

戦略プロポーザル

脳型 AI アクセラレータ

～柔軟な高度情報処理と超低消費電力化の両立～

STRATEGIC PROPOSAL

Brain-Inspired AI Accelerator

- Realizing flexible advanced information processing with ultra-low power consumption -

研究開発戦略センター（CRDS）は、国の科学技術イノベーション政策に関する調査、分析、提案を中立的な立場に立つて行う公的シンクタンクの一つで、文部科学省を主務省とする国立研究開発法人科学技術振興機構（JST）に属しています。

CRDSは、科学技術分野全体像の把握（俯瞰）、社会的期待の分析、国内外の動向調査や国際比較を踏まえて、さまざまな分野の専門家や政策立案者との対話を通じて、「戦略プロポーザル」を作成します。「戦略プロポーザル」は、今後国として重点的に取り組むべき研究開発の戦略や、科学技術イノベーション政策上の重要課題についての提案をまとめたものとして、政策立案者や関連研究者へ配布し、広く公表します。

公的な科学技術研究は、個々の研究領域の振興だけでなく、それらの統合によって社会的な期待に応えることが重要です。「戦略プロポーザル」が国の政策立案に活用され、科学技術イノベーションの実現や社会的な課題の解決に寄与することを期待しています。

さらに詳細は、下記ウェブサイトをご覧ください

<https://www.jst.go.jp/crds/>

エグゼクティブサマリー

「脳型AIアクセラレータ ～柔軟な高度情報処理と超低消費電力化の両立～」とは、次世代の人工知能技術に求められる効率的な学習、外界・環境の認識・理解、予測に基づく判断などの柔軟で高度な情報処理を、人間の脳のエネルギー効率に迫る桁違いの低消費電力で行うことが可能なAI処理に特化した専用のプロセッサ（脳型AIアクセラレータ）の実現を目指すための研究開発戦略である。ここでは、脳の細胞や神経回路などの構造や機能に対する新たな情報処理モデルの構築と、これに基づくアルゴリズム、高速・超低消費電力で演算できる新たな回路・アーキテクチャ、動作に最適なデバイス・材料、これらハードウェアを効率的に動かすソフトウェアなどの研究開発を、脳科学、数学・数理科学、デバイス・材料技術と情報科学を強固に連携させて進める。これにより、脳型AIアクセラレータを実現するとともに、新たな産業の創成、人工知能研究、脳科学研究、材料・デバイス・回路研究の強化を図る。

近年、画像認識、音声認識・翻訳、自動運転、病気の診断、新材料探索、デバイス設計など様々な用途に機械学習・深層学習などによる人工知能（AI）技術が用いられるようになってきている。これらの用途には、これまでコンピュータが行ってきた単純な計算やプログラムされた演算・画像処理・制御とは質的に異なる、認識・予測・判断といった人間に近い高度な情報処理が重要になっており、AI技術のさらなる高度化が期待されている。また、AI処理は主に豊富な計算パワーを有するクラウド上で行われているが、今後は自動車やスマートフォン、IoT機器といったエッジでのAI処理が重要になってくる。現状でもAI処理を行うクラウド・サーバーでは数MWという膨大な消費電力を使っており、クラウド上でのさらなる高度なAI処理の実行、および限られた電力しか使用できないエッジでのAI処理に向けては、コンピュータハードウェアの大幅な消費電力低減が必要である。しかし、これまでコンピュータの計算パワー増大と消費電力低減の原動力となっていた半導体技術／CMOS集積回路は微細化の限界に直面しており、従来型（ノイマン型）の汎用コンピュータの高性能化・低消費電力化は困難になりつつある。これまでのAI技術はこの従来型の汎用コンピュータ上にソフトウェアにより実現されているため、飛躍的な性能向上や低消費電力化を期待するのは難しい。一方、機械学習・深層学習などは脳機能の単純化したモデルに触発され進展を遂げてきた。脳科学自身も脳の病気の原因解明などを目指して研究が進んでおり、分子レベルからニューロン・細胞、神経ネットワーク、脳全体に至る様々な階層での脳の機能が明らかになりつつある。このため、今後は脳科学の知見を活かして人間の脳の構造や機能から新たな情報処理のヒントを得る、あるいはそれらを模倣して、AI処理に特化したハードウェアを開発し、柔軟で高度な情報処理と低消費電力化の両立を目指すことが重要になる。

米国では2008年から、電子回路によって人間の脳と同じ体積、機能、消費エネルギーをもつシステムの構築を目的としたDARPAのSyNAPSEが開始され、欧州でも2013年から、脳のモデル構築に向けたHuman Brain Projectのサブプログラムの中で、ニューロモルフィックコンピュータや次世代のニューロモルフィックチップの開発が進められている。このような取り組みは脳科学の知見を高度な情報処理に活かすという注目すべき取り組みであるが、まだ性能が十分でなく、その優位性や魅力的な応用を示すまでには至っていない。我が国においても、2016年より新学術領域「人工知能と脳科学」や「脳情報動態学」、WPI「ニューロインテリジェンス国際研究機構（IRCN）」などが開始されている。また、2019年には「AI戦略2019」が策定され、必要な研究開発としてデバイス・アーキテクチャ、AIチップ技術が示されている。このように、ハードウェアによるAI技術に対する新たな流れや要求が生まれていることから、短期的にはAI処理の低消費電力化や学習の効率化の要求に応え、長期的には直感的認識やリアルタイムの判断などの要求に応え、従来技術に対する明確な優位性を示す脳型AIアクセラレータ技術の研究開発を戦略的に推進すべきである。

今後取り組むべき研究開発課題を以下に示す。脳科学における脳の構造や情報処理機能の知見から新たな数理モデルを構築し、脳型AIのアルゴリズムを創出し、それを具体的な回路・アーキテクチャ、さらにはデバイス・材料の開発に繋げていく。また同時に、脳の分子レベル・細胞レベルなど低次構造の機能に類似あ

るいはヒントを得た特性を持つデバイス・材料に注目し、そこから新たな回路・アーキテクチャ、アルゴリズムの創出を行う。この2つの流れを融合して脳型AIアクセラレータの開発に繋げていくことが重要である。

(1) 脳科学の知見を活用した情報処理の新たな数理モデル・アルゴリズムの研究開発

脳の構造や活動・機能に対する細胞レベル、神経回路レベル、脳組織（海馬など）レベル、脳ネットワーク全体など各階層での知見から、情報処理に関する新たな数理モデルとアルゴリズムを構築する。短期的には細胞（ニューロン・シナプス）と神経回路の低次レベル、長期的には脳組織と脳全体の高次レベルを対象に、構造的なつながり、ダイナミックな活動、信号伝達メカニズムなどの理解と、力学系理論、ネットワーク理論・トポロジーなどの数学・数理科学を用いて、それらに触発された情報処理の数理モデル・アルゴリズムを構築する。ここで創出するモデルは、必ずしも実際の脳の情報処理機能を正確に反映している必要はなく、新たな情報処理のアイデア、多様なモデルが出てくることが重要である。

(2) 超低消費電力で記憶・演算可能な回路・アーキテクチャ/デバイス・材料技術の開発

脳の細胞レベルの模倣に利用するデバイス・材料としては、シナプスの接続状態を変更したり、一定時間保持したりする機能が必要である。また、3次元的な接続を可能とするデバイス・材料技術、回路の接続を自由に変更する機能を持つデバイスも重要になる。回路・アーキテクチャへの研究開発としては、現在のデジタル回路よりも大幅な低消費電力化が求められる。すでに、不揮発性メモリやメモリスタを用いたニューラルネットワークの試作が行われているが、短期的にはこれらのデバイスの特性向上、デバイス特性を活かしたアナログメモリ、低電力アナログ回路、スパイクニューロン回路のスパイクタイミング制御などの技術開発が必要である。長期的には新たな脳の情報処理モデルから要求されるデバイス・材料特性の実現にむけた、新たな物理現象、ノイズ、ゆらぎなどを積極的利用、その低電力制御回路、3次元配線技術、記憶と処理の一体化に必要なインメモリコンピューティング技術などの開発に取り組む。

