

(続紙 1)

京都大学	博士 (情報学)	氏名	片山 梨沙
論文題目	Computational Modeling and Neural Mechanisms of Hierarchical Inference and Decision-making in Partially-observable Environments (部分観測環境における階層的推論・意思決定過程のモデル化とその神経基盤に関する研究)		
(論文内容の要旨)			
<p>In the real world, decision-making inherently involves uncertainties arising from multiple sources of information. Particularly in environments with hierarchical hidden states, efficient decision-making requires resolving these uncertainties through mutual inference and incorporating higher-order uncertainties into lower-order inference and decision-making. Hierarchical information processing mechanisms, imitating human behaviors, are being studied in the field of machine learning with the aim of applying them to artificial intelligence. In the fields of neuroscience, however, most previous studies have used simple experimental tasks and flat computational models, resulting in the neural mechanisms of hierarchical information processing in complex and uncertain environments faced by humans in the real world, as well as the effect of metacognition on these processes, remaining unclear.</p> <p>In this study, to investigate the neural substrates contributing to hierarchical information processing in partially observable environments, we designed hierarchical inference and decision-making tasks that extend traditional simple partially-observable problems, and conducted behavioral modeling analyses and functional magnetic resonance imaging analyses using computational approaches. In the first study, using the partially-observable maze exploration task, we demonstrated that scene prediction based on uncertain state inference can be decoded from brain activity patterns in the fronto-parietal network, and that the distinctiveness of its neural representation varies depending on the level of uncertainty. In the second study, we developed a new spatial navigation task incorporating a hierarchical partially-observable problems, demonstrating that the information processing in the human brain during this task is reproduced by a hierarchical inference model in which beliefs of two types of hidden states are mutually updated. We also showed that the uncertainties of these multiple hidden state inferences are represented in distinct subregions of the medial prefrontal cortex. In the third study, using a bandit task incorporating contextual inference, we demonstrated that the confidence in higher-order hidden states modulates lower-order decision-making and its metacognitive assessment. Furthermore, it was revealed that the insular-prefrontal network is involved in such decision-making relying on uncertain higher-order information (i.e., inference of the environmental state).</p> <p>Together, the findings in this study suggest that a broad brain network, including the parietal, medial frontal, and prefrontal cortices, is involved in hierarchical information processing in partially observable environments, with distinct brain regions responsible for different components of such processes. Specifically, our findings show that the posterior parietal cortex encodes the content of predictions and inferences, while the medial and anterior parts of the prefrontal cortex encode the value and confidence</p>			

evaluations of one's actions. Additionally, the dorsal anterior cingulate cortex and insular cortex represent the uncertainty of hidden states and modulate cognitive processing through functional connections with other brain regions.

(論文審査の結果の要旨)

近年、ヒトや動物を用いた行動および脳活動計測実験により、意思決定に関与する複数のモジュールが脳内で階層性をなしつつ協調的・競合的に作動し、その統合的出力として行動が決定されていることが示唆されている。ほ乳類の階層的処理機能を模倣する人工回路は、人工知能への応用を念頭に機械学習分野でも研究が行われてきた。しかし、情報工学・神経科学分野ともに、従来研究の多くは比較的単純な実験タスクと計算モデルを用いたものであり、ヒトが実世界で直面するような複雑かつ不確実な環境における階層的情報処理機構や、その処理過程におけるメタ認知の影響は明らかになっていない。本論文は、従来のシンプルな部分観測マルコフ決定過程 (POMDP) 問題を拡張した階層的推論課題および階層的意決定課題と対応する計算モデルを提案し、行動計測と機能的核磁気共鳴図 (fMRI) による脳活動計測によって、ヒト脳内の階層構造をなす複数のモジュールが構成する意決定の神経基盤の解明を目指したものである。具体的には、以下の成果について報告している。

1. 部分観測環境における予測の脳内表象と主観的確信度との関連

ベイズ脳仮説においては、信念から生成される仮想的な予測が、不確実環境での信念更新の要となる。しかし、従来研究では、予測が脳活動パターンに符号化されていることは完全観測環境下で示されているにすぎず、不確実環境においては検証されていなかった。本論文での研究では、部分観測迷路探索課題と fMRI による脳活動計測を用いて、不確実な位置推定に基づくシーン予測が頭頂・前頭ネットワークの脳活動パターンから解読可能であり、予測の脳活動パターンへの符号化の明瞭性が、被験者が報告する主観的不確実性（ここではメタ認知に対応）およびモデルが予測する客観的な不確実性に依存して異なることを示した。また、頭頂葉の異なる領域が階層的に空間的予測に関与している可能性を示唆した。

2. 階層的POMDP問題における信念の不確実性に関与する脳領域の特定

前項で用いた部分観測迷路探索課題では、信念更新に利用される外界情報（シーン）自体は完全観測であった。しかし実世界での観測は、センサの不完全性などにより何らかの不確実性を内包する。これに対し本論文の研究では、古典的POMDPの一つであるTiger problemを組み込んだ新しい部分観測迷路探索課題を考案し、この課題実行中のヒト脳内での情報処理過程が、2種類の隠れ状態が各々の信念形成・更新にもう一方を利用するとした階層的推論モデルにより再現できることを示した。また、この階層的推論過程における複数の隠れ状態推論の不確実性が、前頭前野内側部に表象されていることを示した。

3. 階層的意決定過程におけるメタ認知修飾の神経基盤の特定

実世界の意決定においては、しばしば、環境に対する知識など、観測から信念を形成するために必要な事前情報にも不確実性が内包されている。こうした階層的な不確実性を処理する意決定過程や階層間でのメタ認知の相互作用については、それらの脳神経基盤は明らかになっていない。本論文の研究では、環境に対する状態推論を組み込んだバンディット課題を用いることで、上位階層の不確実性が下位階層の推定・意決定およびそのメタ認知を修飾する脳内情報処理過程を計算モデル化した。またモデルベースの解析により、このプロセスを内側前頭前野-島皮質間ネットワークが担っていることを示した。さらに、行動レベルの個人間差異が、このネットワークを構成する脳領域間の機能的結合の強さによって説明できることを見出した。

まとめると、本論文では、不確実環境における階層的情報処理および関連するメ

メタ認知過程を計算論モデル化として示し、これを実装するヒト脳神経系メカニズムについて調査した。メタ認知を含む階層的情報処理は、近年注目を集める自己意識の神経機序解明の端緒ともなりえ、情報学のみならず、人間科学・神経科学、さらには、精神疾患等に見られるメタ認知異常・意識障害の機序理解やその治療指針の策定など、医療応用にも将来的に貢献しうる。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。

令和6年8月20日、論文内容とそれに関連した事項について試問を行った結果、合格と認めた。また、本論文について、令和7年9月23日以降でのインターネットでの全文公表について支障がないことを確認した。