



RIETI Policy Discussion Paper Series 17-P-031

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

中馬 宏之
経済産業研究所

今井 正治
エイシップ・ソリューションズ株式会社

黒川 利明



Research Institute of Economy, Trade & Industry, IAA

独立行政法人経済産業研究所
<http://www.rieti.go.jp/jp/>

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*

中馬 宏之 (成城大学/経済産業研究所)

今井 正治 (エイシップ・ソリューションズ株式会社)

黒川 利明

要 旨

現行の AI 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模倣型人工知能 (Neuromorphic AI : NM-AI) である。本論の目的は、現時点で実用化に大きな課題を抱えているものの中長期的には AI の本丸として登場すると期待されている NM-AI に関して、まず、なぜいま NM-AI に注目する必要があるのかについて、経済学のみならず計算機科学・半導体集積回路や Neuroscience の視点から検討することである。そして、それらの検討結果に基づき、そもそも NM-AI や BD-AI のインテリジェンス特性とは何かを明らかにしつつ、BD-AI の先にある NM-AI のインテリジェンスと人的資本、したがって Human Intelligence (HI) との補完性・代替性について論じることである。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という言葉为非行動主義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”やその自己変化能に言及しながら、NM-AI 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、人的資本/HI の社会ネットワーク特性 (Human Capital as social network)、そして、“変化と異常への対応”に必須な柔軟で高速の視点切り替え装置としての情動、互いの多様な意図や常識・文化等々の共有による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、という 3つの視点である

キーワード：ビッグデータ型 AI、脳模倣型 AI、ヒューマンインテリジェンス、人的資本、補完・代替性、社会ネットワーク特性、自己変化能

JEL classification : J01、J24、L63

RIETI ポリシー・ディスカッション・ペーパーは、RIETI の研究に関連して作成され、政策をめぐる議論にタイムリーに貢献することを目的としています。論文に述べられている見解は執筆者個人の責任で発表するものであり、所属する組織及び（独）経済産業研究所としての見解を示すものではありません。

* 本稿は、独立行政法人経済産業研究所におけるプロジェクト「人工知能が社会に与えるインパクトの考察：文理連繋の視点から」の成果の一部である。

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

中馬宏之¹、今井正治、黒川利明

1. はじめに

本プロジェクトでは、21 世紀に入って加速的な進化を遂げはじめた人工知能 (AI) が中長期的に社会に与えるインパクトを、経済学的な視座を中心に据えつつも、計算機科学・半導体集積回路の専門家と文理連携して検討してきた。その際、特に焦点を当ててきたのは、中・長期的スパンで大きな変革を社会にもたらすと期待されている AI の本丸、脳 (神経) 模倣型人工知能 (Neuromorphic AI: 以後 NM-AI と呼ぶ) である。²

現行の AI 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 NM-AI である。もちろん、すべての AI 研究が共に McCulloch-Pitts 型 (脳) 神経細胞 (ニューロン) で名高い神経生理学者 Warren Sturgis McCulloch (1889-1969) や Computational Neuroscience の先駆者 Walter Pitts³ (1923-1969)、神経回路の学習則 Hebbian Rule (ヘブ) 則) で名高い神経心理学者 Donald Olding Hebb (1904-1985) などをも嚆矢としている。したがって、程度の差はあれ、両者の志すところは、脳機能の本質把握に基づく AI の実用化であることは間違いない。

実際、Hassabis 他 (2017) や Marblestone 他 (2016)、Schuman 他 (2017) の BD-AI や NM-AI に関する包括的な Survey に示されているように、時代を席卷している BD-AI も大脳新皮質・海馬・視床などの基本動作原理に習うこと無しに発展することが極めて難しくなっている。ただし、これらの著者たちが認めているように、最近になればなるほど AI 研究と

¹ 本 DP 本論部分の内容に誤りがあるとしたら、すべての責任は筆頭著者の中馬に帰する。

² 本プロジェクトでは、数はまだ限られているが、産総研の AI 研究センターの辻井潤一センター長を中核とした専門家、IBM ワトソンの国内開発部隊、EU・Human Brain Project 内の Neuromorphic Computing プロジェクト (ドイツ・ハイデルベルグ大学・Karl Heinz Meier 教授が主査) の BrainScaleS や英国・マンチェスター大学の Steve Furber 教授を中核とする SpiNNaker、米国国防省 (DARPA) Synapse プロジェクトの目玉である IBM 発の TrueNorth (IBM Research of Dharmendra S. Mocha 氏が主査) で中核的な役割を果たしてきた日本 IBM 部隊、認知発達ロボティクスや脳神経外科の世界的研究者 (いずれも大阪大学工学部と医学部)、豊橋技術科学大学視覚認知情報学研究室、北海道大学アドバンス LSI 工学研究グループ、北海道大学情報科学研究科の国内外で著名な研究者、理化学研究所脳科学総合研究センターを代表する著名な脳神経科学者等々への聞き取り調査を実施してきた。ご協力いただいた皆様に、この場をお借りして心からお礼を申し上げたい。

³ Pitts は、青少年時代にストリートギャングに追い回されたり、高校をドロップアウトしたり、シカゴ大学をはじめとしたいくつかの大学の著名な学者の講義にモグリで出席したりと数奇な人生を辿って大成した数学・論理学分野の大天才でもある。例えば、15 歳の時にシカゴ大学での数学者・論理学者・哲学者 Bertrand Russell の講義を聴講して Russell を驚愕させたり、論理実証主義の大御所論理学者 Rudolf Carnap と共同研究したり、サイバネティクス提唱者のあの Norbert Wiener の助手を務めたりと、数々の逸話を残している。詳しくは、Anderson (2017) や英文 WIKI を参照されたい。特に、当初の “The Brain Works by Logic” (Anderson (2017) 7 章のタイトル) が常識だった時代から “The Brain Doesn’t Work by Logic” (同 8 章のタイトル) が常識となっていた時代の移り変わりや Pitts 特有のパーソナリティにも関係していたとされる彼の晩年の状況にもたらしたであろう寂しさには心が痛む。

Neuroscience 研究の専門分化が著しいようである。アイロニカルな現状は何を意味しているのか？ 第三者的には、そのような現象自体も興味深い。

このように一見似通った両者であるが、BD-AI と NM-AI とは、実用化を目指すタイムスパンのみならず脳（神経）模倣の目的やスタイルが根本的に異なっている。例えば、前者は分析的に扱いやすい極めて単純化された人工ニューロンを基本としているが、後者は実ニューロンに忠実に習った様々に複雑な人工ニューロンを用いるという特徴がある（Schuman 他（2017））。後者が Neuromorphic と呼ばれる根本理由でもある。

そして、本論がインテリジェンスという視点から AI の本丸である NM-AI に注目する理由も、想定されているニューロンやその背景となる設計思想の BD-AI との違いが、次節以降、特に第 3 節で示されるように、経済概念である人的資本、したがって Human Intelligence（HI）との補完・代替関係を規定する NM-AI のインテリジェンス特性が BD-AI と比較して大きな違いを中長期的にもたらすと考えているためである。ちなみに、本論では、NM-AI のインテリジェンス特性＝「過去の記憶と現況に基づきながら環境適応していくための世界/社会/環境モデルに基づいた予測/学習能力ならびにその予測/学習能力に基づく自己変可能」という風に見なす（第 3 節で詳述）。

NM-AI の研究開発者達が実ニューロンのより忠実な模倣に拘るのには、十分な根拠がある。その一つは、現行のノイマン/チューリング型（von Neumann/Turing-type）と呼ばれるコンピュータの本質である順次処理（sequential/serial processing）に拘っているのは、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（第 2 節で詳述）。さらに、NM-AI に注目が集まっている背景には、半世紀以上にわたって半導体微細化を牽引してきた Moore の法則が、原子数十個レベルまで微細化が進展してきたことにより、トンネル（量子）効果⁴などの様々な物理限界に直面してきているという厳然とした事実がある。つまり、中期的な現実課題としても、微細化一辺倒（More Moore）ではなく、並列分散処理（parallel distributed processing：PDP）などの設計上の工夫（More Than Moore）の重要性が急速に高まってきている。⁵

以上の様子は、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の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いる 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

⁴ 通常の半導体の世界では、電子はトランジスタがオフ状態では絶縁膜を通り抜けできない。しかし、微細化が進み電子の量子的な振る舞いが支配的になると、スイッチがオフ状態でも絶縁膜を通り抜ける（したがって発熱する）ようになる。このような本来、電子が通り抜けられないはずの絶縁膜を、ある確率で通り抜けてしまうことを電子のトンネル効果と呼ぶ。ちなみに、絶縁膜の通りにくさは High-k 値で示されるが、この High-k 値の高い材料を発見できれば、トンネル効果を遮断ないし軽減できる。High-k/Metal Gate と呼ばれるトランジスタに関わる半世紀ぶりのイノベーションをもたらした世界の R&D システムに関しては、中馬（2011）に詳しい。

⁵ Schuman 他（2017）参照。

なされている。

“The final version of AlphaGo used 40 search threads, 48 CPUs, and 8 GPUs. We also implemented a distributed version of AlphaGo that exploited multiple machines, 40 search threads, 1,202 CPUs and 176 GPUs.” (「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。」 (中馬訳))

そして、おそらくこの多数のコンピュータ・サーバーを束ねた “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、1 CPU 当たりの最大消費電力を 145 ワット⁶、1 GPU 当たりの最大消費電力を 300 ワット (/毎時)⁷とすると、これだけで 23 万ワットなる。したがって、システムメモリや 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。⁸

また、全米での人気クイズ番組 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る 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”⁹ (超並列仮説形成・評価作業) ソフトウェアである。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 なもののだとして下記のような興味深いコメントをしている。¹⁰

“The Watson software uses many kinds of software “experts,” huge initial datasets, massive parallel computation, task-dependent

⁶ 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。

⁷ Nvidia Tesla P100 の値。 <http://images.nvidia.com/content/tesla/pdf/nvidia-tesla-p100-datasheet.pdf> 参照

⁸ ちなみに、富士通によれば、2011 年 11 月 TOP500 リストのベンチマーク計測時のスーパーコンピュータ「京」の消費電力は、12.65989 メガ (百万) ワットである。したがって、人間の脳の消費電力を 20 ワットとすると、人間約 60 万人分にもなる。途方もない消費電力量である。
<http://www.fujitsu.com/jp/about/businesspolicy/tech/k/qa/k04.html>

⁹ http://researcher.watson.ibm.com/researcher/view_group_subpage.php?id=2159。ちなみに、サーバー単体では質問に答えるのに少なくとも 2 時間半かかるが、サーバー群を束ねる DeepQA を利用すると 3 秒以内で済むという (Van Der Made (2013)、Hurwitz 他 (2015))。

¹⁰ Jeopardy 用 Watson のさらに詳しい説明は、Hurwitz 他 (2015) を参照してほしい。

probabilistic inference, associative learning, confidence and correlation estimation, and what the developers call the “integration of shallow and deep knowledge.” Answers are arrived at by estimate and consensus from multiple supporting sources. The rational analysis characteristic of early AI plays only a small role. … Watson is still a long way from a brain-like computer, but it is heading in that direction.”（「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。」（中馬訳、（）内は同名追加）

ただし、消費電力量という視点から見ると、上記の Anderson の最後のコメントの意味が少しだけ分かってくる。というのは、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。¹¹

ちなみに、第 2 節で紹介するが、現在の代表的な NM-AI の電力消費量は、実行可能な演算の得手不得手には大きな違いがあるが、ノイマン/チューリング型の 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)。そして、上記の AlphaGo や Watson の事例からも明らかなように、本来、メモリ・論理並存の脳型処理を敢えてメモリ・論理分離のノイマン/チューリング型コンピュータで本格的に実現しようとする、とてつもない消費電力量が必要となる (Van Der Made (2013))。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 NM-AI 研究グループ (Benjamin 他 (2014)) の下記の事実がとても興味深い。

“A personal computer simulates a mouse-scale cortex model (2.5×10^6 neurons) 9000 times slower than a real mouse brain operates, while using 40 000 times

¹¹ <https://www.forbes.com/sites/markpmills/2011/02/21/ibms-watson-jeopardy-stunt-unleashes-a-third-great-cycle-in-computing/#5b25222a5b43>. ただし、8 万ワット説や 75 万ワット説もある。

more power (400W versus 10 mW). Simulating a human-scale cortex model (2x10¹⁰ neurons), the Human Brain Project's goal, is projected to require an exascale supercomputer (10¹⁸ flops) and as much power as a quarter-million households (0.5 GW). Hence, large-scale neural modeling's potential has hardly been tapped.” (「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。」 (中馬訳))

これらの事例からも、脳型の非ノイマン/チューリング型コンピュータが囑望される理由の一端を垣間見ることができる。

BD-AI と NM-AI の消費電力に関する大きな違いは、前者が現行の順次処理 (sequential/serial processing) を特徴とするノイマン/チューリング型コンピュータと共に急速な発展を遂げてきたのに対して、後者がなかなか実用化への道が開かない並列分散処理 (parallel distributed processing : PDP) を特徴とするハードウェア¹²との親和性がとても高いことにある (Schuman 他 (2017))。もちろん、前者の場合でも、画像や音声データなどの並列処理を得意とする NVidia/Intel/AMD などの GPGPU (General-purpose computing on graphics processing units) や Xilinx/Altera (現 Intel) などの優れた事前・事後のハードウェア再構成性を誇る FPGA (field-programmable gate array) が多用されるので、順次処理だけに依存しているわけではない。ただし、現時点での GPU や FPGA の消費電力は通常の CPU と五十歩百歩なので、20 ワット前後で動作する人間の脳のような NM-AI をとても実現できそうもない。¹³

事実、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”が大きく立ちはだかっている。その性能上のボトルネックは、メモリと PE

¹² 小柳他 (2012)によれば、1970 年代から 80 年代にかけて米国ベンチャ企業による PDP 開発・実用化の大きな流れが起こったという。そして、その流れは Intel X86 をはじめとする安価な汎用 CPU 時代の到来と共に表舞台から消えていった。例えば、NASA に納入された後述する Content-Addressable Memory (CAM) を駆使した Goodyear Aerospace Corporation の STARAN@1972 やその後継の MasPar Computer Corporation の MasPar MP-1@1990 や MP-2@1992 などがあった。この部分は、英文 wiki 参照。

¹³ 最先端の FPGA の一個当たりの平均及び最大消費電力は、例えば、68 億個という驚異的なトランジスタ搭載数を誇る Xilinx UltraScale+の場合、各々 5 ワット強及び 35 ワットである。したがって、GPU や CPU に比べて少なくとも消費電力が一桁低い。(https://www.xilinx.com/support/documentation/sw_manuals/xilinx2014_1/ug440-xilinx-power-estimator.pdf)。

(Processing Elements) 又は Processors/Cores の間のコミュニケーション (経路) 構造に規定される転送・応答速度の隘路に起因する。そのため、Intel Processor に代表されるような Many Cores 型¹⁴になればなるほど、上記ボトルネックが支配的となり、消費電力のより多く (最先端プロセッサでは 90%前後) がメモリへのアクセスで消費されてしまう。

以上のようなノイマン/チューリング型コンピュータの本質的な限界を突破できる可能性が高い PE・メモリ融合型 (メモリ・論理並存/同居型) アーキテクチャを持つ Neuromorphic Chip/Device や NM-AI への期待が高まってきている大きな理由である。¹⁵

本論の目的は、現時点で実用化に大きな課題を抱えているものの中長期的には AI の本丸として登場すると期待されている NM-AI に関して、まず、なぜいま NM-AI に注目する必要があるのかについて社会科学のみならず計算機科学・半導体集積回路や (門外漢ではあるが) Neuroscience の視点から検討することである (第 2 節)。そして、それらの検討結果に基づき、そもそも NM-AI や BD-AI のインテリジェンス特性とは何かを明らかにしつつ、BD-AI の先にある NM-AI のインテリジェンスと人的資本、したがって Human Intelligence (HI) との補完性・代替性について論じることである (第 3 節)。

より具体的には、Numenta (2017)、George (2008)、Minsky (2006)、Hawkins (2004)、松本他 (2003)、Ornstein (1986)、Sloman 他 (1994) 等に習いつつ、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という言葉を用いた視点から再整理・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、NM-AI 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、人的資本/HI の社会ネットワーク特性 (Human Capital as social network)、そして、“**変化と異常への対応**” に必須な柔軟で高速の視点切り替え装置としての情動、互いの多様な意図や常識・文化等々の共有による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、という三つの視点である。¹⁶

人間の脳のようなエネルギー効率で作動する NM-AI の時代が訪れると、次節で詳述するように、その優れた仮説発見・創造の補助能力が人間との協業形態や社会そのものの有り

¹⁴ 例えば、最大の CPU コアを持つ Intel Xeon Processor E5-4669 v4 (インテル推奨価格 7007 ドル/個) の場合、22 個を誇る。<http://ark.intel.com/products/series/91287/Intel-Xeon-Processor-E5-v4-Family> 参照。

¹⁵ IBM 発の最新の NM-AI である TrueNorth (後述) の消費電力レベルを 1 とすると、インテル CPU では、その約 1000 倍になる (http://news.mynavi.jp/articles/2016/12/12/sc16_ibm/、Carothers 他 (2015))。もちろん、平均的な人間の脳が 20 ワット程度で動いている状況と比較すると、既存の脳型 AI に関しても、さらにその 3 桁以上のエネルギー効率向上が望まれている。

¹⁶ 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。ただし、“The question is not whether intelligent machines can have any emotions, but whether machines can be intelligent without any emotions.” (「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なり得るかである。」 (中馬訳)) と述べている。そして、ほぼ 10 年後の Minsky (2006) になると、“Machines cannot be intelligent without any emotions” (この文章は Minsky (2008) には無い) という確信に変わっていて大変に興味深い。

様を一変させる可能性が高い。NM-AI の帰納力が格段に上昇すると、人間の創造力を特徴付けるパーズ (Peirce) 流アブダクション (abduction: 仮説発見) にも貢献できる可能性が出てくるからである。しかも、人的資本/HI との親和性が高いそのような NM-AI の能力がある閾値を一旦超えてくると、人的資本/HI と NM-AI との相乗効果により、NM-AI のレベルが加速度的に向上していく可能性がある。そうなる、人的資本/HI が持つ重要な特徴である“変化と異常への対応力”ならびにそのような対応力自体の変化能力である自己変化能への NM-AI のインパクトを考察する重要性がさらに大きく増大してくる。変化と異常への対応力には、小サンプルでも威力を発揮する帰納力やアブダクション能力が深く関係するからである。

上記の点に関して、松本 (1995) によれば、「生きている脳を解明する (AI) 研究には、脳が獲得したアルゴリズムの解明を目指す研究 (Computational Neuroscience) と、そのアルゴリズム獲得の戦略の解明をめざす研究がある」という。そして、前者は、その応用展開の際に「脳の獲得したアルゴリズムを現在の (ノイマン/チューリング型) コンピュータのプログラム開発に用いる。」(松本 (1995)、() 内は中馬追加)。まさに、いまはときめく AlphaGo や Watson に代表される BD-AI を特徴付けるアプローチと言える。

他方、後者、したがって NM-AI が目指す脳研究は、「“脳がアルゴリズムを自動獲得するためのアルゴリズム” (戦略) を解明し、この戦略を具現化する素子とアーキテクチャ (構築) を見出しシステム化して、このシステムがはたして情報処理のアルゴリズムを自動獲得するかどうかを検証すること」(松本 (1995)) だということ。つまり、脳が獲得したアルゴリズムそのものを 20 ワット前後の消費電力で自動的に産み出すことのできるメタアルゴリズムの解明・実現ということになる。そして、本論の取り組みも、そのようなメタアルゴリズムの実現が夢物語でなくなりつつある現状に深く根ざしている。

なお、メタアルゴリズムとは、特定問題解決のための特定計算手順としてのアルゴリズムではなく、そのようなアルゴリズムをも自ら生成するアルゴリズムである。¹⁷ したがって、完成形としての NM-AI は、得手不得手はあるものの、それ自身が“変化と異常への対応力”ならびにそのような対応力自体の変化能力である自己変化能を持つようになる。

以上のような問題意識に基づいた NM-AI 関連の社会科学的な研究は、現段階では、近未来を扱うという特異性やそのために必要な学際的なハードルの高さも影響し、内外での本格的な研究はこの調査の範囲内では存在しなかった。

2. なぜ NM-AI に着目するのか？

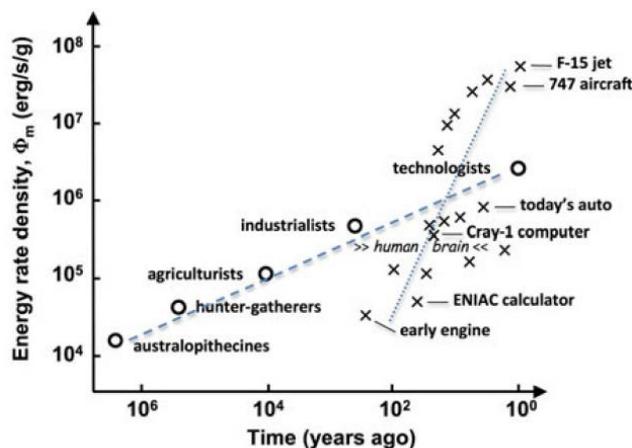
¹⁷ Bostrom (2016) は、そこで提示されている SF 的な技術の早期実現を想定した近未来の構図には様々な異論があり得るが、メタアルゴリズムに関して次のように述べている。“At its later stages, however, a seed AI should be able to understand its own workings sufficiently to engineer new algorithms and computational structures to bootstrap its cognitive performance.” (「“Seed AI<自らのアーキテクチャをも改善できるようになった最初の NM-AI>は、その最終段階になると、自らの振る舞いを十分に理解することができ、その認知能力を自力で改善するための新しいアルゴリズムと計算構造を作り上げるだろう」(中馬訳)

2. 1 エネルギー効率曲線の長期動向と NM-AI 登場の社会的意義

前述のように、現在、20 ワット前後で動作する人間の脳のような NM-AI が囑望されている。そして、そのためには、現行のノイマン/チューリング型コンピュータよりも6桁以上の消費電力の削減が必須である。ただし、そのような格段のエネルギー効率性の向上は、想像以上に難度が高そうである。

A. Chaisson (2013)が明白にした Technological Singularity

この点は、Chaisson (2013)が計算している宇宙規模での諸活動の 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消費量)を示している下図に如実に示されている。この図(縦軸が Energy rate density)に示されるように、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。そして、NM-AIの実現には、このような傾向線からの人類史上2回目の逸脱が要請されているようである。なお、Chaissonは、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。なお、Kurzweilの弱点(早急な単線思考)に関する分かりやすい指摘が Zarkadakis (2015)、科学(含む Neuroscience)・技術史のみならず壮大な地球史・宇宙史的な視点からの示唆に富む警鐘が Anderson (2017)に述べられている。

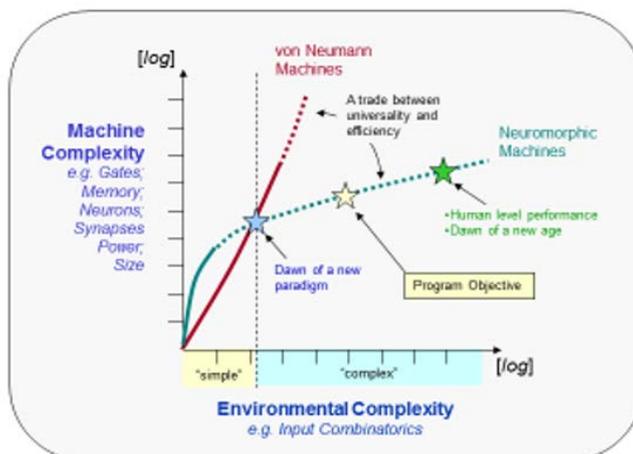


Chaisson (2013)より抜粋

B. 半導体国際ロードマップ委員会(2015)での指摘

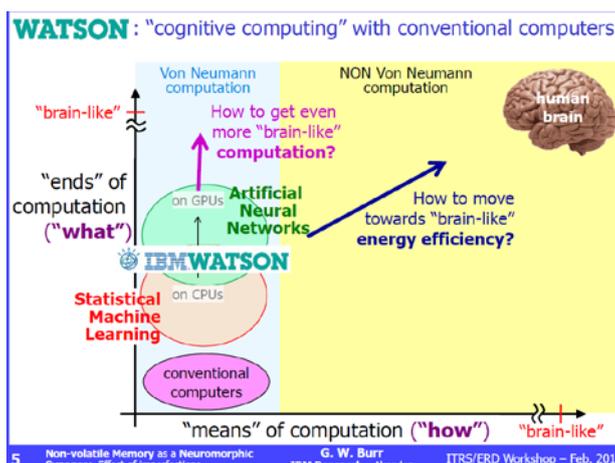
事実、このような現実を前にして、計算機科学の専門家達の多くも、DARPA・SyNAPSEプロジェクトの下図に示されているように、現行のノイマン・チューリング型コンピュータでは、そのエネルギー(消費電力)効率の低さから、近未来社会の莫大なコンピューティング

