

認知アーキテクチャ評価手法の一考察

A Study of Evaluation Method for Cognitive Architectures

市瀬 龍太郎^{1*}

¹ 国立情報学研究所

¹ National Institute of Informatics

Abstract: There are many cognitive architectures available nowadays. However, it is difficult to compare those architectures, because there is no standard measure to evaluate them. In this paper, we propose a method to evaluate cognitive architectures based on CHC model.

1 はじめに

汎用人工知能の作成を行うには、人間の知的な活動をモデル化する必要がある。人間の知的な活動を外界（環境）からの刺激に対して、反応を決定する機構として見なすと、人間の知的な活動をモデル化することが可能となる。このようなモデルは、エージェントモデルと呼ばれ、人工知能を作成する際のモデルとして利用されている [Russel 09]。このモデルにおいては、人間がある環境下において、目、耳などを通して外界を知覚し、口、手などを通して外界に反応するのと同様に、エージェントは、センサーで外界を知覚し、アクチュエータで反応するものとしてみなす。この時、センサーから入力されたデータが、なんらかの変換を施され、アクチュエータへの出力がなされるとみることができる。この変換の部分がエージェントの知的な活動となる。この変換においては、いかなる時系列の知覚に対しても、適切な時系列の出力を返さなければならないため、エージェントの内部に、過去のデータを蓄える部分など、何らかの構成要素と構造を持つ必要がある。このような知的動作を実現するための内部構造を認知アーキテクチャ [Newell 90] と呼ぶ。認知アーキテクチャは、汎用人工知能を作る際に最も見込みのある研究と言われている [Thórisson 12]。しかし、認知アーキテクチャは、知的動作を実現するための汎用的なフレームワークであるため、評価が難しいという問題点がある。そこで本稿では、人間の知的能力を整理分類した CHC モデルとの対比で認知アーキテクチャを評価することを提案する。

本稿の構成は以下のようになっている。まず、第 2 章では、認知アーキテクチャについて説明を行う。次の第 3 章では、従来の認知アーキテクチャの評価手法について述べる。第 4 章では、知的な能力を整理分類

した心理学のモデルである CHC モデルについて述べる。第 5 章では、CHC モデルを用いた評価手法について考察を行い、最後の第 6 章で本論文をまとめる。

2 認知アーキテクチャ

人工知能を実現するための認知アーキテクチャとして、これまでに様々なアーキテクチャが提案されてきた。例えば、ICARUS [Langley 06] においては、目標、必要条件、手段の基本要素を階層的に記憶することで、状況と行動の間を知的に決定することができる。また、CogPrime [Goertzel 13] においては、外界の情報を知識に変換する認識階層、知識を保持する AtomSpace、知識を基に推論や学習を実行する認知プロセスに分け、それらが協働することにより知的な行動を決定することができる。このように認知アーキテクチャにおいては、様々な要素が統合されることにより、知的な活動を実現している。

認知アーキテクチャには、様々な要素が必要となるが、いくつか代表的な要素が知られている。まず、知識や経験の保存方法が必要となる。これは、記憶機構として実現されるが、記憶にも様々な種類があるため、作業記憶、手続き記憶、意味記憶、エピソード記憶などの種類に分類されることが多い。次に学習機構が必要となる。機械学習の分野では、学習はいくつもの手法に分類されている。代表的な手法として、教師あり学習、教師なし学習、強化学習などが挙げられる。次に、課題の解決機構が必要となる。ある問題が与えられた時に、その解決方法を与えるもので、代表的なものとしては、プランニング、推論機構、意思決定などが挙げられる。さらに、外界を認識するための機構も必要となる。そのためには、視覚処理、聴覚処理、言語処理などの他に、メタ認知なども必要になると考えられる。ここでは、いくつかの代表的な要素を挙げたが、その他にも様々な要素が必要となる。このような

*連絡先：国立情報学研究所情報学プリンシプル研究系
〒101-8430 東京都千代田区一ツ橋 2-1-2
E-mail: ichise@nii.ac.jp

要素をどのように統合し、実装するかにより、多様な認知アーキテクチャが生まれることとなる。

3 認知アーキテクチャの評価

認知アーキテクチャに対して、評価を行う試みは、これまでにいくつも行われてきている。一つの手法としては、認知アーキテクチャの動作を人間と対応させることで、人間と同等の動作がなされているかを基準に評価する手法がある。例えば、ACT-R [Anderson 04] においては、記憶、問題解決などにおいて、エラー率、反応時間が人間と等価であるかが検証され、人間の思考と同様の動作であることの傍証としている。また、別の手法としては、多様なアプリケーションを同じ認知アーキテクチャを使って作成することで、評価を行う手法がある。例えば、学校で使われる学習システム、ロボットの制御、自動車の運転などさまざまなタスクに対して、同じ認知アーキテクチャが使えることを示すことで、その認知アーキテクチャが汎用的なフレームワークであることを示すことができる。

