

第8章 心の科学：将来に向けて



カイロの塔

記号処理とニューラルネット (NN) の立場で心の働きやデータについて議論してきた。本論は、認知科学を背景にして AI の基礎を築く取り組みと言えよう。本章では、関連研究との比較によりまず本論を位置づける。その上で本論を含む AI 研究の応用、AI の今後の課題、さらには AI の近未来に向けての提言を議論しよう。

8.1 関連研究との比較

提案理論が妥当かどうか調べる一つのやり方は関連研究と比較、検討することである。ミンスキーは、述べるまでもなく AI の父とも称される先駆的研究者の一人で、この分野を牽引してきた。ジョンソン-レアードは、認知科学に造詣が深く、鋭い洞察力で独自の理論をまとめあげた。そこで本節では、ミンスキーの心の社会 [46] ならびにジョンソン-レアードの心とコンピュータ [11] と比較しよう。重要な項目について3者を比較すると、次の通りである：

心の理論の比較（

著者(ミンスキー、ジョンソン-レアード、岡田)、

表題(心の社会、コンピュータと心、心・脳に宿る心)、

動機(心をもたないエージェントによる心の構成、計算論による認知心理の体系化、脳に宿る心の体系化)、

機能([個性、記憶、感情、発達、推論、表現、言語フレーム、思考の領域など30項目]、

[視覚、学習・記憶・行為、思考、コミュニケーション、意識・無意識]、

[知覚、認識、推考、表出、動作、情緒、記憶、言語])、

必要十分な機能 (-, -, 議論あり)、

アーキテクチャ(自律分散方式を提唱、計算論、

自律分散方式（計算論+NN 論）に追随），
常識（-, -, 語彙を背景に概念体系を網羅的に調査），
シミュレーション（-, -, 個別および総合的）
3者の動機が興味深い。ミンスキーは、こびと論議のような循環論に陥らないよう、心をもたない小さなエージェントから出発して全体を把握しようとした。ジョンソン-レアードは、チューリング機械に基づいて認知心理の世界を俯瞰しようとした。本書は、筆者自身の AI 研究の体系化である。ジョンソン-レアードの取り上げた心理機能と本書には共通性がある：

機能の比較（（ジョンソン-レアード vs 岡田），
（視覚 vs（知覚，認識）），
（（学習・記憶） vs 記憶），
（行為 vs（表出，動作）），
（思考 vs 推考），
（コミュニケーション vs 言語），
（（意識・無意識） vs（情緒，主体性））

残念なことにミンスキーもジョンソン-レアードも心の機能を取り上げるさいなげそれらに注目するかの説明がなく、漏らしていないか十分性について懸念が残る。アーキテクチャについて、本書はミンスキーに追随している。ミンスキーは常識の重要性を認識しているものの自ら調査しておらず、ジョンソン-レアードも意味の表現形式にふれるのみである。以上から本書には、注目機能の十分性、欲求・情緒による主体性、語彙を背景にした常識概念の網羅的調査、コンピュータシミュレーションによる妥当性確認などの点で、特長があるものと認められる。

8.2 AI の応用と課題

8.2.1 計算論の応用

近年、スマートフォンを用いてメニュー選択形式で求める情報を得る、あるいは入学試験においてマークシート方式で回答を選択するなど、与えられた選択肢から答えを得る機会が増えているように見受けられる。このような様式が常態化すると、必要な場面で自らの意見を述べたり、困難な局面や危険な状況を自ら打開したりする力の不足が大いに懸念される。問題解決の能力を培うには、やはり日頃から基礎的な訓練や教育が必要である。

AI教育を実施するための大学の専門課程についてはすでに1990年代に取り組みがなされている。学術会議において知識科学の体系化とカリキュラム編成に関する作業部会が設置され、20名余の委員が2年間にわたって調査、討論し、報告書が答申された[47]。本書の心のモデルの原型が体系化の枠組みとして採用され、それに沿ったカリキュラムも提案された。報告書は学会誌ならびに理工系情報学科協議会にも示され、わが国の大学における情報系専門教育に指針を与えた。そこで以下では、大学の専門教育ではなく、高校の一般教育課程にAIの成果を応用することを考えよう。

