

## 人間機械協調系のための自律ロボット群における 知識獲得・利用法の開発

Knowledge Acquisition and Utilization of Autonomous Robots for Human-Robot Cooperation

2021023



研究代表者

広島大学大学院 工学研究院  
機械システム・応用力学部門

助教

保田 俊行

共同研究者

広島大学大学院 工学研究院  
機械システム・応用力学部門

教授

大倉 和博

### [研究の目的]

近未来において、自律人工物であるロボットと人間が様々な場面で協同作業をする場面に出くわすことが容易に想像される。しかし、現在の知能システムのメインストリームでは、効率・最適指向の組み込み制御に力点が置かれるあまり、Rodney Brooks が必須の性質と指摘した身体性・立脚性などが軽視され過ぎている。ゆえに、本研究では実ロボット群の制御に対して、研究代表者が開発してきた知識獲得と知識利用を適応的に切替える手法の有効性を検証する。特に、これまでに行ってきたマルチロボットシステム (Multi-Robot System : MRS) の実験の拡張として、人間機械協調系に対する創発的問題解決に適用する。

### [研究の内容, 成果]

#### 自律的機能分化に基づく MRS の協調行動獲得とその人間機械協調系への応用

自律的機能分化とは、我々が提案している適応的行動生成のための概念である。これは、複数の均質なロボットで構成される MRS において各ロボットが機能や役割を適応的に見出し、それらを動的に割り当てるといったものである。従来の非均質なロボットで構成される役割が固定の MRS は、ロボットの台数が増減すると

いった環境変化に対し往々にして対応できない。一方、均質なロボット群から状況に応じて自律的に機能分化ができれば、頑健性の向上が期待できる。これまでに、アーム型ロボットの協調荷上げ問題や移動ロボットの協調搬送問題において、均質な複数台のロボットによる協調行動獲得を実現している。

それらの先行研究では、各ロボットはパートナーの台数やタスク、さらには環境変化に関する情報などを与えられていないにも関わらず、自律的機能分化によって協調行動の獲得に成功している。このことから、ロボットはたとえパートナーが人間であったとしても協調的に振る舞えると考えられる。すなわち、人間機械協調系への拡張が期待できる。本研究では、この自律的機能分化に基づくロボット群と被験者との協調行動獲得実験を行う。特に、実験成功後の被験者交代という環境変化に対して適応できるか否かを検証することで、MRS の頑健性について検証を行う。

#### 連続な状態・行動空間を自律的に分割する機構を持つ強化学習法とその拡張

自律的機能分化する MRS は、個々の行動だけでなく他ロボットとの間での役割分担も発現し得る高い知能を持たなければならない。そのような観点から、我々はロボットが自らの経験に基づき連続な状態・行動空間を自律的に

分割する強化学習法 Bayesian-discrimination-function-based Reinforcement Learning (BRL) を研究・開発している。BRL は状態空間をガウス分布によるパラメトリックモデルを用いて記述し、ベイズ判別法により実行する動作を決定する。BRL では、ベイズ判別法により入力データがどの状態にも識別されなかった場合、新しくガウス分布を追加して状態空間を分割する。既存の状態に識別された場合、ガウス分布のパラメータを修正して状態空間の分割を逐次更新する。そのため、未知環境など入出力データが少ない状況でも状態空間を分割することが可能となり、タスク達成に有用な行動を迅速に獲得することができる。

しかし、BRL はガウス分布のパラメータを逐次更新するため実験によっては過学習が発生し、汎化能力が低下するという問題があった。BRL が入力データを充分収集した段階で、その入力データを基にノンパラメトリックモデルを用いて類似状態を記述し、Support Vector Machine (SVM) によって状態空間を分割する手法を提案した。この手法は、SVM を用いることで汎化能力を維持しつつ意思決定することが可能であり、既に獲得した有用な行動をより有効に利用できる。

#### 協調荷上げタスクによる検証

被験者一名とアーム型ロボット二台が、荷物を傾けることなく規定の高さまで持ち上げることがタスクである (図 1)。ロボットは床に置かれており、タスク開始時の荷物の高さは床か

