

Title	Cooperative Behavior Acquisition by Learning and Evolution in a Multi-Agent Environment for Mobile Robots
Author(s)	内部, 英治
Citation	大阪大学, 1999, 博士論文
Version Type	
URL	https://hdl.handle.net/11094/41443
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"〉 大阪大学の博士論文について <a>〉 をご参照ください。

Osaka University Knowledge Archive : OUKA

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

Osaka University

氏名	内 部 英 治
博士の専攻分野の名称	博 士 (工 学)
学位記番号	第 1 4 6 2 1 号
学位授与年月日	平成11年3月25日
学位授与の要件	学位規則第4条第1項該当 工学研究科電子制御機械工学専攻
学位論文名	Cooperative Behavior Acquisition by Learning and Evolution in a Multi-Agent Environment for Mobile Robots (マルチエージェント環境における移動ロボットのための協調行動の 学習および進化的手法による獲得)
論文審査委員	(主査) 教授 浅田 稔 (副査) 教授 白井 良明 教授 池田 雅夫 教授 赤木 新介 教授 古荘 純次

論文内容の要旨

本論文では、マルチエージェント環境下において、複数の移動ロボットに協調行動を獲得するための学習法および進化的手法を提案する。本論文は全7章から構成され、第1章では、多くの分野でのマルチエージェント環境下での学習の問題の重要性を指摘する。第2章では、分散人工知能やロボティクスの分野において、複数学習者の問題を扱った従来研究を様々な観点から分類し、その問題点を指摘する。

第3章では、マルチエージェント環境のような動的環境における状態表現の問題について議論する。マルチエージェント環境で協調行動を獲得するために、各学習者は他の対象物との局所的な相互作用を個別に推定し、次数を含めた状態ベクトルを決定した後で強化学習を適用する手法を提案する。また、学習の初期段階における学習過程を安定化させるための学習スケジュールを提案する。

第4章では、環境との相互作用の複雑さに応じて、学習者がビヘービアをどのように発達させるか、ということを議論し、ビヘービアを効率的に改善する手法を提案する。問題の複雑さを状態ベクトルの次数とみなし、他者のビヘービアに応じて状態ベクトルの次元を適切に変更することで、学習時間を短縮できることを示す。

第5章では、複数タスクを扱う方法について議論する。従来スカラーの評価関数で定式化されていた強化学習をベクトルの評価関数の場合に拡張することで、従来よりも学習を安定化できることを示す。また、各学習者に個人としての目標と全体の目標を与え、それらの多目的最適化を解くことで、複数の学習者の協調行動を実現する。

第6章では前章とは異なり、共進化の過程を通して複数の学習者の行動を同時に創発させる方法について考察する。従来の共進化によるアプローチでは、個体間の関係は競合だけに限定されていたが、ここでは各学習者に遺伝的プログラミングを個別に適用し、協調および競合行動を獲得できることを示す。

近年マルチエージェントの標準問題として提唱されているサッカーゲームに適用し、提案した手法の全ての有効性を検証する。最後に第7章では、本論文のまとめと今後の課題について検討をおこなう。

論文審査の結果の要旨

マルチエージェント環境での複数のエージェントに自律的に学習させる場合、(1)状態表現、(2)学習時間、(3)複数報

酬の割り当て、(4)同時学習、といった解決すべき様々な問題点がある。

申請者の研究では、それらの問題に対して以下に示す手法を提案している。

- (1) マルチエージェント環境では、瞬間のセンサ状況だけでは状態を特定することはできない場合が多く、エージェントは履歴を考慮して学習するための状態空間を構成する必要がある。本研究では、エージェントの過去の観測と行動から、回数も含めて状態ベクトルを推定し、獲得された状態ベクトルをもとに強化学習を適用する手法を提案している。また、学習を安定に行うために、1台ずつ学習するエージェントを指定するスケジューリング法を提案している。
- (2) マルチエージェント環境では、学習者以外の他のエージェントの振る舞いが学習者の学習時間に影響を及ぼす。本研究では、簡単な状況から次第に困難な状況を学習すれば、学習が効率化できるという考えに基づき、学習者側と非学習者側の困難さの制御をおこない、学習の高速化を実現している。学習者側は状態ベクトルの回数を困難さの指標とし、非学習者側は移動速度を指標とすることを提案している。他者のビヘービア（移動速度）と与えられたタスクの達成率を監視しながら、状態ベクトルの次元を変更している。
- (3) 従来、強化学習は報酬をスカラー値で定式化していたために、エージェントに複数の行動を学習させる場合には、環境から与えられる報酬を重み付き線形和を新たな報酬として学習する必要があった。本手法では強化学習法をベクトル報酬の場合に拡張し、複数の行動を調停する手法を提案している。ベクトル化することで、個別の評価と協調した場合の評価を別々に考えることができ、パレート最適性の考えを導入することで、複数行動の調停を実現している。
- (4) 3台のエージェントの行動を同時に獲得させるために、進化的手法の一つである遺伝的プログラミングを用いて、協調および競合行動を実現している。同時学習させる場合には、各エージェントに与えられた問題の困難さが均等であることが重要であることを指摘している。

これらの手法はすべて移動ロボットの簡単なサッカーゲームに適用され、その有効性が検証されている。

以上のように、本論文はマルチエージェント環境で視覚を持った移動ロボットが協調および競合行動を獲得するための手法を提案するとともに、本論文で得られた結果は、様々な分野への発展も期待できると考えられる。よって本論文は博士論文として価値あるものと認める。