



Title	強化学習を用いたエージェントの環境適応能力向上手法に関する研究
Author(s)	小谷, 直樹
Citation	大阪大学, 2011, 博士論文
Version Type	
URL	<a href="https://hdl.handle.net/11094/58406">https://hdl.handle.net/11094/58406</a>
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、 <a href="https://www.library.osaka-u.ac.jp/thesis/#closed">https://www.library.osaka-u.ac.jp/thesis/#closed</a> 大阪大学の博士論文について <a href="https://www.library.osaka-u.ac.jp/thesis/#closed">https://www.library.osaka-u.ac.jp/thesis/#closed</a> をご参照ください。

*The University of Osaka Institutional Knowledge Archive : OUKA*

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

The University of Osaka

氏名	小谷 直樹		
博士の専攻分野の名称	博士(工学)		
学位記番号	第 24590 号		
学位授与年月日	平成 23 年 3 月 25 日		
学位授与の要件	学位規則第 4 条第 1 項該当 工学研究科電気電子情報工学専攻		
学位論文名	強化学習を用いたエージェントの環境適応能力向上手法に関する研究		
論文審査委員	(主査) 教授 谷口 研二 (副査) 教授 谷野 哲三 准教授 高橋 篤司 教授 伊藤 利道 教授 森 勇介 教授 片山 光浩 教授 尾崎 雅則 教授 栖原 敏明 教授 近藤 正彦 教授 森田 清三 教授 八木 哲也		

論文内容の要旨

本論文は、強化学習を用いたエージェントの環境適応能力向上手法に関する研究をまとめたもので、以下の6章から構成されている。

第1章では、人の生活空間で活動するロボットの環境適応能力の向上を達成するために、強化学習を用いて、マルチタスク下においてロボットアームが少ない試行回数で障害物の出現位置に応じた軌道を学習できるようにすることと、超音波アレーセンサが様々な環境に対して適切に計測できるようにすることを課題に採り上げ、本研究の目的を明らかにした。

第2章では、本論文で用いたアルゴリズムの基本的な概念を整理するために、強化学習の概念と理論的背景、代表的な強化学習アルゴリズムの1つであるアクター・クリティック (Actor-Critic) 、そして強化学習の状態空間の自律的構築に用いる教師なしクラスタリングアルゴリズムの1つである適応共鳴理論 (Adaptive Resonance Theory; ART) を示した。

第3章では、マルチタスク下で学習するエージェントの適切な行動獲得に必要な試行回数を削減するために、従来のFuzzy ARTのカテゴリ数増加の問題点を指摘すると共に、その問題点を解決するクラスタリング方法を提案した。障害物回避を伴う2リンクロボットアームのリーチング問題に適用したところ、カテゴリを表す重みベクトルの更新方法の改良によってカテゴリ数が抑制され、適切な行動獲得までに必要な試行回数が削減できていることを確認した。

第4章では、マルチタスク下で学習するエージェントの適切な行動獲得に必要な試行回数を削減するために、交叉、突然変異、淘汰の遺伝的操作を取り入れたFuzzy ARTを提案し、過去のタスクで獲得した政策の利用による学習の効率化の有効性を示した。障害物の位置を変えた10個のタスクに連続的に遭遇する障害物回避を伴う2リンクロボットアームのリーチング問題を与えたところ、提案手法は過去に獲得した政策を用いることによって、新しく遭遇したタスクに対して早期に行動獲得でき、記憶量も抑制できることを確認した。

第5章では、強化学習を用いた超音波アレーセンサの指向性設計ツールが、解析的な設計が困難である超音波アレーセンサの指向性設計問題に対して有効であることを示した。7素子のリニアアレー超音波送信器におけるサイドローブが抑制されたビーム形状の設計に適用し、所望のビーム形状が得られることを確認した。

第6章では、本論文で得られた研究成果を総括し、本論文の結論とした。

論文審査の結果の要旨

本論文では、人と協働する自律ロボットを実現するために、強化学習を用いたエージェントの環境適応能力向上に関する新しい手法を提案している。その主要な成果は以下のとおりである。

(1) マルチタスク下で学習するエージェントの適切な行動獲得までの試行回数を削減するために、従来のFuzzy ARTの状態分割問題を克服する新規クラスタリング方法を提案している。提案手法では、カテゴリの領域を表す重みベクトルを該当カテゴリに分類された観測状態の平均値とすることでカテゴリの急激な成長を抑え、カテゴリ数の増加と誤った状態分割による政策の発散を抑えている。また、適切な行動獲得に不要な価値の低いカテゴリを削除してカテゴリ数の増加を抑えている。本提案手法を2リンクロボットアームのリーチング問題に適用した結果、提案手法は従来法に比べてカテゴリ数を62%削減するとともに、約半分の試行回数で効率的な学習を実現している。

(2) Fuzzy ARTに交叉、突然変異、淘汰の遺伝的操作を取り入れ、過去のタスクで獲得した政策の有効利用により、マルチタスク下で自律学習するエージェントの行動獲得の効率化を図っている。交叉操作では、新しいカテゴリ生成時に、その状態と類似する異なるカテゴリの政策を参照し、新しい政策の初期値を決める。突然変異操作では任意の確率で政策をランダムに変更して、局所解に陥ることを回避している。淘汰操作により、タスク達成に必要なカテゴリを消去して記憶量の抑制を図っている。提案手法による学習効率化の効果を、障害物の位置を変えた10個のタスクと連続的に遭遇する障害物回避を伴う2リンクロボットアームのリーチング問題に対してシミュレーションを行い、従来に比べて試行回数を約60%削減できることを示している。この結果、Fuzzy ARTに遺伝的操作を取り入れることで効率的な強化学習ができることを確認している。

(3) 超音波アレーセンサの指向性パターンは、センサを構成する各素子の利得、応答特性、配置などのパラメータに依存するため、所望の指向性を決めることは極めて困難である。この問題を解決するため、強化学習に基づく超音波アレーセンサの指向性設計ツールを開発している。本指向性設計ツールを7素子のリニアアレー超音波送信器に適用し、所望の方向での超音波の積算エネルギーが最大となる強化学習を行った結果、サイドローブの抑制された超音波ビーム形状が得られることを確認している。

以上のように、本論文はエージェントの環境適応能力の向上のための新しい手法を提案しており、マルチタスク下で学習するエージェント実現のための技術進展に貢献する成果を挙げている。よって本論文は博士論文として価値あるものと認める。