

CHCモデルに基づく対話エージェントのための認知アーキテクチャ

A Cognitive Architecture for Dialogue Agent Based on the CHC Model

市瀬 龍太郎^{*1}

Ryutaro Ichise

^{*1}国立情報学研究所

National Institute of Informatics

There are many cognitive architectures available nowadays. However, those architectures do not pay attention for covering all factors of human intelligence. In this paper, we propose a novel cognitive architecture based on the CHC model, which is a psychological model covering all factors of human intelligence.

1. はじめに

汎用人工知能を構築するためには、人間のような知的活動を行うための内部メカニズムを明らかにする必要がある。このような知的活動を行うための内部メカニズムは、認知アーキテクチャと呼ばれる。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も言われている [Thórisson 12]。そのため、これまでも CogPrime, LIDA など様々な認知アーキテクチャの研究が行われてきた。それらの認知アーキテクチャでは、知的処理に関わる様々なコンポーネントを利用し、それらが協調動作を行うように全体の設計が行われている。しかし、コンポーネントは、各認知アーキテクチャに応じて様々なものが使われており、汎用人工知能を設計するために十分なものが利用されているか不明であるという問題が存在する。

一方、心理学においては、知的活動に関わる要素にどのようなものが存在するかについて、長年、研究が行われてきた。そのような知的活動に関わる要素を整理したものととして、CHCモデル [Schneider 13] がある。CHCモデルは、人間の知的活動を実現する知能を細かい因子に分けてモデル化したものである。そのため、人間の知能を実現するための要素が網羅されていると考えられる。

本稿では、CHCモデルで提示された因子に基づき、認知アーキテクチャを構成することを試みる。しかし、CHCモデルでは、非常に多数の因子がからみ合っているため、全ての因子を対象に認知アーキテクチャ全体を構成するのは、多大な労力が必要である。そのため、本稿では、認知アーキテクチャ全体を構成するための第一歩として、対話エージェントを想定し、そのために必要な要素に関して、認知アーキテクチャの構築を行い、その結果について報告する。

2. CHCモデル

CHCモデルは、Cattell, Horn, Carrollの理論を統合した知能の因子のモデルで、知能が3つの階層から成り立っているとする考え方である。本稿では、Schneiderらの研究 [Schneider 13] に基づいて、CHCモデルについて述べる。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1つあるのみで、知能を集約したものになっている。第2

連絡先: 市瀬 龍太郎, 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, Tel:03-4212-2000, E-mail:ichise@nii.ac.jp

表 1: 結晶性知識 (Gc) を構成する要素

K0	社会における一般知識
VL	語や概念の定義に関する知識
MY	会話の際の文法などに関する知識
LD	会話の際の単語を理解する能力
LS	会話の際の話を理解する能力
CM	会話の際の考えを伝える能力

階層は、16個に分類されている。その中で、主に対話エージェントに関連する要素を列挙すると、以下の4つがあげられる。

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

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

この中で、対話に大きく関連するものは、結晶性知識 (Gc) である。結晶性知識 (Gc) は、第3階層でより詳細に分類がなされており、表1に示した6つの要素から構成されている。

3. CHCモデルに基づく認知アーキテクチャ

CHCモデルは、心理学のモデルであるため、計算機上で実装可能な認知アーキテクチャを作成するためには、CHCモデルを計算科学的な観点から分析する必要がある。そのため「データ」と「処理」の観点からCHCモデルを分析し、知能が4つの観点から計測できることが示されている [Ichise 16]。その分析によると、知能の因子である K0, VL, MY はそれぞれデータ量の多寡により知能が計測可能な要素とされており、これらの要素は知識を示すデータとみなすことが可能である。一方、LD, LS, CM はデータ処理の正確性により知能が計測可能な要素とされており、これらの要素は、データの処理を行うプロセスとみなすことが可能である。

本稿では、エージェントとして、対話の際の言語情報や周りの状況を示す視覚情報、外界からのフィードバックに相当する報酬が与えられ、言語や動作を用いて、外界への働きかけができるものを想定する。上記の考察に基づき、対話処理を行う際のデータフローに着目し、認知アーキテクチャを考えると、図1のような認知アーキテクチャを構成することが可能となる。

