

人工知能への道(4);対話する構造

2021年10月7日 | [岡島義憲の集積回路の明日に向けて](#)

(情報統合技術研究合同会社)

人工知能の父と呼ばれたマービン・ミンスキー氏(1927-2016)は、1961年に、脳内の Models(世界モデル、生成モデル)と Mind(注目と制御を行う装置)の間の対話(内省)が知性(Intelligence)を発現させるという仮説を持っていた(参考資料1)。また、数学者であるチューリング氏に始まる計算複雑性理論も、証明者と検証者からなる二つの計算機構の間の対話構造(Interactive Proof System)の証明能力の高さに注目して来ている(参考資料2)。そこで、本稿では、両者に現れた「対話構造」の類似点と相違点をまとめてみる。AI コンピューティングの装置が持つべき大構造を理解するためである。

ミンスキー氏の対話構造

1961年に書かれた論文(参考資料3、4、5)は、人工知能(AI)開発に向けて同氏が課題と考えた案件が列記されている。イントロ部と結論の章を除くと、その本文は、以下の5章からなる。

1. 検索動作 : The Problem of Search
2. パターン認識 : The Problem of Pattern Recognition
3. 学習システム : The Problem of Learning Systems
4. 計画策定と問題解決 : Problem-Solving and Planning
5. 抽象化情報の抽出とモデル群 : Induction and Models

1章から4章までは汎用的に現れると想定する4種のアゴリズムに関する課題であるが、「それらの整備だけでは知性が発現し得ない」として5章をスタートさせている。

5章の骨子は、以下である。

- ・ 知性の発現には、入力情報からの抽象化情報の抽出(帰納、Induction)が重要な役割を果たす
- ・ 脳の中に2つの自律動作する計算機構である Models(抽象概念群が作る世界モデルと生成モデル)と Mind が存在すると考えると、脳内の動作を理解しやすい
- ・ Mind は、Models と対峙し Models と対話する
- ・ Models を構成する抽象化情報(群)には、サブマシン(群)が対応して存在する
- ・ Models の計算機構の入出力インターフェースも、Mind の計算機構の入出力インターフェースも、Models 内のサブマシンのサポートによって符号化(エンコード)、または、復号化(デコード)されている(従って、Models の回路は、入れ子的(再帰的)構造を持つ)

以下、筆者の想像する処を過剰に含んだ表現となっているかもしれないが、サブマシンや Models は更に次のような特徴を持つと思われる。

- ・ 「モデル(a Model)」やサブマシンの動作を定義するのは、サブマシンが記憶するデータ(パラメータ)群である。モデルやサブマシンは、様々な事象や対象物に対応した時の学習やトレーニングを通じて、それらデータ(パラメータ)群のより良い値を獲得する。

- ・ Models(世界モデル群や生成モデル群)は、「モデル(a Model)」の集合である。
- ・ サブマシン群は、時折、動作を効率化するためにメンテナンスされ、自律的に抽象概念を抽出する。その抽象概念は、新たな「モデル(a Model)」であり、新たなサブマシンによってサポートされるとすべきかもしれない。
- ・ 獲得した抽象概念のサブマシンに関するデータは、画像理解や言語生成のための脳内の符号化(Encoder)や復号化(Decoder)の回路の動作を定義する。
- ・ “a Model”もサブマシンも、複雑に階層化された内部構造を持ちうる。各々の階層が自律的に動作する。
- ・ サブマシンが動作すると、他のサブマシンやそのサブマシン自身の次の動作に影響を与える学習動作を引き起こしうる。

ここでいう学習や抽象化は、1度の処理で完了する訳ではなく特定の集合全体へのオペレーションとなりうる。類似のデータが取り込まれた時、または、何らかの意味で関連データが短期記憶に取り込まれた時に、学習は繰り返され、その繰り返しを通じて、抽象概念(帰納化、Induction)の更新も行われるということであろう。

