

システム情報工学研究科修士論文概要

年 度	平成 25 年度	学位名	修士(工学)
専 攻	知能機能システム	専攻	著者氏名 白井 翼
指導教員氏名 安信 誠二			
論文題目 環境変化の事前知識を取り入れた強化学習に関する研究			
論文概要 強化学習は、環境のモデルを用いずにロボットの試行錯誤の経験から行動規則を獲得する学習の枠組みであり、環境が複雑なモデルであるロボットの行動規則の獲得に有用である。従来の強化学習の多くは静的な環境を前提としたアルゴリズムであるため、環境変化が発生する動的な環境における学習では、膨大な学習時間を必要としたり、学習が破綻することがある。そのため、例えば時間帯によって人通りが異なる場所を巡回するロボットや、天候が変化する屋外で活動するロボットの行動規則を獲得することは困難である。しかし、これらの環境変化は、事前にその変化の仕方を想定することができる。人通りは過去の統計や経験から「おそらく1時から2時の間には人通りが多いだろう」、天候は天気予報から「午後から80%の確率で雨が降るだろう」といったように、内容に不確実性を含みながらも変化の仕方に関して知識が得られる。動的な環境下で学習するロボットが、発生する環境変化に関して予めこのような知識を与えられていれば、環境変化に対応した行動規則を獲得しやすくなると期待できる。そこで本論文では、事前に想定できる環境の変化の仕方を事前知識として取り入れたベイズ適応的 MDP(Bayes-Adaptive MDP:BAMDP)上の強化学習を提案する。提案する強化学習は、事前知識をBAMDPの情報状態:経験から状態遷移確率を推定する確率分布に反映させることで、環境変化に対応した行動規則の獲得を促す。計算機実験で、提案する強化学習を環境変化が発生する強化学習問題に適用し、その有効性を示す。			
審査日	平成 26 年 1 月 29 日		
審査員	(大学名 職名)	(学位)	(氏名)
主査	筑波大学 教授	工学博士	安信 誠二
副査	筑波大学 助教	博士(工学)	澁谷 長史
副査	筑波大学 教授	Ph.D.(工学)	堀 憲之