

脳と生成 AI を巡って

合原一幸

●東京大学特別教授／名誉教授

1 はじめに

筆者は生成 AI に関して特に詳しいわけではないが、本特集は友人の坂上貴之さんが関わられている特集だと編集部からお聞きしたことで、筆者の周りには生成 AI の研究者が割といて門前の小僧の経験は豊富なので、専門外ではあるが本エッセイをお引き受けすることにした。生成 AI の進歩はとても早くていろいろな知識はすぐに陳腐化するため難しいテーマであるが、今思っていることを書き留めておきたい。

画像生成 AI との出会い自体は比較的早かった。2018 年の夏頃だったと思うが、以前分岐図のパターンを使ったドレスで東京コレクションのファッションショーと一緒にやったデザイナーのエマ理永さん、ジャーナリストの永山悦子さんと議論していて、AI を利用したドレスができないだろうか、という話題になった。早速、理研 AIP の杉山将さん、上田修功さんにご相談したら、快く協力してくれることになった。特に清田純さんたちの頑張りで、画像生成 AI の元祖のひとつである GAN (Generative Adversarial Networks) にエマさんが過去に作った約 500 着のドレス画像などを学習させて、エマ理永風ではあるが過去のどのドレスとも違うドレス画像を無数に生成する GAN が開発できた。次に、この GAN の潜在空間をエ

マさんに探索してもらったところ、1% 程度は面白いドレス画像が出てくるということだったので、それらから発想を広げて実際のドレスを新しく創作していただいた。このドレスを用いて 2019 年春の東京コレクションでファッションショーをやってみたところ、幸い望外の好評を博した。学習データを提供したデザイナー本人にとっても面白いドレスパターンが GAN によって生成され得る点が筆者には興味深く、美や創造性に関係する分野においてすらヒトと AI の協働がこういうやり方なら可能かなと感じた。

AI の進歩によって、「ラマヌジャン・マシン」[1]など、数学との接点も増えてきている[2, 3]。最近の生成 AI は画像や音声はもちろんだが、特に(自然)言語を巧妙に扱えるのが驚くべき点である。数学も一種の言語なので、数学のためにうまくカスタマイズした特化型生成 AI を作ることは自然な構想で、研究開発も進んでいるようだ。もちろん、生成 AI には通常確率的な処理が入るので、1 か所でも間違えると全体が間違いとなる数学への応用はそのままでは容易ではない。したがって、高度なプロンプト工学、数学的厳密さを強化するための工夫や生成された出力の注意深い数学的検証などが不可欠となるため、かなり苦勞しそうな気がするが、(本エッセイではこれ以上論じないけれど)今後さらに研究開発が進むだろう。

2 人工知能と人工ニューラルネットワーク

よく知られているように、AI 研究は過去に 2 回のブームとその後のいわゆる冬の時代を経験しており、今回の第 3 次ブームでようやく AI の本格的な社会実装が実現した。他方で、こちらは AI ほど知られていないが(ちょっと悲しい)、人工ニューラルネットワークの研究が歴史的に AI 研究と微妙な関係を持ちながらも、ほぼ独立して行われてきた。この分野は脳や神経細胞の数理モデルを作って研究をする。目的は AI と似たところがあるが、脳自体を理解したいという点がより強い研究のインセンティブになっている。この分野も過去に 2 回のブームと冬の時代を経験しているが、現在の AI を支えている深層学習は超多層の人工ニューラルネットワークを使うので、第 3 次ブームでようやく AI 研究と人工ニューラルネットワーク研究がある程度融合された自然なカタチとなった[4]。

人工ニューラルネットワークの源流となる基本構成要素は、マカロックとピッツによる形式ニューロンモデルで、次式で表される[5]：

$$x(t) = f\left[\sum_{i=1}^M w_i s_i(t) - \theta\right]$$

ここで、時刻 $t \in \mathbf{N}$ 、 $x(t)$ と $s_i(t)$ は時刻 t における出力と i 番目の入力の値で $x(t), s_i(t) \in \{0, 1\}$ 、1 は発火(興奮)状態、0 は非発火(静止)状態、 M は全入力数、 w_i は i 番目の入力のシナプス係数(興奮性シナプスは正值、抑制性シナプスは負値)、 θ はしきい値である。

f は出力関数で、形式ニューロンモデルの場合、次式のヘビサイド関数を用いる：

$$f(u) = \begin{cases} 1, & u \geq 0 \text{ のとき,} \\ 0, & u < 0 \text{ のとき.} \end{cases}$$