高校の教科

初めに高校教育課程における教科を一覧する。文部科学省の示している教科目録[48]に沿って項目をそのまま並べると、次の通りである。

高校教科(国語, 地理と歴史(地理, 歴史), 公民(現代社会, 倫理, 政治と経済), 数学, 理科(物理, 化学, 生物, 地学), 保健体育, 芸術(音楽, 美術, 工芸, 書道), 外国語, 家庭,

情報(社会と情報(

情報活用(メディア, アナログ・デジタル, 表現・伝達),
ネットワーク探索(インターネット, コミュニケーション),
情報化社会の課題(モラル, 安全, 著作権, 個人情報),
望ましい情報化社会))

情報科学(

コンピュータの仕組み(コンピュータ, ネットワーク, データベース),
問題解決とコンピュータ活用(基礎(問題の分析, 考案, 解決手法),
モデル化・シミュレーション, アルゴリズム・プログラム),
情報化社会の科学的理解(社会の情報化, 安全性, モラル))

項目は、教科名というより分野名もあり、一様ではない。議論の都合で細かい名称を付した項目もある[49]。まず情報科学に注目しよう。進化の激しい分野にもかかわらず、かなり整理されていることが伺われる。細かい項目の中に、問題解決とコンピュータ活用が見出される。すでに情報の世界では、知能の基本的能力として問題解決の大切なことが認識され、教科内容として取り扱われている。しかしながら人の問題解決能力を、情報分野のごく一部で取り上げるだけでよいものであろうか。本書の議論から明らかなように、認識、推考および表出の能力は、一般に思考と呼

ばれる能力の基礎である。以下では教科の不足している部分を補い、明確な理念の下に体系化を試みる。

教育の理念を“精神と身体の健全な育成”ととらえて、以下の体系を提案する：

精神・身体（

精神（

知性（

一般（**思考**（認識・推考・表出），

言語（国語，外国語），

個別分野（

数学，

理科（物理，化学，生物，地学，

情報科学（コンピュータ，ネットワーク，データベース，アルゴリズム，プログラム，モデル化，シミュレーション）），

社会（（現代社会，歴史），

（政治と経済，社会と情報），

家庭と地理（家庭，地理）

感性（

芸術（音楽，美術，工芸，書道），

倫理），

身体（保健体育）

精神を大きく知性と感性に分け、知性をさらに一般と個別分野に分けることは自然であろう。一般分野に、従来情報科学で扱われていた問題解決を発展させ、思考として組み込んでおり、逆に*情報科学*の方はコンピュータの技法に特化している。なお個別分野の社会と情報はそのままである。

思考の教科内容

思考では、具体的に次のようなデータ構造あるいはアルゴリズムを取り上げてはいかかであろう：

- ・ データの構造化

現実社会の事物を構成要素とそれらの関係に注目し、フレーム構造として表現する。

- ・ IF-THEN処理
演繹的，機能的あるいはアブダクションの手法で因果関係を追求する。
- ・ パターン照合
一方のパターンから手がかりとなる特徴を発見し，それを根拠にして芋づる式に他方のパターンとの共通性を探る。
- ・ 探索
思い付きや成り行き任せで対処せず，縦優先あるいは横優先の手順に沿って問題全体を系統的に調べ尽くす。
- ・ 試行錯誤
少し進んでは新たな手掛かりを探り，その手掛かりに基づいてさらに進む，という手探りの繰り返しで問題解決に取り組む。

以上の具体例は第2～4章で詳しく論じた。従前から“考える”ことに対する教科の不足を認識している人は少なくなかったはずである。しかし専門分野である認知科学の分野でさえ，思考の分野の複雑さゆえに体系化の試みがなされてこなかった。AIが誕生し半世紀以上の歴史を経た今日，ようやく可能となった提案と言えよう。現在の教育課程は長い伝統と数多の経験ならびに討論を経た成果であり，軽々に修正や軌道変更を加えてよいとは考えていない。しかしAI研究の成果もまた，人の知能の重要な性質をとらえている。その成果がすでに*情報科学*に一部取り入れられているからには，さらなる改善，発展のなされることを期待したい。

