



Title	生物的目的に根差した予測処理の限界
Author(s)	菅野, 裕暉
Citation	研究論集, 25, 19 (左) -33 (左)
Issue Date	2026-01-16
DOI	https://doi.org/10.14943/rjgshhs.25.l19
Doc URL	https://hdl.handle.net/2115/98760
Type	departmental bulletin paper
File Information	04_rjgshhs_25_p019-033_l.pdf



生物的目的に根差した予測処理の限界

菅 野 裕 暉

要 旨

近年の認知科学分野のトレンドである、脳を予測器官として捉える「予測処理」は、知覚や認知プロセスなどを「予測誤差最小化」の観点から統一的に説明する野心的な理論である。この予測処理の能力はこれまで申し分なく発揮されてきたように思われる。現に、認知科学分野における少なくない研究者たちが、その説明能力に魅了されて、予測処理研究を開拓してきたのだ。だが本稿では、予測処理に基づく多産な研究成果にもかかわらず、その射程には限界があることを明らかにする。本稿の試みによって、予測処理が生物的目的「生物が恒常性を維持すること」を内包していることや、また予測処理の射程は生物的目的により制限されていることが明らかになる。

0. はじめに

近年の認知科学分野のトレンドである、脳を予測器官として捉える「予測処理」は、知覚や認知プロセスなどを「予測誤差最小化」の観点から統一的に説明する野心的な理論である (Clark 2024, Hohwy 2020, Piekariski 2021)。予測処理に依れば、脳はボトムアップ的に感覚入力から特徴検出を行う受動的器官ではなく、「内的・外的環境に関する生成モデルに基づいて、トップダウン的に感覚入力を予測し続ける能動的器官」である¹。そんな脳は予測を実際の感覚入力に適合させるべくモデル選択（知覚推論）を行うか、または実際の感覚入力を予測に適合させる

¹ 予測処理のトップダウン過程への言及自体は新規的でない。たとえば、プトレマイオスやアルハゼン、ヘルムホルツといった理論家が、判断や推論によって対象の大きさや形、距離などの知覚情報処理を説明してきた (Hatfield 2002)。だが予測処理の「新奇な貢献は、それがこの考え [トップダウン予測] を極端なまでに強調し、知覚の広範囲に見られる特徴としてトップダウン過程と事前知識の影響を描き出したことにある」 (Wiese & Metzinger 2023, p.3 括弧筆者)。

べく行為選択（能動的推論）を行うことにより、予測と実際の感覚入力 of 誤差（予測誤差）を最小化させることで知覚や運動を生じさせている。そして、これら知覚推論と能動的推論は「無意識的に」「自動的に」達成される。すなわち、脳は無意識的に自動的に予測誤差を最小化させているのである。

予測誤差最小化の観点から「知覚や認知プロセス」を統一的に説明するという、予測処理の能力はこれまで申し分なく発揮されてきたように思われる。現に、認知科学分野における少なくない研究者たちが、その説明能力に魅了されて、予測処理研究を開拓してきたのだ²。そんな予測処理研究の成功は、ホーヴィ（Hohwy 2020 Supplementary Material）が示すように、その適用事例の多さ——たとえば、注意（精度）や学習、情動（ホメオスタシスとアロスタシス）、両眼視野闘争、錯覚・幻覚、精神疾患、想像、自己欺瞞、内的表象、脳-身体カップリング、4E 認知などへの適用——が物語っている。そして今後の研究において、予測処理の射程は、たとえば有機体の進化と系統発生的改良 phylogenetic refinement や、予測処理に基づいた AI やロボットの実装研究へとますます拡張されていくだろう（Pezzulo et al. 2024）。このように一見して、予測処理は「生物に関わる諸現象」を説明し尽くすかのように思えてしまう。

だが、本稿では「（現状の）予測処理の射程には限界がある」、すなわち「（現状の）予測処理の射程は、脳の大統一原理である自由エネルギー原理の射程に捕らわれる」ことを明らかにしたい。自由エネルギー原理は「自己組織化システム（動物や脳）は好ましい状態を維持するために、刻々と変化する内的・外的環境のなかで自由エネルギーを最小化せねばならない」と主張する（Friston 2010, Friston & Stephan 2007, Parr et al. 2022）。こうした生物の適応・生存に焦点を当てた自由エネルギー原理は「生物が恒常性を維持するため（生物的目的）」の原理だと理解できる。（現状の）自由エネルギー原理から派生した予測処理は、「生物的目的に根差した予測処理」なのであって、その射程は「生物的目的に関する諸現象」に限定される、と本稿は主張する。しかし、自由エネルギー原理は生物的目的のみに拘束されず、そして予測処理の射程は拡張されえることを若干示唆したい。

以下では、まず、自由エネルギー原理から予測処理が派生することを概観したい（第1章）。そして、生物的目的に根差した予測処理（第2章1節）の射程が限界を有することを、いかに生物的目的に根差した予測処理が共同体を説明し損なうのかを示すことで明らかにしたい。そのうえで、自由エネルギー原理は生物的目的のみに拘束されず、そして予測処理の射程は拡張されえることを若干示唆したい（第2章2節）。