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実現できそうもないと考えられている。¹⁸



(米国 DARPA・SyNAPSE プロジェクト報告書から)

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)の下記の図が印象的である。この図の縦軸の“what”に相当する脳模倣型コンピュータが望まれている様子、横軸には“how”にあたる脳模倣型コンピュータを実現する方法の難度の高さが示されている。そして、脳模倣型コンピュータが、縦軸・横軸の二重波線で表現されているようにノイマン/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、しかも、脳機能が十分に解明されていないことにより実現手段が未だ見えていないこと等が示されている。



Burr(2015)@ITRS2015年2月のプレゼン資料より

¹⁸ このような意味でも、一世を風靡している Kurzweil (2004)のシンギュラリティ(singularity)仮説実現のハードルは、想像以上に高いと判断される。

C. NM-AI ハードウェア実践例での消費電力効率性

なかなか実用化に至らないNM-AIであるが、では、その消費電力効率性は、試作段階ではどの程度発揮されているのだろうか？ここでは、この点を代表的な NM-AI ハードウェア実践例で見てみよう（詳細は本論の附論を参照）。なお、BD-AI とは異なった設計思想に基づくノイマン/チューリング型コンピュータ上でのNM-AI ソフトウェア実現の興味深い試みについては、どのようなインテリジェンスを目指しているかに直接関わってくるので、主に第3節で触れたい。

C.1 DARPA: IBM-TrueNorth

現在、NM-AI ハードウェアとしては、IBM の TrueNorth (アメリカの DARPA (国防高等研究計画局)・支援プロジェクト) が有名である。このシステムでは、40 億個の (デジタル) ニューロンと 1 兆個の (デジタル) シナプスを 4096 個の TrueNorth (デジタル) チップで動作させる際の電力消費量の目標値が 4kw 未満に設定されている。¹⁹ 4kw は、人間の脳(20 ワット)/全身体(100 ワット)と比べると 200 人/40 人分に相当する。

ちなみに、前述の AlphaGo の駆動には約 1 万 2 千人分の電力消費量が必要であったから、実行できる事柄の種類が異なるため単純比較は参考程度であるが、両者で電力効率が 3 桁異なることになる。²⁰ また、2014 時点での TrueNorth は、54 億個のトランジスタで百万個のニューロンと 2 億 5600 万個のシナプスを実現しており、実ニューロンと同じ (realtime) 速さで動作させる場合、70mw しか消費電力を使用していない。これらの事実から、前節で紹介した Kwabena Boahen グループによる PC 上でのマウス脳の模倣例と比較すると、TrueNorth の消費電力効率の格段の高さをうかがい知ることができる。

ヒトの脳内には 1000 億個以上のニューロン (うち 800 億個は小脳関連) が存在するというのであるから、TrueNorth の目標ニューロン数がもう一桁上がるとヒトの数値に近づく。ただし、そのシナプス数は、ヒトの 150 兆個前後に比べると圧倒的に少ない。この辺りは、TrueNorth の構造特性に由来している。具体的には、TrueNorth の最小単位である Neurosynaptic Core と呼ばれる部分には、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 個付随しているので、Core 全体では 256x256 個 (6 万 5536 個) のシナプスが利用可能となっている。

ところが、1 ニューロンあたり 256 個のシナプスという組合せは、実ニューロンという視点からはやや見劣りがする。というのは、ヒトでは 4 万個/ニューロン、猿では 2 万個/

¹⁹ <http://www.research.ibm.com/articles/brain-chip.shtml>

²⁰ 聞き取りによれば、TrueNorth を用いたソフトウェア作成は、未だ極めて抽象度の低いトランジスタレベル設計に留まっているという。言い換えれば、NM-AI 開発では、誠に残念ながら、TrueNorth のようなハードウェア開発と Numennta 社の HTM (Hierarchical Temporal Memory) のようなソフトウェア開発とがまったく異なったグループによって独立的に行われているという現状である。

ニューロン、マウスでは8千個/ニューロン、ミツバチでも1千個/ニューロン²¹となっているからである(Rolls (2017))。後述するように、1ニューロンあたりのシナプス数が増えれば増えるほど、学習・記憶の空間・時間的な幅と深さが深まるので、認知可能な動的世界の解像度が高まり環境適応の幅と深さも増大する。したがって、NM-AIにとっても、豊富な樹状突起の有無は、重要な評価点なのである(Numenta(2017))。

現行のNM-AIハードウェアで電力消費量が格段に低下する大きな理由は、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、現在主流のコンピュータ回路は、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である。そのため、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。

なお、TrueNorthを含めた多くの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)などが使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。²²

C.2 欧州 Human Brain Project : BrainScaleS と SpinNaker

付論でも紹介しているヨーロッパ Human Brain Project のNM-AIプロジェクトである Heidelberg 大学を拠点とする NM-AI ハードウェア・BrainScaleS は、チップ上ではなく、直径 20cm の半導体ウェーハ(wafer)上に回路を集積化したウェーハスケール集積回路として実装されている。²³ BrainScaleS では、一つのウェーハ上に約 20 万個のニューロンと約 5 千万個のシナプスからなるネットワークを構築できる。そして、その場合の消費電力はわずか 1kW だという(堀尾(2017))。なお、BrainScaleS は、アナログ・デジタル混成の集積回路で構成されている。そのため、消費電力量が、デジタル回路だけの場合に比べて相当に低くなっている(堀尾(2017))。ただし、BrainScaleS のシナプス/ニューロン比率も 250 であるから、TrueNorth のシナプス/ニューロン比率とほぼ同じである。

加えて、BrainScaleS と同一プロジェクト内で並列で走っていた英国・Manchester 大学の SpinNaker プロジェクト (<https://www.spinnaker.io/>) では、専用チップを用いている IBM-TrueNorth プロジェクトとは異なり、最近ソフトバンク傘下に入った ARM 社提供の

²¹ <https://galton.uchicago.edu/~nbrunel/teaching/fall2016/11-intro.pdf> 参照。

²² 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。

²³ <http://brainscales.kip.uni-heidelberg.de/index.html>。本プロジェクトは 2015 年 3 月に終了

ARM968 というプロセッサコアを並列で百万個用いて 10 億個のニューロン網が実現されている。²⁴ つまり、低消費電力で動作する多数の ARM968 が並列分散処理 (PDP) システムとして非同期で動くニューロンシステムを作動させる仕組みである。また、18 個の ARM コア毎に 128M バイトの DRAM (Dynamic Random Access Memory) がシステムメモリとしてパッケージされているので、ハードウェア的にもほぼメモリ・論理共役/同居型の構成になっている。その結果、このような巨大なシステムの動作が、高々 90kW の消費電力で実現可能だとされている (詳細は付論参照)。Rast (2012) によれば、同論文の報告時点では一つのニューロンに 40 個のニューロンが最大連結可能とあるので、シナプス/ニューロン比率は 40 と低い。ただし、ニューロンは ARM コアでソフトウェア的に実現されているので、この比率も可変である。例えば、Stromatias 他 (2013) によれば最大で 326 になっているので、TrueNorth や BrainScaleS とほぼ同じレベルである。

C.3 Stanford 大学 Kwabena Boahen プロジェクト : Neurogrid

上記の三つの NM-AI に比べ、Kwabena Boahen@Stanford 大学発の Neurogrid と呼ばれる大脳 (新) 皮質のシミュレーションに特化した NM-AI では、ニューロン発火の際に重要な役割を果たす細胞体 (Soma) の細胞膜内外で電位差を発生させるタンパク質でできたイオンチャネルの仕組みまでもアナログ的に模倣している。同グループの名声を高めた Neurogrid は、16.5cmX19.1cm のボード上に実装された約 100 万個のニューロンと 400 万個の多重 (共用) 化シナプスからなる NM-AI ハードウェア・システムである。このボード上には、Neurocore と呼ばれる 12mmX14mm の中核チップが 16 個搭載されており、システム全体は僅か 3.1W の消費電力で作動するという。また、この多重化シナプスは、実シナプスに換算すると 60 億個のシナプスにも匹敵する。²⁵ つまり、シナプス/ニューロン比率は 6 千とほぼ先のマウス規模になっており、TrueNorth などと比べる一桁高い。

なお、Khodagholy 他 (2014) には、この Neurogrid をラットの大脳 (頭頂葉) 表面にある感覚野に埋め込んだ生体脳との融合実験とその際の発火データなどが示されている。したがって、最終的には、Brain-Machine Interface (BMI) としての利用も考えられている。事実、Boahen グループでは、人工網膜などの特殊なデバイス開発も行っている。²⁶

さらに、Arthur 他 (2005) では、STDP (spike timing-dependent plasticity)²⁷ と呼ば

²⁴ <http://apt.cs.manchester.ac.uk/projects/SpiNNaker/architecture/>

²⁵ <https://web.stanford.edu/group/brainsinsilicon/neurogrid.html>。この URL にも言及されているが、基本アイデアは、日本の牧野淳一郎教授 (神戸大学/理化学研究所/国立天文台) を中心とした土星の輪などの宇宙現象生成メカニズムの解明に特化したスーパーコンピュータ Grape に依拠している。

²⁶ https://web.stanford.edu/group/brainsinsilicon/kareem_bio.html。なお、人工網膜デバイス開発に関しては、大阪大学の八木康史教授グループが先駆者であるという。実際、聞き取りにより、八木先生と Boahen との接点があったことを確認した。八木研究室の人工網膜プロジェクトには、本論の著者の一人である今井も半導体集積回路の設計の立場から関与していた。

²⁷ Rolls (2017) には、STDP が次のように定義してある。“An interesting time-dependence of LTP and LTD has been observed, with LTP occurring especially when the presynaptic spikes precede by a

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な相互作用を組み込んだ STDP チップを開発している。そして、この STDP チップと事前・事後でソフトウェア的に再構成可能なハードウェアである FPGA 内に実現した実ニューロンに近い興奮型と抑制型の双方を含む自己回帰型 (Recurrent) 階層ニューロンシステム、RAM (Random Access Memory) チップなどを組み合わせたシステムボードを使って、STDP の存在が上記ニューロンシステムの安定性に大きな寄与をすることを示している。ボードには PC とのインタフェース用 USB ポートが備わっているので、PC 上のソフトウェアから制御する一種のアクセラレータ (加速機) ボードとも見なすことができる。ちなみに、TrueNorth や BrainScaleS、SpiNNaker にも STDP が組み込み可能となっている。²⁸

STDP が組み込まれると、ニューロン間の空間構造に加えて時系列構造を組み込むことができるし、このような時空間パターンを自律的に学習して記憶することにより先読み制御も可能になってくる (Rolls (2017))。ただし、Rolls は、STDP の重要性に関しては低頻度の刺激の時にしか成立していないのでは、といった疑問を呈している。さらに、Numenta(2017)では、シナプス前ニューロンとシナプス後ニューロンとのやり取りが高頻度になると実ニューロンによるタイミング検知が難しくなるのではないかとの判断に基づき STDP が組み込まれていない。

C.4 Windows/Linux 用アクセラレータとして実用化された BrainChip (2017)²⁹

BrainChip とは、少なくとも現時点では、主に監視カメラなどの動画情報をインプット情報として、そこからヒトの顔情報などを高速に抽出する NM-AI チップとして期待されている。実際の使用に際しては、BrainChip Studio という Windows/Linux 用ソフトウェアをインストールして BrainChip ボードを操作する。アクセラレータとしての BrainChip ボードは、PC 用として標準的な PCI-Express インタフェースに装着される。

この種のアクセラレータは BD-AI でも実用化が加速しているようであるが³⁰、同チップの開発者である Peter Van Der Made 氏の関連 US 公開特許 (US2010/0076916A1 と US2013/0297537A1) を参照する限り、正真正銘の正統派 NM-AI であることが分かる。具体的には、同特許や Van Der Made (2013) によれば、BrainChip は興奮性及び抑制性ニューロンの混合型であり、自己回帰構造を持つ階層型ニューロン構造 (ニューロン数は約 2 万個で各々

few MS the post-synaptic activation, and LTD occurring when the pre-synaptic spikes follow the post-synaptic activation by a few milliseconds (MS). This is referred to as spike timing-dependent plasticity, STDP.” 「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 (MS) 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。」(中馬訳、() 内は同名追加)。

²⁸ <http://www.artificialbrains.com/darpa-synapse-program>

²⁹ <http://www.brainchipinc.com/products/civil-surveillance-solutions/brainchip-accelerator>

³⁰ 例えば、日経エレクトロニクス 2017 年 10 月号『狙いは「リアルタイム AI」、推論用 AI チップの提案続出：データフロー型や FPGA、DRP に期待』参照。

のニューロンに数千個の樹状突起付随)、先に紹介した Boahen 的な細胞体の膜電位調整機構なども導入されている。さらに、前述の Boahen グループの成果の所で触れた STDP の組込による時空間パターンのオンライン学習が実行可能になっている。これらの仕組みは、実際には、専用チップではなく、事前・事後の再構成が可能な FPGA にソフトウェア的に組み込まれている形である。

ビジネスモデルとしては、このようにして FPGA にソフトウェア的に組み込みが可能な半導体 IP (知的財産権) を、(ソフトバンク傘下となった)英国 ARM 社のように、ソフトウェアとしてさまざまな契約形態で販売したり、ユーザ需要に合わせて込むためのソリューションビジネスを提供したりすることを目論んでいる。もちろん、FPGA ではなく専用半導体チップにも組み込むことができる。Wall Street ジャーナルによれば、今年に入って日本のみずほグループを含めた複数の VC からの資金集めを活発化することに成功している。

気になる NM-AI としての性能であるが、同社の HP には、下記のように記されている。

“In a recent field trial, BrainChip Studio detected, extracted and classified in real-time more than 500,000 facial images during 3 1/2 hours of video across eight different cameras. In a separate trial, it processed 36 hours of recorded video in less than two hours, extracting over 150,000 facial images.” (「最近の実際の現場での試行では、(Windows/Linux で動作する) BrainChip Studio は、8つの異なるカメラで撮影した3時間半のビデオ映像からリアルタイムで50万個の顔の画像を検知・導出・分類した。別の試行では、36時間分の撮影済みビデオを2時間未満で処理し15万個の顔の画像を抽出した。」(中馬訳、()内は同名追加)

この分野の素人であるから判断能力はかなり劣るが、カタログ通りだとすると、現段階の BD-AI では実用化されていないオンライン学習(on-line learning)機能が実現されているし、処理速度も相当に速いのではないだろうか。

しかも、これまで世界中の各所でほぼ死蔵状態だった莫大な数の防犯ビデオ映像等の分析を高速に実行することができるようになるとすると、潜在需要もかなり大きいと思われる。しかも、原理的にはビデオ映像だけではなく、様々な時空間データをオンラインで検知・分析・学習・記憶できるわけであるから、場合によっては大きな社会変革をもたらす可能性もある。

ただし、Van Der Made 氏の特許内容を表面的に参照する限り、フルデジタルということであるが、前述の Boahen グループの(特許化されていないと推測される) STDP ボードや膜電位調整機構をも持つ興奮・抑制混合型ニューロンシステム等々との類似性も高いように感じられる。したがって、この NM-AI チップがどの程度ユニークであり、どのような大きなインパクトを社会に与えることができるのかは、少なくとも門外漢である筆者には判断

が難しい。ただし、NM-AI の発展段階から予想すると、このようなアクセラレータ的な形での NM-AI の登場が、これから急増していくのではないだろうか。

C.5 Intel の Neuromorphic チップ “Loihi”³¹

ハワイ島沖の巨大海底火山を意味する “Loihi”³²は、2017 年 9 月下旬に発表されたばかりの試作チップである。本格発表は同年 11 月ということで、詳細はほとんどオープンにされていないが、当初は、上記の BrainChip と同じく、そのリアルタイム学習性を活かした防犯／監視ビデオ解析用などをターゲットにしているようだ。ただし、発表時のスペックは、いずれもフルデジタル化された 13 万個のニューロンと 1 億 3000 万個のシナプスということであるから、NM-AI としてはかなり小規模のようである。より具体的には、128 個のクラスタ各々に 1024 個のニューロンが含まれる形のようなのだ。構成から判断すると、TrueNorth や SpiNNaker に似ているということだろうか。ただし、シナプス／ニューロン比率は 1000 であるから、その点では、これまで紹介してきた中では BrainGrid や BrainChip に近い。消費電力効率は、非同期でイベント駆動型という特徴を反映して、汎用プロセッサの 1/1000 を達成しているとのこと。

また、インテルの最先端微細化プロセス (14nm) で製造されており、しかも、インテル独自の X86 プロセッサと Loihi とがワンパッケージ化されたユニークな構造をしている。ということは、Windows や Linux などの標準 OS あるいはその簡易版から自在にアクセス可能ということになる。具体的には、X86 プロセッサが Loihi 用の前処理（刺激・イベント情報のスパイク電流フォーマットへの変換とその Loihi への転送など）を担当するという。

インテルでは Neuromorphic チップ開発が、買収済みの Movidius 社や Fulcrum Microsystems 社などからの技術を結集して 2011 年頃に開始されたという。Loihi の本格デビューは 2018 年度上半期との触れ込みであるし、あのインテルが NM-AI にも本腰を入れはじめたとすると、NM-AI 実用化の波が今後次々に訪れてくるのではないだろうか。

2. 2 低消費電力性を誇る NM-AI を支える仕組み：脳模倣の意義を探る

先の Burr (2015) の指摘のように“脳機能が十分に解明されていない”現状ではあるが、NM-AI が格段の低消費電力効率によって“明日の AI を担うとされる”根拠は、門外漢にも十分に理解可能である。その大きな理由は、脳内での処理が、先に述べた非同期・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/ニューロン集団間ネットワーク処理・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（例えば、

³¹ 本節の記事は、<https://newsroom.intel.com/editorials/intels-new-self-learning-chip-promises-accelerate-artificial-intelligence/> 及び <http://www.zdnet.com/article/why-intel-built-a-neuromorphic-chip/> を参照している。

³² https://en.wikipedia.org/wiki/L%C5%8D%CA%BBihi_Seamount を参照。

Rolls (2017)、Eliasmith(2013)など参照)。そして、高度に階層モジュール化されたニューロン/ニューロン集団間ネットワーク処理・決定システムの動作に必須なものが、やや聞き慣れない(後述する)学習・記憶の高次元スパース分散表現(Sparse Distributed Representation: SDR)化ということになる。そして、階層モジュール化とSDR化とが切っても切れない関係にあることは、モジュール設計思想の普遍性という意味でも誠に興味深い。

これらの点をより直感的に理解するために、長くなって誠に申し訳ないが、ニューロンのそもそもの特性(含むBD-AIやNM-AIによるその模倣形態)、脳内におけるニューロン/シナプス主体の学習・記憶方式、それらの学習・記憶方式を利用してニューロン/ニューロン集団で行われる類似性・新奇性判断(パターンマッチング)などについてまず簡単に触れておきたい。

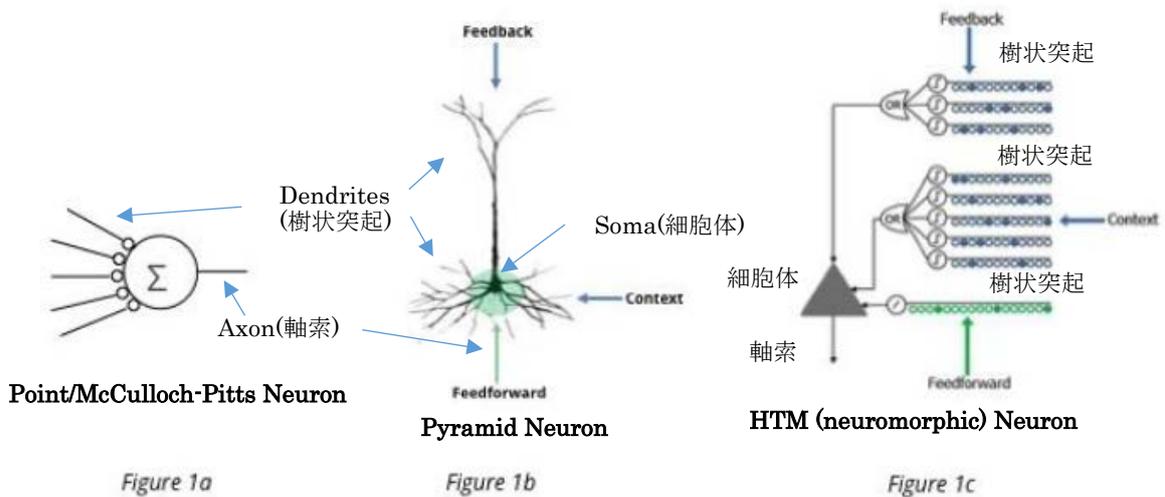
A. 実ニューロンと人工ニューロン

BD-AIで使用される人工ニューロンは一種類のみであり、その中核を成す細胞体(soma)へのインプット部分を構成する樹状突起(dendrite)が数本、そのアウトプット部分を構成する軸索(axon)が終端分岐としての側枝(collateral)を伴わない真っ直ぐな1本という極めて単純な形をしている(下図のFigure 1a参照)。このような単純構造を持つニューロンは、“Point Neuron”(Numenta(2017))とか“McCulloch-Pitts Neuron”(Anderson(2017)、Rolls(2017))と呼ばれ、McCulloch and Pitts(1943)以来、BD-AIの核となる人工ニューロンとして長期間使われてきている(Schuman他(2017))。³³ また、パーセプトロンと呼ばれる多層構造を持つBD-AI(甘利(2006)参照)の場合、インプットの樹状突起が1本に対して軸索が複数で、場合によっては100本を遙かに超えるような場合もある。こうなってくると、ニューロンというより単なるネットワークの端点(Node)という風である。

NM-AIが模倣の対象とする実物のニューロンには大きく分けると全体の80%を占める興奮性ニューロン(excitatory neuron)と20%を占める抑制性ニューロン(inhibitory neuron)とがあり、大きな構造的違いでは3種類、小さな違いまでも含めると数多くの種類がある。しかも、典型的な興奮性ニューロンで全体の75%を占めるピラミッド型錐体ニューロン(pyramidal cells)には、ヒトの場合5000-20000/ニューロンの樹状突起がある(Rolls(2017))。なお、図1-3には、Numenta(2017)のHTM(Hierarchical Temporal Memory)と呼ばれるNM-AIソフトウェアに組み込まれている実ニューロンをより忠実に模倣した人工ニューロンの事例が示されている(各樹状突起の役割については後述)。

³³ 1980年代末の当時AI研究の第一人者だったWinograd他(1987)は、このような単純なニューロン使用に関して“The guiding image was that of a large number of essentially uniform neurons with widespread and essentially uniform potential for interconnections. Work over the years in neuroanatomy and neurophysiology has demonstrated that living organisms do not fit this image.”

