

対話型人工知能シミュレーター (IDES) の基本理論

遠藤智彦[†] 大熊昭広[‡] 高田明典^{†‡}

尚美学園大学芸術情報学部情報表現学科

1. はじめに

これまで「対話型人工知能システム」に関する研究は数多く行われてきた。しかしながら自然知能の特徴的機能である「内言」「意識」「思考の流れの制御」などを機能として模倣する形式の対話型人工知能システムについての提案は少なく、また存在したとしてもそれらの多くは作動原理に関する心理学的モデルの提案にとどまるものであり、システムとしての実装を目指したものは考えにくいものが多かったと言える^{[1][2][3]}。また、「議論」「思考」などの自然知能の局所的特徴としての機能の実装を目指したものは存在するが、それらの多くも、一つの統合された対話型人工知能システムとして提案されているものではないものが多い。

「意識されない」何らかの現象が脳内に存在するという状況を、本研究では、fig-1 のようにとらえる。システムは第0層から第3層までの全4層からなる階層構造を想定したものとなっている。対話者は「外部世界」に存在しており、「文(英語)」を通して本システムとインターフェースをとる。

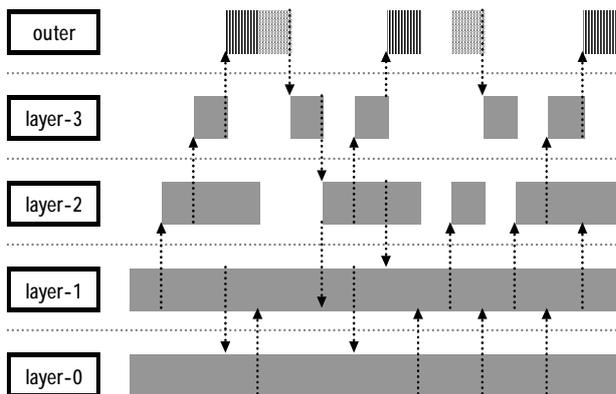


fig-1. A layer model of human thinking process

第0層は「下意識(非論理展開)」, 第1層は「下意識(論理展開)」, 第2層は「前意識」, 第3層は「内言」にそれぞれ擬せられる。モニターできるのは第3層のみであり、したがって一般に「意識」と呼ぶところのものは、第3層の「機能」であると考えられる。

Basic Theories for
Intelligence Driver Enhanced Simulator(IDES)

[†] Tomohiko Endo

[‡] Akihiro Okuma

^{†‡} Akinori Takada

Shobi University
Faculty of Informatics for Arts
Department of Digital Expression

我々はこれまでに、心理学的モデルおよび言語哲学的モデルに基づいた対話型人工知能システム (IDES) の研究を行ってきた。それは、「統合されたシステム」として自然知能の機能を模倣することを目的としたシステムであり、「対話型人工知能システム」の一つの基軸となりうるものであると考えている。

本研究においては、我々がこれまで研究・開発してきたシステムにおいて援用されている「基本理論」に関して、その心理学および言語哲学分野における理論的基盤に一度立ちかえり、検討を加えることを第一の目的としている。さらに、これまで援用してきた基本的理論の一部改良および変更を加えることのできる可能性に関して、心理学および哲学的分野からの知見に基づいて提案を行う。

2. IDES で援用される二つのモデル

対話型人工知能システムである「IDES」の基本機能は大きく二つに分類することができる。一つは「論理展開モデル」であり、もう一つは「動因展開モデル」である。論理展開モデルは、推論を行ったり、外界のユーザーに対して質問を実行したりして、現実を反映したより妥当な世界写像を内部に形成する。論理展開モデルはユーザーが入力した文字情報により、論理空間を構築する。また、その論理空間内における論理矛盾を探索し、それを修正することによって内部写像空間における論理的整合性を追及する。論理展開モデルによって構築される「世界写像」は、いわば「論理地図」である。それに対し動因展開モデルは「基本論理文」を成立させることを目的とし、「最終的なゴール」に到達するための地図として論理展開モデルによって構築されたその「論理地図」を用いて、「ゴール」に向けて進んでいくことを管理・実行する。

3. IDES の論理展開モデル

「IDES の論理展開モデル」としては、これまで if-then 関係に基づく論理探索モデルの使用を基軸として考えてきた。Aが「成立している」ならば、Bが「成立している」とする。ここで、A, B, Cは、論理文をあらわす。IDES では、ユーザーからの自然文の入力により、この「if-then 関係」を収集していく。これらの複数によって構成された論理的關係を基軸とし、「論理地図」を構築する。

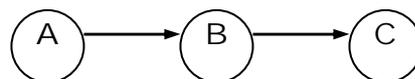


fig-2. A sample of logical structure based on "if-then" relation

「A → B」という論理的关系を基軸におくこのモデルは、いわば IDEs における「論理地図」を構築するためのものである。IDES は後述する動因展開モデルにおいて、この論理地図を利用して「最終的な目的」に到達しようとする。

