

報告番号	※甲	第	号
------	----	---	---

主 論 文 の 要 旨

論文題目 Elucidation of the cognitive computational process underlying the behavior of pursuing the unprofitable targets

(手に入らない対象を追求する行動の基盤にある認知計算過程の解明)

氏 名 菅原 通代

論 文 内 容 の 要 旨

試行錯誤学習において、ヒトを含む動物は肯定的な結果を伴う行動に対して、より多くの反応を示すようになる。効果の法則と呼ばれるこの理論は、後に洗練された形で強化学習理論として定義された。現在、心理学における強化学習は、背景に行動の結果に基づく価値更新を想定しており、ヒトの学習過程の多くの側面を捉えることが示されている。しかしながら、時にヒトは非合理的に見える行動をとる。例えばストーカーは、相手から良い反応が返ってこないにもかかわらず、同じ対象を追い続ける。このような効果の法則に反する追跡行動はなぜ起こるのだろうか。本論文は、選択課題において選択結果から計算される価値更新（学習率の非対称性）と、選択そのものから計算される履歴更新（選択の固執性）を仮定した強化学習モデルを用いて、「手に入らないものを追い続ける行動」の基盤にある認知計算過程を明らかにしようとするものである。

第1章では、行動データに計算論モデリングを使用する意義や強化学習モデルの理論的背景、およびヒトの認知バイアスとしての学習率の非対称性と選択の固執性についての先行研究を概観した。ある対象を持続的に追い続けるということは、言い換えれば同じ行動を繰り返すということである。これまでの研究では、選択行動の反復を説明するために、選択に伴う結果からの情報と、選択したという選択自体による情報のいずれか一方の更新過程のみを定式化したモデルが用いられてきた。前者の学習率の非対称性モデルは、これまでの学習から予測される報酬値と実際の結果として得た報酬値の差分（報酬予測誤差）をどの程度価値として更新するかという重み付けパラメータ（学習率）に非対称性を許容する。この非対称性により、自分にとって都合の良い情報（正の予測誤差）を多く取り入れ、都合の悪い情報（負の予測誤差）は取り入れないといった認知バイアスを表現できると考えられている。この非対称な価値更新は、同じ選択を繰り返す行動を促す。一方、後者の固執性モデルは過去に自分が選んだという情報を選択履歴として表現する更新過程を組み込み、選択履歴に対する重みづけが大きいほど、選択の反復が生じるとする。これらの報告から、「手に入らないものを追い続ける行動」の背景にある認知計算過程と

して、選択の結果と選択自身に由来するいずれの計算過程も考慮する必要があると考えられる。

このように独立した計算過程が類似した行動へと帰着する場合に、2つの過程を計算論モデリングで分離することは可能なのだろうか。仮に適用したモデルに選択行動に大きく影響する計算過程が含まれていない場合（例えばデータに固執性が含まれているにもかかわらず、その影響を無視したモデルを適用した場合）、パラメータ推定にバイアスが生じ、誤った解釈を導く可能性がある。

そこで第2章では、学習率の非対称性と固執性の両方を組み込んだハイブリッドモデルの有用性を検討した。第一に、シミュレーションを行うことで、ハイブリッドモデルが想定する認知計算過程を正しく分離し、測定できることを示した。第二に、オンライン実験により選択行動データを収集し、その選択行動にみられた反復行動は学習率の非対称ではなく選択の固執性をより強く反映していることを明らかにした。第三に、過去に学習率の非対称性を報告している2つの公開データに対してハイブリッドモデルを適用した。その結果、一方の公開データには学習率の非対称性でしか説明できない要素が含まれるが、他方の公開データは選択の固執性だけで十分に説明できることを明らかとした。これらの結果は、選択行動の背景にあるプロセスを正しく推定するためのツールとしてハイブリッドモデルの有用性・妥当性を示すとともに、モデルの誤使用が誤った解釈を生み出す可能性を示唆するものである。

第3章では、第2章で有用性が示された学習率の非対称性と固執性の影響を切り離すことができるハイブリッドモデルを用いて、「手に入らないものを求め続ける行動」の背景にある認知計算過程を実験的に検証した。オンライン実験において、恋愛シミュレーションを模したアバター選択課題を用いて、「手に入らないものを求め続ける行動」を計測し、課題の前後でアバターに対する主観的な価値（魅力度）も調査した。選択課題では、良い結果を返す確率の高い「簡単なアバター」とそれが低い「難しいアバター」を設定した。148名中68名の実験参加者は、良い結果をほとんど返さないにも関わらず「難しいアバター」を繰り返し選択した（追跡群）。この行動傾向がどのような計算過程のバランスで表現できるかを検討するためにシミュレーションを行い、選択の固執性が高い時に、難しいアバターを追求する傾向が出現することを明らかにした。そして、実際の選択データに対してハイブリッドモデルを適用することで、追跡群の方が非追跡群よりも有意に高い選択の固執性を持つことを示した。さらに、選択課題を行う前は各アバターに対する主観的魅力度は同等であったにもかかわらず、追跡群では選択課題後に「難しいアバター」の魅力度が有意に高くなっていた。これらの結果から、選択の固執性が「手に入らない対象を追い続ける行動」を導き、追い続けることが対象をより魅力的に感じさせると結論づけた。

第4章では、本研究を通じて明らかとなった知見を総括し、「手に入らないものを追い続ける行動」の背景にある選択の固執性がどのように生じるかについて考察を行った。本論文では、望ましい結果が得られないにもかかわらず、選択自体が対象をより魅力的に感じさせる点に着目した。我々は選択を行う過程において、選択によって引き起こされる可能性のある結果を想像する。選択によって引き起こされた現実の結果からの学習と、選択中に想像した結果からの学習が並行して生じるといふ二重更新モデル (Dual-updating model) を提案した。そして、この新たな心理モ

デルを検証するための将来展望についても論じた。

「手に入らないものを追い求める行動」の背景にある心理メカニズム論じた本研究は、社会的に問題となる不適応行動の理解と治療だけでなく、情熱や忍耐といったヒトの魅力的な心理過程の解明にも貢献すると考えられる。