



強化学習によるマルチロボットシステムの協調行動 獲得に関する研究

保田, 俊行

(Degree)

博士 (工学)

(Date of Degree)

2006-03-25

(Date of Publication)

2008-04-22

(Resource Type)

doctoral thesis

(Report Number)

甲3705

(URL)

<https://hdl.handle.net/20.500.14094/D1003705>

※ 当コンテンツは神戸大学の学術成果です。無断複製・不正使用等を禁じます。著作権法で認められている範囲内で、適切にご利用ください。



【 3 1 6 】

| | | |
|------------|--------------|---------|
| 氏 名・(本 籍) | 保田 俊行 | (兵庫県) |
| 博士の専攻分野の名称 | 博士(工学) | |
| 学 位 記 番 号 | 博い第414号 | |
| 学位授与の 要 件 | 学位規則第5条第1項該当 | |
| 学位授与の 日 付 | 平成18年3月25日 | |

【 学位論文題目 】

強化学習によるマルチロボットシステムの協調行動獲得に関する研究

審 査 委 員

| | | |
|-----|-----|--------|
| 主 査 | 教 授 | 田浦 俊春 |
| | 教 授 | 小島 史男 |
| | 教 授 | 大須賀 公一 |
| | 助教授 | 大倉 和博 |

対象とするシステムの複雑化に伴い、モデルベースの手法では限界が生じる。それゆえ、異なるアプローチとして、“相互作用”をキーワードとして生物のような知的な振る舞いの実現やその原理の解明を目指したシステム構築に関する研究が多く行われている。ロボット間およびロボットと環境間の相互作用を通して協調的に振る舞うマルチロボットシステムは、この研究の枠組みにおける典型例のひとつである。本研究では、マルチロボットシステムにおいて本来期待される特徴である頑健性を実現するための概念として、状況に応じて役割を生成・割り当てるという“自律的機能分化”を導入した。また、それを実環境で実現するために、計算知能のアプローチのなかでも教師データを必要とせずにオンラインで行動獲得が可能な手法である強化学習に着目した。マルチロボットシステムの自律的機能分化には個々のロボットの自律適応能力を向上させることの重要性を述べた。その実現のために強化学習を適用する場合に対処すべき点として、(a)連続な状態・行動空間の自律的分割、(b)ダイナミクス許容量(頑健性)の増大、(c)ダイナミクスの軽減、および(d)試行錯誤回数の軽減の四つを挙げ、強化学習の機能拡張を行った。以下、各章で得られた研究結果を要約する。

第2章では、ロボット単体レベルで頑健な意思決定を行う手法を提案した。まず、強化学習で問題となることとして、行動収束後に実験を繰り返すことによる頑健性の低下について指摘した。その問題に対処するため、確率ネットワークを用いたロボットの獲得戦略の保存・適用手法を提案した。この手法は、強化学習は環境内を探索して適切な入出力関係を収集するために用い、そのデータを基に確率ネットワークを用いて意思決定機構を構築するというものである障害物の形状が変化する環境における移動ロボットのゴール到達問題において、強化学習器のみを用いたロボットよりも頑健に振る舞うことを。計算機実験、および実機実験で確認した。タスク達成可能な知識を確率ネットワークで表現することで、環境変化が生じて未経験な状態に陥っても、システムの頑健性を維持し柔軟に対処することができた。

第3章では、マルチロボットシステムにおける頑健な自律的機能分化を実現するための強化学習の拡張法を提案した。まず、状態・行動空間の離散化の困難性、および連続空間を取り扱う強化学習の研究例について述べた。続いて、その中からマルチロボットシステムの自律的機能分化の観点から適用する強化学習の位置づけを示した。それに基づき、ベイズ判別法を用いて連続な状態行動空間を自律的に分割する機能を持つ強化学習法の Bayesian-discrimination-function-based Reinforcement Learning (BRL)を取り上げた。そして、BRLにおける過学習問題を指摘した。その後、過学習を抑制してシステムの頑健性を向上させるために、学習過程で有効であったルールを保護することでルール集合の多様性を維持する機構を付加するとともに、そのための有効ルールの選択の指標を提案した。その有効性を検証するために、三台のアーム型自律ロボットによる協調荷上げ問題に適用

して実機実験を行った。その結果、可塑的な自律的機能分化が発現するなど頑健性が向上し、環境変化に対して従来型 BRL よりも安定した大域的秩序を迅速に再形成していることを確認した。

第4章では、BRLを用いたマルチロボットシステムの協調行動獲得において、環境のダイナミクスを軽減することで学習をより安定化することで学習効率を向上させる手法を提案した。まず、マルチロボットシステムに強化学習を適用する場合の問題点を指摘し、それに対処するための従来手法を述べた。その後、次時刻における他ロボットの状態を予測する機構を構築し、その予測機構の出力をBRLの入力の一部として付加する手法を提案した。提案手法を自律移動ロボットによる協調搬送問題に適用した。計算機実験により予測情報を用いない場合よりも提案手法が効率的に学習していることを確認した。また、実環境においても、行動獲得に成功したことからノイズなどの不確定要素にも耐え得る頑健な行動学習を実現していることを確認した。