しかし if-then 関係に基づく論理探索モデルのみでは現実世界の論理的構造との間で妥当な写像を構築することは難しいことが知られている。この状態を回避するために、我々は従来から古典論理の枠を超えて、「様相論理」の機構を IDEs に組み込む可能性に関して検討してきた。「様相論理」とは、Kripke などの研究に代表される論理学の一領域であり、合理的エージェント（環境に関する知識に基づき、達成すべき目標を目指して行為を遂行することの出来るエージェント）の論理モデルの基礎である。現在では Kripke による可能世界意味論が様相論理のスタンダードな意味論となっている^[4]。可能世界の概念は、様相論理のセマンティクスにおける最も基本的な考え方であると言える。様相論理のストラクチャーには、論理式を解釈するための世界がいくつもありえる。この様相論理を取り入れることによって、より複雑な論理展開を行えるようになると考えられる。

論理構造は、複数の if-then 関係が組み合わされたネットワークを構成する。ここにおいて、if A Then B という関係の前件部 (A) を「サブゴール」とであるとする。最終的なゴールの前段階には、いくつかのサブゴールが存在している。この「if-then 関係のネットワーク構造」を、ここでは「論理地図」と呼ぶ (fig-3)。

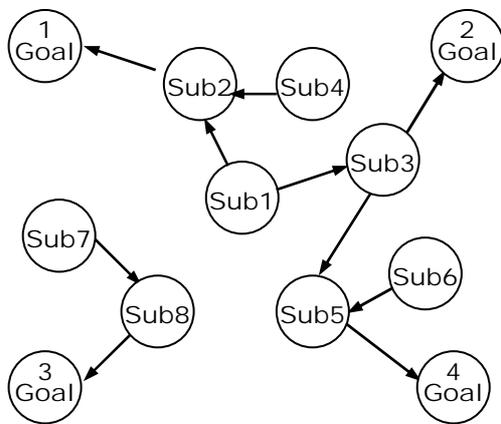


fig-3. A sample of logic network structure.

4. IDEs の動因展開モデル

IDES の「目的」は、あらかじめ開発者によって与えられているいくつかの「ゴール」によって決定される。それぞれの「ゴール」は、論理形式（対象 / 関係 / 対象）で表現された基本論理文であり、IDES はこの基本論理文を「成立させる」ことを動作の目的としている。この「目的」を達成するために、論理展開モデルで構成された「論理地図」上で「最適なゴール」を探索し、そのゴールに到達しようとするための行動モデルを「IDES の動因展開モデル」と呼ぶ。

動因展開モデルは、あらかじめ定められた「目的」もしくは「ゴール」を「成立させる」ために、何らかの出力を行う。それぞれの目的が成立するために必要な条件は「論理展開モデル」によって与えられる。論理展開モ

デルは、最終的な目的を成立させるために必要となる「サブゴール」を構成している。

あらかじめ組み込まれる「目的」は、心理学の分野で「動因」と呼ばれているものに等しい^[5]。初期に組み込まれる「動因」はわずか数種類だが、稼働時間が多くなり、蓄積データが多くなってくると、多くの「サブゴール」が生成されるようになる。

IDES における「最終的な目的」は、層状の構造を持つ IDEs システムの最下層（第 0 層）に設定されている「原則」である。IDES は、この「第 0 層ルール（原則）」から派生したサブゴールを「成立」させていくことによって、このゴールに到達しようとする。

5. 現状と展望

現在、本研究によって得られた「論理展開モデル」および「動因展開モデル」のアルゴリズムが IDEs の機能として一部実装されている。ただし、Kripke による可能世界モデルに関しては、今後の実装に向けてアルゴリズムを精緻化している段階である。

現状においては現在用いられている上述の基本理論が最も妥当なものであると考えるが、IDES の稼働時間が多くなってくること起因するさらなる問題の発生が予測される。実装および実験の結果に基づき、設計の調整および、さらなる改良を加えていく必要があると考える。

6. おわりに

本研究では、対話型人工知能シミュレーター (IDES) を構築する際に使用される「論理展開モデル」と「動因展開モデル」を提案した。本研究は、従来心理学の分野や言語哲学的分野で提唱されてきたモデルの統合的実装を目指してきたものである。当該分野においては、様々な有益かつ有意義なモデルの提唱がなされているが、現在は、すでにそれらを実装し検証していく段階に至っていると考えられる。

参考文献

- [1] A. Barr, E.A. Feigenbaum (編), 田中幸吉, 淵一博 (監訳): 人工知能ハンドブック第 1 巻, 共立出版 (1983).
- [2] Stuart Russell, Peter Norvig, 古川康一 (監訳): エージェントアプローチ人工知能, 共立出版 (1997).
- [3] Marvin Minsky, 安西祐一郎 (訳): 心の社会, 産業図書 (1990).
- [4] 戸田山和久: 論理学をつくる, 名古屋大学出版会 (2000).
- [5] E.C. Tolman, 富田達彦 (訳): 新行動主義心理学, 清水弘文堂 (1977).