

論理的階層の圧縮としての内部観測によるトレードオフの乗り越え

Overcoming tradeoffs by compression of logical hierarchy from internal measurement

高橋 達二

Tatsuji Takahashi

東京電機大学 理工学部 情報システムデザイン学系

School of Science and Engineering, Tokyo Denki University

The importance of real world algorithms and robotics, we need to understand what is it to observe and act from within the world. Internal measurement, which admits that observation is itself action, gives a theory of the internal stance and it can be implemented as self-referential confusion by compression of logical hierarchies. We discuss that the confusion driven by internal measurement enables to overcome the speed-accuracy tradeoff. The confusion is consistent with the theory by Herbert Simon, Amos Tversky and Daniel Kahneman.

1. はじめに

内部観測は、松野孝一郎により導入され [松野 1991, 2000]、郡司ベギオ幸夫によって発展させられた [郡司 2004]。Michael Conrad の適応能、Robert Rosen の理論生物学、Humberto Maturana と Francisco Varela のオートポイエーシス、Otto Rössler の内在物理学などの議論と並行的な、あるいはそれらを引き継ぎ発展させるものと言える。しかしながら、近代科学の原理自体を問題にする難解さとその理論展開の晦渋さから、科学や工学における支持者を集めるには至っていない。そこで本オーガナイズドセッション「内部観測と推論」(OS-24) では、科学と工学に益する研究プログラムとして内部観測を部分的に定式化し、有用な方法論を提唱する準備を行う。

2. 自己言及と混同

内部観測は上に述べたいくつかの類似した理論と同様、有機体をいかに理解するか、すべきか、に関わる。分子生物学は分子機械として有機体を理解することを一つの目的とするが、そうでなく生命とは何かに関する何らかの理論を構築しながら有機体を理解するとすれば、有機体を機械と区別することが必要である。オートポイエーシスは、自律性、個性、境界の自己決定、入出力の不在を生命システムの条件として挙げた。このうち第一の自律性は、自己言及によってモデリングされる。

内部観測のモデルとしてまず、自己言及のパラドクスが単純ながら有効である。自己言及のパラドクスが論理的あるいは哲学的パズルにすぎないのではなく、機械と異なる生命システムの在りようを表しているというのが、内部観測やほかの理論に共通する観点である。機械は力学系で記述される。運動法則と境界条件が明確に確定している場合、あとは与えられた初期条件によって機械は作動を行う。それに対して、生命には自律性が認められる。生命は何を行うかを自ら決定する。この決定に際して、能動的な決定・言及する自己（一人称）と受動的な決定・言及される自己（三人称）の形式的な同一性が保持されながら、実質的な同一性が非自明化する [Gunji 04]。

私は嘘をついている

というそつ文の真偽を問うと、矛盾が発見される。「私は嘘をついている」という文が真であれば、嘘をついていないのだからこの文は偽であるということになるし、偽であれば、私は嘘をついているのだから、この文が真であるということになる。この矛盾を引き起こすのが、上の文の主語である「私」という言葉からの二重の指し示しで、「私」は「私は嘘をついている」の発話者を指しながら、この文中、話中の「私」でもある。後者の場合、主語である「私(の言うこと)」に対して「嘘である」という述語が与えられている。この文における「私」の二重の指し示しに伴う構造を明示化すると、

「私は嘘をついている」と私は言う

となる。すなわち、「私」の意味として、発話者（外側の一人称の私）と、発話の中の主体（内側三人称の私）の二義性があり、両者が同一でなければならないという要請から矛盾が生まれる [Takahashi 12]。そのため両者を厳格に区別するのが論理学の型理論であるが、しかし他方で、私は私である。両者の同一性は無視できず、論理的な型の違い（オブジェクトとメタ）として分けてしまえば、分けてしまったオブジェクトレベルの私とメタレベルの私をどう繋ぐのか、という難問が発生する。オートポイエーシスが、高次の論理の創発と称して弁証法的に無矛盾に両者を繋いでしまうのに対し、内部観測では論理的改装の圧縮、両者の区別と混同のダイナミクスをモデルとして用い、時間を導入する。この混同は、個別の私と場としての私、私の部分をなす発話と私全体、の混同である。この混同がたとえば情報工学にとって有効であることを本論では示し、混同という方法論の有効性を示唆したい。具体的には、内部観測の混同により意思決定を行うことで、不確実性の下で避けがたい既存のトレードオフを破ることができることの概要を示す。

3. 意思決定における自己言及的混同

意思決定においては、いくつかの行動選択肢から一つを選び、その行動を行う。選択においては、それぞれの選択肢の主観的価値を参照し、典型的には主観的価値が最も高いものが優先される。ここで、主観的期待効用理論のような規範に従えば、我々はそれぞれの選択肢を独立に、絶対評価によって価値付けしなければならない。しかし、行動経済学が明らかにしたように、人間のある選択肢に関する価値付けは、その選択肢や状況に関する事前の知識だけでなく、他の選択肢としてその場