8.2.2 ニューラルネット／深層学習の応用

近時 AI 技術は多種多様な分野で急速に応用が進んでいる。例えば，囲碁や将棋の対局システム，公道を走る車の自動運転システム，4千万人のスマートフォン利用者と個別に対話するシステム，インターネットの Web サイトを世界規模で検索して利用者に好みのニュースを配信するシステム，地球規模の気象データを処理する天気予報システム等，枚挙にいとまがない。それらの特徴の一つは，いわゆるビッグデータすなわち大規模なデータを高速かつ大量に処理する点，もう一つは類推の手法を用いる点である。

対局プログラム

ビッグデータの観点から，囲碁や将棋などの対局プログラムを眺めてみよう。α碁は2016年に世界チャンピオンのイ・セドルと5番勝負し，4勝1敗と圧勝した。

α 碁は15万の対局データを記憶装置に蓄え、さらにプログラム同士の3千万回の対局を通じて、囲碁の知識を習得したという。15万という数字は、一人の棋士が毎日4回の対戦もしくは棋譜学習を行い、それを100年間続ける数に相当する。さらに3千万回という対戦訓練は、もはや人の能力からしては想像の範囲を超える。AIの大きな長所の一つは、この人間業では不可能な大規模データあるいは高速の処理にある。

ビッグデータの陰に隠れているが、もう一つ、人間業ではない数字がある。短期記憶の容量に関し、AIは魔法の数字の制約を受けない。このことがどのような意味をもつのか。一般に交互に手を打つ対局において、次の一手には多数の曖昧性がある。候補の手ごとに、後に続く一連の手筋を読まなければならない。すべてを調べつくすことはしばしば天文学的数字となる。そこで有力候補を絞り込むが、棋士の側は天性の才能と長年の訓練とにより核心の知識や事例への、素早い接近あるいは探索を試みる。その際魔法の数字 7 ± 2 の範囲内に候補を絞り込み、優先順位をつけてそれぞれの一連の手筋を深く読む。この範囲を越すと、専門棋士といえども混乱を起こす。それに対しシステムの側は、短期記憶に制限はなく二桁の有力候補でも困難はない。棋士にとってシステムが思いもよらない手を打つと映る原因の一つは、この生理的条件にも起因しているであろう。

従来からもコンピュータは高速、大量かつ正確という点で人の道具として役に立ってきた。これに知能が加わり、分野を限定すると、人より強力な知的処理も可能となる。ビッグデータの処理はその典型的な応用例である。

自動運転

特に乗用車に注目すると、自動運転には大きく三つのステップがある。

1. 運転支援

アクセル、ハンドルあるいはブレーキのいずれか一つあるいは複数の操作を支援する。

2. 部分的な自動運転

自動システムが三つの操作を行うが、必要に応じて人が操作する。

3. 全自動運転

自動システムが常にすべての操作を行い、人は運転に関与しない。

自動化で核心となるAI技術はTVカメラのとらえる画像の認識である。2.2節の庭園の散歩において述べたように、時々刻々移り変わる視界をとらえなければなら

ない。画像認識は深層学習の最も力の発揮できる応用分野であり、以下順を追って要点を述べよう。

ステップ1では、まず道路の連続性の把握が基本である。前方の路面や路側帯の連続性を白線やガードレールなどで追跡する。次は交通規則、すなわち信号、種々の標識、路面上の指示などを読み取り、規則に則って走行する。ちなみに規則は多数の人や車が安全に道路を移動する拠り所である。その次は、規則以外の安全確認である。前方を横切る人や物体、前後左右の車間距離など安全を妨げる人、物、あるいは出来事を認識する。

ステップ2では人とシステム間の操作の切り替えが重要である。特に緊急事態では、システムが画像認識のできないこともままある。それに対し、人は不注意で居眠り運転することもある。切り替えには、一方では人の高い認識力と臨機応変の対応力を活かしてシステムに割り込む、他方ではシステムが素早く正確な操作能力を発揮して人の運転に割り込む、といった協調関係が求められよう。

ステップ3では、自動運転を認める条件あるいは環境をいくつかのレベルで制度化する必要がある。例えば晴れた日に渋滞のない高速道を走行する状況と雨降りの夕方に市街地の一般道を走行する状況では、画像認識の困難さが格段に異なる。

