

## 論文内容の要旨

博士論文題目

Task-relevant Model-based Reinforcement Learning for Contact-rich  
Robotic Tasks

(コンタクトリッチなロボットタスクのためのタスク関連モデルベース強化学習)

氏名

Kuo Cheng-Yu

(論文内容の要旨)

モデルベース強化学習は、エージェントが試行錯誤的に収集した経験データから環境のダイナミクスを学習し、制御方策の改善に活用する強化学習のアプローチである。そのサンプル効率の高さから、ロボットなどのサンプル収集コストの高い実環境タスクに有望とされている。しかしながら、意図したタスクを完了するための目的関数は、ダイナミクスの状態に基づいて設計されるため、タスクに関連する状態変数およびダイナミクスを事前かつ適切に設計する必要がある。

本論文では、ロボットタスクへの実用にとって重要となる、環境や物体との接触が頻繁に生じる（コンタクトリッチな）タスクに焦点を当て、タスク関連モデルベース強化学習の枠組みを提案する。提案する枠組みの主要な特徴は、1)意図したタスクの要件を満たし、2)適切なダイナミクス学習し、3)タスクの実行中における動作の調整が可能となる、適切なダイナミクスの状態変数を設計することにある。

提案する枠組みを2つのコンタクトリッチなタスクに適用し、その有効性を検証した。第一のタスクは、キッチン環境における物体操作タスクである。調理作業の中には、食材を混ぜたり、切ったりといったロボットと物体、環境との物理的接触が不可避な作業が数多く存在する。学習途中のダイナミクスを通じて制御方策を改善する際、学習データ不足によるダイナミクスの予測精度不

足等の原因により、環境と強く衝突する危険な状況が発生し得る。そこで、学習過程における不可避な接触に対する安全性向上を図るために、タスク関連ダイナミクスの状態変数として、確率的回帰モデルの予測分布の統計量である平均および分散を採用した。これにより、制御方策の更新時に予測分散の値を一定以下に制約することで、データ不足に起因される危険な行動の計画を回避できる。また、データ収集の進捗に応じて制約が緩和される仕組みとなり、学習される制御方策の最終的な性能を阻害しないことも利点である。

第二のタスクは、関節に弾性要素を有する2足ロボットの歩行タスクである。2足歩行ロボットは、関節数や浮遊ベースに加えて、関節のばね要素の影響を考慮すると、ダイナミクスの状態変数は高次元となる。この高次元性の課題を克服するため、すべてのバネをエネルギー容器として捉え、アクチュエータをエネルギー源として扱うことで、エネルギー保存の法則に基づいたタスク関連ダイナミクスを採用した。これにより、エネルギーに基づく状態表現が可能になり、モデル次元が削減され、結果として行動計画の計算速度が大幅に向上した。実ロボット実験においても、従来法と比べて高速な制御ループを構築でき、環境変化に頑健な歩行運動を達成できることを確認した。

上記の2つのタスクに対するタスク関連モデルベース強化学習の設計および実験検証により、提案手法の有効性と実用性を確認した。



氏名	Kuo Cheng-Yu
----	--------------

(論文審査結果の要旨)

本論文は、ロボットの実用上で重要となる、環境や物体との接触が頻繁に生じるコンタクトリッチな環境でのタスクに焦点を当て、タスク関連モデルベース強化学習の枠組みを提案している。提案する枠組みでは、モデルベース強化学習における状態変数およびダイナミクスに対して、1)意図したタスクの要件を満たし、2)適切なダイナミクス学習し、3)タスクの実行中における動作の調整が可能となる、を満たす重要性を提唱している。

提案する枠組みの具体事例として、2つのコンタクトリッチなタスクにおいて異なるモデルベース強化学習を設計し、実ロボットを用いたタスクへの適用実験を通じてその有効性を実証した。第一のタスクは、キッチン環境における物体操作タスクである。調理作業には、食材を混ぜる・切るといったロボットと物体、環境との接触が避けられない状況が多々存在する。制御方策の改善において学習中のダイナミクスを利用すると、学習データ不足により予測精度が問題となり、環境との過度の衝突が発生するリスクがある。そこで学習過程での避けられない接触に対する安全性を向上させるため、タスク関連ダイナミクスの状態変数として、確率的回帰モデルの予測分布の統計量である平均と分散を用いた。制御方策の更新する際に、予測分散の値を一定以下に制約することで、データ不足に起因される危険な行動の計画を回避でき、さらにデータ収集の進捗に応じて制約が緩和される仕組みとなっている。第二のタスクは、関節に弾性要素を有する2足ロボットの歩行タスクである。関節の弾性要素を考慮すると、関節数や浮遊ベースに加えダイナミクスの状態変数は高次元化する課題がある。そこですべてのバネをエネルギー容器として扱い、アクチュエータをエネルギー供給源として扱うアプローチを開発した。エネルギー保存の法則に基づくエネルギー表現を含める状態変数およびタスク関連ダイナミクスを定義した。結果として、モデルの次元が削減され、行動計画の計算速度が大幅に向上することを示した。実ロボット実験においても、従来法と比べて高速な制御ループを構築し、環境変化に頑健な歩行運動を達成できることを確認した。

公聴会では、提案手法における目標軌道の設計指針や性能への影響に関する考察、実装に関する説明の不足、一般性や他のロボットタスクへの応用可能性

についての議論など、いくつかの不足事項が指摘された。最終審査では、これら指摘事項への対処や改善について詳細な説明が行われ、論文に適切に修正がなされていることを確認した。これにより本論文の貢献がより明確化された。

本論文は、ロボットのコンタクトリッチな環境でのタスクに焦点を当て、タスク関連モデルベース強化学習の設計指針を提案した。また、その具体事例として、学習過程における接触安全性を考慮可能なコンタクトセーフなモデルベース強化学習、関節に弾性要素を有するロボットにおける次元圧縮を可能とするエネルギー保存則に基づくモデルベース強化学習を開発した。それぞれの手法に対して、実ロボットによる学習実験を実施し、安全または効率的にタスクを達成する方策を獲得できることを実証した。これらの成果に鑑みて、その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。