連絡先: 350-0394 埼玉県比企郡鳩山町石坂. Tel: 049-296-5416

E-mail: tatsujit@mail.dendai.ac.jp

に何が存在するにも依存する、相対評価によるものである。相対評価の典型としては、いわば「反变的」な価値付け、価値の総量が保存していること的前提がある。二つの選択肢 A と B があるとき、A が悪いと分かると、A の価値を下げるだけでなく、B が良く見えてきて、B の価値を上げる。同様に、A が良いと思えば B が悪いと思えてくる。このような評価は上で述べた自己言及的な混同である。個別の選択肢の価値を、その個別の選択肢を含む選択肢集合全体と混同する形で関連づけているのである。ここで矛盾は主観的期待効用理論における独立性公理への違反といった形に弱められている。

4. 混同の効用

内部観測的な混同による意思決定が効果を発揮する一つの場面が、様々なトレードオフである。行動の一つの基準によって評価される単なる最適化問題であれば、問題によって最も効率のよいアルゴリズムを用意することが可能であろう。しかしながら、多くの現実的な場合においては、両立しがたく、かつそれらの間に明確な優先順位を与えられないような複数の基準の評価が必要となる。典型的なトレードオフは、短期的な結果と長期的な結果の間にある。今日一日だけ良ければ良いという、ごく短期優先の生き方と、数十年後を見据えた長期的な結果を優先する生き方では、行動は全く異なってくる。この短期と長期のトレードオフは、世界に不確実性がある場合には速さと正確さのトレードオフの形をとる。

4.1 情報利用と探索のジレンマと速さと正確さのトレードオフ

ここでこの事態の最も単純なモデルとして、 n 台のスロットマシンのいずれか一台の選択を繰り返して当たり報酬を最大化するゲーム、 n 本腕バンディット問題を考える。各マシンの価値 = 当たり確率はゲーム開始時には全く不明であるため、選択を繰り返す中で徐々に明らかにしていくしかない。それまでの知識に基づく情報活用 = これまでにベストだったマシンの選択と、知識の幅を増やすための情報探索 = ベストだったマシン以外の選択は、どちらも必要でありながら定義上同時には両立し得ない (ジレンマ)。「善は急げ」で現在の当たり獲得という意味での速さを優先し探索を止め情報活用を行うと、より良いマシンを見落とす。といて、「急がば回れ」で将来の正確さのために探索ばかりを行うと、探索の間より劣るマシンを選択するはめになりその分の期待損失が蓄積する (トレードオフ, 図 1)。

このトレードオフを導くジレンマは、情報活用と探索が定義上相互排他的であり、同時に実行できないというジレンマに由来する。そうであれば、情報活用と探索の厳格な区別を弱め、部分的であれ情報活用と探索の同時実行を可能とするようなモデルを採用すればよい。このジレンマは、マシンの価値をそれぞれ独立に評価し、選択マシンのフィードバックで一度に一つのマシンの価値のみをアップデートする、絶対評価を行う限り不可避である。そこで、個別の選択肢の評価と情報活用とそれ以外の選択肢の評価と探索を相対評価によって同時に行う。しかし、相対評価と言っても、様々なやり方がある。そこで相対評価の形式を絞るために、認知科学の結果を参照した。

4.2 Cognitively-inspired computing

我々はこれまで、認知科学の分野で人間の推論や意思決定に関する計算論的な研究を構成論的アプローチで行い、特にリスクや不確実性に直面した際の人間の認知の偏りやクセ—認知バイアス—の適応的な意味を明らかにしてきた [Takahashi 11, Takahashi 12, Uragami 11, 大村 12]。その過程で、人間の不

確実性下での、速くかつ正確な効率的行動が人工知能・経営学の基準充足化 (*satisficing*: Simon)、行動経済学におけるプロスペクト理論の期待値の信頼性 / リスクの考慮 (Tversky & Kahneman)、と組み合わせられたそして行動経済学の相対評価 (Kahneman & Tversky)、という 3 つの極めて人間的な行動価値評価の特性によることを解明した ([Takahashi 13], 図 1)。いずれの特性も近年脳科学でも扱われ、まさにバンディット問題を行う実験参加者の脳計測において論じられている。これを経験ベイズ法の形式を持つ行動価値関数、篠原修二の「LS モデル」[篠原 07] を一般化したものとして実装した [Takahashi 10, Kohno 12]。