運転支援車が普及すると、健常者はもちろん、高齢者や障害者に大きな福音となる。健常者の場合は、長距離運転等の専門職の人に利点大きい。また高齢者や障害者に取って信号の見間違いやブレーキ反応などで支援が得られれば心強い。安全面で自動運転には2面性がある。一つはセンサー技術の向上でより安全になる。カメラを前後時には左右にも設置すると、人の視界とは格段に注視力が增加する。もう一つは技術力、特にシステムの画像認識力が未熟で、誤認識してしまうばあいである。科学技術は実社会への応用に常に慎重でなければならないが、人の生命に強く結びつく自動運転においては、特にそれが言える。

機械翻訳

深層学習の機械翻訳への応用を議論しよう。深層学習では事例ベースに基づく類推を行う。類推の方式には種々の変化形があり一概には言えないが、原理についてはすでに2.3節で述べた。以下では機械翻訳の歴史にふれながら、取り組みの要点を述べよう。

機械翻訳の試みはコンピュータの出現とほぼ同時期の1940年代に遡る。伝統的に文法規則に基づく翻訳が模索されたが、細かい面で種々の変遷を重ねてきた。例

えば文法として直接構成要素法かあるいは依存文法かによって、処理アルゴリズムは大きく異なる。また音声言語の一部では確率文法も用いられた。しかし句構造文法の出現によって主流は定まった。書き換え規則を中心に構文解析し、浅い意味内容すなわち格構造を介して訳文を合成する。実用システムに対しては、次のような課題がある。

- (1) 一つの言語の文法を網羅する規則ベース
- (2) 浅い意味の把握に必要な語彙の意味素性と格関係の意味ベース
- (3) 文脈を把握する理論と知識ベース（意味ベースを含む）

(1) に関し 1980 年代には、わが国の大学やコンピュータメーカーの研究開発グループはかなり強力な規則ベースを構築していた。しかしそれに基づく翻訳で学んだことは、いくら規則を追加しても次々に例外的な文が見つかり、一つの言語の文法体系を網羅することの難しさである。そこで登場したのが 1990 年代にかけての事例翻訳である。規則に従うよりも用例に倣って例外をカバーすることを目指すもので、深層学習による翻訳の先駆けでもある。しかし事例翻訳も、大幅には翻訳効率に結びつかなかった。事例通りの入力文は少なく、単語の品詞への置き換えあるいは事例の構造化すなわちある程度の規則化が必要になる。これをあまり進めると規則に基づく翻訳への回帰になり、ジレンマが生じる。また、実用に耐える事例ベースやシソーラスを構築できなかったことも効率改善を阻む原因の一つである。事例翻訳に次いで登場したのが統計的手法で、その基本的な考え方は 3.3 節の統計的処理で述べた。確率で文法をとらえる手法はいったん影を潜めていたが、新たな発想で再登場した。

(2) についてわが国では国の支援によりコンピュータメーカーの共同体が意味辞書の開発を進めた[50]。細かい課題はあるものの一定の成果が得られ、5章で参考にした分類語彙表とともに自然言語処理の研究開発に寄与している。最後の(3)については、次項で述べよう。

深層学習による翻訳は、長い機械翻訳史から見れば緒に就いたばかりである。規則と事例の間のジレンマにどう対処するか。また記号表現された意味ベースを用いるのか、あるいは記号的コネクショニストに類する NN で意味内容を類推できるのか。さらには事例ベースによる類推のみで常識的判断に相当する処理を実現できるのか、今後の取り組みが注目される。

8.2.3 課題と対策

記号処理の AI がやや足踏みしている状況で、深層学習が活況を呈している。AI の課題と、その解決に向けての対策を議論しよう。

記号処理

AI において知識とりわけ常識の重要性が指摘されて、半世紀が経つ。常識は大切かもしれないが全体の解明は不可能に近い、と取り組みの敬遠された時期もある。逆に国策として意味ベースの開発に取り組んだ時期もある。取り組みに濃淡があり、継続性に欠ける。常識ベースの構築がなぜ難しいのか、二つの大きな理由がある。

➤ 常識調査の手順

述べるまでもなく常識の体系は複雑で、曖昧で、かつ巨大である。やみくもに調査しても重箱の隅をつつく類となり、全容の解明につながらない。そのような体系を紐解くには、確たる手順が必要である。その手順に沿って追求していけば時間を要しても全容を網羅できる、という見通しがなければならない。従来の取り組みにはその手順あるいは筋道が欠けていたと言える。

