

ヒューマン・インテリジェンスと ビッグデータ型及び脳神経模倣型 AI： インテリジェンスとは何かを考える¹⁾

中馬宏之

目次

1. はじめに
2. BD-AI と NM-AI の違いを検討する：機能特性という視点から
 - 2-1 BD-AI と NM-AI が想定するニューロンの違い：時空間概念の有無
 - 2-2 BD-AI の実用性 vs. NM-AI の低消費電力性
 - 2-3 NM-AI 実用化を阻むもう一つのボトルネック：大容量“連想メモリ”という難題
 - 2-4 長期的には、BD-AI に加えて NM-AI が必須：マクロの視点
3. BD-AI 及び NM-AI の“インテリジェンス”とは？
4. ヒューマン・インテリジェンス (HI) と BD-AI・NM-AI との違いを探る
5. 結びに代えて

1. はじめに

現行の AI (Artificial Intelligence) には、大きく 2 つのタイプがある。実用化に富み今をときめくビッグデータ型 AI (以下 BD-AI と呼ぶ) と明日の AI を担うとされるがなかなか実用化までに至っていない脳神経/ニューロン模倣型 AI (Neuromorphic AI: 以後 NM-AI と呼ぶ) である²⁾。そして、両者のニューロン模倣の程度には大きな差 (後述) があり、アーキテクチャー (設計思想) も根本的に

1) 本論は、成城大学経済研究所年報用論文として、中馬他 (2018) に大幅な加筆・訂正を加えたものである。

2) このような評価は、最近 Intel Labs 所長に就任した Rich Uhlig にも見られる (IEEE Spectrum 誌上における Ackerman (2018) との対談)

異なっている。なお、本論では、AI という言葉の定義上の混乱を避けるために、世界的に著名な AI 研究者／哲学者である Slowman (1995) に習って、広義に下記の“自己変化する情報駆動型制御システム”と見なしたい。

「AI は人間と動物のインテリジェンスを研究範囲としているので、AI というのは誤った名称だ。私は、それを洗練された自己変化する情報駆動型制御システム (sophisticated self-modifying information-driven control systems) に関する一般研究と理解している。そこには、生命体と人工物、実際のもの と可能性のあるもの (進化してきたに違いないものと造られたもの) とを含む。」(筆者訳)

ただし、BD-AI, NM-AI の両者が志すところは、共に脳機能の本質把握に基づく AI/ML (Machine Learning (機械学習)) の実用化であることに変わらない。実際、BD-AI, NM-AI 共に、McClulloch-Pitts 型ニューロンと呼ばれる人工ニューロンで名高い神経生理学者 Warren S. McCulloch (1889-1969) と Computational Neuroscience の先駆者 Walter Pitts (1923-1969)³⁾、神経回路の学習則 (ヘップ則) で名高い神経心理学者 Donald O. Hebb (1904-1985)、Computational Neuroscience の泰斗 David C. Marr (1945-1980) などを始祖としている。さらに、Hassabis 他 (2017) や Marblestone 他 (2106)、Schuman 他 (2017) の包括的な展望によると、時代を席卷している BD-AI でも、最先端分野では、NM-AI と同じく、大脳新皮質・海馬・視床などの基本動作原理に習ったものがまればではない。つまり、一部に両者の収束傾向すら見られる。

なお、上記の機械学習 (ML) とは、BD-AI 分野の標準的教科書 (Goodfellow 他 (2016)) によれば、「AI システムが、生データからパターンを抽出することで、自分自身の知識を獲得する能力」と定義されている。一方、同書には、集合関係“AI ⊃ ML”が図示してはあるが、肝心の AI システムの定義自体は示されていない。ただし、「我々は、機械学習が、複雑な実世界の環境下で動作する AI システムを構築可能にする唯一のアプローチだと考えている」というこ

3) 良く知られているように、そして驚くことに、von Neumann が現行コンピュータの原型である EDVAC を考案した際の基本アイデアは、McClulloch and Pitts (1943) に基づいていた (太田 (2017))。

となので、AI と ML とは、実質的には区別されていないと言える。

したがって、本論でも、インテリジェンスという言葉の多義性の弊害をできるだけ避けるために、AI と ML とをほぼ同義に扱う。また、さらなる定義上の混乱をさけるために、第 3 節でその正当性は詳述するが、標題で含意するインテリジェンスを“インテリジェンス”とカッコ書きにして

「実世界の変化と異常に対応していくために、自他の過去の記憶と現況に立ち向かう自らのアクション（行動）とを活用しながら、様々な活動の“起因の特定 (Credit Assignment)”と予測 (Prediction) を行う能力ならびにその自己変化能」

と、Slowman 流の生命体・人工物の双方に適用可能な定義を採用したい。なお、定義の中にアクションが含まれるのは、エナクティブ・アプローチ (Enactivism)⁴⁾ の考えに従い、インテリジェンスというものは、それ自体として存在意義をもつものではなく、生命体であれ人工物であれ、実世界の中で自らの行動によって生き抜くための手段であることを強調するためである (Noë (2009), Hohwy (2013), Engel 他 (2016))

さらに、起因の特定とは、Minsky (2006 の第 8 章 5 節) にならって、

「特定の出来事において選ばれた色々なアクション（活動）の中から良好・不良な最終結果に繋がったアクションを関連状況（コンテキスト）やそこに至った学習方法と共に記憶して、さらなる変化と異常への対応に備える作業」

と理解する。このような起因の特定作業は、自他の過去の記憶と現況に立ち向

4) 本論の視点からは、「認知能力とは、実世界モデルを作るのではなく、人々の感覚・運動スキルに埋め込まれているようなアクションを援助・促進するものである。」(Engel 他 (2016), 筆者訳) とする考え方をさしている。なお、伝統的な認知科学では、認知能力＝実世界モデルを作り上げる能力と見なし、アクションとはそのような脳内に構築された実世界モデルに基づいて生みだされるものだとしてきた (Engel 他 (2016))。ところが、そのような脳内モデルは、実世界の動きに受動的に対応するだけではなく、自らのアクションによって積極的に働きかけることによって頻繁に改訂・改良されていくものだと思われる。

かう自らのアクション（行動）とによってしか実行できない。そして、起因の特定の抽象度や幅と深さが妥当であればあるほど、変化と異常に直面した際に、学習したことを新しい状況へ適用する転移学習 (Transfer Learning) を多彩かつ迅速にできるようになる⁵⁾。起因の特定能力が、学び方を学ぶ (メタ学習) 能力の基本であることによる。もちろん、そのような基本能力は、より一般的な (統計的) 因果関係の同定にも必須となる (Pearl (2018))⁶⁾。

このように、両タイプの AI は、“自己変化する情報駆動型制御システム”の実用化という基本目的では似通っており、しかも、一部に収束傾向さえも見られる。ただし、第 2 節で詳しく触れるが、脳神経模倣の度合いは両者で大幅に異なっている。例えば、BD-AI では非発火型 (Non-Spiking) 興奮性ニューロンだけにに基づく微分可能な非線形のニューロン間伝達関数 (ノ活動関数) を用いる。他方、NM-AI では実ニューロンにより近い発火型 (Spiking) の興奮性及び抑制性の二つのタイプのニューロンや両ニューロンに働きかけて微調整を行う調整性ニューロンに基づく非連続な非線形伝達関数を用いる⁷⁾。そして、直ちに理解できるように、発火型であるか否かは、前者の並列順次処理 (parallel serial processing) 特性から、時空間構造の表現のしやすさや情報伝達方向の有無 (有向グラフか否か) に本質的に関わってくる。

このような両者の違いは、素人的には一見僅かなように思える。ところが、次節で詳しく説明するように、非線形伝達関数が微分可能 (非発火型) であるかどうかや、興奮性ニューロンに加えて抑制性・調整性ニューロンが組み込まれているかどうか等々は、発揮される“インテリジェンス”の幅と深さにも非常

5) この点に関する次の Minsky (2006) の引用は、極めて示唆的である：「実際、起因の特定の優良さは、人々がインテリジェンスと呼ぶまるでスーツケースの中身のような種々雑多な特性の重要な側面となりうる。問題への解決策を単に記録するだけだと僅かに似た問題を解くことだけにしか助けにならないが、もしそのような解決策を我々がどのようにして見出したかを記録できれば、それによって、(起因の特定の仕方も学べるので) 我々はもっと広い種類の状況に対処できるようになる。」(Minsky (2007), 筆者訳、() 内は筆者追加)

6) なお、AI/ML の現状を知るという意味で興味深い事実であるが、先の標準的教科書 (Goodfellow 他 (2016)) では起因の特定は考慮外である。逆に言えば、実用的な AI/ML は、未だそのレベルに至っていない。

7) この特徴から、NM-AI は、Spiking Neural Network 型 AI (SNN-AI) とも呼ばれる。BD-AI の非発火型ニューロンネットワークは、SNN に対比する場合、Artificial Neural Network 型 AI (ANN-AI) と呼ばれる。

に大きな違いをもたらす。そして、それらの違いを深く知ることは、社会科学や人文科学にとっても、そもそもヒューマン・インテリジェンス (Human Intelligenc: HI) とは何か、HI は現行の BD-AI や NM-AI に組み込まれている“自己変化する情報駆動型制御システム”とどのように類似していたり相違していたりするの、などの現代的な課題を考えるための絶好の学習機会を与えてくれる。

本論の目的は、上記の視点に基づいて、今をときめく BD-AI と中長期的には AI の本丸として登場すると期待されている NM-AI の双方を取り上げ、そもそも BD-AI や NM-AI の“インテリジェンス”特性とはどのようなものであるのかを、HI と両タイプの AI との補完性・代替性に焦点を当てながら検討することである。

より具体的には、スーツケースの中身のような種々雑多な特性を持ってしまっているインテリジェンス⁸⁾という言葉を用いた非行動主義 /Neuroscience 的な視点から再整理・細分化することを試みる。そして、そのような試みに基づいて、HI の中核をなす“変化と異常への対応力” (小池他 (2001)) やその自己変化能に言及しながら、BD-AI/NM-AI 流の“インテリジェンス”と HI とがどのような包含関係や非包含関係にあるのかについて考察する。考察に際して特に留意するのは、柔軟で高速な視点切り替え装置としての情動 (Ornstein (1986), Franklin (1995), Minsky (2006), Rolls (2018)), 広範囲な協力を生み出すコミュニティ形成装置としての情動 (Minsky (2006), Damasio (2018), Rolls (2018)), 変化と異常に対応するための起因の特定／予測装置としての意識 (Llinas (2001), Friston (2010), Hohwy (2013), Feinberg 他 (2016), Tani (2017)) 等々といった視点である。

2. BD-AI と NM-AI の違いを検討する：機能特性という視点から

前述した人工ニューロン間の非線形伝達関数が微分可能であるかどうかや興奮性・抑制性ニューロン組み込みの有無などの緒特性は、BD-AI と NM-AI の“インテリジェンス”を規定する (何ができて何ができないのかという意味での) 機能特性に非常に大きな違いをもたらしている。この点をより直感的に理解する

8) この表現は、Minsky (2006) に倣っている。

ためには、BD-AI や NM-AI で想定されている人工ニューロンが、実ニューロンとどれほど異なっているかを理解することが早道だと思われる。また、そのことによって、BD-AI 型のみならず NM-AI 型“インテリジェンス”の限界も見えてくる。

2-1 BD-AI と NM-AI が想定するニューロンの違い：時空間概念の有無

下図の中央には、ヒトの場合にニューロン全体の 75%を占めると言われる興奮性錐体ニューロン (excitatory pyramidal cells) の概念図が示されている。この実ニューロンの上部や中央部には、他のニューロンからの刺激・イベント情報をシナプス⁹⁾／スパイン¹⁰⁾経由で受け取る上下に横広がり樹状突起 (dendrite) がある。なお、ヒトの錐体ニューロンには、一つのニューロン当たり 5 千-2 万個の樹状突起がある。また、シナプスの数は、ヒトで 4 万個／ニューロン、猿で 2 万個／ニューロン、マウスで 8 千個／ニューロン、ミツバチでも 1 千個／ニューロン¹¹⁾となっている (以上は、Rolls (2017) 参照)。

これらのシナプス／スパイン経由で樹状突起に集められた他のニューロンか

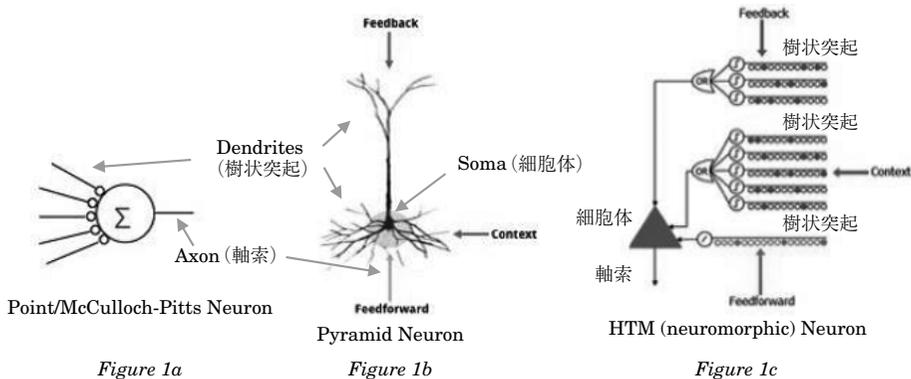


Figure 1a

Figure 1b

Figure 1c

Numenta (2017) より抜粋

9) ニューロンとニューロンとの接合部のこと。電気的な興奮・抑制情報がシナプス前部線維の末端までくると、そこからドーパミン、セロトニン、グルタミン酸、GABA などの各種の化学伝達物質が放出され、それらがシナプス後部膜の膜電位を変化させる。このようにして興奮・抑制情報がニューロンからニューロンに伝えられる。このように、ニューロン間が化学的に繋がって行く方式をシナプス伝達と呼ぶ。

10) 樹状突起には、既存のシナプス以外にもシナプス的な役割をするスパインと呼ばれる棘状の微細な突起が数多く存在する。

11) <https://galton.uchicago.edu/~nbrunel/teaching/fall2016/11-intro.pdf> 参照。

らの刺激・イベント情報は、中央部にある細胞体 (Soma) に集められる。細胞体は、集積された刺激・イベント情報の値がある閾値を超えると、その下方に伸びる軸索に活動電位 (Action Potential) の不連続的な放電・発火 (Discharge/Spiking) を起こす。

なお、実ニューロンには、興奮性錐体ニューロンの他に、シャンデリア細胞 (chandelier cell) とかバスケット細胞 (basket cell) などと呼ばれる抑制性 (inhibitory) ニューロンがある。さらに、両ニューロンを繋ぐ多数の介在ニューロン (Interneuron) や両ニューロンの微調整 (ファインチューニング) のみに関与する少数の調整性ニューロン (modulator neuron) が存在している (Rolls (2017), Luo(2016))。

微に入り細に入りの叙述で申し訳ないが、実ニューロンのネットワークシステムでは、このようなシナプス/スパイン型のコミュニケーションに加えて、ニューロン間・アストログリア¹²⁾間・ニューロン-アストログリア間のギャップ結合 (Gap Junction)¹³⁾と呼ばれるカルシウムイオンの波などを活用した直接接合による非シナプス型のコミュニケーションも積極的な役割を果たしている (Fields (2009), 工藤 (2011), Deutch 他 (2014))。また、細胞外にある局所的な電場・磁場を介した混線現象 (Crosstalk) に基づく直接接合をも包摂した Ephaptic Transmission (表面接触型伝達) と呼ばれる細胞間コミュニケーションの存在が知られている (Kandel 他 (2013: 第 6 版), Deutch 編 (2014))¹⁴⁾。

さらに、軸索の中の細胞骨格を形成している微小管 (Microtubule) を経由して新陳代謝のために必要な各種の細胞内小器官を順送・逆送する軸索輸送 (axonal transmission) といったコミュニケーション手段も存在する (Deutch 他

12) ニューロンとは大きく異なる脳・脊髄内のグリア細胞の一つ。脳・脊髄内のグリア細胞には、大きくは、アストログリア、マイクログリア、オリゴデンドログリアがある。ヒトの場合、脳・脊髄内には、ニューロンの約 6 倍のグリア細胞が存在する。しかも、各々のグリア細胞は、状況に応じて大きく機能や形態を変えていく。また、ミエリン鞘と呼ばれるニューロンを包んで絶縁特性を発揮するオリゴデンドログリアには、脳・脊髄外では同類のシュワン細胞がある。これらのグリア細胞は、最近になればなるほど、対ニューロンという意味でも、その重要性が大きく見直されてきている。以上は、主に Fields (2009), 工藤 (2011) を参照。

13) 例えば、アストログリアの場合、広汎な細胞間ネットワークがカルシウムイオン波を媒介として生成されている。しかも、このようなグリア細胞間ネットワークは、ニューロンネットワークとも密接に同期し合っている (Fields (2009), Kandel 他 (2013: 第 6 版))。

14) ニューロンやグリア細胞等々が渾然一体として存在する網状組織は、neuropil (ニューロピル) とも呼ばれる (Freeman (2001))。

(2014), Luo (2016))。しかも、微小管は、軸索のみならず樹状突起内部にも張り巡らされていて、樹状突起と他のニューロンを繋ぐシナプス／スパインの機能を裏方的に支えている。また、最近では、驚くべきことに、筑波の物質・材料研究機構 (NIMS) の研究グループによって、微小管自体がフラッシュメモリと同じような多ビットのメモリ・スイッチング機能を果たせることも示されている (Sahu 他 (2013))。つまり、ニューロン内の微小管自体が、タンパク質でできた SSD (Solid State Drive) のような役割をも果たすことができるというのである。

このように、実ニューロンのネットワークシステムの効率性や低消費電力性を支える仕組みは、実際には、ニューロン主体の BD-AI や NM-AI の枠組みをも遙かに超えた極めて複雑なものである。したがって、現在開発段階にある NM-AI も、今後の Neuroscience のさらなる発展によって、その設計思想自体が大きく変化していく可能性が高い。

話を BD-AI や NM-AI の伝達 (活動) 関数に戻すと、j 番目のニューロンの伝達関数とは、他の i 番目のニューロンからこのニューロンの樹状突起を経由して届く刺激・イベント情報の強さ (活動水準) を X_i 、j 番目と i 番目のニューロン間のシナプス結合の強さを W_{ji} すると、数式的には下記のような非線形関数 F で簡略に表現されている (Rolls (2017) 参照)。

$$Y_j = F(W_{j1} \cdot X_1 + W_{j2} \cdot X_2 + \dots + W_{ji} \cdot X_i + \dots + W_{jn} \cdot X_n)$$

(Y_j は j 番目のニューロンの活動水準、n は繋がっているニューロンの総数)

そして、BD-AI では、上記の F が、数式的に扱いやすい非発火型の微分可能関数である。他方、実ニューロンや NM-AI では、 F が、活動電位がある閾値を超したときにだけ $Y_j > 0$ となる発火型の非連続な関数となっている¹⁵⁾。

さらに、実ニューロンには、ヒトではニューロン全体の 20% を占めると言われる先の抑制性ニューロンがあり、興奮性ニューロンの暴走 (例えばてんかん症状など) を防ぎながら最適制御に関与している。つまり、実ニューロンシステムは、これらの興奮性・抑制性ニューロンの正負の相互作用、ならびに両者の微調整を司る調整性ニューロンによってシステムの安定性を自動的に高め

15) なお、非発火型のニューロンを採用する場合、少なくとも概念上は、同じく非発火型である前述のグリア細胞の機能の一部と識別が難しくなるようだ。

ている。

そして、その忠実な模倣ぶりにはやや驚きであるが、ほとんどの NM-AI には、実ニューロンと同じく、興奮性と抑制性ならびに調整性のニューロンが組み込まれている。例えば、IBM の TrueNorth, Intel の Loihi, ハイデルベルグ大学の BrainScaleS, マンチェスター大学の SpiNNaker, スタンフォード大学の NeuroGrid, ベンチャー企業 BrainChip Holdings の BrainChip といった代表的な NM-AI がそれに当たる。

ところが、BD-AI には興奮性のニューロンだけしか組み込まれていない。もちろん、当該ニューロンから負値を出力可能にすれば、便宜的ではあるが、抑制ニューロンの振る舞いをさせることはできる。したがって、BD-AI では、NM-AI に比べてニューロンの間が広範囲にわたって極めて密に配線されがちである¹⁶⁾。このような配線構造は、Dense Distributed Representation：“密な分散表現／表象：DDR¹⁷⁾”と呼ばれる。興奮性ニューロンだけしか存在しないと、一つのニューロンがある刺激・イベントに対して反応した場合、その周辺ニューロンの反応性も上昇させ、結果として互いが反応性を次々に高め合って広がっていく暴発型のポジティブフィードバックが産み出されるからである。そして、このことが、BD-AI の高電力消費量の一因ともなっている。

なお、BD-AI では、数式的な扱いやすさを保つために、ニューロンが階層別にグループ化されており、階層グループ間の配線は自由であるが、各グループ内でのニューロン間配線は許されない。しかも、そのニューロンネットワー

16) BD-AI にも、できるだけ過密配線を避けようとの試みが少なくない。例えば、日本の LeapMind (<https://leapmind.io/>) や中国の Cambricon (<http://www.cambricon.com>) などのベンチャー企業の試みでは、消費電力量が一桁から二桁下げられるとしている。特に、後者は有名であり、Huawei のスマートフォンにも搭載されている。

17) 分散表現/表象とは、特定の記憶が特定の一つだけのニューロンに任されるのではなく、刺激・イベントの到着によって同時に活性化した一群のニューロン集団によって記憶される様子をさしている。太田 (2017) は、この分散表現/表象を「複数のニューロン発火で一つの概念を表象するという考え方」、「同期した発火=表象」と明快に表現している。ただし、彼は分散表現/表象を一つの仮説に過ぎないと断言し、次のように続けている。「理論神経科学は、シナプス遅延時間は一定であるという仮定から、同期した発火を表象とみなしてきた。しかしながら、実験的にニューロンのシナプス遅延時間を計測してみると、脳の領域・細胞種によって多様であることがわかって来る。その場合、ある時刻における発火パターンを用いて一様にそれを表象と結び付けることは不適当であるといえる。そこで、空間的な発火パターンではなく、時間的なスケールが多様であることを前提とした発火の連鎖過程についての考察が必要である。」脳機能の奥深さが、垣間見えるようである。

クには実質的に向きがない（無向グラフ）ので、時系列情報を表現できない。このような制限を持つニューロンネットワークは、制限付きボルツマンマシンあるいは多重パーセプトロン (Perceptron) と呼ばれる (Goodfellow 他 (2016))。

他方、脳内ニューロンネットワーク及び NM-AI では、このような過密配線やポジティブフィードバック暴走ができるだけ起こらないようになっている。具体的には、周辺ニューロンの中で最大の出力を誇るニューロン群だけに学習・記憶機会が与えられ、それ以外のニューロンは主に抑制性ニューロンによって発火しないように抑えられる。つまり、多数の周辺ニューロン中の一つ、あるいは、その中のわずかな数のニューロンだけが発火して低消費電力性に貢献するように工夫されている (Luo (2016), Rolls (2017), Anderson (2017))。その結果、対応するニューロン間の配線は、とても疎くなっている。