(「(現行AIの)基本イメージは、広範囲で本質的に一様な配線可能性を持つ数多くの本質的に一様なニューロンからなっている。神経解剖学や神経生理学の長年にわたる研究では、生命体はこのようなイメージに合致していないことを明らかにしている。」(中馬訳、()内は同名追加))とコメントしている。



Numenta (2017) より抜粋：ニューロン名は筆者追加

ニューロンの中心にある細胞体(Soma)は、これらの樹状突起先端のシナプス (synapse) 経路で集められた他のニューロンからの刺激・イベント情報を集積し、その強さに応じて軸索に発火/活動電位 (Action Potential) という形でアウトプットを吐き出す。その際に、興奮性ニューロンから入ってきた刺激・イベントは発火の可能性や強さを高め、抑制性ニューロンから入ってきた刺激・イベントは発火の可能性や強さを低める。シナプス自体は、はじめから興奮型か抑制型化のどちらかに特化している。

さらに、ニューロン発火の時空間パターン³⁴は、ニューロン間シナプス結合の強さと拮がりという形で学習・記憶されていく。その結果、過去の学習・記憶パターンに類似した刺激・イベントが到着する兆候があると、その後に発生するであろう状況の予測もできるようになる。しかも、このような予測精度には、後述するように、樹状突起の豊富さが大きく関わってくる (Numenta (2007)、Rolls (2017))。

このように、ニューロンネットワークシステムには、興奮・抑制混合型ニューロンの正負の相互作用によって、その安定性を自動的に高める仕組みや将来予測をよりの確に行うための仕組みが組み込まれている。³⁵ 先に紹介した TrueNorth、BrainScaleS、SpinNaker、

³⁴ 発火パターンにも、標準 (standard) 型、高速 (fast) 型、連続 (continuous) 型、Burst (突発) 型といった種類があり、それぞれ役割が異なっている (Trappenberg (2010))。

³⁵ 豊泉 (2017) によれば、当初は興奮性ニューロンのみが作動しているが、その後に抑制性ニューロンができてくる臨界期 (Critical Period) と呼ばれる時期になると、興奮性ニューロンがより強い刺激でしか発火しなくなる。豊泉は、このような現象を、臨界期以前=脳の内部の情報を元に神経回路を構築する時期、臨界期=脳の外部の情報を元に神経回路を構築する時期と位置づけ、臨界期以前は遺伝等の内部発火パターンが主流であり、その後に外部起因の発火パターンが優勢になるとしている。実際、ヒトの場合、3歳頃になると、当初興奮性ニューロンしか存在しなかった状態から、抑制性ニューロンが生まれはじめ、当初配線されていなかった脳内に行き渡り始める。そうすると、より高度な学習プロセスによる脳内制御

NeuroGrid、BrainChip はもちろんこと、NM-AI の多くが、このような興奮・抑制混合型のニューロンを模倣している理由である。他方、多くのBD-AI に組み込まれているのは興奮性の単純な Point Neuron だけなので、このような自動安定化機能は存在しない。

B. ニューロン/ニューロン集団での学習・記憶蓄積の仕組み

脳内での特定の学習・記憶は、ニューロンに付随する多数のシナプスを媒介としたニューロン間ネットワーク結合の強さと拡がりという形で刻み込まれていく。その際、特定の記憶が特定のひとつだけのニューロンに記憶・活用されるのではなく、刺激・イベントの到着によって同時に活性化した一連のニューロン集団によって分散表現 (distributed representation: DR) される。また、特定のニューロンは、特定の刺激・イベントの到着が刻んだ記憶だけに結びついているのではなく、その他の様々な刺激・イベントの到着に際して同時に活発化するニューロンネットワークの一員としてもオーバーラップ (重複) して活用される。そして、このオーバーラップの多寡が連想記憶の幅と深さに繋がっていく。

さらに、そのようなネットワークを1成分とするネットワーク、つまり、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、・・・といった複雑な階層構造を経る形で学習・記憶の抽象度を上げていく (例えば、Anderson (2017) 、Eliasmith (2013) など参照)。

上記のような同時に活性化して分散表現を担うニューロン集団は、小集団から大集団に至る数種の柱状モジュールから構成される典型的な階層モジュール構造をしている (Anderson (2017))。具体的には、百個程の(ニューロン)集団から成るミニ柱状モジュール (minicolumn) を基礎単位として、それらを百個ほど束ねた一万個ほどの集団から成る機能柱状モジュール (functional-column)、それらをさらに百個ほど束ねた側頭葉に存在する“顔領域” (facial domain) などの特化領域を形成する百万個程度の集団から成る領域柱状モジュール (Domain Column)、そのまた百個ほどを束ねたブローカー野 (言語野) などと呼ばれる1億個程度の集団から成る領野柱状モジュール (regional column) という風に構成されている。

言い換えれば、大脳(新)皮質では、モジュール・システムアーキテクチャの基本的な特徴である、個々の構成モジュールの独立性が高い、構成モジュール間の相互依存性が少ない、構成モジュール間のインターフェース(シナプス)が標準化されている、階層型のモジュール構造をしている、という典型的な特徴が存在している。

各種柱状モジュールを構成するニューロン集団には、特定の刺激・イベントの到着と共に同時活性化する部分と活性化しない部分とがある。その際に、これらの集団内のニューロンは、活性化しているか否かにかかわらず、一つ一つのニューロンを異なった成分とする高

が可能になる。臨界期は脳の各部位で異なっている。例えば、視覚の臨界期は8歳前後であり、この時までだったら弱視の治療可能である。学習信号が内部だけではなく外部からも数多く入ってくる臨界期以降は世界をより高い解像度で理解できるようになる。実に興味深い進化の仕組みである。

次元ベクトルとみなすことができる。そして、この高次元ベクトルは、高次元分散表現 (distributed representation : DR) ベクトルと呼ばれている (Rolls (2017))。

誠に興味深い事実であるが、上記のニューロン集団という視点で定義された特定記憶を表現する高次元 DR ベクトルは、ほぼ全てスパース表現あるいはスパース符号化 (sparse coding) されているという特徴を持つ。スパース表現/符号化とは、上記の柱状モジュールと言えば、各々の集団を表す高次元ベクトルの半分以上の成分が刺激・イベントの到着に際して活性化/発火しないという状態をさす。そして、情報の抽象度が上がれば上がるほど、ニューロン (モジュール) 集団内で活発に動作するニューロンの数が減ってくる。つまり、スパース分散表現 (SDR) のスパース性が増大していく。ただし、抽象度が上がって行っても、単一のニューロンだけが活性化/発火するといった状態にはならない (Anderson (2017))。

実際、Rolls (2017)によれば、脳内ではスパース性が顕著である。例えば、6層からなるヒトの大脳新皮質の記憶中枢部分である (軸索が終端で枝分かれして自己再帰的に反回してきた反回側枝による) 自己再帰型構造をした第2層・第3層では平均10%、海馬の同様の自己再帰構造をした CA3 という領域では平均4%のスパース率だという。このような特徴を持つニューロン集団を表現する高次元ベクトルが、**高次元スパース分散表現 (SDR) ベクトル**と呼ばれる。

以下では、この高次元スパース分散表現ベクトルを、簡略化して高次元 SDR と呼ぶことにする。高次元 SDR は、次元の高低の違いはあるものの、哺乳類のみならず昆虫にまで共通に用いられており、進化の荒波の中で超長期にわたって生き残ってきている特徴である (Jayaraman 他 (2009)、Lotto (2009))。

なお、前述の興奮性と抑制性の二つのニューロンの存在は、スパース分散表現にも深く関わっている。事実、一つのニューロンがある刺激・イベントに対して反応しやすくなった場合、興奮性ニューロンだけしか存在しないと、その周辺ニューロンの反応性も上昇させ、結果として互いの反応性を高め合うポジティブフィードバック状態 (極端な例ではてんかん症状など) が産み出される可能性がある。脳内では、このような過敏/暴走状態ができるだけ起こらないように、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抑制性ニューロンによって発火しないように押さえられる競合学習 (competitive learning) や側方抑制 (lateral inhibition) と呼ばれる仕組みが働く。その結果、周辺ニューロン中の一つ、あるいは数個だけが発火して、スパース符号化に貢献する (甘利 (2008)、Rolls (2017)、Anderson (2017))。

例えば、前節で紹介した TrueNorth、BrainScaleS、SpiNNaker、NeuroGrid、BrainChip、などにもハードウェアとしてこのような競合学習/側方抑制の仕組みが組み込まれている。また 前述した Numenta (2017) の HTM (Hierarchical Temporal Memory) にも、このような競合学習/側方抑制機能が組み込まれている。逆に言えば、単純な興奮性ニューロンだけが組み込まれた BD-AI では、脳内で重要な役割を果たしているスパース分散表現を内生的に産み出すことはなかなか難しい。この点に関する Anderson (2007) の下記のコメントは、脳内

の構成原理と半導体デバイス設計原理の類似性を彷彿させ、誠に興味深い。

“Connections are expensive biologically since they take up space, use energy, and are hard to wire up correctly. The connections that exist are precious, and their pattern of connection must be under tight control. These issues provide severe constraints for the structure of large-scale brain models. It does not do to just say that everything connects to everything, because it doesn't.” (「(神経回路での)配線は、場所を取り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、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、そうになっていないからだ。」(中馬訳)

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・イベントは、高次元 SDR の形で次々に新たなニューロンネットワークとして学習・記憶されていく。そして、その際には、新たな高次元 SDR が、既にニューロンネットワーク内に高次元 SDR として蓄えられている多くの記憶と比較対照される形での類似性・新奇性判断がなされる。より具体的には、そのような比較対照・判断が、新旧の高次元 SDR の内積値を計算するという形で行われる(Rolls (2007))。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。

脳内では、この値がある閾値を超えるか否かで類似性・新奇性判断が行われているというから驚きである。ちなみに、内積計算は、高次元 SDR の次元には関係せず、それがスペースであればあるほど計算量が少なくて済むので省エネ性が高くなる。共にゼロでない成分同士を掛け合わせるだけの簡易計算で済むからである。しかも、このような比較対照・判断が、様々なニューロンネットワークで並列に実行される。我々の連想が、低速なニューロン流コンピュータに依存しながらも、高速に実行される理由の一つである。

なお、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經由して最終的に脳内のニューロン集団に到着した時にどれほどまで高次元 SDR 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 がやり取りされて抽象度が上昇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(ただし、Olshausen(2004)などは参照)。³⁶ 特に、ネットワークのネットワーク(networks of

³⁶ 例えば、感覚刺激・イベントの大脳皮質への上向き/求心的な(afferent)流れの場合、感覚器官→末梢神経→中枢神経→脳幹→視床→各種大脳皮質→前頭前皮質(短期ワーキングメモリ)→海馬(エピソード記憶化)→各種大脳皮質(意味記憶化)といった経路を経る。なお、嗅覚の場合、視床を經由しない。

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説明が進んでいないようである (Anderson(2017)、Rolls(2017)、Eliasmith(2013))。

C. 低消費電力化や脳内動作の確率過程化に必須の高次元 SDR

では、特定ニューロン集団内の記憶は、なぜそもそもスパース表現/符号化される必要があるのだろうか？そして、SDR(スパース分散表現)で高次元である必要性はどこにあるのだろうか？実は、進化的な視点から見ると、その理由に**低消費電力化への自然淘汰圧**が大きく関わっていることが分かる。

C.1 スパース表現/符号化される必要性

スパース表現/符号化の特徴は、ニューロンが刺激・イベント到着によって活性化して発火する際に、高次元 SDR 内のわずかな比率のシナプス/ニューロンしか活性化させないということであった。その第一の理由は、活性化/発火するニューロンの数が増えれば増えるほど、送信しなければならない活動電位が急増するので、エネルギー消費量が増えてしまうからである。他の条件が一定であれば、エネルギー消費量が減れば減るほど生存確率が上がる。実際、ニューロンを使った処理は、他の細胞に比べるとエネルギー的にも高価である。ヒトの場合でも、あのわずか 1-2 キログラムの重さの脳だけで体全体の 20-25%のエネルギーを消費する (Eliasmith(2013))。

ただし、生存確率は、使用するニューロンの数を減らすだけではなく、情報処理の幅と深さを広げて自らの環境適応力を上げることで増大する。例えば、ヒトの視覚情報の場合、驚きであるが、1秒あたり1ギガビットの画像情報に相当するという。ニューロンは、このような巨大データから生存に貢献する情報をエネルギー効率的に濾し取らなければならない (Olshausen 他(2017))。

その方法の1つは、画像や音声などの形で到着する刺激・イベントを可能な限り単純な基本単位(モジュール)に分割し、より複雑な情報をこれらの基本モジュールの組合せ、さらにそのような組合せの組合せ等々という風にして情報表現・処理能力の幅と深さを拡大することである。そうすれば、学習・記憶パターンの再利用性を高めることもできる。例えば、画像であれば、特定形状の線分だけに分業・特化したニューロン群によるネットワークを構成して学習・記憶するということになる。

そうすれば、特定の刺激・イベントに対して使用されるニューロン数を減らせるし、新しい記憶を形成するために必要な修正箇所を少なくできるし、特化という形で一つ一つのニューロンの処理効率を向上させることもできる。さらに、Olshausen 他 (1996)がシミュレーションで示しているように、基本モジュールがレゴブロックのように単純化されていればいるほど、組合せの自由度が増すので、それらを複雑に組み合わせる形での合成・抽象

化がやりやすくなる。まるでアダムスミス流の市場規模が拡大すればするほど分業・協業の利益が増大していくというロジックを地で行くようなシミュレーション結果である。

なお、SDR は DDR に比べてベクトル単体としては表現力には劣るが、このようなデメリットはニューロン数が多ければ問題にならなくなることは直ぐに理解できる。このニューロン数の問題に関しては、哺乳類どころか、既に節足動物レベルでも解決されている。例えば、節足動物の大部分を占める昆虫類の脳とも言われるキノコ体を構成しているケニオン細胞 (Kenyon cell : KC) は昆虫の嗅覚システムの中で最も多い中枢ニューロンということであるが、ゴキブリでも 330 万個、昆虫類のキノコ体を留めている生きた節足動物の化石であるカブトガニでも 1000 万個ある (Jayaraman 他 (2009)、Lotto (2009))。³⁷

このことから理解できるように、スパース表現/符号化とは、エネルギーを可能な限り節約しながらも、到着した刺激・イベント情報処理の幅と深さを広げるための徹底した階層モジュール化とニューロン間インタフェースとしてのシナプスによって広汎なネットワークを構築するための仕組みだと類推できる。

スパース表現/符号化によってもたらされる階層モジュール構造は、脳構造にさらなる便益をもたらす。それは、SDR を用いると、DDR に比べてニューロン集団間に蓄積される記憶パターン間の相互干渉 (Cross-Talk) が少なくなるので、情報重複のみならず情報転送・応答の際のノイズ (雑音) を大きく削減できるという点である。その結果、記憶容量を大幅に増やすことができる。³⁸ 加えて、ニューロンのスパイク (発火) に関連したノイズは、スパース化によってより少なくなるので、脳内ニューロンネットワークシステムの安定性が増し、より正確な決定/記憶の想起ができるようになる (Rolls (2017))。³⁹

C.2 SDR が高次元化される必要性

SDR が高次元化されると、当然のことながら、その他の条件を一定とすると、表現の自由度が増大する。前述したように、脳内でのニューロン集団視点での類似性・新奇性判断は、SDR 構成するどのニューロン成分が同時に活性化したかを SDR 同士の内積値を比較する形で実行されている。しかも、その際、SDR が高次元化すればするほど発火ノイズやニューロン/シナプスの欠落などによる誤差発生への頑健性が増すので、比較対照される SDR 同士の内積値で表現される類似性・新奇性の判断に影響が及ばなくなる。つまり、信号理論で言われる “graceful degradation (耐障害性)”⁴⁰ が得られるようになる (Kanerva (2014))。

このような耐障害性は、各種脳内処理・決定の確率過程化が実現できるとさらに増大する。実際、Rolls (2017) に示されているように、各々のニューロンは確率的な意味で独立に

³⁷ 昆虫類は節足動物中の陸生・甲殻類から進化してきたと考えられているが、現在の甲殻類の脳には、カブトガニのような生きた化石をのぞくと、キノコ体は存在しない (水波 (2009))。

³⁸ 実際の記憶容量計算例に関しては、例えば、Rolls (2017)、甘利 (2008) など参照。

³⁹ スパイク (発火) 関連ノイズの発生要因に関しては、Rolls (2017) に網羅的に示されている。

⁴⁰ 故障が起きた際に、機能が完全に停止するのではなく、システムに支障をもたらさない形である程度まで正常な機能が維持され続けること。

動作しており、ニューロン間の時系列的な推移は、離散的なスパイク（発火）によって引き起こされるポアソン確率過程 (Poisson Jump Process) ⁴¹に従うと想定できる。したがって、ニューロン依存的な脳内処理は、そもそも確率過程にしたがわざるを得ない。そして、このような処理・決定の確率過程化は、様々な処理上の便益をもたらす(Rolls(2017))。

例えば、SDR が高次元 n の確率変数だとすると、このような SDR が共に $w (< n)$ 個だけがゼロでない場合、任意の 2 つの SDR が同一である確率は、限りなくゼロに近い。例えば Numenta(2017)にしたがって $n=2048$ 、 $w=2$ とした場合でも、一致確率は、52 万 3776 分の 1 になる。このことは、次元 n が大きければ大きいほど、到着した刺激・イベントに対応する SDR の一部分でも既学習・記憶済みの SDR と一致していれば、二つの SDR の一致する確率が増えていくことを意味する。つまり、SDR が高次元確率ベクトルであればあるほど、上記の許容誤差や耐障害性が高まると共に、類似性・新奇性判断の正確さとスピードが増す。

さらに、SDR が高次元化される必要性は、エネルギー節約的な確率決定過程と分かちがたく結びついている。例えば、前述の 1 秒あたり 1 ギガビットもの画像情報が引切り無しに到着するケースを考えてみよう。これほどの情報量になってくると、各時点・時点で到着する空間情報を別々に独立なもののみなして処理しては、いくらスパース表現/符号化しても、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するデータ数を激減させる工夫が必須となる。そのためには、時空間データの中に繰り返して現れる統計学的な関係を利用することが極めてエネルギー節約的となる(以上は Lotto (2009)参照)。統計学的な規則的な関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。

良く知られているように(例えば Numenta (2017)参照)、そして、先のオンライン学習を得意とする BrainChip などの NM-AI 事例を紹介する際にも言及したように、上記のような確率変数としての高次元 SDR を用いない BD-AI は、このような時空間データの迅速な処理が不得意である。

D. 多彩な高次元 SDR 表現に必須な樹状突起

これまでは、SDR に関する叙述は、ニューロン/ニューロン集団中心のものであった。ただし、SDR としての表現の幅と深さは、ニューロン内の樹状突起構造をどれほど脳模倣型にできるかによって大きく規定される。ところが、前述のように、BD-AI 型の人工ニューロンでは、樹状突起そのものの役割にほとんど注意が払われてきていない(Numenta (2017))。以下では、この点の NM-AI との違いについて、高次元 SDR に関連させながら触れてみたい。

⁴¹ 「ポアソン過程 (Poisson process) は、ランダムに (離散) 生起する事象を表す基本的な確率過程で、客の到着や故障の発生、個体の出生など様々な現象のモデル化に使われる。」
(<http://www.orsj.or.jp/~wiki/wiki/index.php/ポアソン過程>)

脳内外から到着した刺激・イベントをより忠実に表現した高次元 SDR を生み出すのが、数多くのニューロンからシナプスを介して入力を受ける樹状突起である。ニューロン間ネットワーク表現としての高次元 SDR やその時系列によって表現される時空間記憶は、これらの樹状突起・軸索終端におけるシナプス結合の強さと幅で特徴づけられるからである。ところが、Anderson (2017)や Rinkus 他(2016)が指摘しているように、BD-AI には、前述のように、競合学習/側方抑制を含むスパース化のための仕組みが組み込まれていない。そのため、スパース (Sparse) ではなく密 (Dense) な分散表現 (Distributed Representation)、つまり、DDR になりがちである。密という意味は、繰り返しになるが、ベクトル成分の半分以上がゼロ以外の値をとるという意味である (Rolls(2017))。

高次元 SDR の獲得には、上図 (Figure 1b) に示されるように、実ニューロンの樹状突起の数が十分に多い必要がある。しかも、空間パターンのみならず時系列パターンを表現できるためには、時間的な継起表現もできるような構造である必要がある (Numenta (2017))。それが、実ニューロンに下記に説明する二種類の樹状突起が存在している理由でもある。

ちなみに、BD-AI で時系列パターンの継起を表現する際には、モデル的には複雑な時系列上の入れ子構造を採る RNN (Recurrent Neural Network) が用いられている。RNN の基本的な仕組みは、BD-AI に時系列概念を組み込むために必須であるが、豊富な時系列情報を組み込もうとすればするほど短期・長期の記憶を保持しておくための複雑な入れ子構造や特殊な外部 (連想) 記憶装置が必須となるので (Graves (2016)、岡野原(2017))、実際の脳内の仕組みに対応して理解することがとても難しくなる。門外漢の感想で誠に申し訳ないが、冒頭で言及した BD-AI と Neuroscience との乖離が、最近になればなるほど顕著になってきている理由の一つだと思われる。なお、RNN についての紹介は本論では割愛するが、浅川 (2017)や坪井ほか(2017)に分かりやすい説明がある。

実ニューロンが空間パターンと時系列パターンとを樹状突起でどのように捉えようとしているかについて、もう少し見てみよう。前述の興奮性ニューロンとして典型的な錐体ニューロンの場合、1つは、細胞体のはるか上部にあって他のニューロンから到着する各種の刺激・イベントをシナプス経由で広範囲に受け取る上方向に扇状に広がった先端樹状突起 (apical dendrites) である。もう一つは、細胞体の近傍にあって主に他のニューロンからシナプス経由で送られてきた学習済みの数多くの刺激・イベントを条件刺激コンテキストとして提供する下方向に扇状に広がった基底樹状突起 (basal dendrites) である。

さらに、前述のように錐体ニューロンのシナプスの数は人間で約 4 万個ということであるが、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シナプスの卵とも呼べるようなトゲ上の微細な突起が上記の樹状突起の表面にシナプス予備軍として数多く存在する。また、シナプスには興奮性のものと抑制性のものがあり、両者が途中で入れ替わることはない。

樹状突起の複雑さは、これだけに留まらない。Numenta(2017)によれば、基底樹状突起の中でも、細胞体に距離的に近い部分 (proximal dendrites) は先端樹状突起とほぼ同じ機

能を持つが、それらの外側の部分 (distal dendrites) は、他のニューロンと呼応して細胞体発火時に既存の学習・記憶結果としてのコンテキスト情報を伝えている。その結果、細胞体が、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))。2 番目と 3 番目の状態が、既存学習に基づく先読みによる準備と考えられる。⁴² ただし、Rolls (2017) が言及しているように、このような樹状突起間の分業状況は、ニューロンの種類によってもかなり異なっており、そもそもまだ十分には分かっていないようだ。

E. 高次元 SDR に基づく類似性・新奇性判断: その低消費電力性の例示

先に言及した NM-AI が模倣に拘る脳内での類似性・新奇性判断の低消費電力性を確認するために、もう少し具体例に立ち入って確認してみたい。着目する事例は、長期記憶が蓄えられている大脳新皮質部分のニューロン集団である。⁴³

大脳新皮質はすべての領域で 6 層の柱状構造をしているが、長期記憶のほとんどは第 2 層と第 3 層に広がる錐体ニューロン間ネットワーク内に蓄積されている。しかも、そこでは、錐体ニューロンの軸索が、記憶容量拡大や配線長最短化 (したがって脳サイズの縮小) 等に必須の自己再帰性を実現するために、反回側枝 (recurrent collaterals) と呼ばれる軸索が終端で枝分かれして自己再帰的に反回する形でニューロン集団が互いにシナプス結合されている。⁴⁴ Rolls (2017) は、このような自己再帰型ネットワークを Self-Associative Attractor Networks と呼ぶ。

このような高次元 SDR の状態は、単純化してしまうと、例えば、アクティブなニューロンを 1、非アクティブなニューロンが 0 となっている要素を持つ下図のような高次元、ここでは仮想的な 1 万次元のベクトル) として具体的に表現できる。このベクトル SDR A、SDR B、SDR C の各々の要素は、柱状構造内の (第 2 層と第 3 層の) 特定ニューロンを指しているとする。そして、この特定ニューロンが刺激・イベント A、B、C に晒されたときに示すニューロン集団の反応パターンが、下図で SDR A、SDR B、SDR C と名付けられている。繰り返しになるが、ベクトル内の特定の成分は、特定ニューロンに対応している。上記の例では、0-1 成分の出現パターンでは、SDR A と SDR B を生み出した刺激同士には 3 箇所と同じニューロンがシナプス結合しているという共通点があるが、両者は SDR C を生み出した刺激・イベントとはどこにも共通点がない。

