

Towards Efficient, Safe, and Reliable Reinforcement Learning Algorithms

李, 焯

<https://hdl.handle.net/2324/7363818>

出版情報 : Kyushu University, 2024, 博士 (情報科学) , 課程博士
バージョン :
権利関係 :



氏 名	李 焯 (リ タク)			
論 文 名	Towards Efficient, Safe, and Reliable Reinforcement Learning Algorithms (効率性、安全性及び信頼性の高い強化学習アルゴリズム)			
論文調査委員	主 査	九州大学	教授	殷 成久
	副 査	広島大学	教授	劉 少英
	副 査	九州大学	教授	趙 建軍

論 文 審 査 の 結 果 の 要 旨

本論文は、強化学習における効率性、安全性、及び信頼性の向上を目的として、サンプル効率の改善、安全なポリシー最適化アルゴリズム、既存モデルの信頼性テストを統合的に研究したものである。著者は、統計モジュールの導入やエピソードデータに基づくメトリクスを活用により効率的なアルゴリズムを提案するとともに、報酬と安全性を同時に最適化する方法を開発し、安全性に配慮した強化学習を実現している。さらに、多様なエージェントの挙動を評価し、潜在的な問題点を特定するためのスケーラブルなテストフレームワークを構築することで、信頼性向上のための新たな手法を提示した。本研究は、強化学習の理論と応用の双方において重要な知見を提供しており、その成果は情報科学分野における学術的および実践的な貢献として高く評価される。従って、本論文は博士（情報科学）の学位に値すると認める。