このような配線構造は、先の“密な分散表現:DDR”に対比して Sparse Distributed Representation：“疎な分散表現:SDR”と呼ばれ、NM-AI の低消費電力性にも大きく貢献している¹⁸⁾。この SDR に関しては、後でも繰り返して触れる。

なお、NM-AI は、多重パーセプトロン型の BD-AI に比べて、ニューロン間の配線の自由度が極めて大きい。例えば、Intel 製 NM-AI である先の Loihi では、BD-AI とは異なり、ニューロン間の配線は、同一階層内でも階層間でも可能であり、さらに（自らの軸索が反回して自らの樹状突起に繋がる形の）自己再帰型の配線をも組み込むことができる (Davies (2018))。階層型にするか否かの決定も可能である。しかも、そのニューロンネットワークには向きがある（有向グラフ）ので、時空間情報の表現力も高い。このような自由度の高い配線構造が可能になると、BD-AI では到底実行できない多彩な連想が可能になる。特に、自己再帰型配線の場合、僅かな初期条件の違いによって異なった連想に

18) この点に関し、先の Intel Labs 所長の Uhlig は、「アルゴリズム的な視点からは、発火型ニューロンは、例えばワン・ショット学習（小サンプル学習）や意志決定を支えている時間軸に沿ったイベント（出来事）処理を行うニューラルネットワークの主要なアプローチである。実装化の視点で見ると、発火現象により、脳神経模倣型設計思想にとって、これらのアルゴリズムによってエネルギー効率上の大きな便益をもたらす高いスパース性（SDR 性）を活用できるようになる。これらの優位性は、製造現場や自動運転車、ロボットのようなエッジデバイス（ユーザーサイドで使用される機器）に大きな価値をもたらす。これらの応用製品では、リアルタイムでの処理や（データなどの）取込が必要なことによる。」(Ackerman (2018), 筆者訳、() 内は筆者追加) と述べていて興味深い。

たどり着けるので、創造性の源泉にもなり得る（甘利 (2008), (2016), Kandel 他 (2013), Rolls (2016) ¹⁹⁾。

脳内神経ネットワーク及び NM-AI で SDR を可能にする上記の Winner-Take-All（勝者独り占め）型の巧妙な仕組みは、競合学習 (competitive learning) や側方抑制 (lateral inhibition) と呼ばれている。なお、生命体の脳の低消費電力性の達成には、進化論的にも SDR が必須であった。この点に関する Anderson (2007) の次のコメントは、BD-AI の限界を知る上でも誠に興味深い。

「(神経回路での) 配線は、場所を取るしエネルギーも使うし正確に配線することも難しいので、生物学的にはとても高価なものである。存在している様々な配線はとても高価なものだし、それらの連結パターンはキッチンとしたコントロールの下で行われなければならない。これらの問題は、大規模な脳モデルにとって厳しい制約を課す。したがって、(BD-AI が想定するような) 全てが全てに繋がっていると言うのは間違っている。現実にも、そうっていないからだ。」(筆者訳, () 内は筆者追加)

繰り返しになるが、上図の Figure 1a に示されているように、BD-AI で使用される人工ニューロンは、ポイント・ニューロン (Point Neuron, 別名“McClulloch-Pitts 型ニューロン”) と呼ばれる極めて単純な構造をした興奮性ニューロンであり、ニューロンとしてはこの 1 種類だけしか組み込まれていない。ポイント・ニューロンは、具体的には、細胞体に入力する単純で同一な多数の樹状突起と一本の軸索からなっている。単純という意味は、それが他のニューロンからの刺激・イベント情報を伝えるだけであるからである²⁰⁾。しかも、前述のように、

19) この点に関する以下の甘利 (2016) の指摘は、とても興味深い。「脳は記憶そのものを蓄えるのではない。これを思い出すための仕掛けを蓄え、ヒントから復元すべき情報を作り出す。だからときには間違えるし、思い出せないことも起きる。思い違いだってある。その代わりに、脳は柔軟である。間違ったヒントや曖昧なヒントからでも答えが出せる。多数のパターンを重ね合わせてしまうから、全体が茫洋としていてどの記憶事項がどこにあるかはわからない。しかし、並列のダイナミックスで働く分散した記憶が実現できるというわけだ。」

20) Boden (2016) は、BD-AI 型のニューロンを“too neat, too simple, too few, and too dry”と特徴付けている。このなかの neat は数式的扱いやすさの優先, simple は単一ニューロンで全ての複雑な並列分散処理を実現, few はヒトの脳と比べてのニューロンの数の少なさ, dry は時間概念・共時性や樹状突起・神経伝達物質・シナプス電流・イオンの流れの生物物理特性を

数式的に扱いやすい非発火型にするために、非線形の伝達関数には計量経済学で多用される多項ロジット関数²¹⁾のような微分可能な関数が多用される。

このような単純なポイント・ニューロンと対比させるために、図右端の Figure 1c には、Numenta (2017) の HTM (Hierarchical Temporal Memory) と呼ばれる NM-AI ソフトウェア²²⁾に組み込まれている人工ニューロン事例が示されている。この人工ニューロンでは、実ニューロンに習って細胞体の遙か上に先端樹状突起 (Apical Dendrite)、周辺に基底樹状突起 (Basal Dendrite) が組み込まれている。また、実ニューロンでは、樹状突起の細胞体からの距離によって役割が異なってくるので、さらに、近接 (Proximal) か末梢か (Distal) かの区別がなされる。役割が異なる大きな理由は、シナプス結合する抑制性ニューロンや調整性ニューロンの質・量が異なっているからである (Luo (2016))。ただし、Numenta (2017) では、簡略化のために先端樹状突起では近接か末梢かの区別はないが、基底樹状突起には区別がある。

具体的には、Figure 1c の HTM では、樹状突起の中の先端樹状突起には他のニューロンからの刺激・イベント情報が入力するが、その基底樹状突起、特にその末梢部分には、実ニューロンに習って細胞体発火時に既存の学習・記憶結果としての様々なコンテキスト情報が周辺の色々な興奮・抑制・調節性ニューロン群から伝わるようになっていく。その結果、HTM の人工細胞体は、実ニューロンと同じく (図中では Feedforward と追記されている) 近接の基底樹状突起などの制御を主に受けて、発火、発火直前、発火前準備、非発火という 4 つの状態をとることができる (Numenta (2017), Deutch 他 (2014))。2 番目と 3 番目の状態が、既存学習・記憶に基づく先読みによる発火直・前の準備と考えられる。

このような先読みのための既存学習成果のコンテキスト回路化は、どのソースからの情報が重要であるかに関して内生的に示された (無意識ベースの) 価

無視していることを意味する。

21) 多項ロジット関数に関しては、<http://www.ier.hit-u.ac.jp/~kitamura/lecture/Hit/08Statsys8.pdf> などに分かりやすい説明がある。

22) NM-AI では、現行のコンピュータ上で利用可能なプログラミングを駆使した NM-AI ソフトウェアの研究開発グループと新奇な設計思想に基づく脳模倣型ハードウェアの研究開発グループとが分かれて存在している。ソフトウェアグループが、ノイマン/チューリング型コンピュータ用の高度に発展したソフトウェア資産を用いるためだと思われる。

値判断による回路化と見なすことができる。そして、このようなコンテキストの回路化が精密かつ多彩に達成できればできるほど、より多彩な幅と深さを持つ起因の特定が迅速にできるようになる。このことから、樹状突起の質と量が、“インテリジェンス”の高低を大きく規定するということが分かる。事実、Richard 他 (2018) によれば、高度な起因の特定機能の発現に際して、先端末梢部分に位置する樹状突起が、行動・知覚に関する予測誤差をニューロンレベルで局所的に修正する（繋ぎ直す）際に決定的な役割を果たすことができるという。

このように、実ニューロンの先読みの精度には、興奮・抑制・調節性ニューロン群からの情報を処理する樹状突起自体の質・量の豊富さが大きく関わっている (Numenta (2007), Deutch 他 (2014), Luo (2016), Rolls (2017), Richard 他 (2018))。さらに、このような豊富さの有無は、次に詳述するように、“インテリジェンス”特性に決定的な違いをもたらす時空間概念の有無にも大きく関わってくる。

事実、脳内や NM-AI の神経回路では、不連続的な伝達関数に特有の発火という形で、時空間パターンがニューロン間配線の強弱・広狭として刻まれていく。ところが、BD - AI の場合、特に DNN (深層ニューラルネットワーク) や CNN (畳み込みニューラルネットワーク) と呼ばれる代表的なモデルの場合、空間的な拡がりパターン化することはできるが、時間の流れ (時系列) をパターン化することができない。そもそも、BD-AI のニューロンネットワークには、扱いやすさを優先した向きがない (無向グラフ) 構造が採用されているからである。さらに、前述のように、その単純な樹状突起特性により、先読み機能などが組み込まれていないことも大きく影響している。

加えて、前述のように BD-AI には抑制性や調整性のニューロンも組み込まれておらず、しかも、扱いやすい微分可能な伝達関数が想定されているので、実ニューロンや NM-AI に備わっている STDP (spike timing-dependent plasticity) と呼ばれるシナプスを起点としたシナプス前ニューロンとシナプス後ニューロンとの時間依存的で巧妙な相互作用を組み込むことができない。繰り返しになるが、STDP 機能は、前述した代表的な NM-AI には例外なく組み込まれている。ちなみに、STDP とは、次のように定義される実ニューロンの仕組みである。

「LTP (長期記憶増強) と LTD (長期記憶抑制) の興味深い時間依存性が次

のような形で観察されている。シナプス前 (のニューロンの) 発火がシナプス後の (ニューロンの) 活性化より数ミリ秒先んじるときには LTP が起き、シナプス前発火がシナプス後活性化の数ミリ秒後に起こるとき LTD が起こる。この現象は、発火タイミング依存的柔軟性、STDP と呼ばれる。」(Rolls (2017), 筆者訳, () 内は筆者追加)

STDP が組み込まれていると、局所的な情報に依存するだけでニューロン間の空間構造に加えて時系列構造の様々なパターンを扱うことができる。しかも、このような時空間パターンが事前に学習・記憶されていれば、同じようなコンテキスト (状況) に出会った時に、前述の先読み制御をより高度なものにすることができる (Rolls (2017), Tavanai 他 (2019))。

実際、ヒト同士の日常会話一つをとってみても、時系列情報の効率的な処理による先読み (含む常識の活用など) が、相互のコミュニケーション効率を上げるために必須の仕組みとなっている (Minsky (2006))²³⁾。統計学的な規則的關係を帰納的に類推するのであれば、計量経済学者がそうしているように、時空間にわたって少なくとも千個あるいは一万個のサンプルが確保できれば、そのような関係を高い確率で同定できるようになるからである。

以上のように、NM-AI では、樹状突起の質と量の豊富さに加えて、発火型ニューロンの STDP 機能などが実現する時空間構造を組み込むことができるので、時系列構造が組み込まれていない BD-AI と比べると、時空間にまたがる起因の特定能力、したがって“インテリジェンス”の高さに大きな違いが生まれる。

ちなみに、BD-AI にも、極めて複雑な入れ子構造の導入が必須ではあるが、自然言語処理用などのモデルとして時間概念を取り入れた RNN (Recursive

23) 時系列情報を効率的に扱えることは、五感情報の中でも特に膨大な視覚情報処理の場合、さらに重要となる。例えば、ヒトの視覚には、驚きであるが、1 秒あたり 1 ギガビットの画像情報が引切り無しに到着する (Olhausen 他 (2017))。これほどの膨大な情報量になると、各時点・時点で到着する空間情報を独立なもののみならず処理しては、前述のエネルギー節約的な“疎な分散表現：SDR”をもってしても、たちまちのうちに処理能力不足に陥ってしまう。したがって、到着する膨大な時空間情報からサンプリング (抽出) するデータ数を激減させる工夫が必須となる。そして、哺乳類のみならず節足動物である昆虫にとってさえ、時空間データの中に繰り返して現れる統計学的な相互依存・因果関係を利用することが極めてエネルギー節約的となる (Jayaraman 他 (2009), Lotto (2009) 参照)。

Neural Network), RNN 内に短期・長期記憶機能を組み込んだ LSTM (Long and Short Term Memory)型 RNN, LSTM に加えて RNN 外にも後述の高速な連想メモリ (Content-Addressable Memory: CAM)を持つ Neural Turing Machine などが考案されている²⁴⁾。ただし、以下で紹介するように、時系列処理用のメモリ総容量がキロビットあるいは高々メガビットと極めて限られているにも関わらず、その消費電力はより大きくなってしまふ。

2-2 BD-AI の実用性 vs. NM-AI の低消費電力性

以上のように、低消費電力性や組み込まれているニューロン諸機能の卓越性という意味では、起因の特定に必須の短期・長期にわたる時系列構造の組み込み易さをも含めて、NM-AI に遥かに大きな利点がある。ところが、実用化に富み今をときめいているのは紛れもなく BD-AI であり、明日の AI を担うとされる NM-AI は現時点では実用化からほど遠い段階にある。なぜだろうか？

その大きな理由の 1 つは、ビッグデータを使いこなして実用に耐える AI/ML 機能を実現するために必要な伝達関数の推定すべきパラメータ数が、関数型の単純さにもかかわらず、あまりに膨大なことによる。そして、NM-AI では、現段階では、最先端のコンピュータを利用しても、このようなビッグデータを、BD-AI のように実用的な時間内で処理できない。

推定すべきパラメータ数の膨大さを示すものとして、例えば、AI 実用化元年とも言われる 2012 年にカナダ・トロント大学の Hinton 教授グループがその卓越した実践性を発揮した AlexNet と呼ばれる BD-AI モデルの場合、伝達関数の推定すべきパラメータ数は実に 6200 万個を超えるものであった。また、SegNet と呼ばれる英国・ケンブリッジ大学の著名な BD-AI モデルでは、自動運転のみならず食肉処理ロボットなどにも応用されつつあるが、BD-AI 本来の DDR (密な分散表現) 特性からする 1 億 3400 万個のパラメータ推定が必要だという²⁵⁾。ただし、このような莫大な数のパラメータ推定は実用的でないた

24) これらの簡単な紹介は、Goodfellow 他 (2016) にある。

25) 各モデルのパラメータ数に関しては、下記を参照。

<https://jeremykarnowski.wordpress.com/2015/07/15/alexnet-visualization/>, <http://mi.eng.cam.ac.uk/projects/segnet/> 及び日経ロボティクス 2018 年 12 月号の「豚の食肉処理ロボットにディープラーニング技術：肉切るナイフの①をセグメンテーションで推定し個体差対応」。

め、様々な工夫によって配線経路が数多く間引かれている。それでも、1470 万個のパラメータ推定が必要だという。

これほどまでに莫大な数のパラメータ推定には、容易に想像できるように、推定の自由度を確保するためのビッグデータと推定処理のための超高速コンピュータの双方の利用が必須となる。しかも、推定の際には、ニューロン間での伝達関数が微分可能でないと、最急降下法 (Steepest Descent Method) や誤差逆伝播法²⁶⁾ (Back-Propagation Algorithm) などの汎用数値計算アルゴリズムが利用できないので、そもそも超高速コンピュータでも実用的な速度での計算ができない。

逆に言えば、BD-AI というイノベーションは、月並みな表現ではあるが、ビッグデータ・高速コンピュータ・(その原型は 1960 年代に開発済みの) 機械学習アルゴリズムの三拍子が揃った現代になって初めて産み出されたものなのである。ただし、Tavanaei 他 (2019), Wu 他 (2018), Severa (2018) によれば、NM-AI でも伝達関数が微分可能でないという弱点を克服するためのイノベーションが実際に産み出されつつある。

なお、BD-AI の分野では、このようなパラメータ推定を行うことを学習／訓練 (training) と呼ぶ。そして、学習／訓練には、高速なコンピュータでも一週間以上を要することは希ではない。一方、BD-AI 応用の山場である学習／訓練過程が終了すれば、一般ユーザーの PC やスマートフォン単体 (Edge: エッジと呼ばれる) でも、推定済み関数の活用 (Inference: 推論と呼ばれる) ができるようになる。その結果、BD-AI は、iPhone や Galaxy などのスマートフォンにみられるように、学習／訓練型 AI と推論型 AI とが分離された形で開発・実装されることが多くなってきている。上記の Tavanaei 他 (2019), Wu 他 (2018), Severa (2018) も、学習／訓練は BD-AI で行い、推論だけを NM-AI に任せる方式の提示が主である²⁷⁾。

26) インプットデータが、多段階の伝達関数で表現される深層神経ネットワーク (Deep Neural Network) を経て生成されるアウトプットデータと可能な限り同一になるように考案されたアルゴリズム。アイデア自体は、日米の研究者によって 1960 年代前半に開発されたと言われる。詳しくは、Goodfellow 他 (2016) を参照されたい。

27) もちろん、グーグル翻訳・音声認識エンジンや iPhone の Siri のような莫大な学習／訓練が必要な自然言語処理用プログラムの場合、推論段階でも依然としてインターネット経由で Google や Apple のサーバにアクセスする必要がある。いずれにせよ、このような状況であるため、我々が思い描く AI の特徴であるリアルタイム (即時) での学習／訓練・推論は、現行の BD-AI に関しても、夢のまた夢なのである。

では、ビッグデータの前では NM-AI の登場余地が全くなくなっただのかというと、そうではない。実際、BD-AI に比べて現時点で大きな処理速度上のハンディキャップを持つ NM-AI であるが、多くの研究開発者達が実ニューロンのより忠実な模倣に拘っていることには十分な理由がある。その一つは、現行のノイマン／チューリング型と呼ばれるコンピュータの申し子である BD-AI に拘っていても、中長期的に地球規模での電力供給がとても追いつかなることがほぼ確実であることによる。その様子は、BD-AI 実用化に最大の貢献をしてきたノイマン／チューリング型コンピュータが必要とする莫大な消費電力に着目するとより直感的に理解できる。

例えば、世界の囲碁チャンピオンを次々になぎ倒して引退した Google の AlphaGo を支えているコンピュータ・システムを眺めてみよう。AlphaGo の仕組みを詳細に伝えている Nature 論文 (Silver 他 (2016)) では、そこで使われているノイマン／チューリング型コンピュータに関して下記のような記述がなされている。

「AlphaGo の最終版では、40 の検索スレッドと 48 個の CPU 及び 8 個の GPU を (持つマシン) 使った。また、我々は、40 の検索スレッドと 1202 個の CPU 及び 176 個の GPU で特徴づけられる複数マシンに分散して動作する AlphaGo の分散版も実装した。」(筆者訳)

そして、おそらくこの“1202 CPUs and 176 GPUs”というスペックに基づいたと思われる 2017 年 7 月 27 日の日経新聞記事は、「人間の脳の消費エネルギーは思考時で 21 ワット。一方のアルファ碁の消費電力は 25 万ワットとされてきた。約 1 万 2 千人分だ。」と強調している。実際、1CPU 当たりの最大消費電力を 145 ワット²⁸⁾、1GPU 当たりの最大消費電力を 300 ワット (／毎時)²⁹⁾ とすると、これだけで 23 万ワットなる。したがって、システムメモリ³⁰⁾ や

28) Intel Xeon ES-2600 V4 の値。https://ark.intel.com/ja/products/91755/Intel-Xeon-Processor-E5-2697-v4-45M-Cache-2_30-GHz 参照。

29) Nvidia Tesla P100 の値。<http://images.nvidia.com/content/tesla/pdf/nvidia-tesla-p100-datasheet.pdf> 参照。

30) 「オペレーションシステム (OS) が使用するコンピュータ・システム内の記憶領域。OS の中核部分であるカーネルやデバイス管理情報、管理するウィンドウなどのインターフェース情

HDD/SSD などの周辺機器なども勘案すると 25 万ワットは順当な推定値だと見なせる。

また、全米の人気クイズ番組 Jeopardy で用いられる様々なクイズ形式の難問にも迅速な高正答率を誇った IBM Watson は、同じ BD-AI でも囲碁という用途に特化した AlphaGo とは大きく異なる極めて実用的な仕組みを持っている (Hurwitz (2015), Anderson (2017) など)。中でも Watson をユニークにしているのが、IBM 独自開発の自然言語処理技術が組み込まれた DeepQA : “a massively parallel hypothesis generation and evaluation task” (超並列仮説形成・評価作業) ソフトウェアである³¹⁾。

ただし、Watson の消費電力が、これまたすごい。実際、IBM Research の資料によれば、Jeopardy で使用された 2011 年当時の Watson の中核は、90 台からなる IBM Power 750 サーバ群であり、16 テラバイトの DRAM (Dynamic Random Access Memory), 4 テラバイトのディスク, 2880 個の Power 7 コア (Power-CPU 360 個相当), 80 テラフロップス (1 秒間に 80 兆回の演算可能) を誇るコンピュータであると記されているからである。同資料に Watson の当時の使用電力量は見つけられなかったが、Forbes の記事では 20 万ワットとされているので、上記の AlphaGo とほぼ同じである。

他方、前述した IBM の TrueNorth, Intel の Loihi, ハイデルベルグ大学の BrainScaleS, マンチェスター大学の SpinNaker, スタンフォード大学の NeuroGrid, ベンチャー企業 BrainChip Holdings BrainChip といった代表的な NM-AI の電力消費量は、開発段階にあるものの、いずれもノイマン/チューリング型の少なくとも 1/1000 ほどになっている。この値は、NM-AI の演算速度を実ニューロンに合わせたり、アナログ回路を多用したりすると、さらに 1-2桁下がってくる (堀尾 (2017) など)。例えば、上記の BrainScaleS や NeuroGrid

報などが書き込まれる。」 (<https://kotobank.jp/word/system%20memory-1689353>)

31) NM-AI の世界的な研究者として名高い Anderson (2017) は、Watson を AlphaGo とは大きく異なる Brain-like なものだとして次のような興味深いコメントをしている。「ワトソンのソフトウェアには、数多くの“先端”ソフトウェア、巨大な初期データセット、超並列計算、作業依存的な確率推定、連想学習、信頼推定・相関推定、そして開発者達が” (Know-How 的な) 浅い知識と (連想や推定などが必須の洗練された専門的な) 深い知識との統合“とよぶものが使われている。初期の AI を特徴づける合理的分析は、ほんの僅かな役割しかない。…脳のようなコンピュータからは遥かに遠いものだが、方向はそちらに向いている。」(筆者訳、() 内は筆者追加)

では、アナログ回路が多用されている。最近では、このアナログ化への試みが、IBM のような大企業³²⁾、Mythic (<https://www.mythic-ai.com/>)、Syntiant (<http://www.syntiant.com>) などのベンチャー企業によっても精力的に行われてきている。