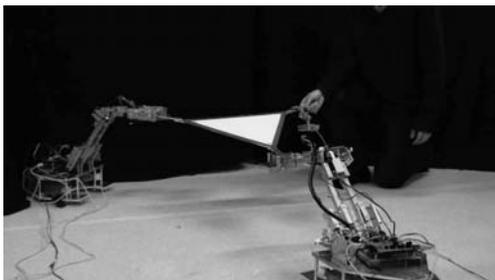


図 1 ヒトとロボット群の協調荷上げタスク

ら約 100 mm であり、タスク達成に必要な荷物の高さは床から約 450 mm である。実験の特徴は以下のように表される。

- 各ロボットは二関節のアーム型ロボットで、各関節にサーボモータとポテンシオメータが取付けられている。
- 各ロボットに搭載されている PC や CPU は同一のものを用いる。すなわち、各ロボットは均質である。
- 各ロボットが知覚できるセンサ情報は、関節の角度と先端の荷物の傾きのみであり、直接的に他ロボットや被験者の状態を知覚できない。明示的な通信を用いていない。
- 各ロボットはそれぞれ独立した学習器により非同期で制御される。
- 持ち上げの成否は、被験者の荷の把持部と各ロボットのアーム先端に取り付けられた IR センサによって判定する。ただし、IR センサはロボットと独立に制御されており、IR センサの情報をロボットが知覚できない。さらには、ノイズによって正しくゴール判定が行われない場合があり、評価系にも不確定な要素を含んでいる。
- 被験者に対し、これまでの研究で観察された振る舞いのうちのリーダ役を担うように指示する。被験者は常に各ロボットのアーム先端を下に見るように、アームを高く保つ行動をとる。荷の傾きの保ち方などについては被験者に一任する。

【ロボットの設定】自作したロボットはアーム部分、台車、制御系から構成される。アーム部分は台車にマウントされている。各関節にはサーボモータが取り付けられており、0.35°の精度で、0~120°の範囲で制御できる。また、実際の姿勢についての情報を得るために各関節にポテンシオメータを装備している。台車の車輪は受動車輪であり、全方向に回転する。

【制御器の設定】入力各関節角度と荷の傾き、

出力はサーボモータの角度変化量である。入出力の1サイクルを1ステップとし、70ステップ経過するかタスクを達成するまでを1エピソードとする。エピソード終了後はロボットを初期位置に戻し、エピソードを更新する。ロボットへの報酬はタスク達成時に与えられ、罰は荷を10°程度以上傾けた場合に与えられる。10エピソード連続でタスク達成したとき、学習成功と定義する。BRLのパラメータはこれまでの研究で得た推奨値を用いた。

【実験結果】まず、基礎実験として、ロボット一台と被験者一名で実験を行う。学習成功後、学習結果を複製したロボットを一台新規投入し、追加実験を行う。

基礎実験において、全10試行中、従来型・拡張型BRLのいずれでも9試行が学習成功した。学習成功までには、従来手法で36.67、拡張手法で31.44エピソードを要した。このことから、提案したSVMによる知識利用法により、学習が高速化できた。そのうち、9試行全てでロボットを一台追加した場合でも再び学習成功することができた。

図2に提案手法を用いた基礎実験での学習履歴の一例を示す。13エピソードまでは一度のみタスク達成をしているが、それ以外では多くの罰を受けている。14エピソード以降は連続してタスクを達成している。

続いて、基礎実験と追加実験におけるアーム

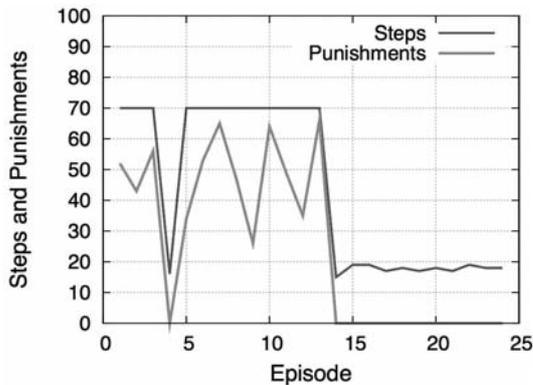
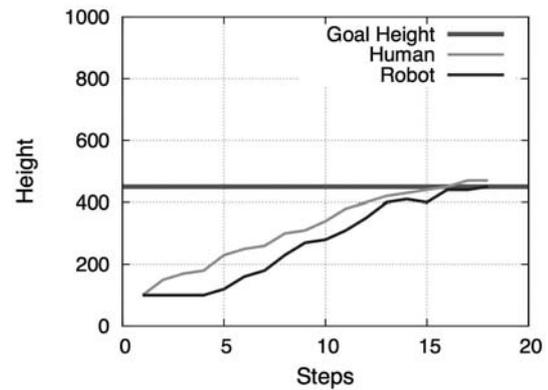
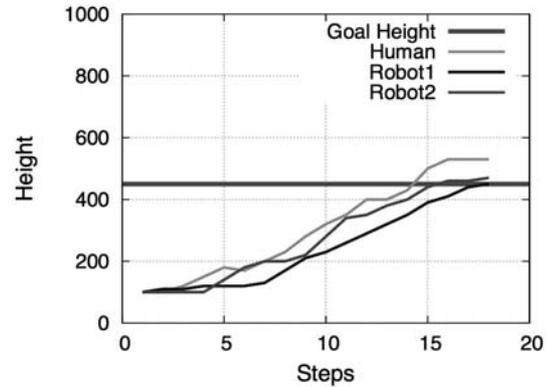