上記のようなアプローチを使って、汎用人工知能に向けた認知アーキテクチャを評価すると、大きな問題点がある。まず、反応速度が人間と同等であることを基準に考えるアプローチは、内部の処理プロセスが人間と等価であることを保証する上では、有用である可能性がある。しかし、汎用人工知能においては、人間と同じ知的な能力を持つことが目的であり、内部機構まで等価であることは必要ではないため、評価の基準としては不十分である。また、多様なタスクで評価を行う場合には、評価するタスクに特化した特化型人工知能の集合体と区別が付かない。特化型人工知能の作成と汎用人工知能の作成は本質的に異なるとする汎用人工知能の中心仮説 [ゲーツェル 14] を考慮すると、このアプローチも汎用人工知能に向けた認知アーキテクチャの評価には向かないこととなる。

人工知能を評価する際には、タスクではなく、能力で比較の方が望ましいとの提案 [Hernandez-Orallo 14] がなされている。複数の認知アーキテクチャを能力などに基いて評価したものとして、Samsonovich がまとめた表 [Samsonovich 15] がある。その表では、ACT-R など 26 個の認知アーキテクチャに対して、いくつかの基準で比較、評価を行っている (2015 年 12 月現在)。比較に使われている指標として、認知アーキテクチャで利用されることが多い作業記憶などのコンポーネント、学習アルゴリズム、問題解決などのパラダイムなどが使われている。また、Thórisson らの研究 [Thórisson 12] では、自律性という観点から、学習、メタ学習、即応性、資源管理に分け、ACT-R など 9 個の認知アーキテクチャについて比較、評価を行っている。これらの研

究においては、能力に着目しているが、能力の分類に関して統一的な指標が使われておらず、厳密な比較が難しいという問題点がある。

4 心理学における知的能力

人間の知的な能力に関しては、心理学の分野において長年研究されてきている [三好 10]。Spearman は、知能に 2 つの要素があるとして、共通して関与する一般因子 g と、特定部分に関与する特殊因子に分けて論じた。その後、Cattell は、一般因子を詳細化し、流動性知識、結晶性知識に分けたが、Horn がさらに多くの因子を取り出し、視覚的知能など 10 個の能力に分けている。さらに、Carroll は、知能が 3 つの階層から成り立つとし、第 1 階層に一般因子、第 2 階層に 8 つの因子、第 3 階層に 70 個の因子があるとした。そして、Cattell, Horn, Carroll の理論を統合し、現在では、CHC モデルとして、知的能力を 3 階層の分類で表現できている。

4.1 CHC モデル

本節では、Schneider らの研究 [Schneider 13] に基づいて CHC モデルを概観する。CHC モデルでは、知能を「一般 (General)」、「広範 (Broad)」、「限定 (Narrow)」の 3 つの階層に分類している。第 1 階層は「一般知能 (g)」がただ一つあるのみで、知能を集約したものになっている。第 2 階層は、以下のように、16 個に分類されている。