形式ニューロンは、時間と入出力がともに離散値の論理関数のようなモデルである。実際、論理

関数としての万能性があり、任意の論理関数が、十分たくさんの形式ニューロンから成る人工ニューラルネットワークで実現できる。

ところが深刻な難点があって、出力関数の微分値が $u \neq 0$ で 0 になるので、多層ニューラルネットワークの中間層のニューロンの学習がうまくいかない。この問題点をいち早く見抜いたのが甘利俊一先生で、微分可能な単調増大出力関数を持つアナログニューロンを用いて中間層の学習を可能とする理論を世界にさきがけて提案された[6, 7]。

アナログニューロンの出力関数としては、次のシグモイド関数

$$f(u) = \frac{1}{1 + \exp(-u/\varepsilon)}$$

が長年使われてきた(ε は増大特性の急峻さを決める正值パラメータ)が、最近ではより単純化した次式の ReLU (Rectified Linear Unit) 関数などが広く使われている：

$$f(u) = \max(0, u).$$

3 生物ニューロンのスパイクングダイナミクス

それにしても、膨大な数の上記のような単純な人工ニューロンモデルを用いた超多層ニューラルネットワークに、膨大な電力を費やした計算資源と学習データを注ぎ込めば、高機能の生成 AI が構築できてしまうとは実に予想を越えた成果であるが、同時に生成 AI は脳とはあまりにかけ離れている感がある。この点を、少し論じてみよう。

前節で、発火状態を形式ニューロンでは 1、アナログニューロンでは正の実数値(発火頻度に対応)で表したが、実際の生物ニューロンでは、活動電位と呼ばれる電気スパイクの生成が、この発火状態表現の基盤となっている。20 世紀前半の神経科学では、この活動電位の発生機構を理解することが最重要課題のひとつであった。そして、この課題は、ケンブリッジ大学の E. D. エイドリア

ン、A.L. ホジキン、A.F. ハクスレイという3名のノーベル賞学者によって見事に解明された。特に、ホジキンとハクスレイの仕事は数理的でもあり、ヤリイカ巨大軸索膜における電気スパイク生成過程を、以下のホジキン-ハクスレイ方程式として定量的に定式化してノーベル生理学医学賞を受賞した[8]：

$$\begin{aligned} C \frac{dV}{dt} &= I - 120.0 \, m^3 h (V - 115.0) \\ &\quad - 40.0 \, n^4 (V + 12.0) - 0.24 (V - 10.613), \\ \frac{dm}{dt} &= \frac{0.1(25 - V)}{\exp\left(\frac{25 - V}{10}\right) - 1} (1 - m) \\ &\quad - 4 \exp\left(\frac{-V}{18}\right) m, \\ \frac{dh}{dt} &= 0.07 \exp\left(\frac{-V}{20}\right) (1 - h) \\ &\quad - \frac{1}{\exp\left(\frac{30 - V}{10}\right) + 1} h, \\ \frac{dn}{dt} &= \frac{0.01(10 - V)}{\exp\left(\frac{10 - V}{10}\right) - 1} (1 - n) \\ &\quad - 0.125 \exp\left(\frac{-V}{80}\right) n. \end{aligned}$$

ここで、 V は神経膜電位、 m, h, n は神経膜のナトリウムコンダクタンスとカリウムコンダクタンスの時間変化を記述する現象論的変数である。ニューラルネットワークを作るときは、神経膜電流 I に他のニューロン群からの入力を導入すればよい。

筆者の実験研究の恩師である松本元先生は、この神経膜の発火(興奮)現象が、外部入力によって引き起こされる、平衡構造(静止状態)から散逸構造(発火状態)への転移であるという説を、非線形非平衡物理学の観点から提唱された[9]。そして、この転移の背景に、静止状態を表す平衡点におけるサブクリティカル・ホップ分岐が存在することが、ヤリイカ巨大軸索膜における実験研究とホジキン-ハクスレイ方程式やその修正数理モデルの

理論研究であきらかとなった[10]。ここで分岐とは、システムのパラメータ値の変化に伴って、解の構造が定性的に変化する現象である。

A.L. ホジキンは実験神経科学者であるが、なんと上記の数理構造に関連した神経応答特性について、とても早い時期に深く探究している。素晴らしい慧眼である。彼は、1948年の論文[11]で、神経膜を定常電流で刺激した際に電流を次第に大きくしていくと反復的に発火するようになるが、その特性を調べると、0に近い低周波数で発火が始まるクラスⅠと不連続的に高周波数で突然発火が始まるクラスⅡの2種類に主として分けられるという実験結果を発表している。そして、はるか後年の数理研究で、各々のクラスは、典型的にはサドルノード・オン・インバリアントサークル分岐とホップ分岐で特徴づけられることが示された[12]。