➤ 抽象概念からの特徴抽出

概念が抽象化するにつれ本質的な特徴を見抜き、陽に記述することが難しくなる。具象概念の場合はいくつもの具象的特徴の束で把握できるが、抽象概念の場合は一つあるいは二つの抽象的特徴しかなく、その記述はしばしば他の抽象概念による言い換えに陥る。抽象化には 6.1 節で述べたように過程があり、その成り立ちにまで配慮して特徴を抽出しなければならない。従来は、その姿勢や洞察の訓練が欠けていたと言える。

本書では、終始一貫して言語の窓から心と脳をとらえてきた。第 5 章は、その窓から常識概念の体系とらえた縮図である。常識には、基本概念、それらの合成概念、ならびに種々の概念を結びつけたエピソード概念の 3 種類があり、これらで網羅できる。統計的には約 3 千 3 百の基本概念とたかだか数千の標準的エピソード概念がある。エピソード概念の抽象度は格段に高いが、それには筋書きと制度という枠組みがある。常識の構造はフレーム構造で記述でき、基本概念からエピソード概念まで入れ子構造で統一的に表現できる。また文脈は、大きなフレーム構造を種々の角度から文の列で表現したものとして把握できることを、3.3 節の理解と生成で示唆した。

次に抽象概念については、最も抽象度の高い真・善・美についてそれぞれ本質的

な特徴の抽出を 6.2 節で試みた。特に美については、芸術的創作といえども 2.3 節の試行錯誤のアルゴリズムが適用できることを示したうえ、計算論の立場で感受性、発想力ならびに表現技術という、芸術面での 3 要素を指摘した。また 3.3 節の文章の理解と生成のアルゴリズムならびにフレーム構造は、文学作品の理解や生成にも通じる。

ブリタニカ百科事典は、18 世紀の後半に初版が刊行され、現在では人の知識が 50 万もの項目で体系化されていると聞く。250 年もの編纂の歴史とくらべれば、AI は歴史そのものが五分之一に過ぎない。本書の提案に沿って常識データ、とりわけ数千のエピソード概念の分析データを積み上げれば、21 世紀の後半にはブリタニカの記事を意味理解できる常識ベースの原型が視野に入るのではなからうか。

NN

初めに NN だけを用いて心全体の処理を実現できるかを考えてみよう。一つの課題として変数処理に注目する。3.2 節の記憶で述語論理の形式にふれた。記号処理では例えば“X は飛ぶ”という性質／述語を $\text{fly}(X)$ と表し、システム内のデータやプログラムで使用する。変数処理は、この X に一つの定数を代入することが基本である。さらに、例えば鳥を代入して不適切なら飛行機と、適切な定数が見出せるまで代入を繰り返す、単一化と呼ばれる変数処理がある。加えて言語処理では単一化に基づく限定節文法と呼ばれる高次の解析手法もある。NN では、4.1 節で紹介した記号的コネクショニストと呼ばれる高次処理の研究において、格構造の要素に制約を満たす一つの定数を代入する処理が見出される。しかし限定節文法のような高次の変数処理は未だ見当たらない[51]。

NN ではシステムを大きくするとブラックボックス化の問題の生じることはすでに述べた。そのためシステムの分割・構造化、換言するとモジュール化を避けて通れない。モジュール化を行うのであれば、当然プログラムとの結合も可能になる。したがって変数処理を初め高次の記号処理をすべて NN で実現することは、可能性の議論は別として、非現実的である。ちなみに第 7 章の自律分散システムは、低次の処理では NN のモジュール化を極限まで進めて知覚や動作を受け持つ。それに対し高次の処理ではプログラムを用いて変数を始め高次の推考や言語を受け持つ。このような協力関係が大切と考える。