ちなみに、人間の脳は、20 ワット前後で動いていると言われるが、一般的な生命体の脳では、現行コンピュータよりも 10^{-10} 乗程度の低消費電力性を誇るといふ (Rhine (2018))。この点に関しては、アフリカ生まれの鬼才 Kwabena Boahen@Stanford 大学が率いる先の NeuroGrid 研究開発グループ (Benjamin 他 (2014)) の指摘する下記の事実がとても興味深い。

「パーソナルコンピュータは、マウス規模の脳モデル (= 250 万個のニューロン) をシミュレートする際に、4 万倍 (400 ワット対 10 ミリワット) もの大きな電力を必要とするにもかかわらず、実際のマウスの脳よりも 9 千倍遅い。(欧州の) Human Brain プロジェクトのゴールである人間規模の脳モデル (= 200 億個のニューロン) をシミュレートする際には、エクサスケール (1 秒間に 100 京 (京 = 1 万兆) 回の演算能力) のスーパーコンピュータと (それを動かすための) 40 万世帯分に匹敵する電力消費量 (= 5 億ワット) とが必要になると予想されている。そのため、大規模なニューロンモデルの潜在力は、ほとんど利用できていない。」(筆者訳)

では、なぜ BD-AI では消費電力効率が NM-AI に比べて極端に悪いのだろうか？それは、BD-AI を支えているノイマン／チューリング型コンピュータには、良く知られた“フォンノイマン・ボトルネック (VNB)”が大きく立ちはだかっているからである。また、上記の Boahen らが教えてくれているように、VNB が不可避な BD-AI では、前述した非発火型の単純なニューロンに依存せざるを得ない。

現行のコンピュータでは、ソフトウェアプログラムによって命令 (Instruction) ・データがメモリからプロセッサに呼び出されて (= fetch) 演算処理が実行され、その結果が再度メモリに戻されるというパターンが限りなく繰り返される (Stokes (2010))。しかも、実行される命令のひとつ一つは極めて単

32) <https://www.techspot.com/news/77687-ibm-announces-8-bit-analog-chip-projected-phase.html>

純なものなので、プロセッサの速度やプログラムの複雑性が増せば増すほど、プロセッサ・メモリ間のやり取りが極端に増加する。そして、高速化傾向の著しいプロセッサの動作速度に比べて、現行の各種メモリの動作速度は技術的な限界からなかなか上げられない。その結果、高速なプロセッサであればあるほど VNB が深刻になり、消費電力の大部分がプロセッサ・メモリ間のやり取りに消費されてしまう（以上は、Hennessy 他 (2016) 参照）。

他方、NM-AI を支える実ニューロンにより近い発火型の非線形伝達関数を持つ人工ニューロンは、プロセッサとメモリの融合・同居型アーキテクチャを前提としているので、そもそも VNB が発生しない。その理由を、松本他 (2003) は、次のように説明している。

「フォン・ノイマン型デジタル・コンピュータではメモリの 1 番地の内容からプロセスし計算を実行する。従って、ここでのメモリの役割はデータ（プログラムもデータと見做される）の一時格納であり、プロセッサを可変にする為の補助装置である。これに対し、脳はメモリベース・アーキテクチャ（メモリ主体型方式）である。脳は、脳が獲得したアルゴリズムを神経回路の構造やその活動などの変化として学習によって固定化し記憶するので、脳のアルゴリズムは一種のルックアップ・テーブル（計算処理を配列の参照処理で置き換えて効率化を図るために作られた連想配列³³⁾）に貯えられたメモリとして存在する、と考えることができる。脳への入力情報は、このルックアップ・テーブルからどの答えを引きだすかの検索情報として用いられる。脳が答えを引きだす（出力する）と、引きだした答のアルゴリズムは、出力依存性学習によって、自動的に書き変わる。」（松本他 (2003), 230 頁, () 内は筆者追加）

なお、ニューロンはプロセッサとメモリの融合・同居型と表現されると相当に複雑なように感じられるが、計算的には、掛け算と足し算が行われている程度である (Rolls (2017))。より具体的には、ヒトの脳内では、各々のニューロン／ニューロン集団に脳内外から引切り無しに到着する新しい刺激・イベント

33) <https://it-words.jp/w/E383ABE38383E382AFE382A2E38383E38397E38386E383BCE38396E383AB.html>

が、ニューロンひとつ一つを構成要素とする 1 万を遙かに超える高次元ベクトル（先の SDR）の形で次々に刻み込まれていく（分散表現されていく）。しかも、この高次元 SDR では、先の Winner-Take-All 型の仕組みなどによって、成分のほとんどの値が 0 になっている。つまり、先の高次元な疎分散表現（高次元 SDR）となっている。

さらに、新たに到着した高次元 SDR として表現される新情報が記憶されるに足るかどうかは、既に松本の言うルックアップ・テーブルに蓄えられている多くの高次元 SDR とのベクトル内積計算を実行する形で判断される。そして、両者の内積値がほとんど 0 であればほぼ同一、それが大きな値をとれば新奇と判断される。しかも、このような類似性・新奇性判断が、様々なニューロン群内で並列に実行される。そして、新奇なものであればあるほど、海馬³⁴⁾などのルックアップ・テーブルに新たに記憶される部分が多くなる³⁵⁾。ベクトル内積とは、高校数学で登場したベクトルの同じ成分同士を掛け合わせて合計する操作である³⁶⁾。

34) 正確には、海馬内の自己再帰型配線で大局的に繋がっている CA3 という領域に蓄えられている。また、連想は、より広くは、海馬をも含んだ内部側頭葉記憶システムに代表されるように、大脳新皮質（その自己再帰的な記憶領域は、6 層からなる新皮質中の第 2 層と第 3 層）の各所とも繋がっている。以上は、Kandel 他 (2013), Rolls (2016) を参照。

35) 以上のような高次元 SDR の構造上の特性から、各刺激・イベントに反応した SDR 同士を足しあわせれば、それらの刺激の意味のある和集合も定義できる。例えば、刺激・イベント情報 A, B, C に対応する SDR A, SDR B, SDR C を足しあわせてできた SDR H は、三者を重ね合わせた刺激に対応する SDR になる。甘利 (2008) や (2016) によると、脳内のルックアップ・テーブルでは、情報節約のために、このような重ね合わせ（圧縮）情報/“多重分散記憶”だけが保存・活用される。そして、新たな刺激・イベントが到着して特定ニューロン間の結合ベクトルが形成されると、この新規パターンと上記の情報圧縮された記憶済みパターンとの類似性・新奇性判断が行われる。素人目には、上記のような情報圧縮が行われてしまうと実行される類似性・新奇性判断に混乱が起きそうであるが、そのような混乱を避ける巧妙なエネルギー節約的な仕組みも脳内には備わっている（甘利 (2008) や (2016)）。具体的には、先の高次元 SDR ベクトルの次元が高ければ高いほど、重ね合わせパターンを構成している各々のパターンが互いにはほぼ直交してくることによる（Kanerva (1988)）。確かに、このような直交性があれば、新規の高次元 SDR と情報圧縮された既存の高次元 SDR との内積を取れば、重ね合わされた多くのパターンの中の似通ったパターン以外のパターン・ベクトルとの間の内積値がゼロとなる。その結果、情報圧縮されたパターンと比較した新奇度がすぐに分かるので、類似性・新奇性判断を高速実行できる（甘利 (2008) や (2016)）。進化は、なんというエネルギー節約型の仕組みを生み出したのだろうか。

36) 以上は、Rolls (2007)。

また、五感情報を受容する感覚器とそれらを末梢神経で受けて中枢神経を経由して最終的に脳内のニューロン群に到着した時にどれほどまで高次元 SDR 化されているのか、脳内ニューロン階層間ではどのような高次元 SDR が上下の階層でやり取りされて抽象度が上昇・下落して行くのか、等々については、門外漢のため、現段階では十分に調べ切れていない (Olshausen (2004) などは参照)。特に、ネットワークのネットワーク (networks of networks)、そのまたネットワーク…といった形で各種の刺激・イベントが統合 (トップダウンで下がってくれば分割) されて行く仕組みは、結び付け問題 (binding problems) と呼ばれ、neuroscience 的にも試行錯誤的な形でしか解明が進んでいないという (Anderson (2017), Rolls (2017))。なお、多くの NM-AI には、先の類似性・新奇性判断機構に加えて、程度の差はあるが、上記のような結び付けの仕組みも組み込まれている (Eliasmith (2013), Rinkus 他 (2016) などに例示)。

以上では触れなかったが、現行の NM-AI によって電力消費量が格段に低下するもう一つの大きな理由は、TrueNorth を含む NM-AI の場合、刺激・イベントが到着した場合だけにチップが駆動する形 (Event-Driven / イベント駆動型) であること、そのために半導体回路が非同期回路となっていること、の 2 つである。ヒトを含む生命体は非同期で動いているが、現在主流のノイマン / チューリング型コンピュータは、ほぼ全てが同期回路、つまり、回路全体に同一のクロックを行き渡らせる形で各回路ブロック間の同期を取る形の回路を用いている。このようなことから、同期回路の消費電力は、通常の動作状況では、非同期回路に比べて格段に大きい³⁷⁾。

37) 太田 (2017) によれば、この同期回路方式こそ、von Neumann (1945) が、McCulloch and Pitts (1943) にヒントを得て EDVAC に組み込んだ重要なメカニズムだという。この辺りは極めて深遠なので、ちょっと長くなって申し訳ないが、太田 (2017) からの次のような引用を提示しておきたい。「一つのニューロンにおける論理演算 (加算演算) は、複数のニューロンが同期して発火することによってもたらされる同期したシナプス入力によって行われる。論理演算を終えたニューロンは、再び他のニューロンと同期的に発火して次の回路へと同期したシナプス入力が渡される。この同期的発火の連鎖の成立は、すべてのニューロンの入力においてシナプス遅延時間が一定であるという仮定に依存している。この仮定に注目して、現代のコンピュータの理論的基盤を作り、マカロック・ピッツの業績を世に広めたのがフォン・ノイマンである。フォン・ノイマンは、真空管素子をニューロンに、素子間の二値パルス信号の伝達をシナプス伝達とみなした。マカロック・ピッツに倣って、パルス信号の遅延時間はマイクロ秒オーダーで一定に制御された。これによって複数ラインのビット信号が同期化し、数と命令が表象された。この同期的表象の元で、メモリから読み込まれた命令によって

加えて、現行 NM-AI 自体にも、さらなる低消費電力化を達成するための工夫が導入されつつある。例えば、TrueNorth を含めた前述の多くの NM-AI のニューロンには、高速だが消費電力と設置面積が共に大きな SRAM (Static Random Access Memory) が使われている。SRAM は電気が供給されないとメモリが消えてしまう揮発性メモリ (volatile memory) であるが、不揮発性で SRAM に比肩する速度を誇り設置面積もより小さな ReRAM (抵抗変化型ランダムアクセスメモリ) や MRAM (磁気抵抗メモリ) などが活用できるようになると、消費電力量をさらに下げることができる³⁸⁾。

2-3 NM-AI 実用化を阻むもう一つのボトルネック：大容量“連想メモリ”という難題

NM-AI の BD-AI に対する卓越性は、前節で言及したように、桁違いの低消費電力性だけに留まらない。起因の特定に必須の短期・長期にわたる時系列構造の組み込み易さ、それらを活用したヒト・レベルに近づき超えて行けるほどの“インテリジェンス”実現にも繋がっていく。そして、先の Tavanaei 他 (2019) などに示唆されている発火型の非線形伝達関数に基づいたビッグデータ処理というイノベーションが実現すれば、ヒトをも越える“インテリジェンス”を誇る NM-AI 実用化の時代が間近にやってくるような期待を抱いてしまう。

実際、前節でも、BD-AI に比べて質・量共に豊富な樹状突起を持つ発火型の NM-AI であれば、高い先読み精度などによって BD-AI をはるかに上回る“インテリジェンス”を組み込める可能性について言及した。事実、ヒト並の起因の特定能力が実用的なレベルで獲得できれば、各ニューロンには、頻繁に到着する刺激・イベント情報の関連ニューロンへの伝達機能だけではなく、短期・局所的のみならず長期・大局的な記憶機能と学習プロセス自体を俯瞰しながらの学習 (メタ学習) 機能を組み込むことができる可能性が高まる。もちろん

処理内容を分岐させることが可能な逐次処理システムすなわちコンピュータが実現された。」

38) http://nice.sandia.gov/documents/2015/Marinella_150225%20Nice%202015%20HAANA%20Final.pdf や <https://www.tohoku.ac.jp/japanese/2016/12/press20161214-03.html> などを参照。ちなみに、Samsung は、2019 年前半にギガビット前後の容量を持つ MRAM の市販を開始している。https://www.mram-info.com/samsung-starts-shipping-28nm-embedded-mram-memory?utm_source=feedburner&utm_medium=email&utm_campaign=Feed%3A+mram-info+%28MRAM-Info%3A+MRAM+tech+news+and+resources%29

ん、その際には、先の松本他 (2003) に言及されている海馬などにあるルックアップ・テーブルが必須である。

さらに、NM-AI に上記の多彩な記憶機能とメタ学習機能を組み込めるようになれば、リアルタイムで変化する巨大なルックアップ・テーブルを活用・再生成しながら、松本他 (2003) の“出力依存学習”が実行できるようになる。そして、そのようなことができるようになれば、NM-AI にとって、リアルタイムでの高度なゼロ・ショット学習 (未経験事例への転移学習: Zero-shot Learning) やワン・ショット学習 (小サンプル学習: One-shot Learning) が現実のものとなる。

ところが、ヒト・レベルに近づき超えて行く NM-AI 実用化には、そうは問屋が卸さないもう一つのボトルネックが厳然として存在する。実用化を阻む大きなボトルネックとは、デジタル革命の担い手である現行の半導体デジタルメモリ技術である。というのは、短期・局所的のみならず長期・大局的な記憶機能を組み込んだ巨大なルックアップ・テーブルを作り上げることは、既存の半導体メモリではとても難しいからである。そして、このボトルネックが解消されない限り、ヒトに比肩する上記のゼロ・ショット学習やワン・ショット学習の実用化など、NM-AI にとっても夢のまた夢なのである。

なお、このような海馬などにある巨大なルックアップ・テーブルの組込によってヒト・レベルに近づき超えて行く NM-AI を実現しようとするアプローチは、連想メモリネットワーク (Associative Memory Network) とか高次元計算論 (hyperdimensional computing), ベクトル記号基本設計論 (Vector Symbolic Architecture) などと呼ばれており、NM-AI とは異なる扱いがされているケースも多い。例えば、Boden (2016) では、「この一派は、自らを AI 研究者とも呼ばないので、隠れた AI 研究グループだ」と紹介されている。ただし、連想メモリネットワークの BD-AI に比較した卓越性に関しては、その脳機能上の重要性から、脳神経科学分野の標準的教科書である Kandel 他 (2013) でとても詳しく紹介されている。

また、Intel Lab の下記 URL では、このような高次元 SDR を確率変数ベクトルとして扱う方法論 (Hyperdimensional Computing) を、NM-AI と区別して Probabilistic Computing (確率コンピューティング) と呼んでいるようである。ただし、NM-AI の本質は、決定論的なデジタル性ではなく曖昧さを許すアナログ性にあるとすると、その行き着いた先では、インテルの意味での NM-AI と

確率コンピューティングとが融合していくのだと思われる。敢えて言えば、前者が学習・訓練、後者が推論に相当するからである。

<https://www.intel.com/content/www/us/en/research/neuromorphic-computing.html>

<https://www.intel.co.jp/content/www/jp/ja/research/neuromorphic-computing.html>

半導体デジタルメモリ技術上の大きなボトルネックの存在を理解するには、現行のノイマン／チューリング型コンピュータにおけるメモリの扱い方を理解することが重要だと思われる。この種のコンピュータは、整数値などで指定されるアドレス（番地）とそのアドレスに記憶されているコンテンツ（内容）を明確に分離したアーキテクチャに基づいて長足の進歩を遂げてきたからである。この点を少し例示してみよう。

初期の 8 ビットのノイマン／チューリング型コンピュータの場合、最大で 2 の 8 乗ビット = 256 ビット分のシステムメモリ領域を扱えるが、この 256 ビットには識別のためのアドレスが定められている。例えば、アルファベットの整数 a には十進法の +127 という数値が割り当てられていて、コンピュータが丁度この a の content（したがって、+127）がプログラム実行上必要になったとしよう。その場合、コンピュータのポインタと呼ばれる機能に a のアドレス（特定の番地）が提示されると、127 に対応する (+) 符号あり 8 ビット表現 (01111111) が呼び出される。そして、現行のコンピュータでは、どのような複雑な数値計算であっても、最終的には 1 次元の二進法表現された数値同士の演算に帰着する。

このようにアドレス（番地）とコンテンツ（内容）が分離されている理由は、現行のコンピュータでは、先のフォン・ノイマン・ボトルネック (VNB) の所で触れたように、プロセッサとメモリとが分離されていることに深く関わっている。そして、プログラムに書かれた命令・データを実行する際には、1 次元の 2 進法演算を実行する限り、アドレスを指定してメモリ部分から命令・データというコンテンツをプロセッサに呼び出す方式の方が格段に低コストで効率も良い。微細化によってコンピュータの演算速度が急速に増大する状況では、特にその便益が高まる。

他方、プロセッサ・メモリ融合・同居型のメモリベース・アーキテクチャを

基本とする NM-AI では、先の松本が教えてくれた実際の脳神経システムとほぼ同じようにふるまう。具体的には、到着した感覚・刺激情報は、高次元 SDR 形式の検索情報コンテンツ＝アドレスとして学習・記憶済みのルックアップ・テーブルに提示され、環境適応に必要な類似性・新奇性判断が引き出される。そして、類似性が高ければ転移学習が、新奇性が高ければゼロ・ショット学習やワン・ショット学習が作動し始める。

ところが、類似性・新奇性判断の際にアドレスとコンテンツとが分離されている方式のルックアップ・テーブルを使っているのは、1次元の2進法演算実行ステップ数が莫大になるので、恐ろしく時間がかかってしまう。例えば、先に触れたベンチャー企業 BrainChip Holdings の BrainChip は、筆者の知るかぎり、現在唯一の実用化されている NM-AI であるが、このチップは現行 PC に備わっている PCI-Express スロットに刺して利用される。したがって、BrainChip 側で行われているのは、実ニューロンに近い STDP 機能を備えた発火型の人工ニューロンを使った学習・訓練だけである。そして、このような学習・訓練結果は PCI-Express バスを通じて PC 内の HDD や SSD 等に蓄えられ、様々な推論用途に活用される (Vab Der Made 他 (2017))。また、Intel の NM-AI である先の Loihi も、上記の学習・訓練部分だけが同チップ内の Neuromorphic Processor で行われ、推論は同一チップ内に組み込まれた同社の PC 用 X86 プロセッサ・コア部分で行われる (Davies 他 (2018))。

したがって、両 NM-AI 共に、ルックアップ・テーブル的な機能は、アドレスとコンテンツとが分離されている旧来のフォンノイマン/チューリング型のコンピュータに依存している。そのため、先に引用した Kwabena Boahen @Stanford 大学グループの「パーソナルコンピュータは、マウス規模の脳モデル (=250 万個のニューロン) をシミュレートする際に、4 万倍 (400 ワット対 10 ミリワット) もの大きな電力を必要とするにもかかわらず、実際のマウスの脳よりも 9 千倍遅い。」となってしまうという限界は、十分に解決されているわけではない。

NM-AI の諸機能を高速に実現するために必要になるのが、アドレスとコンテンツが同一の連想メモリ (CAM: Content Addressable Memory)³⁹⁾である。CAM の

39) 文献によっては、連想メモリ (AM: Associative Memory) と CAM とを区別している。例えば、Rusch その他 (2018) では、前者ではアドレスとコンテンツの部分一致が通常であるが、

場合、コンテンツがアドレスそのものになっているので、感覚・刺激入力情報としての高次元 SDR に関連した環境適応に相応しいコンテンツをルックアップ・テーブルから高速に読み出せる。しかも、内容的に近い記憶同士を場所的にも可能な限り隣接する形で配置しておけば、メモリアクセスがさらに高速になる。このような類似記憶の近接性こそ、特定分野の知識・ノウハウを芋づる方式で数多く記憶しているプロが、一を知って十を知るといような素人をあつと言わせる連想力を披露できる理由でもある (Brogliato (2014))。

加えて、関連記憶の近接性便益を活用できる CAM は、現行のノイマン／チューリング型コンピュータでは達成できない耐障害性 (graceful degradation) という極めて有用な特徴をも備えることができる。例えば、現行のスマートフォンや PC などのコンピュータの中では、アルファベットの小文字列 c は、c: 01100011 と表現されている。このとき、後ろから 3 桁目の数値が何らかの原因 (例えばアルファ線などの宇宙線) で 0 から 1 に誤って変わってしまったとしよう。そうすると、g: 01100111 となり、まったく異なった文字 (g) になってしまう。そして、計算プロセスの中で一箇所でもこのような誤りが起きると、計算全体にまったく意味が無くなってしまう。これに対して、先の CAM 活用型の NM-AI では、少い箇所での誤りであれば、扱える高次元 SDR の次元が上げれば上がるほど、かけ算や足し算が行われているだけでなので、処理結果への影響が無視できるようになる。

このように良い所だらけの CAM であるが、現状の半導体デジタルメモリ技術で CAM を実現しようとする、製造上とても高価になり、消費電力もかな

後者では両者が完全一致したものと規定している。ただし、本論では、Kanerva (1988) や (2014) に従い、CAM と AM とは同一だと見なす。実際にも、0 - 1 の識別に加えて一部分を無視する機能 (Don't Care (ドントケア) と呼ばれる) を持つ TCAM (Ternary CAM) が一般化してきているので、コンテンツ・アドレスとして全部を使うか一部を使うかは選択可能である。NM-AI 用を含めた CAM の現状に関する包括的なサーベイは、Karam 他 (2015) を参照した。なお、Karam 他によれば、この論文執筆時点の CAM の最大容量は 100 メガビット = 12.5 メガバイトと報告されており、フラッシュメモリや DRAM に比べると極端に少ない。ただし、最近では、日本発の Axonerve (<https://axonerve.jp/>) のような大容量・高パフォーマンス・低価格のハードウェア TCAM が登場してきている。開発責任者の大塚寛治教授 @明星大学によると、ソフトウェア CAM を二桁以上上回る 1ms (千分の一秒) の処理速度でのワンショット・ラーニングが可能になるといふ。この技術は、NM-AI/BD-AI の世界が一変させる可能性を秘めているのではないだろうか。

り大きくなってしまふ。CAM 式ルックアップ・テーブルに格納されているメモリと到着した感覚・刺激情報との比較のたびに、すべての記録済みメモリとの並列同時比較が必要になるので、回路構造がとても複雑になるからである。したがって、メモリ容量をたやすくは増やせない（以上は、Sharma (2003) 参照）。そのため、現状では、CAM は、未だ高速 CPU のキャッシュメモリ⁴⁰⁾とか高速なネットワーク・ルーターなどのごく一部に小容量でしか組み込まれていない⁴¹⁾。