(3) 脳型AIアクセラレータの開発

上記の新たな脳の数理モデル・情報処理技術と超低消費電力回路・デバイス・材料技術を融合させることにより、短期的にはニューロモルフィックコンピューティングおよびリザバーコンピューティングのハードウェア技術の高度化を進め、深層学習チップに比べた優位性を示す。長期的には脳の高次機能を模した新たな数理モデル・アルゴリズムに対応したアクセラレータの設計やそれを効率的に利用するためのソフトウェアなどに関する研究開発を行う。特に、重要な応用領域と考えられるサービスロボットなどの実空間上のエッジ応用に向けて、多くのセンシング情報の中から重要な情報を選択して認識・判断し、判断した結果を複数のアクチュエータに効率的に伝えるなど、これまでのAI処理になかった機能の開発が必要である。さらに、短期的または長期的に開発する脳型AIアクセラレータの性能や潜在的な能力を社会に示すために、特定の応用を想定したフラグシップチップを開発していくことも重要である。

本研究開発を進めていくためには、魅力的な応用領域（例えばロボット）を決めて長期的な視野で我が国が先導するシナリオを描き、脳科学、数学・数理科学、情報科学、ナノテク・材料技術の異分野の研究者が日々議論できる環境の整備、研究開発拠点の構築が必要である。また、産学間の連携および海外との連携によるチップ作製や応用領域探索のエコシステムの構築、得られた知識や技術の蓄積、コア技術の知財化・国際標準化も含めたオープン・クローズ戦略の策定、などを進めていく必要がある。さらに、日本の研究者の裾野を広げるために、異分野の学会間の連携を進め、新たなコミュニティの形成や他分野に興味を持つ人材の育成が重要である。これらを後押ししていくためには、ファンディングが重要であり、短期目標と長期目標を設定し、関連府省が連携した施策の推進が望まれる。

Executive Summary

“Brain-Inspired AI Accelerator - Realizing flexible advanced information processing with ultra-low power consumption -” is a research and development strategy to realize a dedicated processor specialized for artificial intelligence (AI) processing (brain-inspired AI accelerator) that can perform flexible and advanced information processing, such as efficient learning, recognition and understanding of the external world and environment, and decision making based on predictions, which are required for next-generation artificial intelligence technology, with ultra-lower power consumption, which is almost as low as the energy efficiency of the human brain. Here, research and development of new information processing models for the structure and function of neurons and neural circuits, algorithms based on these models, new circuits and architectures capable of high-speed and ultra-low-power operation, optimal devices and materials for such an operation, and software to efficiently operate such hardware will be promoted in combination through strong collaboration with brain science, mathematics and mathematical science, device and material technology, and information science. This will enable a brain-based AI accelerator, as well as the creation of new industries, and the strengthening of artificial intelligence research, brain science research, and materials, devices, and circuits research.

In recent years, AI technologies such as machine learning and deep learning have been used in various applications such as image recognition, speech recognition and translation, automatic driving, disease diagnosis, new material exploration, and device design. These applications require advanced information processing similar to that of human beings, such as recognition, prediction, and judgment, which are qualitatively different from simple calculations, programmed operations, image processing, and control that computers have performed to date. Therefore, AI processing is expected to advance further. In addition, while AI processing is mainly done in the cloud with abundant computing power, AI processing at the edges, such as in automobiles, smartphones, and IoT devices, will become more important in the future. Currently, cloud servers that perform AI processing consume a huge amount of power (several MW). To perform more advanced AI processing in the cloud and AI processing at the edges, where limited power is available, it is necessary to drastically reduce the power consumption of computer hardware. However, it is becoming more and more difficult to achieve both such advanced AI processing and low power consumption. Semiconductor technology/CMOS integrated circuits, which have been the driving force behind the increase in computing power and the reduction in power consumption of computers, are facing the limits of microfabrication, and it is becoming difficult to improve the performance, power consumption, and efficiency of conventional (Neumann-type) general-purpose computers. Since AI technology to date has been realized by software on these conventional general-purpose computers, it is difficult to expect dramatic improvement of performance or significant reduction of power consumption. On the other hand, machine learning and deep learning have been inspired by simplified models of brain functions and have made progress. Brain science itself is also advancing in research to elucidate

the causes of brain diseases, and the functions of the brain at various levels, from the molecular level to cells (neurons), neural networks, and the entire brain, are being clarified. For this reason, it will be important in the future to use the knowledge of brain science to obtain hints for new information processing from the structure and functions of the human brain, or to mimic these hints to develop hardware (accelerators, chips) specialized for AI processing, aiming for flexible and advanced information processing and low power consumption.

In the U.S., DARPA's SyNAPSE project was launched in 2008 with the goal of building a system with the same volume, functionality, and energy consumption as the human brain using electronic circuits. In Europe, the development of neuromorphic computers and next-generation neuromorphic chips has been underway since 2013 in a sub-program of the Human Brain Project to build models of the brain. While these efforts are noteworthy to apply the findings of brain science to advanced information processing, their performance is not yet sufficient to demonstrate their superiority and attractive applications. In Japan, the Grant-in-Aid for Scientific Research on Innovative Areas, "Correspondence and Fusion of Artificial Intelligence and Brain Science" and "Brain information dynamics underlying multi-area interconnectivity and parallel processing", as well as the WPI International Research Center for Neurointelligence (IRCIN) were launched in 2016. In addition, in 2019, the AI Strategy 2019 was formulated, which indicated device architecture and AI chip technology as necessary research and development target. Thus, as new trends and demands for hardware-based AI technology emerge, we should strategically promote the research and development of brain-inspired AI accelerator technologies that meet the demands for low power consumption and high efficiency of AI processing and learning in the short term, and for intuitive recognition and real-time decision making in the long term, and that demonstrate clear advantages over conventional technologies.

Research and development issues to be addressed in the future are listed below. We will build new mathematical models based on the knowledge of brain structure and information processing functions, create algorithms for brain-type AI, and lead them to the development of specific circuits and architectures, as well as devices and materials. At the same time, we will focus on devices and materials that have characteristics similar to or inspired by the functions of lower-order structures such as the molecular and cellular levels of the brain, and create new circuits, architectures, and algorithms from them. It is important to integrate these two flows into the development of a brain-inspired AI accelerator.

(1) Research and development of new mathematical models and algorithms for information processing using the knowledge of brain science

We will develop new mathematical models and algorithms for information processing based on our knowledge of brain structure, activity, and function at the cellular level, neural circuit level, brain tissue level (e.g., hippocampus), and overall brain network level. In the short term, we will focus on the lower levels, neurons and synapses, and neural circuits, and in the long term, on the higher levels of the brain architecture and the entire brain, to understand structural

connections, dynamic activities, and signal transmission mechanisms, and to use mathematical sciences such as dynamical systems theory, network theory, and topology to construct mathematical models and algorithms for information processing inspired by them. The models created here do not necessarily have to accurately reflect the actual information processing in the brain, but it is important that new information processing ideas and various models emerge.

(2) Development of circuits/architectures and device/material technologies capable of storage and operation with ultra-low power consumption

Devices and materials to be used to mimic the cellular level of the brain need the ability to change the state of synaptic connections and hold them for a certain period of time. Device and material technologies that enable three-dimensional connections, and devices with the ability to freely change circuit connections will also be important. Research and development of circuits and architectures will require significantly lower power consumption than current digital circuits. Prototypes of neural networks using non-volatile memory and memristors have already been developed. In the short term, it is necessary to improve the characteristics of these devices, and to develop technologies for analog memories that take advantage of device characteristics, low-power analog circuits, and spike timing control for spiking neuron circuits. In the long term, to realize the device and material characteristics required by the new information processing models of the brain, we will actively utilize new physical phenomena, noise, and fluctuations, and develop low-power control circuits, three-dimensional wiring technology, and in-memory computing technology necessary for integrating memory and processing.

(3) Development of brain-inspired AI accelerator

By integrating the new brain mathematical models and information processing technologies described above with ultra-low power circuits, devices, and materials technology, we will advance the hardware technology of neuromorphic computing and reservoir computing in the short term, and show its superiority over deep learning chips. In the long term, we will conduct research and development on the design of accelerators for new mathematical models and algorithms that mimic the higher functions of the brain, and software for efficient use of these accelerators. In particular, for edge applications in real circumstances, such as service robots, which are considered to be an important application field, it is necessary to develop functions that have not been available in AI processing to date, such as selecting important information from a large amount of sensing information, recognizing and judging it, and efficiently communicating the results of the judgments to multiple actuators. Furthermore, it is important to develop flagship chips for specific applications to demonstrate to society the performance and potential capabilities of brain-inspired AI accelerators to be developed in the short or long term.

