

特集「脳科学とAIのフロンティア」にあたって

山川 宏

(株式会社ドワンゴ, 全脳アーキテクチャ・イニシアティブ)

森川 幸治

(パナソニック株式会社)

岡本 洋

(富士ゼロックス株式会社)

寺島 裕貴

(NTTコミュニケーション科学基礎研究所)

大森 隆司

(玉川大学)

1. はじめに

深層学習をはじめとした人工知能 (AI) 技術の急速な進展は, AI 技術の適用範囲を押し広げ, 社会に対して大きなインパクトを与えつつある. この状況下でさらに AI 技術を深化させるために, 私達は何を考えなければならないか. この問題意識をもとに本特集「脳科学とAIのフロンティア」を企画した.

AI 研究の目的の一つとして, ヒトのような形での知能の実現は含まれるだろう. 例えば汎用 AI 研究者である Pei Wang 氏は AI の理論構築においては, いかにしてヒトのように機能するかを記述する科学と, いかにヒトのように振る舞うかを設計・開発するための工学の両面に支えられる必要があるとしている [Wang 12].

一方で, AI において脳の知見を生かしたいという基本的な欲求は以前から存在した. 脳の知見を AI に活用することは Brain-inspired Computing と呼ばれ, その意義は自明にも思える. しかしこれまでは, AI が脳に学べる機会は必ずしも多いとはいえなかった. それはなぜであろうか.

本特集における私達の立場は, 近年の技術的状況の変化により, AI が目指すべきすべての機能を包含する「脳」

をより深く知ることができるようになり, 知能研究の方向性を見いだすフロンティアがついに出現しつつあるというものである. そうしたフロンティアに関わるさまざまなテーマを扱う記事を最先端の脳研究者にご寄稿いただくことで, 新しい研究領域の将来性を感じてもらうことにある.

2. 基本構図: AI の目指すものと脳の関係

2.1 ニューラルネットワークの意義

AI 一般においては, 直接的に神経科学の知見を活用する Brain-inspired Computing を行うことは難しい. そこで AI と脳科学との関係を図 1 の「AI-ニューラルネットワーク-神経科学の進展図」から見てみる. すると, AI や認知科学と神経科学の間にはニューラルネットワークという分野が存在し, それが橋渡しを担うことで, AI は脳から学び得る.

こうした代表事例としては, 神経細胞を単純な非線形関数としてモデル化したマカロック・ピッツモデル [McCulloch 43] をベースとした人工ニューラルネットワークがある. そして, 深層学習の起点となったネオコグニトロン (1979 年発表) は脳の初期視覚野の情報処

AI-ニューラルネットワーク-神経科学の進展図

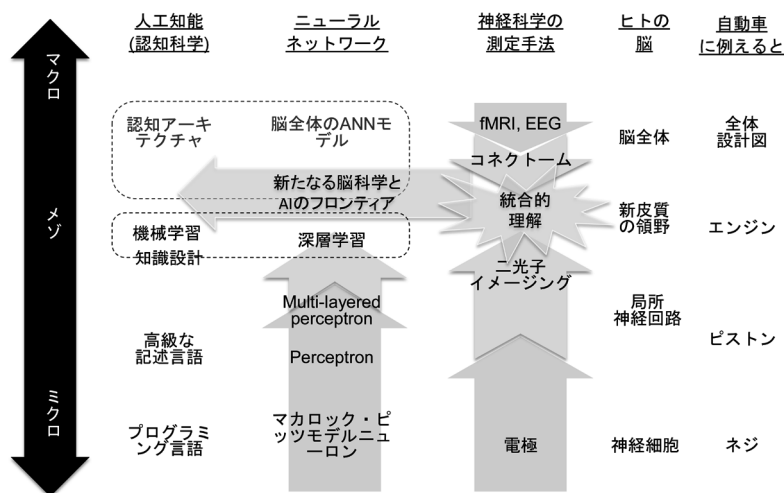


図 1 AI-ニューラルネットワーク-神経科学の進展図

理にヒントを得ている。ネオコグニトロンはニューラルネットワークという、脳の神経回路に近い技術をベースにつくられているからこそ、視覚情報処理のアーキテクチャから学ぶことができ、現在においては深層学習という形で一般物体認識などの機能を実現している。

つまりAIが脳に学ぶためには、脳に似せてAIをつくるというニューラルネットワーク技術の存在が必要となっている。

2.2 ミクロとマクロ

AI研究は、主にヒトの振舞いにひそむ知能を工学的に再現することを目的としている。脳に学んだAIをつくるために、図1の右側に記した脳のミクロからマクロにわたるスケール内のどの範囲についての理解が必要であろうか。

