

主体感の自由エネルギーモデルとデザインへの応用の検討

行為の七段階理論を用いた考察

Free energy model of sense of agency and study of its application to design • Discussion using seven-stage theory of action

谷山建作¹⁾ 柳澤秀吉²⁾

Taniyama Kensaku¹⁾ Yanagisawa Hideyoshi¹⁾

1) 東京大学

Abstract: Sense of agency plays crucial role in motivating user exploration, especially in adaptive and evolving technologies. This study extends the comparator model by integrating uncertainty through the free energy principle, proposing how prediction errors and uncertainty shape agency. By relating this

framework to Norman's seven-stage action theory, we consider how agency influences user engagement with new tools and products. This perspective contributes to designing interactions that foster curiosity and sustained exploration in evolving digital environments.

Key Word: Sense of Agency, Free Energy Model, Motivation

緒言

主体感とは、自分が行動の主体である感覚を意味する (Gallagher, 2000)。例えば、機械の自動化の文脈において、機械が主体的に行動を引き起こすようになると、人間の主体感は減少すると考えられる。主体感は、行動する楽しさや行動の結果に対する責任感を引き起こす (田中ほか, 2016; Bigenwald & Chambon, 2019) ため、機械の自動化は楽しさや責任感を減少させると考えられる。機械の自動化と主体感の両立は課題である。機械の自動化は主体感の錯覚も引き起こすと考えられる。主体感の錯覚とは、実際は自分が行動の主体であるが機械が主体と感じる、または、実際は機械が行動の主体であるが自分が主体と感じる錯覚である。主体感の錯覚は事故を引き起こす (O' Toole, 2016; Conradi, 2016)。機械の自動化が引き起こす主体感の錯覚の解消も課題である。

以上のように、機械の自動化の文脈において主体感のデザインは重要である。主体感のデザインのためには、主体感のモデルが有効である。主体感のモデルとしては、コンパレータモデル (Frith et al., 2000) がある。コンパレータモデルは、運動制御の内部モデルを用いた主体感の定性的なモデルである。一方、機械の自動化と主体感の両立や機械の自動化が引き起こす主体感の錯覚の解消などの課題を解決するためには、定量的かつ錯覚の原因である不確実性を考慮したモデルが必要と考えられる。著者らは、コンパレータモデルをもとに、不確実性を考慮した主体感の数理モデルとして、自由エネルギーモデルを構築した (Taniyama & Yanagisawa, 2025)。自由エネルギーモデルは、運動制御の内部モデルをもとにしているため、身体の運動や機械の操作における主体感のデザインに応用できると考えられる。

一方、運動制御だけでなく、より一般的な行為においても主体感は重要と考えられる。例えば、新しい道具や製品の使い方の探究において主体感は動機付けになると考えられる。道具や製品の使用におけるユーザの行為を説明する理論としては、ノーマン (1990) が提唱した行為の七段階理論がある。主体感のモデルと行為の七段階理論を対応付けることができると、より一般的なユーザの行為における主体感のデザインに応用できる。

本稿では、まず、主体感のコンパレータモデルと著者らが構築した自由エネルギーモデルを紹介する。次に、これらの主体感のモデルと行為の七段階理論の対応付けを試みる。そして、デザインにおける主体感の意味を考察する。

コンパレータモデルと自由エネルギーモデル

コンパレータモデルと自由エネルギーモデルの対応を図1に示す。コンパレータモデルによると、人間はゴールを達成するために望まれる状態をイメージし、逆モデルを用いて運動指令を生成する。そして、運動によって実際の状態を変化させ、感覚フィードバックから実際の状態を推定する。この時、順モデルを用いて状態を予測する。この予測される状態と推定される実際の状態の差 (予測誤差) が主体感を減少させる (Frith et al., 2000)。

著者らは、自由エネルギー原理を用いて、自由エネルギーモデルを構築した (Taniyama and Yanagisawa, 2025)。自由エネルギー原理とは、知覚・学習・行動は自由エネルギーというコスト関数を最小化するように決まるとする脳の情報理論である (Friston, 2006)。自由エネルギーは、予測誤差の情報量表現である。著者らは、知覚によって最小化された自由エネルギーがコンパレータモデルにおける予測誤差に対応すると考えた。予測される状態を意味する事前分布と推定される実際の状態を意味する尤度関数を正規分布としてモデリングすると、知覚によって最小化された自由エネルギーは事前分布と尤度関数の期待値の差 (予測誤差)・事前分布の分散 (予測不確実性)・尤度関数の分散 (観測不確実性) によって記述される。

著者らは、自由エネルギーが主体感を減少させると考えた。そして、自由エネルギーモデルから、(1) 予測不確実性・観測不確実性が増加すると予測誤差が主体感に与える負の影響は増加する、(2) 予測誤差が増加すると予測不確実性・観測不確実性が主体感に与える影響は負から正に変化する、と仮説を立てた。これらの仮説をボタン押し実験によって検証した (Taniyama and Yanagisawa, 2025; 番場, 柳澤, 2021)。

コンパレータモデルと行為の七段階理論

行為の七段階理論を図2に示す。ユーザの行為は、ゴールの形成・行為のプラン・行為系列の詳細化・行為系列の実行・外界の状態の知覚・知覚したものの解釈・ゴールと結果の比較の七段階のサイクルとして単純化できる (ノーマン, 1990)。この七段階は、それぞれ、(1) 何を達成したいか? (2) 代替となる行為系列は何か? (3) 今どの行為ができるのか? (4) それをどうやってやるのか? (5) 何が起こったのか? (6) それは何を意味するのか? (7) それで良いか? 私はゴールを達成したのか? というユーザの質問を提供する。これらのユーザの質問に答える