4.3 トレードオフの乗り越え

n 本腕バンディット問題で言えば、図 1 の特性 A : 基準充足化はある期待値、たとえば 50% の当たりという基準を上回る選択肢を探すものである。効用の最大化は一般に非常に困難でありうるが、人間は自らの得る効用の最大化や最適化に拘泥し続けず、ある基準を満足する基準充足化を素早く行うことが知られている。特性 B : 信頼性とリスクの考慮は、相対的サンプルサイズつまり信頼性を勘案したマシン価値の評価を行うことで適切な判断を可能とする。行動経済学のプロスペクト理論で論じられるように、基準 = 参照点を決め、基準以下のマシンしか見当たらなければリスクを追求して不確実なマシンの探索に賭け、基準以上のマシンが見つかったらリスクを回避して探索を抑える。特性 C : 行動価値の相対評価は、限定的ながら情報活用と探索の同時実行を可能とする。マシンの価値の間にシーソーゲーム的な相互作用、すなわち相対評価を行い、これまで最良のマシンを選択 = 情報活用して外れた際にはそのマシンの価値を下げるだけでなく、それ以外のマシンの価値もまた上げる。これによりジレンマが弱められ、次の選択においては情報活用がすなわち探索となる。この 3 特性を持つ LS は図 1 のように速さ = 序盤性能と正確さ = それ以後の性能を両立し、図 2 にあるように、従来のトレードオフの限界を破った [Takahashi 11]。さらに、可能な行動数 (スロットマシン数) が増えれば増えるほど、従来のアルゴリズムよりも優れた結果をもたらすことが分かっている ([Kohno 12], 図 3)。

5. 結び

ビッグデータの普及とロボットの発達につれて自律的に情報を探索・活用するエージェントの必要性が非常に高まっているが、情報活用と探索のジレンマとそれが導く速さと正確さのトレードオフという問題が存在する。一般に仮想的探索が不可能である上に、探索空間が広大なほどエージェント数や探索時間などのリソース制限がより深刻になるため、GA などの従来有効な手法は利用できない。そのため、できるだけ単純で特定の問題構造に依存しない汎用性と、様々なシステムに容易に組み込める可搬性を持ちながらこのトレードオフによりよく対処できるモデルの開発が喫緊の課題である。

内部観測に基づくモデルは、速さと正確さのトレードオフを体現する最も基本的な問題である n 本腕バンディット問題において従来のトレードオフを乗り越え、この課題に答えうるものと言える。バンディット問題は不確実性下での繰り返し意思決定と学習、あるいは強化学習一般の最も基本的な問題でもある。そのため、我々はこれをロボットにおける運動の強化学習による獲得 [Uragami 11, Uragami 13] やモンテカルロ木探索による囲碁やリバーシの AI [Nishimura 12, Nishimura 13] に用い、すでに部分的な成功を収めつつある。ロボット運動学習では物理空間における多状態と非マルコフ性、モンテカルロ

木探索では超巨大な探索空間における多行動を有効に扱うことができ、内部観測の現実的な有効性の証拠であるといえる。

謝辞

本研究は、東京電機大学総合研究所研究 Q13K-03 と Q11K-02 として行ったものである。なお、実施と執筆において東北大学電気通信研究所共同プロジェクト (H22/B08)、日本学術振興会科学研究費補助金 25730150 の資金援助を受けた。

参考文献

- [Gunji 04] 郡司ペギオ 幸夫: 原生計算と存在論的観測 生命と時間、そして原生, 東京大学出版会, 2004.
- [Kohno 12] Kohno, Y., Takahashi, T.: Loosely symmetric reasoning to cope with the speed-accuracy trade-off. In *Proceedings of SCIS-ISIS 2012, Kobe, Japan, November 20-24, 2012*, 1166–1171, 2012.
- [Matsuno 91] 松野孝一郎: プロトバイオロジー 生物学の物理的基礎, 東京図書, 1991.
- [Nishimura 12] 西村友伸, 大用庫智, 高橋達二: 可変参照型緩対称性推論のモンテカルロ木探索での効果, ゲームプログラミングワークショップ 2012 論文集, 2012(6), 191-196 (2012-11-09)
- [Nishimura 13] 西村友伸, 大用庫智, 高橋達二: ヒトの適応的認知特性によるモンテカルロ木探索の効率化, 本論文集, 1L4-OS-24b-5in. (2013)
- [大村 12] 大村英史, 柴山拓郎, 高橋達二, 澁谷智志, 岡ノ谷一夫, 古川聖: 人間の因果推論による認知バイアスに基づいたメロディ生成システム, 知能と情報, 24, 5, 954–966. (2012)
- [Oyo in preparation] Oyo, K., Takahashi, T.: A cognitively-inspired heuristics for two-armed bandit problems: The loosely symmetric (LS) model, in preparation.
- [篠原 07] 篠原修二, 田口亮, 桂田浩一, 新田恒雄: 因果性に基づく信念形成モデルと N 本腕バンディット問題への適用, 人工知能学会論文誌, 22(1), 58–68, 2007.
- [Takahashi 10] Takahashi, T., Nakano, M., Shinohara, S.: Cognitive symmetries: Illogical but rational biases, *Symmetry: Culture and Science*, 21, 1–3, 275–294. (2010)
- [Takahashi 11] Takahashi, T., Oyo, K., Shinohara, S.: A Loosely Symmetric Model of Cognition, In: *Lecture Notes in Computer Science (Springer Proceedings of the 10th European Conference on Artificial Life (ECAL 2009))*, Springer, vol. 5778, pp. 234–241. (2011)
- [Takahashi 12] Takahashi, T., Gunji, Y.-P., Symmetrizing Object and Meta Levels Organizes Thinking, *BioSystems*, 107, 95–105. (2012)
- [Takahashi 13] Takahashi, T.: The adaptive combo of human cognitive biases—Satisficing, comparative valuation, and risk attitude—, 本論文集, 2J1–2.
- [Uragami 11] Uragami, D., Takahashi, T., Alsubeheen, H., Sekiguchi, A., and Matsuo, Y.: The Efficacy of Symmetric Cognitive Biases in Robotic Motion Learning, *Proceedings of the 2011 IEEE International Conference on Mechatronics and Automation August 7–10, Beijing, China*, pp. 410–415. (2011)
- [Uragami 13] 浦上大輔, 高橋達二, アルスビヒーンヒシャム, アルアルワンアリー, 関口暁宣, 松尾芳樹: 対称性推論と運動学習の分節化, 本論文集, 1L3-OS-24a-5.