最もマクロなスケールは、一人のヒト全体、つまり脳の全体と身体を含んだものになるであろう。さらに全体を設計する粒度では、その直下のメゾスコピックレベルの設計図に基づいて組み立てる必要がある(車に例えれば、エンジンなどの機能部品を組み合わせる粒度)。しかしさらにミクロなレベルを考えるなら、神経細胞内の代謝システムまで及ぶかもしれない。いずれにしても、将来において脳の仕組み、つまり、脳内のネットワーク構造と神経活動、およびその学習規則などのすべてが神経科学で解明されていれば、それを基づいた形でのシステムを再構築すればヒトのような知能をつくれる見込みは高い。

脳構造の階層性を踏まえようとして、どの程度に詳細な粒度からボトムアップにモデル化すれば知能を構築できるかは自明ではない。現在の工学的なニューラルネットワークは、先に述べたマカロック・ピッツモデルのように空間的に広がりをもたない非常に単純なニューロンモデルである。果たしてそれで十分であろうか。こうした疑問は常に付きまとう。

今世紀に入って神経科学がメゾスコピックレベルに進出したことで、AIへ影響を与える形でフロンティアが生じてきた。

3. 20世紀までの脳科学とAIの限界

前世紀においては、以下のような実情によりAIが脳に学ぶには限界があった。

3.1 ミクロ-マクロギャップの存在

20世紀の神経科学におけるミクロな神経活動の測定手法は、図1に示すように、主に脳に挿入された電極などによる測定であり、得られる情報の範囲は極めて限定されていた。脳全体の活動を捉える手段としては、以前は脳波計のみしかなかったが、20世紀の後半にはfMRIが出現し、不十分なが脳全体にわたるおおざっぱな活動は捉えられるようになった。

いずれにしても、20世紀の神経科学においては、脳のごく一部を詳細に測定するミクロな研究と、脳の全体

をおおざっぱに捉えるマクロな研究とが乖離しており、その間をつなぐメゾスコピックな領域が埋められていなかった。こうした状況では、ミクロレベルからマクロレベルにわたって、一貫したヒントを脳から得ることは難しかった。

つまり20世紀の神経科学はミクロな理解とマクロな理解の間にギャップが存在しており、AI研究が求める脳の機能的理解に役立てるにははなはだ力不足であった。

3.2 脳機能を再現するためのNNモデルが不足していた

脳を起点として、ヒトの認識行動レベルを理解するためには、脳における新皮質の領野や、基底核、海馬、小脳などといった各脳器官について計算機能を含めてモデル化する必要がある。

これまでの研究は、工学的に実現されている技術が脳内にも存在することが示される研究事例が多かった。これら研究は主に脳を理解することを目的としており、計算論的神経科学と呼ばれる。例えば、図1に示すように、小脳の機能をフィードフォワード制御モデルとしてパーセプトロンで実装したり、大脳基底核の機能を強化学習の遅延報酬計算を行うニューラルネットワークとして実装したりといったケースである。しかしながら人間において特に発達し、知能の汎用性にも関わる大脳新皮質のモデル化は十分に進んでいなかった。

4. 今世紀における進展

4.1 ミクロ-マクロギャップを埋めつつある神経科学

今世紀に入り、神経科学は大きな進展を遂げ、時間的・空間的なミクロ-マクロギャップを埋める新しい技術が開発されつつある。一つには光遺伝学や二光子顕微鏡の進展により、実験動物が何らかのタスクを行っている最中に、ある脳部位における1000個程度の神経の活動を同時に測定できるようになった。また、複数の脳部位において神経活動の測定を同時に行うような研究も盛んになってきている。

また脳全体のおおざっぱな接続構造を表すメゾスコピックレベルのコネクトームが¹⁾齧歯類を中心に明らかにされており、さらに精緻化が進んでいる。こうした情報に基づけば、これを参照アーキテクチャとみなして脳型AIを設計する際のガイドとして用いることが可能になってきている。

4.2 深層学習：マクロ化したニューラルネットワーク

今世紀に入り、脳を模することができる計算モデルの進展が著しい。つまり深層学習の発展により、人工ニューラルネットワークとして実装できる機能が、物体認識などといった特定の機能をもつ部品レベルに到達した。すでに、畳込みニューラルネットワーク(CNN)による一般物体認識は視覚の側頭葉経路と対応付けられている[Yamins 16]。ちなみに、これまではAIと呼ばれなかったニューラルネットワークが、第三次AIブームにおい