To promote this research and development, it is necessary to determine an attractive application field (e.g., robotics), draw up a scenario in which Japan will take the lead from a long-term perspective, establish an environment in which researchers in different fields of brain

science, mathematical science, information science, and nanotechnology/materials technology can discuss on a daily basis, and build a common research and development base. It is also necessary to establish an ecosystem for chip fabrication and application field exploration through collaboration between industry and academia as well as with overseas partners, accumulate the knowledge and technologies obtained, and formulate an open/closed strategy that includes intellectual property and international standardization of core technologies. Moreover, to broaden the base of researchers in Japan, it is important to promote cooperation among academic societies in different fields, form new communities, and develop human resources who are interested in other fields. Funding is important to boost these efforts, and it is desirable to set short-term and long-term goals and promote measures in cooperation with relevant government ministries.

目次

1	研究開発の内容	1
2	提案を実施する意義	4
2.1	現状認識および問題点	4
2.2	社会・経済的効果	10
2.3	科学技術上の効果	12
3	具体的な研究開発課題	16
3.1	脳科学の知見を活用した情報処理の 新たな数理モデル・アルゴリズムの研究開発	17
3.2	超低消費電力で記憶・演算可能な回路・ アーキテクチャ/デバイス・材料技術の開発	21
3.3	脳型 AI アクセラレータの開発	23
4	研究開発の推進方法および時間軸	29
4.1	脳科学、情報科学、数学・数理科学、 ナノテク・材料との密な連携	29
4.2	脳科学・数理モデルからと材料・デバイス・ 回路から攻める体制	31
4.3	技術・知識の蓄積とチップ化のエコシステムの形成	31
4.4	産業応用を見据えた新たなコミュニティの形成、 産学連携、国際連携	32
4.5	人材育成	34
4.6	時間軸	35
付録 1	検討の経緯	39
付録 2	国内外の状況	43
付録 3	専門用語解説	47

1 | 研究開発の内容

脳科学の知見を情報処理のハードウェア技術として工学的に利用することで、次世代のAI技術として求められる効率的な学習、外界の直感的な理解、予測、判断、意思決定などの柔軟で高度な機能を、低消費電力で実行できる専用のプロセッサ（脳型AIアクセラレータ）の研究開発を行う。ここでは、ニューロン・シナプスレベルや神経回路レベルでの構造・機能を模倣する「ニューロモルフィック（neuromorphic）」、「ニューロミメティック（neuromimetic）」、脳にヒントを得る「ニューロインスパイアード（neuro-inspired）」、脳の高次機能の知見を反映する「ブレインモルフィック（brainmorphic）」、などの言葉を厳格に区別せず、これらによるハードウェア（チップ）全体を総称して「脳型AIアクセラレータ」という言葉を用いている。その実現に向けて今後取り組むべき研究開発とその推進の流れの概念図を図1-1に示す。脳科学における脳の構造や情報処理機能の知見から新たな数理モデルを構築し、それを脳型AIのソフトウェア・アルゴリズム、回路・アーキテクチャ、さらにはデバイス・材料の開発に繋げていく。それと同時に、脳の分子レベル・細胞レベルなど低次の構造・機能に類似あるいはヒントを得た特性を持つデバイス・材料に注目し、そこから新たな回路・アーキテクチャ、脳型AIのソフトウェア・アルゴリズムを創出していく。これら2つの方向の研究開発が相互に影響を与え、新たな発想や連携を促進していくことで、新たな脳型AIアクセラレータの研究開発を進めていく。この中で今後取り組むべき具体的な研究開発課題としては以下の3項目になる。

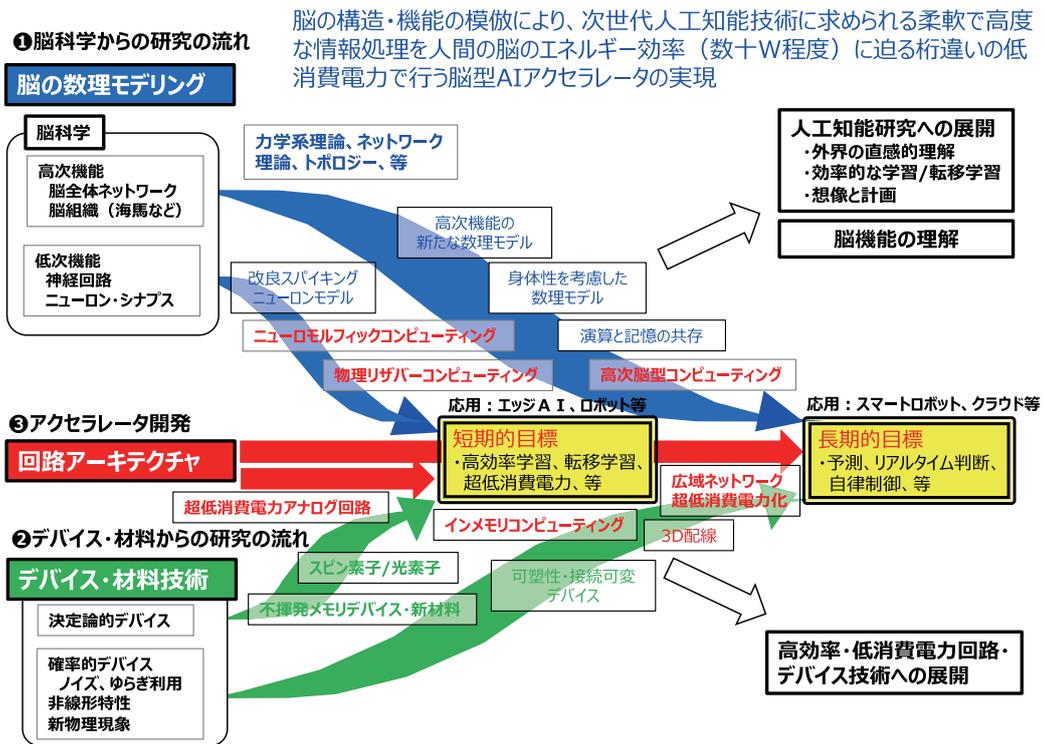


図1-1 脳型AIアクセラレータ研究開発の概念図

(1) 脳科学の知見を活用した情報処理の新たな数理モデル・アルゴリズムの研究開発

脳の情報処理機能を模倣・ヒントにした脳型AIアクセラレータの実現には、脳の情報処理の数理モデルとそれに基づくアルゴリズムの構築が重要である。脳の構造や機能には、低次レベル（細胞（ニューロン・シナプス）レベル、神経回路レベル）、高次レベル（脳組織（海馬など）レベル、脳全体ネットワーク）があるが、

短期的には高効率な学習などを目標にして脳の低次レベル、長期的には直観的理解やリアルタイムの判断などを目標に脳の高次レベルを対象に、構造的なつながり、ダイナミックな活動、信号の伝達メカニズムなどの知見に触発された情報処理の数理モデル・アルゴリズムを構築する。これには、これまでの脳科学の様々な知見を情報処理の視点で見直してみることや、最新の脳科学の研究成果について脳科学者と議論・意見交換をして理解を深めることが重要である。

脳科学の知見を数学・数理モデルで記述する例としては、ニューロンのスパイク発火を非線形の微分方程式を扱う力学系理論、脳全体のマクロな活動の時間発展を統計力学的に記述するものなどがある。脳活動のダイナミクスが数学により抽象化できると、同じような複雑系からなる様々な物理系との比較が可能となり、物理リザーコンピューティングなどの発展にも繋がる。また、トポロジー・ホモロジーの手法は脳情報ネットワーク全体の時系列活動を大域的に調べることができ、特にネットワーク情報のトポロジー的な特徴を抽出してマップ化するパーシステントホモロジーは、脳活動のネットワーク的な特徴を抽出する有益な方法であることから、その時間発展の検討ができる手法の開発が望まれる。ただし、ここで創出するモデルは、実際の脳の情報処理機能を正確に反映させているかどうかを議論する必要はなく、正確に再現するよりも多様な数理モデルが出てくることが望ましい。

