

論文内容の要旨

博士論文題目

Goal-Oriented Representations of the External World:
A Free-Energy-Based Approach

(目的指向的な外界の表現に関する研究：自由エネルギーからのアプローチ)

氏名 大塚 誠

The central goal for living organisms or intelligent systems is not to model the world as accurately as possible, but to act in their environments to achieve some goals. However, efficient decision making in the real world requires successful encoding of noisy, high-dimensional sensory inputs and representation of the implicit constraints in the environmental dynamics. In this dissertation, we explored the methods for goal-directed sensory representation and decision making in partially observable Markov decision processes (POMDPs) with high-dimensional, noisy, multimodal sensory inputs and unknown dynamics.

First, we investigated whether free-energy-based reinforcement learning (FERL), which was known to handle Markov decision processes (MDPs) with high-dimensional states and actions, can handle an easy class of POMDPs, in which the detection of true state behind the noisy, high-dimensional observation reduces the problem MDPs. Using a novel “digit-floor” task, we verified that reward- and action-dependent sensory coding in the distributed activation patterns of hidden units despite large variations in the sensory observations for the hidden state. Second, the FERL was combined with recurrent neural networks to handle POMDP problems that require dynamic combination of sensory inputs. Using both low-dimensional bit patterns and high-dimensional binary images, we verified that the dynamic task-structure was implicitly reflected in the time-varying hidden unit activations.

These results show that our dynamic extension of FERL can construct distributed representation of the external world autonomously while solving realistic sequential decision making problems. This approach, which is compatible with Friston’s free-energy principle, may provide the basis for biologically plausible models of representation learning in the brain.

(論文審査結果の要旨)

生物の目的は、何らかを達成するために行動することであり、外界を正確に表現することではない。しかしながら、効率的な行動選択を行うためには、高次元でノイズの多いセンサ入力からの情報抽出や、外界のダイナミクスに関する何らかの情報表現が必要となる。本論文では、自由エネルギーを用いた強化学習手法を持つ目的指向的なセンサ符号化の特性を明らかにした上で、ダイナミクス未知の環境下で高次元センサ入力と報酬の経験だけを頼りに部分観測問題を解く機械学習手法を提案している。

本論文では、まず始めに、高次元入出力のマルコフ決定過程 (MDP) 問題を扱うことができる自由エネルギーを用いた強化学習 (FERL) 手法を、手書き文字を入力として用いる digit floor タスクに適用し、その特性を調べている。その結果、真の状態に関する情報が、報酬や最適行動に依存する形で隠れ層の活性パターンに分散表現されていることを示した。また、ノイズに対する強いロバスト性や、報酬を抜いたときにはセンサ情報のみに依存したコーディングが行われることも示した。その次に、自由エネルギーを用いた強化学習手法を、エルマン型のリカレントネットワークと組み合わせることにより、部分観測問題を扱えるように拡張している。その結果、ダイナミクス未知の環境下でも、センサ入力と報酬経験だけを頼りに部分観測問題が解けることを示した。この際、タスクに関係する情報が隠れ層に動的に表現されていることも示した。また、この手法が低次元入力 (ビットパターン) だけではなく高次元入力 (手書き文字) をセンサ情報として用いたタスクにも有効なことを示した。

自由エネルギーを用いた強化学習手法をリカレントネットワークと組み合わせた拡張により、実世界での逐次行動選択に近いタスクを解くことができただけでなく、タスクに応じて外界からの入力がダイナミックに分散表現されていることも示唆された。このアプローチは、自由エネルギー原理を、視覚系などでの特徴抽出のモデルとして提唱されている deep belief network を用いて実現したものであり、生物学的に尤もらしい状態表現獲得モデルとなる可能性もある。これらの研究は、機械学習の主要分野である強化学習の手法開発という工学的貢献のみならず、神経科学への貢献にも期待される。以上のように、工学における手法開発のみならず周辺分野への波及効果もあり、博士 (工学) の学位論文として価値あるものと認める。