

二重過程統一脳理論による行動理解：教育・学習データへの適用

MUNEO KITAJIMA (北島 宗雄)^{1,a)}

Abstract: 人間は、環境（外部と内部）から情報を取り入れ、処理し、環境に働きかけるということを、生きている間、絶え間なく繰り返している。その結果として外部から観測されるのが、「行動」である。人間の行動を理解することは、情報が入力されてから出力されるまでの流れを理解することに他ならない。情報の流れの制御は主として脳によって行われる。情報の流れの結果は、ニューラルネットの結合の仕方に影響し、記憶として蓄積される。そして、入力された情報によって記憶の一部が活性化し行動の仕方に影響を及ぼす。時間の矢は一方向にしか進まないで、行動・記憶のサイクルは、個人個人に固有の時間発展をする。ただし、情報処理の仕方にはさまざまな強い制約がかかるので、全く自由に時間発展をするというわけではない。特に、無意識・自動的な情報処理と、意識・計画的な情報処理が、情報の流れを大きく制約する。無意識処理・意識処理を、過ぎ去った出来事に対して行う、あるいは、これから起こることに行う、というプロセスが、どのように実行されるのかが、行動・記憶サイクルの動き方に大きな影響を及ぼす。教育・学習という行動モードでは、入力される情報に操作が行われ、出力される情報は教育・学習の効果を計ることを目的として処理される。本講演では、人の一般行動の理由を説明することができる、我々が構築した意識・無意識過程を基盤に据えた並列分散処理方式の統一脳理論について概説し、教育・学習という行動モードで収集されるデータを分析する場合に留意すべき事項を示す。

1. はじめに

人間の行動を理解するためのアプローチには、分析的なアプローチと構成的なアプローチがある。分析的なアプローチでは、人間の行動をデータ化し、分析を行い、その中に規則性を見出すことが試みられる。一方、構成的なアプローチでは、仮説を構築し、収集したデータと仮説との整合性が検討される。前者がボトムアップによるマイニング型の手法であるのに対し、後者はトップダウンによるモデル検証型の手法である。我々^{a1)}は、近年の脳研究に係るさまざまな分野（情報科学・ロボティクス、人工知能・認知科学、神経科学、言語学、複雑系、心理・行動科学・文化人類学、生命科学・進化科学、哲学）での研究が著しい発展を見せていることから、それらの成果を統合することによって、脳内で生じている事象の時間発展を陽に組み入れた脳内情報処理に関する統一的な理論の構築を試みた。

図1に関連研究の系譜を示した。構築した理論（二重過程統一脳理論）をひとことで表現すれば、

Allen Newell が彼の著書“Unified Theories of Cognition [2]”に著した理論を時間制約のある日常行動選択の理論に拡張したもの

となる。脳の仕組みの解明が進み、多少の考え方の違いは生じているが、Newell の脳の認知科学的解析は、ほとんどその

まま継承している。Newell の理論を軸として、最新の研究成果を組み入れることによって拡張を行った。具体的には、心理学、文化人類学、進化生命科学などの人間の行動研究の関連分野を調査し、それらを最新の複雑系の科学がもたらした知見を参考にふるいにかけて整理し Newell の理論に組み入れた。理論拡張の重要なポイントは、生命界のような複雑系に表出する非線形性と階層性をどのように取り入れ、体系化するかであった。これらの概念は、Prigogine[3]により発展した非平衡熱力学と散逸構造理論においてもたらされたものである。

二重過程統一脳理論に基づいて構築した TK 脳モデルを図2に示した。図には、感覚器官を介して環境から情報が取り込まれ、記憶機能（自律神経系）と演算処理機能（体性神経系）の相互の動きによって身体行動が生み出され、環境が更新される行動サイクルが示されている。時間制約の強さによって異なった行動モードでサイクルは実行される。具体的に取る行動は、設定された行動目標（例えば、競争に勝つこと、他者に貢献すること、目標を達成すること、など）に応じて異なってくる。