(2) 超低消費電力で記憶・演算可能な回路・アーキテクチャ/デバイス・材料技術の開発

脳の細胞レベルのニューロン・シナプスを模倣した代表的なものとして、スパイクングニューラルネットワークがあり、以前よりデジタル回路を用いて研究開発が進められているが、さらなる高効率化と低消費電力化とともに、大規模化のための新たなデバイス・材料の開発、およびそれを効果的に利用できる回路技術の開発を行う。デバイス・材料技術としては、相変化メモリ（PRAM）、磁気抵抗メモリ（MRAM）、抵抗変化メモリ（ReRAM）、強誘電体メモリ（FeRAM）などの不揮発メモリやメモリスター（memristor）を用いて、シナプスの接続状態を変更したり、一定時間保持したりする機能が必要である。また、3次元的な接続を可能とするデバイス・材料技術も必要である。さらに、神経回路レベルを模倣する場合には、回路の接続を自由に変更できるような機能を持つデバイス（例えば、原子スイッチ）も重要になる。

深層学習よりも簡単な構造で時系列データを取り扱い易いものとして、リザーコンピューティングが注目されているが、その中でも非線形の物理現象を用いた物理リザーコンピューティングでは、望まれる物理特性に合ったデバイス・材料（スピンドバイス、強誘電体トランジスタ（FeFET）、フォトニックデバイスなど）の開発を行う。

さらに、新たにつくられる脳のモデルがデバイス・材料に要求する特性によっては、現状のデバイス・材料で使われている物理現象では実現できないものもあると考えられる。例えば、ニューロン・シナプスを模倣するデバイスとして、スパイク状の出力特性を持つものや、3次元的に針（導線）が伸びていくもの、結合の強さが信号で変わっていくものなどが期待される。新たな物理現象や、ノイズ、ゆらぎなどを積極的に利用したデバイスの研究も必要である。

回路・アーキテクチャへの研究開発としては、脳科学の知見から数理モデル、アルゴリズムを経て要求されるものと、新デバイス・材料側からその動作に適切なものとして要求されるものがある。いずれにしても、多少計算精度が落ちてでも現在のデジタル回路よりも低消費電力化が求められるため、超低電圧回路、アナログ不揮発メモリ回路、ストカスティック演算回路などの新たな回路技術を開発する必要がある。脳の記憶と処理の一体化という面に注目すると、現在の深層学習チップで研究が盛んになっているインメモリコンピューティング技術の進展が望まれる。

(3) 脳型AIアクセラレータの開発

上記の新たな脳の数理モデル・情報処理技術と超低消費電力回路・デバイス・材料技術を融合させることにより、脳型AIアクセラレータの開発を行う。短期的にはアクセラレータの有力な候補としてこれまで研究開発が進められてきたニューロモルフィックコンピューティングや、リザーバーコンピューティングの高度化を進めるとともに、長期的には脳の高次機能による新たな数理モデル・アルゴリズムに対応したコンピューティング（高次脳型コンピューティング）のチップ設計に関する研究開発を行う。これまでの脳の神経細胞・ネットワークのモデルに基づくニューロモルフィックコンピューティングや、リザーバーコンピューティングの高効率化・低消費電力化を進めることで、従来の深層学習チップよりも効率的な学習や低消費電力化が可能であることを示す。また、脳の高次機能を模倣したアクセラレータの設計・開発により、身体・環境の把握と運動制御、予測・判断をリアルタイムに行えることを示す。実空間・エッジ応用に向けた新たなAI技術として、センサやアクチュエータとの接続技術、応用・利用場面に従ってアクセラレータへの最適なタスクの割り当てを可能とするソフトウェア技術などの研究開発も行う。

さらに、脳型AIアクセラレータの性能や潜在的な能力を社会に対してアピールできるように、特定の応用を想定したフラグシップチップを開発して、アクセラレータ開発の進捗具合や実使用への段階が分かるようにしていく。これにより、企業や一般の人たちも目標に対する現状の技術レベルを判断し、その将来の可能性を知ることができる。

上記の研究開発を効率的に進めていくためには、長期的な視野に立ち、脳科学、数学・数理科学、情報科学、ナノテク・材料技術の異分野の連携を促進して新たなコミュニティを形成することや、魅力的な応用領域（例えばスマートロボットなど）を決めて、産業界を巻き込み我が国が世界に先んじて脳型AIアクセラレータの基盤技術開発と産業化を進められるシナリオを描くことが重要になる。そのためには、脳科学の知見を新たな情報処理技術、次世代の人工知能技術、チップ化技術に活用することを強く意識し、コンピュータやプロセッサ開発に携わった経験を持つリーダーが必要であり、その元で脳型AIアクセラレータの開発を目指す異分野連携のチーム編成が重要である。ここでは、日々議論ができる実行体制の構築、この体制の中で得られた知識や技術の蓄積、コア技術の特許化と国際標準化の戦略、海外との連携によるチップ化や応用領域の探索などのエコシステムの構築、異分野の学会間の連携と新たなコミュニティの形成、他分野に興味を持つ人材や次世代のリーダーとなる人材の育成などが必要になる。この研究開発体制を支える上では、研究開発拠点の形成や効果的なファンディングが重要である。これまで行われてきた脳型AI関係の新学術領域、世界トップレベル研究拠点プログラム（WPI）などの研究開発拠点、JSTのCREST・さきがけ「革新的コンピューティング」、NEDOプロジェクト「AIチップ・次世代コンピューティング」などの研究成果を活用したり、これらと連携したりすることのできる仕組みが必要である。研究課題、短期的目標および長期的目標を関係府省で共有し、新たな学術的成果や技術を生み出すとともに、企業の積極的な参加を促し実用化・産業化に結び付けられる府省連携による施策の推進が望まれる。

2 | 提案を実施する意義

2.1 現状認識および問題点

(1) 柔軟な高度情報処理と低消費電力化の要求

画像認識、音声認識・翻訳、自動運転、病気診断、新材料探索、デバイス設計など様々な用途に人工知能（AI）技術が用いられるようになってきた。特に、2010年ごろから機械学習・深層学習を用いたAI技術の発展が著しく、将棋や囲碁でプロ棋士を打ち負かす強さ（アルファ碁）を示したほか、自動運転、顔認証、病理画像解析、バイオインフォマティクスやマテリアルズインフォマティクスなどのデータ科学応用、ロボット実験など、様々な分野で利用されるようになってきた。これらの用途は、これまでコンピュータが行ってきた単純な計算やプログラムされた演算・画像処理・制御とは質的に異なる。今後も、機械学習・深層学習の高性能化・高効率化が求められるとともに、さらに現状のAI技術では難しい高度な情報処理への期待も高まっている。例えば、少ないデータから効率的に学習することや、数少ない例と過去の記憶から素早く新しい概念を学ぶこと、物事を一般化して他の似たタスクに転移させて実行する学習（転移学習）、風景を個別の物体に分解し互いの関係を推察するといった実世界についての直感的な理解、周りの環境の変化とこれまでの経験やシミュレーションからこれから起こることを予測すること、そこから適切な判断や柔軟な意思決定をすることなどは、まだ十分にはできていない。これらの柔軟で高度な情報処理を可能とする次世代のAI技術に関しての問題点や具体的な研究開発課題については、CRDSの戦略提言「第4世代AIの研究開発 ー深層学習と知識・記号推論の融合ー」（CRDS-FY2019-SP-08）に詳細に記載されている。

