

## 〔短期在外研究〕

研究者	国立研究開発法人理化学研究所 基礎科学特別研究員 磯村 拓哉	2176101
共同研究者／所属・職名	Karl Friston / Full professor, Wellcome Trust Centre for Neuroimaging, UCL	
滞在期間	2018年1月5日～3月30日	
滞在地	London, United Kingdom	
研究課題	自由エネルギー原理に基づく個体間相互学習の情報熱力学的理解 Information theoretical approach to understand mutual learning between agents under the free-energy principle	

## 概 要：

身近な人と話すとき、我々はそれぞれの人の思考を区別して想像することができます。この優れた能力は、聞いている声を出した人を特定し、その精神状態を推測するための仮説を切り替えることで行われます。この能力をモデル化するために、我々はベイジアン・フィルタリングとモデル選択を統合した新しい学習手法を開発しました。本研究は、近年着目されている内部モデル仮説および自由エネルギー原理に基づいており、生理学的に妥当な複数個体間での相互学習の数理モデル、あるいは社会的相互作用を通じた社会環境への適応の数理モデルを構築することを大きな目的としています。本手法では、人工知能は複数の内部生成モデルを並列に保持し個々のモデルを基に知覚データ（例えば音声）から隠れ状態変数を推論する一方で、オンラインのベイズモデル選択により個々の仮説の尤もらしさを評価し、知覚データを生成した可能性の高いモデルのみがモデルパラメータを更新します。我々は本手法を用いて、人工知能に複数の鳥の歌を区別し学習させることに成功しました。このデモンストレーションでは、6羽のゼブラフィンチの歌をランダムに入力し訓練したところ、入力された歌を生成した可能性の高いモデルを的確に選択し、生成モデルのパラメータを正しく学習することで、人工知能にその歌を真似して歌わせることができました。この結果は、人工知能が複数の生成モデルを並列に保持し環境に応じて使い分ける学習方法の有用性を示唆しています。本成果は、複数の個体が互いに学習し合う状況における内部状態・パラメータの時間発展に関する一般的法則性の導出に貢献できると期待されます。