にもかかわらず、ヒト・猿・マウス並の“インテリジェンス”を誇る NM-AI 実用化のためには、少なくとも 1 万次元を越える疎な分散表現 (SDR) で表されるアドレス空間を高次元 CAM で実現することを試みる必要がある (Kanerva (2014))。このことは、ヒトのピラミッド型ニューロンの場合、各ニューロン当たり 5 千-2 万個の樹状突起と 4 万個のシナプスが付随していることから類推できる。ちなみに、この各ニューロン当たりのシナプス数は、猿では 2 万個、マウスでは 8 千個、ミツバチで 1 千個だった。

このような高次元 SDR を CAM で扱うとなると、誰もがそのあまりに高い技術上の壁に立ちすくんでしまふ。事実、1 ビットの 1 万次元だとしても、 2^{10000} ビットという天文学的なアドレス空間内で高次元 SDR のベクトル内積処理等を実施しなければならない。そのようなことは、とても実現できそうもない。事実、現在の iPhone・iPad などの最新 64 ビット OS (iOS) ですらも、 2^{64} ビット (16 エクサバイト=1600 万テラバイト) までのシステムメモリしか扱えない。しかも、現行の実用的なスマートフォンでのシステムメモリの容量は、

40) 「CPU とメインメモリーのデータのやり取りを高速化するために、CPU 内部に設置されたメモリ」 (<https://kotobank.jp/word/キャッシュメモリ-2703>)

41) このような CAM の難しさに関して、松本他 (2003) や Anderson (2017) は、次のように言及している。「脳型コンピュータの能力決定においては、実時間処理可能な人工神経細胞の結合数が 1 つの重要な要素である。例えば、視覚パターンの網膜レベルでの特徴抽出では、1000 万程度の結合数が必要であり、大脳新皮質での視覚パターン認識では、数 10 億から 100 億程度の結合数が必要であろうと思われる。従って、脳型コンピュータの工学実現には、工学実現のイメージ化、その要素技術の工学モデル化のみならず、シリコン半導体技術の革新的開発を伴うことが必要である。後者では、従来のプロセッサ主体型の半導体技術と共に、脳型コンピュータの為にメモリ主体型半導体技術の飛躍的進展が必要とされる。」(松本他 (2003)), 「(NM-AI 実現にとって) 最も技術的に難しい問題はメモリである。数テラバイトのメモリが必要になるはずだし、この量のメモリを (半導体) デバイスに統合するのは難易度が高い。」(Anderson (2017), 筆者訳, カッコ内は筆者追加)

せいぜい 4~8 ギガバイトである。

したがって、NM-AI は、少なくとも、現行の半導体デジタルメモリ技術を前提とする限り、ヒト並みどころかミツバチ並みの“インテリジェンス”さえ、100 年経っても実現できそうにない⁴²⁾。となると、20 ワット前後で動作しているヒトの脳内でどのような仕組みで CAM 風の仕組みが実現されているのか？興味深い問いであるが、現状はほとんど見当がつかない状況のようである (Natarajan (2018))。とすると、いったい全体、高次元 SDR を多用しているヒトを含む生命体の脳では、類似性・新奇性判断などを実行する際に、どのような仕組みで高次元 SDR を駆使した記憶の呼び出し・書き込み処理が実際に行なわれているのだろうか？

ここで登場する一つの有力な現行デジタル技術依存型のアイデアが、Kanerva (1988) の Sparse Distributed Memory (SDM) という小脳からヒントを得たコロンブスの卵的な提案である。もちろん、ミツバチの 2^{1000} ビットは言うに及ばず、 2^{100} ビットの CAM でも、そのままではソフトウェア的にすら実現が難しい。ところが、例えばだが、100 年間を秒に直すと 3.154×10^9 秒 $\div 2^{32}$ 秒であるから、人生の“実際に起きた”出来事を全て記憶する際には、 2^{50} ビット程度でも多過ぎる。そして、Kanerva (1988) が着目するのは、この点である。つまり、NM-AI での CAM メモリ処理では、 2^{1000} ビットや 2^{10000} ビットの巨大メモリ空間を事前に想定する必要があるとしても、事後に記憶場所 (Storage) として使われるのは、人生 100 年間分でも、事前の想定メモリ空間内の無視できるほどわずかな領域に過ぎない、という事実である。

さらに、高次元 SDR ひとつひとつが独立な正規分布をするような確率変数としての扱うことができれば、事後的に生みだされた SDR データのほとんどが、平均値の周りに存在するはずである⁴³⁾。正規分布では平均値の両側 3σ (=標準

42) もちろん、アナログメモリを使うという選択肢もある。ただし、その際には、エラー発生率の抑制やエラー訂正機能などが必須になるので、そのためのメモリ新素材の選択、メモリ回路の難度・大きさ・省電力性などが問題になる。ただし、その場合でも、対応する NM-AI は、デジタル・アナログ混載の半導体チップとなる。実用化の試みに関しては、Demler (2018) や下記の IBM や Mythic の URL を参照されたい。

<https://www.techspot.com/news/77687-ibm-announces-8-bit-analog-chip-projected-phase.html>,

<https://www.mythic-ai.com/technology/>

43) Kanerva (1988) は、現実的な仮定に基づいて、そのような状況が起きることを数学的に証明している。なお、Maren 他 (1990) によれば、このような直交性に関する数学的証明は、遙

偏差) 内に 99.7%の値が入ってしまうからである。

とすると、Kanerva の計算に従えば、1 万ビットの高次元 SDR で表現される刺激・イベント情報を 1000 万個、あるいは 100 億個を事後的に記憶する場所が必要であるとしても、各々 1 万 x1000 万ビット = 125 億 5000 万バイト = 12.5 ギガバイト ($\approx 2^{34}$)、1 万 x100 億ビット = 1.25 テラバイト ($\approx 2^{40}$) 程度の事前準備で十分となる。最近の Windows 10-PC では個人レベルでも 2 テラバイトの SSD が数万円で入手可能であるが、このレベルだと、高次元 SDR の 100 億個分でもゆうゆう収納できるのである。

このように、高次元 CAM 実現のために事前には天文学的なメモリ空間が必要だとしても、事後的にはその僅かなメモリ空間を使用するだけで済む。実際、Kanerva (1988) によれば、 2^{10000} の高次元 SDR を扱うために巨大なメモリ空間が事前に必要だったとしても、事後的には、現行の半導体技術でも取り扱い可能な 2^{33} や 2^{40} のメモリ空間さえ事前に確保できていれば良い。そして、Kanerva (1988) は、そのことを数学的に証明すると共に、その実用化方法も示唆している。

さらに、Numenta (2017) や Brogliato (2014) は、現行のコンピュータ上にミツバチ並みの 1000 次元や 2000 次元の高次元 CAM を Kanerva 流 SDM としてソフトウェア的に実現している⁴⁴⁾。では、開発段階の NM-AI ハードウェアに関しては、どうだろうか? この点に関しては、門外漢のため十分な評価はできないが、日本の長瀬産業 (<https://www.nagase.co.jp/>) が自社開発・販売している Axonerve があり、32 ビットから 1152 ビットまでの長さの検索キーが可能とされている。しかも、このような連想メモリを使った検索が、FPGA (field-programmable gate array)⁴⁵⁾ と呼ばれるハードウェア上に実装された ASIC とテラ

か以前の 1960 年代の後半に、理化学研究所の甘利俊一氏グループによって既に成し遂げられていたという。ただし、その数学的な難解さを理解できた同分野の専門家が乏しかったとされている。正直、小生には、当該論文がほとんどフォローできなかった。

44) より具体的には、Hash 関数 (<https://ja.wikipedia.org/wiki/ハッシュ関数>) と呼ばれるソフトウェアで実現された連想配列関数が用いられている。明星大学の大家寛治先生によれば、この型のハッシュ関数は、Google などの検索エンジンにも多用されているという。また、ソフトウェア的な連想メモリを使って実際にビジネスを行っている Saffron Technology (https://en.wikipedia.org/wiki/Saffron_Technology, Intel が 2015 年に買収) のような会社もある。そして、同社の検索エンジンは、世界の主要 1000 社で活用されているとされている。

45) 製造後にも回路構成をソフトウェアで自在に再設定できる集積回路。大手ベンダーは

バイト級の SSD を使って極めて高速に実行されているという。そして、もし検索キーのビット数をもう一桁増やせるとすると、高度なゼロ・ショット学習やワン・ショット学習ができることを意味するので、その社会への影響は計り知れないのではないだろうか。

なお、Merolla 他 (2014) によれば、IBM の NM-AI チップ：TrueNorth の場合、一つのチップ内には、4096 個の Neurosynaptic Core と呼ばれるものが含まれている。具体的には、最小単位であるこの Core に 256 個のニューロン、256 個の軸索と 256 個の樹状突起、そしてそれらの各々にシナプスが 256 個付随しており、Core 全体では 256×256 個 (6 万 5536 個) のシナプスが利用可能となっている。そして、例えば、各々のコアの 256 個のニューロン群が SDR を構成していると仮定すると、実際の Core あたりに利用可能なシナプス用メモリは 65.536 キロビット (2^{16} ビット) である。CAM として理論上必要なアドレス空間は 2^{256} であるが、Kanerva 的な SDM の仕組みが導入されていれば、素人判断で誠に申し訳ないが、先の Kanerva の計算例 (2^{10000} 対 2^{33} , 2^{40}) から類推すると、もし TrueNorth に CAM を組み込んだ場合、このような容量で済むのかもしれない⁴⁶⁾。

なお、時代を席卷している BD-AI でも、最先端分野では、CAM の活用を必須とするモデルが登場してきている。例えば、先に触れた Graves 他 (2014) や Graves 他 (2016) の“Neural Turing Machine”と呼ばれる BD-AI では、メモリ容量的にはほんの僅か⁴⁷⁾であるが、海馬内の CA3 領域と呼ばれる巨大な連想メモリ (CAM) 貯蔵庫／装置に倣った外部メモリ機構が組み込まれている。このような CAM 形式の外部メモリが備わっていると、新奇性の高い物事に遭遇した場合でも、そこに過去の記憶の断片に類似したものがあれば、ルックアップ・テーブルと比較対照する形で過去のエピソード全体を迅速に復元し、新規

Xilinx や Intel 傘下の Altera である。

46) また、Davies 他 (2018) によれば、インテルの NM-AI チップである Loihi では、一つのチップが 128 個の Core で構成されており、各々の Core に 1024 個のニューロンが含まれている。そして、各ニューロンでは 1000 個のシナプスが利用可能である。したがって、1024 個のニューロンが SDR を構成していると仮定すると、CAM として理論上必要なアドレス空間は 2^{1024} であるが、実際の Core あたりに利用可能なシナプス用メモリは、16 メガバイト = 1.28 ギガビット ($< 2^{27}$) となっている。

47) 多くが高々メガビット前後容量の CAM なので、2 の 20 乗から 25 乗、つまり 20 次元から 25 次元程度。

な事例と過去の事例との間に共通する類似性を見出すことができるようになる。つまり、低消費電力性の発揮は極めて難しいが、起因の特定の幅と深さを掘げられるので、ゼロ・ショット学習やワン・ショット学習が可能になる⁴⁸⁾。もちろん、そのような学習の幅と深さは、CAM 容量次第である。事実、Graves 他 (2016) には、ロンドンの地下鉄ネットワークを事例としたゼロ・ショット学習やワン・ショット学習の様子が示されている。このように、BD-AI の進化も、高次元 CAM の実用化に大きく左右される⁴⁹⁾。

2-4 長期的には、BD-AI に加えて NM-AI が必須：マクロの視点

BD-AI に加えて NM-AI が強く望まれる根本的な理由は、宇宙物理学者である Chaisson (2013) の下図が説得的である。Chaisson は、実データに基づいて宇宙規模での諸活動の長期的な Energy rate density (amount of energy per unit time per unit mass : 単位質量・単位時間当たりのエネルギー不変消費量) を計算している。そして、この図には、縦軸にその Energy rate density が、そして、横軸に現代 (10⁰) から 100 万年前まで遡る形での時代が示されている。

図によれば、我々の社会は、第一次産業革命前後から過去 100 年超にわたって傾きの急な特定のエネルギー効率 (対数) 直線上に沿って進化してきている。もちろん、BD-AI も、このエネルギー効率直線上に乗っている。他方、NM-AI が実用化できれば、このような傾向線からの人類史上 2 回目の下方へのシフトが可能になる。NM-AI の活用により、少なくとも 4~5 桁高いエネルギー効率の達成が可能になるからである。

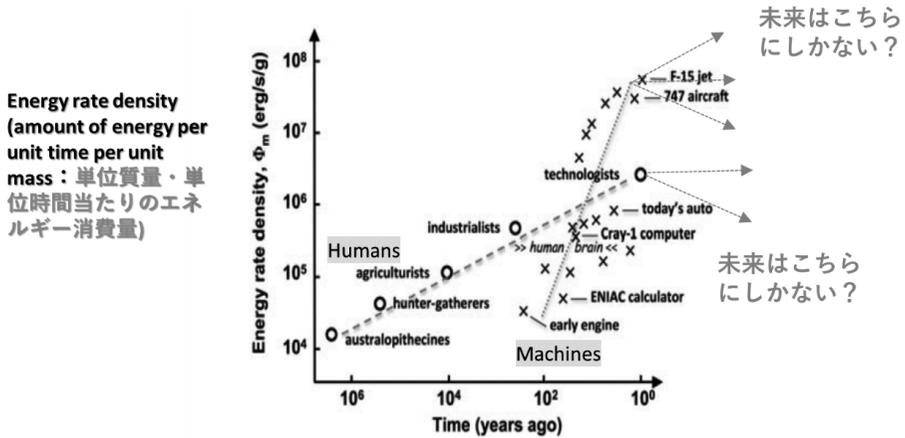
なお、Chaisson は、上記の過去 100 年超の傾向に現れている現象を Technological Singularity (TS) と呼んでいる。今流行りの Kurzweil (2004) 流のシ

48) Ke 他 (2018) は、RNN での起因の特定を行うために、潜在的には数千あるいは数百万ステップの過去に遡ることのできる BPTT (back-propagation through time) と呼ばれるアルゴリズムの改良版 (SAB: Sparse Attentive Backtracking) を提唱している。ただし、そのメカニズムは、筆者には未だ十分に理解ができていない。

49) このことを象徴的に示しているのは、IBM が Watson を駆使して目指している AI Doctor の実用化が现阶段で苦渋の道を辿りつつあるという事実である。その本因は、一言で表現すれば、人間のドクターが得意とする Zero-Shot/One-Short/Few-Shot Learning ができない点に尽きる。詳細は下記を参照されたい。

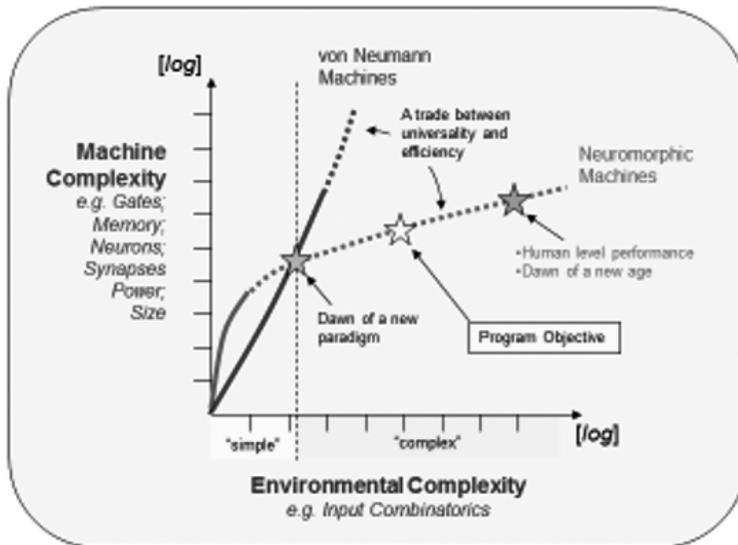
<https://spectrum.ieee.org/biomedical/diagnostics/how-ibm-watson-overpromised-and-underdelivered-on-ai-health-care>

ンギュラリティ（教）には歴史的証拠が乏しいが、TS には十分な根拠があることを強調するためである⁵⁰。



Chaisson (2013) より抜粋, 矢印・日本語書込は筆者

実際、このような現実を前にして、計算機科学の専門家達の多くも、DARPA・SyNAPSE プロジェクトの下図に示されているように、現行のノイマ



(米国 DARPA・SyNAPSE プロジェクト報告書から)
https://www.theregister.co.uk/2011/08/18/ibm_darpa_synapse_project/

50) なお、より一般向けの視点からの Kurzweil 批判に関しては、Zarkadakis (2015) が説得的である。

ン・チューリング型コンピュータでは、そのエネルギー（消費電力）効率の低さから、近未来社会の莫大なコンピューティング需要には到底応えられないと予測している（2015 年の ITRS Emerging Research Device (ERD) Meeting での Srinivasa (2015), Roy (2015), Burr (2015) など参照）。言い換えれば、そのような未来の世界・社会は、現行半導体のエネルギー効率の延長線上では、とても実現できそうもないのである。

このように将来期待の大きい NM-AI であるが、冒頭で述べたように、明日の AI を担うとされるがなかなか実用化までに至っていない。この点に関しては、IBM の Almaden 研究所で NM-AI 実現を模索するグループに属する Burr (2015) の指摘が印象的である。彼は、近未来社会の莫大なコンピューティング需要にどの程度の計算能力で (“What”), どのような方法で (“How”) 対処すべきかという問いかけをしている。彼によれば、前者に対しては、近未来の計算需要は、IBM-Watson を含むノイマン／チューリング型ハードウェアの延長線上ではとても実現できないこと、ところが、現段階ではその先の実現手段が未だ見えていないとする。他方、後者に関しては、“脳機能が十分に解明されていない”現状ではあるが、脳のような格段のエネルギー効率を誇るコンピュータによってしか近未来はやってこないのではないかと主張している。つまり、What, How の両方の視点から考えても、その行き着くところは脳模倣型のコンピュータであろうという見解である。その大きな理由は、前節でも既に指摘したように、脳内での処理が、前述した時空間に広がる非同期・イベント駆動型という特徴に加えて、アダムスミス流の分業・協業の利益を実現するために高度に階層モジュール化されたニューロン／グリア間を電磁氣的・化学的に繋ぐネットワーク処理・決定システムの構築、そのような処理・決定システムの確率過程化、確率過程化された処理・決定システム自体の並列分散処理化、といった低消費電力化に徹した巧妙な特徴を持つことによる。

なお、最近では、開発段階ではあるものの、Intel の Configurable Spatial Accelerator (CSA) や Graphcore の Intelligence Processing Unit (IPU) などに代表される BD-AI と NM-AI の良いとこ取りを狙ったハイブリッドなバルク並列処理型 AI も登場してきている。この方式では、GPU 等と比べて同じ消費電力で 2 - 3 桁の計算能力とエネルギー効率の向上が見込めるといふ⁵¹⁾。なお、バルク並列処理型とは、プロセッサとシステムメモリを近接させた 1 つの

ブロックをバルク (Bulk) と呼び、1000 個を越えるプロセッサを含む各々のバルクで同時並列処理 (Compute Phase と呼ばれる) を行い、バルク間で同期が必要な際には同時並列処理をすべて停止してバルク間での情報交換処理 (Exchange Phase) だけを交互に行う方式のことである。したがって、What ベクトルと How ベクトルの双方を組み合わせているという意味で、上記の Burr 流脳型コンピュータへの進化経路を辿る選択肢の一つだという解釈が可能である。

3. BD-AI 及び NM-AI の“インテリジェンス”とは？

AI 関連の文献には、インテリジェンスの定義を行動主義的な視点から与えようとするものが数多い。例えば、AI がチェスや将棋／囲碁で人間のチャンピオン達を遥かに凌ぐことができる、高速道路のみならず一般道路でも自動運転ができる、通販サイト・融資の窓口、果ては医療の現場でプロフェッショナルからの質問に的確に答えることができる、という風に。実際、大勢の人間中心主義な思考を好む人々は、AI (コンピュータ) にヒト並あるいはヒト以上のことができるか否かを問うことを好む。

そして、そのような問いの正当性を主張するために、人類が生んだ大天才の一人である数学者・コンピュータ科学者 Alan Turing が提示した下記のイミテーションゲーム風の“チューリング・テスト”を引き合いに出すことが常套化している。

<チューリング・テスト> 「もしも、コンターが人間の質問者を欺して自分も人間だと思込ませることができれば、その際は、定義から、コンピュータは知性的でなくてはならない」(引用は Sloman (1995) から、筆者訳)

ところが、世界的に著名な AI 研究者／哲学者である Sloman (1994) が的確に指摘するように、このような人間行動主義的な視点に基づいた問いを発する人々は、生命体・人工物にかかわらず、“インテリジェンス”とはそもそもどの

51) 以上は、Morgan (2018), Anadiotis (2018) を参照。

ようにして生み出されるものであるかに関する本質的な議論を飛ばしがちである。計算機科学の泰斗 Joseph Weizenbaum 自らが 1960 年代半ばに創ったチャットボット“Eliza”の当初には意図していなかった人々からの大きな反響に当惑してしまったように、彼らに「人間だと思ひ込ませる」手段は数多く存在すると分かっているのに、である⁵²⁾。

本論にとっても、人間行動主義的なインテリジェンスの定義に拘ってはいは、AI と HI との補完・代替性という現実的な課題を論じる際に先に進めなくなってしまう。この点を理解するために、上記の Sloman (1995) の的確な指摘を以下に引用してみよう。

「我々が他人の心の状態を判断する際の手がかりとなるのはせいぜい行動だから、人によっては、心の状態や動きに関する事柄が単にシステム全体の行動能力や傾向によって定義可能だと考える。同じく、(Dennett 流の) ‘意図的な態度’は、心の内部の動きを無視している。もしインテリジェンスが行動の生まれ方に依存するとするなら、これは間違いに違いない。(チューリング自身はあまりに賢かったので彼のテストをクリアできるインテリジェンスとか理解力の尺度として提示できなかった。彼は、単にそれ (=チューリング・テスト) を、巨大な‘ルックアップ・テーブル’ (機械的な探索指示表／表引きテーブル) を使わないで実現できる筈だと彼が考えた技術的な挑戦課題として提示したに過ぎない。) (筆者訳, (=) の部分だけ筆者追加)

Sloman の指摘するように、BD-AI であれ NM-AI であれ、両者のインテリジェンスを人間行動主義的な視点から定義することは、動く標的を追い求めるようなものであるから、極めて危険であると再認識できる。さらに、Sloman が上記で述べているように、人間行動主義的な視点に立脚する限り、イルカやシャチ、鯨などの鯨類、ゴリラやチンパンジーなどの霊長類や鳥類、はてはミツバチなどの昆虫類にもヒトに比肩する高度な“インテリジェンス”が備わっていることを扱えない。また、異なった生活圏に適応してヒトとは異なった驚く

