

## 論文内容の要旨

博士論文題目

部分観測システム同定の応用と脳型情報処理に関する研究

氏名 吉田 和子

(論文内容の要旨)

ヒトを取り巻く自然環境は、直接観測できない事象を含み、時間と共に変化する。このような複雑な環境においても、ヒトは環境の特性を学習し、その変化を予測することによって最適な行動を取る。つまり、動的且つ部分観測環境におけるシステム同定を行い、それに基づいて最適意思決定問題を解いていると考えられる。観測に基づくシステム同定とその予測問題は、認知科学的な問題としてだけではなく、工学的問題としても興味深い。本研究は、部分観測システム同定を機械学習と脳型学習の両方の枠組みで捉え、理論的研究と実験的研究の両側面から明らかにしようとするものである。本論文は、融合領域研究であるシステム脳科学の立場から論じたものである。

まず初めに、統計的手法による非線形力学系のシステム同定問題について議論する。ここで、システム同定とは、未知のシステムの入出力データを用いて、そのデータを再現するようにダイナミクスを模倣することである。本研究では、低次元カオスシステムを対象として取り上げ、関数近似器を用いてカオス力学系のダイナミクスの学習を行う。部分観測問題を取り扱うために、例えば1次元時系列データから元の状態空間での位相構造を構成する手法として、2種類の埋め込み法を用いる。初めに、一般に用いられる遅れ座標埋め込み手法を用いた結果から、部分観測においてもシステムを同定できることが示される。次に、この手法を発展させた IIR フィルタを用いた積分埋め込み法が提案される。学習データにノイズを付加した実験結果から、積分埋め込み手法は従来のものと比べてより近似精度が高く、さらにノイズに頑強であることが確認される。

次に、部分観測環境におけるシステム同定とそれに基づく意思決定法について議論する。機械学習において、最適性を報酬に基づき定義すると、確率的環境における最適意思決定問題は、マルコフ決定過程 (MDP) として定式化される。直接観測できない変数(隠れ変

数)が存在する環境は、部分観測 MDP としてモデル化する。本研究では、隠れ変数をもつ環境でのモデル同定強化学習法を提案する。モデル同定強化学習では、過去の経験から環境を同定し、それに基づいて意思決定を行う。本研究では、環境の同定に忘却効果を導入したベイズ推定が用いられる。また、動的に変化する環境を扱うため、環境の変化を検出しそれに基づいて行動様式を変化させる手法を提案する。本手法を隠れ変数を持つ迷路探索問題に適用した結果、従来の手法よりも環境の変化にうまく適応できることが示される。

最後に、モデル同定強化学習を実現する脳の情報処理モデルを提案する。相互作用によって環境システムを同定し、それに基づいて意思決定を行うという学習過程は、ヒトの学習法としても妥当である。本研究では、モデル同定強化学習で用いる主な関数と、脳、特に前頭前野の機能を対応付けた、強化学習脳内モデルが提案される。このモデルでは、背外側前頭前野が環境モデルの保持と操作に、前部前頭前野が直接観測できない環境変数の推定に関わるとみなされる。また、行動選択を行う部位として、帯状回が想定される。このモデルを検証するために、核磁気共鳴画像を用いた脳計測実験を行ったところ、提案したモデルと矛盾しない結果が得られた。

(論文審査結果の要旨)

部分観測システム同定とは、対象システムから部分的に観測されるデータをもとに、そのダイナミクスを同定することである。この問題は、システム制御などの工学分野において重要である一方で、脳内情報処理機構でもあると考えられる。本論文は、システム脳科学の観点からシステム同定問題を取り上げ、アルゴリズムやモデルの開発といった理論的研究と、モデル検証のための認知科学的研究の両面からの研究をまとめている。本論文の成果は以下のように要約される。

1. 部分観測システム同定は、ヒトのコミュニケーションにおいて重要になる。この問題意識から、長距離依存構造のあるカオス時系列を対象として、システム同定と時系列予測を行う手法を開発した。カオス時系列はその初期値鋭敏性から長期予測が不可能であるが、高速な統計的学習手法と大量の学習サンプルを用いることにより、実用上問題がないくらいに限界を広げられることが示された。また、部分観測データから元の力学系を再構成するための積分埋め込み手法を提案した。シミュレーション実験により、提案手法は従来の手法よりも近似精度が優れており、ノイズへの耐性が非常に高いことを示している。
2. 経験に基づいた環境の同定と最適意思決定を行う工学的手法として、モデル同定強化学習を取り上げ、動的な環境に適した手法を開発した。隠れ状態変数の推定にはベイズ推定を用いた。ベイズ推定に無情報事前知識と忘却効果を導入することで、環境への適応性が促進をはかっている。さらに、学習戦略を規定するメタパラメータを環境に応じて決定する制御機構を開発した。実験により、開発手法が従来手法よりも環境への適応能力が高いことを示している。また、従来任意に決められていたパラメータをエージェントの状態から自動的に決める手法は新しく、人間の労力を減らすことができるという応用面での波及効果も大きい。
3. 開発した学習アルゴリズムに対する生物学的に妥当な脳型情報処理モデルを提案し、非侵襲脳計測機器を用いた検証実験を行った。実験結果は、提案モデルの妥当性を示唆するものであった。近年、脳神経系の分子機構が数理的な学習アルゴリズムと関連付けられることが指摘されており、理論モデルと脳内メカニズムの対応付けが注目されている。本研究は、脳機能をトップダウンとボトムアップの両側面から理解しようとするものであり、その課題設定は高く評価できる。

本論文の成果は、未知の環境のモデル化とそれに基づく最適意思決定問題について重要な知見を与えるものであり、工学、特にシステム制御と学習理論の分野において学術上、および応用上の寄与が少なくない。また、脳情報処理機構を異分野の手法や知見を融合することによって解明しようとする研究アプローチは、理学、あるいは学際領域である脳科学研究の発展に大いに貢献するものである。よって、博士(理学)の学位論文として価値あるものと認める。