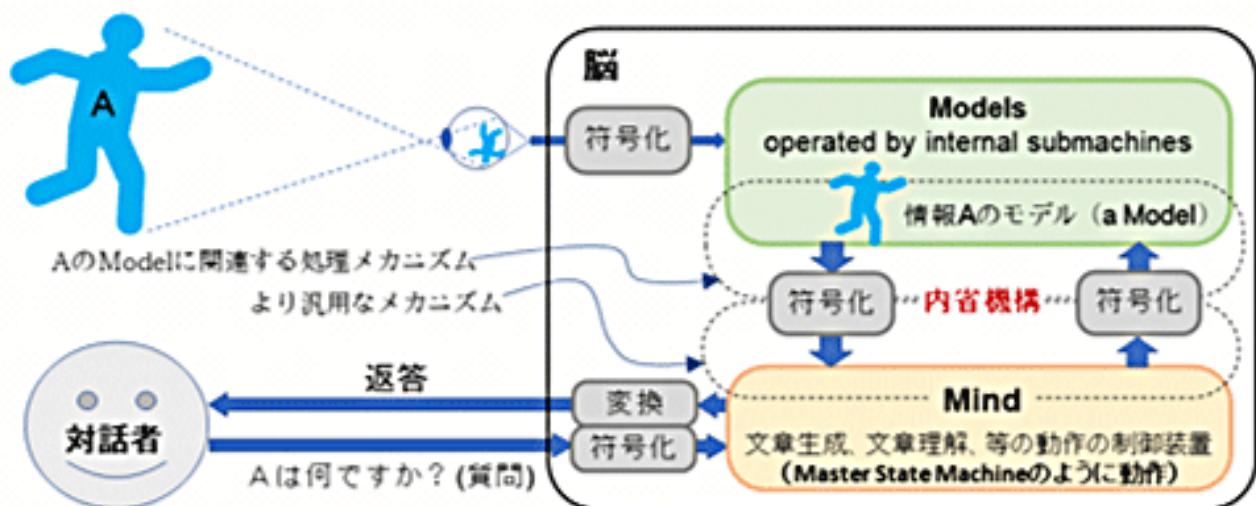


図1 ミンスキー氏の考える脳の中の Models と Mind の間の内省動作； 外部の事象(A)に関する情報は、脳内に符号化(エンコード)されて“a Model”を生成する。もし、対話者から、「Aは何ですか?」と問われると、Mind は、Models を検索し、返答された内容を元に解答文章を作成して、対話者に返答する。
出典: 参考資料 2~4 を元に、筆者が作成

Models については、以下のような表現での説明もある(図1)。

「脳は、脳内のいくつかのサブマシン(もしくは、彼の脳内に存在する Model)の振る舞いデータを符号化して出力することができる。サブマシン群や Models は、脳が、対話者からの質問に応答する主体である」。

筆者の理解では、サブマシンとは、次の通りである(注 1)。

- ・Models (世界モデル群や生成モデル群)を表現するハードウェア・プラットフォームであり、
- ・学習によって獲得した「内部データ」を元に、次の入力情報を処理する計算機構

Mind と“State Machine”の違いについて

Mind は、参考資料 4 の中で、以下のように登場する。

「もし、脳内の知識を担う部分である Models と新たな質問を生成したり思考プロセスを管理したりする汎用装置部分である Mind の二つを、解剖学的に識別することができるのであれば話しはしやすい。しかし、知識を担う部分と、エンコードを行う部分を解剖学的にも、機能面でも切り分けることができるとは思えない。なぜなら、知識は、エンコード(符号化)プロセスにも、デコード(復号化プロセス)にも使われるからだ」

つまり、Mind は、「Models のサポートの元に、新たな会話文章を生成したり、思考プロセスを管理したりする汎用装置」としている(注 2)。

Mind の動作は状態遷移回路(State Machine)に似るが、状態遷移回路とは異なる。脳内の特に Models 内のサブマシンは、各々自律的に動作しており、それぞれに「状態」が定義されるためである。各サブマシンが自律動作するため、脳全体に対して単一の「状態」をいうことはできない。

しかしながら、Mind は、Models に質問(実態としては、検索要求)を送ることによって、特定の情報(a Model)を切り出し、そのサブマシンの状態を読み出すことができる。これは、特定の情報(a Model)への「注目(Attention)」である。

Attention とは

Attention は、パターン認識の章(第 2 章)で登場する表現である。ミンスキー氏は、

「対話者が、脳に、『A は何ですか?』と尋ねると、Mind は、その質問をエンコードし、A に関するモデル(“a Model”)を検索し、“a Model”の状態を読み出す(図 1)」