⁴² 例えば、東京にある勤務先周辺であれば仕事仲間に合えば彼らの顔を迅速に識別できるが、勤務先から遥か離れたロンドンの地下鉄の中で東京にある勤務先での仕事仲間会っても、彼らがロンドンにきているという事前情報が無ければ、彼らの顔をなかなか識別できない。素人風のおおざっぱな解釈で申し訳ないが、このように、先端樹状突起で受けた一般の顔情報が、基底樹状突起で受け取ったコンテキスト (先読み) 情報と合わさって特定の顔情報として迅速に認知される、という風なことだと思われる。

⁴³ もちろん、脳内では、様々な箇所でも種類の高次元 SDR が生み出されており、それらの箇所でのスパース性もかなり異なっている (Rolls (2017))。

⁴⁴ この辺りの説明は、主に Rolls (2017)、甘利 (2008) や Anderson(2016)を参照。

良く知られているように、歯状回は、ヒトが大人になってもニューロンが新生する数少ない箇所の一つとして有名である。この歯状回には、包括的な感覚情報の連合を司る側頭葉の中の嗅内皮質と呼ばれる領域から貫通線維と呼ばれる経路を通して新しい刺激・イベントが次々に到達する(Rolls (2017))。そして、そのような刺激・イベントが嗅内皮質内の10個のニューロン集団によって歯状回に伝えられたとしよう。そうすると、この歯状回では、この10個のニューロン集団に対してその約10倍の100個のニューロン集団が(ランダムに)割り当てられる。つまり、10個のニューロンで表現されていたパターン情報が、100個のニューロンに拡散する形でスパース分散表現(SDR)化がさらに高次元化される形で格段に強化される(Poldrack 他(2010))。

その結果、嗅内皮質でのオリジナルなパターンが、歯状回でさらに細分化されたパターンとして表現され、当初の高次元SDR(ベクトル)の直交性が増大するのである(Poldrack 他(2010)、Greve 他(2015)、Finnegan (2017)、Rolls (2017))。そして、この直交性が増したパターン情報が、苔状繊維と呼ばれる経路を通して先の反回側枝によって形作られる**大局(グローバル)構造とポジティブフィードバック特性**を持つ海馬内のCA3記憶領域に転送されて自己(Self)視点のエピソード記憶として蓄積されていく。また、非自己視点のより抽象的で一般的な意味記憶は、海馬内でCA3→CA2→CA1という経路を辿り、そこから側頭葉内の嗅内皮質/海馬台を経由して側頭葉内に長期記憶されていく。

なお、上記で大局構造という意味は、CA3内のどこから出発しても、CA3領域内の全ての記憶にアクセスできるという意味での大局的な連結の仕方をしているということである。また、ポジティブフィードバック特性とは、特定の記憶断片を基点に記憶のアクセスがより広域にどんどん広がっていくという意味である。⁴⁵ ちなみに、大脳新皮質の場合、海馬内のCA3領域とは大きく異なり、ネットワーク構造は大局的ではない。つまり、ネットワーク構造に局所的に閉じた部分が存在するので、特定の記憶断片を辿っていくだけでは辿り着けない記憶が多々存在する。⁴⁶

直交ベクトルとは線形代数学の話と想像していたら、脳内で多用されている……。その巧妙な低消費電力化の仕組みに驚嘆させられる。なお、前述したように、確率変数としての分散表現が高次元化しがちであり、しかも、それらの高次元SDRの間に直交性が生まれがちとなる数学的な理由に関しては、Kanerva (1988)に示されている(後述)。さらに、Kanerva(1988)の理論上の提案(“Sparse Distributed Memories”(SDM)と呼ばれる)が、ヒトの実際の脳内記憶システムと酷似していることに関しては後ほど一部紹介する。なお、SDMに関する分かりやすい解説は、Brogliato 他(2014)に詳しい。

F. 言語習得過程に例示される高次元SDR間の結び付け問題(Binding Problem)

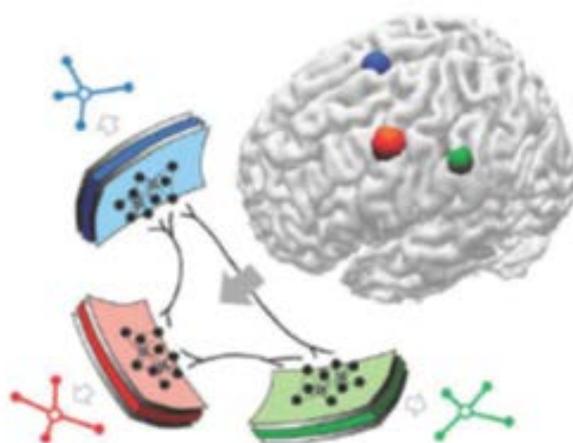
⁴⁵ 大局的なポジティブフィードバック構造をしている海馬CA3領域では、てんかん症状などの脳障害が起きやすい(Rolls (2017))。実際、てんかんの多くは、側頭葉てんかんだと言われている。

⁴⁶ Rolls (2017)によれば、海馬起因の側頭葉てんかんに比べて大脳新皮質起因のてんかんが少ないのは、このようなハードウェア構造のためだという。

甘利 (2008)、Rolls (2017)、Anderson (2017)などが繰り返して強調しているように、脳内外から刺激・イベントが到着したときには、特定の一つのニューロンだけではなく、同時に活性化した多数のニューロンに分散表現される形で学習・記憶されていく。そして、ニューロン間にまたがるネットワークの強さと広がり、表現される刺激・イベントの抽象度が高まれば高まるほど、ミニ柱状、機能柱状、領域柱状、領野柱状のモジュール内外に拡大していく。

したがって、概念の抽象度が低ければ、脳内記憶の基本となるミニ柱状モジュール内に存在しているニューロンが相対的に数多く発火してシナプス結合を形成していく。しかも、その事象の頻度が高ければ高いほど、そのようなシナプス結合が強固なものとなるので、連想の速度が増す。そして、概念の抽象度が増せば増すほど、ミニ柱状、機能柱状、領域柱状、領野柱状などのモジュールにまたがったニューロンがシナプス結合による結びつけられて行く。

Potts Attractor Network のイメージ図



(Russo, Pirmoradian and Treves (2012) , 32 頁を参照)

本図の使用許可は Treves 博士より直接いただいた。

上記の階層的に構成されるニューロンネットワーク間の結合の強さと幅が広がって行く様子は、Rolls (2017)の言語習得過程に関する仮説に沿いながら理解してみると誠に興味深い。Rolls によれば、大脳皮質内には名詞・形容詞・副詞等が一括りに集められた互いに疎な一塊の数ミリ程度の名詞であれば1万個程度の単語を記憶可能な(先の機能)柱状モジュールが互いの独立性を保つためにスパースな SDR として数多く存在しており、そのような機能柱状モジュールの一つ一つに、各々のモジュール特性を記したメタコード (place

code)⁴⁷ が付随している。人々は、共通言語や共通学習機会に接することによって、これらの機能柱状モジュール間に先の Place Codes を利用しながら類似の連結パターン(上図の Potts Attractor Networks と呼ばれる)を無意識に刻み込んでいくというのである。

なお、同様の高次元 SDR 間の結び付け問題 (Binding Problem) に関するより脳模倣的な解決法やその実践例に関しては、その難解さから未だ十分に咀嚼できていないが、Eliasmith(2013) がとても興味深い。⁴⁸

2. 3 NM-AI の実用化を難しくしている様々な要因

門外漢のために詳しく触れることはできないが、2.1 節で紹介した Burr(2015) の“How”に関連した Neuroscience 上の難しさの一端を、先に触れた海馬関連の事例で垣間見ることはできる。さらに、NM-AI 実用化を難しくしている要因には、Neuroscience 上の難しさだけではなく、半導体設計/コンピュータ・アーキテクチャ上の難しさも存在する。そのような大きな一因が、高次元 SDR 動作に必須の高次元 Content-Addressable Memory (CAM、連想メモリとも呼ぶ) 実現の難しさにあることが良く知られている。

なお、後ほど詳述するが、CAM とは、コンテンツ (Content 内容) 自体がアドレスとなっているメモリをさす。アドレスとコンテンツの関係に関する卑近な例では、郵便番号がアドレス、郵便番号の住所にある建物の詳細な特徴がコンテンツである。したがって、CAM では、建物の詳細な特徴(ベクトル)自体がアドレスかつコンテンツとして扱われ、郵便番号は使用されないということになる。

加えて、このような詳細な特徴をアドレスとして使い、特徴が似通ったもの同士であればあるほど近くに配置するような記憶方式を採用するとしよう。そうすると、ある建物の特徴が提示された時に、その建物と可能な限り類似のものを探す類似性・新奇性判断では、アドレスとコンテンツが分離されたケースに比べ、類似な成分座標を持つベクトルをその近傍で探せば良いので、より迅速な探索・判断が可能になる。しかも、このような近傍での迅速な探索・判断が並列に実行できる場合、そのための時間をさらに短縮できる。

他方、現行のノイマン/チューリング型コンピュータでは、冒頭で触れたフォンノイマン・ボトルネックが示すように、アドレスとコンテンツを明確に分離したアーキテクチャに

⁴⁷ Place Code は、次のように定義されている。“In most cortical systems, information is encoded by which neurons are firing, and how fast they are firing. This is called place coding, for each neuron has its place in the cerebral cortex.” (「多くの大脳システムにおいて、情報はどのニューロンがどれほど速く発火しているかによってコード化されている。このことを Place Coding と呼ぶ。その理由は、各々のニューロンが大脳皮質の特定場所を占めているからである」)。“the place code present in the cerebral cortex is one with a sparse distributed representation, in which a small portion of the neurons is active, each tuned to a set of stimuli in a different ways, with the profile of firing rates across a population of neurons conveying the information.” (「大脳皮質の Place Code はまばらに分散表象されており、そのような表象内では、小さな一部分のニューロンが、一連の刺激に対して異なった方法で、そして、特定情報を伝える一群のニューロンにまたがった (一定時間内の) 発火率特性を示しながら、活発に活動している。」)いずれも、中馬訳。

⁴⁸ <https://www.youtube.com/watch?v=1A02g1EgcWE> には実践例 (“Spaun”) がアニメーションで示されている。

基づいて長足の進歩を遂げてきた。そのため、アーキテクチャが大きく異なるメモリ・論理デバイス同居型ハードウェアとの親和性が高いNM-AIにとって、ノイマン/チューリング型コンピュータのそびえ立つ壁を打破することは容易なことではない。というのは、現行のノイマン/チューリング型コンピュータでは64ビットアドレスに基づく演算処理が主体であるが、メモリ・論理デバイス同居型コンピュータではコンテンツ自体がアドレスとなっているため、これまで言及してきた1万次元を超える高次元SDRそのものをアドレスとして扱える未踏の演算処理（High-Dimensional Computing）が必要となってくるからである（Kanerva（2014））。

また、64ビットアドレス空間でも 2^{64} ビット（16エクサバイト＝1600万テラバイト）という巨大空間になるのに、 $2^{10,000}$ ビット空間で高次元SDRをそのままアドレス・コンテンツとして演算処理するなどということは、一見する限りは、夢のまた夢の先でもとても実現できそうもない。この点に関して、松本他（2003）やAnderson（2017）は、下記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。」

“The most difficult technical problem is memory. Several terabytes of memory would be required, and integrating this memory into the device would be challenging.”（「（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⁴⁹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）、中馬訳、カッコ内は同名追加）

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・新奇性判断を実行する際にどのような仕組みでメモリ処理・演算処理が行なわれているのだろうか？

この節では、門外漢のために簡単な事例ベースの紹介に留まってしまいが、NM-AI実用化を図る上でのNeuroscience上の難しさと半導体設計/コンピュータ・アーキテクチャ上の難しさの双方について少しだけ触れてみたい。

⁴⁹ 1テラバイト＝約1000ギガバイト＝1兆バイト

A. Neuroscience 上の難しさ：エピソード記憶模倣の事例

我々人間を含む動物は、過去に記憶していた断片やその類似断片に遭遇すると、それらの断片を含む全体の記憶やその類似記憶を、上記の高次元 SDR の類似パターンマッチングなどを駆使して、ほぼ自動的に復元すること (Pattern Completion と呼ばれる) ができる。例えば、テレビドラマを見ていてそこに田舎の山間の風景が出てくると、それをきっかけに自分の 10 歳の頃に親しかった友人達と野山を駆け巡っていたこと、その中の一人と喧嘩をしてしまった時の様子や苦い思い等々が次から次に思い出される、という風に、次々に連想が起きてくる。⁵⁰

このような連想記憶は一瞬の出来事に基づく自己 (self) を起点とした空間と時間の拮がりを持つ記憶 (One-Shot Learning と呼ばれる) であり、一般的にエピソード記憶と呼ばれている (LeDoux (2015))。エピソード記憶は、ほとんどが前述した海馬内の CA3 領域内に形成されている自己再帰型ネットワーク (Self-Associative Attractor Networks) の中に保存されている。そして、場合によっては、前述した側頭葉などに長期保存されている非自己的で一般的な“意味記憶”をも呼び出してより精緻な Pattern Completion が実現される (Rolls (2017))。

この海馬内 CA3 領域の特徴は、錐体ニューロン軸索が終点で分岐して反転してきた反回側枝によって自己再帰的に生成された前述の柱状モジュールから構成されている点である。しかも、CA3 領域は、大局的に配線されたポジティブフィードバック構造をしている (Rolls (2017))。繰り返しになるが、大局的という意味は、CA3 内のどこから出発しても、CA3 領域内の全ての記憶にアクセスできるという意味である。また、ポジティブフィードバック構造とは、特定の記憶断片を基点に記憶のアクセスがより広域にどんどん広がっていくという意味である。

先の BD-AI を取り上げている Hassabis 他 (2017) や Neuroscientist として名高い利根川 (2017)、Rolls (2017)、Anderson (2017) 等に述べられているように、現在の Neuroscience の到達レベルでは、海馬内の CA3 を含む様々な部位間の役割分担や大脳新皮質と海馬あるいは関連した大脳新皮質と視床等々との相互情報交換の仕組みが十分に解明されていない。そのためだろうか、それらの相互依存性を NM-AI としてソフトウェア的にさえ実装するまでに至っていない。

例えば、最近、利根川グループ (Kohara 他 (2014)) は、海馬内 CA2 領域の新機能、歯状回から CA2 ならびに CA2 から CA1 への新たな連結路などを発見した。それまでは、CA2 領域は、CA3 領域から CA1 領域への単なる通り道であるとされ、通常の教科書的な説明には明

⁵⁰ より詳細な連想メカニズムに関しては、甘利 (2008) や甘利 (2016) が直感的で分かりやすい。特に、連想記憶が、特定事象に関与したニューロン集団の中でどのように記録され、それらの記憶からどのようなメカニズムで復元されるかに関する説明は目から鱗である。

示さえされていない。⁵¹

また、BD-AI 及び NM-AI に関する包括的なサーベイを行っている Hassabis 他 (2017) や Schuman 他 (2017) は、エピソード記憶/ One-Short Learning/On-line Learning を実用化することの難しさについて、各々下記のように述べている。

“In the future, it will be interesting to harness the benefits of rapid episodic-like memory and more traditional incremental learning in architectures that incorporate both of these components within an interacting framework that mirrors the complementary learning systems in mammalian brain.” (「将来、高速なエピソード様記憶と従来型の徐々に増えていく学習 (記憶) との二つの便益を、哺乳類の脳における相互補完的な学習システムを映し出す相互作用構造の中に、これらの 2 つの (記憶) 構成要素を包摂したアーキテクチャで結びつけることは興味深いものとなるだろう。」(Hassabis 他 (2017) から抜粋、中馬訳))

“One of the key reasons neuromorphic systems are seen as a popular post-Moore’s law era complementary architecture is their potential for on-line learning; however, even the most well-funded neuromorphic systems struggle to develop algorithms for programming their hardware, either in an off-line or on-line way.” (「ニューロモルフィック型システムがポピュラーなムーアの法則後の補完的アーキテクチャとして見なされている一つの重要な理由は、オンライン・ラーニング (即座の学習) に対する潜在性である。ところが、最も予算豊富なニューロモルフィック型システムでさえも、そのハードウェアをオフライン型あるいはオンライン型でプログラムするためのアルゴリズム開発に苦慮している。」(Schuman 他 (2017) から抜粋、中馬訳))

このように、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の解明が十分に進んでいない。NM-AI のみならず BD-AI にとっても、Neuroscience との共進化が強く望まれている理由の一端である。

B. 半導体設計/コンピュータ・アーキテクチャ上の難しさ：CAM の事例

本節では、まず CAM の仕組みを提示しながら、それが高次元 SDR と極めて親和的であるにもかかわらず、実用化上大きな難問を抱えていることを直感的に提示してみたい。しかも、高次元 SDR の扱いが必須な NM-AI にとって、実用化の際にこの難問が大きく立ちはだ

⁵¹ 要旨は、http://www.riken.jp/pr/press/2014/20140219_1/参照。「CA2 は、CA1 や CA3 と性質が全く異なり、記憶している過去の環境と現在置かれている環境のわずかな違いにより、活動パターンを変化させることが明らかになりました。」等々。なお、最新の教科書 VAN Ooyen 他編 (2017) には、CA2 の機能が取り上げられていた。

先に説明したように、これらの高次元 SDR のそれぞれの要素は、特定ニューロン集団の 1 万個中の特定ニューロンに対応しているとする。したがって、両者の数値の 1 で示されている共通箇所は、刺激・イベント A でも刺激・イベント B でも同じニューロンが活性/シナプス結合されたことを意味する。そして、脳内では、高次元 SDR 間での 1 の重なりが多ければ多いほど、両事象の類似性が高いと判断されるようになっている。つまり、到着した刺激・イベントに対応する content と address が一体化した高次元 SDR がメモリ・論理演算子一体型の計算機としてのニューロンにそのまま提示され、それが過去に学習・記憶された高次元 SDR と内積値を求める形で比較対照される形で類似性/新奇性判断が並列に実行される。

さらに、このようなニューロン集団での同時並行処理が時間の流れの中で次から次に継起していくので、到着した高次元 SDR A がニューロン集団の活性/結合パターン (= content) を生み出し、その content を基点として過去の類似パターンマッチング (計算) 機構が駆動されベストマッチしたパターンが選択/想起される。さらに、その選択/想起されたパターンを基点としたベストマッチパターンの選択/想起、・・・といったプロセスが芋づる式に続いて行く。このような Content が Content を次々に直接起動していくという振り舞い(連想)が、20 ワット前後で作動するヒトの主要な脳活動を支えている。

では、このような処理を通常のコンピュータで行うとするとどうなるだろうか？ まずは、到着した刺激・イベントを記憶するアドレスを成分毎に指定、次にそのアドレスにポインタを移動してコンテンツを各々の成分ごとに数値として書き込み、書き込まれたそのコンテンツを過去に既に学習・記憶されているコンテンツと比較対照するために過去のコンテンツに割り当てられている該当アドレスをコンテンツ毎及びその成分毎に指定してメモリから成分毎に読み込みこんでベクトル化、それらを CPU に提示してベクトル成分毎に比較演算、そのような逐次比較演算結果をメモリにアドレス指定して書き込み、逐次演算終了後にもっとも類似度の高いものを選択、・・・といった長い長い逐次演算プロセスが必要になる。しかも、メモリ構造的に類似性の高いもの同士が近傍に集められてはいないから、アドレスが似ているからと言って類似なパターンを直ぐ近くで探せるわけではない。

他方、高次元 SDR をコンテンツ&アドレスとして使えるような実ニューロンのようなメモリ・論理同居型のコンピュータが利用できれば、上記のような極度に面倒な処理の多くが一発で、しかも並列に行えるので、少なくともベクトル内積値を用いた類似性・新奇性判断に関しては、例え各々の処理速度がかなり劣っているとしても、高速な Neuman-Turing 型コンピュータに対しても勝るとも劣らない性能を格段の低消費電力で発揮できるようになる。

B.3 高次元 CAM の光と影：高耐障害性・高速、高消費電力・高コスト

さらに、高次元 CAM は、content と address が完全に分離されている現行の Neuman-Turing 型コンピュータでは達成できないいくつかの好ましい特徴を持っている。その一つが、高次元 SDR の必要性に関して言及した耐障害性である。この点を具体例で見てみよう。

例えば、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てしまうと、計算全体にまったく意味がなくなってしまう。これに対して、先のCAM型計算機であるニューロンの場合、SDR内の一箇所だけ誤りが起きる程度のものであれば、内積値が1だけ減るだけなので、高次元であればあるほど類似性・新奇性判断には影響しなくなる。そのため、NM-AIの場合、CAM型であることにより、障害に対してとても頑健性が高い。

このように高次元CAMが利用可能であれば、高性能を誇るNM-AIの実現が直ぐにでも実現できそうである。しかも、高次元SDRを直接扱うCAM型の一対対照比較は、一見簡単な操作のように思える。

ところが、実際には、現在のコンピュータの限界を遥かに超える操作が行われている。現在では、スーパーコンピュータのみならずスマートフォンやPCでも最新のものは(iPhoneならiOS、WindowsマシンならWindows 10といった)64ビットOSで動いている。したがって、上記のようなベクトル同士を直接比較する操作は、せいぜい64次元で間に合うものに対してしかできない。もちろん、addressとcontentが分離されていても、ハッシュ関数と呼ばれるようなものや仮想化などのソフトウェア的な工夫をすれば、64次元を越えた疑似直接比較は可能である。ただし、そのためには、大がかりな回路上の工夫も必要になるし、そのための消費電力もかなり大きくなったり、実行速度が極端に遅くなったりする。

もちろん、低次元ベクトルを活用したCAMは、現行のNeuman-Turing型コンピュータでも高速化を図るためにアチコチで使われている。例えば、Intel製CPUのキャッシュレジスタ（メモリ）と呼ばれるCPUが直接データのやり取りを行う部分とかCisco製の高速ルーターでネットワーク機器に割り振られているMACアドレス/IPアドレスのスイッチング（切り替え）を高速に行う際に活用されている。

ただし、これらのCAMは、低次元とは言え、現行のアーキテクチャでは、PCやスマートフォンなどのシステムメモリとして使用されているDRAM(Dynamic Random Access Memory)などとは異なり、記憶されているコンテンツ各々の箇所で並列同時比較するための比較回路(comparatorと呼ばれる)が組み込まれている等々による回路上の複雑性が大きい。そのため、既存の半導体製造・設計技術を前提とする限り、とても高価だし、比較のたびにすべての記憶コンテンツとの並列同時比較が行われるので消費電力もとても大きい(Sharma(2003))。もちろん、フラッシュメモリやReRAM、MRAMなどの不揮発性メモリを使って消費電力を抑える試みも出てきているが、回路上のみならず製造プロセス上の難度も未だかなり高い(Dieny(2017)他)。

したがって、Neuman-Turing型コンピュータならびに既存の半導体製造・設計技術を前提とする限り、先の1万次元のSDRを扱える高次元CAMの実現など夢のまた夢なのである。

つまり、そのような高次元の CAM を実現するには、半導体設計・製造プロセス技術上の Revolution が必須となっている (Kanerva (2014))。

B.4 待望されている Kanerva(1988)流高次元 CAM の実現

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 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。脳内で動作する CAM アーキテクチャが解明されて安価で低消費電力の高次元 CAM が幅広く使えるようになり、それを使いこなすための設計ツールが利用可能になれば、現在の微細化技術を利用してより高度な NM-AI の実現が、これまでに紹介してきた低消費電力性を含む様々な便益を生むので、一気に進む可能性があるのに……である (Kanerva (2014))。

ただし、このような手探り状態の現状の中、異彩を放っているのが Kanerva(1988) のアイデアである。したがって、以下では、Brogliato 他(2014)や Franklin(2001)に習いながら、Kanerva の提唱する高次元 SDR に親和的な CAM である Sparse Distributed Memory (SDM) ⁵²について、ほんの少しだけ触れてみたい。