実際の生物の脳を見ると、クラスⅠは興奮性ピラミダルニューロン、クラスⅡは抑制性介在ニューロンによくみられるので、このような分岐構造が脳における異なる種類のニューロンの働きと関係している可能性が高い。また、ほかにもピッチフォーク分岐を伴うニューロンモデルもあり、このタイプのニューロンから成るネットワークは、たけすえ武居弘樹さん、稲垣卓弘さん、稲葉謙介さんたちによって、光電子ニューラルネットワークとして実機が構築され、その非線形ダイナミクスによる情報処理能力が解析されている[13, 14]。

これら一連の研究をはじめとした神経発火過程を考慮する数理モデルは、活動電位の電気スパイクを陽に取り扱うので、スパイクングニューラルネットワークと呼ばれる。これらのモデルも前節で述べた人工ニューロンや人工ニューラルネットワークの一種とも言えるが、ヒトを含む生物のニューロンやニューラルネットワークにより近いので、神経科学研究との関係がもっと深くなる。

このように生物ニューロンや関連数理モデルを

掘り下げてみると、非線形ダイナミクスが重要な役割を果たしているように思えてくる。生物知能と生成 AI のような人工知能は、系としてどちらも複雑系の典型例であるが、基本構成要素と全体のネットワーク構造はあまりに大きく異なっている。この違いをどう見ればいいのかだろうか？

特に気になるのは、基本構成要素の複雑さの大きな違いである。こう感じるのは、筆者が松本先生の下で生物ニューロンの動的複雑さを若いときに見てしまったからかもしれない。個人的には、なんとかこの差を埋められないかと人工ニューロンにカオスダイナミクスを導入したカオスニューラルネットワークを数理的に研究してきた[15, 16]。また、堀尾喜彦さん、河野崇さん、森江隆さんたちが、高性能のカオスニューロチップを作ってくれたので実装も可能である[17]。さらに、金子邦彦さんや津田一郎さんたちによって、カオス要素からなる複雑系の多彩な機能が示されているし[18]、ReLU 関数を使ってカオスニューロンモデルを作るのも容易である[19]上に、最近のカオスとリザーバー AI 計算が多面的に結びついてきた[20-22]ので、今後の大きな発展が期待される。

ただし、ホジキン-ハクスレイ方程式のような詳細数理モデルやカオスニューロンを生成 AI のニューロン数と同じくらい膨大に使って、脳の構造も考慮してネットワークを作っても、本当の脳のような高次機能が再現できるとはとても思えない。ヒトを含む生物脳の数理モデル化は、この分野にとって依然としてはるかかなたの夢なのである。

4 生物知能と人工知能の協働

上記の意味で、生成 AI は脳とはかけ離れているが、他方で深層学習以降の AI の進歩には本当に目を見張るものがある。特に一般の人々に強い印象を与えたのは、囲碁の AI があまりにも強くなって、プロ棋士も勝てなくなったことであろう。

ところで、AI が圧倒的に強くなった通常の囲碁とは異なる、滝久雄さんが考案したペア碁という面白い囲碁がある。2 組の男女ペアがひとつの碁盤を使って対局するもので、自分のパートナーとも相談をしてはいけないのがポイントである。すなわち、パートナーの打つ手を見て、その心を読まないといけないので、ペア間の心のコミュニケーションが鍵となる。2017 年に渋谷でペア碁世界選手権が開催された折に、世界トップ棋士^かの柯潔さんや井山裕太さんたちにご協力いただいて、囲碁 AI とプロ棋士のペア同士で対局していただいたことがある。まあ、試しにやってみた程度だったが、結構盛り上がり、その後テレビの囲碁番組でも同様の企画が行われたりしているので、それなりのインパクトはあったようだ。

脳自体を数理的に再現する AI を創ることは少なくとも当面はとても無理だと思うので、現在考えるべき問題は、ヒトと AI の協働の在り方を探ることだろう。いろいろな協働様式があり得るが、まず考えるべきは、冒頭のファッションショーの例のように、AI をヒトのツールとしていかに使いこなすかという問題である。AI にとっては一見やさしそうだが実際にはとても難しかったり、逆に囲碁のように AI には難しいと思われていたがやってみると意外にうまくいったりする例が多数あるので、このレベルでも今後産業界を筆頭にさまざまな研究開発が行われることが期待されるし、大きなビジネスチャンスが生まれるだろう。