52) Weizenbaum の述懐は、Weizenbaum (英訳 2015, 独語オリジナル 2006) の「Eliza Today」の章に詳しく記されている。また、本書に展開される Minsky の AI 賛美論に対するアンチテーゼは傾聴に値する。

べき高度な行動パターンを示すこれらの動物たちに共通する“インテリジェンス”の特性についてさえ言及できない。彼らの思考方法や行動パターンがヒトと大きく異なっていることは事実であるが、彼らがヒトと共通する高度な“インテリジェンス”を保有していることは誰しも認めている。

例えば、McFadden 他 (2014) には、渡り鳥のヨーロッパコマドリが、伏角コンパス (inclination compass) と呼ばれる地表に沿った方位としての「極向き」を決定する量子力学的な仕組みが組み込まれた特殊なコンパスを使って移動することが紹介されている。しかも、このヨーロッパコマドリは、磁場の水平成分を使って北の方角を判定するロブスターやサケなどが持っている極性コンパス (polarity compass) と呼ばれるコンパスと上記の伏角コンパスを状況によって切り替えているらしい⁵³⁾。人間中心主義の浅薄さを実感させられる事実である。

したがって、NM-AI の HI へのインパクトを考えるには、NM-AI が実現しようとしているインテリジェンスとはどのようなものであるかに関して、人間行動主義的な視点からではなく、そもそも NM-AI が実世界で起こる物事や事象をどのように捉えようとしているのかという内側の視点から検討することが必須である (Hawkins (2006))。その理由は、「(世界を) 理解しているか否かは外側から見える行動では測定できない」(Hawkins (2006)) からであり、「我々の心がどのようにして動いており、どのように上手く動かなくなるか、適正に動かなくなるかについて十分に理解するには、設計思想 (アーキテクチャ) ベースの概念が重要である」(Sloman (1994), () 内は筆者追加) からである。

では、NM-AI のインテリジェンスを、具体的にどのようにして外部に現れた行動に頼らず内部のメカニズムとして定義すれば良いのだろうか？ 先の Hawkins (2006) や Hawkins の元共同研究者である George (2008) は、そのような試みを、大脳新皮質の階層構造の意義・意味に着目しながら、次のように明瞭で小気味良い定義を提示していて注目に値する⁵⁴⁾。

その要点は、NM-AI の主目的は大脳新皮質の作動原理に習うことであるか

53) <http://www.natureasia.com/ja-jp/reviews/highlight/10618>。

54) Hawkins は NM-AI のソフトウェア的な実践で世界をリードしている開放型ビジネス形態の Numenta 創設者、Hawkins の共同研究者でもあった George は Google/Facebook/Amazon 等の名だたる投資家から巨大な開発資金を獲得している閉鎖型ビジネス形態の Vicarious 創設者でもある。

ら、NM-AI のインテリジェンスを大脳新皮質が生み出すと考えられるインテリジェンスに限定してはどうかという提案である。限定する大きな理由は、大脳新皮質を持つのは哺乳類だけだからとしている。そのために、彼らは、まず大脳新皮質が進化の結果として階層構造をしている意義・意味に着目する。傾聴に値すると思われるので、ちょっと長くて申し訳ないが、彼らの主張を引用してみたい。ちなみに、彼らが大脳新皮質に限定するもう一つの理由は、意識上・意識下の情動（後述）などの仕組みを解明して実装・実用化することは夢のまた夢と考えているからでもある。

「大脳新皮質には世界に関するモデルが備わっているのだから、あなたは、世界について考え、その中で動き回り、将来について予測したりすることができる。あなたは、生まれながらにして言語や家や音楽のことを知っているのではない。大脳新皮質には賢い学習アルゴリズムが備わっているのだから、どのような階層構造であろうと自然に見いだせるし把握できる。（大脳新皮質の中に）構造がないと、我々は混乱状態、あるいは大混乱状態にさえ陥ってしまう。」（Hawkins (2004), 筆者訳、()内は筆者追加）

「学習に関する No-Free-Lunch (NFL) 定理によれば、全ての学習問題に関して他を圧倒する生まれながらの秀逸性を誇る学習アルゴリズムは存在しない。もしあるアルゴリズムが特定の問題に優れているとすると、それは、そのアルゴリズムが当該問題に適した仮定を有効に使っているだけのことである。他方、Neuroscience の成果によれば、大脳（新）皮質は、視覚・聴覚・体知覚のような異なった感知に際して同じアルゴリズムを使用しているようである。また、大脳（新）皮質が階層構造をしていることも良く知られている。このような観察結果に関する尤もらしい説明は、異なった領域や感覚種からのデータが、それらの表面上の違いにかかわらず、基本的に同じ統計的な特徴を持つことができるということである。その理由は、世界に関するデータ生成メカニズムが、物理法則や自己組織化の法則にしたがって基本的に階層構造をしている筈だからである。もしそれが本当たるとすると、階層構造に基づいた学習の効率性の背後に潜む極意を、世界の構造を学ぶことによって解き明かすことができることになる。」（George (2016), 筆者訳、()内は筆者追加）

したがって、Hawkins (2006) や George (2008) を拡大解釈すれば、大脳新皮質が階層構造をしているのは、第一にヒトや他の動物を取り囲む実世界が各種の物理・化学法則に支配された階層構造をしていることの反映である。第二に、そのような実世界の中で発生する物事や事象により効果的に適応しながら生き抜くためには、脳を持つ生命体にとって、実世界の階層構造の主観的な写し絵としての脳内階層モデルを、自他のアクションによって生みだされる体験・経験に基づいて日々改良改善していく必要があるからである。そして、第三に、Engel 他 (2016) を援用して拡大解釈すれば、それらの脳内階層モデルの予測値との対比という形で変化と異常にいち早く気づいて対応し、さらに、そのような気づきや対応の良否判断に基づいて脳内モデルの予測精度をたかめるような階層モデルの改良改善を行う必要があるからである。

その意味で、NM-AI が実装するインテリジェンスとは、短く表現すれば、次のように定義できる。「インテリジェンスとは、人間のような行動によってではなく、(大脳(新)皮質内の)階層メモリ構造の予測能力によって測定される」(Hawkins) 筆者訳)、「インテリジェンスとは、過去との類似性によって将来を予測する脳の潜在的な能力である。」(Hawkins (2006), 筆者訳)。

つまり、NM-AI に実装されるインテリジェンスとは、その完成形においては、Hawkins-George に倣って、冒頭で提示した本論の“インテリジェンス”=「実世界の変化と異常に対応していくために、自他の過去の記憶と現況に立ち向かう自らのアクション(行動)とを活用しながら、様々な活動の“起因の特定(Credit Assignment)”/予測を行う能力ならびにその自己変化能」と定義できるのである。そして、このインテリジェンスの定義は、ヒトであろうとなかろうと、そして、生命体であろうとなかろうと関わりなく使用できるので、Slowman (1995) 流の AI = “自己変化する情報駆動型制御システム”を特徴付けるインテリジェンスとしての意義・意味を持つことができると考えられる。

なお、自己変可能が上記の“インテリジェンス”の定義に不可避免的に付随するのは、個々のニューロン/ニューロン集団内外に張り巡らされた強弱を持つ配線構造パターンにおいて、類似性・新奇性判断が行われるたびにシナプス結合の重みが自動的に変化するからである。そして、そのことによって、新たな世界/社会/環境モデルに基づいた各自の予測能力が生み出されていく。つまり、

自己変可能とは、先の松本 (2006) が教えてくれるように、「脳が答えを引きだす (出力する) と、引きだした答のアルゴリズムは、出力依存性学習によって、自動的に書き変わる」というメタアルゴリズム (アルゴリズムのアルゴリズム) の特性そのものだと言える。

このような出力依存学習に特徴付けられる“自己変化する情報駆動型制御システム”の自己変化能について、Creps (2017) は、20 世紀初頭に活躍した現代の脳神経科学者とも見間違えるほどの見識を誇るフランスの偉大な哲学者アンリ・ベルグソンの発想を紹介しながら、下記のような興味深い事実を指摘している。

「だから、この過程 (詩篇の暗記仮定) が繰り返されるそれぞれの時間は、記憶しようとしているのが一人の人間であるとき、一個人の歴史にだけある瞬間であり、その個人だけの記憶に存在し、そのようなものとして想起できる。例えば“p”という文字でつかえた時のこととか、二行目を忘れた時、等々。文字“p”のところまできて、それを再認すると、今では過去になってしまっている“p”でつかえた、その個人だけの瞬間を思い出す。この詩篇はひとたび暗記されれば、素晴らしい精度で復唱されることができるかもしれないが、そのプロセスを担った人間は自分自身が、アルゴリズムが絶対に真似することのできない方法で、変化しているのだ。」 (Creps (2017), 203-204 頁)

「アルゴリズムがそのパターン認識の精度を向上させていくときに、各実行について記録していくためのログファイルは、いわば連続的に存在していて、リアルタイムにアクセスすることができ、コンピュータ・ストレージにおける“今”の中にトークンとして記録される。しかし、アルゴリズムは目の前にあるものと記録されたものを照合することしかできない。これは人間における再認経験と比べればはるかに劣るものである。」 (Creps (2017), 204 頁)

確かに、少なくともヒトの場合、完了型としての暗記事項そのものだけでなく、時間軸にそったその試行錯誤過程も記憶として保持される。Creps (2017) の言うように、前者はコンピュータ上で操作対象となる Storage として

の記憶だが、後者は文脈なども付加された連想するための仕掛けとしての記憶 (Memory) でもあり、起因の特定の際の有力な助っ人にもなる。そして、BD-AI では Storage としての記憶しか扱えないが、ヒトや完成形としての NM-AI では、CAM さえ十分に利用できれば、コンテキスト情報に基づいて上記のログファイルとの照合も可能なので、両タイプの記憶保持が可能となる (Plate (2003))。

なお、Hawkins (2006) や George (2008) が大脳新皮質という時には、海馬と視床とを含めている⁵⁵⁾。ところが、脳内で最も Neuroscience 研究が進んでいるとされる海馬でさえ、NM-AI 実用化に必須の機能解明が十分に進んでいない (Kohara 他 (2014), 利根川進グループ (2017), Hawkins 他 (2018) など参照)。さらに、Sherman 他 (2013) によれば、視床の機能解明は、海馬のみならず大脳新皮質などに比べても、未だほとんど進んでいない。とすると、海馬・視床・大脳新皮質の緊密な結び付け問題 (binding problems) に関わっているヒト並みの大容量 CAM が可能とする連想記憶の NM-AI への実装は、今後、想像以上に困難を極めるのかもしれない。

4. ヒューマン・インテリジェンス (HI) と BD-AI・NM-AI との違いを探る

前節では、本論が冒頭で提示した“インテリジェンス”の定義 (下記に再掲) の妥当性が、Sloman 流の“自己変化する情報駆動型制御システム”としての広義の AI を前提とする場合でも、結構高いことを確認した。したがって、この節では、本論の定義に基づいて、第 2 節の分析結果をふまえながら HI, BD-

55) 大脳新皮質内での類似性・新奇性判断が行われる時に、後者の新奇性判断がなされると、直ちに海馬に情報が転送されるからである。その意味で、海馬は、類似性・新奇性判断という機能上は、大脳新皮質の上位に位置する (Hawkins (2006))。さらに、6 層構造を成す大脳新皮質では、大脳新皮質内の異なった部位間の情報転送・応答のほとんどが視床を経由して行われる。そのため、視床無しでは、大脳新皮質の階層構造のメリットを活かすことができない。しかも、外部からのセンサー情報も嗅覚以外の全てが視床を経由して大脳新皮質の主に第 4 層に伝えられるし、大脳新皮質からのセンサーへのアウトプット情報は多くが視床を経由して末梢神経まで降りていく。したがって、Hawkins (2006) や George (2008) が指摘するように、大脳新皮質の機能を NM-AI として実現するには、海馬と視床が共に必須なのである。以上は、Rolls (2017) と Hawkins (2006) を参照している。

AI, NM-AI の類似点や相違点を明確にしてみたい。

(本論の定義) “インテリジェンス” = 「実世界の変化と異常に対応していくために、自他の過去の記憶と現況に立ち向かう自らのアクション (行動) とを活用しながら、様々な活動の“起因の特定 (Credit Assignment)”と予測を行う能力ならびにその自己変化能」

三者 (HI, BD-AI, NM-AI) に、何ができて何ができないのかという意味での機能特性の程度を探るには、適当な試金石を活用することによって本論“インテリジェンス”の解像度をもう少し高める必要がある。この節では、そのような試金石として、はじめに、下図に示されるような Pearl (2018) の“インテリジェンス”の三分割法 (“LADDER OF CAUSATION”) に習いたい。そして、その次に、その解像度をさらに高めるために、後述する Minsky (2007) の六分割法を援用したい。ちなみに、Pearl は、ベイジアンネットワーク (Bayesian network)⁵⁶⁾ 及びその発展系として反実仮想推論 (思考実験) までをも自動で組み込むことができる統計的因果推論 (Statistical Causal Inference) のパイオニア・泰斗である。

Pearl は、“インテリジェンス”のレベルを、同図の上から下に向かう形で Association (連想) レベル, Intervention (介入) レベル, Counterfactuals (反実仮想) レベルの三つに分けている。1. の Association (連想) レベルとは、ビッグデータを受身的に“観察”する方式で、「～だったら～する」という“If-then ルール”で示される関数/アルゴリズム⁵⁷⁾を、それらの背後にある因果関係には触

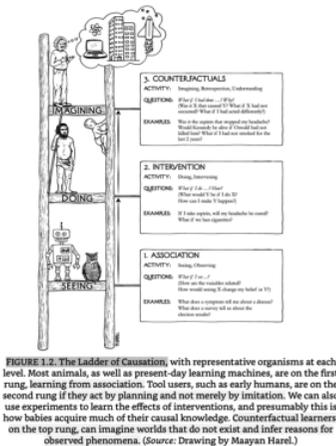
56) 「ベイジアンネットワーク (Bayesian Network) とは、「原因」と「結果」の関係を複数組み合わせることにより、「原因」「結果」がお互いに影響を及ぼしながら発生する現象をネットワーク図と確率という形で可視化したものです。過去に発生した「原因」と「結果」の積み重ねを統計的に処理し、『望む「結果」に繋がる「原因』や『ある「原因」から発生する「結果』を、確率をもって予測する推論手法ともいえます。この考え方は人がさまざまな出来事や他人の振る舞いを予測するときの考え方に倣ったものといえます。近年、IT、特にインターネットがより人間的に使いやすくなってきている背景には、ベイジアンネットワークを活用した推測エンジンの活用が盛んになってきたことがあります。』 (<http://www.hitachi-hri.com/keyword/k052.html>) なお、Pearl (2018) が繰り返し強調しているように、ベイジアンネットワークでは、確率分布に裏打ちされた有向グラフを用いるものの「原因」と「結果」がお互いに影響を及ぼす構造になっているため、因果関係を内生的に決定することはできない。

57) 特定の問題を解く手順を、単純な計算や操作の組み合わせとして明確に定義したもの。BD-AI では、莫大な数の“If-then ルール”に基づいて、当てはまりの良さそうなルールの一つ抽

れないままで推定するレベルである。実際、Pearl (2018) が繰り返し警告しているように、既存ビッグデータを観察するだけで“If-then ルール”の背後にある因果関係を導出することは極めて難しい。ビッグデータであっても、因果関係の同定に必要なすべてのデータが揃っている状況を望むことはできないからである。

したがって、このような欠損データの問題を克服して因果関係を導出するためには、2. の Intervention (介入) が必須となる。そうすれば、観察対象をより深く理解するための積極的な現地踏査や操作実験によって、ビッグデータをそのまま受け身的に活用することの限界を克服できる。現地調査／操作実験という自らの積極的なアクションを起こす際には、鶏が先か卵が先かの問題が残っているが、少なくとも観察データが生み出される様々な因果関係に関する素朴な仮説を創造・保持していることが必要となる。仮説のない闇雲な現地調査／操作実験の繰り返しでは、より説明能力の高い仮説になかなか辿り着けないからである⁵⁸⁾。ちなみに、Pearl (2018) は、“自己変化する情報駆動型制御システム”としての AI/ML に因果関係の特定ができるか否かを問うテストを、先の Turing Test の簡易版だとして“Mini Turing Test”と呼んでいる。

Judia Pearl (2018)による“インテリジェンス”試金石



◆ Pearlの主張：「インテリジェンスの本質は、現象に基づいて因果関係を類推/起因の特定できること」

1. **Association (連想) レベル**では、莫大なデータの“観察(seeing)”に基づき複雑な関数/アルゴリズムを、因果関係を理解しないブラックボックス状態のまま推定
2. **Intervention (介入) レベル**では、“観察”だけではなく観察対象に介入して変化させる（実験する/Doing）形で推定
3. **Counterfactuals (反実仮想) レベル**では、観察データを説明する様々な因果関係を内生的にイメージ、確率的に最も当てはまりの良い因果関係を生成/学習

出すことに対応している。

58) そもそもこのような素朴な初期仮説をどのようにして“自己変化する情報駆動型制御システム”に組み込めるのか、という疑問は残る。この点に関しては、Clark (2016) は、幼児が文法を習得していく自己再帰 (Bootstrapping) 的な仕組みの重要性を指摘している。

三つ目の Counterfactuals (反実仮想) レベルでは、AI/ML が、1. と 2. 双方の観察データに基づいて様々な因果関係仮説を自ら複数作成し、各々の妥当性を反実仮想状況下でも検討することができる。加えて、確率的に最も当てはまりの良い因果関係にたどり着くこともできる。ちなみに、反実仮想とは、抗がん剤 A を使用した癌患者さん達の事例ということであれば、「もしこれらの患者さん達が抗がん剤 A を使用しなかった場合にどうなったのか」とか「もしこれらの患者さん達が抗がん剤 B を使用した場合にどうなったのか」といった想定をすることに相当する。「自分/Aさんが～をしなければ、こうならなかったのに」といった後悔や自己/他者批判をできることも、このレベル 3 になって初めて可能となる。

まず、BD-AI を俎上に上げてみよう。単純な樹状突起機能しか持たず、時空間表現にも弱い BD-AI では、既に前節で確認したように、そもそも冒頭の意味での起因の特定ができない。つまり、「特定の出来事において選ばれた色々なアクション (活動) の中から良好・不良な最終結果に繋がったアクションを関連状況 (コンテキスト) やそこに至った学習方法と共に記憶して、さらなる変化と異常への対応に備える作業」ができない。そのため、ビッグデータ内に潜む因果関係を内生的に同定することができない。

実際、BD-AI で起因の特定に利用できるのは、伝達関数の極めて局所的な微分係数情報しかない (Richard 他 (2018))。そして、当該ニューロンネットワークへの入力値と出力値の間の違い (誤差と呼ばれる) の平均二乗誤差を最小にするような微分係数を生成できる伝達関数の数千万個を超えるパラメータが推定される。その結果、利用可能なビッグデータ全体の中でだけで辻褄が合うような一種の相関関係の特定が行われるに過ぎない。ところが、多彩な起因の特定には、短期・局所的のみならず長期・大局的な記憶機能と学習プロセス自体を俯瞰しながらの学習 (メタ学習) 機能が必須である。この点に関する Rolls (2017) の指摘は、極めて深遠である。

「実際、例えば (BigData 型 AI に多用される) 誤差逆伝播法や強化学習などによって特徴付けられる (神経) ネットワークによって達成される非常に複雑な関数関係は、それらが、学習・訓練後に固定化されたルールに従って作動するというものであり、しばしばとても頑迷で融通が利かない。

対照的に、多段階で一回限りの言語に基づいた計画や手続きを修正するためには、推論や計画の際の複数ステップやおそらくそれらに関連したエピソードを思い起こしたりするだけで、エラーにもっとも繋がりがちな連鎖を識別できるようになる。このことが、記憶を明示的に呼び起こすことができる宣言記憶システムと意識との緊密な関係が存在する理由の一つである。」(筆者訳、()内は筆者追加)

したがって、特定の問題に直面した時に BD-AI に高々できることは、莫大な数の“If-then ルール”をビッグデータから抽出し、内生的に抽出された特徴量⁵⁹⁾に基づいて当てはまりの良いルールを一つ抽出することに過ぎない(Darwich (2018))。この意味で、BD-AI では、アルゴリズムが、あるいは関数近似がすべてなのである。たしかに、BD-AI は、Pearl (2018) 流に表現すれば、利用可能なビッグデータを受け身的に活用しているので、レベル 1 には到達している。ところが、初期時点で素朴な仮説を念頭に置き、この仮説の精緻化(含む変更)のために実世界に積極的に働きかけて必要なデータを自ら生み出せるレベル 2 には到達していない。

もちろん、BD-AI がレベル 1 で達成可能な Performance (能力) は、AlphaGO や Watson に例示されるように、おそろしいほどに高度で実用的なものである。しかも、手持ちの If-then ルール数が膨大になればなるほど、次々に特定分野の能力でヒトをも越えるようになる。そして、これまで大勢の人々が極めて複雑だと思い込んでいたヒトの知覚・認知に関わる活動が、実は、因果関係モデルや理論なしに、単に関数(アルゴリズム)の近似でアッと驚くまでに実現できることを明らかにした。

しかしながら、“If-then ルール”だけからは、たとえその選択肢が莫大な数に達しようとも、起因の特定／因果関係の導出ができないので、ゼロ・ショット学習やワン・ショット学習を実行できるようにはならない。そのため、パフォーマンスと“インテリジェンス”を混同して BD-AI に頼り切ってしまうと、BD-AI と HI との代替・補完関係に関する見極めを誤る可能性が高くなる。言い

59) 特徴量とは、BD-AI が内生的に生み出す分析対象を認識する際に使用する定量的な特徴。このような特徴量には、ヒトにも容易に理解できるものからまったく理解できないものまで数多く存在する。

換えれば、起因の特定／因果関係の導出ができる HI とそれが苦手な BD-AI との関係は、本質的に補完的なものであると言える。

では、NM-AI は、どのレベルの“インテリジェンス”に達することができるのだろうか？ 現行の NM-AI は、発火型伝達関数の制約により、そもそもビッグデータ処理ができないので、レベル 1 にも達していない。では、このビッグデータ処理制約が突破され、高容量の CAM が組み込まれた完成形としての NM-AI は、どのようなレベルまで行けるのだろうか？

レベル 1 の突破は、BD-AI を数桁越えるエネルギー効率で可能である。ただし、レベル 2 に達することは、想像以上に困難を極めそうだ。もちろん、このレベルでは積極的なアクションを行使することが必須なので、NM-AI という人工物であろうと、少なくとも介入を行う際のセンサー系や駆動系（アクチュエーター系）を備えたロボットのような構成物を必要とする。言い換えれば、起因の特定／因果関係の同定が可能な NM-AI は、ソフトウェア的な仕組みだけでは実現できない。