2. 脳モデルの3世代

脳モデルの発展の流れは、三つの世代に分けて考えることができる。前述の TK 脳モデルは、第二世代のモデルである。

第一世代の脳モデルは、米国の認知科学会（Cognitive Science Society）や人工知能学会（Association for the Advancement of Artificial Intelligence, AAI）を中心とした研究により構築された。これらのモデルは、現象を近似的に再現する

¹ 長岡技術科学大学, Nagaoka University of Technology, Nagaoka, Niigata 940-2188, Japan

^{a)} mkitajima@kjs.nagaokaut.ac.jp

^{a1)} 豊田誠, 北島宗雄

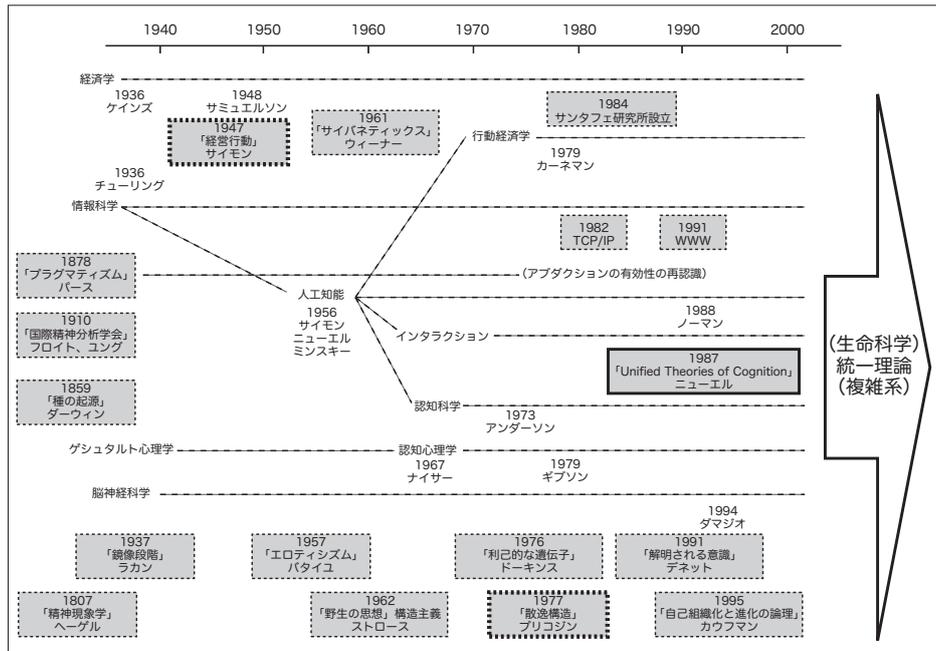


Fig. 1 関連研究の系譜 (I)の図 1.2.1) .

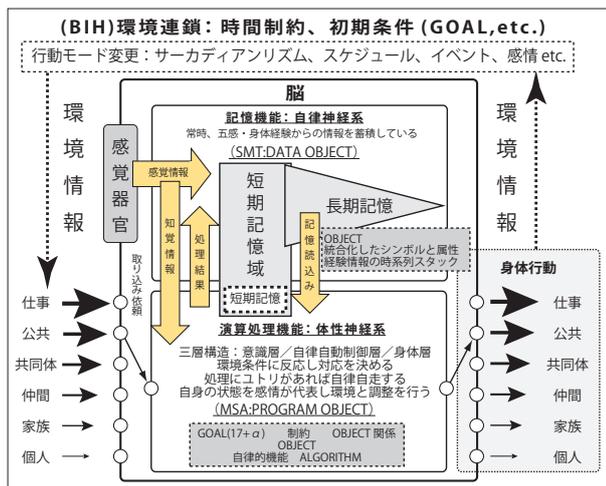


Fig. 2 TK 脳モデル (I)の図 1.2.7) .