また、今後のAI技術の利用として期待されるのは、Society 5.0の実現、あるいは高度なサイバーフィジカルシステム（CPS）の構築であるが、クラウドーエッジ間の回線負荷を軽減するとともに、エッジでのリアルタイム処理を要する用途に対応するためにも、クラウド上のサイバーでの情報処理だけでなく、実空間（フィジカル空間）上のエッジでの情報処理が重要になってくる。また、クラウドに上げずにエッジ側で処理することは、健康情報など個人に紐付く情報をそのままクラウドに上げることで生じるプライバシーの問題等の負の影響を減らす上でも重要である。エッジにおいては、ロボットや自動車、IoT機器などへのAI機能の搭載が重要になってくると考えられ、上記のAI機能に加えて、個々の応用に合わせた個別学習、多種多様なセンシング情報を統合した周辺環境および身体状態の認識、周辺環境変化の予測、リアルタイムの自然言語処理、危険回避や行動目的に合わせたリアルタイムの判断、アクチュエータへの指令など身体の自律的な制御などの新たなAI機能も必要になってくる。このような実空間での利用に際しては、高度な情報処理機能の面だけでなく、限られた電力、バッテリー容量での使用といった面も考慮しておく必要があり、ソフトウェアだけでなくハードウェアも含めた検討が不可欠になる。これに関連して、戦略プロポーザル「IoT時代のセンサ融合基盤技術の構築」（CRDS-FY2019-SP-10）がセンシングの観点からエッジでのAI処理の必要性について言及している。

このような次世代AI技術の研究開発に向けては、いくつかの方法が考えられる。現在の深層学習に代表される機械学習の手法を基本にしてそこに新たな手法・技術を付加して高度化する方法、人間の脳を模倣したシステムを構築する方法がある。その他にも、量子コンピュータの利用や、自然の中で計算される過程の利用（自然計算）などの新たなコンピューティングアルゴリズムの利用なども考えられる。これらの中には、次世代AI技術に対する適合性や優位性の根拠になるものが示されていないものもあるが、今後の動向には注意しておく必要がある。

現在の機械学習の主流である深層学習においては、層の数や隠れ層の性質、各層の重みの付け方、逆伝

搬などの工夫といったパラメーターチューニングだけでなく、トランスフォーマー (Transformer)、自己教師あり学習 (Masked Modeling)、深層生成モデル (Deep Generative Model)、ニューラルODE (Neural Ordinary Differential Equations) などの革新的なモデル・手法が次々に発表されており、認識や学習の精度向上は今後も図られていくと思われる。しかし、深層学習ネットワークの根本的な変更をしない限りは効率的な学習・転移学習や外界の直感的理解、想像・予測・判断といった高度な機能の実現は容易ではないと考えられ、他の機械学習についての検討も必要である。例えば、再帰型ニューラルネットワーク (RNN: Recurrent Neural Network) の一種であるリザバーコンピューティングは、重みをランダムに固定し、出力部の重みだけを学習で決めることで、時間に依存する情報処理が可能になっている。この方式はネットワークの構造を簡素化し、非線形の様々な物理現象を利用することができ、時系列予測や音声認識などの時系列の情報処理への展開が期待されることから研究開発が活発になっており、次世代AI基盤技術の一つとして注目すべきものである。

人間の脳の単純な情報処理モデルである深層学習・機械学習の限界が見え始めている状況ではあるが、実際に高度な情報処理を行っている人間の脳の構造や機能にさらに学ぶ (模倣する) ことが一つの確実な方法である。深層学習の基礎はニューラルネットワークであり、これは脳にある膨大な数の神経細胞 (ニューロン) と結合部位 (シナプス) のつながり、および入力情報 (信号) が一定の値を超えた際のニューロンの発火によって信号を伝搬するという、脳の仕組みを基にしている。最近では実際の脳のシナプスの活動に近いモデルとして、スパイク信号の頻度やタイミングなどで表現するスパイクニューロンモデルが注目されている。そのモデルに沿ったニューロモルフィックチップが試作されるようになってきており、深層学習に迫る精度を高速・高効率で行ったことや、教師無し学習を非常に少ない回数で高精度にできたことなど、深層学習を超えるような結果も報告されるようになってきている。しかし、まだ人間の脳に比べるとニューロン数とシナプス数が3桁ほど低く、ニューロン1個あたりのシナプス数も少ないといった問題もあり、従来のコンピュータシステムに対して優位性を示すためにはさらなる大規模化、3次元的なシナプス接続の実現などが必要になる。また、このような脳の細胞レベル、神経回路レベルの研究だけでなく、大脳皮質や海馬などの脳組織、脳のネットワーク全体の知見から情報処理モデルを学んで脳全体の機能をソフトウェアで再現しようという研究も行われている。例えば、認知に関わる大脳皮質などの複数のニューロンのダイナミックな活動を計測してふるまいを抽出し、そこから学習のモデル、意思決定のモデルが構築されている。このように、次世代AI技術に求められるAI技術の高度化に向けては、脳の構造や機能を模倣あるいはヒントを得て新たな情報処理モデルを構築していくという流れに注目しておくことは非常に重要である。

一方で、現在の深層学習・機械学習を基本とするAI技術においては、消費電力の問題も顕在化しており、処理の効率化・低消費電力化の要求が高まっている。現在のAI技術は主に高性能サーバー上でソフトウェアにより深層学習 (畳み込みニューラルネットワーク (CNN: Convolutional Neural Network) など) を行うことで実現されている。膨大な数のデータを学習させて、画像認識などの処理を行っており、膨大な電力を消費している。例えば、2018年から稼働した人工知能技術開発のための高速計算基盤「ABCI (AI Bridging Cloud Infrastructure)」は積和演算を高速に行うGPUを大量に実装している世界最大級の超省電力のAIインフラであるが、それでも最大2.3MWの電力を消費している。今後の、AI応用の大幅な拡大のためには、システムの拡張を図る必要があるが、冷却システムの上限を考えると、全消費電力の大幅な増加は困難と考えられる。一方、詳細は後で述べるが、現在の情報処理性能向上・低消費電力化の担い手である半導体集積回路はすでに加工寸法が10nm程度になってきて微細化の限界に直面しており、従来のCMOSデジタル回路の微細化・高集積化によるプロセッサの大幅な高性能化・低消費電力化は難しくなっている。

このため、従来のノイマン型のプロセッサだけでなく、深層学習用のプロセッサも含めて、高度な次世代のAIの機能を低消費電力で実行できる専用のプロセッサ（アクセラレータ）が必要であり、アルゴリズムや回路アーキテクチャ、デバイス・材料などの技術レイヤーに跨る新たなコンピューティング技術の研究開発が重要になってきている。

最近では、TPU（Tensor Processing Unit）など深層学習専用のプロセッサを開発して、処理速度の向上と消費電力削減が行われるようになってきたが、大幅な消費電力低減には至っていない。深層学習用のプロセッサの高速化・低消費電力化に関して、研究開発レベルでは、高い精度で演算が可能な32ビットや64ビットのデジタル回路の代わりに、4ビット、2ビットなど演算精度を落としたデジタルCNN回路や、アナログ回路を使ったもの、様々な物理現象を利用したリザバーコンピューティングなどが報告されるようになって、低消費電力化の研究開発が活発になってきている。このような研究開発において、従来のデジタル技術だけ、あるいはアナログ回路との併用による低電力化には限界も見えてきており、不揮発メモリなど不揮発機能を持ったデバイス・材料の積極的な利用や、回路を大幅に単純化・低電力化することが可能な物理リザバーコンピューティングなどの研究開発が重要になってきている。

このような低消費電力化の要求は、実空間上のエッジでのAI処理にとっては、さらに重要である。限られた使用電力、バッテリー容量の中でのAI処理は、クラウド・サーバーで利用されているCPU+GPUベースのものよりも数桁の消費電力低減が必要である。現状での深層学習の利用を考えた時でも、計算負荷の大きな処理を専用ハードウェアであるアクセラレータで実行することが必要であり、次世代の脳を模倣したAI技術においても効率的な演算を実行するためには応用領域を特定した専用のアクセラレータという形態が必須になる。