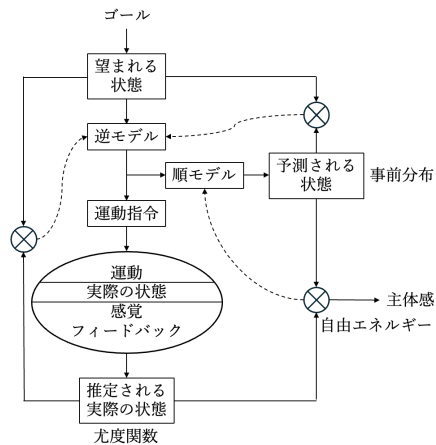


図1 コンパレータモデルと自由エネルギーモデル

ために、フィードフォワード情報とフィードバック情報をデザインする必要がある。

行為の七段階理論は、コンパレータモデルにおいて、望まれる状態と推定される実際の状態の比較までの過程を七段階に分割したものと解釈できる。行為の七段階理論とコンパレータモデルの違いは、コンパレータモデルが望まれる状態と予測される状態の比較と予測される状態と推定される実際の状態の比較も記述する点である。フィードフォワード情報は状態の予測を支援するため、望まれる状態と予測される状態の差（誤差1）は、フィードフォワード情報によって最小化できると考えられる。一方、フィードバック情報は実際の状態の推定を支援するため、状態の予測が正確である場合は、予測される状態と推定される実際の状態の差（誤差2）はフィードバック情報によって最小化できると考えられる。誤差1と誤差2が最小化されると、望まれる状態と推定される実際の状態の差が最小化され、ゴールが達成される。

デザインにおける主体感

行為の七段階理論の特徴は、サイクルとして記述される点である。近年、スマートフォンのようなソフトウェアを搭載した機能がアップデートする製品が一般的になり、ユーザは実行と評価のサイクルによって使い方を探究する必要がある。このような製品のデザインにおいて、主体感は探究の動機付けとしての意味を持つと考えられる。主体感は行動を促進する、または、内的な報酬として探索行動と利用行動の意思決定に影響すると考えられている（温ほか、2023）。

著者は、主体感を制御するために自由エネルギーモデルを構築した。そして、予測不確実性と観測不確実性が主体感に与える影響を明らかにした。予測不確実性と観測不確実性は、それぞれ、フィードフォワード情報とフィードバック情報の不確実性によって制御できると考えられる。これに加えて、主体感が動機付けに与える影響を明らかにできると、ユーザの製品の使い方の探究を促進できると考えられる。

まとめ

主体感は、楽しさや責任感をもたらすだけでなく、機械の操作の安全性にも影響を与える重要なユーザの認識である。この主体感は、主に運動制御の文脈において研究されているが、より一般的なユーザの行為の文脈においても重要な意味を持つ。特に、アップデートする製品に対するユーザの使い方の探究において、主

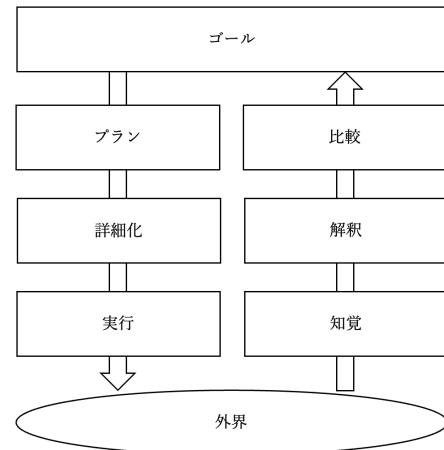


図2 行為の七段階理論

体感は動機付けとなると考えられる。主体感が動機付けに与える影響を明らかにできると、フィードフォワード情報とフィードバック情報を用いて主体感をデザインすることによって、ユーザのプロセス体験を向上できることが期待される。

参考文献

- Gallagher, S. (2000). Philosophical conceptions of the self: implications for cognitive science. *Trends in cognitive sciences*, 4(1), 14-21.
- 田中克則, 佐久間壮, 見澤英明, 宮下由佳. (2016). 感性工学に基づく Driving Pleasure に関する研究. 自動車技術会論文集, 47(5), 1147-1152.
- Bigenwald, A., & Chambon, V. (2019). Criminal responsibility and neuroscience: No revolution yet. *Frontiers in Psychology*, 10, 1406.
- O' Toole, R. (2016, July 6). What we know about fatal Tesla accident. *Cato Institute*.
- Conradi, K. (2016, September 6). Report on the serious incident to Saab 2000, G-LGNO Approximately 7 nm east of Sumburgh Airport, Shetland 15 December 2014. *Air Accidents Investigation Branch*.
- Frith, C. D., Blakemore, S. J., & Wolpert, D. M. (2000). Abnormalities in the awareness and action. *Philosophical Transactions of the Royal Society of London B*, 355(1404), 1771-1788.
- Taniyama, K., & Yanagisawa, H. (2025). Sense of Agency as a Function of Prediction Errors and Uncertainties: Free Energy Model and Experimental Evidence. *International Journal of Affective Engineering*, 24(1), 1-13.
- ノーマン, D. A. (1990). 誰のためのデザイン?: 認知科学者のデザイン原論. 新曜社.
- Friston, K., Kilner, J., & Harrison, L. (2006). A free energy principle for the brain. *Journal of Physiology-Paris*, 100(1-3), 70-87.
- 番場雅典, 柳澤秀吉. (2021). 情報利得を用いた操作主体感の数理モデリング (応答遅れ時間を用いた検証). 日本機械学会論文集, 87(893), 20-00035.
- 温文, 濱田裕幸, 鈴木陽介. (2023). 運動主体感は人間にとって何を意味するのか? 基礎心理学研究, 42(1), 53-64.