ことを目指した 単一構造の 仮想モデル であり、人の特定な 行動の予測が可能 である .

第二世代の脳モデルは、脳の神経の仕組みに基づいた並列分散処理方式のモデルであり、人の一般行動の理由を説明することができる。第一世代のモデルとは異なり、実脳機能モデルを探求する中で構築されてきた。脳モデルは、認知アーキテクチャと呼ばれ、代表的なものとして、Soar[4]、ACT-R[5]がある。また、人間の認知行動プロセスが意識のプロセスと無意識のプロセスから成り立っていることを取り入れた研究の大きな成果として、意識・無意識の二重過程理論である Kahneman の Two Minds [6], [7]、二重過程脳モデルである我々の MHP/RT [8], [9] が挙げられる。これらの脳モデル構築の拠点のひとつは、軍隊使用のための新技術開発および研究を行うアメリカ国防総省の機関である“アメリカ国防高等研究計画局”(Defense Advanced Research Projects Agency, DARPA) のプロジェクトのひとつである Biologically Inspired Cognitive Architectures (BICA) に端を発して組織された BICA 学会である。多くの脳モデルが開発され、DARPA 周辺では実質的に

脳の解明は終わっている .

第一世代の脳モデルが対象とできるような現象は線形であり、その範囲では決定論的に予測可能である。第二世代の脳モデルは非線形な現象を対象としているが、この領域になると予測は原理的に不可能なので、モデルは将来の可能性を確率論的に予測(ただし、初期状態の事前確率は必要)する。これらのモデルは、動的構造モデルと呼ぶことができる。複雑系で観察される動的現象は、循環的複雑系において形成された構造に環境変動が作用して生じる現象である。脳の解明が第二世代に移行したことにより、人の理解は、これまでの現象的な静的解析から構造的な動的解析に移り、そのことで、単なる「仮説的な推測の状態」から「多様な可能性の推論」へと変化した。

第三世代の脳モデルは、脳の神経の仕組みを詳細に再現した完全シミュレーション可能なモデルであり、その構築にはこれから少なくとも 10 年はかかるものと考えられる。

3. 二重過程統一脳理論 (the unified dual-process brain theory)

3.1 概要

図 3 は、図 2 に示した TK 脳モデルがある動作モードで動いているときに行動が生じる様子を示している。環境情報が入力されると 5 つの自律システムが働き身体行動が生じ、その結果、環境の状態が更新される。同図に示した脳の中の情報の流れを、脳内の主要な自律システムが協調して作り出すプロセスを以下に示す：

- (1) 環境情報が知覚情報処理自律システムに入力される
- (2) 知覚情報処理自律システムでは、感覚情報フィルターを通して OBJECT 切り出しを行う。そして、その結果結果が、記憶処理自律システム、自律自動制御処理自律システム、意識処理自律システムに入力される。
 - 記憶処理自律システムでは、入力された情報がトリガーとなり、記憶ネットワーク内に発火連鎖が起きる