第5章では、強化学習における試行錯誤の回数を極力削減したいという基本的要求から、行動空間の探索効率の向上を目的として BRL の拡張を行った。従来型 BRL におけるルール生成時の行動の決定法における問題を指摘した。機能拡張として、行動選択における事後確率の計算において新たなしきい値を設定し、近傍のルールのパラメータを基に新ルールを生成する手法を付加した。提案手法を一台の移動ロボットによる光源到達問題に適用し、行動獲得および獲得後の環境変動に対する再学習に関する計算機・実機実験を行うことで提案手法の基本性能を検証した。その後、協調搬送問題の計算機・実機実験においても同様の実験を行った。いずれの実験でも従来型 BRL と比較して短いエピソードでの行動獲得を確認し、提案手法の有効性を示した。

第6章では、強化学習によるマルチロボットシステムの協調行動獲得問題において、システムの環境変動への頑健性の向上のために、より迅速な再学習を目指して BRL の拡張を行った。i)ルール発火の情報エントロピを入力の一部として用いて行動の安定性を考慮したルールを構成する、ii)エントロピに基づいて学習過程で有益であったルールを削除から保護するというアプローチを提案した。提案手法を三台のロボットの協調搬送問題に適用し、学習集束後の環境変動として未学習のロボットとの交換をしてシステムの挙動を観察した。保持しているルールを有効に利用するとともに、エントロピの変化量を基に再学習へ効率的に移行することで新しい協調行動を獲得していることを確認した。また、5章で提案した手法と融合することで、行動獲得のパフォーマンスがさらに向上することを確認した。

本研究では、以上のように五つのアプローチで、自律的機能分化を実現するために強化学習ロボットの自律適応能力の向上手法を提案し、それらの有効性を示した。対象としたマルチロボットシステムは、一般的に用いられる非均質なのではなく均質なものであった。そのため、機能分化にいたるまでにコストがかかることやタスク達成の効率が劣ると

(氏名： 保田俊行 NO. 3)

いった問題がある。しかし、本来、マルチロボットシステム環境は動的で複雑なものであり、将来の人間の生活環境などでの運用を考えた場合は、想定しないような状況に陥る可能性が十分に考えられる。そのような状況では、非均質なマルチロボットシステムではその非均質性が新しい協調関係を阻害する場合があるといえ、自律的機能分化を発現することが重要となる。本研究ではロボット同士の協調に特化していたが、ロボットが均質性に起因する自由度の高さから、協調する相手はロボットに限定しないアプローチとも捉えることができる。つまり、将来的に、より多入力・多出力のシステムにおける迅速な行動獲得が可能となれば、様々な性格の人間との共同作業においても、自律的機能分化を通して柔軟に役割を発現することが期待できる。以上より、本研究で提案した強化学習による自律的機能分化を通じた協調行動獲得は、頑健なマルチロボットシステムのみならず、人間機械協調系への展開など様々なレベルの相互作用を有するシステム構築における一針を示したといえる。