² 予測処理、自由エネルギー原理や予測符号化、予測誤差最小化などの研究者も含む。

1. 予測処理とは

本章では、まず1節にて、予測処理を説明するうえで欠かせない、自由エネルギー原理に対して概略的説明を与える。ここで、自由エネルギー原理は生物にとって何を意味するのか、そして自由エネルギー最小化はいかにして達成されるのかを確認する。2節では、予測処理に対して概略的説明を与える。ここで、いかに予測処理は自由エネルギー原理に基づくか、そして予測処理は生物にとって何を意味するのかを確認する。

1-1. 自由エネルギー原理

(はじめに)において、自由エネルギー原理は「生物が恒常性を維持するため(生物的目的)」の原理である、と述べた。本節では、自由エネルギー原理と生物的目的の密な関係について、自由エネルギー原理の導出過程から確認していく。そして、その導出過程と関連させつつ、自由エネルギー原理は生物にとって何を意味するのか、また自由エネルギー最小化はいかにして達成されるのかを確認していく。

自由エネルギー原理の関心は、「生物は刻々と変化する外的・内的環境のなかで、いかに適応し生存するのか」という問いにある(Friston 2010)。ここで「エントロピー」の概念を導入しよう³。エントロピーは、熱力学第二法則に従えば、熱や仕事が環境とやり取りされない閉鎖系にて時間経過とともに増大してゆく「乱雑さ」の指標のことである。すなわち、閉鎖系におけるエントロピーの増大は、流体の真空への自由膨張が不可逆的であるように、ある状態から別の状態への不可逆的な変化なのである⁴。だが、これは種特異的な、何らかの好ましい内的・外的状態を維持する生物にとって実に不都合極まりない。空気に満ちた大地に立つ人間が(魚とは異なり)水中でその正常な状態を維持できないように、生物によっては、エントロピー増大は避けるべき事態なのである。というのも、それにより「それらシステム[生物]はそれ自身を環境の種特異的な領域へと制限できず、究極的には相転移(たとえば死)を被るかもしれない」(Friston & Stephan 2007, p.429 括弧筆者)ためである。それゆえに、自由エネルギー原理の関心は、「生物はいかにエントロピー増大に抵抗して、その好ましい秩序を維持するのか」へと言い換えられる。

熱力学第二法則は、雪や石という、熱や仕事が環境とやり取りされない閉鎖系に該当するの

³ 自由エネルギー原理が「生物の適応・生存に焦点を当てた原理」、すなわち「生物が恒常性を維持するため(生物的目的)の原理」であることを強調するため、あえて熱力学第2法則から説明したい。

⁴ エントロピーは「不可逆性の定量的な尺度」(田崎 2000, p.104)である。この不可逆性は、ある状態Aから別の状態Bへの変化は可能だが、その逆であるBからAへの変化は起こりえない性質を意味する。不可逆性の具体例としては、牛乳にコーヒーを入れた場合、時間経過に伴ってコーヒー牛乳ができるが、そのコーヒー牛乳は初期状態へと変換できない事態がそうである。

であって、生物はその適用の範囲外である。熱や仕事を環境とやり取りする開放系である生物は、エントロピーを環境と交換することで、また自らの環境での位置やその関係を変化すべく積極的に環境へ働きかけることで、エントロピー増大を防ぐことができる⁵。ここで、非生物と生物の違いを理解すべく、フリストンの例を紹介しよう。

非生物的な自己組織化システムの範例として雪片をとり、その雪片が環境に作用できるように翼を授けてみた。通常、雪片は落下していき、その温度によって誘拐が引き起こされるところの境界面に直面するだろう。対して、その高度を維持し、また自らの温度を制御する雪片は、質的に認識可能な形態のままでいつまでも存続するだろう。(Friston & Stephan 2007, p.423)

すべて生物は自らの行為によりエントロピー増大を防いで、好ましい状態を維持している。しかし、生物が自らの秩序を維持すべくその増大を防いでいる、エントロピーとは何か？

エントロピーは「サプライザル」の時間平均である。サプライザルは負の対数証拠、すなわち「ある特定のエージェントとその行為を考慮したときの、ある感覚入力が起こりにくさ」として解釈される (cf. Friston & Stephan 2007)。サプライザルは、証拠が小さな値をとるとき増大するが、これは頻繁に起こる事象の生起を伝える情報よりも、ほとんど起こりえない事象の生起を伝える情報の方が情報量は大きい、というシャノンの情報理論に基づいている (cf. 乾 et al. 2021)。それゆえに、サプライザルは「ほとんど起こりえない事象 (感覚入力) が生じたときの驚き」と理解できる。たとえば、36.5℃前後の体温を維持する人間にとって、「灼熱の曝露による超高温の体温」を伝える感覚信号はほとんど起こりえない事象であり、高いサプライザルを有する。サプライザルの時間平均がエントロピーであるならば、平時サプライザルを低い状態に維持すれば、それに伴ってエントロピーも低下する。それゆえに、エントロピー増大に抗う生物はサプライザル減少に努めねばならない。

だが、すべて生物にとってサプライザルは「直接評価不可能な」情報理論量なのである。というのも、サプライザルは「感覚入力の生起に関わりうる場合すべて」を考慮するのだが、感覚入力は様々な原因で生じるがゆえに、脳の神経回路では計算困難となるためである。そこでサプライザルの代替指標として、生物が直接評価できる「自由エネルギー」がある。そして生物、すなわち「環境と〔定常〕状態にある自己組織化システムは、自由エネルギーを最小化せねばならない」⁶ (Friston 2009, p.1 括弧筆者)。これまでの展開を踏まえるなら、生物は好まし

