

自律的制御学習法

工学部 知能機械工学科 山田 訓, 西山 陸(山田研究室)

Keywords: 強化学習, 要素配置型強化学習

1. 研究目的

強化学習とは、制御結果の良し悪しをだけを用いて学習する自律的な制御学習法である。色々な課題の制御を学習するためには、課題毎に学習中に必要な位置に要素を配置する要素配置型強化学習が有効である。要素配置型強化学習の竹馬型二足歩行ロボット制御と3リンクマニピュレータ制御への適用を検討した。

2. 要素配置型強化学習の改良と学習結果

従来の要素配置型強化学習を竹馬型二足歩行ロボット制御に適用したところ、制御を学習することができなかった。従来の要素配置型強化学習では、評価の推定値の誤差が大きく、周囲に要素を配置していない場合に、新しい要素をその時点での状態に配置する。従来の要素配置型強化学習で学習した場合の要素の分布(図1左上)と状態の分布(図1右上)を比較すると、周囲の転倒寸前の状態には要素が配置されているが、中心付近の歩行中の状態にはほとんど配置されていない。竹馬型二足歩行ロボットの制御を学習するためには、歩行中の状態の違いを認識する必要があるが、従来の方法では、歩行中の状態付近には要素が配置されていないので、学習できなかったと考えられる。そこで、制御中の状態の履歴も要素を配置する際の候補にする方法(履歴あり)を試したところ、要素の分布(図1左下)が変化し、歩行中の状態付近にも要素が配置され、制御を学習することができた。状態の履歴を要素配置の候補とする要素配置型強化学習を3リンクマニピュレータ制御の学習に適用したところ、目標の位置で静止する制御を、従来の要素配置型強化学習(履歴なし)より効率よく学習できた(図2)。

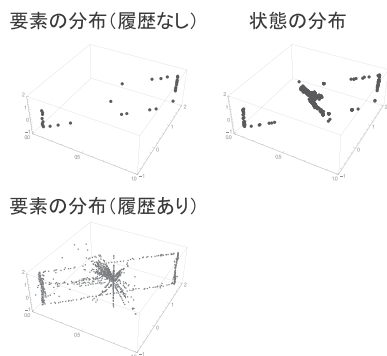


図1 竹馬型二足歩行ロボットの要素配置型強化学習の要素と状態の分布

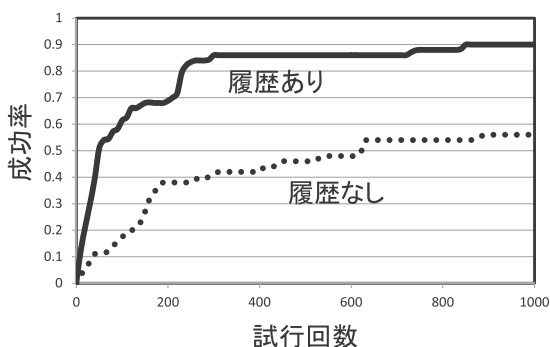


図2 3リンクマニピュレータの要素配置型強化学習による学習曲線

3. 応用の可能性

状態の履歴を要素配置の候補とする要素配置型学習を用いることにより、学習の効率化と実用的な制御への応用が可能になる。今後は、家事や介護等の実用的な制御課題への適用を検討したい。この研究は、文部科学省・私立大学戦略的研究盤形成支援事業「QOL 向上を目指す支援技術のイノベーション研究拠点形成」の一環である。

連絡先: TEL: 086-256-9529 E-mail: yamada@are.ous.ac.jp