との表現で、Mind を用いている。

Mind が「Models を検索して特定のサブマシンの状態を読み出す(特定の情報を切り出す)」動作を Attention と呼んでいると思えるが、その動作は検索を用いたパターン認識動作としても理解されているのだろうと思われる。

“Attention”によって、Mind は、特定のサブマシンから出力を受け取り、その他のサブマシンの出力を(ほとんど)無視することになる。従って、Mind から見ると、注目(Attention)によって読み出された特定“a Model”の状態は、Models 全体の「状態」であるように見えるだろう。その読み出された「状態」にて「Models の状態」を定義すれば、Mind は、“Models”の側の状態遷移とすることもできるだろう。

とすると、Mind は、「特定のサブマシンの状態をエピソードとして読み出し得た情報を元に、表現やシナリオを生成する回路」となりうるのではないだろうか? そのように仮定し、Models や“a Model”と、「表現

(シナリオやエピソード)」の関係を図示すると図 2 のようになる(注 3)。

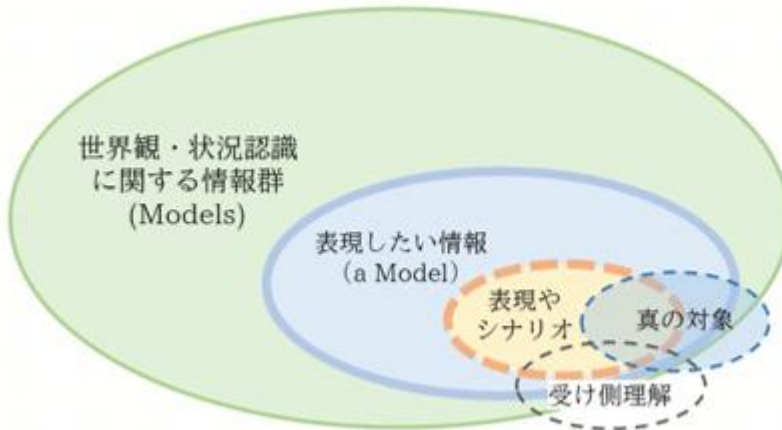


図 2 Models と“a Model”と「表現」の関係 Mind は、Models の中から、特定の“a Model”に注意を注ぎ、その中の情報を用いて、「表現」を組立て生成しようとする； 出典:参考資料 4 を元に、筆者が作成

計算複雑性理論における Interactive Proof System との比較

Models と Mind との間の対話構造(内省)は、計算複雑性理論でいう Interactive Proof System の証明者と検証者の間の対話構造に似る(図 3、参考資料 2)。