⁵ 本稿の「自己組織化システム」は「生物的な自己組織化システム」と解釈されたし。

⁶ フリストン (Friston 2009) は定常状態ではなく「平衡」状態とある。だが、生物は環境とエネルギーをやり取りし、また自らの秩序を維持するような「非平衡」「定常」状態である (cf. 吉田 & 田口 2025)。

い状態（秩序）を維持すべく自由エネルギーを最小化せねばならない，となる。この主張こそが「自由エネルギー原理」なのである。

生物が最小化させる自由エネルギーは、「カルバック・ライブラーダイバージェンス（KL 情報量）」と「サプライズ」の和である（cf. Parr et al. 2022）。KL 情報量は「近似事後分布と真の事後分布の差」である。ここで KL 情報量を説明したい。すべて生物はサプライズを減少させるためにも，自らの内的・外的環境を適切に捉える必要がある。しかし，生物は感覚入力（たとえば光の波長や音波）を受け取るのみであり，感覚入力の原因である隠れ状態（たとえばリンゴやサイレン音）を推論しなければならない。そこで近似ベイズ推論が行われ，生物（の脳）は隠れ状態の近似事後分布を推論して，それを真の事後分布へと近づけていく（Knill et al. 2004）。KL 情報量は近似事後分布と真の事後分布の差であって，それらに差分があるとき正の値をとる（KL 情報量 ≥ 0 ）。自由エネルギーは，KL 情報量とサプライズの和であるため「サプライズの上限」なのである。それでは，いかにして自由エネルギーは最小化されるのだろうか。

自由エネルギーは無意識的推論によって最小化される。より正確に言えば，KL 情報量には「知覚推論」が，またサプライズには「能動的推論」が関与する。知覚推論は感覚入力をうまく説明できるようにモデル選択を行うことで，能動的推論は予測に合うように感覚入力を選択することで達成される（Friston 2009, Parr et al. 2022）。そこでは「知覚と行為が同じ（推論的）性質をもつという認識にとどまらない。そして，一般に想定されているような2つの異なる目的をもつというよりも」，自由エネルギーの最小化に向けて「むしろ知覚と行為の両方が協働すること」（Parr et al. 2022, p.24）が含意されている。そこで以下では，これら知覚推論と能動的推論を詳しく見ていきたい。

まず KL 情報量はモデル選択により最小化される（知覚推論）。すなわち，感覚入力の隠れ状態に関する生成モデル⁷を備えている脳は，感覚入力をうまく説明できるようなモデルを選択することにより，実際の感覚入力に合うような予測を行うことができ，結果として KL 情報量を最小化する。たとえば，仮説 H1「本物のトラが存在する」と，仮説 H2「作り物のトラが存在する」がある。もし脳が H1 を選択していた場合，その形状のほかに運動や音，匂いなどの感覚入力が予測されるだろう。対して，H2 は形状のみ予測をもたらす。ここで H1 が正しいとしよう。この場合，H1 を選択した脳は実際の感覚入力をうまく説明できるが，H2 を選択した脳は運動や音，匂いといった感覚入力を説明できない（そして KL 情報量最小化の失敗へと導く）。以上が，知覚推論による KL 情報量の最小化に関する説明である。それは，KL 情報量の最小化に伴って自由エネルギーを減少させるものの，サプライズの増減には一切関与していない。

⁷ 生成モデルとは「原因から観察された結果（データ）への確率的写像」（Friston 2009, p.293）である。脳は生成モデルに依拠して，隠れ状態に関する最も可能性の高い仮説から感覚入力を推論している。

そこで次に、能動的推論の説明へと進むことにしよう。

サプライズは行為選択により最小化される（能動的推論）。すなわち、脳は行為によって実際の感覚入力を変化させて、予測に実際の感覚信号を適合させることにより、サプライズを最小化させる。能動的推論は「世界に働きかけてデータの生成方法を変えて、自己がもつモデルのもとで最もサプライズの小さいデータを選択することによって、モデルを目的に適合させる」（Parr et al. 2022, p.25）。何らかの好ましい状態（秩序）内にある生物は、そんなサプライズ最小化した状態を自らの行為によって維持している。たとえば、人間は水中や真空状態（高サプライズ状態）を避けることによって、空気に満ちた大地に立つ状態（低サプライズ状態）を維持している。このように能動的推論はサプライズ自体を最小化する。そうであるならば、知覚推論ではなく、まさに能動的推論が「生物の恒常性を維持すること（生物的目的）」を本質的に担っているといえよう。

これまで自由エネルギー原理の導出過程を概観してきた。エントロピーやサプライズ、自由エネルギーとの関連で問題となったのは、すべて「生物は刻々と変化する外的・内的環境のなかで、いかに適応し生存するのか」という問いだった。つまり、生物が恒常性を維持するため（生物的目的）の方策が繰り返し問題となっていたのだ。そんな生物の適応・生存に焦点を当てた問題に対して、自由エネルギー原理は「知覚推論と能動的推論に基づき、自由エネルギーを最小化させることによって」と答えるだろう。次節では、自由エネルギー原理から派生する「予測処理」を概観したい。予測処理は、知覚推論と能動的推論に着目して、知覚や認知プロセスを「予測誤差最小化」の観点から統一的に説明するのである。

