

研究課題別事後評価結果

1. 研究課題名： 頑健性と安全性の性能限界を明らかにする深層強化学習

2. 個人研究者名

小林 泰介（情報・システム研究機構 国立情報学研究所 助教）

3. 事後評価結果

本研究は、自律ロボットなどの環境を探索する自律型 AI に焦点を当て、実世界における信頼性を改善するための深層強化学習を研究開発している。とくに、深層強化学習の信頼性に大きく関与する頑健性（外乱への補償能力）と安全性（制約条件を満たす能力）の2つに着目し、モデルベース強化学習に基づいて、以下の3つの観点から、これらを積極的に保証する新しい機構を開発した。

(1) 確率的予測モデルの表現力向上に関しては、予測の最悪ケースを考慮可能な最適化問題を導出し、そこに潜むバイアス・バリエーションのトレードオフを高効率に調整可能なメタ最適化手法を開発した。さらに、確率分布が持つ表現能力を効率的に高めつつも、分布の平均的挙動を解析的に算出可能な制約付き確率分布モデルを開発した。(2) 実時間内での安全性保障に関して、高次元空間での低効率な予測を避けるため、制御可能な限界まで低次元空間に圧縮するためのスパース表現学習を開発し、ロボットマニピュレータによる実証実験で有用性を実証した。さらに、制約条件を破綻させない準最適解を高速に獲得するため、行動系列候補から失敗を積極的に回避する更新則を持つモデル予測制御を開発・実装して、GPU を用いずに CPU のみの計算環境においても、円滑な人とロボットのインタラクションを実現可能なことを実証した。(3) 頑健性の制約付き最大化に関しては、不等式制約付きの最適化問題を適切に扱うため、不等式制約の正則化項への変換手法を開発し、これをロボットアームに実装し、学習時に経験しなかった未知物体の把持や人とのインタラクション時の適応的な振り舞いを実現することができた。予測モデルの頑健性を最大化するために、敵対的学習において予測精度の低下率に不等式制約を設ける新たな枠組みを開発した。

これらの成果は、設計時から頑健性と安定性の要求を新たな最適化問題として組み込み、そのための理論に裏付けられた最適化手法を設計することで、深層強化学習器の信頼性を向上させた。単なる新奇な理論的アイデアの提案を行うだけにとどまらず、粘り強く、かつ、確実に、実ロボットにおける実装と実証実験を進め、人とロボットの円滑な相互作用の実現における提案手法の有用性を示した。まとめると、実世界で人と協調して働くロボットに要求される頑健性と安全性を、タスクの設計段階から数理的にモデル化し、性能保証をもつ最適化手法を設計することで、実世界の信頼されるロボットの開発・運用に際して重要な課題の一つの解を示した研究である。本さきがけ研究は、上記の課題に関して、個別のタスクに対する要素技術の開発を対象としたものである。今後は、将来出現すると予想される人と協調して働くさまざまな汎用のロボット運用基盤へも対応できるように、本研究で得られた方法論の一般化と汎用化にも取り組むことを期待したい。