

感情の価値計算システム仮説にもとづく前頭葉推論モデルの検証

Validation of prefrontal inference model derived from emotion as a value calculation system hypothesis

宮田 真宏*¹
Masahiro Miyata大森 隆司*²
Takashi Omori^{*1} 玉川大学大学院 工学研究科 ^{*2} 玉川大学 工学部
Graduate School of Engineering, Tamagawa University College of Engineering, Tamagawa University

There were many modeling studies on the emotion. But most of them were phenomenological and don't approach to the brain and/or cognitive mechanism. In this study, we discuss a possibility of its computational modeling based on an idea that an emotion in wider sense, which we call feeling, is an appearance of a value calculation for an action decision. Therefore, in this study at first, we focused on the elements of emotions, and discuss on a set of computational model for each of the elements. Next, we discuss on a model of an intuitive inference in prefrontal cortex that is, we assumed, one of the brain emotion system elements. We applied Monte Carlo method for the inference modeling and discuss for a validity of the hypothesis.

1. はじめに

近い将来に生み出される AI 技術を用いた製品の応用例として対人インタラクションが挙げられる。対人インタラクションでは AI は他者の認識状態や意図を推定し、その結果に応じた意思決定を行わねばならない。そのため、人の意図や目的(意思決定)の背後にある人の価値計算システムの理解は、高度な対人インタラクション AI の実現には欠かせないステップである。

一方で、人は価値と同様に感情によっても意思決定を行う。感情は時として不合理なこともあるが、日常生活の多くの場面では合理的な価値計算と同等の意思決定を導いている。とすると、意思決定に関して感情は価値計算と類似の役割を持つように見える。そこで本研究では、感情とは動物が行動する際の意思決定に用いる価値計算システムの表出であると仮定し、その計算モデル化の可能性を検討する。本稿では、感情と情動は基本的には同じものと考え、議論する。

2. 従来の感情モデル

感情に関してはこれまでも多くのモデルがある[Ekman 1997][Russell 1980]。しかし、それらの多くは感情を現象として捉えて表した記述モデルであり、脳や認知のメカニズム、さらにその計算論的意味について迫るものではない[Samsonovich, 2013][Chernavskaya, 2015]。また計算論的なモデルも提案されているが、まだ単純な段階にある[Vallverdú, 2016]。

また戸田は、人間の複雑な感情から行動決定に至るまでの心的過程に対して定性的な説明理論を構築している[戸田 1992]。その理論では、人の感情は自らがおかれている場面に価値を割り振って意思決定に至るまでの計算過程であるとしている。その説明は非常に魅力的ではあるが、感情の具体的な処理過程については述べていない。

一方で[Koelsch 2015]は、人間の広義の情動を生存・安全・愛着・経済的価値の 4 要素に分解し、これらを個別の脳部位に対応付けて説明を試みている。この研究では、情動に経済的価値も含まれており、眼窩前頭葉がその機能を司るとしている。これらのモデルの背後にある脳としての認知・計算的なメカニズム、さらにその役割はどのようなものであろうか。

3. 感情＝価値計算システム仮説

3.1 感情と価値計算システムとの関係

感情は意思決定に大きく影響する。また我々の感情表出は他者の意思決定にも影響する。現在、経済学や脳科学では人の意思決定は報酬(価値)の計算に基づくとされている[大竹 2012]。これより我々は、感情とは価値計算の結果の表出であり、感情の背後には脳の価値計算システムがあると想定する。価値と感情表出の対応は比較的固定であり、それがゆえに価値伝達システムとしての確に機能できる。感情のメカニズムを理解するには、むしろ価値システムの理解が先に必要であろう。