前述したように、確率変数としての分散表現 (DR) には、スパース化・高次元化しがちとなる進化論的理由が数多く存在する。そして、Kanerva (1988) では、まず、高次元 SDR 化しがちな理由やそれらの間に直交性が生まれがちとなる理由が、二項分布の特性に言及しながら数学的に説明されている。なお、SDR 間の直交性は、類似性・新奇性比較の際の情報圧縮に不可欠だった (甘利 (2008)、Rolls (2017))。

具体的には、下記のようなロジックである。Kanerva (1988) では、簡単化のために、分散表現ベクトル (DR) が、一般性を失うことなく、その成分としての各々のニューロンが p の確率で 1、 $(1-p)$ の確率で 0 となる独立な確率変数だとされる。そして、このようなニューロンの n 個からなる集団うちの d 個が 1 となる分散表現ベクトル (DR(d) と呼ぶ) の特徴を検討してみよう。良く知られた数学上の性質から、この DR(d) は、平均値が np 、分散が $np(1-p)$ の二項分布に従う。そして、良く知られているように、 n が大きくなればなるほど、したがって高次元 DR であればあるほど、この二項分布に従う変数 x を $z = (x - \text{平均値}) / \text{標準偏差}$ と変換すると、標準正規分布 (平均値 0、標準偏差 1) に近づいていく。

したがって、このような高次元 DR は、 n が大きくなればなるほど、平均値 np の周りに密集していく。したがって、 $p=0.5$ とすると、ほぼ全ての高次元 DR がスパースな DR、つまり SDR となる。さらに、前述したように、Rolls(2017)によれば、大脳新皮質の第 2・3 層では $p=0.1$ 、海馬の CA3 領域では $p=0.04$ であるから、 $n=10,000$ としたときでも、前者では 1000 個、後者では 400 個しか 1 の値を含まないような高次元 SDR になっている。しかも、前者の平均値と標準偏差 (σ) が 1000 と $\sqrt{np(1-p)}=30$ 、後者の平均値と標準偏差が 400 と 20 となる。とすると、正規分布では平均値の両側 3σ 内に 99.7% の値が入ってしまうので、ほと

⁵² SDM に関しては、wiki が異様に詳しい。Kanerva 氏ご本人が書いておられるかのようでもある。
https://en.m.wikipedia.org/wiki/Sparse_distributed_memory

んどの高次元 SDR (ここでは $n=10,000$) が前者では 1000 ± 90 個、後者では 400 ± 30 個の 1 をしか含まないことになる。そして、任意に選んだ二つの高次元 SDR は、両者が同一である確率は限りなくゼロに近いし (Numenta (2017))、互いにほぼ独立なのである。

ただし、Kanerva (1988) がイントロダクションでも言及しているように、各々の SDR がほぼ独立であるにもかかわらず、任意の二つの高次元 SDR の間に数個の SDR を連想によって介在させると、独立な高次元 SDR 間が結びつくようになる。この点に関して、Kanerva は、下記のような尻取りゲームのような例で最初の質問の答えに 14 ステップ後に辿り着く例を示している。当初の質問が、文字面、同音異義語、同音、同色等々に起因する白昼夢のような論理薄弱な連想によって、最後のもっともらしい答えに連なっていく例である。

Why are fire engines painted red? 消防車はなぜ赤色なのか？

Firemen's suspenders are red, too. 消防士のサスペンダーも赤だ。

Two and two are four. 2×2 は 4 だ。

Four times three is twelve. 4×3 は 12 だ。

Twelve inches in a foot. 12 インチは、1 フットだ。

A foot is a ruler. 1 フットは、(12 インチの) 物差しだ。

Queen Mary was a ruler. マリー女王は支配者だった。

Queen Mary sailed the sea. クイーンマリー号は、海洋を航海した。

The sea has sharks. 海にはサメがいる。

Sharks have fins. サメにはヒレ (フィン) がある。

The Russians conquered the Finns. ロシア軍がフィンランドを征服した。

The Russians' color is red. ロシアの色は赤だ。

Fire engines are always rushin'! (赤い) 消防車はいつも急いでいる。

So that's why they're painted red! だから消防車は赤色なんだ！

上記の事例は、言語を駆使する時として合理的だがほとんどの場合に非合理的(直情的)であるヒトの創造性の源を垣間見せてくれている。ところが、こういう風な豊かな連想を実現するには、高次元 SDR が必須、つまり、 $2^{1,000}$ ビットや $2^{10,000}$ ビットのアドレス空間で高次元 SDR をそのままアドレス&コンテンツとして演算処理する必要がある (Rolls (2017), Anderson (2017)、Numenta (2017)、Kanerva (2014) など)。現行のノイマン/チューリング型コンピュータでは 64 ビットアドレスに基づく演算処理が主体であるが、64 ビットアドレス空間でも 2^{64} ビット (16 エクサバイト=1600 万テラバイト) という巨大空間になってしまうのである。したがって、現状の半導体製造・設計技術を前提とする限りは、ハードウェア上のみならずソフトウェア上でもとても実現できそうもない。

ここで登場するのが Kanerva (1988) の Sparse Distributed Memory (SDM) というコロンプスの卵的なアイデアである。ちなみに、筆者の知る限り、Numenta (2017) や

Brogliato(2014)では1000次元や2000次元のCAMがソフトウェア的に工夫した形で試用されているようである。

もちろん、 $2^{1,000}$ 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難しいことには変わりはない。さらに、100年間を秒に直すと 3.154×10^9 秒 $\approx 2^{31.55}$ 秒であるから、人生の出来事を全て記憶する際にも、 2^{100} ビットでも多過ぎるほどである。そして、Kanervaが着目するのは、この点である。つまり、 $2^{1,000}$ ビットや $2^{10,000}$ ビットの巨大アドレス空間を想定しても、実際に記憶場所(Storage)として使われるのは、100年間分でも、そのようなアドレス空間内の無視できるほどわずかな領域だという点である。

しかも、独立な確率変数としての高次元SDRでは、たとえその次元が1万($n=10,000$)でも、先のロジックで、ほぼ全てのSDRが平均値10000pの 3σ 内に入ってしまう。とすると、1万ビットで表現される刺激・イベント情報を100万個、あるいは100億個記憶しておく場所が必要であるとしても、1万 \times 1000万ビット=125億5000万バイト=12.5ギガバイト($\approx 2^{33.5}$)、1万 \times 100億ビット=1.25テラバイト($\approx 2^{40.2}$)程度となる。筆者(中馬)のWindows 10-PCでは2テラバイトのSSDが使われているので、このSSDだと100億個用でも多過ぎるほどだということになる。

となると、Kanerva(1988)が提唱しているように、高次元SDRでの類似性・新奇性判断のための内積演算を、 $2^{10,000}$ 個の巨大なアドレス空間を仮想空間(Virtual Space)と見なして、そこでの演算を 2^{33} や 2^{40} の実空間での演算に置き換える仕組みを活用すれば良いようだと言外漢にも類推できる。そして、Kanerva(1988)が、実際にそのようなSDMが導入可能なことを数学的に証明すると共に、現行の半導体製造技術を使ってどのような形で実用化可能となるかも提示している。その詳しい方法は本論では十分に説明できないが、Brogliato他(2014)などを参照すると、現状でも $2^{1,000}$ ビットの高次元SDRであれば、この仮想空間と実空間との間の変換が実用的な段階に達していることを確認できる。

ちなみに、KanervaのSDMの場合、到着した刺激・イベント情報は、実空間の特定箇所だけに書き込まれるのではなく、そのような箇所の近傍にも同じく一様に書き込まれる。そして、そのことによって、メモリの一部に障害が発生しても、読み込みの際には近傍のメモリを含む平均値が使用されるので、耐障害性を発揮できる。さらに、このような配置方法から、メモリストレージの中では、類似(非類似)した記憶同士は近く(遠く)に配置されることになる。そのため、到着した刺激・イベントに類似したパターン発見速度やそのパターンによって呼び起こされる連想の幅の広がりやその速度が増す。

したがって、特定分野の知識・ノウハウを芋づる方式で密に記憶しているプロであればあるほど、一を知って十を知るというような素人をあつと言わせる連想力を誇る。他方、類似パターンだとして選ばれた記憶が本来の記憶パターンと異なる箇所に誘導されてしまうと、思い出せそうで結局思い出せないといった現象(Tip of Tongue現象)も起こる。さらに興味深いのは、そのような刺激・イベントが書き込まれる箇所は、特定の刺激・イベント

情報だけに使われるのではなく、複数の刺激・イベント情報用にも使われる。そのため、文字通りスパース分散表現 (SDR) そのものとして記憶される。

したがって、Kanerva 流 SDM が高次元で、しかも、ハードウェア的に実現できるようになればなるほど、NM-AI の実用化が一気に加速することは間違いない。しかも、この高次元 SDR の便益は、NM-AI のみならず、時空間に広がるエピソード記憶やオンラインでの学習が不得意な BD-AI にも大きな便益をもたらす。というのは、この Kanerva の SDM のアイデアは、Graves 他(2016)などの事例や岡野原(2017)に言及されているように、まだ低次元での応用であるが、最先端 BD-AI でも必須となってきたからである。

なお、SDM は、現行のアドレスとコンテンツが一对一で決定論的に決まるロジックではなくビット欠けなどを許す確率的に決まる確率論理を使うものである。そのため、おそらくであるが、大規模回路などのハードウェアとしての実現方法のみならず、それらを使いこなすためのソフトウェア上の工夫にも革命的なアイデアが要請されているようである (Kanerva (2014))。

C 高次元 CAM が作動する NM-AI の難しさ

本節で検討してきた脳内記憶に関する仕組みから類推されることは、コネクショニズム⁵³ が主張するように、人間の情報処理は、ノイマン/チューリング型コンピュータとは異なり、「心の中でシンボル (言語) 操作をする」という形では行われていないのではないかということである。これまで見てきたように、動的に生み出される連想間の時空間的な拡がりや深さこそがインテリジェンスの中核だと類推できるからである。松本他(2003)や甘利(2016)は、この点を下記のように巧妙に表現している。

「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 (表引きテーブル) 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

⁵³ 「多数の基本要素 (ニューロンであろうとなかろうと) の結合による相互作用で情報処理が (並列に) 進んでいくという側面を重視しこれによってどのような情報原理が実現できるのかを問う立場」 (甘利 (2008)、5 頁)。実際、我々が保有する様々な知識はニューロン同士の結合の強さなどによって表現され、脳内の様々な場所に分散して保存されている。コネクショニズムでは、そのことを“分散表現/表象 (distributed representation)”と呼ぶ。そして、Numenta (2017)は、この分散表象こそ、コンピュータには発揮できない柔軟性・頑健性を脳 (特に大脳新皮質) が保有する本質的な要因だとしている。

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)

「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、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、並列のダイナミクスで働く分散した記憶が実現できるというわけだ。」（甘利(2016)、121 頁)

そして、現行の Neuman-Turing 型コンピュータは、「思い出すための仕掛け」が脳とは大きく異なっているために、そのアーキテクチャに基づくだけでは分散並列処理と確率論理で特徴づけられる NM-AI 用の CAM(「ルックアップ・テーブルに貯えられたメモリ」)を容易には実現できない要因だと言えそうである。この点に関して、Kanerva (1988) と共に高次元 SDR の有用性を実現するためのアイデアを提示している下記の Plate (2003) の指摘は、まだ実現にはほど遠いようであるが、極めて興味深い。そのような CAM 導入により「思い出すための仕掛け」をより実際の脳に近づけ、NM-AI の飛躍方向を垣間見ることができるからである。

“Much of the interest in connectionist models for higher-level processing stemmed from dissatisfaction with the limitations of symbolic rule-based systems. These include brittleness, inflexibility, difficulty of learning from experience, poor generalization, domain specificity, and sloth of serial search in large systems. Initial results with connectionist models seemed to suggest potential for overcoming these limitations. Various connectionist models have tantalized researchers with many attractive properties: pattern completion, approximate matching, good generalization, graceful degradation, robustness, fast processing, avoidance of sequential search, parallel satisfaction of soft constraints, context sensitivity, learning from experience, and excellent scaling to larger systems.”

(「高次レベル処理に対するコネクショニストモデルへの興味の多くは、シンボル操作に基づくシステムの限界に満足できないことから生まれてきた。それらには、脆さ、柔軟性の欠如、経験学習の難しさ、貧弱な一般性、領域特殊性、大きなシステムでの逐次検索の遅さなどが含まれる。コネクショニストモデルによる初期の成果は、これらの限界を克服できる潜在性を示しているよ

うに見えた。様々なコネクショニストモデルは、研究者達に、多くの魅力的な特性を見せびらかしてきている。(記憶断片からの) 完全記憶復元、近似マッチング、良好な一般化、障害耐用性、頑健性、迅速処理、逐次検索回避、並列でのソフト制約充足、コンテキスト感応性、経験学習、大規模システムへの優れたスケール性などである。)]

“However, no single connectionist model has had all these properties and researchers have found it very difficult to perform higher-level reasoning tasks in connectionist models. The difficulties can in part be traced to two characteristics of higher-level reasoning tasks. The first is that the temporary data structures required for higher-level reasoning are often complex, and cannot be represented in common connectionist representation schemes. The second is that many higher-level reasoning problems appear to at least sometimes require sequential processing, and connectionist models lack the procedural controls necessary to control sequential processing. In this book I focus on how some of the structures that are required for higher-level reasoning can be represented in a distributed fashion.” (「ところが、コネクショニストモデルの一つとしてこれらのすべての特性を持っていないし、研究者達はコネクショニストモデルでは高次の推論を行うことはとても難しいということを見つけた。それらの難点は、一部は、高次推論作業の二つの特性に帰着させることができる。その一つは、高次推論に必要とされる一時的に利用されるデータ構造がしばしば複雑であるため、共通のコネクショニスト的な表現形式で表現することができない。もう一つは、多くの高次の推論問題には、少なくとも時として逐次処理が必要になるが、コネクショニストモデルには逐次処理制御に必要な手続き的制御が欠けている。この本では、(したがって) 高次の推論問題に必要とされるある種の構造が、どのようにして分散的な方法で表現可能であるかに焦点を合わせる。」(いずれも中馬訳)

3. 人的資本との関係でNM-AIのインテリジェンスをどう捉えるか？

この節では、NM-AIが今後社会に与えると考えられる大きなインパクトについて、特に人的資本/HIが持つべき重要な特徴である**変化と異常への対応力**ならびにそのような対応力自体の**自己変化能**へのインパクトについて、そもそもNM-AIの**インテリジェンス**とは何かといった問いに留意しながら考察してみたい。繰り返して申し訳ないが、注目するのは、人間のインテリジェンス(HI)ではなく、完成形としてのNM-AIのインテリジェンスである。そのことによって人的資本/HIとNM-AIとの補完性や代替性をクローズアップするためである。

AI 関連の文献には、インテリジェンスの定義をチューリング・テストのような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）人々は、AI（コンピュータ）に人間並あるいは人間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“チューリング・テスト”を引き合いに出すことが常套手段化している。その傾向は、（人類にとっては？）誠に皮肉なことに、著名な哲学者 John R. Searle の強烈な反チューリング・テスト論である“中国語の部屋（Chinese Room）”ストーリーでさらに注目を浴びることになってしまった。⁵⁴

<チューリング・テスト> “If a computer can fool a human interrogator into thinking that it too is a person, then by definition the computer must be intelligent.”（「もしも、コンピュー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、中馬訳）

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大きな反響に当惑してしまったように、人々に「人間だと思込ませる」手段は数多く存在すると分かっているのに、である。⁵⁵

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、NM-AI と人的資本との補完・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（1994）の的確な指摘を以下に引用してみよう。

“AI is misnamed, for its purview includes human and animal intelligence. I see it as the general study of sophisticated self-modifying information-driven control systems, both natural (biological) and artificial, both

⁵⁴ 本論では“中国語の部屋”の話は本質的ではないので割愛する。詳しくは、Hawkins（2006）や公開されている Searl の論文：<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.120.749&rep=rep1&type=pdf> を参照されたい。

⁵⁵ 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、本書に展開される彼の AI 賛美論に対するアンチテーゼは傾聴に値する。

actual and possible (including what might have evolved or might be made).”

（「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**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、実際のものど可能性のあるもの（進化してきたに違いないものと造られたもの）とを含む。」（中馬訳）

“Since all we have to go on in judging mental states of others is behaviour, some people think that concepts of mental states and processes are definable simply in terms of behavioural capabilities and dispositions of the whole system, independently of how it works internally, e.g. what its architecture is. Similarly, the ‘intentional stance’ ignores internal processing. If intelligence depends on how behaviour is produced this must be wrong. (Turing himself was too intelligent to propose passing his test as a criterion for intelligence or understanding. He merely offered it as a technological challenge he thought could be met without using an HLT (Huge Lookup Table).”（「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総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ー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。))

(中馬訳、()内一部同名追加)

いずれにせよ、NM-AI 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、極めて危険であることが理解できる。さらに、上記で Sloman が述べているように、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等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し、異なった生活圏に適応してヒトとは異なった驚くべき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。⁵⁶ 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

⁵⁶ 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

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ことではないだろうか。⁵⁷

したがって、NM-AI の人的資本/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 が外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 (Understanding cannot be measured by external behavior)」(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 (Design-based concepts are important for full understanding of how our minds work and how they may fail to work, or work inappropriately.)」(Sloman (1994)、()内は中馬追加)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？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 創設者の Hawkins (2006) や 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 Hawkins の元部下 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。その要点は、NM-AI の主目的は大脳新皮質⁵⁸の作動原理に習うことであるから、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。⁵⁹

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 (<http://www.natureasia.com/ja-jp/reviews/highlight/10618>)。人間中心主義の浅薄さを実感させられる事実である。

⁵⁷ 過去の記憶と現況に基づく将来の予測力の違いは、哺乳類では配線構造、従ってシステムアーキテクチャの違いに帰着する部分も多いようだ。例えば、Rolls (2017) には、霊長類とラットの味覚処理に関する脳内配線上の大きな違いが紹介されており、なぜ前者が後者に比べて味の奥深さを堪能できる“食通”であるかが理解できる。

⁵⁸ 教科書によって定義が異なるので我々門外漢は大変に混乱するが、大脳は、大きくは進化的に新しい大脳新皮質 (neocortex/isocortex) と進化的に古い大脳辺縁系/不等皮質 (Limbic System/Allocortex)、さらに大脳基底核 (cerebral basal ganglia) に分けられる。辺縁系/不等皮質には、広義・狭義とあるようで、前者の場合、脳梁辺縁を前後に回り込んでいる帯状回、海馬体 (海馬、CA1-CA4、歯状回、海馬台)、扁桃核、嗅脳、脳弓その他が含まれる。辺縁系/不等皮質をさらに大脳原始/原皮質 (Archaeocortex/Archicortex) ならびに大脳旧/古皮質 (Palaeocortex/Paleocortex) に分けているものも少なくない。大脳基底核には、線条体 (被殻、尾状核)、淡蒼球、視床下核、黒質などが含まれる。このような意味では、松本他 (2003) が重ねて強調しているように、Hawkins などが主張する大脳新皮質に宿るとされるインテリジェンス以外に、情動や直観などを含むより広義のインテリジェンスも存在する。

⁵⁹ 因みに、鳥類や爬虫類、硬骨魚類などには大脳皮質がないとされてきたが、最近では無層型のものが存

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。⁶⁰

“You can think about the world, move around in the world, and make predictions of the future because your cortex has built a model of the world. One of the most important concepts in this book is that the cortex’ s hierarchical structure stores a model of the hierarchical structure of the real world. The real world’ s nested structure is mirrored by the nested structure of your cortex.” (「あなたは世界について考え、その中を歩き回り、将来予測を行える。それは、あなたの大脳(新)皮質が、世界に関するモデルを構築しているからだ。この本の中で最も重要な概念の一つは、そのような大脳(新)皮質の中の階層構造が現実世界の階層構造のモデルを格納しているということである。現実世界の入れ子構造が、あなたの大脳(新)皮質の入れ子構造に映し出されている。」(Hawkins (2004)、中馬訳))

“The design of the cortex and the method by which it learns naturally discover the hierarchical relationships in the world. You are not born with knowledge of language, houses, or music. The cortex has a clever learning algorithm that naturally finds whatever hierarchical structure exists and captures it. When structure is absent, we are thrown into confusion, even chaos.” (「大脳(新)皮質には世界に関するモデルが備わっているので、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で、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)、中馬訳))

“But with its hierarchical structure, invariant representations, and

在することがわかってきている。つまり、これらの動物の大脳は、層構造ではなく、哺乳類の各層に類似した独立の一塊として存在している。無層型と呼ばれる理由である(以上は、渡辺(2010))。

⁶⁰ “Being human and being intelligent are separate matters. An intelligent machine need not have sexual urges, hunger, a pulse, muscles, emotions, or a humanlike body. A human is much more than an intelligent machine.” (「人間であることとインテリジェントであることは別問題だ。インテリジェントな機械には、性欲、飢え、脈拍、筋肉、情動あるいは人間らしい肉体は必要ない。人間は、インテリジェントな機械以上のものなのだ。」(Hawkins (2006)、中馬訳))

prediction by analogy, the cortex allows mammals to exploit much more of the structure of the world than an animal without a neocortex can.”（「ところが、哺乳類は、大脳（新）皮質内の階層構造と恒常化された表現⁶¹ならびに類推による予測を使って、大脳（新）皮質を持たない動物よりも世界の構造をさらにいっそう活用できる。」（Hawkins（2004）、中馬訳）

“The No Free Lunch (NFL) theorems⁶² for learning teach us that no learning algorithm has an inherent superiority over another learning algorithm for all learning problems. If an algorithm is superior for a particular problem, it is only because the algorithm exploits assumptions that are suitable for that problem. On the other hand, results from neuroscience tell us that the neocortex might be using the same algorithm for different tasks like visual, auditory and somatosensory perception. It is also well known that the neocortex is organized as a hierarchy. A plausible explanation for these observations is that data from different domains and sensory modalities, despite their apparent superficial differences, can have the same underlying statistical properties. This could be because the data generation mechanisms of the world have an underlying hierarchical structure due to the laws of physics and self-organization. The neocortex, through evolution, might have discovered this fact and could be exploiting it to learn efficient models for hierarchically structured data. If that is the case, then the secrets behind the efficiency of hierarchical learning can be unlocked⁶² by studying the organization of the world.”（「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

⁶¹ 拡大・縮小・変形等の操作によっても認知対象の同一性が揺るがなくなっている脳内の表現／表象。

⁶² <http://www.no-free-lunch.org/>

である。もしそれが本当だとすると、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)、中馬訳)

したがって、Hawkins (2006) や George (2008) によれば、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外界が各種の物理・化学法則に支配された (モジュール) 階層構造をしていることの反映である。⁶³ 第二に、そのような外界の中で繰り返し発生する物事や事象に対してより効果的に適応しながら生き抜くためには、外界の階層構造の (間主観⁶⁴的な) 写し絵としての脳内階層モデルを自他の経験に基づきながら作り上げる必要があるからとなる。そして、第三に、それらの脳内階層モデルを基準として、そこからの対比で変化と異常にいち早く気づいて対応したり、さらに、そのような気づきに基づいてモデルのさらなる精緻化を継続的に成し遂げていくことによって適応能力 (自己変化) の向上を達成できるからである。

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、下記のように定義できる。

“Intelligence is measured by the predictive ability of a hierarchical memory, not by humanlike behavior” (「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) 中馬訳)、

“Intelligence is the capacity of the brain to predict the future by analogy to the past.” (「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)、中馬訳)

“The point is not to label some species as intelligent and others as not intelligent. Memory and prediction are used by all living things. There is just a continuum of methods and sophistication in how they do it.” (「重要な点は、ある生物種に知性があるが他の生物種には知性が無いとレッテルを貼ることではない。記憶と予測とは、全ての生命体によって活用されている。彼らの実践方法やそれらの巧みさには連続した流れが存在しているだけなのである。」(Hawkins (2006)、中馬訳)

⁶³ Rolls (2017)は、加えて、神経回路ネットワークに弱エルゴード性 (weak ergodicity) が成立していることが必要だと強調している。弱エルゴード性とは、平たく言えば、あるシステムの振る舞いに関して、極めて長期間観察することによって得られるその平均特性 (時間平均特性) と極めて短い期間に空間内で幅広く観察することによって得られる平均特性 (空間平均特性) とが同じになるという統計学的な性質である。

