

論文内容の要旨

博士論文題目

Reinforcement Learning for Formal Language Instruction-Driven Robot Autonomy
(形式言語命令によって駆動されるロボットの自律性のための強化学習)

氏 名

畠中 渉

(論文内容の要旨)

工場やプラントなどの社会インフラにおいて、移動を伴う人間の作業をロボットで自動化する需要が高まっている。形式言語の一つである線形時相論理 (Linear Temporal Logic, LTL) は、シンボリックイベントと時間演算子によって多様なタスクの仕様を表現でき、強化学習と組み合わせることで作業工程と達成基準を順守するロボットの学習を可能にする。しかし、多くの従来研究では真の環境状態が得られることを前提にしており、現実環境において適用できる場面が限られていた。

そこで本論文では、ロボットが移動しながら実世界アプリケーションを想定した LTL タスクに従うための学習フレームワークを提案する。提案手法は、LTL によって定義された様々なタスクの下で、1) タスク達成判断に含まれる不確実性による誤認識と不必要なクエリによる不確実性の増大、2) タスクの達成仕様の変更ごとに必要な再学習による学習コストの増加、そして 3) 部分観測環境におけるロボットの移動と知覚行動の同時学習の最適化困難性 の 3 つの課題に対処する。

まず、1) に対しては、判断結果によって分岐する LTL タスクに対し、不確実性によって分岐しうる LTL を信念としてモデル化することで、不確実性の違いに応じて異なる指示に従うことを可能にするための学習手法を提案する。結果として、ロボットは含まれる不確実性が低い場合にはその判断を信じ、高い場合にはより保守的で安全な行動を取る、などの信頼性の高い行動を選択することを学習することができる。2) では、LTL タスクを構成する各イベントの達成仕様空間を新たに定義し、その条件付けられた特徴抽出を学習することで、LTL とイベントの仕様空間の両方を満たすための行動を学習する手法を提案する。これにより、方策の再学習無しに多様な達成仕様に即座に対応することが可能になる。最後に 3) では、カメラなどの外界センサを持ったロボットにおいて、移動行動に対する知覚行動

の貢献を定量化することで内部報酬として利用し，増大した行動空間において効率的な探索による同時学習を可能にするアルゴリズムを提案する．

提案手法は 3 次元シミュレーション環境におけるナビゲーションシナリオとロボットアームを用いた点検シナリオにおいて有効性が評価され，本論文が LTL タスクに従うロボット学習の実世界アプリケーション適用の一端を担うことが期待される．

(論文審査結果の要旨)

本論文は、ロボットが移動を伴う作業を行う実世界アプリケーションを想定し、線形時相論理 (Linear Temporal Logic, LTL) で記述された一連のタスク指示に従う行動方策の強化学習フレームワークを提案している。LTL を用いた従来研究の多くは真の環境状態が得られるなどの非現実的な仮定を置くため、現実環境において適用できる場面が限られていた。そこで、本論文では実世界アプリケーションを志向し、LTL によって定義された様々なタスクの下で、1) タスク達成判断に含まれる不確実性による誤認識と不必要なクエリによる不確実性の増大、2) タスク達成仕様の変更ごとに必要な再学習による学習コストの増加、3) 部分観測環境におけるロボットの移動と知覚行動の同時学習の最適化困難性、の3つの課題を設定し、各課題を解決する学習手法を提案した。

1) については、判断結果によって分岐する LTL タスクに対し、不確実性に応じて分岐しうる LTL を信念としてモデル化することで、不確実性の違いに応じて異なる指示に従うことを可能にする学習手法を提案した。2) では、LTL タスクの各イベントの達成仕様空間を新たに定義し、その条件付けられた特徴抽出を学習することで、LTL とイベントの仕様空間の両方を満たす行動を学習する手法を提案した。3) では、カメラなどの外界センサを持ったロボットにおいて、移動行動に対する知覚行動の貢献を定量化することで内部報酬として利用し、増大した行動空間において効率的な探索による同時学習を可能にする手法を提案した。本研究は、LTL 記述によるタスク指示に基づく行動方策学習の枠組みを実世界アプリケーションに導入するための新しいアプローチを開拓している。個別の提案手法も、実世界応用時の課題を抽出し、実装可能な解決技術を提示したことから、新規性および有用性の観点で、高い学術的価値を有する。

公聴会では、実世界のタスクに適した LTL の設計方針や、LTL の記述可能範囲、LTL から行動を生成する利点、近年台頭した大規模言語モデルとの関係性などについて、審査委員との質疑が行われた。最終審査において、各質問への回答および論文への修正について説明が行われ、適切な対応がとられたことを審査員全員で確認した。

本論文はロボットが移動を伴う作業を行う実世界アプリケーションを想定した LTL タスクに従うための学習フレームワークの提案、およびそのロボット点検などの現実的なタスクのシミュレーションなどに対する評価実験を実施していることに鑑み、新規性および有用性の観点から一定の学術的意義があるものと評価できる。よって、本論文は博士 (工学) の学位論文として価値あるものと認める。