その先にあって前向きな夢を感じるのは、AI とヒトが互いの持ち味を活かし補完し合うパートナーとして協力する可能性である。上述のペア碁の例も、この問題を考えたくて企画したものである[4]。たとえば、囲碁 AI と筆者のペアが、柯・井山ペアに勝つようなことが起きるだろうか？ これは、たぶん今の AI の圧倒的強さなら、悪手を連発するいわば内なる敵の筆者の存在にもかかわらず、実質的には AI 対筆者を含めた 3 人との

勝負となって結果的に AI・筆者ペアが勝ちそう
だ。したがって、AI とヒトとの対等に近い協働
の仕組みを研究するためには、もっと適切な研究
プラットフォーム設計が必要である。AI が心を
持つのは困難でも、ヒトの心のある程度理解でき
るとその応用はもっと大きく広がるように思える。
さらに、複数のヒトと複数の AI の集合知を考え
る基盤ともなり得よう。

東京大学国際高等研究所ニューロインテリジェ
ンス国際研究機構では、「人の知性はどのように
して生じるか？」という究極の難問をターゲット
に、特に脳の臨界期、自発神経活動、社会学習な
どの観点から、実験、臨床、数理を統合するチー
ムサイエンスが進められている。生物の発達過程
の知見が今の AI には欠けているし、親子を典型
例とするヒト同士の相互作用による社会学習は、
上記のヒトと AI との多様な協働を実現するため
にも、本質的な知見を与えてくれるものと思われ
る。

5 おわりに

なんとなく焦点が絞れないエッセイになってし
まったが、脳と生成 AI とを巡って今考えている
ことを思いつくままに書いてみた。もちろん、ま
だまだ考えるべき論点は多く残っている。最後に
一つだけ挙げるとすれば、電力消費量の問題であ
る。生成 AI の開発や巨大データの維持はもとよ
り、今後の膨大な AI の社会実装を考えると、そ
のために必要な電力量の問題が実はとても深刻に
なる。おそらく将来のエネルギー問題や省エネを
考える上で、避けては通れない切実な課題になる
だろう。

この点でも、脳と AI の対比はとても参考にな
る気がする。ヒトの脳は約 1000 億のニューロン
から成ると推定されているが、わずか 20 W で動
いている。同サイズの人工ニューラルネットワー
クをデジタルシミュレーションすると 7.9 MW 必

要だという[23]。この意味で、3 節で述べたス
イキングニューラルネットワークやリザバー計算
が有する低消費電力性が今後大いに注目されると
思われるが、行数も尽きたのでこの問題は他の機
会に譲りたい。いずれにしろ、生成 AI の予想も
しなかった急激な進歩により、脳と AI を巡る議
論も今後より深くかつ新しい展開を見せることが
期待できよう。

●参考文献

- [1] G. Raayoni et al., Generating conjectures on funda-
mental constants with the Ramanujan Machine,
Nature, 590, 67-73, 2021.
- [2] A. Davies et al., Advancing mathematics by guiding
human intuition with AI, *Nature*, 600, 70-74, 2021.
- [3] T. H. Trinh et al., Solving olympiad geometry without
human demonstrations, *Nature*, 625, 476-482, 2024.
- [4] 合原一幸編著『人工知能はこうして創られる』, ウエ
ッジ, 2017.
- [5] W. S. McCulloch and W. H. Pitts, A Logical Calculus of
the Ideas Immanent in Neural Nets, *Bull. Math.
Biophys.*, 5, 115-133, 1943.
- [6] S. Amari, A Theory of Adaptive Pattern Classifiers,
IEEE Trans. Electr. Comp., 3, 299-307, 1967.
- [7] 甘利俊一『深層学習と統計神経力学』, サイエンス社,
2023.
- [8] A. L. Hodgkin and A. F. Huxley, A Quantitative
Description of Membrane Current and its Application
to Conduction and Excitation in Nerve, *J. Physiol.*,
117, 500-544, 1952.
- [9] 松本元『神経興奮の現象と実体(上)』, 丸善, 1981.
- [10] K. Aihara and G. Matsumoto, Temporally Coherent
Organization and Instabilities in Squid Giant Axons, *J.
Theor. Biol.*, 95, 697-720, 1982.
- [11] A. L. Hodgkin, The Local Electric Changes Associat-
ed with Repetitive Action in a Non-medullated Axon,
J. Physiol., 107, 165-181, 1948.
- [12] E. M. Izhikevich, *Dynamical Systems in Neuroscience:
The Geometry of Excitability and Bursting*, The MIT
Press, 2007.
- [13] T. Inagaki, et al., A Coherent Ising Machine for 2000-
node Optimization Problems, *Science*, 354 (6312),
603-606, 2016.
- [14] T. Inagaki et al., Collective and Synchronous Dynam-
ics of Photonic Spiking Neurons, *Nature Comm.*, 12,
2325, 1-8, 2021.