て、AIと呼ばれ始めたのは、それが実現できる機能が画像認識や音声認識といった特定の機能レベルまでマクロ化したことによるものと思われる。

さて本稿の観点からは、深層学習の発展がもたらしたインパクトには二つある。

一つ目は、深層学習は、実用的に動作する計算機能をもつ新皮質モデルの初の候補となったことである。こうして大脳基底核、小脳に続いて新皮質についてもAIに組み込める部品がようやく現れたことになる。

二つ目に、現在の深層学習の進展により、十分にデータが得られるタスクや状況に限定すれば、多くの認知機能が人工ニューラルネットワークによって実装できることがわかってきた。となれば、ニューラルネットワークを用いて脳のようなAIを実現しようとする場合に、多くの脳器官においてその基本単位はマカロックピッツのニューロンモデルで十分であるという可能性が格段に高まったといえる。

5. 歴史上初めて現れた脳科学とAIのフロンティア

AIが実現したい能力はもともと、認知行動機能レベルの知能であった。神経科学では今世紀に入ってマクロとミクロの両方から進展があり、メゾスコピックレベルにおいてギャップを埋めるための知見が急増した。そこで得られた知識は、認知機能を実現できる程度に大規模化したニューラルネットワークと相性が良い。このような技術進展を背景に、今ようやく神経科学の知見をAI構築に役立てる形でのフロンティアが生じてきた。

つまり、神経科学とニューラルネットワークの二つの分野がメゾスコピックレベルで進展することにより、AIの統合的理解に役立つ状況が初めて生まれつつある。この流れがAIに生み出し得るメリットとして想定されるのは以下に述べるアーキテクチャ化と理論化の二つである。

アーキテクチャ化とは、機能モジュールを組み合わせ、それらを連携するアーキテクチャ上でさまざまな認知行動機能が実現される。これが脳における複数の領域での脳活動データの連携として解釈される。また近年では、各モジュールを微分可能とすることで、多数のモジュールを連携させた学習を可能とすることが多く試されている。またアーキテクチャ導入における典型例は、具象から抽象への階層性などである。さらに、システム内のある部分の神経活動や計算処理を、一步離れた視点から捉えるメタレベルからの認識、制御、学習などが起こってくる。こうしたアーキテクチャの視点は、知能の汎用性実現に関わり、その実現を目指す全脳アーキテクチャ・アプローチにおいても重要な観点である。

脳が高度に複雑なシステムであることを踏まえれば必ずしも、比較的簡単な理論によって知能が記述されるには限らない。しかしながらヒトが知能の本質を理解するためには、できるだけ単純な理論によって知能を定式化したい。こうした試みとしては、強化学習の枠組みを拡

張した万能AIや、行動によって選択肢の多様性を拡大しようとする知能の方程式、などがある。また[Friston 06]らは、脳との関係性を議論しつつ、一種の自由エネルギーのような量の最小化として知能を捉えようとしている。一様な構造をもつ新皮質が多様な機能を獲得する能力に基づいて知能の汎用性にアプローチする方向性も多い[Numenta 10, 山川 17]。

6. 各記事の位置付け

こうして、AIとして興味の高い比較的マクロな領域で脳科学とニューラルネットワークの融合によるフロンティアが出現し、AIの現状の限界を突破するために脳を参考とし得る状況が生まれてきた。

そこで私達は、今後のAIの発展に向けて、深層学習と組み合わせ、深層学習を改良し、深層学習を超えるために脳から何を学ぶべきかを探ることを目的として本特集を企画し、著者の皆様には、進展著しい神経科学の知見がどのようにAIの発展に寄与し得るかのヒントを与えていただけるようお願いした。

以下ではこうした観点から本特集の各記事を位置付けたい。

甘利：もうちょっとだよなー、ディープラーニング

まず初めに、長年にわたって数理脳科学を牽引し続けてこられた甘利俊一先生に、近年の深層学習の進展を中心に据え、過去・現在・未来をどのように見渡しておられるのか解説していただいた。現在の深層学習の根幹を支える確率勾配降下学習・自然勾配学習の起源から、現在進行中の理論の解説、さらには人類の未来までも見渡した原稿は、多くの人に深い示唆を与えるに違いない。

なお、標題がいささかふざけているのは、担当委員からの依頼を反映したものであることを申し添えておく。

岡本：全脳ネットワーク分析—コネクトームのリバースエンジニアリング—

脳の配線図の解明、すなわちコネクトームの進展は、先に述べたように近年の神経科学の大きな成果である。特にメゾスコピックなレベルのコネクトームは、ヒトのような知能を実現するための大きなヒントとして期待される。この記事では近年のコネクトーム解析について基礎から詳しく解説した。

