

論文審査の結果の要旨および担当者

報告番号	※	第	号
------	---	---	---

氏 名 遠山 朝子

論 文 題 目 Reinforcement Learning in Humans: Rethinking
Psychologically Plausible Model-Free and Model-Based Algorithms

(ヒトにおける強化学習: 心理学的に妥当なモデルフリー・モデルベース
アルゴリズムの再考)

論文審査担当者

主 査 名古屋大学大学院情報学研究科 教授 大平 英樹

副 査 名古屋大学大学院情報学研究科 教授 田邊 宏樹

副 査 名古屋大学大学院情報学研究科 准教授 片平健太郎

論文審査の結果の要旨

本論文は、ヒトの意思決定を表現する計算論モデルとして近年盛んに研究されている、モデルフリー・システムとモデルベース・システムという2つの過程を含む強化学習モデルについて、ヒトの情報処理容量の制約を考慮したアルゴリズムに基づく新たなモデルを提案し、その妥当性を実験とシミュレーションによって実証的に検討したものである。

モデルフリー・システムは経験によりボトムアップ的に状態の価値を更新する仕組みであり、モデルベース・システムは環境の構造をトップダウン的に表象する内的モデルを利用して状態の価値を計算する。これら2つのシステムの機能バランスは、認知能力の個人差、あるいは発達障害や精神疾患にも関連するとして注目を集めている。Dawら(2011)が提唱した計算論モデルは、これら2つのシステムの動作原理を表現するものとして広く認められている。しかし彼らのモデルでは、モデルフリー・システムとモデルベース・システムにおける状態の価値が並行して計算されると仮定されていること、特にモデルベースの価値計算には状態の遷移に関する情報と次の段階における状態の価値が必要となることから、計算負荷が高いことが問題であると指摘できる。情報処理能力に制約のあるヒトの脳が、本当にDawら(2011)のモデルと同様な計算を行っているか否かは未だ検討されていない。

そこで本論文では、第1章において先行研究の知見に基づいて問題点を整理した上で、第2章で上記の2つのシステムを融合させ、価値の計算は基本的にはモデルフリー・システムで行われるが、モデルベース・システムがトップダウン的にその計算を重み付けると仮定する新たな計算論モデルを提案した。さらに第3章では、モデルフリー・システムのモデルを改善する試みを行った。従来のモデルでは、選ばれなかった選択肢の価値はそのまま維持されると仮定されていた。しかしヒトの記憶には忘却が働くため、選ばれなかった選択肢の価値は一定のデフォルト価値に回帰していくという仮定を導入した。第4~6章においては、ヒト参加者を対象にした実験研究、数値計算によるシミュレーション研究を行って、新たに提案したモデルが従来のモデルより妥当性が高いことを実証した。第7章では総合的な考察を行い、今後の研究の展望について議論した。

本論文は、意思決定の計算論的研究における最先端の問題を扱っており、そこにおいて標準的と考えられてきたモデルとは異なる発想に基づいた新たなモデルを提案し、その妥当性と優位性を実証的に示したという点で、高い学術的価値と独創性を有すると評価できる。本論文で提案したモデルが、実際の脳における情報処理とどの程度対応しているか、また精神疾患の理解や治療などにおいてどの程度有益であるかは、さらに慎重に検討されねばならないが、これは今後の課題である。

よって本論文の提出者遠山朝子君は、博士(心理学)の学位を授与される資格があるものと判断した。