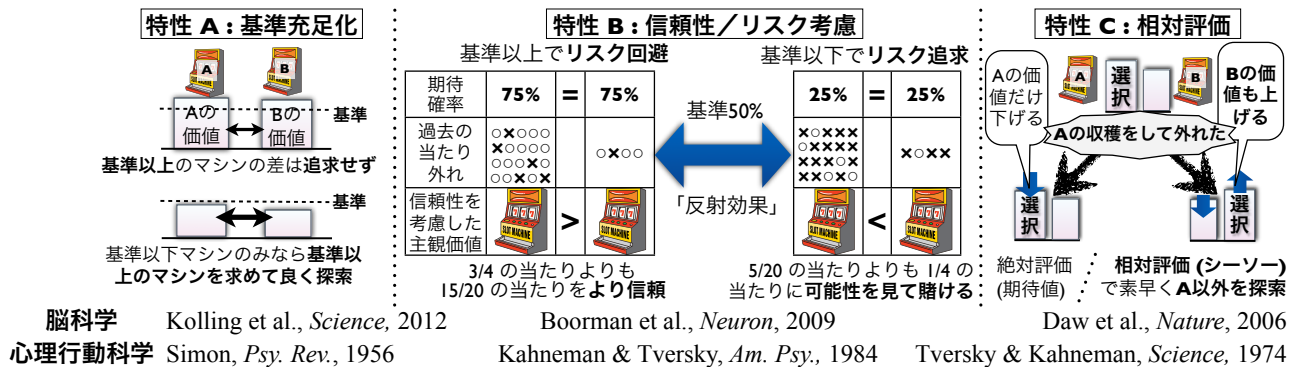


図 1: 人間的な状態行動価値評価の 3 つの特性

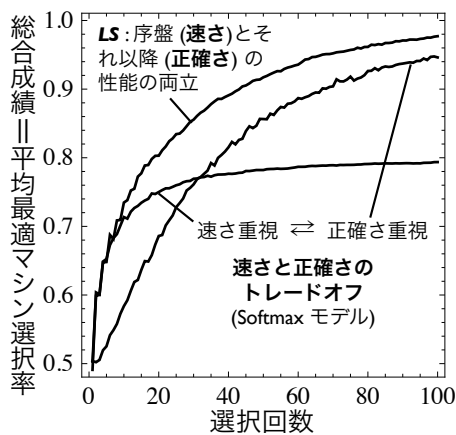


図 2: マシン数 2 での最適マシン選択率の時間発展におけるトレードオフ

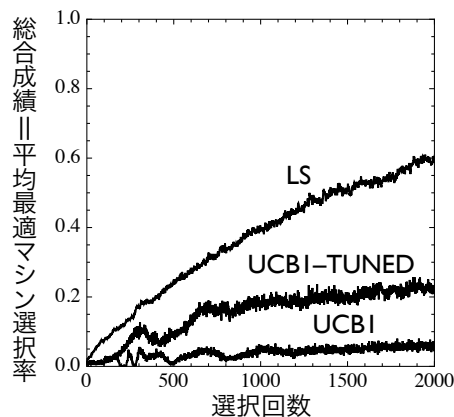


図 3: スロットマシンが 100 台の場合の最適マシン選択率の平均の時間発展