田中・坂上：推移的推論の脳メカニズム—汎用人工知能の計算理論構築を目指して—

ヒトが得意とする学習の般化は、現在のAIがまだまだ不得意とする領域である。私達が報酬に関連する情報をどのように抽象化して学習しているのか、計算理論の進展を解説していただいた。

佐藤：ヒト海馬の神経回路モデルの構築

知的な行動を実現するためには、過去の経験を抽

象化して蓄えて取り出せるようにしておく機能, すなわち記憶が必要不可欠である。これに深く関わる脳器官である海馬の計算モデルについて解説していただいた。ノーベル医学・生理学賞が記憶に新しいように, ナビゲーション機能などの多様な機能がどのように関わって実現されているのか, 理解を深める助けとなるだろう。

西田・西本: 意味認知と脳内情報表現

脳に入ってくるさまざまな情報が行動に結び付くまでの間には, さまざまな記号とそれにひも付いた意味処理が介在していることを私達は知っている。このメソスコピックなレベルの情報が脳内でどのように表現されているのかを近年明らかにしつつある, データ駆動型の解析について解説していただいた。

中原: 社会知性を実現する脳計算システムの解明: 人工知能の実現に向けて

ヒトの知性の深みは, 複数個体間の相互作用, すなわち社会性を考えたときに最も顕わになる。そのために必要と考えられている他者を適切に抽象化して自己の中に捉えるシステムについて, 計算論・強化学習の基本的な考え方から丁寧に解説していただいた。

大澤: 知能への学問の分解と統合—人工知能+認知科学+神経科学異分野交流会—

最後に, 脳とAIの接点において活発な若手の動きが生まれつつあることを紹介するために, 「人工知能+認知科学+神経科学異分野交流会」の報告記事を全脳アーキテクチャ若手の会元代表にお願いした。人工知能学会だけでなく, 日本神経回路学会・日本認知科学会・脳科学若手の会などと広く関わる学問領域での議論の雰囲気を感じてもらいたい。

7. おわりに

私達は, 脳科学の知見とAIの知見の橋渡しになるような境界領域のテーマを扱うオーガナイズドセッション「脳科学とAI」を, 2011年の全国大会より継続してきており, 近年は毎回10件以上の投稿が続いている[森川17]。特にOSを開始してからの6年間の間に, 以上に述べたような技術的背景から神経科学の進展と人工ニューラルネットワークが織りなすフロンティアがAI研究との接点を拡大している。こうした流れは「第1回次世代脳型人工知能研究会—Beyond Deep Learning」にもつながってきている。

脳を参考としながらAIを構築することは, 神経科学への理解や, ニューラルネットワークとしてAIを実装するなどの制約があり必ずしも簡単ではない。しかしながら本特集を通じて, 神経科学とニューラルネットワークの進展で生じつつあるAI研究の新たなフロンティアについて知っていただき, 皆様ご自身の研究などに役立てていただければ幸いである。

◇ 参考文献 ◇

- [Friston 06] Friston, K., Kilner, J. and Harrison, L.: A free energy principle for the brain, *J. Physiol.*, Vol. 100, No. 1-3, pp. 70-87 (2006)
- [McCulloch 43] McCulloch, S. W. and Pitts, W.: A logical calculus of the ideas immanent in nervous activity, *The Bulletin of Mathematical Biophysics*, Vol. 5, No. 4, pp. 115-133, Kluwer Academic Publishers (1943)
- [森川 17] 森川幸治, 岡本洋, 山川宏: 脳科学とAI, 特集「2017年度人工知能学会全国大会(第31回)」, 人工知能, Vol. 32, No. 6, pp. 911-914 (2017)
- [Numenta 10] Numenta, Inc., ed.: *Hierarchical Temporal Memory Including HTM Cortical Learning Algorithms* (2010)
- [Wang 12] Wang, P.: Theories of artificial intelligence—meta-theoretical considerations, *Theoretical Foundations of Artificial General Intelligence*, Wang, P., Goertzel, B., eds., *Atlantis Thinking Machines*, Vol. 4, Chapter 16, Atlantis Press (2012)
- [山川 17] 山川宏, 荒川直哉, 高橋恒一: 全脳アーキテクチャに必要な新皮質マスターアルゴリズムの検討, 2017年度人工知能学会全国大会(第31回)予稿集, 3K1-OS-06a-5 (2017)
- [Yamins 16] Yamins, D. and DiCarlo, J. J.: Using goal-driven deep learning models to understand sensory cortex, *Nature Neuroscience*, Vol. 19, pp. 356-365 (2016)