



Title	学習環境の動的変動を考慮したロボットの強化学習法に関する研究
Author(s)	山口, 智浩
Citation	大阪大学, 1997, 博士論文
Version Type	
URL	https://hdl.handle.net/11094/40299
rights	
Note	著者からインターネット公開の許諾が得られていないため、論文の要旨のみを公開しています。全文のご利用をご希望の場合は、 https://www.library.osaka-u.ac.jp/thesis/#closed 大阪大学の博士論文について ご参照ください 。

The University of Osaka Institutional Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

氏 名	山 口 智 浩
博士の専攻分野の名称	博 士 (工 学)
学 位 記 番 号	第 1 2 8 0 3 号
学 位 授 与 年 月 日	平 成 9 年 1 月 27 日
学 位 授 与 の 要 件	学 位 規 則 第 4 条 第 2 項 該 当
学 位 論 文 名	学習環境の動的変動を考慮したロボットの強化学習法に関する研究
論 文 審 査 委 員	(主査) 教 授 谷 内 田 正 彦 (副査) 教 授 井 口 征 士 教 授 西 田 正 吾 教 授 北 橋 忠 宏

論 文 内 容 の 要 旨

本論文は、実環境で知的に振る舞う学習ロボットシステムの実現について述べたものである。ここで学習とは、システムが状況に応じて自己の能力を改善する機能を意味する。特にロボット学習とは、ロボットの動作によって生じる環境とのインタラクションを通して、自己の能力を改善する学習を指す。例えば産業用の組み立てロボットにおいて、作業対象が変わるごとに動作を人間がプログラミングしなおすのが繁雑な場合に、ロボットが知覚した部品の大きさや形状の変化に対し、組み立て動作の結果から自己の動作を柔軟に修正する機能である。このような従来の学習ロボット研究は、大半が、以下に示すように学習環境と実行環境とが分離されているオフライン学習である。

- ・学習環境：学習データを事前に収集し、バッチ的に帰納学習を行う。
- ・実行環境：実環境でロボットの実行は、学習済みデータで行動する。

その前提として、以下の2点が仮定されている。

- 1) 学習環境が静的かつ既知と仮定し、最適な行動が完全に同定できる。
- 2) 学習環境=実行環境と仮定し、学習データが実行環境で役に立つ。

これらの前提は、人工知能での学習理論に代表されるように、計算機上での閉じた世界やシミュレーション環境(仮想環境と呼ぶ)では一般に成立する。しかしながら、ロボットを現実世界で動かそうとした場合には、実行環境が閉じていないため、上述の前提が成り立たず様々な問題が発生する。それゆえ、産業用ロボットが動作可能な空間は、人工的に管理された工場等の静的かつ既知な環境に限定されてきた。

その原因は、実環境は静的でなく、全て既知でもない、という点にある。すなわち、実環境や未知環境下では上述の前提が成り立たないためオフライン学習は役に立たず、実行環境=学習環境を前提とするオンライン学習能力が必要となる。

こういった問題点に対して本論文では、実環境の本質が、開いた環境における状況の多様性ととらえ、これらに対応するために必要なシステムの振る舞いが知的であると考え。その基本的学習戦略は、学習環境の動的変化に追従するための高速なオンライン学習能力である。本論文では、学習環境が動的に変動する以下の4つの場合について、

それぞれ強化学習の高速化手法を提案する。

1) 学習環境≒実行環境の場合

→仮想環境での準備学習結果を実環境で洗練する準備洗練強化学習法：SARA

2) 学習環境=実行環境=実環境の場合

→副報酬の自動生成による、実環境での実ロボット実時間強化学習法：RtRRL

3) 未知の状況変化への対応

→同型性に基づく強化学習法：Iso-RL による多様な挙動の自律的生成

4) マルチエージェント学習環境において、学習個体自身が環境を変動させる場合

→適応型模倣による複数個体の強化学習の効率化：MMRL

学習環境の動的変化を前提としたロボットの強化学習研究は、実環境で知的に振る舞う実ロボットシステムを目指す上で極めて重要な課題であり、本論文はその一部をなし得たと確信する。

論文審査の結果の要旨

強化学習手法には、静的環境やマルコフ決定過程を仮定した上で理論的な解析が行われるものが多い。しかし従来、上述の仮定を満たさない実環境や動的環境では理論的解析が困難なため、強化学習法を実ロボットに適用し、学習環境の動的変動を直接扱う研究はほとんど行われてこなかった。

本論文では、動的な環境で行動するロボットの学習を効率化・高速化し、かつ環境の変動に対応するための強化学習法を提案し、これらを学習環境が動的に変動する4つの場合に適用することにより、実世界でのオンライン学習を実現するための方法と評価に関する研究についてまとめている。

3章では、学習環境と実行環境が異なる従来のオフライン学習の問題点を検討し、シミュレーション学習と実ロボット学習とを組み合わせ、効率化にオンライン学習を行う学習システムを提案している。次に4章では、これを用いて、シミュレーションでの学習結果を実環境で洗練することにより、全体での学習時間を抑えながら実環境に適応する学習法を提案している。5章では、未知環境のようにシミュレーション環境が学習環境として利用できない場合に、実環境のみでの高速な学習を可能とする実ロボット学習法を提案している。4、5章では実ロボットを用いた学習実験の結果についても述べており、提案している学習手法の実現性、有効性が確認されている。

6章では、未知の状況変化に対しては、あらかじめ既存の学習で得られた挙動のバリエーションを効率良く生成する手法を提案している。7章では、マルチエージェント学習環境において、学習個体自身が環境を変動させる場合について検討し、学習結果を個体間で効率良く伝播させ、しかも強化学習とのバランスをとることにより、学習結果の均質化を抑えながら複数個体の強化学習を効率化する手法を提案している。

6、7章ではシミュレーションによる検証実験によって提案している。学習手法の実現性、有効性の確認及び従来手法との比較が行われ、学習手法の効率化に大きな効果があることも確認されている。

以上のように、本論文で提案された学習手法を適用することにより、学習環境の動的変動下での学習が可能となり、ロボティクス、人工知能などエージェントの学習を利用する多くの研究分野の向上に寄与するところ大であり、高く評価される。よって、本論文は学位（工学）論文として価値あるものと認められる。