以上述べたように、次世代のAI技術としては、効率的な学習、外界の直感的な理解、予測、判断などの柔軟で高度な情報処理と低消費電力化の両立が必要になってきており、脳の構造や機能を模倣したり、演算のヒントを得たりすることで、新たな情報処理のモデルの構築、そこからの新たなアルゴリズム・ソフトウェアの研究とともに、計算負荷の大きなところを専用のハードウェアで実行するアクセラレータの開発を進めていくことが重要になってきている。特に、短期的な視点では、現在クラウドやサーバー上のソフトウェアで実行している深層学習などによるAI処理の低消費電力化や学習の高効率化、エッジにおける効率的な学習と推論の桁違いの低消費電力化の問題に対処するように、スパイクニューロンモデルに基づくニューロモルフィックコンピューティングや物理リザバーコンピューティングなどの研究開発を加速して、脳型AIアクセラレータを実現していくことが必要になっている。また長期的な視点では、AI技術としてまだ実現されていない外界の直感的認識や予測、判断などの課題に対して、脳の組織や全体のネットワークのダイナミックな活動の高精度計測により高次機能の新たな知見の獲得と理解を深め、例えばゆらぎ学習のような新たな情報処理モデルを構築し、それに対応する回路・アーキテクチャ、デバイス・材料を開発し、さらに柔軟で高度な情報処理が可能な新たな脳型AIアクセラレータへと繋げていくことが重要である。

(2) 脳科学の進展

深層学習をはじめAI技術は、人間の知能の特性や振る舞いから学ぶことで発展してきた。脳の仕組みの解明を目指す脳科学は、生物学的に脳を理解しようとする実験的研究と、脳のアルゴリズムやモデルを研究する理論的研究に分かれている。実験的研究は、観測する構造や現象のスケールが分子レベルから個体レベルまで多くの階層に渡るといふ大きな特徴がある。近年の計測技術の進歩とともに、脳の分子・細胞的理解が進み、さらに理論研究と実験的研究が相互に知見をやりとりすることによって、脳科学研究は大きく進展し、

分子レベルからニューロン・細胞、神経ネットワーク、脳全体に至る様々な階層での脳の機能が明らかになりつつある。1980-90年代に日本で行われた、実験科学に基づく小脳の比較的単純な神経回路についてのフィードバック制御理論研究は実験・理論脳科学の融合研究の嚆矢のひとつだったが、深層学習の時代になって、視覚情報処理や物体認識などの様々な脳の高次機能を模した深層学習が提唱されている。2016年に脳の活動と畳み込みニューラルネットワークを対応づける研究が、2020年には自然言語処理に対応した脳の活動の研究が発表された。[出典：調査報告書, ドライ・ウェット脳科学, 2020. <https://www.jst.go.jp/crds/pdf/2019/RR/CRDS-FY2019-RR-06.pdf>]

神経細胞内部の構造を使った計算原理・デバイスの可能性が示唆されている。神経細胞の樹状突起内の情報処理から得られた理論モデルを使い、主成分分析や独立成分分析などが分散的に効率よく行えることが理論的に示された。ニューロンへの入力に単純に係数を掛けて足し合わせたものが閾値を超えるとシグナルを出力するというのが深層学習で使われているニューロンモデルであるが、実際のニューロンはもっと複雑な情報処理をしている。樹状突起から入力情報が細胞体に集約した結果、細胞体で活動電位が発生すると、その電位が樹状突起側へ逆方向に情報が戻され（バックプロパゲーション）、樹状突起の興奮状態を修飾することがわかっている。これを理論に組み込んで、樹状突起の枝ごとに行われる局所的な計算の結果として最終的なニューロンの出力が決定され、さらにニューロンの中で出力側から入力側へのアナログのフィードバックも大きな影響をもつとするモデルが提唱されている。すなわち、一つのニューロンの中でリカレント的な振る舞いがある。こうした個々のニューロンレベルの研究により、より脳の情報処理に近い効率のよい素子開発が有効ではないかと考えられている。

脳の細胞レベル、局所神経レベルの活動の計測に関する具体的な研究としては、二光子顕微鏡技術、超小型内視鏡システム、蛍光プローブ技術などを用いて、生きた脳深部の神経活動を細胞レベルで読み出すものや、行動中の動物の脳深部から長期間の活動記録を行うことで、空間・時間的な活動を詳細に把握するものがある。また、シナプス結合構造の全貌を電子顕微鏡レベルで網羅的に読み出す「コネクトーム」解析も注目される。

脳の局所神経回路に目を向ければ、例えば大脳皮質には抑制性のニューロンが多数存在し、興奮性ニューロンに対し、入力側である樹状突起を抑制するニューロンと、出力側の軸索起始部を抑制する異なる種類の抑制ニューロンが存在している。この解剖学的知見を基に、理論家のジャオ・ジン・ワン（ニューヨーク大学）は、それぞれの抑制ニューロンが局所回路への入力のセクター、あるいは出力のゲートとして働いているということ提唱している。こうした入出力のゲーティングは、脳の回路構造がフレキシブルに切れたりつながったりする処理の一端を担っていると考える研究者もいる。このような脳の局所回路の知見はデバイス設計に有益と考えられる。

このように、脳の細胞レベルと局所神経回路レベルでは部分的ではあるが新たな知見とその情報処理モデルが提案されてきており、短期的な視点としてはこれら脳の低次機能の研究をさらに進展させていき、これまでの情報処理モデルの改良・精緻化を図っていく必要がある。また、長期的な視点では海馬などの脳組織レベルや脳全体のネットワークなどの高次機能についての理解を深めていくことも重要である。磁気共鳴機能画像法（fMRI：functional Magnetic Resonance Imaging）による特定の脳組織や脳全体の活動の計測、計測データのAIによる解析などにより、脳の高次機能の理解を進める必要がある。将来に起こることの予測や適切な判断を行う場合には、過去の経験、社会倫理、社会通念、歴史、現在の環境、立場など様々な要素の情報を集めて、総合的に情報処理していくものと考えられ、長期的な研究が必要になると思われるが、脳の高次機能を参考にした情報処理モデルの構築が望まれる。

(3) 集積回路の微細化の限界

ムーアの法則をベースにした集積回路の微細化の開発速度は数十nmノード（実際の最小パターン寸法ではなく、プロセス技術の世代を表す数値）といわれる頃からそのスケーリング速度が減速してきた。構造パターンの露光に用いるArFエキシマレーザ波長が193nmであるので波長より微細な構造を実現するにはマスクパターンに各種の工夫が必要になり、一層分の構造パターンを形成するのに多重に露光するなど、プロセスが複雑化し製造コストも増加している。このように技術的には難しくなっているが、2018年には7nm、2020年には5nmノードの量産が始まっている。最近では波長が13.5nmのEUV（Extreme ultraviolet）リソグラフィ装置が量産に使われるようになり、今後もEUVの改良や素子構造の開発が進み3nm、2nmなどより微細化が進むと思われる。素子構造に関しても現在の微細プロセスに使われているFinFET（Fin Field-Effect Transistor）では性能向上が困難になってきており、GAA（Gate All Around）トランジスタやCFET（complementary FET）など新しい構造の研究が行われている。以前は微細化に伴い素子あたりのコストも低下し、高集積化と同時にコストパフォーマンスも向上していたが、最近ではプロセスや素子構造の複雑化などが原因となり微細化を進めることでコストは逆に上昇する傾向が認められる。そのため最先端CMOSプロセスが利用できるのは真に大量生産が期待できる一部のアプリケーションに限定されるようになってきた。また、高集積化、高速クロック化に伴いチップの消費電力が増加し発熱問題なども顕在化している。

エッジ用LSIに関しては、少量多品種製品の市場が大きいことや、大がかりな冷却機能が使えないことなどから、上記微細化の流れとは異なる戦略が求められる。その戦略の方向性は高速化、小型化、低消費電力化、低コスト化である。そのため、従来型（ノイマン型）ではないアプローチが必要になり、世界的に新たなコンピューティング技術を探る動きが盛んになっている。フォン・ノイマン・ボトルネック問題が顕在化してきた2009年には、IEEEにRebooting Computing Initiativeが設立されており、新たなコンピューティング技術を模索する動きが加速している。その一つの方向として脳科学の知見を生かした人間の脳の構造や機能を模倣し、柔軟で高度な情報処理を低消費電力で行うハードウェア技術の開発があり、非ノイマン型デジタルコンピューティング、量子コンピュータとともに世界的にも注目されている。日本においても、これまでAIに関する研究は主に情報処理学会、人工知能学会で行われてきたが、2019年にはデバイス・材料分野をカバーする応用物理学会でもAI応用を目指すフォーカストセッション「AIエレクトロニクス」が創設され、デバイス・材料のAI応用に関する活動が活発化している。

