森利枝
2. 発表標題 Character-level CNNを用いたディプロマ・ポリシーマッチングテストの大規模調査結果との比較
3. 学会等名 計測自動制御学会 システム・情報部門 学術講演会2019
4. 発表年 2019年

1. 発表者名 小玉直樹、原田拓、宮崎和光
2. 発表標題 経験強化型学習を用いた分散深層強化学習手法の提案
3. 学会等名 計測自動制御学会 システム・情報部門 学術講演会2019
4. 発表年 2019年

1. 発表者名 宮崎和光、高橋望、森利枝
2. 発表標題 ディプロマ・ポリシーと学位に付記する専攻分野の名称の整合性に関する研究 - 大規模調査結果の分析 -
3. 学会等名 電気学会C部門大会
4. 発表年 2019年

1. 発表者名 Kazuteru Miyazaki、Nozomi Takahashi、Rie Mori
2. 発表標題 Research on Consistency between Diploma Policies and Nomenclature of Major Disciplines : Deep Learning Approach
3. 学会等名 7th International Conference on Information and Education Technology (ICIET 2019) (国際学会)
4. 発表年 2019年

1. 発表者名 Naoki Kodama、Kazuteru Miyazaki、Taku Harada
2. 発表標題 A Proposal for Reducing the Number of Trial-and-Error Searches for Deep Q-Networks Combined with Exploitation-Oriented Learning
3. 学会等名 17th IEEE International Conference on Machine Learning and Applications (ICMLA 2018) (国際学会)
4. 発表年 2018年

1. 発表者名 Kazuteru Miyazaki、Masaaki Ida
2. 発表標題 Consistency Assessment between Diploma Policy and Curriculum Policy using Character-level CNN
3. 学会等名 Joint 10th International Conference on Soft Computing and Intelligent Systems and 19th International Symposium on Advanced Intelligent Systems (SCIS&ISIS 2018) (国際学会)
4. 発表年 2018年

1. 発表者名 Daisuke Shiraishi、Kazuteru Miyazaki、Hiroaki Kobayashi
2. 発表標題 Proposal of Detour Path Suppression Method in PS Reinforcement Learning and Its Application to Altruistic Multi-agent Environment
3. 学会等名 International Conference on Principles and Practice of Multi-Agent Systems (PRIMA 2018) (国際学会)
4. 発表年 2018年

1. 発表者名 Daisuke Mizuno、Kazuteru Miyazaki、Hiroaki Kobayashi
2. 発表標題 On Stable Profit Sharing Reinforcement Learning with Expected Failure Probability
3. 学会等名 Biologically Inspired Cognitive Architectures Meeting (BICA 2018) (国際学会)
4. 発表年 2018年

1. 発表者名 Kazuteru Miyazaki、Naoki Kodama、Hiroaki Kobayashi
2. 発表標題 Proposal and Evaluation of an Indirect Reward Assignment Method for Reinforcement Learning by Profit Sharing
3. 学会等名 IntelliSys 2018 (国際学会)
4. 発表年 2018年

1. 発表者名 小玉直樹、原田拓、宮崎和光
2. 発表標題 非ブートストラップ手法を利用した深層強化学習アルゴリズムの提案
3. 学会等名 第46回 知能システムシンポジウム
4. 発表年 2019年

1. 発表者名 宮崎和光、高橋望、森利枝
2. 発表標題 Character-level CNNを用いたディプロマ・ポリシーマッチングテスト
3. 学会等名 計測自動制御学会 システム・情報部門 学術講演会2018
4. 発表年 2018年

1. 発表者名 小玉直樹、原田拓、宮崎和光
2. 発表標題 深層強化学習アルゴリズムRainbowとProfit Sharingベース学習の結合
3. 学会等名 計測自動制御学会 システム・情報部門 学術講演会2018
4. 発表年 2018年

1. 発表者名 宮崎和光
2. 発表標題 経験強化型学習XoLに関する最近の発展
3. 学会等名 計測自動制御学会 システム・情報部門 学術講演会2018
4. 発表年 2018年

