

未来社会創造事業 探索加速型  
「超スマート社会の実現」領域  
終了報告書(探索研究)

令和元年度  
終了報告書

平成30年度採択研究開発代表者

[研究開発代表者名：潮 俊光]

[大阪大学大学院基礎工学研究科・教授]

[研究開発課題名：形式手法を用いたデータ駆動階層型管理システムの設計]

実施期間：平成30年11月15日～令和2年3月31日

## § 1. 研究実施体制

### (1)「潮」グループ(大阪大学)

- ① 研究開発代表者:潮 俊光 (大阪大学基礎工学研究科、教授)
- ② 研究項目
  - ・時相論理仕様に対する階層型管理エージェントの強化学習
  - ・ネットワーク化制御システムの深層学習制御
  - ・ユースケースの探索(ドローン, コネクティッドカー, ローバ, ネットワーク化制御など)

### (2)「松原」グループ(神戸大学)

- ① 主たる共同研究者:松原 崇 (神戸大学システム情報学研究科、助教)
- ② 研究項目
  - ・分散強化学習の協調動作に向けた意思決定のモデル化
  - ・上位管理エージェントの指令に従う効率的な分散強化学習

## § 2. 研究実施の概要

時相論理で書かれた仕様を満たす方策を制御対象のデータから学習する階層型強化学習法を開発した。時相論理を満たす振る舞いのみを受理する一般化 Rabin オートマトンの受理条件の構造に対応させて3階層から構成される。一般化 Rabin オートマトンでは、複数の受理条件があり、いずれの受理条件も有限回しか訪問できない状態集合と無限回訪問する状態集合の対で記述されている。この受理条件の特徴を利用し、最上位層ではどの受理条件を採用するかを学習、中間層では訪問の順番の学習、そして最下位層では各時刻でのアクション選択の学習を行う。この3階層からなる学習によって学習の完全な自動化が達成できた。さらに、時相論理を受理する新しいオートマトンとして注目されている limit-deterministic Büchi automaton を使った新しい強化学習法を提案し、割引率を1に十分近く取れば時相論理仕様を満たす方策を学習できることを証明した。

ネットワークによって制御対象と制御器との間にデータ転送遅延がある場合、この遅延は確率的に変動するので、ネットワークを利用した強化学習は失敗することがあった。この問題を解決するために、転送中のデータも状態と見なした拡張状態を定義し、それを用いた強化学習法を提案した。ネットワークでの遅延に変動があっても学習できることをシミュレーションによって確認した。この研究成果のユースケースとして、コネクティッドカーによる編隊制御の制御則の学習に応用し、データ遅延の下でも前方車両との車間距離を目標値に安定化する最適制御則の学習を達成した。また、時相論理で与えられたデータ収集タスクを行う移動ロボットの最適プランニングを、ジャミング攻撃の下で学習するロボット監視システムを提案し、ジャマーの移動に適応した再プランニングが迅速に行えることをシミュレーションにより確認した。

階層的なシステムにおいては、エージェントは上位エージェント、同レイヤーのエージェント、そしてシステムの管理下でない人間エージェントとも協調動作をする必要がある。そこで、深層学習を用いて強化学習に模倣学習や逆強化学習を組み合わせることで、人間エージェントの意思決定や目的をデータ駆動的に学習してモデル化し、他のエージェントと統一的に扱える土台を開発する。そのために、強化学習と人間の模倣学習を組み合わせることで、性能と人間らしい振る舞いを両立するような学習法を提案した。離散行動空間のゲーム Gopher 及び自動車運転シミュレータ Torcs において、提案手法を検証し、人間や人間エージェントの単なる模倣よりも高い性能を得つつ、感性評価において人強化学習よりも間らしいと判定される動きを実現した。

### 【主要論文】

1. 山倉 佑馬, 潮 俊光, 線形時相論理仕様を満たす階層的制御への強化学習の応用, 電子情報通信学会和文論文誌 A 分冊, vol. J102-A, no. 10, pp. 254-263, 10 月 2019.
2. Ikumi Ichikawa and Toshimitsu Ushio, Application of reinforcement learning to adaptive control of connected vehicles, Nonlinear Theory and its Applications, IEICE, vol. 10, no. 4, pp. 443-454, October 2019, DOI:10.1587/nolta.10.443.
3. Rousslan Fernand Julien Dossa, Xinyu Lian, Hirokazu Nomoto, Takashi Matsubara, and Kuniaki Uehara, A Human-Like Agent Based on a Hybrid of Reinforcement and Imitation Learning, Proc. of The 2019 International Joint Conference on Neural Networks (IJCNN2019), Budapest, Jul. 2019.