流動性推論 (Gf)：推論、量的推論に関する能力

短期記憶 (Gsm)：作業記憶の能力

長期貯蔵と検索 (Glr)：記憶の想起、お互いに関連付けて記憶する能力、およびそれを検索する能力

処理速度 (Gs)：知覚速度や、読み書きの速度など処理の速さに関する能力

反応、意思決定速度 (Gt)：簡単な意思決定や反応を素早くこなす能力

精神運動速度 (Gps)：身体を素早く動かす能力

結晶性知識 (Gc)：言語発達、コミュニケーションなどの文化的知識

領域固有知識 (Gkn)：特定の領域に関する知識

読み書き (Grw)：文書の読み書きに関する知識

量的知識 (Gq)：数学に関する知識

視覚処理 (Gv)：視覚的な推論や記憶などに関する能力

聴覚処理 (Ga)：聴覚的な判別や記憶などに関する能力

嗅覚能力 (Go)：嗅覚的な検知や記憶に関する能力

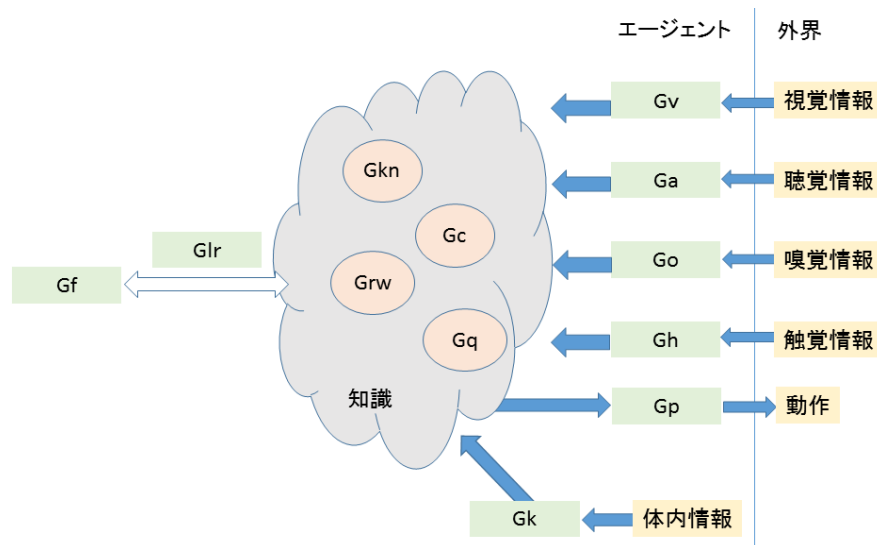


図 1: CHC モデルの認知アーキテクチャ的解釈

触覚能力 (Gh) : 触覚的な検知や処理に関する能力

運動感覚能力 (Gk) : 身体の状態の認識能力

精神運動能力 (Gp) : 身体を動かす能力

この中で、Gf, Gsm, Glr, Gs, Gt, Gps は、領域非依存の一般的な能力と位置づけている。そして、Gc, Gkn, Grw, Gq は、獲得した知識に関する能力とし、残りの Gv, Ga, Go, Gh, Gk, Gp は、感覚運動に関連した能力としている。

CHC モデルでは、上記の階層の下に第 3 階層として、より詳細な能力が割り当てられているが、本稿では省略する。

5 考察

本稿では、認知アーキテクチャの評価に、人間の知的能力の評価にも利用される CHC モデルを利用することについて考察する。そのために、CHC モデルをエージェントアプローチ的な認知アーキテクチャとの対応で捉え直し、整理すると、図 1 のように考えることが可能となる。

Gv, Ga, Go, Gh の 4 つは、外界から得られる情報を、エージェント内部で使われる表象に変換する部分に関する能力となる。これらの能力は、視覚、聴覚、嗅覚、触覚を通して得られた情報から、適切な部分を切り出し、内部で利用できる形態に正確に変換する能力と考えることができる。例えば、猫の画像が与えられた際に、猫がいるという表象に変換するのが Gv の役割になる。ただし、犬ではなく、正確に猫に変換できる能力だけが評価されるのではなく、場合によっては、色や形なども正確に表象に変換することが重要な能力となる。なぜならば、後ほど、これらの情報が思考な

どに使われる可能性があるからである。Gk は、Gv などと共に、感覚運動に関連した能力と分類されているが、エージェント自身の状況进行处理する能力と言える。例えば、Gk で扱われる、自分の筋肉の状態といった情報は、人間の内部から生まれる情報であり、外界から得られる視覚情報などとは区別される情報となる。そのため、内部のセンサーから生ずる情報を Gv などと同様に処理する能力が Gk であるととらえられる。

CHC モデルでは、知識に関する能力は Gc, Gkn, Grw, Gq で構成されている。これらは、獲得した知識の量と質を表していると考えられる。獲得した知識を蓄えている部分は、表象や概念、スキルなどを保持していると見なせる。但し、CHC 理論においては、これらを区別しておらず、Gc, Gkn, Grw, Gq に集約している。なお、ここでは、図示するために、Gc 等は、それぞれが独立なものとして記述したが、実際には分散表象のようなもので表現されることもあり得るであろう。

Gf は、帰納、推論、学習に関連する能力である。これらは、エージェント内の知識を書き換えることによりなされるため、表象、概念、スキルの書き換えを行う能力として捉えることができる。その書き換えをいかに正確かつ的確にできるかが重要な能力となる。メモリに関する能力 Gsm, Glr は、知識の記憶容量、および、記憶の検索性能と考えることが可能である。通常、計算機においては、外部記憶装置も含めれば、記憶容量の制限は、ほぼ無いと言える。そのため、Gsm に関しては、記憶容量の問題となるため、計算機上に実装される認知アーキテクチャにおいては、能力の評価に意味がないものであろう。一方、Glr に関しては、関連する情報を正確に探し出してくる検索性能が関わる。そのため、Gf が知識を取り出す際に、大きな影響