8.4 将来の展望

人には心があり、それぞれ主体性を持つ。自らの意志で思想を持ち、発言し、行動する。AIも人と同様の心を持つのであろうか。

8.4.1 AIは心を持つか

二つの考えを紹介しよう。

▶ 科学技術

哲学や論理学を背景にして弱いAIと強いAIというとらえ方がある[9]。

サーレは、AI研究者が取る立場として“弱いAI”の立場と“強いAI”の立場を区別した。弱いAIは、心の働きに関する仮説を定式化し、テストするための強力な道具としての機能だけを計算機に認め、計算機が人間の心の働きを正確に模倣する能力をもつという主張だけを行う。それに対し、強いAIは、適切にプログラムされた計算機は実際に心をもちうると考える立場であり、物事を理解しているとか、信じているとかいった、文字通りに心的な状態を計算機に帰属できると考える立場である。……

デジタル人工知能学事典, 2.1節より抜粋

プログラムは当然ハードウェアとしてのコンピュータを前提にしている。NNでさえも実際はコンピュータ上のシミュレーションである。強いAIの考えは、無機質なコンピュータが心を持つことを認めている。

▶ 文化芸術

一方文化芸術の分野を背景にすると、人の感動を伴わない文学や美術には意味がない。コンピュータがあたかも情緒を持つかのごとくに振る舞ったり、一見芸術的と見える作品を出力したり、人間性のシミュレーションのできることは否定しない。しかし、生身の芸術家の非凡な感性、それに裏付けられた訓練、そして試行錯誤を繰り返す創作過程を重く見て、そのような背景を持たないコンピュータの出力を芸術作品と受け止めない人も少なくない。

8.4.2 人間社会と心を持つ人工物

初めに、心の階層の立場から強いAIについて議論しよう。具体例として小学1年生程度の知能をもつロボット/強いAIを取り上げ、生命現象あるいは生理機能に焦点を当てよう。ロボットは、無機質なので当然生理機構をもたない。しかし強

い AI なので重い病気になると命を落とすなど、生命や生理に関する常識的な会話はできる。この状況は、生理機構に接地するはずの概念的特徴が下位レベルに向けて接地先を失っているが、上位レベルに向けては言語の意味素性の役割を果している、と解釈できる。このロボットが小学校に入学すると、自身は給食を食べないし夜睡眠も取らない、また怪我をして痛いと言わないし出血もしない。しかし他の生徒が給食を食べる、睡眠を取るなどの意味は理解できる。結局クラスの中で生徒たちとロボットの間には、仲間として交流できても人間生活の基本的事項を共有できないので、相互にかみ合わない部分の存在することを否定できない。心の一部、とりわけ生命現象という人間生活の基盤部分が通じ合わないことは、取りも直さずロボットが人と同様の心の持ち主でないことを意味する。このように無機質なロボットの振る舞いに心を認めると、人間社会に歪が生ずるのではなかろうか。

科学技術が進歩し、いつか蛋白質も合成できる日がこよう。かりに人間と同様の生理的機能をもつ人工物が合成可能になったとする。その人工物は、生命を授かり、栄養を摂取して生育し、知覚し、運動し、そして人間同様の生活をするであろう。その間資源も消費する。学校に通い、大人になってようやく一人前になるが、中には落ちこぼれる物も出てくる。このような状況は人工物が人間社会に割り込んでいると言える。それだけではない。万一、心身共に人間と同様の人工物が人間社会で共生したとする。それらは生命体であり、場合によっては生きる権利を主張し、社会における市民権の問題が生じるであろう。

AI が無機質な道具として人間社会を支援するのか、あるいは生命をもつ仲間として人間社会に入り込むのか。このような議論は、人間とはなにかという本質を問いかける。筆者は、真理の探求や医療技術の開発など特殊の目的の場合は別にして、人間同様の人工物が社会に共生する意義はない、と考える。読者はいかがであろう。

最後に、読者のみなさんにお尋ねしたい。本書の議論を通じて心と脳の働きと仕組みについて理解してもらえたであろうか。はしがきでいくつか質問事項を列挙したが、それらの疑問も解けたであろうか。本書が、皆さんの心と脳の科学に対する知的好奇心を刺激し、さらなる探求に意欲を燃やす一助となり得たならば、筆者にとって何よりの喜びである。なお筆者の勉強不足により万一誤り、不適切な記述、あるいは独善的な思い込み等のあるときは、どうぞ指摘や批判をお願いしたい。また率直な意見もお聞かせ願いたい(=>okada@ai.dialog.jp)。検討して今後の改善につなげたいと考える。