したがって、以下では、完成形としての NM-AI には、身体 (Body) が備わっているとして議論を進めていきたい。ただし、実現できる身体の程度は、本論では、Hawkins (2006) が下記で提案する大脳新皮質で実現できる程度のものであると考える。その理由は、後ほど自己意識や意識上・意識下の情動に関して述べる際に詳しく触れたい。

「まず、人間の心は大脳新皮質だけではなく古い脳の情動システムと複雑な人体によっても創られている。人間になるためには、大脳新皮質だけではなく、その全ての生物学的な機構が必要となる。人間のように全ての問題について語り合う（チューリング・テストに合格する）ためには、実際の人間の経験と情動のほとんど及び人間らしく生きることが必要になる。インテリジェンスのある機械 (AI) とは、大脳新皮質と複数の感覚器官相当部分が必要になるだろうが、その他は任意だ。AI が人間のような身体で足を引きずって歩き回るのを見るのは面白い事だろうが、そのことは、我々が AI に人間のような情動システムや人間のような経験を組み込まない限り、ほんのわずかにも人間のような心を持つことはないだろう。そんなこと（試み）は極端に難しいことだし、私には、まったく的外れに思え

る。」(Hawkins (2006) 筆者訳)，

「インテリジェンス，すなわち大脳新皮質のアルゴリズムと，旧脳の感情的な衝動，つまり恐怖，妄想，欲望などが混同されている。インテリジェンスを備えた機械は，このような感情を持っていない。個人的な野望をいまくことがない。財産も，社会的な評価も，肉体的な満足も望まない。欲張ったり，悪癖を身につけたり，機嫌が悪くなったりすることもない。辛抱強く訓練をすれば，人間の感情をまねた反応ができないことはないだろう。だが，インテリジェンスを備えた機械がもっとも効果的に活用される分野は，生身の人間が知性を発揮しにくい領域，すなわち，特殊なセンサーが必要とされる活動や，退屈に感じられるような活動だ。一般的に，これらの活動には感情の入り込む余地がほとんどない。インテリジェンスを備えた機械は単一用途の地味なシステムから，きわめて強力で超人間的な知能システムまで，多岐にわたることだろう。」(Hawkins (2006), 伊藤文英訳)

たしかに，NM-AI は，多彩で豊富な樹状突起や STDP による時空間に拡がるニューロンネットワークを扱えるので，高容量の CAM が利用可能であれば，リアルタイムでの出力依存学習に加えて，ヒト並みのゼロ・ショット学習やワン・ショット学習を実行できる。その結果，幅と深さに溢れる起因の特定能力をも発揮できる。なお，小容量の CAM でも，これらの学習スタイルが実践可能なことに注意してほしい。

ただし，既存のビッグデータが生みだされるメカニズムへの十分な理解無しには，NM-AI にとって，そもそもどのようなレベル 2 の積極的介入が望ましいかが分からない。また，どの方法が望ましい介入方法であるかについては，介入決定を行う際のセンサー系や介入行動を行う駆動系の正確さや多彩さ，介入の際のコンテキストやタイミングだけではなく，介入主体の主観的な判断にも依存する。

さらに，実数体で表現される実世界へのシンボル操作による介入方法は，連続系／アナログ (A) から離散系／デジタル (D) への変換であるから，無限に存在する。しかも，このような AD 変換・DA 変換を担っているのが脳を含む身

体であるという“身体化された心 (Embodied Mind)”的な解釈⁶⁰⁾が妥当だとすると、レベル 2 の難度はさらに上昇する (Craig (2014), Tani (2016), Jasanoff (2018), Damasio (2018))。ちなみに、冒頭の認知科学領域を席卷しているエナクティブ・アプローチは、“インテリジェンス”の生成・発揮には身体が必須だとしている (Noë (2009), Hohwy (2013), Engel 他 (2016))。

なお、Pearl (2018) は、データに基づいて循環型ではない有向ネットワーク図⁶¹⁾が描ければ、その図を構成する基本パターンを用いて自動的に因果関係の自動導出を可能とするアルゴリズムを開発している。その結果、ロボット的な身体を持つ“自己変化する情報駆動型制御システム”が、このようなレベル 2 の操作の自動化によって、近未来的にもレベル 3 にまで到達できると断言している。もちろん、そのための計算が実用化できるほどの長さの時間やエネルギー効率におさまるかは別問題である。

ところが、そのためには、AI＝“自己変化する情報駆動型制御システム”に行为主体性感 (Sense of Agency), 自我意識 (self-awareness), 自己内省 (Self-reflection), 共感力 (empathy), 自由意志 (free will), 意識 (consciousness) などが必要になると強調している。事実、前述の「自分/Aさんが～をしなければ、こうならなかったのに」といった後悔／批判には、これらの要件が必要になるはずだ。

ただし、これらの高難度の必須要件は、いきなり列挙されており、しかも、各々の項目の意味があまり明確ではない。その結果、残念ながら、NM-AI の“インテリジェンス”特性を本論が望むレベルで識別することができない。この難点を克服して有用性を発揮してくれるのが、Minsky (2007) の“インテリジェンス”に関する 6 階層の概念である⁶²⁾。

60) Tani (2016) や Creps (2017) によれば、メルロ＝ポンティ (Maurice Merleau-Ponty) は、このような AD・DA 変換器としてのからだと心をつなぐ概念としてキアスム (Chiasme) という造語を編み出している。また、驚くことに、Bergson (1903) には、AD・DA 変換器としてのからだという解釈が明示されている。さらに、身体化された心の必要性については、ニューロンレベルの超解像を誇る fMRI 開発で有名な MIT の Jasanoff (2018), 特にその最終稿が圧巻である。

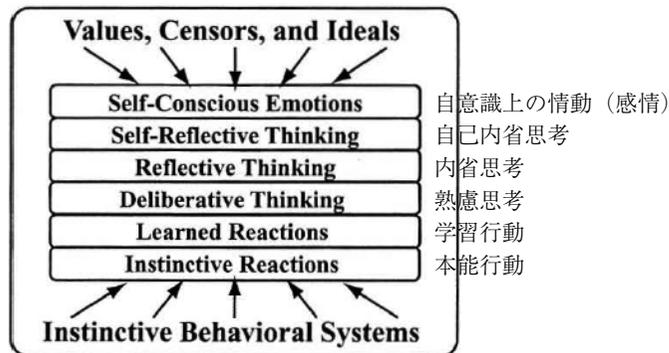
61) 例えば、下記を参照。

<https://understandingsociety.blogspot.com/2016/02/causal-diagrams-and-causal-mechanisms.html>。

62) なお、Minsky (1985) では、情動の役割に関しては触れつつも、それが意識上・意識下のものであるかどうかの区別もハッキリしていない。ところが、「問題はインテリジェントマシン (AI) が情動を持つことができるかではなくて、マシンは情動無しにインテリジェントに

Minsky は、インテリジェンスの要素を、少なくとも下図に示される 6 つに分けて考えることを提案する。具体的には、1) 本能行動 (Instinctive Reactions), 2) 学習行動 (Learned Reactions), 3) 熟慮思考 (Deliberative Thinking), 4) 内省思考 (Reflective Thinking), 5) 自己内省思考 (Self-Reflective Thinking), 6) 自己意識上の情動／内省 (Self-Conscious Emotions/Reflection) の 6 つである。そして、インテリジェンスの高さは、本能行動→学習行動→熟慮的思考→内省思考→自己内省思考→自己意識上の情動という風に上昇していくとする。

図中の本能行動とは、DNA 等に事前に組み込まれている行動パターンであり、直ぐに理解できる。一つ上の学習行動とは、自他の色々な経験を通じて獲得した“If → then rule”をさす。次のレベルの熟慮思考とは、いくつかの採るべき本能・学習行動を思い浮かべ、それらの効果について比較考量する能力だと規定されている。この規定を文字通りに解釈すれば、先の Pearl のレベル 3 と大差がないように思える。ところが、どうも、このレベルでは、意識上ではなく意識下で作動する“インテリジェンス”を想定しているように思える。



Minsky (2006) のインテリジェンスに関する 6 階層

Minsky には申し訳ないが、この熟慮思考レベルに関する彼の叙述は、かなり混乱している。例えば、ある箇所では、「いくつかのとるべきアクションを考慮し、各々の効果を心に思い浮かべ、その後それらの選択肢を比較する」とか、「各々のアクションがもたらすであろう結果を比較するある種の方法が必要である」（いずれも筆者訳）とあるので、意識的な行動が前提とされている

なり得るかである。」(筆者訳) と述べている。そして、ほぼ 10 年後の Minsky (2008) になると、“Machines cannot be intelligent without any emotions” (この文章は Minsky (2008) には無い筆者の文章) という確信に変わっていて大変に興味深い。

ように思える。実際、比較考量のためのシミュレーションは、意識しなければできない筈である (Hohwy (2016))。

ところが、Minsky は、別の箇所では、内省思考レベルに至ってはじめてレベル 3 の熟慮思考の結果を意識的に思い浮かべられると述べている。そして、内省思考が現れるのは、予測が当たらなかつたり、障害に遭遇したり、必要な知識にアクセスできなかつたりしたときだと追記している。ということは、内省思考が必要なレベルとは、変化と異常への対応ができなくなったときということになる。したがって、以下では、熟慮思考レベルは、意識下での“インテリジェンス”だと見なしたい。さらに、Minsky は、敢えてこの意識に“自己”を含めていない。したがって、ここでの意識とは、第三者的なメタの視点という程度のものだと類推できる。

内省思考のワンランク上のレベルが自己内省思考であるが、この二つの区別も、文字通りにはなかなか難しい。ただし、両者の大きな違いは、自己が登場するということに加えて、自己内省思考が、「最近の思考について考慮するだけでなく、そのような思考をした主体についても思考する」(筆者訳) こととか、後者の際に、自己を第三者的な視点からモデル化して(自分の中に複数存在する)自己の動機とか目的とかを巨視的に自己評価できること、などとしている。そのため、Minsky の自己内省思考レベルでは、自分の採った行動が自分の理想とするもの(目的)に値するか否かについて思考することが可能である。したがって、この自己内省思考レベルになってはじめて、先の Pearl の行為主体性感 (Sense of Agency) や自我意識 (self-awareness)、自己内省 (Self-reflection) を必要とする彼のレベル 3 に等しくなると考えて良いようだ。ということは、このレベルに達すると、はじめて Pearl の意味での因果関係を、反実可能状況をも思い浮かべながら導出できるということになる。

Minsky では、この自己内省思考レベルの上に、さらに自意識上の情動レベルをもってくる。彼が説くその大きな理由は、変化と異常に直面した際に、情動に起因する自分の中に複数存在する自己モデル⁶³⁾間の変更や複数自己モデル

63) 人々のパーソナリティは、喜怒哀楽等々の情動モードの種類やそれらの大小によって、平常時のもの(主人格)から大きく乖離する。さらに、そのような情動モードの変化によって人々が“いの一番”に注目する事柄や気になる事柄、それらの事柄へのアクセス速度すらも大きく変化する。Minsky (2007) や Ornstein (1986)、Franklin (1995) は、このような様々な状況

への新メニュー追加を行えるレベルが必要だということにある。自己モデルの変更や追加ができるのは、自分の採った行動に対する良心の呵責、プライド、羞恥心、傲慢な心等々の感情／情動の湧出やそのことによって価値観／目的自体を維持・打破・変革することができるからである。その結果、意識上の情動（感情）は、複数自己間の柔軟で高速な自動切り替え装置としても作動する⁶⁴⁾。

さらに、情動（感情）モードに応じた自己モデルを複数保有・追加できるということは、ほぼ同じ要領で他者モデルを複数保有・追加できるということになる。そして、そのことによって、コミュニティ形成に必須の心の理論（Theory of Mind：他者の心の状態、目的、意図、知識、信念、志向、疑念、推測などを推測する心の機能）を各自が組み込めるようになる。つまり、Minsky 流の柔軟で高速な視点切り替え装置としての情動は、Pearl が先に彼のレベル 3 の要件の一つとして列挙していた共感力（empathy）を媒介に、互いの多様な意図や常識・文化等々の共有によって広範囲な協力を生み出すために必須のコミュニティを生み出すためにもなくてはならないものにもなる⁶⁵⁾。

以上の考察から、Pearl のレベル 3 には、多少の拡張解釈が必要なながら、Minsky のレベル 4, 5, 6 が渾然一体となった形で含まれていると理解できる。

下で発生する情動モードチェンジは、複数自己の考え方との親和性が極めて高いとする。たしかに、我々は、大脳新皮質内に蓄積されている様々なエピソード記憶や意味記憶に意図的にアクセスしようとしても、きっかけがないとなかなかイメージ通りのものにアクセスできない。実際、我々のエピソード記憶は、情動記憶とも呼ばれるように、自らの情動モードによってアクセス速度やアクセス可能な幅と深さが大きく異なってくる。さらに、Rolls (2017) が教えてくれるように、大脳新皮質内に蓄えられている意味記憶には、特定の記憶断片を辿っていくだけでは辿り着けないものが多々存在する。意味記憶のネットワークには、局所的に閉じた部分が存在するからである。ところが、エピソード記憶が蓄えられている海馬内の C3 領域は、全体が大局的に結びついているので、記憶の断片を手がかりに全体復元が可能である。しかも、色々な意味記憶は、Minsky (2007) が意味ネットワーク (Semantic Network) と呼ぶ形で情動記憶と芋づる方式で繋がっている。したがって、情動モードの違いによって、呼び起こされる意味記憶自体も大きく左右される。

64) この点に関する Neuroscience に基づいた説得的な説明は、Rolls (2018) が圧巻である。例えば、Rolls は下記のように述べている：「有限の長さのニューロンネットワーク内においてニューロン発火に起因する確率的な揺らぎによって起こる意思決定プロセスの確率的な振る舞いは、同じようなインプットに対してであってさえも、異なった時には異なった決定やメモリ連想が行われるといったことを含む数多くの利点をもっている。」「したがって、脳内には唯一人の意志決定者が存在するのではなく、異なったタイプの意思決定には多彩な意思決定プロセスが存在する。」(Rolls (2018) 第 11 章から、筆者訳)。

65) この点は、Rolls (2018) にも、Neuroscience 的な視点から繰り返し強調されている。

したがって、“インテリジェンス”の解像度を鮮明にするという本論の主要目的にとって、Pearl の 3 段階基準に加えて、Minsky の 6 段階基準を用いることには十分な意義がある。

さらに、Pearl 基準と Minsky 基準の整合性についての理解は、Pearl の三段階基準と冒頭で触れたエナクティブ・アプローチには深い関係があることに気づくことによってより鮮明になる (Hohwy (2016) 参照)。実際、Pearl のレベル 1 では、「利用可能なビッグデータを受動的に受け入れる = 実世界の動きに受動的に対応する」ということが含意されている。そして、レベル 2 では、自らのアクションによって実世界に積極的に働きかけることによって、起因の特定に基づいた因果関係の同定にまで踏み込もうとしている。ただし、自らの積極的なアクションを駆使して起因の特定／因果関係の同定を行うためには、アクションを行う側に行為主体性 (Agency) やそのような主体的行為を行っているという自らの感覚 (行為主体性感: Sense of Agency) が必要になってくる。起因の特定／因果関係の同定には、自らのアクションの是非を内省的に比較考量しなければならないからである。

また、そのような高い自己内省力を伴う行為主体性感があれば、実際に行ったアクションだけではなく、仮想的なアクションをも含めた選択肢の比較考量を事前・事後にできるようになる。例えば、事後的に「あのアクションをとらなければ良かった」と後悔するには、実際のアクションの結果と仮想的なアクションの期待される結果を比較できなければならない。そして、このような反実仮想レベルのインテリジェンス無しには、実世界に適応するための因果関係モデルに基づく予測力を素早く、そして十分に発揮できない。

では、上記の Minsky 流の 6 段階のインテリジェンス基準では、完成形としての NM-AI の“インテリジェンス”をどのように特徴づけることができるだろうか？

Pearl (2018) の試金石に基づいた分析から、すでに Minsky のレベル 1 と 2 はクリアしている。また、Minsky の熟慮思考レベル 3 になると、意識下でいくつもの採るべき本能・学習行動を思い浮かべ、それらの効果について比較考量できる。この点に関して、NM-AI は、経験事例の転移学習のみならずゼロ・ショット学習 (未経験事例への転移学習) やワン・ショット学習 (小サンプル学習) を可能にする起因の特定能力を持つので、このレベル 3 はどうにかク

リアできると考えられる。判断に迷いはじめるのは、内省思考レベル 4 である。その最大の理由は、このレベルが必要になるときは、予測が当たらなかったり、障害に遭遇したり、必要な知識にアクセスできなかつたりしたとき、より広くは、変化と異常が発生したときという規定に関わってくる。しかも、Minsky によれば、このレベル 4 は、意識上で作動する。繰り返しになるが、この時点の意識とは、第三者的なメタの視点という意味だと解釈できる。

では、Minsky の 6 段階のどの辺りまでクリアできれば、変化と異常を発見し、加えてその対応ができるようになるのだろうか？この問いに答えるには、そもそも変化と異常とはどのようなものであるかを短く説得的に提示する必要がある。本論では、それを 1980 年代末当時の AI 研究のトップスターでありながらその限界の前に突如 AI 研究から身を引いた Winograd 他 (1987) が “Breakdown”⁶⁶⁾ と呼ぶ下記のような状況に重ね合わせてみたい。

「(コンピュータプログラムとして組み込まれた) インテリジェンスの本質は、(同プログラムが対処しなければならない) 問題が単純に事前定義されていない形で起きる時や解決策を模索すべき状態空間が定まっていない時に適切に行動できることである。問題空間内での合理的な探索は、そのような空間自体が作りだされるまで可能ではないし、(既存の) 形式構造がそのような (未知の) 状況に事実上どの程度対応しているかでその有用性が決まる。」(Winograd 他 (1987), 筆者訳, () 内は筆者追加)

Winograd 他 “Breakdown” は、AI/ML という視点から考えると、そのプログラム設計者が事前に想定していた対応すべき問題に関する考察の系 (システム化領域: systematic domains⁶⁷⁾) を凌駕する問題が発生した時に起こる。その意味で、“Breakdown” とは、「脳がメモリしているアルゴリズムの先読みでは予測しえないような事柄」(松本他 (2003)) ということになる。

たしかに、このような “Breakdown” = 変化と異常の発見や対応には、前述の

66) Winograd 他 (1987) では、この点を正当化するために、フッサール以降の現象学の泰斗ハイデッガーまでを引き合いに出している。

67) もっと正確には、「システム化領域とは、プログラムが (適切に) 機能する状況との当該プログラマーの解釈を具現化した領域」(Winograd 他 (1987), 筆者訳) とされている。

ゼロ・ショット学習やワン・ショット学習が必要になる。ただし、これらの学習能力は、あくまでも変化と異常への対応ができるための必要条件に過ぎない。

では、変化と異常への対応は、先の Minsky のインテリジェンスに関する 6 段階のどの辺りからできるようになるのだろうか？まず、1) 「本能行動」、2) 「学習行動」までのインテリジェンスでは無理である。このレベルでは、“If → then rule”メニューにない対応はできないので、変化と異常の発見すらできない。また、3) 熟慮思考レベルでは、1 部の発見はできるかもしれないが、限定的なものに留まる。というのは、選んだ行動の妥当性に関する判断ができないからである。このような判断ができるのは、4) 内省思考レベルである。ところが、このレベルには、第三者的なメタの視点という意味での意識の存在が前提とされる。

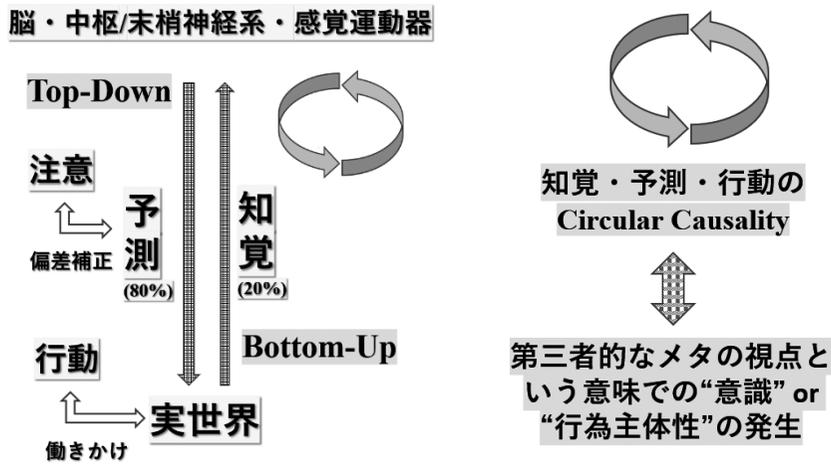
なお、NM-AI にとって、内省思考レベルのインテリジェンスが備わっていれば、選んだ行動の妥当性に関する類似性・新奇性判断が可能なので、変化と異常の発見はできそうである。事実、類似性・新奇性判断によって新奇だと判断した場合には、(人工)海馬システムや(人工)視床システムと連繋してエピソード記憶として蓄えることができる。さらに、自他の過去の記憶と現況に立ち向かう自らのアクション(行動)とを活用し、様々な活動の起因の特定もできる。しかも、ヒトの五感を遥かに越えるセンサー等を組み込むことによって、ヒトを超える変化と異常への発見をもできる可能性が出てくる。

では、そもそも NM-AI に第三者的なメタの視点という意味での意識を近未来において組み込める時代が来るのだろうか？あるいは、現段階において、このレベルでの意識を NM-AI に組み込もうとするどのような試みがあるのだろうか？素人である筆者には、そのような問いに十分に答えることはできないが、認知発達ロボティクス領域⁶⁸⁾におけるロボットに第三者的なメタの視点とい

68) この分野の世界を代表する浅田(2017)によれば、「認知発達ロボティクスは、構成的手法を用いて人間の認知発達過程を理解することを目的とする。核となるアイデアは、物理的埋め込み(身体性)と社会的相互作用で、それらは、他者を含め環境との相互作用を通じて情報を構造化する」としている。構成的(constructive)という意味は、実際にモノとして造ってみせるということである。浅田グループによる赤ちゃんロボット等の構成的な試みについては、<http://www.er.ams.eng.osaka-u.ac.jp/asadalab/>を参照されたい。なお、この場をお借りして、筆者の2回の聞き取り調査に御協力いただいた大阪大学大学院工学研究科知能・機能創成工学専攻の浅田稔先生、ならびに浅田先生への御紹介をいただくと共に一緒に聞き取り調査に御参加いただいた大阪大学大学院情報科学研究科名誉教授(現京都情報大学院大学教授)の

う意味での意識を実際に組み込もうとする Tani (2017) の試みは注目に値する。その基本アイデアは、Tani が“Circular Causality”と呼ぶ Friston (2010) の Free Energy 原理（下記に説明）に類似したものである。その基本アイデアは、下図のような形で表現できる⁶⁹⁾。