従来の AI では知能の要素として、視覚や聴覚などの感覚情報処理、それらを用いた予測や意思決定などの情報処理、さらにそれらを用いた行動生成の処理、を考えてきた。それに対して本研究では、知覚情報処理の背後にある報酬系、すなわち価値について考える。本稿で考える価値計算システムでは、新皮質の知的情報処理系に感覚情報が入力され、その認識の結果は直ちに脳の多様な価値を計算する処理系に伝達されると想定する。我々の考える感情の価値には Koelsch の提案した「生存」「安全」「愛着」「経済的価値」の 4 種を想定する。しかし、それらに対応する脳部位(脳幹・辺縁系・海馬系・前頭葉系)は彼らの考える部位よりも広く考える。これらの部位を統合すると、大脳皮質下の脳部位をかなりの部分を含んでいる。

3.2 価値に基づく推論のモデル

本研究では、これらの価値処理系の各部位で計算された多様な価値が、前頭葉系で行われるとされてきた探索推論に利用されるモデルを検討する。本研究で想定する推論は、各脳部位における神経活動の脳波などによる伝播を媒体として、(1)大脳新皮質において現在の状態を認識し、(2)現在の状態と現状で可能な行動の組み合わせから予測される特徴群の洗いだしを行い、(3)その予測された特徴集合を内部的に再認識して、(4)各価値計算の脳部位における計算結果に基づいて予測・認識の枝刈りを行なう、という基本サイクルを考える。

そして、次の基本サイクルはその予測状態を起点としてより高い価値状態の探索を継続する山登り法的な過程が反復すると想定する。このメカニズムと価値の高い予測を作り出した脳内領

野を選択的に活性化するメカニズムが組み合わさるなら、高速かつ探索的な脳内の情報循環系路の制御が実現できよう(図1).

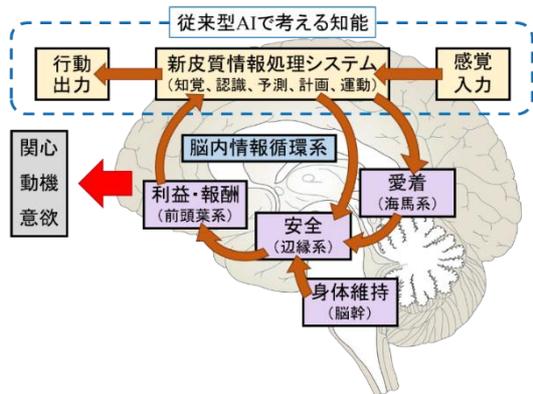


図1 感情=価値計算システム仮説の概要

3.3 価値計算システムモデル

感情を価値計算システムの表出であるとしたとき、その計算メカニズムはどのようなものであろうか。本稿では価値推論の方法と、全体システムを統合、制御する方法について議論する。本研究では神経細胞というデバイスだからこそ可能な瞬時かつ並列的な探索手法として、モンテカルロ法による確率探索を考える。このモデルでは、脳内の神経興奮の連想的伝搬が条件付き確率探索のモンテカルロ法的な実装になっており、そこに価値システムからの価値の興奮制御が加わって実質的な探索の枝刈り(確率分布の低下)をする、という推論方式を用いる。本手法の妥当性はトイモデルにて示されている。[宮田 2016]

4. 仮想環境におけるシミュレーション

4.1 前頭葉における複数の価値推論モデル

シミュレーションでは、環境の一部のみに強化学習にて期待報酬が学習済みである状態に、前頭葉で行なう確率的な推論を組み込むことで、より効率的な行動選択の実現を目指す。

行動決定場面では、すべての状況が経験・学習済みであることはあり得ない。そのため、各瞬間の状況を判断してそれに応じた行動決定戦略を選択することが重要である。そのため本シミュレーションでは、仮想環境中で複数の欲求が次々と発生し、それに対する価値計算過程のモデル化と評価を行う。シミュレーションはゲームエンジン Unity による仮想環境上に人工生命 LIS(Life in Silico)を置き[LIS 2016]、行動探索を行なった。

