

論文内容の要旨

博士論文題目

Leveraging Human Risk Sensitivity for Enriching Imitation Learning

(人間のリスク鋭敏性の活用に着目した模倣学習の強化)

氏名

Oh Hanbit

(論文内容の要旨)

人間は、危険を察知し適切な行動を選択することで高度な作業スキルを発揮している。一方で、そのような人間の行動特性は、人間の教示からロボットの作業スキルを学習する模倣学習 (IL) においては考慮することが難しい。標準的な IL は、人間の実演データから状態と行動の対応関係をブラックボックス的なアルゴリズムで獲得するため、危険に鋭敏な行動を駆動する根本的な特性 (リスク鋭敏性) を見落としてしまう。このことが模倣学習の適用性と汎化性能を制限し、ロボットが人間のようなスキルを獲得することを妨げていると考えられる。そこで、本論文では、人間の行動心理学の知見に基づいて IL アルゴリズムを設計する新しいメカニズムを提案し、IL を強化するためにリスク鋭敏性を捉え、その活用に着目する。具体的には、人間のリスク鋭敏性として、人間の速度-精度トレードオフを考慮する。人間は環境の理解に基づいてリスクを感知し、状態のリスクに応じて速度を調整する (例: リスクの高い状況では遅くなる)。そのため、人間が示す速度教示データから環境のリスク情報を推定することができると考えられる。

提案メカニズムを使用した2つの適用事例を示す。1)人間のリスク鋭敏性に関連する危険予測器を導入し、オンポリシーILの枠組みの安全性を向上させる。産業用のペグ挿入タスクなど、クリアランス制限の厳しい作業では、ロボットが学習中に未熟な作業方策を用いることは衝突リスクを高める。これに対し、危険予測器を導入することで、ロボットが危険な状態に遭遇した際に人間の介入・教示を能動的に要請可能にする。こ

れによりロボットの方策は反復的に最適化され、安全な対話型模倣学習が実現される。

2)人間のリスク鋭敏性から導出される外乱注入モデルを導入することで、外乱注入型ILの実演可能性を強化する。人間の制御コマンドに人工的な外乱を注入し、拡張された実演データを収集することで、頑健な方策を学習できるが、外乱が強すぎると人間の实演可能性が低下してしまう。本論文では、環境の衝突リスクに応じて外乱のレベルを適度に調整する外乱注入モデルを提案する。これにより危険のある領域では外乱が小さく調整され、実演可能性と外乱への頑健性を両立した模倣学習が実現される。これらの事例では、様々なクリアランス制限のあるタスクのシミュレーションと実ロボット実験において他の学習方法と比較を行った。実験の結果から、提案手法によりリスク鋭敏性を取り入れることで、模倣学習の訓練プロセスの安全性が向上することを確認した。また、安全かつ効率的に収集された訓練データにより、獲得される方策のタスク性能も向上することを確認した。

(論文審査結果の要旨)

本論文では、熟練者による実演データからロボットスキル（行動ルール）を学習する模倣学習（Imitation Learning: IL）に焦点を当て、人間の行動心理学の知見に基づく新たなアプローチを検討している。とりわけ、本論文では、環境に潜む様々な危険を認識し、自身の行動選択に反映させる人間のリスク鋭敏性に着目した。具体的には、危険に応じて動作速度を調整する速度-精度トレードオフの特性に基づいて、動作速度から人間が認識するリスク情報を推定し、それを活用することで模倣学習の効率性や汎化性を高めるアプローチを提案した。

その概念検証として、二つの適用事例を示した。一つ目の事例は、リスク鋭敏性に基づく環境との衝突リスク予測器を導入し、オンポリシー型 IL の安全を向上させる取り組みである。ペグ挿入タスクのような物体と環境の間のクリアランスが厳しい作業に対して、リスク予測器を導入し、ロボットが危険な状態に遭遇した際に人間の介入・教示へと主導的に切替える IL アルゴリズム DPIIL を提案した。二つ目の事例は、人間のリスク鋭敏性を考慮した外乱注入モデルを導入し、外乱注入型 IL の実演可能性を向上させる取り組みである。人間の行動コマンドにリスクに応じてレベルを調整した外乱注入を行うことで、頑健な方策を獲得する新しい外乱注入型模倣学習アルゴリズム BDI を提案した。これらの手法は、人間の行動特性を考慮し模倣学習の性能を向上させる新たなアプローチを開拓しており、今後の発展性が期待できる。さらに、実機・被験者実験の結果に裏付けられた有効性と実用性など、高い学術的価値を有すると考えられる。

公聴会では、タイトルと内容の一貫性、環境情報活用の有無、教示・訓練中の衝突回避の保証、教示インターフェースが及ぼすリスク推定への影響などの項目について審査委員との質疑が行われた。その後の最終審査において、すべての質問への回答と論文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。

本論文は人間の行動心理学の知見に基づき、行動のリスク鋭敏性に焦点を当て、教示行動から環境に潜むリスクを推定し、これを模倣学習に活用することで学習の効率性や汎化性を高めるアプローチを提案した。提案アプローチの概念検証として二つの模倣学習システムへの適用事例を示した。シミュレーションおよび実ロボットを用いたタスク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。