

## 論文内容の要旨

博士論文題目

Learning and Decision-Planning in Partially Observable Environments

(部分観測環境における学習と意思決定に関する研究)

氏名

藤田 肇

本研究は、部分観測環境での方策学習と意思決定に関するメカニズムを理学的・工学的な両面から明らかにすることを目指した。まず、理論研究における成果として、部分観測性を持つ大規模な未知環境におけるモデル同定型強化学習法を提案した。本手法に基づく強化学習エージェントは、部分観測性、マルチエージェント系、大規模な状態空間という3つの困難な条件を持つ環境において方策の学習を行う。想定した条件を満たす応用課題としてカードゲーム Hearts を取り上げ、提案手法に基づいてこの実環境における戦略の獲得を目指した。このゲームでは相手が所持するカードは観測できないため、非観測となるカードを推定しなければならない。また、4人のプレイヤーが競合するこのゲームは複雑なマルチエージェント系となるため、相手プレイヤーのモデルを同定した上で、効果的に行動予測を行う必要がある。さらに、このゲームは巨大な状態空間を持つため、予測と推定に伴う計算困難性を回避するための近似解法が必要になる。本研究では、この近似として期待観測状態を用いる手法とサンプリングを利用する手法を提案した。提案手法に基づく学習エージェントの性能を、計算機シミュレーションにより評価した結果、上級レベルの強さを持つルールベースエージェントや人間のプレイヤーよりも優れた戦略を獲得でき、複雑な部分観測環境における強化学習法として有用であることを示した。

次に、脳神経科学に関する研究成果として、部分観測環境におけるヒトの意思決定モデルを提案した。被験者に対して部分観測迷路課題を課し、課題遂行中の脳をfMRIによって撮像した。被験者は迷路内において自身の周囲に関する限られた観測を得ることができるが、真の位置を知ることはできない。したがって、課題を効率よく達成するためには、部分的な観測情報から迷路内における自身の位置を推定する必要がある。また、こうした部分観測状況においては、推定の精度を高めるための探索行動と、現在の推定に対する確信に基づく搾取行動が競合するため、二つのモードを適切に切り替えながら意思決定を行わなければならない。提案モデルは部分観測迷路課題に対する被験者の意思決定過程を適切に説明でき、隠れ状態の推定と探索・搾取の戦略切り替えに関わる脳機能の解明に有用であることを示した。

## (論文審査結果の要旨)

実世界において、周囲を取りまく環境に関する情報は感覚を通して知覚できる不完全な入力に依存する。そのため、生物や知的システムにとっては、隠れた情報を推定によって補いながら学習と意思決定を行う機構が不可欠である。この、生物の脳内で実現されている高次認知機能の仕組みを解明し、これを利用して真に有用な計算機システムを開発する試みは、理学的にも工学的にも大きな意義がある。本論文では、環境に関する情報が部分的にしか得られない状況を模した二つの課題を用いることで、上記の本質的なメカニズムを明らかにすることを目指している。本論文の主な成果は以下のように要約される。

1. 部分観測性を持つ大規模な未知環境におけるモデル同定型強化学習法を提案した。本手法を、複雑な実問題のひとつであるカードゲーム“Hearts”へ応用した結果、上級レベルの強さを持つルールベースエージェントや人間のプレイヤーよりも優れた戦略を獲得できることを示し、本手法が部分観測性を有する実環境において学習と意思決定を行うための有効なアルゴリズムであることを証明した。
2. 部分観測迷路課題における被験者の意思決定モデルを構築し、認知課題の遂行に関する情報処理機構に関する仮説を検証した。提案モデルは部分観測迷路課題に対する被験者の意思決定過程を適切に説明でき、隠れ状態の推定と探索・搾取の戦略切り替えに関わる脳機能の解明に有用であることを示した。

一般的なゲームは問題の規模が巨大であるため多くの隠れ状態を含み、複数人でのプレイを仮定することから協調・競合関係が発生する。そのように複雑な実環境において適応的学習を試みた例は少なく、本研究が工学分野に与えるインパクトは大きい。また、認知実験によって理論的な知見を実験的に検証することを目指しており、異なる二つの分野を結ぶ試みとして今後も成果が期待できる。本論文では、脳内で行われている情報処理、特に環境変動の予測と非観測状態の推定を伴う高次の認知機能の仕組みを統計的手法に基づいて明らかにすることを主眼としており、人間の知能の謎を解き明かすという理学的な興味を惹くだけでなく、その成果は、複雑度の高い機械をユーザの認知モデルに基づいてトップダウンに制御する機構の開発など工学的な応用も考えられる。以上のように、意義深い成果が得られており、博士（工学）の学位論文として価値あるものと認める。