強化学習によるロバストな行動の獲得

泉田　啓，藤井 信治，松本 勉，
岡野 裕三，高田 浩平，藤原 直史
金沢大学大学院　自然科学研究科

1. 研究の背景と目的

宇宙ロボット

宇宙飛行士に代わって，宇宙ステーションなどの大型宇宙構造物の組立やメンテナンスを行なわせたい．

宇宙ロボット実現への課題とアプローチ

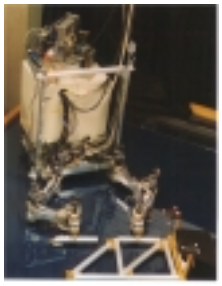
- 環境の不確定性・変動に対する適応的な作業達成

⇩

強化学習の適用

⇩

- 強化学習は学習速度が**遅い**



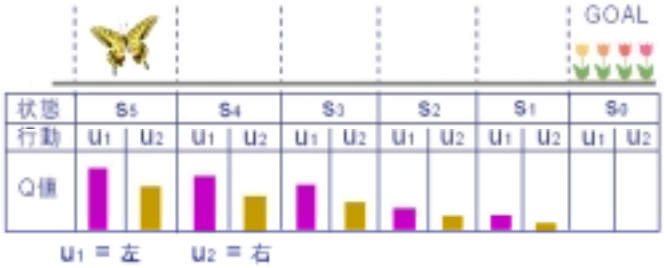
模擬実験システム

研究目的

強化学習の実用化に向けた**学習時間の低減化**

2. 強化学習とは

○　1回の移動にコスト1が必要とする．

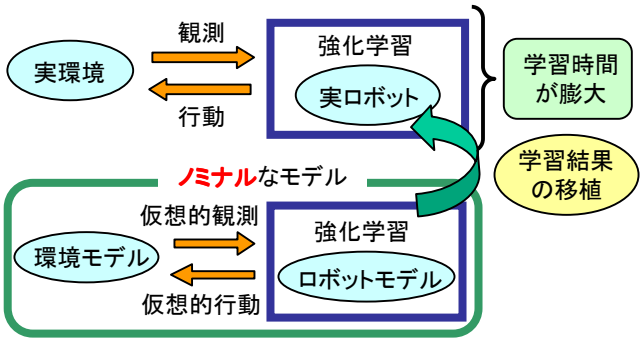


任意のQ値 → 試行錯誤 → **最適なQ値**

Q値の小さな方の行動選択 → **最適な行動**

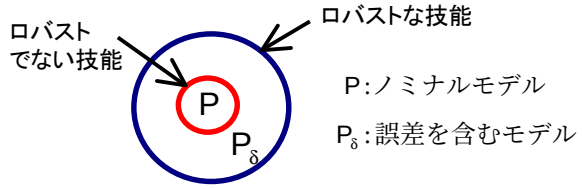
学習とは最適なQ値と最適な行動を見つけること

強化学習の通常のアプローチ



3. ロバストな技能の獲得

ロバストな技能とは



タスクを達成可能な範囲

Peg-in-Hole タスク

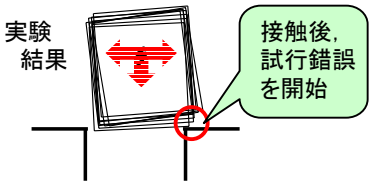
ペグを同サイズのホールへ挿入する組立作業の代表的なタスク

タスクの特徴

- 形状精度の誤差や弾性変形などの不確定性や変動がある
- 接触状態が非常に多い（2次元でも数十の接触状態あり）
- 各状態の遷移が僅かな力の差に依存する
- 接触状態の違いは，運動方程式中の拘束条件式の違いとなり，制御則を変える必要が生じる．

通常のアプローチによる実験

視覚誤差などに起因した**僅かな誤差**によるタスク失敗



実環境での修正

⇩

学習時間が膨大に

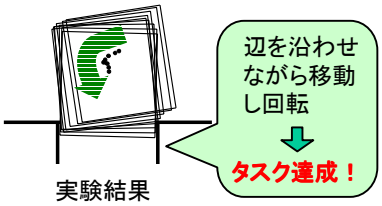
人間の場合は

誤差に対して**ロバストな技能**で楽々挿入達成

どうすればこの技能を獲得できるのか？

本研究のアプローチ

ノミナルなモデルに**予測可能な誤差**を追加して事前に学習



実環境での試行錯誤の低減化

⇩

学習時間短縮