1-2. 自由エネルギー原理から予測処理へ

予測処理は、知覚推論と能動的推論に依拠した、認知活動の中核に「予測誤差」を据える理論である⁸（Clark 2013, 2015, 2023; Hohwy 2013, 2020）。すなわち、予測処理は、知覚や認知プロセスを「予測誤差最小化」の観点から統一的に説明する。予測処理に拠れば、脳は生成モデルに基づきトップダウン的に感覚入力を予測し、その予測と実際の感覚入力の誤差を伝達するボトムアップ信号に考慮しつつ、予測誤差の最小化を試みている。そして、予測誤差が最小化された段階で、認知主体には（程度の差あれど）現実が反映された、身体や外界に関する経験が生じる。予測誤差最小化は「知覚推論」と「能動的推論」により達成される。まず知覚推論は

⁸ 認知活動の中核に「予測誤差」を据える理論は、予測符号化や予測誤差最小化、バイズ脳仮説、能動的推論、予測する心などがある。すべて脳を「予測脳」と看做す点では同じであるが、たとえば予測符号化は能動的推論に関して何も語れず、また予測誤差最小化はホーヴィの過激な神経表象・構成主義的な解釈（Hohwy 2013, 2016）を連想しかねない。対して、予測処理は知覚推論に加えて能動的推論の要素も含み、また立場中立的な理論なのである。これら全体像を把握するためにも、近年の包括的研究として、Hohwy（2020）やPiekarski（2021）、Wiese & Metzinger（2023）を参照されたい。

モデル選択により、すなわち「実際の感覚入力をうまく説明し除けるような生成モデルを選択する」ことにより予測誤差を最小化させる。そして能動的推論は行為選択により、すなわち「予測に実際の感覚入力を適合させるべく、行為によって実際の感覚入力を変化させる」ことにより予測誤差を最小化させる。そして予測処理では、これら知覚推論と能動的推論による予測誤差最小化が、知覚や認知プロセスを統一的に説明するのである⁹。

ここで予測処理を適切に理解するため、自由エネルギー原理との同異について言及する。まず予測処理と自由エネルギー原理の共通点は、脳を「能動的な予測脳」と看做す点である。生物の脳は生成モデルに基づく予測と実際の感覚入力（たとえば光の波長や音波）をもとに、隠れ状態（たとえばリングやサイレン音）を推論している。それゆえ、脳はもはや感覚入力の特徴検出に秀でた受動的器官ではなく、むしろ感覚入力を予測し続ける能動的器官である。このことに関連して、知覚（知覚推論）と行為（能動的推論）が推論の性質を共有することも、予測処理と能動的推論の共通点として挙げられる。だがしかし、予測処理と能動的推論とで、「何を目的として知覚推論と能動的推論が協働しているのか」は相異なっている。すなわち、予測処理では「予測誤差（KL 情報量）」の最小化が、そして自由エネルギー原理では「自由エネルギー（KL 情報量とサブライザルの和）」の最小化が目的なのである¹⁰。この相違点は能動的推論に顕著に現れる。たとえば、街中でトラらしき物体に直面した場面を想定しよう。ここで予測処理の場合、もしあなたの脳が当該物体に対する確証を得られていないならば、あなたは当該物体に対する更なる確証を求めて、無謀にも当該物体へと突き進むであろう。対して自由エネルギー原理の場合、もし脳が当該物体に対してある程度確証があるならば、あなたは当該物体から早々と逃げ出すであろう。これらは予測処理と自由エネルギー原理の力点（目的）の相違を示している。

⁹ より適切に述べるならば、「知覚推論と能動的推論、精度予測に基づく、脳の各階層間において生じる予測誤差の全体的な最小化」が知覚や認知プロセスを統一的に説明するのである。ここで若干であるが、

「階層的予測処理」と「精度予測」に関して説明を与えたい。階層的予測処理とは、 N 階層は $N-1$ 階層へと予測を伝達し、 $N-1$ 階層は N 階層へと予測を伝達するように、予測処理は各階層間において生ずることを意味する (Lee et al. 2003; Mumford 1991; Rao et al. 1999)。また精度予測とは、予測や感覚入力の孕む偶然性（曖昧さ）に関する推論を意味する (Feldman & Friston 2010; Friston 2009)。偶然性（曖昧さ）の逆数である精度は、それが高いほど予測や感覚入力（予測誤差）の影響力を増大させるもので（逆も然り）、予測と予測誤差の均衡を生み出すのに必要である。たとえば、感覚入力（予測誤差）の精度が低いならば、予測誤差の影響力は小さく、多少の予測誤差の存在は無視されるため、予測の影響力が増すのである。

¹⁰ 自由エネルギー原理は「（予測処理の意味での）能動的推論」を持ち出すために言及されることもある。このときは、脳の予測は自由エネルギー（特にサブライザル）の最小化ではなく、「予測誤差の最小化」を目的としている点に注意したい。

ここで上記を踏まえながら、予測処理の議論に頻繁に登場する「幻覚」をケースとして、予測処理を実際の知覚や認知プロセスから理解していく。まず以下の場面を想定されたい。