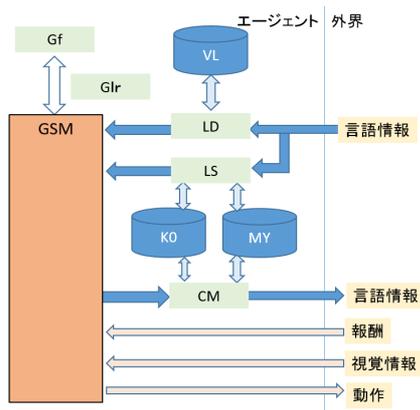


図 1: 対話エージェントの認知アーキテクチャ

この認知アーキテクチャにおいて、外界から与えられた言語情報は、2つのコンポーネント LD, LS により処理される。LD は、会話の際に単語を理解する能力である。単語を理解するためには、語や概念の定義などを入力された言語情報と結びつける必要がある。そのため、語や概念の定義に関する知識を格納する VL から情報を引出し、今後の処理に必要な情報を作業記憶となる Gsm に置く^{*1}。また、LD では、知識を VL から取り出すのみならず、言語情報から新たな知識を獲得し、VL に追加する機能も必要となる。一方、LS は、会話自体を理解する能力である。そのために、文法を解釈し、意味を捉えたり、社会的な文脈などを考慮した処理が必要となる。この処理では、一般知識 K0 や文法知識 MY のデータが使われる。LS の処理では、LD の処理と同様に、K0, MY から知識が取り出すのみならず、新たな知識の学習が行われる必要がある。LS の処理結果は、LD と同様に、作業記憶となる Gsm に置かれる。

作業記憶となる Gsm に置かれた情報は、記憶を検索、変換する能力である Glr や、推論・量的推論に関する能力である Gf によって、更なる処理が行われ、Gsm の情報が更新される。更新された情報を基に、CM が言語情報の出力を決める処理を行う。CM では、LS と同様に K0, MY の情報が使われ、出力する言語情報が決定される。

このエージェントにおいては、その他に、報酬や視覚情報などの処理が必要となる。これらの情報は、処理された結果が Gsm に置かれ、CM の処理の際に利用される。しかし、その処理の詳細なモデル化に関しては、対話処理の中核ではないため、今回は割愛した。同様に、動作についても、処理を割愛している。

4. 考察

本章では、このような認知アーキテクチャにおいて、実際どのような動作が行われるのかについて考察を行う。考察を行うために、本稿では、対話システム SHRDLU [Winograd 71] における状況を想定する。SHRDLU では、積み木を操作する環境を仮定して対話を行うシステムである。人間は積み木のある環境で、操作を指示することで対話を行っていく。

SHRDLU においては、以下のような対話を行う。

人：大きな赤い積み木を持ち上げて。

*1 CHC モデルにおいては、Gsm は、作業記憶の容量が知能の因子になることを示しているが、本稿では、単に作業記憶の領域として用いている。

計算機：了解。

この認知アーキテクチャにおいて、上記の対話を行った場合を考察する。「大きな赤い積み木を持ち上げて」という言葉が入力された際に、LD においては、VL に格納された知識とのひも付けが行われる。「積み木」が何を指すのか、「赤い」が何を指すのかといった情報が関連付けられ、Gsm に置かれる。また、それと同時に、LS の処理により、意味の理解が行われる。この場合には、「掴む(積み木)、赤い(積み木)、大きい(積み木)、移動(積み木、上)」といった形式で対話の内容が理解され、Gsm に置かれる。同時に、視覚的な情報が処理され、今の状況が Gsm に置かれる。Gsm に置かれた情報は、Gf, Glr により移動が可能か否かなどの推論が行われ、その情報が Gsm に置かれる。この対話の場合には、実行可能であると判断がなされる。その情報に基づき、CM が「了解」という、新たな言語情報を生成し、外界に出力する。それと同時に、積み木を持ち上げるという動作が外界で行われる。

近年では積み木世界ではなく、バーチャルワールドを用いた対話システムなども検討されている [Mikolov 16]。しかし、本稿で検討した認知アーキテクチャは、学習機構を持つため、そのような環境にも対応が可能であると考えられる。本稿では、記号主義的な方法で動作の説明を試みたが、実装を行う際には、コネクショニズム的な方法を用いることも考えられる。しかし、言語は、記号であるため、コネクショニズム的な方法を利用する際には、言語と分散表象知識をつなぐ機構を備える必要がある。

5. まとめ

本稿では、対話エージェントを対象として、CHC モデルに基づいた認知アーキテクチャの構成を試みた。この認知アーキテクチャは、CHC モデルに基づくため、人間の知能が網羅されることが期待できると同時に、それぞれのコンポーネントが知能の要素に基づくため、性能評価が可能という利点がある。今後は、本稿で詳細化していない、Gf, Glr や、報酬、視覚情報、動作の処理を詳細化していくと共に、実装を行っていく予定である。

参考文献

- [Ichise 16] Ichise, R.: An Analysis of the CHC model for Comparing Cognitive Architectures, in *Proceedings of the 7th Annual International Conference on Biologically Inspired Cognitive Architectures*, pp. 239–244 (2016)
- [Mikolov 16] Mikolov, T., Joulin, A., and Baroni, M.: A Roadmap towards Machine Intelligence (2016), arXiv:1511.08130v2
- [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)
- [Winograd 71] Winograd, T.: Procedures as a Representation for Data in a Computer Program for Understanding Natural Language, Technical Report AITR-235, Massachusetts Institute of Technology (1971)