⁶⁴ 間主観性に関しては <https://plaza.umin.ac.jp/kodama/ethics/wordbook/intersubjectivity.html> 参照。

つまり、NM-AI に実装されるインテリジェンスとは、過去の記憶と現況に基づきながら環境適応していくための各自の世界/社会/環境モデルに基づいた予測能力ならびにその学習能力に基づく自己変可能なのである。自己変可能が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としてのアルゴリズムが、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（“効率、影響度”⁶⁵）と拡がり自動的に変化させる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るからである。したがって、自己変可能とは、冒頭で引用した松本他(2003)が教えてくれる「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」というメタアルゴリズムそのものなのである。

なお、彼らが脳新皮質という時には、海馬と視床とを含めている。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、脳新皮質の上位に位置する(Hawkins(2006))。さらに、6層構造を成す脳新皮質では、脳新皮質内の異なった部位間の情報転送・応答は常に視床を経由して行われる。そのため、視床無しでは、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脳新皮質の主に第4層に伝えられるし、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、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。

ところが、前節でも紹介したように、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の解明が十分に進んでいない。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・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エピソード記憶の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。

ここまで来ると、完成形としてのNM-AIのインテリジェンス特性がだいぶクリアーになってくる。ただし、NM-AIが実装するインテリジェンスの提示だけでは、そのようなインテリジェンスと人的資本/HIとの補完性・代替性に関する考察を進めることはできない。そして、このような考察の方向性を示してくれているのが、ヒトの意識上・意識下の情動/価値判断をも含む松本他(2003)とMinsky(2007)の脳機能に関する研究成果/仮説である。

詳細は以下で紹介するが、松本他(2003)は、上記のNM-AIで実装される(脳新皮質型)インテリジェンスを“知・情・意”の中の知とし、ヒトを含む哺乳類の諸活動を支える脳活動の特性を理解するには、その役割の大きさから“情・意・知”と見なすことの重要性を強調する。他方、Minsky(2007)も、詳細は後述するが、ヒトを含む生命体の低位から高

⁶⁵ 甘利(1989)の表現。

位に至るインテリジェンス活動を、本能、学習、熟慮、内省、自己内省、(意識上の)情動の六つであると見なし、大脳新皮質型インテリジェンスを越える NM-AI の実現には、少なくともこれらの6階層構造を持つ実体、つまり彼が呼ぶ Emotion Machine (EM) が不可欠であるとする。その背景を成すのは、意識上・意識下の情動とは、変化と異常への対応に必須な柔軟で高速の視点切り替え装置であり、互いの多様な意図や常識・文化等々の共有による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との主張である。⁶⁶

以下では、両者の考え方を座標軸に据えて、具体的な実装事例に沿った形で BD-AI や大脳新皮質型 NM-AI が現在の延長線上で到達可能なインテリジェンス特性を相対化し、変化と異常への対応及びその自己変化能という視点から、人的資本/HI との補完性・代替性を考察してみたい。

松本他(2003)が“知・情・意”を“情・意・知”と並び替えることに関しては、十分な脳神経科学的な根拠があるという。後述する Minsky の6階層モデルとの親和性が高いので、そのロジックを追ってみたい。なお、松本他(2003)では、大脳を大脳新皮質と大脳古皮質に分けている。文面から判断すると、大脳古皮質には、大脳辺縁系・大脳基底核までも含まれているようだ。

脳自身が到着した刺激・イベント情報に意識下でどれだけ注意を向けるかどうかは、大脳古皮質、したがってその主な原動力である情(動)が決定する。そして、それが快情動であれば脳が大きく活性化して学習(意)欲が増大し、大脳新皮質での類似性・新奇性判断機構が作動して新たな(知)識がより急速に増えていく。前節で紹介したように、類似性・新奇性判断機構の作動は、階層モジュール構造を形成しているニューロン間ネットワークの重み付けと拮抗りの自動的な変化を伴うので、それまでに脳が生み出したアルゴリズムを改変する形で世界/社会/環境を理解するための新しいアルゴリズム(Hawkins-George 流には世界/社会/環境モデル)が自動的に生み出される。一方、不快情動であれば、学習意欲が減退するため、学習速度が低下する。もちろん、不快であればあるほど、生存を維持するための様々な忌避行動などが学習・選択される。⁶⁷

言い換えれば、脳で目標選択ができないと情動が活性化されにくくなるので、その旺盛な学習意欲、したがってニューロン回路網の興奮状態が十分に喚起されない。そうすると、大脳新皮質が十分に活性化されないの、世界/社会/環境により良く適応するための新たなアルゴリズム(世界/社会/環境モデル)が生まれにくくなる。このように「大脳皮質が古皮質と新皮質の二重構造から成っていて、古皮質側から新皮質にいわば目標を設定する」(松本他(2003))。⁶⁸ その際、大脳古皮質での目標設定は、到着した刺激・イベントに対

⁶⁶ Weizenbaum (2015)は、Minsky (2006)には進化の過程で“バグだらけの人間”を越える情動をも備えたもっと完成度の高い AI の出現が可能とする不遜なアイデアが前提とされているとして彼を強烈に批判している。

⁶⁷ この段落部分の内容は、松本他(2003)に依存している。

⁶⁸ 大脳新皮質と視床とを繋ぐ前者の側に属するインタフェース的な位置を占める島皮質 (Insular Cortex) は、意識上・意識下の情動を生み出す主要部位とされ、ヒトの多彩な情動表現を支えている (Damasio (2010),

する意識下での粗い意味処理に基づき素早く行われる。他方、新皮質では、そのような大脳古皮質の目標設定によりの確な対応をするため、意識下・意識上で時間をかけた緻密な情報処理を行なわれ、それが古皮質側にフィードバックされる。⁶⁹

以上のように、動物の脳では、たとえヒトの場合でも、大脳新皮質が大脳古皮質に従属する形になっており、その様子を松本他（2003）は“情・意・知”と表現するのである。ただし、容易に理解できるように、“情・意・知”という構図だけでは、NM-AI と人的資本/HI との補完・代替関係を十分に論じることはできない。

その大きな理由の一つは、松本他（2003）で暗黙に仮定されている単一自己の考え方である。Minsky（2007）が教えてくれるように、単一自己を想定すると、自らの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なかなかアクセスできない。特に、エピソード記憶は、情動記憶とも呼ばれるように、ヒトやその他の動物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、エピソード記憶再生中にアクセスされる大脳新皮質内の意味記憶ネットワークは、Rolls（2017）が教えてくれるように、局所的に閉じた部分が存在する（大局的ではない）。そのため、特定の記憶断片を辿っていくだけでは辿り着けない記憶が多々存在する。したがって、辿り着ける意味記憶の幅と深さは、情動モードの違いによっても大きく規定されると考えられる。

実際、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、人々のパーソナリティは平常時のもの（主人格）から大きく乖離する。さらに、そのようなモードチェンジ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高いとする。⁷⁰ そして、Minsky（2007）によれば、このような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しているのが情動であり、このような切り替え装置無しには、変化と異常への対応に必須な柔軟で高速の視点切り替えができないとしている。したがって、松本他（2003）の“情・意・知”仮説は、Minsky 流 Emotion Machine と合わせて考えることにより、人的資本/HI に関する

Craig（2015））。その意味では、この部分の二分法は、必ずしもキッチリと成立してないようだ。

⁶⁹ このような二分法は、Daniel Kahneman を嚆矢とする行動経済学が好む認知の二重過程理論（dual process theory）を彷彿させる。ただし、Kahneman（2012）には、情=進化上の古い部分という価値判断が見え隠れするので、依然として“知・情・意”が想定されているのではないだろうか。なお、Anderson（2017）によれば、このような思考傾向は認知科学者にも一般的に当てはまるという。

⁷⁰ Minsky は、複数自己仮説の正当性を、下記のようなヘルマン・ヘッセの『荒野の狼』からの引用を提示する形でも主張している。“For there is not a single human being . . . who is so conveniently simple that his being can be explained as the sum of two or three principal elements. . . . Harry consists of a hundred or a thousand selves [but] it appears to be an inborn and imperative need of all men to regard the self as a unit. . . . Even the best of us shares the delusion.”（「というのは、都合の良いほど単純なためにその存在が 2 つ 3 つの主要な要素（人格）が合わさったものとして説明できるという人間など一人もいないからだ。（主人公）ハリーは、百あるいは千の（複数）自己からなっているが、全ての人達が自己が単一だと見なすことを生まれながらで必須のことだと考えているようだ。たとえ最も優れた人達でもそんな幻想を持っている。」（中馬訳）、（）内は同名追加）

るより緻密なアプローチへの応用が可能となる。

もう一つの大きな理由は、“情・意・知” 仮説では、人的資本/HI には社会ネットワークとして特性 (human capital as social network) が備わっていることを扱えないからである。実際、人的資本・HI の希少性は、社会ネットワークの中でしか生まれない。⁷¹ HI の根幹を成す知識とは、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共有という形で埋め込まれているからである (戸田山(2002))。

人的資本・HI には、中馬(2017)で詳しく論じられているように、個々人の人的資本の構成要素モジュールを繋ぐインタフェース (“内部インタフェース”) と個々の人的資本を繋ぐ/組み合わせるインタフェース (“外部インタフェース”) の二つが存在する。そして、人的資本の“変化と異常への対応力” やその自己変化能は、個々人の人的資本の構成要素がモジュール化されていればいるほど、内部及び外部インタフェースの標準化度が高ければ高いほど、そして、インタフェースを介した情報転送・応答速度が速ければ速いほど、個々人の内外で高まっていく。

さらに、Minsky 流の柔軟で高速な視点切り替え装置としての情動は、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である。言い換えれば、先に紹介した Hawkins-George 流の大脳新皮質型インテリジェンスだけでは、社会ネットワーク性が顕著な人的資本/HI の変化と異常への対応力やその自己変化能をより効果的に生み出すことはできない。この点に関して、Hawkins(2006)は、大脳新皮質型のインテリジェンスの限界をふまえ、そのHI との高い親和性によって両者間で強固な補完関係が生み出されるのではないかとコメントしている。⁷² 再度長い引用で申し訳ないが、傾聴に値する的確な判断だと思われる。

“First, the human mind is created not only by the neocortex but also by the emotional systems of the old brain and by the complexity of the human body. To be human you need all your biological machinery, not just a cortex. To converse like a human on all matters (to pass the Turing Test) would require an intelligent machine to have most of the experiences and emotions of a real human, and to live a humanlike life. Intelligent

⁷¹ Sloman 他(2017)も同じような議論を、自分が知らないことを知っていることの重要性を強調しながら、興味深い形で展開している。

⁷² ただし、この点に関しては、Bostrom (2014) の次のようなハードボイルドな AI 観も存在する。
“Artificial agents need not have any of the social emotions that bind human behavior, emotions such as fear, pride, and remorse. Nor need artificial agents develop attachments to friends and family. Nor need they exhibit the unconscious body language that makes it difficult for us humans to conceal our intentions.” (「AI 型ロボットは、人間の行動を縛っている恐れとかプライドと良心の呵責とかといったいかなる形の社会的な情動を持つ必要はない。さらに、AI 型ロボットは、我々人間が意図を隠すことを難しくしている無意識の身体表現も示す必要は無い。」(中馬訳) 他人への情動的及び認知的な Empathy、したがって Theory of Mind (心の理論) が備わっていなければ、表面上に現れた行動だけからは人々の意図は十分には分からない筈だが……。また、こういう所にも、行動主義的な把握の仕方への考え方の違いが出ていて興味深い。

machines will have the equivalent of a cortex and a set of senses, but the rest is optional. It might be entertaining to watch an intelligent machine shuffle around in a humanlike body, but it will not have a mind that is remotely humanlike unless we imbue it with humanlike emotional systems and humanlike experiences. That would be extremely difficult and, it seems to me, quite pointless.”（「まず、人間の心は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、脳新皮質だけではなく、あなた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。知能のある機械（AI）とは、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）中馬訳）

“They are based on a conflation of intelligence— the neocortical algorithm— with the emotional drives of the old brain— things like fear, paranoia, and desire. But intelligent machines will not have these faculties. They will not have personal ambition. They will not desire wealth, social recognition, or sensual gratification. They will not have appetites, addictions, or mood disorders. Intelligent machines will not have anything resembling human emotion unless we painstakingly design them to. The strongest applications of intelligent machines will be where the human intellect has difficulty, areas in which our senses are inadequate, or in activities we find boring. In general, these activities have little emotional content. Intelligent machines will range from simple, single-application systems to very powerful superhuman intelligent systems, but unless we go out of our way to make them humanlike, they won’ t be.” 「インテリジェンス、すなわち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。個人的な野望をいだくことがない。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)

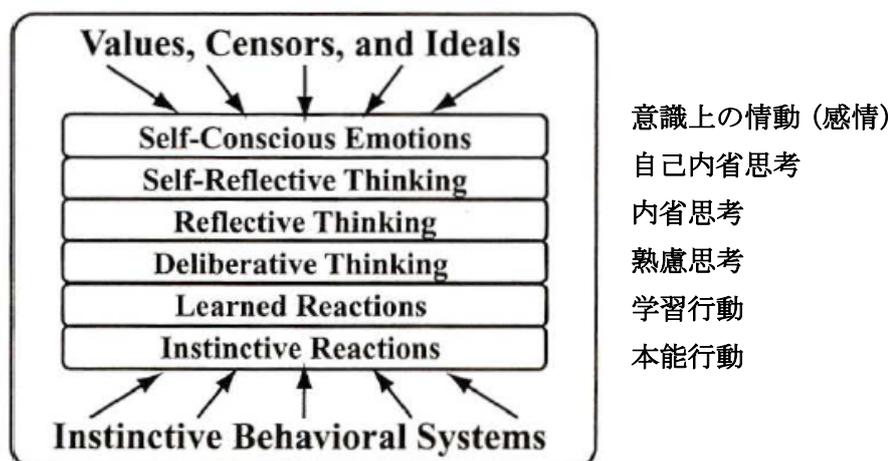
それでは、最後に、柔軟で高速な視点切り替え装置及び広範囲な協力を生み出すコミュニティ形成装置としての情動特性と人的資本/HI の変化と異常への対応力やその自己変能力との関係をクローズアップするために、NM-AI と対比させる形で Google-Deepmind 社の俊才達が編み出している様々な BD-AI に組み込まれていると信じられているインテリジェンスを、下記に紹介する Minsky のインテリジェンスに関する 6 階層モデルと対比させる形で腑分けしてみよう。

Minsky (2007) では、インテリジェンスがどのように捉えられているのだろうか？それを示すのが、下記の引用である。

“Indeed, the quality of our credit assignments could be important aspects of the suitcase of traits that people call “intelligence.” For merely recording solutions to problems will help us only to solve somewhat similar problems, whereas if we can record how we found those solutions, that could further enable us to deal with much broader classes of situations. (「実際、(どのソースからの情報が重要であるかに関して我々が示す価値判断である) 信頼度付与(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)、中馬訳、()内は同名追加)

Minsky は、インテリジェンスという言葉を上記で Suitcase Word(何もかもが十把一絡げにされた言葉)としているが、その中に価値判断に基づく起因の特定機能が含まれていることからも明らかなように、先に紹介した松本他(2003)の脳古皮質の(素早い意識下)機能を含んでいる。そして、我々が意識上・意識下という風に使う時の意識という言葉の Suitcase Word 性を払拭するために、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) 熟慮的思考 (Deliberative Thinking)、

4) 内省思考(Reflective Thinking)、5) 自己内省思考(Self-Reflective Thinking)、6) (意識上の) 情動(Self-Conscious Emotions/Reflection)の六つである。



Minsky(2006)のインテリジェンスに関する6階層

本能行動は、DNA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。**学習行動**は、自他の色々な経験を通じて“If→then rule”(場合場合に応じて学んだ行動)をさす。**熟慮的思考**とは、いくつかの採るべき本能・学習行動を思い浮かべてそれらの効果について比較考量する(意識上・意識下の)思考能力をさす。脳内でのシミュレーションもここに入る。**内省思考**とは、行った行動について考察し、それが妥当であったか否かを思考するための能力をさす。ここまでのインテリジェンスには、自己が含まれていない。**自己内省思考**とは、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能力をさす。最後の**意識上の情動(感情)**とは、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/情動やそれらによって価値観/目的自体を維持・打破・変革することができる能力である。そして、松本他(2003)の“情・意・知”仮説と同じように、インテリジェンスの高度さは、本能行動→学習行動→熟慮的思考→内省思考→自己内省思考→(意識上の)情動という順に上昇していくとする。

上記のMinsky流の6段階のインテリジェンス“座標”を基準にすると、完成形としてのNM-AIの脳皮質型インテリジェンスをどのように相対化できるだろうか?その際の相応しい試金石となるのが、“**変化と異常への対応力**”、その次が同対応力の**自己変化能**だと思われる。では、“**変化と異常への対応力**”をどう定義すれば良いか?その際に、BD-AIとNM-AIの違いをより鮮明にする定義であるようにしたい。より具体的には、人的資本/HIとの相応しい対比という意味で、本論では、当時のBD-AIの限界を説得的に指摘しているWinograd他(1987)が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“Breakdown” = “変化と異常”とするのが良さそうだ。彼らのBreakdownの深遠だが小難

しい部分を捨象すると、それは、下記のように定義できる。

“The essence of intelligence is to act appropriately when there is no simple pre-definition of the problem or the space of states in which to search for a solution. Rational search within a problem space is not possible until the space itself has been created, and is useful only to the extent that the formal structure corresponds effectively to the situation.” (「(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)、中馬訳、() 内は同名追加)

つまり、Winograd 他(1987)は、AI プログラムの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(彼らはシステム化領域(systematic domains)⁷³と呼ぶ)を凌駕する問題が発生した時を“Breakdown”、したがって“変化と異常”の発生と定義している。したがって、このような意味での“変化と異常”とは、松本他(2003)流に言えば、大脳新皮質的なインテリジェンスにとって「脳がメモリ主体型情報システムであることから、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」ということになる。

実際、「コンピュータはプログラムに従って入力情報を処理し出力するが、プログラムが予め備わっていない事柄には全く対応できない。コンピュータの計算汎用性の能力は、コンピュータが対応する事柄に対し予めプログラムを用意しておくことで引き出すことができる。しかし、実社会で起こる全ての事柄に対しプログラムを予め完備しておくことは、極めて難しい。実際この事が難しいので、現在最大公約数的な対応で処理しようとしている。」(松本他(2003))。

このような意味での Breakdown” = “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。3)の「熟慮的思考」はどうだろうか？この段階では状況に応じて記憶済みの“If→then rule”のどれを適用可能であるかのシミュレーションまでできる。ただし、“If→then rule”メニューにない対応はできない。したがって、3)レベルのインテリジェンスでも無理である。ただし、4)の「内

⁷³ より具体的には“systematic domain, which embodies the programmer’s interpretation of the situation in which the program will function” (「システム化領域とは、プログラムが(適切に)機能する状況だとの当該プログラマーの解釈を具現化した領域」(Winograd 他(1987)、中馬訳、() 内は同名追加)。

省思考」レベルのインテリジェンスでは選んだ行動の妥当に関する判断が可能なので、Breakdown” = “変化と異常” の発見はできそうである。

とすると、完成形としての NM-AI の Breakdown” = “変化と異常” への対応力はどうか？パターン認識に秀でる NM-AI には類似性・新奇性判断ができるし、前述したように、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、海馬に蓄積された新しいエピソード記憶は、海馬と大脳新皮質との間でのフィードフォワード&フィードバックを通して、大脳皮質内に構造化/概念化された意味的記憶として蓄えられていく (Rolls (2017))。

さらに、外界からの情報や新皮質内で産み出された情報入出の中継基地としての視床は、Rolls (2017)によれば、ハイパスフィルターの役割も果たす。ハイパスフィルターとは、高周波数のみを通すフィルターであるため、例えば、画像で言えば、画素値の変化が大きいところ(細部の情報)が高周波領域にあたり、画素値の変化が小さいところ(背景など)が低周波領域にあたる。したがって、視床は、前節で触れた海馬内の歯状回と同じように、変化と異常を自動的に強調する役割も担っている。

ただし、この点も前節で触れたように、(人工)海馬と(人工)視床とを含む大脳新皮質型インテリジェンス、したがって NM-AI のこのような変化と異常への対応力は、“メモリ主体型情報システム”(松本他(2003))が備えることのできる CAM (Content-Addressable Memory) 容量に大きく規定される。言い換えれば、最低限 1 万ビット長の CAM が安価で大量に利用できるようにならない限り、Breakdown” = “変化と異常” への対応は夢のまた夢なのである (Kanerva (2014))。逆に言えば、Kanerva 流高次元 CAM が組み込まれた場合の NM-AI の「内省思考」力は、ヒトを遙かに超えた Breakdown” = “変化と異常” への対応力を発揮できる可能性が出てくる⁷⁴

上記の NM-AI に関して述べたことは、BD-AI のインテリジェンスには、どの程度当てはまるのだろうか？この問いに答えることは、Hassabis 他(2017)にサーベイされているように、BD-AI 自体が、Neuroscience の急速な発展と共に NM-AI の目指すところと極めて類似しはじめているので、なかなか難しい。もちろん、前節で紹介した単純な深層ニューラルネットワークモデル (DNN) では、どうか Minsky の 3) 熟慮思考もどきまでであることは容易に理解できる。そもそも、前節で紹介した海馬的な On-Line/One-Shot 学習もできないし、時空間的な構造も扱えない。そもそも、仕組みとして CAM 的な機能に依存した On-Line の類似性・新奇性判断などを実行するにはできていないからである。

ただし、Neural Turing Machine (NTM)で有名になった Graves 他(2014)などになって

⁷⁴ なお、LeDoux(2015)は、人間並のエピソード記憶が(人工)海馬と(人工)視床とを含む大脳新皮質型インテリジェンスに達成可能であることに関しては、自己意識が不可避としている。そして、人間以外では、哺乳類にさえ自己意識は無いのではないかという示唆をしている。となると、LeDoux 流に言えば、人間並の Breakdown” = “変化と異常” への対応力は、少なくとも上記 Minsky(2006)の第 5 層以上のインテリジェンスが必須ということになる。ただし、彼自身も、犬や猫のみならずカラスなどの鳥類やミツバチなどの昆虫類にもエピソード記憶もどきがあることを認めている。

くると、DNNの発展系である時空間的な構造を扱えるRNN (Recursive Neural Network)⁷⁵を用いて短期記憶(一種のワーキングメモリ)を備え、僅かな容量ではあるが、外部メモリとしてCAM的なものも利用されてきている。そのため、Toyモデル的ではあるが、On-Line/One-Shot学習などもできるようになっている。⁷⁶ 同じ構図は、Atariの特定シューティングゲームで得た経験則を他の異なるAtariシューティングゲームに汎化できるといったより高度な状況でも可能となっている(Oh他(2015))。さらに、最近では、Graves(2016)のように、海馬的なエピソード記憶システムを使って記憶間の類似性(Minsky流の“Panalogy”:Parallel Analogy)を見出して問題解決を行う実践例が、ロンドンの地下鉄ネットワーク事例などを用いて示されている。このような試みは、実践例(存在証明)としてのインパクトの強さから、大勢の人々にアピールしている。⁷⁷

したがって、本節の脈絡からすると、完成形としてのNM-AIと既存のBD-AIのインテリジェンスの違いは、冒頭で述べた低消費電力を可能とする仕組みを欠いているという特徴を別とすると、Breakdown”=“変化と異常”への対応力に関しては、敢えて主張できる段階ではないようだ。ただし、素人判断で誠に申し訳ないが、今後さらに複雑な“変化と異常”への対応力が望まれるようになってくると、言い換えれば、Breakdown復旧のための考察の幅と深さが急速に拡大してくると、高次元スパース分散表現(SDR)に基づく確率論理を実行可能な低消費電力性を誇るNM-AIには、消費電力効率で六桁も劣るBD-AIではとても対抗できないだろう。Neuman-Turing型コンピュータに力づくで実装されたBD-AIに比べると、脳模倣型の階層的なモジュール構造を特徴とするNM-AIのスケールメリットの方が極めて大きいと判断されるからである。同じことは、前述したように、未だ解明が十分に進んでいない海馬や未解明部分だらけの視床を組み込んでいく際の親和性の高低に関しても当てはまる筈である。さらに、Kanerva流CAMに関する半導体設計・製造プロセス技術上の革命が起こってくると、あくまで素人判断ではあるが、その順風は、そのようなCAMとの親和性を誇る高次元SDRを中核に据えたNM-AIの方に大きく吹く可能性が高い。