ある朝の日、私は減入程の山積みの仕事に囲まれながら、いまだ寝室で眠りについていた。そして、物足りなさそうに眠りから覚め、なにか優雅な鳥の囀りをふと耳にした。いや少なくとも、当初、私にはそう聞こえた気がした。しかし、すぐにそれが誤りだと判明した。……私は鳥の囀りの幻覚を経験していたのだ。(Clark 2023, p.3)

この「鳥の囀り」幻覚では、まったく鳥の囀りに関する感覚入力を受け取っていないのに、なぜ認知主体は鳥の囀りを経験したのかが問題となる。予測処理に拠れば、この「鳥の囀り」幻覚は予測誤差最小化の観点から説明できる現象なのである。ただし、正常な知覚や認知プロセスの場合とは異なり、異常な予測誤差最小化の過程（仕方）の観点からではあるのだが。そこで、まず正常な知覚プロセス（1から3の繰り返しと4）を確認する。

1. 脳は生成モデル（仮説）に基づいて感覚入力を予測する
2. 予測と実際の感覚入力を比較する
3. もし予測誤差が生じた場合、予測誤差は予測（仮説）を修正・変更する
4. 予測誤差が最小化されたとき知覚経験が生じる

この正常な知覚プロセスでは、もし脳が「鳥の囀り」仮説に基づき感覚入力を予測する場合、予測（「鳥の囀り」感覚入力）と実際の感覚入力（非「鳥の囀り」感覚入力）とで誤差が生じる。この予測誤差は予測（仮説）を修正・変更する。その結果として、脳は非「鳥の囀り」仮説に基づき感覚入力を予測して、予測誤差を最小化する（認知主体は「鳥の囀り」を経験しない）。だが異常な知覚プロセスでは、脳は「鳥の囀り」仮説に対して非常に高い確証度を与えて、予測の影響力を実際の感覚入力の影響力よりも非常に大きいものとした。その結果として、脳は予測誤差の影響力を軽視して、予測誤差最小化の状態を成立させる（認知主体は「鳥の囀り」を経験する）。ただし、認知主体は窓から外を見るなどして、たとえば「鳥がいない」といった、無視できない感覚入力を得ることで、予測（仮説）を修正・変更することができる。すなわち、正常・異常によらず知覚や認知プロセスは、知覚推論と能動的推論による「予測誤差最小化」の観点から説明できるのである。そして、幻覚など異常な知覚や認知プロセスは、ただ予測誤差最小化を達成する過程（仕方）に（程度の差こそあれ）違いがあるに過ぎない。

「鳥の囀り」幻覚の例では、鳥の囀りを原因とする音波といった外受容感覚入力に関して、脳は生成モデルに基づき予測して、予測誤差を最小化することで「知覚」を生じさせていた。すなわち、外受容感覚入力の予測が知覚を生じさせるのである。そして同様に、関節や筋の

動きに関する固有受容感覚入力（運動）の予測は「運動」を、心拍や胃腸の動きに関する予測は「情動」を生じさせる。それゆえに、たとえば情動も、ボトムアップ的に内受容感覚入力を処理することではなく、トップダウン的に内受容感覚入力（と外受容感覚入力）を予測して、予測誤差が最小化されることで生じるもの、と述べよう（Seth 2013）。そして、幻覚などの知覚障害が異常な予測誤差最小化の過程（仕方）の観点から説明できるように、「運動障害」や「情動障害」も異常な予測誤差最小化の過程（仕方）の観点から説明できる。繰り返したが、すべて知覚や認知プロセスは予測誤差最小化の観点から理解できるのであって、異常な知覚や認知プロセスはただ予測誤差最小化を達成する過程（仕方）に違いがあるに過ぎない。

これまで自由エネルギー原理から派生した予測処理を概観してきた。前節で見たように、自由エネルギー原理は、知覚推論と能動的推論による「自由エネルギー」の最小化によって、「生物の秩序維持や生存」に対して説明を与える原理であった。また本節では、予測処理は、知覚推論と能動的推論による「予測誤差」の最小化によって、「生物の知覚や認知プロセス」に対して説明を与えるプロセス理論であることが明らかとなった。しかし、自由エネルギー原理と予測処理は、“知覚推論”と“能動的推論”を共有こそすれど、何を最小化しているのか（自由エネルギーか予測誤差）や、何を説明しているのか（「生物の秩序維持や生存」か「生物の知覚や認知プロセス」）に関して相異なっているように思われる。

2. 生物的目的に根差した予測処理

本章では、まず1節にて、それぞれの相異なる特徴にも関わらず、いかに自由エネルギー原理から予測処理が派生しており、関連づけられるかを考察する。ここで、それぞれの前提に生物的目的「生物が恒常性を維持するため」があることを明らかにする。2節では、生物的目的に根差した予測処理が共同体を説明し損なっていることを示すことによって、現状の予測処理の射程が限界を有することを立証する。そしてそのうえで、予測処理（更には自由エネルギー原理）の射程が拡張されえることを若干示唆したい。