- [15] K. Aihara et al., Chaotic Neural Networks, *Phys. Lett. A*, 144(6/7), 333-340, 1990.
- [16] K. Aihara, Chaos Engineering and its Application to Parallel Distributed Processing with Chaotic Neural Networks, *Proc. of the IEEE*, 90(5), 919-930, 2002.
- [17] たとえば, Y. Horio, K. Aihara and O. Yamamoto, Neuron-Synapse IC Chip-Set for Large-Scale Chaotic Neural Networks, *IEEE Trans. Neur. Netw.*, 14(5), 1393-1404, 2003.
- [18] 金子邦彦, 津田一郎, 『複雑系のカオス的シナリオ』, 朝倉書店, 1996.
- [19] H. Tamura et al., ReLU Chaotic Neuron Model for Deep Chaotic Neural Networks, *SEISAN-KENKYU*, 70(3), 45-47, 2018.
- [20] T. Kanamaru, T. K. Hensch and K. Aihara, Maximal Memory Capacity Near the Edge of Chaos in Balanced Cortical E-I Networks, *Neural Computation*, 35(8), 1430-1462, 2023.
- [21] Y. Sakemi et al., Learning Reservoir Dynamics with Temporal Self-Modulation, *Com. Phys.*, 7(29), 1-11, 2024.
- [22] Y. Ebato, et al., Impact of Time-History Terms on Reservoir Dynamics and Prediction Accuracy in Echostate Networks, *Sci. Rep.*, 14(8631), 1-13, 2024.
- [23] A. Mehonic and A. J. Kenyon, Brain-Inspired Computing Needs a Master Plan, *Nature*, 604, 255-260, 2022.

[あいはら かずゆき]

例と演習で学ぶ 確率論

中島 誠・著

A5・432 頁・定価：4,400 円（税込）
ISBN 978-4-06-535634-0

測度論を用いた確率論の新定番となる入門書。確率論の王道をしっかり押さえた丁寧な解説と豊富な例題・演習問題が初学者にぴったり。序章で確率論の具体的なモデルも紹介し、確率論を「体験」できる構成。



新刊

暗号の理論と技術

國廣 昇・編著 量子時代のセキュリティ理解のために
安田雅哉 / 水木敬明 / 高安 敦 / 高島克幸 /
米山一樹 / 大原一真 / 江村恵太・著

A5・336 頁・定価：3,740 円（税込）
ISBN 978-4-06-535635-7

次世代の暗号技術の標準化に向け、暗号理論とその技術を学べる入門書。暗号理論の数学的基礎から応用技術まで、幅広い内容をバランスよく学べる。暗号理論を学ぶ学生や研究者、セキュリティ関係の技術者に最適。



新刊

面倒なことは ChatGPT にやらせよう

カレーちゃん / からあげ・著

A5・320 頁・定価：2,750 円（税込）
ISBN 978-4-06-534290-9

ChatGPTでExcel・PowerPointやデータ分析、画像生成まで！かんたん雑用丸投げ術で仕事が楽になる！大好評で累計4万部突破です。



東京都文京区音羽 2-12-21
<https://www.kspub.co.jp/>



KODANSHA



編集
販売

☎03(3235)3701
☎03(5395)4415

合原一幸●あいはら・かずゆき
1954年、北九州市生まれ。東京大学特別教授／名誉教授。ヒトを含む生物のダイナミズムとそれを生み出すダイナミクスに興味がある。昨年末、高校入試目の孫娘のにわか家庭教師という大役を急遽任され、それ以来無事務めをはたし終えた今でも、毎朝数学関連のYouTubeを見るのが日課となりました。素晴らしい動画を作られている先生方に感謝いたします。ただ、孫娘にギターのFコードの弾き方を教えた時の方が、数学を教えた時よりもっとうれしそだったのは、まあしょうがないですかね。