| | | | |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------|-----|--------|
| 氏名 | 保田 俊行 | | |
| 論文 題目 | 強化学習によるマルチロボットシステムの協調行動獲得に関する研究 | | |
| 審査 委員 | 区 分 | 職 名 | 氏 名 |
| | 主 査 | 教授 | 田浦 俊春 |
| | 副 査 | 教授 | 小島 史男 |
| | 副 査 | 教授 | 大須賀 公一 |
| | 副 査 | 助教授 | 大倉 和博 |
| | 副 査 | | |
| 要 旨 | | | |
| <p>制御対象とするシステムが複雑化するに伴い、モデルベース手法では限界が見え始めてきた。これに対して、システムがおかれた環境内で作動することによる相互作用を基軸とした適応的振る舞い獲得の研究が広く行われるようになってきた。その典型例の一つに、ロボット間およびロボットと環境間の相互作用を通して協調的に振る舞うマルチロボットシステムがある。本論文では、このマルチロボットシステムにおいて期待される特徴であるシステムの頑健性を飛躍的に向上させるには、状況に応じて役割を生成・割り当てる“自律的機能分化”が重要であると指摘するとともに、これを実機へ実装するために、計算知能のアプローチのなかでも教師データを必要とせずにオンラインで行動獲得が可能な手法である強化学習に着目した。そのために、強化学習をマルチロボットシステムに適用するにあたって対処すべき点として、(a)連続な状態・行動空間の自律的分割、(b)頑健性の増大、(c)環境ダイナミクスの軽減、および(d)試行錯誤回数軽減の四つを挙げ、強化学習の機能拡張をはかるとともに、計算機実験および実機実験を通して、その有効性を検証している。</p> <p>第1章では、現在の研究動向とともに本論文の背景と注目点について説明している。</p> <p>第2章では、ロボット単体レベルで頑健な意思決定を行う手法を提案している。強化学習の脆弱性は主に過学習の影響に由来することから、これに対処するため、確率ネットワークを用いて強化学習で獲得した知識を保存することおよび適用方法について解説している。提案手法の有効性を、障害物の形状が変化する環境における移動ロボットのゴール到達問題を例にして、強化学習器のみを用いたロボットよりも頑健に振る舞うことを計算機実験および実機実験を通して検証している。</p> <p>第3章では、マルチロボットシステムにおける頑健な自律的機能分化を実現するための強化学習の拡張法を提案している。まず、状態・行動空間の離散化の困難性、および連続空間を取り扱う強化学習の研究例について指摘している。続いて、その中からマルチロボットシステムの自律的機能分化の観点から適用する強化学習の位置づけを示し、ベイズ判別法を用いて連続な状態行動空間を自律的に分割する機能を持つ強化学習法のBayesian-discrimination-function-based Reinforcement Learning (BRL)を取り上げている。そして、BRLにおける過学習問題を指摘し、これを抑制してシステムの頑健性を向上させるために、学習過程で有効であったルールを保護することでルール集合の多様性を維持する機構を付加するとともに、そのための有効ルールの選択の指標を提案している。その有効性を検証するために、三台のアーム型自律ロボットによる協調荷上げ問題に適用して実機実験を行っている。その結果、可塑的な自律的機能分化が発現するなどして頑健性が飛躍的に向上し、環境変化に対して従来型 BRL よりも安定した大域的秩序を迅速に再形成することを確認している。</p> <p>第4章では、BRLを用いたマルチロボットシステムの協調行動獲得において、環境ダイナミクスを見かけ上軽減することで学習をより安定化することにより学習効率を向上させる手法を提案している。まず、マルチロボットシステムに強化学習を適用する場合の問題点を指摘し、次時刻における他ロボットの状態を予測する機構を構築しその予測機構の出力を BRL の入力の一部として付加することにより、強化学習問題環境のマルコフ性を向上させてより安定な強化学習を行う手法を提案している。提案手法を自律移動ロボットによる協調搬送問題に適用し、計算機実験により提案手法の効果を確認している。また、実環境においても行動獲得に成功したことから、ノイズなどの不確定要素にも耐え得る頑健な行動学習を実現していることを確認している。</p> | | | |

氏名

保田 俊行

第5・6章では、第4章で提案した制御器を基に、BRLの機能拡張を行うことでロボットの自律適応能力のさらなる向上をはかっている。第5章では、強化学習における試行錯誤の回数をできるだけ削減したいという基本的要求から、行動空間の探索効率の向上を目的としてBRLの拡張を行っている。まず、従来型BRLにおけるルール生成時の行動の決定法においては非効率なランダム探索が採用されていることを指摘し、これを、行動選択における事後確率の計算において新たなしきい値を設定し、類似する行動ルールがある場合にはそれらのパラメータを基に新ルールを生成する手法を付加することを提案している。これを一台の移動ロボットによる光源到達問題に適用し、行動獲得および獲得後の環境変動に対する再学習に関する計算機実験および実機実験を行うことで提案手法の基本性能を検証している。その後、実用規模の協調搬送問題の計算機実験、および実機実験においても同様の実験を行い、提案手法の有効性を確認している。

第6章では、強化学習によるマルチロボットシステムの協調行動獲得問題において、システムの環境変動への頑健性の向上のために、より迅速な再学習を目指したBRLの拡張法を提案している。それらは、i)ルール発火の情報エントロピを入力の一部として用いて行動の安定性を考慮したルールを構成すること、およびii)エントロピに基づいて学習過程で有益であったルールを削除から保護すること、のふたつからなっている。この提案手法を三台のロボットの協調搬送問題に適用し、学習収束後の環境変動として未学習のロボットとの交換をしてシステムの挙動を観測し、その有効性を確認している。また、5章で提案した手法と融合することで、行動獲得のパフォーマンスがさらに向上することも検証している。

第7章では、本論文におけるまとめが示されている。さらに、自律的機能分化は、均質なマルチロボットシステムだけでなく非均質なものにも、さらには人間機械系への展開も可能であると指摘している。

以上のように五つのアプローチで、自律的機能分化を実現のために強化学習ロボットの自律適応能力の向上手法を提案し、それらの有効性を計算機実験および実機実験を通して確認している。本研究はマルチロボットシステムについて、自律的な協調行動の獲得を実現するために強化学習を用いて自律的機能分化機能を持たせることによる自律ロボットの知能化手法を研究したものであり、頑健な知能化システムの構築手法について重要な知見を得たものとして価値ある集積であると認める。よって、学位申請者の保田俊行は、博士(工学)の学位を得る資格があると認める。