2-1. 生物的目的に根ざす

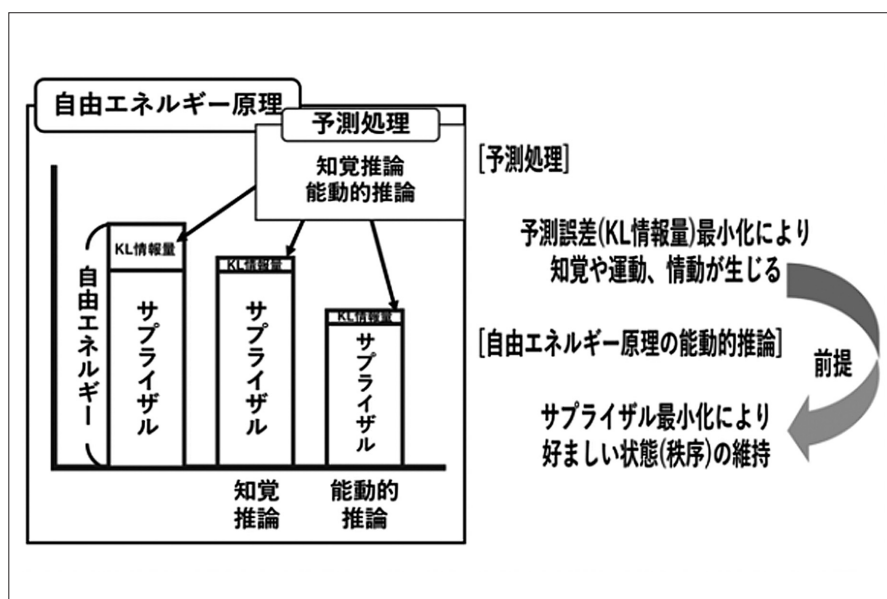
自由エネルギー原理では「生物の秩序維持や生存」を「自由エネルギー」最小化によって、また予測処理では「生物の知覚や認知プロセス」を「予測誤差」最小化によって説明する。このような相異なる特徴ゆえに、自由エネルギー原理が「脳の大統一原理」なのであって、ゆえに予測処理は自由エネルギー原理から派生している、という事実が不鮮明になっている。しかし実際には、明らかに、自由エネルギー原理は予測処理の根元に位置づけられる（cf. Pezzulo et al. 2024）。というのも、

1. すべて知覚や認知は自由エネルギー最小化につながる
2. 自由エネルギー原理は知覚や認知に関する諸理論を統合しうる

そうであるならば、「知覚や認知プロセス」に関する理論である予測処理が自由エネルギー最小化に関係していることを示すことで、予測処理が自由エネルギー原理から派生していることが明らかになる。そして同時に、予測処理は自由エネルギー原理の目的、すなわち生物的目的「生物が恒常性を維持するため」を前提とすることが明らかになるだろう。しかし、そもそも予測処理は自由エネルギー最小化に関係しているのか。

そこで、予測処理の「予測誤差最小化」は自由エネルギー原理の「自由エネルギー最小化」に関係しているのか、という問いに答えたい。この問いへの応答として、「自由エネルギーは KL 情報量とサプライザルの和であって、予測誤差は「KL 情報量」である。それゆえに、そもそも予測処理は KL 情報量（予測誤差）を最小化している。」があるだろう。この応答は間違っているではないが、しかし完全解でもない。すなわち、自由エネルギー原理と予測処理は、それぞれ別々の目的から KL 情報量を最小化している、といった可能性を排除しきれない。たとえば、哲学書を「日常的な研究のために」読む人と「日常的な教養のために」読む人は、共通する哲学書により関係づけるのではなくて、哲学者と教養人として区別するのがよい。そこで、予測処理は自由エネルギー原理の目的、すなわち生物的目的「生物が恒常性を維持するため」に基づいて KL 情報量（予測誤差）を最小化しているのか、という問いに移ろう。

この問いへの応答は YES である。すなわち、予測処理は生物的目的に基づいて予測誤差を最小化している、と答えたい。第 1 章 1 節では、(自由エネルギー原理における意味での) 能動的推論がサプライザルを最小化し、好ましい状態（秩序）を維持していることを確認した。この能動的推論が「生物の恒常性を維持すること（生物的目的）」を本質的に担っているのだ。だがしかし、知覚や運動、情動が「予測誤差最小化」により生じるからこそ、生物は「恒常性を維持するために」(自由エネルギー原理における意味での) 能動的推論を行えるのである。すなわち、能動的推論は「予測誤差最小化」を前提としてサプライザルを最小化させている。それゆえに、能動的推論の前提を担保する予測処理は、自由エネルギー原理と同じように、生物的目的を内包している、といえよう。



「予測処理」が生物的目的を内包するゆえに、「正常な予測誤差最小化の過程」と「異常な予測誤差最小化の過程」が生物的目的を内包する、とても言い換えよう。というのも、予測処理（予測誤差最小化の過程）自体が生物的目的を内包しているためである。この否定は、たとえばガン患者の「生存する」「好ましい状態を保つ」という目的を否定するのに等しい。そうではなく、ガン患者や遷延性意識障碍の患者でも、ただ「ヒト」であるがゆえに「生存する」「好ましい状態を保つ」という目的を有するのだ。また同様にして、「正常な予測誤差最小化の過程」や「異常な予測誤差最小化の過程」も、ただ「予測処理（予測誤差最小化の過程）」であるがゆえに生物的目的を有するのである。ここで注意すべきことは、生物的目的を内包するのは「異常な予測誤差最小化の過程」であって、「鳥の囀り」幻覚ではない。異常な予測誤差最小化の過程において「鳥の囀り」幻覚は生じる。だが、前者はプロセスであって、後者は現象である。ゆえに、「異常な予測誤差最小化の過程」と「鳥の囀り」幻覚を混同することはカテゴリー錯誤なのである。