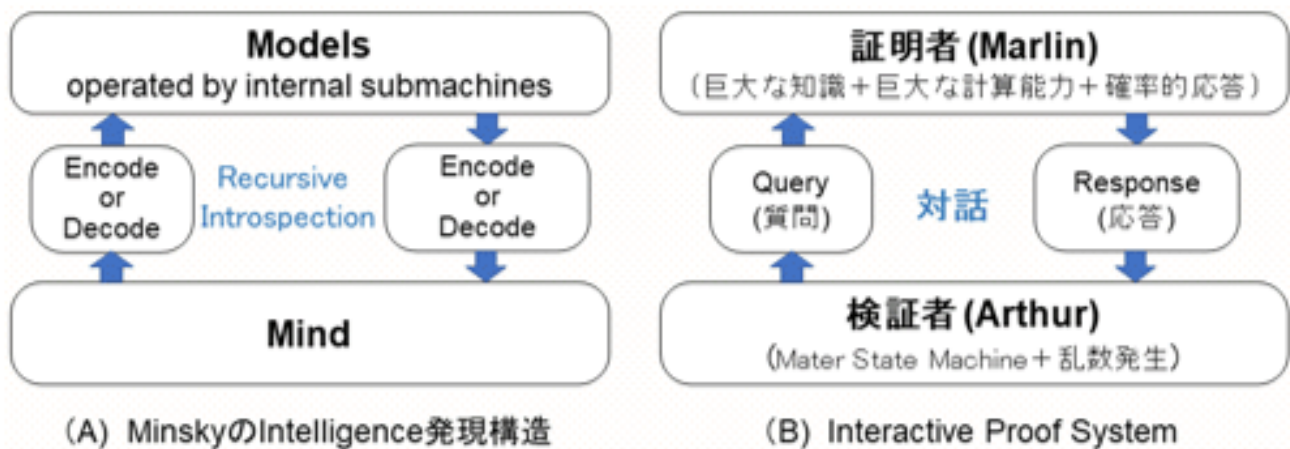


図 3

ミンスキー氏の知性 (Intelligence) 発現の機構と、計算複雑性理論でいう Interactive Proof System の構造の比較 共に、知識のための計算機構 (Models や証明者) と、文章作成のための計算機構 (Mind や検証者) を持ち、両者の間の対話にて適切な情報やシナリオ探すというオペレーションを行う 出典:参考資料 1 と 2 を元に筆者が作成

共に、知識のための計算機構 (Models や“証明者”) と、文章作成のための計算機構 (Mind や“検証者”) を持ち、両者の間の対話にて適切な情報やシナリオを探すというオペレーションを行う。

Interactive Proof System では、命題証明の効率を最大限に高める観点から、証明者と検証者の間の役割分担を図 3(B) のように設定するが、ミンスキー氏の知性発現構造における Mind の役割もエネル

ギー効率の最大化である可能性が大きい。両者が、「乱数に基づく質問の作成」や「確率的応答」を採用するとの点も類似点である(ニューラルネットワークも、確率を応答しうる)。

相違点としては、Interactive Proof System は、命題の証明の装置であるため計算が停止することを重視するが、一方、「脳の場合は、生命活動であり、生ある限り計算が完了する必要はない」という点が上げられる。サブマシンの各々は自律して動作するためループ状態を維持したり出力を迷い続けたりしても、それらのサブマシンを Attention の対象から外すことは支障がないだろう。

脳内の複数のサブマシンによって質問への応答が生成されるとの考えは、前野隆司氏の受動意識仮説に似る(参考資料 6)。前野隆司氏は、一般向けの解説書の中で、脳の中のニューラルネットワークを「大勢の小人」と表現し、「脳の中の傍観者が注意を払って意識した『小人』の働きが意識に上る」と言い、また、「意識される『小人』の陰で、意識されない大勢の『小人』が機能している」という。

自律的、並列に動作する大量の“a Model”や“小人”をエミュレートする人工知能用のコンピュータは、現代の中央制御式のコンピュータとはハードウェアとは大きく違っているべきだと筆者は思う。

注

1. 機械学習や深層学習では、「生成モデル」と「世界モデル」は区別され、「生成モデルによって世界モデルが形成される」とするが、Minsky のモデルは、それら両方のモデルを包含しているように思える。
2. ミンスキー氏の Mind には、日本語でいう『心』と『思考』を合わせた意味も込めていると思うが、「意識する」の意味合いで“Be conscious”という動詞的表現も使われており、“Be conscious”には受動的、Mind には主体的の能動的な機能を言い含め、使い分けていると思われる。
3. 脳は、入力情報である「真の対象」に似るように「表現」を制御しようとするかもしれないが、両者の間には一定の差があるのが通常である。また、「表現」を受け取る他者の理解も、「表現」とは多少なりとも差のある内容となるだろう。その誤差は、非建設的に作用することが多いかもしれないが、時には、建設的に作用することもある。

Models や“a Model”は階層構造を持つが、その下層の階層は共通の抽象概念を用いる。従って、図 2 は、決して、「独立要素の集合」の包含関係を意味していないことに注意が必要である。

参考資料

1. 岡島義憲、「人工知能への道(2):Intelligence とは何か」、セミコンポータル、(2021/07/28)
2. 岡島義憲、「人工知能への道(3):対話する構造の重要性」、セミコンポータル、(2021/08/26)
3. Minsky, M., “Steps toward Artificial Intelligence”, *Proceedings of the IRE Contents*, Vol. 49, 1961.
4. Minsky, M., “Step Toward Artificial Intelligence”, in *Computers and Thought*, published by McGraw-Hill, pp. 406-450, 1963.
5. Minsky, M., “Step Toward Artificial Intelligence”, in Marvin Minsky Homepage in MIT Media Lab.
6. Minsky, M., “[Matter, Mind and Models](#)”, *Proc. International Federation of Information Processing Congress*, vol. 1, pp. 45-49, 1965
7. 前野隆司、「脳はなぜ『心』を作ったのか」、ちくま文庫、2010 年