4.2 二層構造モデル

本研究のモデルは「地図-価値連合」および「最適価値選択」の二層構造である。「地図(場所)-価値連合」を表す層は、空間地図の脳内表現に相当し、脳では例えば海馬系や前頭前皮質が想定され、その一部の場所に対して過去の経験に基づいた価値(予測報酬)と実環境からの報酬を得る機能を持つ。脳では大脳基底核と脳幹や間脳がこれに相当しよう。シミュレーションでは、初期の価値分布はQ学習で獲得した。

「最適価値選択」を表す層は、脳内でのニューロンの発火による連想的な脳活動を表現する。外部からの価値評価ができない新奇入力に対して、エージェントは現在状態から価値が最大となる状態(ゴール地点)への確率並列探索を行う。それに対応する脳神経細胞の興奮伝搬を表現するため、モンテカルロ法による事前分布付き確率連想を用いた。事前分布は一層の地図の知識であり、そこで連想を反復することで、エージェントの頭

の中での地図探索を表す神経興奮が広がり、「地図-価値連合」層が保持しているその位置の価値情報が利用可能な状態への予測の到達を実現する(図2)。

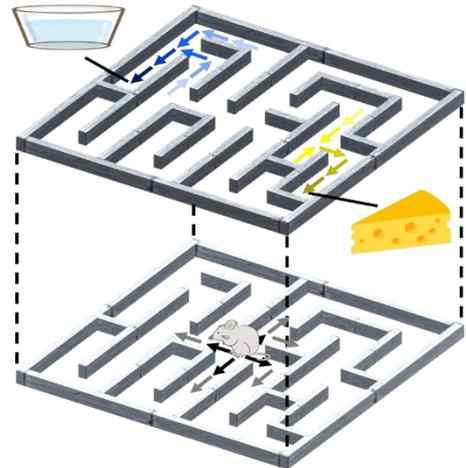


図2 二層構造モデルの概要

5. まとめ

本稿では感情の価値計算システムモデルと前頭葉推論モデルに基づく価値の導出モデルについて述べた。本研究で用いる仮想環境は、単純なタスクから複数の価値を同時に計算するような複雑なタスクまで実現できる。多様なシミュレーション環境でモデルの振る舞いを検証することで、感情と価値システムの計算論的意義の理解が進むことを期待する。なお、本研究は文部科学省科研費 15H01622 の助成を受けた。支援に感謝する。

参考文献

- [Ekman 1997] P. Ekman et al: What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS), Series in Affective Science, Oxford University Press, 1997.
- [Russell 1980] J. A. Russell: A circumplex model of affect, Journal of Personality and Social Psychology, vol.39, pp.1161-1178, 1980.
- [Chernavskaya 2015] Chernavskaya, : An architecture of the cognitive system with account for emotional component. Biologically Inspired Cognitive Architectures. , BICA Journal, 2015, doi:http://dx.doi.org/10.1016/j.bica.2015.04.009
- [Samsonovich, 2013] Samsonovich : Emotional biologically inspired cognitive architecture. Biologically Inspired Cognitive Architectures. , BICA Journal, 2013. doi:http://dx.doi.org/10.1016/j.bica.2013.07.009
- [Vallverdú 2016] Vallverdú : A cognitive architecture for the implementation of emotions in computing systems. Biologically Inspired Cognitive Architectures, BICA Journal 2016, doi:http://dx.doi.org/10.1016/j.bica.2015.11.002
- [戸田 1992] 戸田正直: 感情, 東京大学出版, 1992.
- [Koelsch 2015] S. Koelsch et al: The quartet theory of human emotions: An integrative and neurofunctional model, Physics of Life Reviews, Vol. 13, pp. 1-27, 2015.
- [大竹 2012] 大竹文雄ら: 脳の中の経済学, ディスカヴァー携書, 2012.
- [宮田 2016] 宮田, 大森: 感情の価値システムとしてのモデル化の試み, 第33回日本認知科学会 O3-1, 2016.
- [LIS 2016] LIS (Life In Silico), http://wba-initiative.org/1036/.