なお、ここまでは、Minskyの4)「内生思考」段階のインテリジェンス、したがって(広義の意味での)大脳新皮質型インテリジェンスの話であった。では、5)の「自己内省思考」や6)「意識上の情動(感情)」とBreakdown”=“変化と異常”への対応力との関係、つまり、そのようなインテリジェンスをも備えている人的資本/HIとの補完性・代替性はどのような形で推移していくと考えられるだろうか。

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)

⁷⁵ DNNを数多く用いた何層にもわたる複雑な入れ子構造をしており、高性能コンピュータが利用可能でなかった時代には、その実用化は夢物語であった。空間的な広がりしか持たないDNNを、RNNの複雑性に依存するが、現況とその時系列的な推移まで扱えるように拡張されている。ただし、外部記憶装置を持たない場合、依然として類似性・新奇性判断は中々できない。なお、1980年初頭に萌芽を見たRNNの発展や各種のRNNに関するコンサイスな説明は、浅川伸一(2017)を参照した。

⁷⁶ 本論文に関しては、岡野原(2017)がコンサイスな形で解説している。

⁷⁷ 進藤(2017)によれば、Google-Deepmind社は、このようなエピソード記憶をも組み込んだBD-AIの枠組みをGoogle Street Viewを有効活用する形で自動運転技術に応用しようと試みはじめているようである。

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。数学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。では、完成形としての 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）自体に合致している否かを判断して Breakdown” = “変化と異常” への対応ができるだろうか？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不可能である。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、Breakdown” = “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ない。そして、敢えてそれを駆使し続けてしまうと、過学習⁷⁸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、Breakdown” = “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人的資本/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

⁷⁸ 甘利(1989)は、過学習に関してこう述べている。「汎化に関連して興味あるのは、過学習の問題である。学習用の信号を固定して、繰り返し学習させると、この信号に対する応答は良くなる。しかし学習をさせすぎると、これらの例題信号に対してはきわめて良好だが、未学習のパターンの識別にはかえって結果が悪くなる。これは、回路網のパラメータの数が多いとき、学習用のパターンの個々の特殊性に適合するように、一般性のない性質にまで合わせて学習してしまうからである。このため、こうした特殊性を持たない一般の未学習のパターンに対してはかえって能力が落ちる。受験勉強など、特殊の問題に対する勉強のしすぎは良くないのである。」

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情動なのである。

この点に関する Winograd 他(1986)の下記の言明は、“Breakdown” = “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。

“As we have pointed out in earlier chapters, a person is not permanently trapped in the same kind of blindness. We have the potential to respond to breakdown with a shift of domains in which we enter into new commitments. Understanding is not a fixed relationship between a representation and the things represented, but is a commitment to carry out a dialog within the full horizons of both speaker and hearer in a way that permits new distinctions to emerge.”(「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。合意するという事は、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。」(中馬訳、()内は同名追加)

言い換えれば、大脳新皮質型インテリジェンスでも、それを越えた広い意味でのインテリジェンスでも、それらは個々人の脳内だけに存在しているわけではなく、社会を構成する様々な人々の間の社会ネットワーク構造の中にも埋め込まれている。つまり、個々人のインテリジェンス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ると同じように、そのようなインテリジェンス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の人的資本/HIの威力は、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ようになる。つまり、ALVA NOË(2009)や河野哲也(2006)、浅田(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拡がっている」と表現できる・・・。

このように拡大解釈すると、変化と異常への対応力/自己変化能には、大脳新皮質型インテリジェンスに埋め込まれたものと Minsky 型の情動を含む広義のインテリジェンスに埋め込まれたもの(したがって人的資本/HI)との二種類が存在することになる。そして、両者は、Minsky 的な Emotion Machine が夢物語である限り、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（松本・Minskyの意味で）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。

4. 結びに代えて

本論では、人的資本/Human Intelligence (HI) と人工知能 (AI)、特にAIの本丸とも言われる脳型人工知能 (NM-AI) との代替・補完性について、そもそもインテリジェンスとは何かという視点に拘りつつ検討した。今をときめくビッグデータ型AI (BD-AI) ではなく敢えてNM-AIに焦点を合わせたのは、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に代わって、中長期的にはNM-AIこそがAIの中核となっていくと予測されるからである。特に、完成形としてのNM-AIは、ヒトの脳に匹敵する20ワット前後の低消費電力で大脳新皮質に酷似した強力な学習・記憶・帰納力を発揮しはじめると考えられる。

もちろん、その際に、BD-AIもさらに進化していくと考えられる。⁷⁹ ただし、現行のBD-AIよりも6桁ほどの消費電力効率を誇るヒトの脳にも比肩するNM-AIの実現は、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、人的資本/HIに有史以来の衝撃的なインパクトを与えはじめるとは間違いない。NM-AIが、現在のPCやスマホとは大きく異なり、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ような対応力自体の自己変化能力を格段に高めるので、とてつもなく大きなインパクトを大衆レベルでもたらすからである。そして、第3節で言及したように、現行の勉強秀才を支える“知・情・意”仮説が、感性に溢れた大衆を支える“情・意・知”仮説の前に色褪せる時代が到来するかもしれない。いずれにせよ、完成形としてのNM-AIがもたらす相変化は、その構図は酷似しているが、メインフレームコンピュータの時代からPC/スマホの時代への相変化の比ではないだろう。

本論では、上記のような現状認識や将来予測の妥当性を検討するために、まず、なぜいまNM-AIに注目する必要があるのかについて社会科学のみならず計算機科学・半導体集積回路やNeuroscienceの視点から検討した。検討結果を一言でまとめると、門外漢がこんなことを言うとAI分野のプロ達に叱られてしまうが、現行のPower Hungry(電力多消費型)なノイマン/チューリング型コンピュータに依存するBD-AIは、既に現代社会に大きな相変化をもたらしつつあるが、中長期的な視点からはNM-AIへの橋渡しの役割を果たしているのではないだろうか、ということになる。そして、NM-AIが中長期的に実用化に至らないとしたら、こちらこそ誠に失礼な表現で申し訳ないが、エネルギー制約という大きな壁の前で、

⁷⁹ 例えば、マイクロソフトCEOのSatya Nadella(2017)が繰り返して強調しているように、BD-AIが量子コンピュータとの相補性をさらに高めていくと考えられるので、BD-AIの重要性が低下することはない筈である。ただし、それは、Google、Microsoft、IBM、Amazonなどのビッグプレイヤーによる上から目線のクラウドコンピューティング環境提供というよりも、NM-AIを駆使することによって多能化・多様化した個人や企業・組織の活動への御用聞き的な目線からのクラウドコンピューティング環境提供という形に変わっていくのではないだろうか？いや、Nadellaが繰り返し唱えているように、そうなってほしいものだ。

AI 研究が再度“冬の時代”に突入してしまう可能性も未だある。⁸⁰

そのような感想に至った根拠は、詳細は本論の第2節を参照していただきたいが、非同期・イベント駆動型でありメモリ・論理同居型のニューロンという非ノイマン/チューリング型コンピュータに組み込むべき数々の低消費電力化の仕組みの一端を理解したことによる。より具体的には、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確率過程化されたニューロン/ニューロン集団間ネットワーク処理・決定システム、そのようなシステムを支える哺乳類のみならず節足動物の昆虫類にも組み込まれている耐障害性に富んだ高次元スパース分散表現 (SDR) を駆使した並列分散処理機構等々である。

もう一つの根拠は NM-AI の実用化が中々達成できない決定的な要因のいくつかについて知り得たことによる。特に、予想通りではあったが、Neuroscience 上の発展 (解明) 速度が NM-AI の大きな律速要因となっていた。そのため、完成形としての NM-AI の実現可能性には、今後の Neuroscience のさらなる飛躍的発展が必須のようだ。さらに、Neuroscience に勝るとも劣らない律速要因が、半導体製造・設計上の隘路であることには正直驚いた。特に、メモリ・論理同居型のニューロン・コンピュータに必須の Kanerva (1988) 流 CAM (Content-Addressable Memory) の実現は、現行の Moore の法則を支えてきた半導体微細化のための製造・設計技術をそのまま延長する限り、プロ中のプロでも立ち尽くしてしまうほど極めて難しい課題であることを再認識した。

そして、高次元 SDR を駆使するために必須の高次元 CAM 及び高次元コンピューティング (Kanerva (2014)) の実現・普及なしには、NM-AI の本領であるエピソード記憶/One-shot Learning/リアルタイム学習がエッジ (最終ユーザー) レベルでなかなか実現できない。となると、変化と異常への臨機応変な対応力やその自己変化能を中核で支える帰納・アブダクション (補完) 能力もなかなか飛躍的には高まらない。そして、CAM 実現に必須の未解明現象の突破には、門外漢の印象で誠に申し訳ないが、AI、Neuroscience、計算機科学、集積回路製造・設計、果ては量子物理学⁸¹などの幅広い領域にまたがる世界規模での共同研究が必須……。そして、この領域でも、基礎的な Neuroscience/Computational Neuroscience 研究の希少性が大幅に高まることが予見できる。

では、完成形としての NM-AI と人的資本/HI との補完性・代替性についてはどうだろうか？本論では、この難問に取り組むため、まず、そもそも NM-AI が中長期的に実現しようと

⁸⁰ 同様のコメントが、Zarkadakis (2015)にも述べられている。ただし、NM-AI に関する理解、特にその情動との関係についての理解が本論と大きく異なっているためだろうか、NM-AI への畏怖が異様に大きい。ちなみに、同書では、情動と心の理論 (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) との密接な関係にはたびたび言及しているが、Minsky (2006)に関する言及はない。さらに、心の形成に関する大脳新皮質の統合的な役割が、松本他 (2008)やHawkins (2004)等に比べて異様に強調されている。著者が人工知能の研究者だったからだろうか、あるいは西洋思想と東洋思想の違いが出ているのだろうか……。筆者には、NM-AI 及び大脳新皮質の役割に関する理解が根本的に間違っているからのように思える。

⁸¹ 門外漢にはほとんど理解不能ではあるが、Rinkus (2012)のように、高次元 SDR と量子力学との親和性が極めて高いとする理論仮説も存在する。

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ついて考察した。その際に理解したことは、チューリングテストのような人間行動主義的な定義の限界だった。直ぐに理解できることであるが、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等にも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し、異なった生活圏に適応してヒトとは異なった驚くべき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。

となると、完成形としての 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か？そこで重要な視点を与えてくれたのが Hawkins (2004) や George (2008)、Sloman (1994) だった。その眼目は、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するということである。実際、BD-AI のみならず完成形としての NM-AI ですらも、大脳辺縁系や大脳基底核を基点とする意識下・意識上の情動や直観、あるいはそれらの背後にある自己・自意識を組み込むことなどは近未来的には夢のまた夢であり、そもそも超長期的にも AI に情動や自己・自意識を組み込む意義・意味があるのかという素朴な疑問も残る。

だとすると、大脳新皮質では何が行われているのか？この問いに答えるために Hawkins (2004) や George (2008) が提案するのは、なぜ大脳新皮質が階層構造をしているのか？という問いに答える形でインテリジェンスを生み出す内部メカニズムの本質を探ることである。

彼らによ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外界が各種の物理・化学法則に支配された（モジュール）階層構造をしていることの反映である。第二に、そのような外界の中で繰り返し発生する物事や事象に対してより効果的に適応しながら生き抜くためには、外界の階層構造の写し絵としての脳内階層モデルを自他の経験に基づきながら作り上げる必要があるからとなる。そして、第三に、それらの脳内階層モデルを基準として、そこからの対比で変化と異常にいち早く気づいて対応したり、さらに、そのような気づきに基づいてモデルのさらなる精緻化を継続的に成し遂げていくことによって適応能力（自己変化）の向上を達成できるからである。

したがって、NM-AI に実装される大脳新皮質型インテリジェンスとは、過去の記憶と現況に基づきながら環境適応していくための各自の世界/社会/環境モデルに基づいた予測能力ならびにその学習能力に基づく自己変可能と定義できるのである。

では、このような NM-AI の完成形としての大脳新皮質型インテリジェンスは、人的資本 / Human Intelligence (HI) とどのような補完・代替関係にあるのか？この難問に立ち向かうために依拠したのが Minsky (2006) の HI に関する 6 階層モデルだった。Minsky は、ヒトを含む生命体の低位から高位に至るインテリジェンス活動を、本能、学習、熟慮、内省、自己内省、(意識上の) 情動の六つであると見なす。そして、AI への情動や自己・自意識の組み込み込み込みが可能だと信じている Minsky は、大脳新皮質型インテリジェンスを越える NM-AI の実現には、少なくともこれらの 6 階層構造を持つ実体、つまり彼が呼ぶ Emotion Machine

(EM) が不可欠であるとする。その本質的な理由は、意識上・意識下の情動とは、変化と異常への対応に必須な柔軟で高速の視点切り替え装置であり、互いの多様な意図や常識・文化等々の共有による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との興味深い視点である。

なお、本文の繰り返しになるが、本能行動は、DNA に組み込まれている行動パターンである。学習行動は、自他の色々な経験を通じて “If→then rule” (場合場合に応じて学んだ行動) をさす。熟慮的思考とは、いくつかの採るべき行動を思い浮かべてそれらの効果について比較考量する思考能力をさす。脳内でのシミュレーションもここに入る。内省思考とは、行った行動について考察し、それが妥当であったか否かを思考するための能力をさす。自己内省思考とは、自分の採った行動が自分の理想とするものに値するか否かについて思考する能力をさす。最後の意識上の情動 (感情) とは、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/情動やそれらによって価値観/目的自体を維持・打破・変革することができる能力である。そして、先の松本他 (2003) の “情・意・知” 仮説と同じように、インテリジェンスの高度さは、本能行動→学習行動→熟慮的思考→内省思考→自己内省思考→ (意識上の) 情動という順に上昇していくとする。

では、完成形としての NM-AI は、上記の Minsky 流インテリジェンス (HI) のどの辺りまで行き着くと言えるだろうか? 本論では、それを、変化と異常への対応力という視点から分析してみた。分析の際には、Winograd 他 (1987) の “Breakdown” 概念を試金石として用いた。彼らによれば、“Breakdown” の発生、したがって “変化と異常” の発生とは、AI プログラムの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 domains)) を凌駕する問題が発生した時である。

“Breakdown” の発生、したがって “変化と異常” への対応は、先の Minsky 流インテリジェンスの 1) 「本能行動」、2) 「学習行動」まででは無理である。3) の「熟慮的思考」でも “If→then rule” メニューにない対応はできない。ただし、4) の「内省思考」は選んだ行動の妥当に関する判断が可能なので、“変化と異常” の発見はできる。では、完成形としての 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) 自体に合致している否かを判断して Breakdown” = “変化と異常” への対応ができるだろうか?

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するためには、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不可能である。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、Breakdown” = “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。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、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。さらに、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。そして、それを可能にするのが、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くと、変化と異常への対応力/自己変化能には、大脳新皮質型インテリジェンスに埋め込まれたものと Minsky 型の情動を含む広義のインテリジェンスに埋め込まれたもの (HI) との二種類が存在することになる。そして、両者は、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と予想される。

最後に本論での未検討課題について触れておきたい。BD-AI 及び NM-AI は、過去に蓄積された知識や学習経験に基づいて直面する既知・未知の問題を発見・解決するために有効な汎用/汎化則を帰納的に導出しようとする。そのため、導出過程がブラックボックス化し、事前には予測不可能であるが事後的には有害な帰納的認知バイアスや過学習結果をも生み出す。つまり、帰納プロセス特有の“ヒュームの呪い”⁸² (戸田山(2002)) が以前にも増して顕在化してくる。そのため、認識論や倫理学の視点からも、NM-AI のインテリジェンス特性に関する精査や規制が不可欠となる。NM-AI 研究に社会科学的な視点も必要となる理由の一つだと考えられる。本論では、これらの点についてほとんど触れることができなかったが、今後の重要な検討課題としたい。

⁸² Popper (1962)には、“even after the observation of the frequent or constant conjunction of objects, we have no reason to draw any inference concerning any object beyond those of which we have had experience.” (「頻繁なあるいは一定の共起する対象が観察されとしても、それがどのような対象であれ、それらの対象を我々が経験してきたという範囲を越えて一般化できるという根拠はない。」(中馬訳)) という Hume からの引用が示されている。

参考文献

- 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コネクショニズム』、東京大学出版会(初版は1989)
- 甘利俊一他 (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
- Arthur, John V. and Kwabena Boahen, “Learning in Silicon: Timing is Everything,” *Advances in Neural Information Processing Systems*, 18, pp.75-82
- 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
- Ben Varkey Benjamin, Peiran Gao, Emmett McQuinn, Swadesh Choudhary, Anand R. Chandrasekaran, Jean-Marie Bussat, Member IEEE, 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
- Brogliato, Marcelo S., Daniel M. Chada and Alexandre Linhares (2014), “Sparse distributed memory: understanding the speed and robustness of expert memory,” *Frontiers in Human Neuroscience*, April, Volume8, Article 222, pp.1-11
- Bostrom, Nick (2014), *Superintelligence: Paths, Dangers, Strategies*, Oxford University Press
- 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 Steinhardt, Springer Science & Business Media, Heidelberg, New York, Dordrecht, London
- 中馬宏之 (2017)、「人工知能が社会に与えるインパクトの考察：文理連繋の視点から」、経済産業研究所ディスカッションペーパー、2017年5月、17-P-015 (<http://www.rieti.go.jp/jp/publications/summary/17050003.html>)
- 中馬宏之(2011)、「半導体産業における国際競争力低下要因を探る：ネットワーク分析の視点から」、一橋大学経済研究所『経済研究』、62巻3号、225-240 頁

- Craig, A.D. Bud (2015), *How Do You Feel? An Interoceptive Moment with your Neurobiological Self*, Princeton University Press, Princeton and Oxford
- Damasio, Antonio (2010), *Self Comes to Mind: Constructing the Conscious Brain*, Pantheon Books, New York 2013
- Dieny, Bernard and Lucian Prejbeanu (2017), “Magnetic Random-Access Memory,” in *INTRODUCTION TO MAGNETIC RANDOM-ACCESS MEMORY*, ed. By BERNARD DIENY, RONALD B. GOLDFARB and KYUNG-JIN LEE, pp.101-164, IEEE Press, Wiley
- Eliasmith, Chris (2013), *How to Build a Brain: A Neural Architecture for Biological Connection*, Oxford University Press
- Finnegan, Rory, Mark Shaw and Suzanna Becker (2017), “Restricted Boltzmann Machine Models of Hippocampal Coding and Neurogenesis,” in Ooyen, Arjen Van and MARKUS Butz-Ostendorf ed. *The Rewiring Brain: A Computational Approach to Structural Plasticity in the Adult Brain*, pp.443-461
- Fisher, Anna V. (2016), “Mechanisms of Induction Early in Development,” in Marie T. Banich and Donna Caccamisse ed. *Generalization of Knowledge: Multidisciplinary Perspectives*, Routledge, pp. 89-112.
- Franklin, Stan (1995), *Artificial Minds*, MIT Press
- 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)
- 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
- Greve, Andrea and Richard Henson (2015), “What We Have Learned about Memory from Neuroimaging,” in Donna Rose Addis, Morgan Barense, and Audrey Duarte ed., *The Wiley Handbook on the Cognitive Neuroscience of Memory*, pp. 1-20
- Griffiths, Thomas L. (2016), “Bayesian Models as Tools for Exploring Inductive Biases,” in Marie T. Banich and Donna Caccamisse ed. *Generalization of Knowledge: Multidisciplinary Perspectives*, Routledge, pp. 135-156.

- Hassabis, Demis, Dhharshan 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*, 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
- Hurwits, Marcia Kaufman and Adrian Nowles (2015), *Cognitive Computing and Big Data Analytics*, John Wiley & Sons, Inc.
- 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
- Johnsen, Sönke and Kenneth J. Lohmann (2005), “The Physics and Neurobiology of Magnetoreception,” *NATURE REVIEWS, NEUROSCIENCE, VOLUME 6, SEPTEMBER*, pp.703-712
- Kahneman, Daniel (2012), *Thinking, Fast and Slow*, Penguin
- 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*
- Keigo Kohara, Michele Pignatelli, Alexander J Rivest, Hae-Yoon Jung, Takashi Kitamura, Junhyup 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_2/参照)
- Khodagholy, Dion, Jennifer N Gelinasi, Thomas Thesen, Werner Doyle, Orrin Devinsky, George G Malliaras and György Buzsáki (2014) , “NeuroGrid: recording action potentials from the surface of the brain” *Nature Neuroscience* 18, 310-315
- 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
- 小柳義夫・中村宏・佐藤三久・松岡聡 (2012)、『岩波講座・計算科学・別巻『スーパーコンピュータ』』、岩波書店
- 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
- Lotto, Baeu R. (2009), “Statistical Analysis of Visual Perception,” in Larry R. Squire ed., *Encyclopedia of Neuroscience*, pp.381-386

- 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>)
- 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
- McFadden, Johnjoie and Jim Al-Khalili (2014), *Life on the Edge: The Coming of Age of Quantum Biology*, Crown Publishers/Penguin.
- Mikolov, Tomas, Kai Chen, Greg Corrado, and Jeffrey Dean (2013), “Efficient Estimation of Word Representations in Vector Space,” arXiv:1301.3781v3 [cs.CL] 7 Sep, pp. 1-12
- 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
- 水波誠 (2009)、「微小脳と巨大脳：自然は多彩な脳を生み出した」、科学、Vol.79、No.6、636-641 頁
- Nadella, Satya (2017), *Hit Refresh: The Quest to Rediscover Microsoft’s Soul and Imagine a Better Future for Everyone*, Harper Business
- Numenta (2017), *Biological and Machine Intelligence (BAMI): A living book that documents Hierarchical Temporal Memory (HTM)*, <http://numenta.com/assets/pdf/biological-and-machineintelligence/BAMI-Complete.pdf>
- 岡野原大輔(2017)、「Differentiable Neural Computers：外部記憶を備えたニューラルネットワーク」、日経ロボティクス、2917年1月号、36-38 頁
- Oh, Junhyuk, Xiaoxiao Guo, Honglak Lee, Richard Lewis, and Satinder Singh (2015), “Action-Conditional Video Prediction using Deep Networks in Atari Games,” arXiv, arXiv:50708750, pp. 1-26
- Olshausen, 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
- Olshausen, Bruno A. and David J. Field (1996), “Emergence of simple-cell receptive field properties by learning a sparse code for natural images,” Nature, vol. 381, 13, June, pp.607-609

- Orunstein, Robert Evan (1986), *Multimind: A New Way of Looking at Human Behavior*, Houghton Mifflin
- Ooyen, Arjen Van and Markus Butz-Ostendorf ed. *The Rewiring Brain: A Computational Approach to Structural Plasticity in the Adult Brain*, Academic Press
- Plate, Tony A. (2003), *Holographic Reduced Representation: Distributed Representation for Cognitive Structures*, CSI Publications
- Poldrack, Russell A., VALERIE CARR, and KARIN FOERDE 他 (2016), “Flexibility and Generalization in Memory Systems,” in Marie T. Banich and Donna Caccamise ed. *Generalization of Knowledge: Multidisciplinary Perspectives*, Routledge, pp. 53-70.
- Popper, Karl R. (1962), *Conjectures and Refutations: The Growth of Scientific Knowledge*, Routledge; 2 版 (2014/5/1)
- Raikov, Ivan, Mark Plitt and Ivan Soltesz (2017), “Network Models of Epilepsy-Related Pathological Structural and Functional Alterations in the Dentate Gyrus,” in Ooyen, Arjen Van and MARKUS Butz-Ostendorf ed. *The Rewiring Brain: A Computational Approach to Structural Plasticity in the Adult Brain*, pp. 485-503
- Rast, Alexander D., M. Mukaram Khan, Xin Jin, Luis A. Plana, and Steve B. Furber (2012), “A Universal Abstract-Time Platform for Real-Time Neural Networks,” in A. Ravishankar Rao and Guillermo A. Cecchi ed. *The Relevance of the Time Domain to Neural Network Models*, Chapter 8, pp.135-157
- 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
- Rinkus, Gerard(2012), “Quantum Computation via Sparse Distributed Representation,” *NeuroQuantology*, Volume 10, Issue 2, pp.311-315
- Rolls, Edmund T. (2016), *Cerebral Cortex: Principles of Operation*, 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
- Russo, Eleonora, Pirmoradian, Sahar and Alessandro Treves (2013), “Encoding Words into a Potts Attractor Network,” in J. Mayor and P. Gomez (eds.), *Computational Models of Cognitive Processes: Proceedings of the 13th Neural Computation and Psychology Workshop (NCPW13)*, World Scientific Press, Singapore, pp. 29-42. (http://people.sissa.it/~ale/Pirm_Bilbao)