1. 発表者名 宮崎和光、井田正明
2. 発表標題 Character-level CNN を用いたディプロマポリシーとカリキュラムポリシーの整合性判定
3. 学会等名 システム研究会 インテリジェント・システム (FAN2018)
4. 発表年 2018年

1. 発表者名 小玉直樹、原田拓、宮崎和光
2. 発表標題 2つのエピソードを持つ経験強化型深層強化学習手法の提案
3. 学会等名 平成30年 電気学会 電子・情報・システム部門大会
4. 発表年 2018年

1. 発表者名 Naoki Kodama、Kazuteru Miyazaki、Hiroaki Kobayashi
2. 発表標題 Proposal of reward sharing method based on safety level and verification of its effectiveness in multi-agent environment
3. 学会等名 SICE Annual Conference 2017 (国際学会)
4. 発表年 2017年

1. 発表者名 Kazuteru Miyazaki
2. 発表標題 Proposal of a Deep Q-network with Profit Sharing
3. 学会等名 2017 Annual International Conference on Biologically Inspired Cognitive Architectures (BICA 2017) (国際学会)
4. 発表年 2017年

1. 発表者名 宮崎和光、高橋望、森利枝
2. 発表標題 学習機能を利用したディプロマ・ポリシーマッチングテストの性能改善
3. 学会等名 第45回知能システムシンポジウム
4. 発表年 2018年

1. 発表者名 小玉直樹、宮崎和光、小林博明
2. 発表標題 経験強化型学習を利用したdeep Q-networkの学習加速化手法の提案と有効性の検証
3. 学会等名 第45回知能システムシンポジウム
4. 発表年 2018年

1. 発表者名 宮崎和光
2. 発表標題 Character-level CNNを用いたテキスト分類に関する一考察
3. 学会等名 電気学会 システム研究会 (ちよだプラットフォームスクウェア 会議室504)
4. 発表年 2018年

1. 発表者名 水野大介、小林博明、宮崎和光
2. 発表標題 将来成功・失敗期待確率を用いた報酬分配型強化学習に関する研究
3. 学会等名 電気学会 システム研究会 (ちよだプラットフォームスクウェア 会議室504)
4. 発表年 2018年

1. 発表者名 宮崎和光
2. 発表標題 深層学習と強化学習 - 経験強化型学習を組み込んだ深層強化学習の評価 -
3. 学会等名 第61回システム制御情報学会研究発表講演会 (SCI ' 17) (招待講演)
4. 発表年 2017年

1. 発表者名 小玉直樹、宮崎和光、小林博明
2. 発表標題 予想失敗確率を組み込んだ新たな罰利用法の提案とマルチエージェント環境下での有効性の検証
3. 学会等名 平成29年電気学会 電子・情報・システム部門大会
4. 発表年 2017年

1. 発表者名 白石大介、宮崎和光、小林博明
2. 発表標題 Profit Sharingにおける迂回系列抑制法のマルチエージェント環境下での有効性の検証
3. 学会等名 計測自動制御学会 システム・情報部門 学術講演会2017
4. 発表年 2017年

1. 発表者名 宮崎和光、小玉直樹、小林博明
2. 発表標題 EFP利用による罰回避を実現したProfit Sharingの現状と課題
3. 学会等名 計測自動制御学会 システム・情報部門 学術講演会2017
4. 発表年 2017年

1. 発表者名 小玉直樹、宮崎和光、小林博明
2. 発表標題 経験強化型学習を組み込んだ深層強化学習DQNwithPSの改良と有効性の検証
3. 学会等名 計測自動制御学会 システム・情報部門 学術講演会2017
4. 発表年 2017年

1. 発表者名 宮崎和光、森利枝、高橋望
2. 発表標題 ディプロマ・ポリシーと学位に付記する専攻分野の名称の整合性に関する研究 - 深層学習による接近 -
3. 学会等名 電気学会 システム研究会 機械学習研究の最新動向
4. 発表年 2017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
--	---------------------------	-----------------------	----