これまで予測処理と自由エネルギー原理の関係に関して考察してきた。それにより、予測処理は能動的推論の前提を担保する「ための理論」であることが明らかになった。すなわち、予測処理は自由エネルギー原理のための理論である。これは、自由エネルギー原理が「脳の大統一原理」であること、また予測処理が自由エネルギー原理から派生することを意味する。そして、自由エネルギー原理が予測処理の根元にあるがゆえに、予測処理は「生物的目的に根差した予測処理」なのである。これが意味することは、「予測誤差の階層構造のいちばん上には、生き物の存在〔好ましい状態（秩序）の維持〕という、その生き物固有のゴール/目的が常にある」

(吉田&田口 2025, p.185 括弧筆者)ということである¹¹。まさに、生物的目的に根差した予測処理は「生物的目的に関する諸現象」を説明する理論なのだ。

2-2. その限界について ― 共同体の観点から ―

現在に至るまで、生物的目的に根差した予測処理（以下、生物-予測処理）は、知覚や運動、情動（ホメオスタシスとアロスタシス）、注意（精度）、学習、両眼視野闘争、錯覚・幻覚、精神疾患、想像、自己欺瞞、内的表象、脳-身体カップリング、4E 認知といった、数多の知覚や認知プロセスに適用されてきた（cf. Hohwy 2020 Supplementary Material）。そしてこれからも、知覚や認知プロセスに関して、実り豊かな研究が展開され続けていくであろう。しかし、そんな実績十分で前途有望である、生物-予測処理の広大な射程は、その目的「好ましい状態（秩序）の維持」や対象「秩序や認知プロセス」に縛られすぎて、暗黙理に「“生物の”知覚や認知プロセス」の範囲内に限定されている。それゆえに、生物-予測処理は、たとえば認知的ニッチや言語、共同体といった、「生物的目的を超えてる現象（以下、X）」をそれ自体の観点から理解することが非常に困難である。以下では、予測処理と X の関係に関して（かなり思弁的ではあるが）考察していきたい。

これまで生物-予測処理は X（生物的目的を超えてる現象）を論じてはいた。ただし、それらの専らの関心は、いかに生物が X と作用するなかで最適化された予測を行えるかであって、環境（X）の積極的な側面は言及されるだけであった。たとえば、認知的ニッチ（Constant et al. 2019）研究では、生物による環境の「予測しやすい状態」への整えに関心があって、環境自体が予測することにはそれ程関心がいかなかった。また言語創発（Taniguchi 2024; 谷口 2024）研究では、いかに外的表象の共有が行われるのかに関心があって、創発された言語が我々に与える影響はそれ程関心がいかなかった。これら研究は X を X 自身の観点からではなく、「生物的目的」の観点から理解しようとしているのである。

生物-予測処理は X（生物的目的を超えてる現象）を X 自身の観点から理解できるだろうか。生物-予測処理では、知覚や認知プロセスは生物的目的「好ましい状態（秩序）の維持」の観点

¹¹ 生物的目的は「ホメオスタシス」と「アロスタシス」により達成される（吉田 & 田口 2025）。ホメオスタシスとは「生物の安定性についての静的なメカニズム」である。これは（生命維持に重要な）生理学的な適正值との誤差が生じたとき、あらかじめ設定された適正值に調整するメカニズムである。たとえば、体温が 40℃ あるとき汗をかく場合や、血糖値が低いとき食事をする場合など。対して、アロスタシスとは「予測に基づく変化を通じた安定性を達成するメカニズム」である。これは適正值との誤差が生じるに先立って「生じるであろう誤差」を予測して、あらかじめ設定された適正值を維持するメカニズムである。たとえば、水中に潜るに先立って「血中酸素濃度の現象」を予測し、潜る前に大きく空気を吸い込む場合など。予測処理はこれらホメオスタシスとアロスタシスにより生物的目的を達成するのである。

から理解された。このことを踏まえるならば、Xを「(Xの) 好ましい状態(秩序)の維持」の観点から眺める必要が生じてくる。それゆえに、Xが「なぜ」予測したり、影響を与えるのか、について考えてみる必要がある。

ここでX(生物的目的を超えてる現象)を「共同体」として考えたい。以下で見るように、共同体こそが「好ましい状態(秩序)の維持」を無意識的に自動的に行っているように思われる。そんな共同体においては、「敬意を示す」「神様への信仰心をもつ」「空気を読む」といった、共同体自体と関連づければ十分に理解したとは言えなそうな現象が生じる。ここで、共同体の秩序がこれら諸現象により維持されているとしよう。このとき「共同体の秩序維持」からの逸脱——たとえば集団主義的な共同体において個人主義的行為(自分勝手な行為)を行くこと、また個人主義的な共同体において集団主義的な行為(甘える行為)を行うこと——を回避しなければならない。そして、たとえば「恥じる」(情動)は、それによって共同体規範への服従と、全体への調和へと本人を仕向けることで集団の崩壊を防いでいる、と理解できる。