- 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
- 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 Graepell and Demis Hassabis (2016), “Mastering the game of Go with deep neural networks and tree search,” *Nature*, Vol. 529, 28 January, pp.484-489
- 進藤智則(2017)、「グーグル DeepMind が移動ロボの技術に進出：ディープラーニングで SLAM 機能獲得、Street View 利用」、日経ロボティクス、2017 年 11 月号、5-12 頁。
- 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
- Sloman, Steven and Philip Fernbach (2017), *The Knowledge Illusion: The myth of individual thought and the power of collective wisdom*, Macmillan; Main Market 版
- Srinivasa, Narayan (2015), “Neuromorphic Electronics for Energy Efficient Processing,” ITRS ERD Presentation, Stanford University, 02/26/15
- 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>)
- Stromatias, Evangelos, Francesco Galluppi, Cameron Patterson and Steve Furber (2013), “Power analysis of large-scale, real-time neural networks on SpiNNaker,” *International Joint Conference on Neural Networks (IJCNN)*; 04 Aug 2013-09 Aug; IEEE; pp.1570-1577.
- Trappenberg, Thomas P. (2010), *Fundamentals of Computational Neuroscience*, 2nd, Oxford University Press
- 戸田山和久(2002)、『認識論』、産業図書
- 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/)

- ・豊泉太郎 (2017)、「数理モデルでつなげる脳の仕組み」、理化学研究所脳科学総合研究センター編 (Kindle 版)、『つながる脳科学：「心のしくみ」に迫る脳研究の最前線』第5章
- ・坪井裕太・海野祐也・鈴木潤(2017)、*深層学習による自然言語処理*、講談社
- ・Van Der Made, Peter (2013), *Higher Intelligence: How to Create a Functional Artificial Brain*, Vivid Publishing
- ・渡辺茂(2010)、『鳥脳力:小さな頭に秘められた驚異の能力』、化学同人
- ・Weizenbaum, Joseph with Gunna Wendt (2015), *ISLANDS IN THE 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
- ・Zarkadakis, George (2015), *In Our Own Image: Savior or Destroyer? The History and Future of Artificial Intelligence*, Pegasus Book, New York, London

補論： BD-AI と NM-AI の現況を概観する （今井正治）

この補論では、本論でのインテリジェンスに関する議論をより確かにするために、AI 実現のためのコンピュータに要求される処理内容および性能、現在提案されている AI 向きコンピュータの種類を計算モデルに基づいた分類、電力効率の高い AI 向きコンピュータの実現に必要なコンピュータの低消費電力化への試みを紹介しておきたい。また、少しだけであるが、AI 向きの計算機アーキテクチャの発展の方向性についても触れる。

AI 技術が現在実用化に近づきつつある理由の一つは、半導体集積回路技術の進歩の結果、非常に高性能でエネルギー効率の良いコンピュータが実装可能になったからである。AI 技術の中で現在実用化が急速に進みつつあるのは、極めて単純化された人工ニューラルネットワーク（Artificial Neural Network：ANN）の計算モデルに基づく深層学習（ディープラーニング）技術を用いた BD-AI である。

BD-AI の計算モデルでは、単純化された多数の人工ニューロンの振舞いをシミュレートする必要がある。そのため計算量が膨大となり、従来のコンピュータでは実時間で問題の処理が不可能であった。しかし、半導体集積回路技術の進歩により高性能なスーパーコンピュータが実装可能となり、この計算モデルに基づく BD-AI の実現が可能となった。BD-AI の他の応用分野、たとえば、証券取引、医療診断、囲碁や将棋などのゲームなどにおいても、高性能なコンピュータがなければ実用化は不可能であったと言って良い。

AI 応用システムは、特定の問題を解く事を目的としており、「専用 AI」と呼ばれている。これに対して、与えられた問題の解決方法を発見できるヒト型の AI は「汎用 AI」と呼ばれて区別されている。ニューラルネットワークに基づく計算モデルは、ヒトの脳のニューロンのネットワークをある抽象度でモデル化したものであり、個々のニューロンの計算モデルは比較的単純である。そのため、この計算モデルでパターン認識のような計算を実行できることは直感的に理解しやすいが、この計算モデルを用いて汎用 AI を実現する方法は自明ではない。

そのため、汎用 AI を実現するためのアプローチとして、ニューロモルフィック・コンピューティングと呼ばれる脳の構造や機能に基づく計算モデルを用いたシミュレーション手法や計算機システムも研究されている[Asai2014]。この計算モデルのうちで、最も抽象度の低い計算モデルでは、イオンなどの生化学的な振舞いを扱っているために、アナログ的な解析手法が必要となり、ニューロンの発火も時系列信号として扱われている。このようなモデルをデジタル・コンピュータで精密に扱うためには浮動小数点演算を行う必要があり、計算コストが一気に増加してしまう。また計算に必要な電力も増加することになる。

AI を実現するためには、膨大な計算を効率良く実行できる高性能なコンピュータが必要なのは当然であるが、コンピュータで消費される電力は非常にクリティカルな問題である。ヒトの脳の消費エネルギーは電力換算で約 20W と言われているが、スーパーコンピュータ「京」の消費電力は 2011 年当時で約 13MW であり、このままではモバイル環境で使用可能なポータブルな AI 応用システムの実用化は不可能である。したがって、このような AI 応

用システムの実用化のためには電力効率を大幅に改善することが必要である。

以上のような観点から、本節では、現在専用 AI で用いられているニューロモルフィック・コンピューティングを効率扱うためのコンピュータ・システムを紹介する。また、低消費電力化を達成するためのアーキテクチャ上の工夫についても言及する。最後に、低消費電力化を更に進めるための最先端の半導体デバイス技術を紹介し、そこで発生する統計的なゆらぎの問題とその解決方法についても展望する。

S.1 AI のためのコンピュータの開発状況

現在、世界の先進工業国では、AI のための高性能コンピュータの開発が活発に行われている。

A. 米国での開発状況

米国では、多数の企業が AI 関連のビジネスを行っている。まず IBM 社は自社の計算サーバ上のソフトウェアとして Watson という名称のコグニティブ・コンピューティング・システムを開発してユーザに提供している [Watson]。それ以外にも、ニューロモルフィック・コンピューティング用（主にエッジ用）に、TrueNorth という VLSI を開発してきた。

また、Microsoft 社、Google 社、Amazon 社も計算サーバを用いた AI サービスを提供している。Google 社は自社の Tensorflow という AI システムを構築するためのツールキットを無償で提供している [Tensorflow]。さらに Tensorflow の実行を加速するアクセラレータとして TPU (Tensor Processing Unit) を発表している [TPU]。

従来は画像処理用のプロセッサである GPU (Graphical Processing Unit) を開発していた NVIDIA 社は、GPU をベースにした AI 用アクセラレータとして従来の CUDA シリーズに加えて Tesla シリーズの VLSI を販売している [NVIDIA]。それ以外にも、複数の組織による深層学習用アクセラレータの共同開発のプロジェクトが実施されている。

B. 欧州での開発状況

欧州では、EU が Human Brain Project (HBP) という名称の包括的な 10 年計画の科学技術プロジェクトを 2013 年に開始した。このプロジェクトには、AI の研究および AI 応用を目的とした 2 つのサブプロジェクトが含まれている。一つは、BrainScaleS という名称のプロジェクトで、ドイツのハイデルベルグ大学を中心にして実施されている。この研究グループは、脳の生化学的なモデルに基づくシミュレーションを高速に実行するために、アナログ・デジタル混載の専用 LSI を開発している。現在の BrainScaleS システムは 20 枚のウェハから構成され、約 400 万個のニューロンと約 10 億個のシナプスから構成されるシステムのシミュレーションが可能である。

もう一つは、SpiNNaker という名称のプロジェクトで、英国のマンチェスター大学を中心にして実施されている。SpiNNaker では 50 万個の ARM プロセッサから構成されるメニーコ

アプロセッサ・システムで5×5のクラスタから構成され、4億6千万（460M）個のニューロンと4600億（460B）個のシナプスから構成されるシステムのシミュレーションが可能である。

C. 中国での開発状況

中国は、スーパーコンピュータの開発に力を入れており、2016年6月のスーパーコンピュータの世界ランキングTOP500で中国製のスーパーコンピュータ「神威太湖之光」が一位を獲得したことは記憶に新しい。中国科学院（CAS）は、2014年に深層学習向きのDRAM混載型のアクセラレータとしてDaDianNaoを発表している[DaDianNao]。

D. 韓国での開発状況

韓国では、KAISTを中心にして、深層学習用のアクセラレータの研究開発を行っている[KAIST]。

E. 日本での開発状況

日本では、「京」のようなスーパーコンピュータの開発が中心であり、欧州のような国家レベルでの本格的なニューロモルフィック・コンピューティング用アクセラレータの開発は開始されていない。また民間企業での研究開発も活発とは言えない。日本に関して特筆すべき研究開発事例としては、PEZY Computing社がSatsuki（阜）、Shoubu（菖蒲）などのスーパーコンピュータを独自に開発していることである[PEZY]。これらのスーパーコンピュータは、2016年のスーパーコンピュータの省エネルギーランキングであるGreen500で1位および2位を獲得している。また、これらのスーパーコンピュータは理化学研究所に納入され、研究に用いられている。

S.2 AIのための計算モデル

現在研究開発が行なわれているAI応用システムの主流は、ニューラルネットワークの計算モデルに基づいている。まず、ニューロンそのものの計算モデルは抽象度レベルに関しては次のように分類できる。

A) 積和演算型モデル

たとえば最も抽象度の高い計算モデルは、各ニューロンは個々のシナプスからの入力に係数を掛けて積和演算を行い、演算結果が閾値を超えると発火するといった単純なモデルになっている。この演算モデルはAI第2次ブームの1980年代から用いられてきた。現在は、深層学習用のアクセラレータの基本演算のモデルとして用いられている。

B) 時系列型演算モデル

TrueNorthやBrainScaleSなどでは、ヒトの脳のニューロンでの生化学的な振舞いに近い計

算モデルを採用している。このモデルではスパイク状の発火信号が時系列的に発生して伝搬するメカニズムを扱っている。ニューロンは受け取った発火信号の個数によって、ある確率で発火するような仕組みになっている。

C) 生化学的動作モデル

BrainScaleS などでは、脳のニューロンのイオンなどの生化学的なレベルでの振舞いをアナログ回路でシミュレーションしている。このレベルでのシミュレーションを行うことにより、脳科学での新しい知見が得られる可能性が高いからである。また、ニューロンを接続するネットワークのモデルに関しては、ヒトの大脳皮質の構造にしたがう 6 層のモデルや記憶を司る海馬のモデルを組合わせた複雑なモデルも採用されている。このような複雑なモデルを用いることによって、ヒトの脳の働きにできるだけ近い機能を実現しようとしているわけである。

ニューロンそのものの計算モデルが複雑になり、ヒトの脳の構造に近い複雑なネットワークモデルを採用すると、ヒトの脳の働きに近い機能が実現できると期待される。しかし、計算モデルが複雑になると、演算が複雑になるだけでなく計算に必要な消費電力も増大することになる。ヒトの脳は約 1500 億個のニューロンから構成されていると言われてるので、従来のノイマン型のスーパーコンピュータでヒトの脳全体のシミュレーションを行うためには膨大な電力が必要になる。

S.3 Neuromorphic Computing のためのコンピュータ・アーキテクチャ

現在我々が用いているノイマン型のコンピュータでニューラルネットワークに基づく深層学習型の計算モデルを実行するには様々な限界がある。最大の問題はノイマン型コンピュータのアーキテクチャにあり、メモリと演算装置を接続するバスが演算性能を制限するあい路となっている。この問題は、フォンノイマン・ボトルネックと呼ばれている。すなわち、ノイマン型コンピュータでは、プログラムを構成する命令がバスを経由して読み出される。また、データはメモリからバスを経由して演算装置に読み込まれて演算が実行され、演算結果はバスを経由してメモリに書き出されるという動作を逐次的に繰り返しているからである。

一方、脳の構造をモデル化したニューラルネットワークでは、個々のニューロンは並列に演算を行っており、演算対象となるデータは軸索からシナプスを経由して並列にニューロンに入力される。ノイマン型のコンピュータは、このような並列性の非常に高い計算モデルのシミュレーションに適していない。また、メモリアクセスに必要な消費電力がコンピュータの消費電力全体の非常に大きな部分を占めているので、電力効率の点からもノイマン型コンピュータがニューラルネットワークのシミュレーションに適していないことは明らかである。すなわち現在広く使われているノイマン型のコンピュータは、AI 時代の主力演算装置とはなりえない。

もちろん、ノイマン型のコンピュータが使われなくなるわけではないが、その役割は逐

次的なアルゴリズムの実行やシステムの制御を行うためのコンピュータとなるであろう。また、後で述べるニューラルネットワークのシミュレーションに適したアクセラレータを制御するためのフロントエンドとして、ヒトや他の装置との間のインタフェースの役割を果たすことになると考えられる。しかし、AI 時代の演算装置としての主要な価値は、アクセラレータに存在することになる。

GPU (Graphical Processing Unit) は、もともとはその名のとおりグラフィックス (画像) の処理を行うことを目的に開発されたコンピュータである。現在では、高い並列処理性能を生かしてより広い分野の計算に適した GP-GPU (General Purpose GPU) がニューラルネットワークのシミュレーションのためのプラットフォームとして用いられている。GP-GPU は多数のプロセッサ (PE) から構成されているが、PE 当たりのメモリ量には制限があるので、密な結合を持つニューラルネットワークのシミュレーションを実行するためには演算のスケジューリングを工夫する必要があり、プログラミングが難しい。そのため、ニューラルネットワークのシミュレーション用のアクセラレータとして最適かどうかは疑問である。

これらの理由から、脳の構造と機能を模倣した計算方式であるニューロモルフィック・コンピューティングが提案されている。このような計算方式を採用することによって、ニューラルネットワークのシミュレーションが効率良くかつ少ない消費エネルギーで実行できるようになると期待されている。

S.4 低消費電力化へのアプローチ

ニューロモルフィック・コンピューティングを実現するために様々なアプローチが提案されている。下記のアプローチはそれぞれ単独もしくは組み合わせて実施されている。

A. デバイスレベルでのアプローチ

(1) 時系列型演算モデルとクロスバー回路の採用

CNN (畳込み型ニューラルネットワーク)、DNN (ディープ・ニューラル・ネットワーク) などの従来型のニューラルネットワークでは、ニューロン間のデータは実数値として伝搬される。従来型のニューラルネットワークでは隣接するニューロンの層の間では多数の実数値が並列に受け渡されて演算が行われるので、演算量も多く、消費電力も多くなる。また各層に多数のニューロンが含まれている場合には、データの伝搬経路も複雑になるという問題もある。これに対して、TrueNorth や BrainScaleS などでは、ニューラルネットワークでの演算はパルス状のスパイクの伝搬によって行われる。スパイクはクロスバー回路によって伝搬されるので、演算は簡潔であって消費電力が少ない。またデータの伝搬経路も簡潔で拡張性が高い。

(2) 不揮発性メモリ (メモリスタ) の利用

不揮発性メモリである ReRAM の技術を応用した素子 (メモリスター) を採用し、シナプスの重みを抵抗値に対応付ける手法である。これにより、乗算結果はアナログ回路の電流値

として表現できる。

B. 演算レベルでのアプローチ

(1) 演算精度の制限

多くの汎用コンピュータでは、32ビットまたは64ビット単位での数値演算を行っている。従来のニューラルネットワークのシミュレーション手法では、ネットワーク内部のデータは単精度（32ビット）または倍精度（64ビット）の浮動小数点数で表現して計算を実行していた。データのビット数が増えると、加算や乗算に必要なエネルギーも増えるので、認識結果に大きな影響が現れない程度にデータのビット数を減らすことによって計算に必要なエネルギーを削減することが可能になる。最近の研究では、半精度（16ビット）の浮動小数点数での演算を行うシステムも提案されている。同様にして、浮動小数点数の代わりに固定小数点数で演算を行うことによって消費電力を削減される方法も採用されている。この方法は、デジタル信号処理では昔から用いられていた方法である。

(2) 0スキップと枝刈り (pruning)

積和演算を実行する場合、係数の値が0の場合にはその係数に関わる乗算と加算をスキップしても認識結果には影響はない。また、係数の絶対値が他の係数の絶対値よりも非常に小さい場合にも、その係数に関わる乗算と加算をスキップしても認識結果にはほとんど影響がないと考えられる。これによって演算に必要な消費エネルギーを削減できる。

(3) 係数の圧縮 (Compression)

(ノイマン型の) コンピュータでは、演算よりもメモリアクセスによる電力消費が支配的である。ニューラルネットワークのシミュレーションで用いられる乗算の係数を、データ圧縮技術を用いてビット数を減らしてからメモリに格納し、演算の直前で伸長を行うことによってメモリアクセスの際に消費されるエネルギーを削減できる。もちろん、データの圧縮・伸長に必要なエネルギーが圧縮されたデータの読み出しに必要なエネルギーの削減量を上回らないような方法を採用する必要がある。

C. 実装レベルでのアプローチ

(1) FPGA の利用

FPGA (Field Programmable Gate Array) は、論理機能やデータの転送経路をユーザが使用現場 (Field) でプログラムできる論理デバイスである。一般にFPGAを用いることにより、汎用プロセッサ上で実行するよりも少ないエネルギーで演算を実行可能である。インテル社がFPGAメーカーのアルテラ社を買収した背景には、データセンターでの計算サーバの消費電力を低減させる方法を提供しようとしていると考えられる。

(2) 高密度不揮発性メモリの利用

先にアナログ回路向きのメモリストタ技術について言及したが、デジタル回路では、MRAM、相変化メモリ (PCM) などの次世代のメモリ技術の利用も提案されている。これらのメモリ

を論理素子と混載することにより、メモリアクセスのオーバーヘッドと消費電力を削減できる。

D. アーキテクチャ・レベルでのアプローチ

(1) マルチプロセッサ・アーキテクチャ

潜在的に演算の並列度が高いプログラムでは、多数のプロセッサを用いて演算を並列化することにより、演算に必要な実行サイクル数を削減できる。その結果、プロセッサの動作クロック周波数を下げることが可能になり、より低い電源電圧でプロセッサを駆動できることになる。一般に集積回路の消費電力は電源電圧の2乗に比例するので、電源の低電圧化によって消費エネルギーが大幅に削減できることになる。このアプローチは、AIに限らずサーバ向けの高性能コンピュータで採用されてきた手法でもある。

(2) 専用アクセラレータ (ASIC)

高性能かつ低消費電力の専用演算システムを実現する方法として、ハードウェアで演算を実行する専用アクセラレータを用いる方法は古くから採用されている。ニューロモルフィック・コンピューティングに適した専用アクセラレータの提案も行われている。専用アクセラレータのアーキテクチャが確定していれば、最先端のASIC(特定用途向き集積回路)技術を用いて実装することによって、高い演算効率が達成可能である。しかし、アーキテクチャが確定していない場合や取り扱う問題に応じてアーキテクチャ・パラメータを動的に変更したい場合などには、先に述べたFPGAを用いて実装する意味がある。

(3) 特定応用分野向きプロセッサ (ASIP)

専用エンジンと同程度の演算性能と電力性能を達成し、かつ拡張性を達成する方法として、特定応用分野向きプロセッサ (ASIP) 技術の利用が考えられる。ASIPは個別の応用ではなく、特定の応用分野に対して命令セットアーキテクチャを最適化したプロセッサである。先に述べた専用エンジンでは演算の制御はステートマシンを用いてハードウェアで実行されるので、拡張性や処理の柔軟性に限界がある。しかしASIPは与えられた命令列(プログラム)によって処理を行うので、専用エンジンと比較して拡張性や柔軟性の点で優れている。システムの拡張性と柔軟性を確保する方法としてはFPGAの利用も候補になり得るが、FPGAはASICと比較すると面積、動作周波数、消費電力の点でほぼ一桁劣っている。半導体の微細化が限界に近づき、ムーアの法則がもはや成立しなくなる時代になると、FPGAのこのデメリットはより顕著となる。これまでFPGAは最先端半導体製造技術を用いて製造されてきたからである。

本節ではAI用のコンピュータがどうあるべきかについて考察を行ってきた。汎用AIを実現するためには、脳の構造と機能を模倣したニューロモルフィック・コンピューティングを効率良く実行できるコンピュータが必要であると考えられている。しかし、従来のフォンノイマン型のアーキテクチャは、ニューロモルフィック・コンピューティングに適しておらず、大規模なシミュレーションを行うためには膨大な電力が必要となり、この点が最大のネ

ックになりつつある。この問題を解決するためには、ニューロモルフィック・コンピューティング用のアクセラレータが必要になる。

この種のアクセラレータを開発するためには、生理学、脳科学、認知科学、情報科学、半導体工学、計算機アーキテクチャなどの幅広い分野の研究者からなる研究開発プロジェクトを構成する必要がある。日本の従来型の産業に見られがちなシーズ指向のアプローチでは限界があり、研究者は自分の専門以外に周辺の科学技術の知見を獲得しておくことが必要になる。

補論の参考文献

- [Asai] 浅井哲也, 「ニューロモルフィック工学・脳型機械学習ハードウェアの行方」, 日本神経回路学会誌, Vol. 22 (2015) No. 4 p. 162-169. 2015年12月
- [BrainScaleS_1] <http://brainscales.kip.uni-heidelberg.de/index.html>
- [BrainScaleS_2] van Albada, S., Helias, M. and Diesmann, M., *Integrating brain structure and dynamics on supercomputers*, Springer Cham Heidelberg New York Dordrecht London, ISBN: 978-3-319-12083-6 (print), 978-3-319-12084-3 (electronic)
- [DaDianNao2014] Yunji Chen, Tao Luo^{1,3}, Shaoli Liu¹, Shijin Zhang¹, Lijiang He^{2,4}, Jia Wang¹, Ling Li¹, Tinashe Chen¹, Zhubei Xu¹, Jinghui Sun¹, Olivier Temam², "DaDianNao: A Machine-Learning Supercomputer", in Proceedings of the 47th IEEE/ACM International Symposium on Microarchitecture (MICRO' 14), IEEE, 2014.
- [HBP] <https://www.humanbrainproject.eu/en/silicon-brains/>
- [KAIST] Auyeung Sim; Jun-Seok Park; Mindy Kim; Dinghying Bae; Yingjie Choi; Lee-Sup Kim, "A 1.42TOPS/W deep convolutional neural network recognition processor for intelligent IoE systems", 2016 IEEE International Solid-State Circuits Conference (ISSCC), Pages: 264 - 265, Jan 2016.
- [NVIDIA] <http://www.nvidia.co.jp/page/home.html>
- [PEZY] https://ja.wikipedia.org/wiki/PEZY_Computing
- [SpiNNaker] <http://apt.cs.manchester.ac.uk/projects/SpiNNaker/>
- [Tensorflow] <https://www.tensorflow.org/>
- [TrueNorth] Paul A. Merolla, John Arthur, Rodrigo Alvarez-Icaza, Andrew S. Cassidy, Jun Sawada, Nabil Imam, Steven K. Essert, Myron D. Flicker, Dharmendra S. Mocha, "A million spiking-neuron integrated circuit with a scalable communication network and interface", Science, Vol. 345 no. 6197 pp. 668-673, Aug. 8, 2014.
- [TPU] Norm Joppa, "Google supercharges machine learning tasks with TPU custom chip", Google Cloud Platform Blog, May 18, 2016.
- [Watson] <https://www.ibm.com/watson/jp-ja/what-is-watson.html>