この図の左半分には、上部の「脳・中枢／末梢神経系・感覚運動器」と下部の「実世界」との間に上からの矢印と下からの矢印が示されている。上からの矢印は脳から中枢／末梢神経系を通じて感覚運動器に伝達される予測、下からの矢印は実世界を感覚器で捉え末梢／中枢神経系を經由して脳に届く知覚を意味している。そして、実世界に関する予測にしたがって運動器を駆使した行動を起こすことでも知覚が発生する。しかも、この際に、予測が当たらないと、予測誤差も認知される。予測誤差が認知されると、そのような誤差をできるだけ迅速に解消するための選択肢を抽出するための注意が喚起され、その結果、より精度の高い予測が再準備される。そして、ヒトや脳を持つ生命体は、この上下の予測・知覚ループを素早く回すことで、予測誤差が最小になるように行動していると見なすことができる。ちなみに、生命体とは、このような予測誤



(筆者作成)

今井正治先生に心から御礼を申し上げたい。

69) Rolls (2016), 特にその第5章第6節第10項では、NM-AIでも可能な（自らの軸索が反回して自らの樹状突起に繋がる形の）自己再帰型の神経回路を二組以上組み合わせると、自己監視 (Self-Monitoring) 機能をシミュレーションによって示せること、そのことによって起因の特定を行うためのメタ学習が可能になること、このような多段階構造が意識の発生に繋がる可能性が高いこと、などが強調されている。

差最小化行動によって自由エネルギー⁷⁰⁾の最小化を図りながら行動しているとするのが、Friston 流 (Minimum) Free Energy 理論の骨子である。

そして、Tani (2017) によれば、図の右半分に示されているように、第三者的なメタの視点という意味での意識や行為主体性 (Agency) は、このような知覚・予測・行動のループが (Minimum) Free Energy 理論にしたがってグルグル回る (“Circular Causality”) 過程で発生すると説く。Tani の場合、このことを脳の代わりにの BD-AI (時空間構造を組み込んだ各種 RNN) ソフトウェア、感覚・運動器の代わりにのセンサー系とアクチュエータ系を備えたロボットを作成し、このロボットに様々な課題を与えて自律学習させていく⁷¹⁾。もちろん、ロボットの行為主体性はある程度まで確認できるが、第三者的なメタの視点という意味での意識は直接観察できないので、ロボットの事後的な行動パターンやその結果の善し悪しからその存在を推し量るという間接的なアプローチを採用している。

残念ながら、筆者は、Tani (2017) のこのような試みの妥当性を評価できるほどの知識を持ち合わせていない。ただし、Windows OS の“タスクマネージャー”のような同一コンピュータ上で動いている各種プログラムを第三者的なメタの視点から可視化する仕組みは既にあちこちに存在している。したがって、Tani の自律型ロボットで試みられているのは、そういうタスクマネージャー的なメタの機能がどのようにすれば自律的に生みだされるかということだと思われる。その意味では、NM-AI に、メタの視点という意味での意識や行為主体性を近未来に組み込むことは思ったよりも容易なのかもしれない。

なお、Rolls (2017) は、上記のタスクマネージャー的なプログラムと Tani (2017) がめざしている“意識”との決定的な違いについて下記のような興味深い指摘をしている。

「コンピュータプログラムの中には、監視機能を持つものもある。この

70) 簡単には、物質または系のもつ内部エネルギーのうち、仕事に変わりうるエネルギーと定義されている。

71) NM-AI の代わりに時空間構造を表現できる RNN が使用されている理由は、NM-AI 用開発言語の乏しさや NM-AI そのものの利用可能性の低さだと思われる。また、このような存在証明を目的とする段階では、低消費電力性に拘る理由がほとんどない。

ことが、より高次の言語的な思考プロセスだと見なすべきだろうか？私のこの点に関する現時点での意見は、それらのプログラムが記号システム内のシンボルがどのようにして実世界の意味と結びつけられるかという問題（記号接地問題）に関与しない形でシステムの動作を修正するための固定的なルールにしたがって作動しているので、そう見なすべきではない、である。」（筆者訳、（）内は筆者追加）

Tani 流の Circular Causality の具現化によって、たとえそれが原始的なものであっても、第三者的なメタの視点という意味での意識や行為主体性を自律的に生み出すことができたとしたら、人工物としての“自己変化する情報駆動型制御システム”に内省思考という新たなレベルの“インテリジェンス”が付け加わることになる。ところが、Tani (2017) のこのような試みが成功裏に終わろうとも、そして、そのような素朴なレベルの意識に加えて自己意識や行為主体性感をも備えた“自己変化する情報駆動型制御システム”が生みだされようとも、Winograd 他の意味での変化と異常への対応は、なかなか実行できそうもない。なぜだろうか？

その理由を明らかにするために、Minsky の 5) の「自己内省思考」や 6) 「意識上の情動（感情）」と変化と異常への対応力との関係を考察してみよう。たしかに、5) の「自己内省思考」レベルでは、自分の採った行動が自分の理想とするもの（目的）に値するか否かについて思考できる。したがって、特定の倫理観や信念・社会規範にしたがった行動が自己目的という視点から実行できるということになる。数式的な表現では味気ないことになるが、特定の目的関数に沿った最適化が可能であるかどうかという風にも表現できる⁷²⁾。

では、先のメタの視点という意味での意識や行為主体性に加えて自己意識や行為主体性感をも備えた NM-AI は、そのアルゴリズムやその自己変化能であるメタアルゴリズムを駆使して、その目的関数（特定の倫理観や信念・社会規範）

72) その意味では、ヒト対ヒト、ヒト対マシンの過去の対戦記録に依存することなく、“強化学習”によってのみから AlphaGo を上回る能力を発揮できるという AlphaGo Zero の場合も、目的関数自体を自ら設定・変更することはできない (<https://deepmind.com/blog/alphago-zero-learning-scratch/>)。なお、強化学習とは ML の一種で、AI が導出した結果を現実の事象に適用する際に、良い結果になったか悪い結果になったかの情報を AI 自信にフィードバックして与えられた目的関数の値がより高まっていくように学習させていく手法。

自体に合致している否かを判断し変化と異常への対応ができるだろうか？なお、メタアルゴリズムとは、先の松本他 (2003) の「脳が答えを引き出す（出力する）」と、引き出した答のアルゴリズムは、出力依存性学習によって、自動的に書き変わる」というアルゴリズムのことである。

その答は、限定的には Yes である。限定的とせざるを得ない大きな理由は、目的関数自体を設定できるためには、6)の「意識上の情動（感情）」が必須であることによる。これは、BD-AI のみならず NM-AI でも、共に大脳新皮質型インテリジェンスしか組み込まれていないので、たとえロボットのような身体を備えていたとしても不可能である⁷³⁾。限定的となるもう一つの決定的な理由は、目的関数自体を変更せずに特定の視点からアプローチし続ける形では、変化と異常への対応ができない場合が数多く存在するからである。

つまり、Winograd 他 (1987) の言葉を借りれば、視点を自在に切り替えることができないと、いかにメタアルゴリズムを駆使しても、単純に事前定義されていない形の事象が起きる時や解決策を模索すべき状態空間が定まっていない時に適切に行動できない。そして、敢えてそれを駆使し続けてしまうと、過学習状態に陥ってしまう危険性を孕む（甘利 (2008/1989), 甘利他 (1997), Fisher (2016)）。言い換えれば、目的関数自体の設定・変更には、Minsky 流の柔軟で高速な視点切り替え装置としての意識上の情動が必須となるのである。

実際、6)の「意識上の情動（感情）」レベルのインテリジェンスでは、特定の

73) この点に関して、Hohwy (2016) は、意識上・意識下の情動が生みだされるためには、たとえそれがロボットのような人工物であったとしても、センサー系・アクチュエータ系による実世界に関する因果関係の外受容的推測 (exteroceptive inference) のみならず、己の体内に関する好不調の原因に関する内受容的な推測 (interoceptive inference) が可能なメカニズムを導入すれば、体内で前述の Circular Causality が回り出すので、そのような仕組みが意識上の情動をも生みだせるのではないかと、そして、そのような情動が心の理論も可能にするのではないかと、文献を提示しながら述べている。同じようなことが Friston (2016) においても、世界的な脳神経科学者である Anil Seth の一連の試みに言及しながら、「内受容的な推測とその己の情動とおそらく他人の情動を推測するという（未だ）想定上の役割」（筆者訳）と述べている。つまり、両者共に、生命体であれ人工物であれ、自律神経系の組込とそのフィードバック・フィードフォワード機能の活用により情動の発生が可能になることを示唆している。たしかに、Damasio (2018) や Craig (2016) の脳神経科学的な情動研究でも、意識上・意識下の情動と自律神経系、特に後者の外受容と内受容、そして体内深部の骨格や骨格筋などからの固有受容 (proprioceptive) とは密接に連係していることが明らかになってきている。したがって、NM-AI への情動の組み込みも夢物語ではないのかもしれない。

倫理観や信念・社会規範の変更さえも自律的に可能なので、世界／社会／環境状況に応じて一端設定した目的関数自体を変更することができる。6)レベルのインテリジェンスが利用可能であると、自閉的な状況に陥りがちな単一自己ではなく、新視点を新たに組み込んで自己メニューを拡大し、さらに既存の複数自己間の柔軟な切り替えが可能になるからである。

さらに、変化と異常への対応に必要なことは、目的関数の設定・変更だけではない。対応の難度が増せば増すほど、社会を構成する様々な人々が時空をまたいで織りなしてきた社会ネットワーク構造の中に埋め込まれている共有知識をより深くより幅広く総動員する必要性が高まる。そのような社会化された知識は、社会ネットワークとして数多くの HI に埋め込まれている（分散表現されている）ので、それらを速やかに動員するためには、6)の「意識上の情動（感情）」レベルの“インテリジェンス”が必須となる。そのためには、自閉的な自らの殻を破って社会ネットワーク（コミュニティ）内で協力し合って対応策を探索できる“インテリジェンス”が不可欠となるのである。そして、それを可能にするのが、Minsky (2006) が教えてくれる、互いの多様な意図や常識・文化等々の共有によって広範囲な協力を生み出すコミュニティ形成装置でもある意識上の情動なのである。

この点に関する Winograd 他 (1986) からの下記の引用は、変化と異常への対応には社会ネットワーク内での対話による新しい状況の識別に必須な新概念の誕生（記号創発）が必須であることを強調していて、極めて示唆的である。

「以前の諸章でも指摘したが、人間は、同じ類いの盲目性にずっと閉じ込められ続けることはない。我々は、新しくコミットする（責任を持って身を委ねる）ことができる活動領域に移ることで Breakdown（変化と異常）に対処する潜在力を持っている。合意するのは、表現するものと表現されるものとの間の固定的な関係に関してではないのだ。合意するということは、話し手と聞き手とが（互いに）コミットして（そのような固定的な関係に）新たな差異を見つけ出せるように（各々が）思考可能な全ての範囲にわたって対話することなのだ。」（筆者訳、()内は筆者追加）

言い換えれば、Minsky の第 6 層まで具備した“インテリジェンス”では、も

はやそれらが個々の“自己変化する情報駆動型制御システム”内だけに留まるのではなく、社会を構成する様々な“自己変化する情報駆動型制御システム”の間の社会ネットワーク構造の中にも埋め込まれるようになる。個々人の“インテリジェンス”，したがって HI が脳を含む身体ネットワーク上にサブネットワークとして分散表現されていると同じように，HI は社会ネットワークの中で大勢の人々の間にも分散表現されている。そして，そのような社会ネットワークとして HI の威力は，ヒトが 6)の「意識上の情動（感情）」レベルの“インテリジェンス”を保有することによってはじめて発揮できるのである。その意味で，ALVA NOË (2009) や河野哲也 (2006)，浅田稔 (2017) の言葉を借りると，「(ヒューマン) “インテリジェンス”は，個々人の脳を含む体の外にも広がっている」。

このように拡大解釈すると，変化と異常への対応力を発揮するための“インテリジェンス”には，Minsky の第 6 層までが必須となるのである。そして，Minsky 的な Emotion Machine が夢物語である限り，HI と BD-AI，NM-AI は，本能行動・学習行動・熟慮思考・内省思考では代替関係をさらに強めていくかもしれないが，“インテリジェンス”が自己内省思考，その次の自意識上の情動（感情）と上昇するにつれて，より顕著な補完性を強めて行くのではないだろうか。

なお，最後になるが，Hohwy (2016) は，意識上・意識下の情動が生みだされるためには，たとえそれがロボットのような人工物であったとしても，センサー系・アクチュエータ系による実世界に関する因果関係の外受容的推測 (exteroceptive inference) のみならず，己の体内状態に関する主に自律神経系経路の内受容的な推測 (interoceptive inference) が第三者的な視点から観察可能なメカニズムを導入できれば，体内でも前述の Circular Causality が回り出すので，そのことが自意識上の情動（感情）をも生みだせるのではないかと，そして，そのような情動（感情）が心の理論も可能にするのではないかと，文献を提示しながら述べている。同じように，Friston (2016) も，世界的な脳神経科学者である Anil Seth の一連の試みに言及しながら，「内受容的な推測とその己の情動とおそらく他人の情動を推測するという（未だ）想定上の役割」（筆者訳）について言及している。つまり，両者共に，生命体であれ人工物であれ，自律神経系の組込とそのフィードバック・フィードフォワード機能の活用により情動（感

情)の発生が可能になることを示唆している。

たしかに、Damasio (2018) や Craig (2016) の生命体に関する研究でも、自律神経系、特に後者の外受容と内受容、そして体内深部の骨格や骨格筋などからの固有受容 (proprioceptive) の三者は、意識上・意識下の情動と密接に連係していることが明らかになってきている。その意味では、実装方法の解明が進むことにより、情動の NM-AI への組み込みも夢物語ではないのかもしれない。そして、そのようなレベルでは、HI と NM-AI との代替関係が支配的になっていく可能性もあるのかもしれない。ただし、Damasio などの研究成果に関しては、Rolls (2016) のように否定的な見解を強烈に述べている脳神経科学者もいる。

4. 結びに代えて

本論では、今をときめく BD-AI と中長期的には AI の本丸として登場すると期待されている NM-AI の双方を取り上げ、そもそも BD-AI や NM-AI の“インテリジェンス”特性とはどのようなものであるのかを、Human Intelligence (HI) と両タイプの AI との補完性・代替性に焦点を当てながら検討した。今をときめく BD-AI に加えて未だ実用化に至っていない NM-AI を同列に扱ったのは、5-10 年後に後者の実用化が活発化しはじめると、既に社会に大きなインパクトを与えている BD-AI との補完性を高めつつ、中長期的には NM-AI が AI の中核となっていくと予測されるからである。

そして、現行の BD-AI よりも 6 桁以上の消費電力効率を誇る NM-AI が実用化の域に到達するようになると、あらゆる人々が肌身離さず身につけているスマホのような形で人々の生活になくしてはならないものとなって行くだろう。そうなってくると、HI に有史以来の衝撃的なインパクトを与えはじめることは間違いない。起因の特定／因果関係の同定に秀でた NM-AI が、その高度なリアルタイム学習、ゼロ・ショット学習 (未経験事例への転移学習：Zero-shot Learning) ワン・ショット学習 (小サンプル学習：One-shot Learning) の機能を発揮することにより、人々の日常生活／業務で直面する変化と異常への対応力やその自己変化能を格段に高めてくれるからである。しかも、そのようなインパクトは、広く大衆レベルに及んでいくはずである。

上記のような視点から本論がまず試みたのは、BD-AI と NM-AI の脳神経模倣に関する設計思想の詳しい検討と、そのような設計思想の違いがもたらす“インテリジェンス”特性の違いを明らかにすることであった。分析結果としてやや衝撃的なことは、BD-AI には冒頭の意味での起因の特定ができないため、ビッグデータ内に潜む因果関係を内生的に同定することができない事実の再確認だった。ところが、手持ちの「～だったら～する」という“If-then ルール”の数が膨大になればなるほど、因果関係モデルや理論モデルなしに、これまで大勢の人々が極めて複雑だと思い込んでいたヒトの知覚・認知に関わる活動までが、実は、ヒトがアツと驚くまでのレベルに達することのできることも同じように衝撃的だった。ただし、この衝撃的な BD-AI のパフォーマンス獲得には、現行のフォン・ノイマン／チューリング型コンピュータの泣き所であるフォン・ノイマンボトルネック (VNB) に起因する驚くほどの高消費電力性という対価を支払う必要のあることも確認した。

次に確認したのは、このような限界を持つ BD-AI に比べ、一見ではあるが、NM-AI がいかに魅力的なものであるかであった。具体的には、VNB の桎梏から解放されることによる桁違いの低消費電力性、多彩で豊富な樹状突起や STDP (spike timing-dependent plasticity) が可能にする短期・長期にわたる時空間構造の組み込み、短期・局所的のみならず長期・大局的な記憶機能と学習プロセス自体を俯瞰しながらの学習 (メタ学習) 機能、多彩な起因の特定能力がもたらす高度なゼロ・ショット学習やワン・ショット学習機能、等々の BD-AI が足下にも及ばない魅力の数々である。

ただし、このような NM-AI の潜在力は、発火型のニューロン伝達関数に起因するビッグデータをなかなか上手く扱えないという弱点により、少なくとも現時点においては、NM-AI の魅力が絵に描いた餅のような状況になっていることも確認した。さらに、そのようなビッグデータ制約をどうにか打破できたとしても、アドレス (番地) とコンテンツ (内容) が一体化したメモリである高次元 CAM 無しには、ヒト・レベルに近づき超えて行けるほどの“インテリジェンス”の実現も夢物語に終わる可能性の高いことを確認した。そして、現行の半導体メモリ技術を持ってしても、高次元 CAM の実現可能性がとてつもなく低いことも確認した。なお、分析の中では、このようなビッグデータ制約や CAM 制約などのボトルネック解消のイノベーションが起りつつあること

にも具体的に触れた。

最後に確認したことは、本論の“インテリジェンス”の定義の正当性を議論した後に、今をときめく BD-AI とビッグデータ制約も CAM 制約もクリアした完成形としての NM-AI のインテリジェンスは、HI とどのような補完・代替関係にあるのか、であった。この現代的な課題を明らかにするために、ペイジアンネットワーク及び統計的因果推論のパイオニア・泰斗である Pearl (2018) の“インテリジェンス”に関する 3 段階モデルと AI の父と称される Minsky (2006) の同 6 階層モデルを“インテリジェンス”評価の試金石として援用した。

分析によれば、BD-AI は、Pearl (2018) が教えてくれているように、彼のレベル 1 (Association レベル) にしか達していないことを、主にその人工ニューロン特性を精査する形で再確認した。ちなみに、このレベルでは、ビッグデータを受身的に観察し、「～だったら～する」という“If-then ルール”で示される莫大な数のアルゴリズムを、それらの背後にある因果関係には触れないままで推定するだけである。つまり、起因の特定／因果関係の同定は、少なくともヒトに理解できる形では全く行われていない。

一方、上記の二つのボトルネックを解消した完成形としての NM-AI も、多彩な樹状突起や STDP などの時空間表現力に優れた人工ニューロンが組み込まれているものの、未だ Pearl のレベル 2 (Intervention レベル) を辛うじてクリアできるか否かの状況であることを確認した。このレベルでは、“If-then ルール”の背後にある因果関係を導出するために欠損データを、NM-AI 自らが実世界に介入して収集できなければならないからである。なお、このレベルでは積極的なアクションを行使することが必須なので、NM-AI は、介入を行うために必須のセンサー系や駆動系 (アクチュエーター系) を備えたロボットのような身体化された実体である必要がある。言い換えれば、起因の特定／因果関係の同定が可能な NM-AI は、既存の認知科学のドグマを次々に破壊しつつあるエナクティビズム (enactivism) が教えてくれているように、ソフトウェア的な仕組みだけでは実現できない (Noë (2009), Hohwy (2013), Engel 他 (2016))。

そして、身体化された NM-AI が Pearl のレベル 2 をどこまでクリアできるかを詳細に検討するために先の Minsky 流の試金石を導入し、同時に両試金石の包含関係を明らかにした。具体的には、Pearl のレベル 3 が、Minsky のレベル 4 (内省思考), 5 (自己内省思考), 6 (自意識上の情動 (感情)) の三つに分解可

能なことを示した。

その結果、NM-AI が誇る上記の魅力的な諸特性に加えて、同 AI に第三者的なメタの視点という意味での意識や行為主体性を組み込むことができれば、Minsky のレベル 4 (内省思考) までは、どうにかクリアできることをまず確認した。逆に言えば、この付加条件がなければ、Pearl のレベル 2 も覚束なくなる。なお、本論では、上記の元型的な意識の実現可能性について、主に認知発達ロボティクス分野で有名な Tani (2017) の試みを、今をとときめく Friston (2010) の Free Energy 原理との関連を指摘しながら紹介した。

Minsky によれば、レベル 4 は、予測が当たらなかつたり、障害に遭遇したり、必要な知識にアクセスできなかつたりしたとき、より広くは、変化と異常が発生したときに作動する。そのため、本論では、身体化された NM-AI がさらに Minsky のどのレベルまで到達可能であるかを確認するために、1980 年代末当時の AI 研究のトップスターでありながらその限界の前に突如 AI 研究から身を引いた Winograd 他 (1987) の“Breakdown”概念に習って変化と異常＝「脳がメモリしているアルゴリズムの先読みでは予測しえないような事柄」(松本他 (2003)) と位置づけ、NM-AI に、このような変化と異常の発見や対応がどこまで可能となるかを検討した。そして、身体化された NM-AI が Minsky のレベル 6 まで実現できれば、そのレベルに達することができることを確認した。

ただし、分析結果によれば、大脳新皮質型“インテリジェンス”の申し子である NM-AI では、Tani (2017) 流の意識やそれを遙かに超える自己意識 (Self Awareness) + 行為主体性感 (Sense of Agency) が組み込まれたとしても、Winograd 他 (1987) 流の変化と異常への対応はとてできそうもないこと、従って Minsky のレベル 6 にまでは到底到達できそうもないことを確認した。事実、Breakdown”＝“変化と異常”への対応には、その難度が増せば増すほど、社会を構成する様々な人々が時空をまたいで織りなしてきた社会ネットワーク構造の中に埋め込まれている共有知識をより深くより幅広く総動員する必要性が高まる。しかも、変化と異常への対応には新状況の識別に必須な新概念の誕生 (記号創発) も必要となる。そして、それらを可能にするのが、Minsky (2006) が教えてくれた互いの多様な意図や常識・文化等々の共有によって広範囲な協力を生み出すコミュニティ形成装置でもある情動なのである。

以上のような分析に基づき、変化と異常への対応力を発揮するための“インテリジェンス”には Minsky のレベル 6 が必須だとすると、Minsky 的な Emotion Machine が夢物語である限り、HI と BD-AI, NM-AI の三者は、本能行動・学習行動・熟慮思考・内省思考では代替関係をさらに強めていくかもしれないが、インテリジェンスが自己内省思考、その次の意識上の情動（感情）と上昇するにつれて、より顕著な補完性を強めて行くのではないだろうか、と結論づけた。