(4) 脳の構造・機能を模倣した新たなコンピューティング技術への期待と新たな動き

脳の構造・機能を模倣した新たなコンピューティング技術は、世界各国で注目が集まっている研究分野の一つであり、約10年以上前から各種プロジェクトが進められている。例えば米国国防高等研究計画局（DARPA：Defense Advanced Research Projects Agency）によるSyNAPSEプログラムや欧州のHuman Brain Projectが有名である。SyNAPSEでは、米国の大学を中心に数十の研究機関が参加し、ニューロモルフィックチップの開発やそのアプリケーションへの実用化の活動を行っている。2014年には、IBMが開発したデジタル型のニューロモルフィックチップTrueNorthが発表され話題となった。Human Brain Project（2013年～）は、EUが助成する脳研究の大型研究支援プログラムであり、欧州内外を含めて100を超える大学・研究機関が参加し、神経科学、コンピューティング、脳疾患などの融合的な研究開発が行われている [出典：国立研究開発法人情報通信研究機構. 欧州における脳情報関連技術の研究開発動向, 2018. https://www.nict.go.jp/global/lde9n2000000bmum-att/re201803_1.pdf]. このような欧米の取り組みは脳科学の知見を高度な情報処理に活かすという意味で注目すべきものであるが、ニューロン・

シナプスレベル、神経回路レベルの脳の低次機能のモデルについての取り組みであり、また人間の脳に比べると桁違いにニューロン数・シナプス数が少なく、その優位性や魅力的な応用を示すまでには至っていない。短期的にはこのような研究を推進し、ニューロン数・シナプス数を増加させて、魅力的な応用分野を探していくことで良いと思われるが、長期的には直観的認識やリアルタイムの判断ができるように、脳の高次機能に基づくモデルの構築とそれを実現する回路・アーキテクチャ、デバイス・材料の研究開発を進めることが望まれる。

日本では、脳科学の知見を情報処理に活かしていこうという取り組みとして、2016年から新学術領域「人工知能と脳科学：人工知能と脳科学の対照と融合」、「脳情報動態学：脳情報動態を規定する多領域連関と並列処理」、WPI「ニューロインテリジェンス国際研究機構（IRCIN）」などが開始されている。2つの新学術領域は脳型AIのハードウェア化までを目的としたものではないが、脳の情報処理ダイナミクスの可視化や情報処理機能の理解を深めるという意味で注目される。このWPIは脳機能の理解が第1目的としてあるが、脳模倣のニューロモルフィックチップの開発も視野に入れており、脳科学、情報科学、数学・数理科学、回路・アーキテクチャなどの異分野の研究者を集めた研究拠点としても注目される。また、2018年から深層学習およびデジタルのアーキテクチャ研究を中心に新たなコンピューティング技術の開発を目指す、内閣府SIP第2期「フィジカル空間デジタルデータ処理基盤」、NEDOのプロジェクト「高効率・高速処理を可能とするAIチップ・次世代コンピューティングの技術開発」、JSTのCREST「Society5.0を支える革新的コンピューティング技術」、さきがけ「革新的コンピューティング技術の開拓」などが開始されている。これらのプロジェクトの中の一部に、脳型のコンピューティングに関わるテーマが含まれている。

このように脳型AIのハードウェアに関する期待や活動が活発になっているが、脳に学んだ新たなAI用のアクセラレータがすぐに実用化されるような状況にはまだない。実際に試作されたニューロモルフィックチップで、人工ニューロン・シナプスの大規模集積がまだできておらず、その優位性が示せていないことや、脳に学んだ情報処理モデルで高度な情報処理と低消費電力化を示せていない、脳のニューロン間の接続のように三次元的で複雑な接続を実現する方法が見つからないことなど、まだ多くの科学的・技術的課題がある。脳型AIハードウェアの研究で先行している欧米でもまだ模索している段階である。しかし、脳を模倣したアクセラレータの研究開発は正しい方向性と考えられる。現在のノイマン型のコンピュータは汎用ではあるが、半導体集積回路の微細化限界のために飛躍的な性能向上は期待できず、特定の機能に特化して高性能化を図るドメインスペシフィックなコンピュータが必要になっている。その代表例が量子コンピュータであり、暗号解読、最適化問題など特定の情報処理については計算論的にその優越性が裏付けされており、その研究開発が世界的に活発化している。一方、脳型のコンピュータについては脳の機能が解明されていないので計算論的な裏付けはないが、人間の脳が数十Wの電力で、学習、認識、判断といった高度な情報処理を行っているという事実があり、将来の性能・応用を予測する上でこれほど確実なものはない。それを反映して、2020年の固体素子回路国際学会（ISSCC2020）におけるIBMの基調講演で、今後のコンピューティングではビット（デジタル）とニューロンと量子ビットが重要になるとしており、また国際電子デバイス会議（IEDM2020）のサムソン電子の基調講演で、今後は機械学習、ニューロモルフィックコンピューティング、量子コンピュータが重要になるとしており、脳型の情報処理には産業界も注目している。

我が国の研究リソースは限られているので、日本にとっての脳型AIアクセラレータ応用として重要な産業領域を見据えて、科学技術的にも強い研究領域に集中するなどの戦略的な取り組みも必要である。Society 5.0に代表されるサイバーフィジカルシステム（CPS）では、実世界上のエッジ側での技術が重要になると考えられ、日本が科学技術的にも産業的にも強いセンサ技術、ロボット技術、デバイス・材料技術を活用した戦略が求められる。脳科学、数学・数理科学においては海外がリードしている状況にあるが、日本にもこ

これらの分野でトップクラス人材が存在しており、上記の日本の強い技術と情報科学を密に連携させ、独自の脳の情報処理モデルの構築、高効率デバイス・回路、脳型AIアクセラレータを開発し、実世界AI技術の強化を図ることが期待される。取り組みの方向性としては、(1)で挙げた短期的および長期的な視点でのAI技術の課題に対して明確な目標を定めて取り組んでいくことだと考えられる。つまり、短期的には脳の細胞・神経回路レベルのモデルを用いて、現在ソフトウェアで実行しているAI処理の低消費電力化や学習の高効率化の課題に取り組み、長期的には脳の高次の機能についての理解を深め、直感的認識やリアルタイムの判断などの課題に対処していくことが望まれる。

2019年には国の戦略として「AI戦略2019 ～人・産業・地域・政府全てにAI～」が策定され、AI戦略2019における研究開発戦略として、世界をリードする次世代AI基盤技術の確立を目的に、「基礎理論」、「高品質かつ信頼できるAI」、「AIのシステム・コンポーネント」とともに、「AIのためのデバイス・アーキテクチャ」の研究開発の必要性が記載されており、脳を模倣した情報処理への期待も高い。先の3つに対しては、CRDSから「第4世代AIの研究開発 一深層学習と知識・記号推論の融合一」(CRDS-FY2019-SP-08)、「AI応用システムの安全性・信頼性を確保する新世代ソフトウェア工学の確立」(CRDS-FY2018-SP-03)、「複雑社会における意思決定・合意形成を支える情報科学技術」(CRDS-FY2017-SP-03)などの提言がすでに出されており、これらのソフトウェア系の提言に加えて、それと相補的となるハードウェアに係る研究開発の提言をまとめて、日本のAI戦略全体を先導することが望まれる。

2.2 社会・経済的効果

脳型AIアクセラレータが開発・実用化されることで、汎用プロセッサと連携して高度なAI処理を高速・低消費電力で行えるようになり、現在のAI技術を使ったクラウドを中心にした応用分野の利用拡大だけでなく、実空間上のエッジでの新たな多くの応用領域に展開される。例えば、以下に示すような少子高齢化による労働力不足の問題、高齢者の介護問題、安全な移動、医療費削減に向けた健康の維持、効率的な社会インフラの維持管理などの社会的問題の改善・解決に貢献すると考えられる。もちろん、このような社会的問題の改善・解決には脳型AIアクセラレータだけでできるわけではなく、システム全体のデザインやソフトウェア技術、規制緩和などが必要であり、これらの関係者と連携した取り組みが重要である。