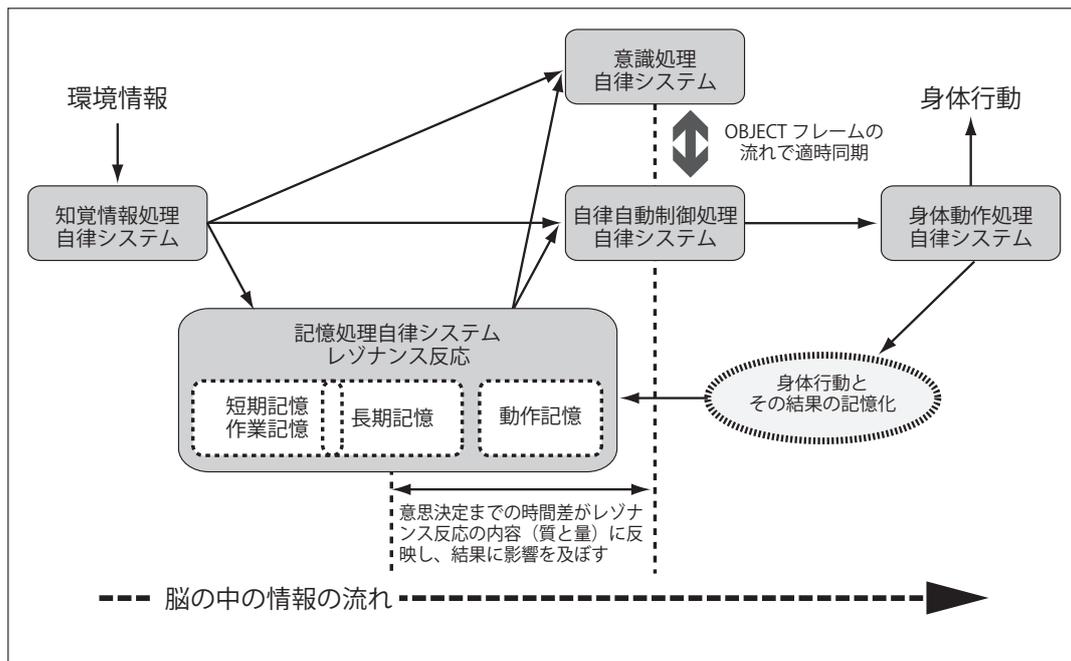


Fig. 3 脳内主要自律システム間協調機構の概要 ([1]の図 1.4.1) .

- 自律自動制御処理自律システムでは、入力された知覚情報と共鳴する記憶を使って無意識処理を進める
 - 意識処理自律システムでは、入力された知覚情報と共鳴する記憶を使って意識的な処理を進める（言語を使った推論など）
 - 自律自動制御処理自律システムと意識処理自律システムは、次々と生成される OBJECT フレームの流れを利用して、適時に同期される
- (3) 身体動作処理自律システムでは、自律自動制御処理自律システムからの出力を受け取り、身体行動を生じさせる
- (4) 記憶処理自律システムは、身体行動とその結果を記憶する

全体は、自律システムの集合体であり、そのことは、システム間に主従関係がなく、一方が他方を制御できる仕組みにはなっていない。記憶も自律システムであり、要請に応じて検索結果を返すために利用される静的なデータベースではない。記憶システムから意識処理、自律自動制御処理に、記憶に格納されている情報が渡されることをレゾナンス反応と呼んでいる。

3.2 行動と記憶のサイクル

図4は、図3が示している行動に含まれている記憶の内容を詳細かして示している。

行動生起に関わる4つの自律システム「知覚情報処理自律システム」「自律自動制御処理自律システム」「意識処理自律システム」「身体動作処理自律システム」のそれぞれには、行動時の活動を記憶するための多次元記憶フレーム (Multi-dimensional Memory Frame) が形成されている。同図に示したように、知覚情報処理に対応する知覚多次元記憶フレーム (PMD-Frame) は、意識処理に対応する言語多次元記憶フレーム (WMD-Frame) と関係多次元記憶フレーム (RMD-Frame)、自律自動制御処理に対応する行動多次元記憶フレーム (BMD-

Frame)、身体動作処理に対応する運動多次元記憶フレーム (MMD-Frame) と重なっており、知覚をトリガーとして記憶内部に活性が伝播するような仕組みになっている。

記憶を司るニューラルネットワークは、言語、関係、行動・運動のそれぞれが、知覚多次元記憶フレーム (PMD-Frame) と結合した3層構造になっている。この階層ネットワーク内の情報の流れは、トップダウンな“言語 → 関係 → 行動・運動”という流れと、ボトムアップに無意識な行動・運動を意識化することによってニューラルネットワークを構築する逆方向の流れがある。図では、前者は、機能的流れ構造 (Functional Flow Structure) として、後者は、図では、サイクリックネットワーク構造発達 (Evolving Cyclic Network Structure) として示している。