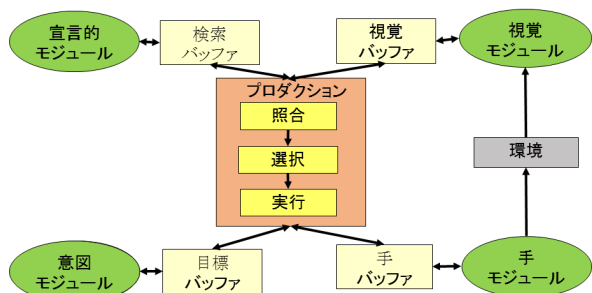


図 2: ACT-R の認知アーキテクチャ

があり、認知アーキテクチャの評価上、重要な能力になると考えられる。Gs, Gt, Gps は、処理の速度に関連する能力となる。認知アーキテクチャにおいては、処理速度は、基本的に使用する計算機的能力に大きく依存し、認知アーキテクチャ自体に依存する部分は少ないと考えられる。そのため、これらの能力に関しては、認知アーキテクチャで議論する必要性は低いと考えられる。

図 1 で整理した項目を用いることで、認知アーキテクチャの能力を図 1 に関連付けることが可能となる。そのため、それぞれの能力をどのように実現しているかが明確になり、お互いの認知アーキテクチャを評価することが可能になると考えられる。例えば、ACT-R [Anderson 04] のアーキテクチャは、図 2 のように書くことができる。この時、視覚モジュールが図 1 における Gv の行っている処理に相当し、環境に応じて、Gkn や Gq などを視覚バッファ上に生成していると解釈できる。ACT-R の評価手法の一つとして、Gv の能力の観点からは、視覚モジュールがいかに正確に Gkn や Gq などを生成できるかという点で、評価することが可能となる。また、異なる認知アーキテクチャに対しても、どのモジュールがどの能力を実現しているかで比較、評価が可能となるであろう。

6 おわりに

汎用人工知能を作成する際に、認知アーキテクチャを考えていくことは欠かせない。しかし、様々な認知アーキテクチャに対して、どのような能力が必要であるかの体系的な議論がこれまでになされてこなかった。本稿では、心理学で使われる CHC モデルを認知アーキテクチャ的に解釈しなおすことで、認知アーキテクチャの能力との対応付けが可能であることを示し、人間の知的な能力の評価と同様に、認知アーキテクチャの能力の評価が可能であることを示した。今後の課題としては、様々な認知アーキテクチャをこのフレーム

ワークに載せることで比較するとともに、粒度が粗いと考えられる推論能力の評価に関する流動性推論 (Gf) の詳細化を行っていく必要があると考えられる。

参考文献

- [Anderson 04] Anderson, J. R., Bothell, D., and Byrne, M. D.: An Integrated Theory of the Mind, *Psychological Review*, Vol. 111, No. 4, pp. 1036–1060 (2004)
- [Goertzel 13] Goertzel, B., Ke, S., Lian, R., O’Neill, J., Sadeghi, K., Wang, D., Watkins, O., and Yu, G.: The CogPrime Architecture for Embodied Artificial General Intelligence, in *Proceedings of IEEE Symposium on Computational Intelligence for Human-like Intelligence*, pp. 60–67, IEEE (2013)
- [Hernandez-Orallo 14] Hernandez-Orallo, J.: AI Evaluation: past, present and future, <http://arxiv.org/abs/1408.6908> (2014)
- [Langley 06] Langley, P.: Cognitive Architectures and General Intelligent Systems, *The AI Magazine*, Vol. 27, No. 2, pp. 33–44 (2006)
- [Newell 90] Newell, A.: *Unified Theories of Cognition*, Harvard University Press (1990)
- [Russel 09] Russel, S. and Norvig, P.: *Artificial Intelligence: A Modern Approach*, Prentice Hall, 3rd edition (2009)
- [Samsonovich 15] Samsonovich, A.: Comparative Repository of Cognitive Architectures, <http://bicasociety.org/cogarch/>, (Accessed 2015-12-09) (2015)
- [Schneider 13] Schneider, J. and McGrew, K.: The Cattell-Horn-Carroll (CHC) Model of Intelligence v2.2, <http://www.iapsych.com/chcv2.pdf> (2013)
- [Thórisson 12] Thórisson, K. R. and Helgasson, H.: Cognitive Architectures and Autonomy: A Comparative Review, *Journal of Artificial General Intelligence*, Vol. 3, No. 2, pp. 1–30 (2012)
- [ゲーツェル 14] ゲーツェル ベン: 汎用人工知能概観, *人工知能*, Vol. 29, No. 3, pp. 228–233 (2014)
- [三好 10] 三好 一英, 服部 環: 海外における知能研究と CHC 理論, *筑波大学心理学研究*, Vol. 40, pp. 1–7 (2010)