共同体の目的に根差した予測処理は、マルコフブランケットに基づいて理解できる。マルコフブランケットのマルコフブランケットによれば、ある階層の自律した個体は、いくつか集まることで高次のまとまりを形成する。たとえば、分子やたんぱく質は集まることで細胞となり、細胞は集まることで人間になり、人間は集まることで社会や国家となる。ここで分子やたんぱく質は「熱力学第二法則」に従う、また生物は「自由エネルギー原理」に従う、文化も「それに相当する原理」に従うと考えてもよいだろう。どうして生物的自由エネルギー原理によって、共同体を説明し尽くせることができようか。これらを考慮するならば、共同体の目的に根差した予測処理はそれ程不自然な考えではなさそうだ。

3. おわりに

本稿では、導出過程から自由エネルギー原理と予測処理を概観することで、(1)自由エネルギー原理が「生物的目的」を内包することや、(2)自由エネルギー原理が予測処理の根元にあること、それゆえに(3)予測処理も「生物的目的」を内包していることを確認してきた。また(4)生物的目的に根差した予測処理の射程は「生物的目的に関わる諸現象」に限定され、生物的目的を越えてた現象を「生物的目的」から理解すること、だが(5)予測処理は共同体の目的に根差した予測処理へと拡張されうる可能性を秘めていることを明らかにした。

(かんの ひろき・哲学倫理学研究室)

参考文献

- ・ Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science.

- Behavioral and brain sciences, 36(3), 181–204.
- ・ Clark, A. (2015). Radical predictive processing. *The Southern Journal of Philosophy*, 53, 3–27.
 - ・ Clark, A. (2024). *The experience machine: How our minds predict and shape reality*. Random House.
 - ・ Constant, Axel; Clark, Andy; Kirchhoff, Michael & Friston, Karl J. (2022). Extended active inference: Constructing predictive cognition beyond skulls. *Mind and Language* 37(3): 373–394.
 - ・ Feldman, H., & Friston, K. (2010). Attention, uncertainty, and free-energy. *Frontiers in Human Neuroscience*, 4, 1–23.
 - ・ Friston, K. (2009). The free-energy principle: a rough guide to the brain? *Trends in cognitive sciences*, 13(7), 193–301.
 - ・ Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature reviews neuroscience*, 11(2), 127–138.
 - ・ Friston, K. (2013). Life as we know it. *Journal of the Royal Society Interface*, 10(86), 20130475.
 - ・ Friston, K. J., & Stephan, K. E. (2007). Free energy and the brain. *Synthese*, 159(3), 417–458.
 - ・ Hatfield, G. (2002). Perception as unconscious inference. D. Heyer, D. (ed.), *Perception and the Physical World: Psychological and Philosophical Issues in Perception*. John Wiley and Sons. 113–143.
 - ・ Hohwy, J. (2013). *The predictive mind*. OUP Oxford.
 - ・ Hohwy, J. (2016). The self-evidencing brain. *Noûs*, 50(2), 259–285.
 - ・ Hohwy, J. (2020). New directions in predictive processing. *Mind & Language*, 35(2), 209–223.
 - ・ Parr, T., Pezzulo, G., & Friston, K. J. (2022). *Active inference: the free energy principle in mind, brain, and behavior*. MIT Press.
 - ・ Knill, D. C., & Pouget, A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. *TRENDS in Neurosciences*, 27(12), 712–719.
 - ・ Lee, T. S., & Mumford, D. (2003). Hierarchical Bayesian inference in the visual cortex. *JOSA a*, 20(7), 1434–1448.
 - ・ Mumford, D. (1992). On the computational architecture of the neocortex: II The role of cortico cortical loops. *Biological cybernetics*, 66(3), 241–251.
 - ・ Pezzulo, G., Parr, T., & Friston, K. (2024). Active inference as a theory of sentient behavior. *Biological Psychology*, 186, 108741.
 - ・ Piekarski, M. (2021). Understanding predictive processing. A review. *AVANT. Pismo Awangardy Filozoficzno-Naukowej*, (1), 1–48.
 - ・ Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature neuroscience*, 2(1), 79–87.
 - ・ Seth, A. K. (2013). Interoceptive inference, emotion, and the embodied self. *Trends in cognitive sciences*, 17(11), 565–573.
 - ・ Taniguchi T (2024), Collective predictive coding hypothesis: symbol emergence as decentralized Bayesian inference. *Frontiers in Robotics and AI*, 11, 1–20.
 - ・ Wiese, W., & Metzinger, T. (2023). Key Concepts of Predictive Processing. Curado, M., & Gouveia, S. S. (Eds.). *Predictive Minds: Old Problems and New Challenges*. Vernon Press.
 - ・ 乾敏朗, & 阪口豊. (2021). *自由エネルギー原理入門：知覚・行動・コミュニケーションの計算理論*. 岩波書房.

- ・ 田崎晴明. (2000). 熱力学 現代的な視点から. 培風館.
- ・ 谷口忠大. (2024). 集合的予測符号化に基づく言語と認知のダイナミクス：記号創発ロボティクス
の新展開に向けて. 認知科学, 31(1), 186-204.
- ・ 吉田正俊. & 田口茂. (2025). 行為する意識 エナクティヴィズム入門. 青土社.