このようにして構築されている記憶は、実時間制約がかかる中で、知覚情報をトリガーとして活性化される (同図左下)。一旦、発火が始まると、関連した記憶の部分が連鎖的に発火する。すなわち、分散記憶システム (Distributed Memory System) で発火場所に近いところが連鎖発火する。発火して活性化した部分が、レゾナンス反応を介して意識処理や自律自動制御処理で利用される。

人間は、誕生後、知覚・運動のサイクルを絶え間なく繰り返し、感覚神経からなるニューラルネットワークと運動神経からなるニューラルネットワークの連合である介在神経のネットワークを、個々の自律処理システムに対応する多次元記憶フレームとして形成し続けている。

4. 二重過程統一脳理論から捉えた教育・学習データ

教育・学習活動に付随して生徒が生成するデータが、学生が図4に基づいて行動することによって生成されたと考えることにしよう。そうした時に、生成されたデータをどのように読むことができるのかを、説明していく：

- 教材や教示などの形式で外部から情報が感覚情報フィル

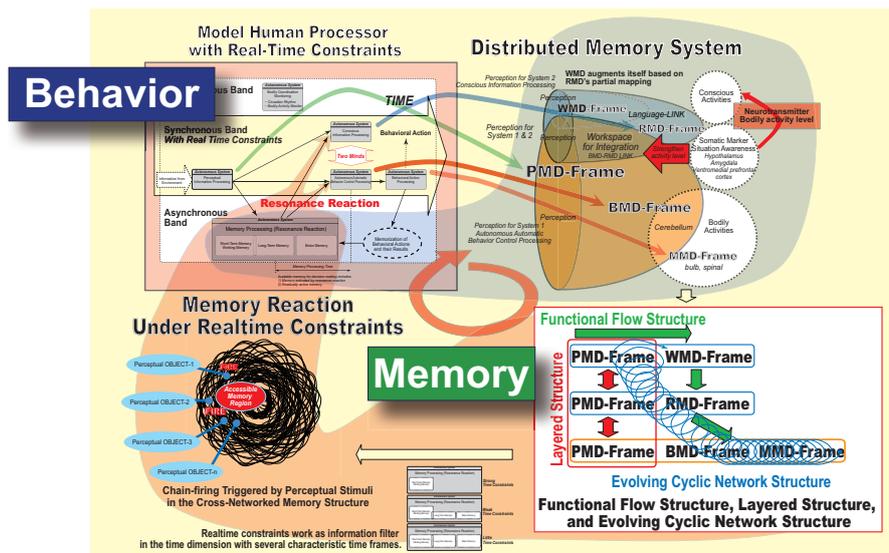


Fig. 4 行動と記憶のサイクル

ターを介して図4に示したシステムに入力されると、切り取られた OBJECT をトリガーとしてそれまでに形成されていた分散記憶システムの内部で発火連鎖が生じる。発火連鎖によって活性化されるのは、その OBJECT に関係して「記憶されている事象の系列」である。それには、言語、関係、行動、運動が含まれていて、活性化された記憶がレゾナンス反応により、意識・無意識の二重過程で利用され、行動が生起する。「記憶されている事象の系列」の長さが長ければ長いほど、遠い将来までの予測ができていくことになる。生徒が生成するデータが、行動データとしては同じであっても、それが属している行動系列の長さは異なっている可能性がある。長ければ長いほど、予見性の高い行動の一部とみなすことができる。

- 行動データとして直接に現れるわけではないが、行動の結果はネットワーク構造の更新に利用される。行動した結果の知覚表象に言語ラベルを付与することにより、その行動に関連したネットワークの活性状態は、意識と関連付けられる。これを形成することにより、図の右下の緑の線に沿った言語による行動の活性化が可能となる。

