

氏 名 (本籍)	和地 瞭良
学 位 の 種 類	博 士 (工学)
学 位 記 番 号	博 甲 第 10015 号
学位授与年月日	令和 3 年 3 月 25 日
学位授与の要件	学位規則第4条第1項該当
審 査 学 術 院	理工情報生命学院
学位論文題目	Safe Reinforcement Learning for Reliable Systems (高信頼性システムのための安全性を考慮した強化学習)
主 査	筑波大学 教授 博士 (工学) 佐久間 淳
副 査	筑波大学 教授 理学博士 北川 博之
副 査	筑波大学 教授 博士 (工学) 福井 和広
副 査	筑波大学 准教授 博士 (工学) 秋本 洋平
副 査	筑波大学 准教授 博士 (情報理工学) 馬場 雪乃
副 査	九州工業大学 准教授 博士 (工学) 長 隆之

論 文 の 要 旨

審査対象論文は、累積報酬（効用）を最大化する政策の最適化を目指す強化学習アルゴリズムについて、エージェントの安全性の保証という観点から一連の研究を行っている。

第一章では、強化学習エージェントにおける安全性保証の問題についてその背景と研究の意義を論じるとともに、この論文における貢献をまとめている。

第二章では、この論文の背景となる強化学習の数理的基礎を導入している。

第三章では、強化学習エージェントの目的関数である報酬と、その一連の行動における安全性の両方を理論的に保証するアルゴリズムを提案している。提案手法は、学習段階だけでなく推論段階においても、エージェントが安全な行動のみを実行しながら、ほぼ最適に近い政策を達成することが高い確率で保証されることを示している。

第四章では、強化学習はソフトウェア構築だけでなく、ソフトウェア検証にも有用であるとの観点から、マルチエージェント敵対的強化学習を用いたエージェント検証手法を提案している。提案法は、敵対的強化学習によって訓練されたエージェントを用いて、(RL ベースまたはルールベースの) エージェントを検証する方法である。

審 査 の 要 旨

【批評】

博士論文では、全体に渡り、強化学習の手法を活用し、安全性が要求されるアプリケーションにおいて自律エージェントを適切に動作させるための方策を構築し、検証する方法について議論している。第三章に記載された研究では、制約のある MDP において安全性を保証しつつ、最適に近い累積報酬を達成する強化学習の問題において、高い確率で、安全性を理論的に保証しつつ累積報酬に関してはほぼ最適に近い政策が得られることを証明しており、また計算機実験によってその有効性を示していることから、強化学習エージェントの安全性保証において有用な提案であると言える。

第四章に記載された研究では、自動運転のためのエージェントを対象とし、敵対的強化学習によって訓練された複数のエージェントを用いて、ルールベースのエージェントを挙動を検証する方法を提案しており、実用性の高い提案であると言える。

総じて、累積報酬を最大化する方策の最適化を目指す強化学習アルゴリズムにおける安全性保証の問題を総合的に議論した優れた研究成果であり、博士論文として十分な内容であるといえる。

【最終試験の結果】

令和 3 年 2 月 9 日、理工情報生命学術院において、学位論文審査委員の全員出席のもと、著者に論文について説明を求め、関連事項につき質疑応答を行った。この結果とシステム情報工学研究群情報理工学位プログラムにおける達成度評価による結果に基づき、学位論文審査委員全員によって、合格と判定された。

【結論】

上記の学位論文審査ならびに最終試験の結果に基づき、著者は博士（工学）の学位を受けるに十分な資格を有するものと認める。