

氏 名	いの うえ ひろ やす 井 上 寛 康
学位(専攻分野)	博 士 (情 報 学)
学 位 記 番 号	情 博 第 207 号
学位授与の日付	平 成 18 年 3 月 23 日
学位授与の要件	学 位 規 則 第 4 条 第 1 項 該 当
研究科・専攻	情 報 学 研 究 科 シ ス テ ム 科 学 専 攻
学位論文題目	経験一般化を行うルールベース学習に関する研究

論文調査委員	(主 査) 教 授 片 井 修 教 授 熊 本 博 光 教 授 喜 多 一
--------	--

論 文 内 容 の 要 旨

本論文は強化学習における経験活用の体系化とその実現を目的として、経験活用の重要な側面である経験一般化に焦点をあてて論じたものである。経験一般化とは、経験が得られた状況だけに留まらず、異なる状況下においてもこの経験を活用することを指している。本論文は強化学習の分野において経験一般化を(1)状態集約による経験一般化、(2)エージェント間共有による経験一般化、(3)内部状態推測による経験一般化、の3つに整理し、それぞれを代表的なルールベース学習システムである学習分類子システムによって統一的に実現している。本論文は7つの章から構成されている。

1章は序論であり、本論文内容の背景、意義、目的、および論文構成を示している。

2章は論文内容の中心課題である経験一般化について論じている。経験一般化が、状態集約、エージェント間共有、および内部状態推測の3つに分けられることについて述べ、それらの関係を明らかにしている。またそれらを学習分類子システムによって実現する根拠を明らかにしている。さらに、これら3種の経験一般化が学習分類子システムによりどのように実現するかについての基本方針を述べている。

3章は、本論文で取り扱う学習分類子システム (Learning Classifier System (LCS)) について、(1)その基本的な構造と働き、(2)(1)を具体的に実現した Zeroth-level Classifier System (ZCS)、(3)(2)の問題点を改善した Extended Classifier System (XCS)、について詳述している。

4章は、状態集約による経験一般化について、3章で紹介の XCS をベースとしてその実現法を明らかにしている。従来より、XCS においては環境が確定的状態遷移するときのみ状態集約による経験一般化は有効であることが知られている。これは、より複雑な環境下では状態集約による経験一般化が適切にできないためであり、本論文ではその改善法を明らかにしている。すなわち、付加的なメモリ (Q-Table) を用いて XCS における分類子の適応度が正しく求まるように改善し、マルチエージェント環境のような複雑な環境下でも従来の XCS より優れた性能が実現可能であることを実験により示している。

5章は、エージェント間の経験共有による経験一般化を扱っている。すなわち、あるエージェントの経験を、そのエージェントだけではなく他のエージェントでも活用することを目指している。先行研究では、一様はエージェント群全体での経験一般化が効果的であることを示しているだけであり、エージェント間での方策の多様性が要求されるようなときには対応できない問題点が指摘されてきた。本論文ではエージェント間での経験共有の形態を動的に変更する手法を、学習分類子システムに基づいて提案している。エージェントに与えられた機能的制約や環境によって、従来手法が実際に局所最適解に陥ることのあること、また提案手法がそのような場合でも有効であることを明らかにしている。具体的には空間と時間が離散化されたサッカーゲームについてのシミュレーション実験により検証を行っている。

6章は、内部状態推測による経験一般化を扱っている。これは、人間とインタラクションするエージェントにおける経験一般化であり、直接は観測されない内部状態が変化するために、エージェントの最適な方策も変更する必要があるような

複雑な環境である。このような環境下では、1つの状態空間で環境を捉えることは合理的ではないため、複数の方策を用意するとともに、複数の状態空間をインタラクション結果に基づいて識別する手法が有効であることを明らかにしている。人間を複数の有限状態機械からなると仮定し、複数の方策を切り替える方法の有効性をシミュレーション実験により示している。

7章は結論であり、本研究で得られた成果を総括し、今後の研究課題を展望している。

論文審査の結果の要旨

本論文は、強化学習における経験活用法の体系化とその実現を目指し、経験一般化に焦点をあてて論じたものである。強化学習において経験一般化を定義し、それを3種に分類している。さらにこれらを、代表的なルールベース学習システムである学習分類子システムによって実現する方法を提案している。得られた成果は以下のように要約される。

1. 学習分類子システムにおいては、従来より、状態集約による経験一般化が確定的状態遷移環境下で有効であることが示されていたが、確率的状态遷移環境下あるいは部分知覚環境下さらにはマルチエージェント環境下等のより複雑な環境下でも有効であるように学習法を拡張した。

2. 上記の改善は効用の期待値を求めるという単純な方法に依拠しており、高い実用性を有することを明らかにした。

3. 従来のエージェント間の経験共有による経験一般化では、常にエージェント群全体で経験共有しているのに対し、経験共有の形態を動的に変更する方法を提案し、その有効性を明らかにした。従来の方法では局所最適解に陥る場合でも提案手法はそれを避けることが可能であることを示した。

4. エージェント間の経験共有による経験一般化の方法として、その有効性を検出しつつ共有する方法と、役割分担の情報を利用して局所的に共有する方法を新たに提案した。さらに役割分担の情報により探索パターンを効果的に絞り込むことができることを明らかにした。

5. 人間の内部状態変化を想定しそれに対応した経験一般化を行う方法を構築した。これはインタラクションを通じてそれら内部状態を識別するものであり、内部状態を識別しないエージェントより高い効用を得ることができることを確認した。

6. 内部状態推測による経験一般化の方法として、過去の状態空間で得た経験の有効性を報酬量によって検出する方法、および過去の状態空間で得た経験の有効性を環境のモデルによって検出する方法の2つを提案した。環境のモデルを用いる方が過去の状態空間の有効性を並列に検出できるため、より高い効用の得られることを明らかにした。

以上のように本論文は、強化学習における経験一般化法を整理体系化し、それらを学習分類子システムによって統一的に実現する方法を示したものであり、その成果は情報学の展開に寄与するものである。よって、本論文は博士（情報学）の学位論文として価値あるものと認める。また、平成18年2月1日実施した論文内容とそれに関連した試問の結果合格と認めた。