5. おわりに

BICA を中心に認知アーキテクチャの研究が進められてきているが、人の脳の研究と自律システム（自律ロボット）の研究は分離し、異なる流れとなっている。自律システムに利用できる技術はさまざまな人工物の中に組み入れられ、人間の活動を支援することが目指されている。人工物は、人間がデザインしたシステムや環境が含まれ、その中に教育・学習も含まれる。我々は、人工物が人間の脳の発達にどのような影響を及ぼすのかを二重過程統一脳理論に基づいて検討し、真に人間の発達・活動を支援するシステムが満たすべき要件をガイドランとして示している [10]。これにより、教育・学習活動に付随して生成されるデータから、教育・学習のためにデザインされた人工物が、その目的を達成しているかどうかを確認することができるようになってきている。

References

- [1] 北島宗雄, 内藤耕 (編著): 消費者行動の科学. 東京電機大学出版局, 東京 (2010).
- [2] Newell, A.: *Unified Theories of Cognition (The William James Lectures, 1987)*, Harvard University Press, Cambridge, MA (1990).
- [3] プリゴジン, I. and スタンジェール, I.: 混沌からの秩序, みすず書房 (1987).
- [4] Laird, J. E., Newell, A. and Rosenbloom, P. S.: Soar: An architecture for general intelligence, *Artificial Intelligence*, Vol. 33, pp. 1-64 (1987).
- [5] Anderson, J. R.: *How can the Human Mind Occur in the Physical Universe?*, Oxford University Press, New York, NY (2007).
- [6] Kahneman, D.: *Thinking, Fast and Slow*, Farrar, Straus and Giroux, New York, NY (2011).
- [7] Evans, J. S. B. T. and Frankish, K.(eds.): *In Two Minds: Dual Processes and Beyond*, Oxford University Press, Oxford (2009).
- [8] Kitajima, M. and Toyota, M.: Simulating navigation behaviour based on the architecture model Model Human Processor with Real-Time Constraints (MHP/RT), *Behaviour & Information Technology*, Vol. 31, No. 1, pp. 41-58 (online), DOI: 10.1080/0144929X.2011.602427 (2012).
- [9] Kitajima, M. and Toyota, M.: Decision-making and action selection in Two Minds: An analysis based on Model Human Processor with Realtime Constraints (MHP/RT), *Biologically Inspired Cognitive Architectures*, Vol. 5, pp. 82-93 (online), DOI: <http://dx.doi.org/10.1016/j.bica.2013.05.003> (2013).
- [10] Kitajima, M. and Toyota, M.: Guidelines for designing artifacts for the dual-process, *Procedia Computer Science, BICA 2015. 6th Annual International Conference on Biologically Inspired Cognitive Architectures* (2015).
- [11] Kitajima, M.: *Memory and Action Selection in Human-Machine Interaction*, Wiley-ISTE, 1 edition (2016).
- [12] 豊田誠, 北島宗雄: 生命体自律活動協調場理論: 幸福感に満ちた社会であるために (改訂版) *The Organic Self-Consistent Field Theory*, ON-BOOK.ltd, Tokyo (2010).
- [13] 豊田誠, 北島宗雄: 脳の自律システムの仕組みと性質: 行動の基準は効率から幸福・満足へ (時間制約下での動的人間行動モデル), ON-BOOK.ltd, Tokyo (2008).
- [14] 北島宗雄, 豊田誠: *CCE (Cognitive Chrono-Ethnography) の実践的概説 - 認知科学に基づく人の行動生態の調査手法*, ON-BOOK.ltd, 東京 (2011).
- [15] Kitajima, M. and Toyota, M.: The Role of Memory in MHP/RT: Organization, Function and Operation, *Proceedings of ICCM 2012: 11th International Conference on Cognitive Modeling*, pp. 291-296 (2012).
- [16] Kitajima, M. and Toyota, M.: Two Minds and Emotion, *COGNITIVE 2015: The Seventh International Conference on Advanced Cognitive Technologies and Applications*, pp. 8-16 (2015).

- 本文中で引用してはいないが、我々の研究を理解する上で重要な書籍や論文も掲載した