図2 学習履歴の一例

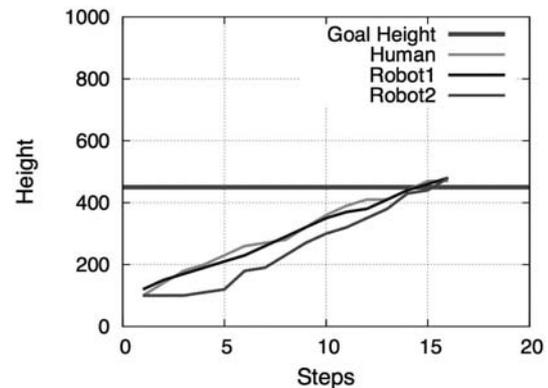
先端高さの推移の例を図3に示す。(a)からロボットはリーダ役の被験者に追従していることがわかる。その後、追加実験では追加投入されたロボットは元のロボットと同一の知識を持つが、(b)と(c)に示すように、同じ行動をするわけではない。また、エピソードの更新後には別のルールを用いたり、異なる方法でルールを切替えていることが確認できた。



(a) 14エピソード



(b) 24エピソード



(c) 25エピソード

図3 アーム先端高さの推移

【考察】 その他の実験として、未学習のロボット二台と被験者一名での協調荷上げタスクなどを取り扱った。その際、被験者との行動獲得後に (i) 同じ役割の別被験者と交代, (ii) 被験者はそのまま異なる役割を変更するなどといった環境変化を行った。被験者が担当するリーダー・フォロワといった役割に関する知識を持たないにも関わらず、ロボットは様々な状況に適應することを確認できた (図 4)。学習の

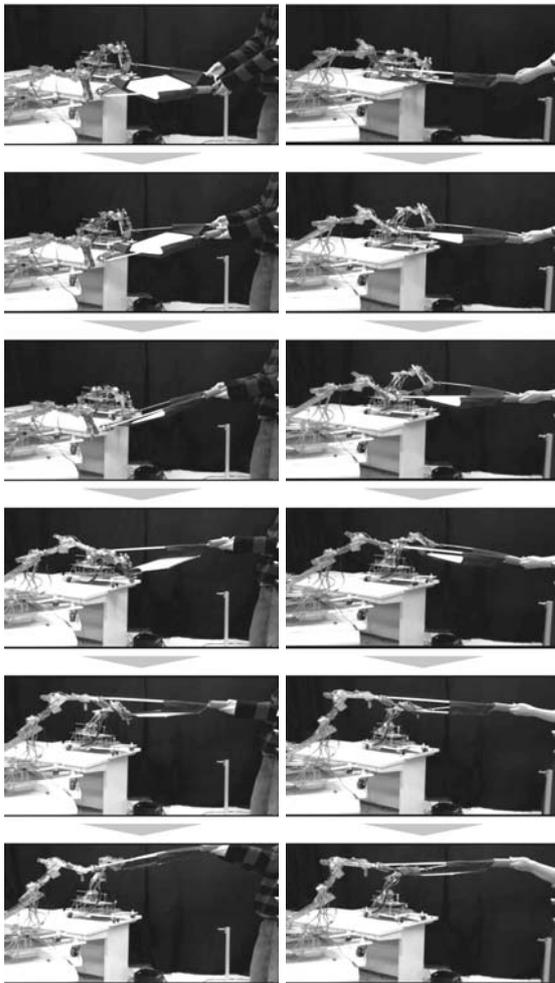
過程で、未経験な状況に遭遇すると新しいルールを生成し、既存のルールを利用した場合にはパラメータを更新する。特に、既存ルールの利用において拡張型 BRL は有効であった。このような知識獲得・利用を通して、様々なパートナーと協調することを確認した。これらの結果から、自律的機能分化に基づく人間機械協調系という新しい知能システム設計論の実現可能性を示すことができた。

#### [今後の研究の方向, 課題]

MRS を構築するロボットの台数やそれらと協働する被験者数の増加を予定している。さらに、より多くの被験者との実験を繰り返すことで協調行動獲得に対する BRL の汎化性能評価を行う。また、本実験では評価の対象外とした獲得行動の効率性や正確性、そしてユーザビリティなどについても、将来的にはより詳細に議論していきたい。

#### [成果の発表, 論文等]

1. T. Yasuda, S. Nomura and K. Ohkura: Self-Organized Task Allocation between Reinforcement Learning Robots and a Human Partner, International Journal of Advancements in Computing Technology, Vol. 4, No. 22, pp. 230-238, 2012.
2. S. Nomura, T. Yasuda and K. Ohkura: Validation of Robustness of Reinforcement Learning Robots Cooperating with a Human, Proceeding of the 16th Asia Pacific Symposium on Intelligent and Evolutionary Systems, pp. 158-163, 2012.
3. T. Yasuda, K. Kage and K. Ohkura: Robust Reinforcement Learning Technique with Bigeminal Representation of Continuous State Space for Multi-Robot Systems, Proceedings of the SICE Annual Conference 2012, pp. 1552-1557, 2012.



(a) リーダの被験者 (b) フォロワの被験者  
図 4 スナップショット