参 考 文 献

- ・ Ackerman, Evan (2019), “Intel Labs Director Talks Quantum, Probabilistic, and Neuromorphic Computing,” IEEE Spectrum, https://spectrum.ieee.org/tech-talk/computing/hardware/intel-labs-director-talks-quantum-probabilistic-and-neuromorphic-computing?utm_source=techalert&utm_campaign=techalert-03-07-19&utm_medium=email
- ・ Anderson, James A. (2017), *After Digital: Computation as Done by Brains and Machines*, Oxford University Press
- ・ 甘利俊一 (2016), 『脳・心・人工知能：数理で脳を解き明かす』, 講談社
- ・ 甘利俊一 (2008), 『神経回路網モデルと接続ニズム』, 東京大学出版会（初版は 1989）
- ・ 甘利俊一・村田昇・Klaus R. Müller (1997), 「学習の数理モデル —汎化能力と過学習—」, 37-53 頁, 外山敬介・杉江昇編『脳と計算論』に所収, 朝倉書店
- ・ Anadiotis, George (2018), “The AI chip unicorn that's about to revolutionize everything has computational Graph at its Core,” <https://www.zdnet.com/article/the-ai-chip-unicorn-that-is-about-to-revolutionize-everything-has-computational-graph-at-its-core/>
- ・ 浅川伸一 (2017), 「リカレントニューラルネットワーク」, 人工知能学会編『人工知能学大辞典』, 共立出版, 519-524 頁
- ・ 浅田稔 (2017), 「認知発達ロボティクス」, 人工知能学会編『人工知能学大辞典』, 共立出版, 1048-1053 頁
- ・ Benjamin, Ben Varkey, Peiran Gao, Emmett McQuinn, Swadesh Choudhary, Anand R. Chandrasekaran, Jean-Marie Bussat, Rodrigo Alvarez-Icaza, John V. Arthur, Paul A. Merolla, and Kwabena Boahen (2014), “Neurogrid: A Mixed-Analog-Digital Multichip System for Large-Scale Neural Simulations,” *Proceedings of the IEEE*, Vol. 102, No. 5, pp. 699-716
- ・ Bergson, Henri (1903), *An Introduction to Metaphysics*, (英訳版), G.P. Putnam’s sons.
- ・ Boden, Margaret A. (2016), *Artificial Intelligence: A Very Short Introduction*, Oxford University Press, Oxford
- ・ Brogliato, Marcelo S., Daniel M.Chada1 and Alexandre Linhares (2014), “Sparse distributed memory: understanding the speed and robustness of expert memory,” *Frontiers in Human Neuroscience*, April, Vol. 8, Article 222, pp. 1-11
- ・ Burr, Geoffrey W. (2015), “Non-Volatile Memory as a Neuromorphic Synapse: Effect of imperfections,” in ITRS ERD Presentation, Stanford University, 02/26/15
- ・ Chaisson, Eric J. (2013), “A Singular Universe of Many Singularities: Cultural Evolution in a Cosmic

- Context,” in *Singularity Hypotheses: A Scientific and Philosophical Assessment*, ed. by Eden, H. Amnon, James H Moor, Johnny H. Soraker, and Eric Steinhart, Springer Science & Business Media, Heidelberg, New York, Dordrecht, London
- ・ 中馬宏之・今井正治・黒川利明 (2017), 「人的資本 /Human Intelligence と脳模倣型人工知能 /Neuromorphic AI: インテリジェンスという視点から」, 経済産業研究所, ディスカッションペーパー, 17-P-031
 - ・ Clark, Andy (2016), *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*, Oxford University Press, Oxford, New York, Auckland, Cape Town, Dar es Salaam, Hong Kong, Karachi, Kuala Lumpur, Madrid, Melbourne
 - ・ Craig, A.D. Bud (2015), *How Do You Feel? An Interoceptive Moment with your Neurobiological Self*, Princeton University Press, Princeton and Oxford
 - ・ Creps, David (2017), 「『物質と記憶』と深層学習」(齋藤俊太訳), 平井靖史・藤田尚志・我孫子信編『ベルクソン『物質と記憶』を診断する: 時間経験の哲学・意識の科学・美学・倫理学への展開』, pp. 196-224 に所収。
 - ・ Damasio, Antonio (2018), *The Strange Order of Things: Life, Feeling, and the Making of Cultures*, Pantheon Books, New York
 - ・ Darwich, Adnan (2018), “Human-Level Intelligence or Animal-Like Abilities?” *Communications of the ACM*, vol. 61, Number 10, pp. 56-67
 - ・ Davies, Mike, Narayan Srinivasa, Tsung-Han Lin, Gautham Chinya, Yongqiang Cao, Sri Harsha Choday, Georgios Dimou, Prasad Joshi, Nabil Imam, Shweta Jain, Yuyun Liao, Chit-Kwan Lin, Andrew Lines, Ruokun Liu, Deepak Mathaikutty, Steve McCoy, Arnab Paul, Jonathan Tse, Guruguhanathan Venkataramanan, Yi-Hsin Weng, Andreas Wild, Yoonseok Yang, Hong Wang (2018), “Loihi: a Neuromorphic Manycore Processor with On-Chip Learning,” *IEEE Micro*, Vol. 38, Issue 1, pp. 82-89
 - ・ Demler, Mike (2018), “Mythic Multiplies In A Flash: Analog In-Memory Computing Eliminates DRAM Read/Write Cycles,” *The Linley Group · Microprocessor Report*, August 27, <https://www.linleygroup.com/mpr/article.php?id=12023>
 - ・ Deutch, Ariel Y. and Robert H. Roth (2014), “Pharmacology and Biochemistry of Synaptic Transmission Classical Transmitters,” in Byrne, John H., Ruth Heiderberger, and M. Neal Waxham ed. *From Molecules to Network: An Introduction to Cellular and Molecular Neuroscience*, Academic Press, Amsterdam, Boston, Heidelberg, London, New York, Paris, San Diego, San Francisco, Singapore, Sydney, Tokyo, @ pp. 207-237.
 - ・ Eliasmith, Chris (2013), *How to Build a Brain: A Neural Architecture for Biological Connection*, Oxford University Press
 - ・ Engel, K. Andreas, Karl J. Friston, and Danica Kragic (2016), “Introduction: Where’s the Action ?” in Andreas K. Engel, Karl J. Friston and Danica Kragic ed., *The Pragmatic Turn: Toward Action-Oriented View*, The MIT Press, Cambridge, Massachusetts, London, England.
 - ・ Feinberg, Todd E. and Jon M. Mallatt (2016), *The Ancient Origins of Consciousness: How the Brain Created Experience* (邦訳名『意識の進化的起源カンブリア爆発で心は生まれた』(2017)), MIT Press, Cambridge, Massachusetts, London.
 - ・ Feinberg, Todd E. and Jon M. Mallatt (2018), *Consciousness Demystified*, Cambridge, Massachusetts, London.
 - ・ Fields, Douglas R., (2009), *The Other Brain: From Dementia to Schizophrenia, How New Discoveries*

- about the Brain Are Revolutionizing Medicine and Science* (邦訳名『もう一つの脳：ニューロンを支配する陰の主役「グリア細胞」』(2018)), Simon & Schuster, New York, London, Toronto, Sydney
- ・ Fisher, Anna V. (2016), “Mechanisms of Induction Early in Development,” in Marie T. Banich and Donna Caccamise ed. *Generalization of Knowledge: Multidisciplinary Perspectives*, Routledge, pp. 89-112.
 - ・ Franklin, Stan (1995), *Artificial Minds*, MIT Press
 - ・ Freeman, Walter J. (2001), *How Brains Make Up Their Minds*, Columbia University Press, New York
 - ・ Friston, Karl J. (2010), “The free-energy principle: a unified brain theory?” *Nature Reviews Neuroscience* vol. 11, Issue 2 (February), pp. 127-38
 - ・ George, Dileep (2008), *How the Brain Might Work: A Hierarchical and Temporal Model for Learning and Recognition*, Dissertation submitted to the Department of Electrical Engineering and the Committee on Graduate Studies of Stanford University
(http://alpha.tmit.bme.hu/speech/docs/education/02_DileepThesis.pdf)
 - ・ Goodfellow, Ian, Yoshua Bengio, and Aaron Courville (2016), *Deep Learning* (邦訳名『深層学習』(2018)), The MIT Press, Cambridge, Massachusetts, London, England
 - ・ Graves, Alex, Greg Wayne and Ivo Danihelka (2014), “Neural Turing Machines,” arXiv: 1410.5401v2, pp. 1-26, <https://arxiv.org/pdf/1410.5401.pdf>
 - ・ Graves, Alex, Greg Wayne, Malcolm Reynolds, Tim Harley, Ivo Danihelka, Agnieszka Grabska-Barwińska, Sergio Gómez Colmenarejo, Edward Grefenstette, Tiago Ramalho, John Agapiou, Adrià Puigdomènech Badia, Karl Moritz Hermann, Yori Zwols, Georg Ostrovski, Adam Cain, Helen King, Christopher Summerfield, Phil Blunsom, Koray Kavukcuoglu and Demis Hassabis (2016), “Hybrid computing using a neural network with dynamic external memory,” *Nature* 538, 27 October, pp. 471-476
 - ・ Hassabis, Demis, Dharshan Kumaran, Christopher Summerfield, and Matthew Botvinick (2017), “Neuroscience-Inspired Artificial Intelligence,” *Neuron*, Volume 95, Issue 2, 19 July 2017, Pages 245-258
 - ・ Hawkins, Jeff with Sandra Blakeslee (2004), *On Intelligence: : How a New Understanding of the Brain Will Lead to the Creation of Truly Intelligent Machines* (邦訳名『考える脳 考えるコンピューター』(2005)), Times Books
 - ・ 堀尾喜彦 (2017), 「ニューラルネットワークのハードウェア化」, 人工知能学会編『人工知能学大辞典』, 共立出版, 546-551 頁
 - ・ Hennessy, John L., David A. Patterson (2017), *Computer Architecture, Sixth Edition: A Quantitative Approach*, Morgan Kaufmann
 - ・ Hohwy, Jakob (2016), “Prediction, Agency, and Body Ownership,” in Andreas K. Engel, Karl J. Friston and Danica Kragid ed., *The Pragmatic Turn: Toward Action-Oriented View*, The MIT Press, Cambridge, Massachusetts, London, England, pp. 109-120 in *Cognitive Science*
 - ・ Hohwy, Jakob (2013), *The Predictive Mind*, Oxford University Press.
 - ・ Hurwits, Marcia Kaufman and Adrian Nowles (2015), *Cognitive Computing and Big Data Analytics*, John Wiley & Sons, Inc.
 - ・ Jasanoff, Alan (2018), *The Biological Mind: How Brain, Body, and Environment Collaborate to Make Us Who We Are*, Basic Books, New York
 - ・ Jayaraman, Vivek and Gilles J. Laurent (2009), “Olfactory System: Circuit Dynamics and Neural Coding

- in the Locust,” in Larry R. Squire ed., *Encyclopedia of Neuroscience*, pp. 187-196
- ・ Kandel 他 (2013: 第 6 版) Eric Kandel, James Schwartz, Thomas Jessell, Steven Siegelbaum, A.J. Hudspeth (2012), *Principles of Neural Science* (邦訳名『カandel 神経科学』(2014)), Fifth Edition, McGraw-Hill Education/Medical
 - ・ Kanerva, Pentti (1988), *Sparse Distributed Memory*, MIT Press
 - ・ Kanerva, Pentti (2014), “Computing with 10,000-Bit Words,” in Proceedings of 52nd Annual Allerton Conference on Communication, Control, and Computing
 - ・ Kohara, Keigo, Michele Pignatelli, Alexander J Rivest, Hae-Yoon Jung, Takashi Kitamura, Junghyup Suh, Dominic Frank, Koichiro Kajikawa, Nathan Mise, Yuichi Obata, Ian R Wickersham and Susumu Tonegawa (2014), “Cell type-specific genetic and ontogenetic tools reveal hippocampal CA2 circuits,” *Nature Neuroscience*, Feb; 17(2), 269-79. http://www.riken.jp/pr/press/2013/20131219_2/
 - ・ 小池和男・中馬宏之・太田聡一 (2001), 『もの造りの技能—自動車産業の職場で』, 東洋経済新報社
 - ・ 工藤巖 (2011), 『脳とグリア細胞：見えてきた脳機能の鍵を握る細胞たち』, 技術評論社
 - ・ Kurzweil, Ray (2012), *How to Create a Mind: The Secret of Human Thought Revealed*, Viking (Penguin Group)
 - ・ LeDoux, Joseph (2015), *Anxious: Using the Brain to Understand and Treat Fear and Anxiety*, Viking
 - ・ Llinas, Rodolfo R. (2001), *I of the Vortex: From Neurons to Self*, A Bradford Book, the MIT Press, Cambridge, Massachusetts, London, England
 - ・ Lotto, Baeu R.(2009), “Statistical Analysis of Visual Perception,” in Larry R. Squire ed., *Encyclopedia of Neuroscience*, pp. 381-386
 - ・ Luo, Liquin (2016), *Principles of Neurobiology*, Garland Science, Taylor & Francis Group, New York, London.
 - ・ Marblestone, Adam H., Greg Wayne, and Konrad P. Kording (2016), “Towards an integration of deep learning and neuroscience,” *Frontiers in Computational Neuroscience*, 14 September, pp. 1-64 (<http://www.biorxiv.org/content/early/2016/06/13/058545.full.pdf>)
 - ・ Maren, Alianna J. (1990), “Laterally-Connected Autoassociative Network,” Alianna J. Maren, Craig T. Harston and Robert M. Pap ed., *Handbook of Neural Computing Applications*, Academic Press, San Diego, New York, Boston, London, Sydney, Tokyo, Toronto, pp. 125-140
 - ・ 松本元 (1995), 『愛は脳を活性化する』, 岩波科学ライブラリー, 岩波書店
 - ・ 松本元・辻野広司 (2003), 「情と意が脳を創る」, 松本元・小野武年編『情と意の脳科学—人とは何か—』, 培風館
 - ・ McClulloch, Warren S. and Walter H. Pitts (1943), “A Logical Calculus of the Ideas Immanent in Nervous Activities,” *Bulletin of Mathematical Biophysics*, vol. 5, pp. 115-133
 - ・ McFadden, Johnjoie and Jim Al-Khalili (2014), *Life on the Edge: The Coming of Age of Quantum Biology*, Crown Publishers/Penguin.
 - ・ Merolla, Paul A., John V. Arthur, Rodrigo Alvarez-Icaza, Andrew S. Cassidy, un Sawada, Filipp Akopyan, Bryan L. Jackson, Nabil Imam, Chen Guo, Yutaka Nakamura, Bernard Brezzo, Ivan Vo, Steven K. Esser, Rathinakumar Appuswamy, Brian Taba, Arnon Amir, Myron D. Flickner, William P. Risk, Rajit Manohar, and Dharmendra S. Modha (2014), “A million spiking-neuron integrated circuit with a scalable communication network and interface,” *Science*, vol. 345, Issue 6197, 8 August, pp. 668-673.

- ・ Minsky, Marvin (1985), *Society of Mind*, Simon & Schuster, Inc.
- ・ Minsky, Marvin (2006), *The Emotion Machine: Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind*, Simon & Schuster, New York, London, Toronto, Sydney
- ・ Morgan, Timothy Prickett (2018), “Intel’s Exascale Dataflow Engine Drops X86 and von Neumann,” <https://www.nextplatform.com/2018/08/30/intels-exascale-dataflow-engine-drops-x86-and-von-neumann/>
- ・ Natarajan, Sanjay (2018), “Unsticking Moore’s Law,” <https://semiengineering.com/unsticking-moores-law/>
- ・ Noë, Alva (2009), *Out of Our Heads: Why You Are Not Your Brain, and Other Lessons from the Biology of Consciousness*, Hill and Wang, A DIVISION OF Farrar, Status and Giroux, New York
- ・ Numenta (2017), *Biological and Machine Intelligence (BAMI): A living book that documents Hierarchical Temporal Memory (HTM)*, <http://numenta.com/assets/pdf/biological-and-machine-intelligence/BAMI-Complete.pdf>
- ・ 太田宏之 (2017), 「空間的神経表象から時間的圧縮過程へ」, 平井靖史・藤田尚志・我孫子信編『ペルクソン『物質と記憶』を診断する：時間経験の哲学・意識の科学・美学・倫理学への展開』, pp. 226-248
- ・ Olhausen, Bruno A. and Christopher J. Rozell (2017), “Sparse codes from memristor grids: The adjustable resistive state of memristors makes it possible to implement sparse coding algorithms naturally and efficiently,” *Nature Nanotechnology*, Advance Online Publication, www.nature.com/naturenanotechnology
- ・ Ornstein, Robert Evan (1986), *Nultimind: A New Way of Looking at Human Behavior*, Houghton Mifflin
- ・ Pearl, Judea and Dana Mackenzie (2018), *The Book of Why: The New Science of Cause and Effect*, Allen Lane, New York
- ・ Plate, Tony A. (2003), *Holographic Reduced Representation: Distributed Representation for Cognitive Structures*, CSI Publications
- ・ Richard, Blake A. and Timothy P. Lillicrap (2018), “Dendritic solutions to the credit assignment problem,” *Current Opinion in Neurobiology*, vol. 54, pp. 28-36
- ・ Rinkus, Gerard, Greg Leshner, Jasmin Leveille, and Oliver Layton (2016), *Sparse Distributed Representation & Hierarchy: Keys To Scalable Machine Intelligence*, AFRL (Air Force Research Laboratory) -RY-WP-TR-2016-0030, United States Air Force
- ・ Rolls, Edmund T. (2016), *Cerebral Cortex: Principles of Operation*, Oxford University Press, Oxford, New York
- ・ Rolls, Edmund T. (2018), *The Brain, Emotion, and Depression*, Oxford University Press, Oxford, New York
- ・ Roy, KAUSHIK (2015), “Neuromorphic Computing With Cellular Array Of Magneto-Metallic Neurons & Synapses,” in ITRS ERD Presentation, Stanford University, http://cspin.umn.edu/events/dac_presentations/KaushikRoy.pdf
- ・ Rusch, Miles and Jan M. Rabaey (2018), “The Design of an Analog Associative Memory Circuit for Applications in High-Dimensional Computing,” Technical Report No. UCB/EECS-2018-72, <http://www2.eecs.berkeley.edu/Pubs/TechRpts/2018/EECS-2018-72.html>
- ・ Sahu, Satyajit, Subrata Ghosh, Kazuto Hirata, Daisuke Fujita, Anirban Bandyopadhyaya (2013), “Multi-level memory-switching properties of a single brain microtubule,” *Applied Physics Letters*, vol. 102, issue 12, pp. 102-106

- ・ Schuman, Catherine D., Thomas E. Potok, Robert M. Patton, Douglas Birdwell, Mark E. Dean, Garrett S. Rose, and James S. Plank (2017), “A Survey of Neuromorphic Computing and Neural Networks in Hardware,” arXiv: 1705.06963v1, pp. 1-88
- ・ Severa, William M., Ryan Dellana, Craig M. Vineyard, and James B. Aimone (2018), “Whetstone: An accessible, platform-independent method for training spiking deep neural networks for neuromorphic processors,” <https://www.sysml.cc/doc/71.pdf>
- ・ Sharma, Ashok K. (2003), *Advanced Semiconductor Memories: Architectures, Designs, and Applications*, IEEE Press, Wiley-Interscience
- ・ Sherman, S. Murry and Ray W. Gullery (2013), *Functional Connections of Cortical Area: A New View from the Thalamus*, MIT Press
- ・ Silver, David, Aja Huang, Chris J. Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel and Demis Hassabis (2016), “Mastering the game of Go with deep neural networks and tree search,” *Nature*, Vol. 529, 28 January, pp. 484-489
- ・ Sloman, Aaron (1994), “Semantics in an Intelligent Control System,” *Philosophical Transactions: Physical Sciences and Engineering*, Vol. 349, No. 1689, *Artificial Intelligence and the Mind: New Breakthroughs or Dead-Ends?* (Oct. 15, 1994), pp. 43-58
- ・ Srinivasa, Narayan (2015), “Neuromorphic Electronics for Energy Efficient Processing,” ITRS ERD Presentation, Stanford University, 02/26/15
- ・ Stokes, Jon (2006), *Inside the Machine: An Illustrated Introduction to Microprocessors and Computer Architecture*, No Starch Press
- ・ Streat Lennard G., Dhireesha Kudithipudiy, and Kevin Gomez (2016), “Non-volatile Hierarchical Temporal Memory: Hardware for Spatial Pooling,” arXiv: 1611.02792 v1, pp. 1-11 (<https://arxiv.org/pdf/1611.02792.pdf>)
- ・ Tani, Jun (2017), *Exploring Robotic Minds Actions, Symbols, and Consciousness as Self-Organizing Dynamic Phenomena*, Oxford University Press
- ・ 利根川進グループ (2017), 「記憶を思い出すための神経回路を発見 — 海馬の二つの局所回路が記憶の書き込みと想起を分担している —」 (http://www.riken.jp/pr/press/2017/20170818_2/)
- ・ Trappenberg, Thomas P. (2010), *Fundamentals of Computational Neuroscience*, 2nd, Oxford University Press
- ・ Tavanaei Amirhossein, Masoud Ghodratiy, Saeed Reza Kheradpishehz, Timoth Masquelierx and Anthony Maida (2019), “Deep Learning in Spiking Neural Networks,” *Neural Networks*, Volume 111, pp. 47-63, <https://arxiv.org/abs/1804.08150>
- ・ Van Der Made, Peter AJ and Anil Shamrao Mankar (2017), “Neural Processor Based Accelerator System and Method,” United States Patent Application Publication, US 2017/0024644 A1
- ・ Weizenbaum, Joseph with Gunna Wendt (2015), *Islands In Tth Cyberstream Seeking Havens Of Reason In A Programmed Society*, Litwin Books
- ・ Winograd, Terry and Fernando Flores (1987), *Understanding Computers and Cognition: A New Foundation for Design*, Addison-Wesley Publishing Company
- ・ Wu1, Yujie, Lei Deng, Guoqi Li, Jun Zhu and Luping Shi (2018), “Spatio-Temporal Backpropagation for Training High-Performance Spiking Neural Networks,” *Front. Neurosci.*, <https://www.frontiersin.org/>

中馬宏之：ヒューマン・インテリジェンスとビッグデータ型及び脳神経模倣型 AI：インテリジェンスとは何かを考える

articles/10.3389/fnins.2018.00331/full

・ Zarkadakis, George (2015), *In Our Own Image: Savior or Destroyer? The History and Future of Artificial Intelligence*, Pegasus Book, New York, London

(ちゅうま・ひろゆき 成城大学社会イノベーション学部教授)