- 人間と同じように細かな部品・物体の認識、状態の判断、作業計画策定ができるロボットが実用的に使われるようになり、工場や農作業などで人間と一緒に作業することができるようになり、労働力不足の緩和・解消に繋がる。
- 知的なロボットが家庭にも入っていくようになり、高齢者、障がい者などが介護の補助や、他者の助けなしに自立した生活を送ることができる。家庭用としては空間移動支援や清掃支援、あるいは入浴支援や排泄支援まで、あらゆる生活支援をするロボットの進展が期待される。
- 車だけでなくパーソナルモビリティなどの自動運転が可能になることで、高齢者、障がい者も安全にどこへでも行けるようになる。

- ウェアラブル機器へのAI機能の導入で、個人の日々の各種データ（体温、心拍数、呼吸、呼気など）の取得、変化の簡易分析から、適切な生活のアドバイス、疾病の早期発見、体調、病状の変化の把握が可能になり、生活上の安心感が増加する。また、個別化医療の進展、健康寿命延伸にも貢献する。
- 橋梁、トンネル等の社会インフラの重要な箇所に自律的な動作をするスマートセンサを配置し、状態の変化や異常を検知して必要な情報だけをクラウド側に送ることにより、センサ端末の長期的な稼働を可能とし、社会インフラの効率的な保守管理に貢献する。

また、一つの応用領域で脳型AIアクセラレータの有効性が示されれば、他の応用領域にも急速に広がると思われる。例えば、産業用ロボット応用において、硬い材質だけでなく柔らかい材質のもの、変形したものでも安全に取り扱うことができ、人との安全な協働作業ができるようなアクセラレータが開発できれば、家庭内で働くようなサービスロボット、介護ロボットにも波及していくと考えられる。

さらに、これらの社会的課題への貢献は、経済的な貢献にもつながる。特に、ロボット産業や自動車産業は日本の強い産業分野であり、インテリジェント化によりこれらのさらなる強化も期待される。また、脳型AIアクセラレータを日本が世界に先駆けて開発・製品化できれば、半導体集積回路の新たな価値を生み出し、日本の半導体産業の再興にもつながっていくと期待される。

コンピュータに求められる機能が高度化、複雑化する中、新たなコンピュータへの期待は高く、経済的な効果として、以下に示す市場規模の大きな幅広い分野での貢献が予測される。（括弧の中の金額はその年における世界市場規模）

- ロボット（約21兆円@2020年、出展：IDC）
 - ・産業用ロボット（人との安全な協働作業、柔らかな材質のもの扱い）
 - ・サービスロボット（人との安全・安心な共生、環境変化の判断と的確な動作、自然言語処理）
 - ・家庭用サービスロボットについては2025年に約2.9兆円の市場予測（出展：富士経済）
- 自動運転機能付き自動車（約9.1兆円@2025年、出展：ReportBuyer）
 - ・事故を防ぐ自動運転機能（車・人・環境の認識、リアルタイムの判断、最適な操作）
- 健康・医療ウェアラブルデバイス（約2.3兆円@2020年、出展：IHS Technology）
 - ・健康管理（日々の各種データの取得・分析、生活のアドバイス）
 - ・診断データの分析（病気の原因の予測）
- 社会インフラ保守（国内社会インフラIT市場：約0.6兆円@2024年、出展：矢野経済研究所）
 - ・社会インフラ状態把握（センサ端末からの各種データの取得、現状分析、劣化予測）
 - ・修理・建て替えの順番最適化（組み合わせ最適化問題）
- ホームセキュリティ（約5.7兆円@2022年、出展：Lucintel）
 - ・監視カメラ、センサの高機能化（画像認識、異常の検知、イベント予測）
- ニューロモルフィックチップ（約1.2兆円@2026年、出展：Future Market Insights）

- ・認識、判断、組み合わせ最適化等の柔軟で高度な情報処理の超低電力・高速処理

2.3 科学技術上の効果

(1) 人工知能研究への展開

現在のソフトウェアベースの人工知能処理の負荷の重い一部の処理を、高速・低消費電力で実行できる脳型AIアクセラレータが開発されれば、人工知能技術開発の基盤となるABCIのような計算センターの高性能化・低消費電力化に貢献できると思われる。これにより、多くのAI処理が実行できるようになり、人工知能の研究が促進されると期待される。

また、脳型AIアクセラレータの研究開発を進める中で、脳型情報処理の新たな数理モデルが提案され、そのアルゴリズムが開発できれば、ハードウェア化する前にソフトウェアとして人工知能技術に取り込むことができ、その有用性や問題点を調べることで、人工知能研究にも刺激を与えられられる。

さらに、深層学習の限界や人工知能の残された問題（①学習に大量の教師データや計算資源が必要なこと、②学習範囲外の状況に弱く、実世界状況に臨機応変な対応が出来ないこと、③パターン処理は強いが、意味理解・説明等の高次処理は出来ていないことなど）に対して、脳型AIアクセラレータの研究開発を進める過程において、これまでは人工知能研究とは直接的な関係のない脳科学者や、回路・アーキテクチャおよびデバイス・材料の研究者と人工知能の研究者が直接的に話すことにより、脳科学や回路・デバイスの研究者の人工知能研究への視野を広げて積極的貢献を促すことができ、人工知能の残された問題などに対する新たな発想に結びつくことも期待される。

(2) BMI 技術や脳機能理解の進展

脳型AIアクセラレータに関する研究開発において、神経回路の理解や新たな情報処理モデルの創出、ニューロモルフィックコンピューティングのチップ開発などが進むことにより、今後のBrain-Machine Interface (BMI) の基盤技術になることが期待される。BMIは脳の情報をリアルタイムに読み取りモデル化することで、脳と機械を相互作用させ、心身機能の補綴や改善、拡張の実現を志向する新しい脳科学の分野である。脳の運動情報処理に媒介してロボットアームやロボットレッグを円滑に操作するサイボーグ技術や脳内の情報処理過程に仲介してその演算を効率化させるニューロモジュレーション技術、失われた神経ネットワークを再建する人工神経接続技術など、生物器官の機能不全や欠損に対する工学的なソリューションとして実用化が進んでいる。また、四肢麻痺の人が補助具を使い歩くことができるようにするリハビリ技術において、神経回路を理解することにより、神経回路の再活性化のプログラムをAI駆動神経刺激デバイスを介して動かす、ということが実証されつつある。サービスロボットなどへの応用に向けた脳型AIアクセラレータは、多くのセンサ情報から自分や環境の状況を認識し、予測に基づいた判断によりアクチュエータに適切な指示を出す機能も備えられられるため、このようなBMI技術の発展に寄与すると期待される。

さらにこのような直接的効果だけでなく、間接的な効果として脳機能の理解の進展にも繋がると考えられる。生命科学はビッグデータ時代に入り、これまでの仮説検証型研究からモデルフリー、すなわちデータ駆動型による現象の理解へとパラダイムがシフトしつつある。脳科学においても、生物学的な脳の「完全な」理解はヒトの認知能力では不可能ではないか、という考え方が根強いが、脳計測から得られるビッグデータをそのまま

モデルフリーで高次元のパラメタにフィッティングすることで、脳の理解が進むのではないか、という意見も出てきている["direct fit for nature", Hasson et al., Neuron 105:416-434 (2020)]。脳型AIアクセラレータの研究開発において、ダイナミックな脳活動の空間および時間的に高精度な計測を行い、解析用の大量のデータを取得することになるので、これらの取得データは高次元パラメタへのフィッティングに適用でき、脳機能の理解に貢献できると考えられる。また、脳型AIアクセラレータの研究開発の中で、神経回路や脳組織レベルの情報処理モデルの構築が進むことにより、DeepMind社（米国）、Open AI（米国）、NPO法人「全脳アーキテクチャ・イニシアティブ」（日本）などが目指している汎用AIの実現を加速することが考えられる。

(3) 数学・数理科学の新たな展開

脳科学からの知見は、圏論などの数学基礎論から、トポロジー・ホモロジー、情報理論・情報幾何学、力学系理論・カオス理論、確率・統計理論など広範な数学・数理領域の研究者に刺激を与えて以下のような様々な成果が得られている。

AIおよびニューロコンピューティングは、ここ50年の間に2度のブームがあり、現在3度目のブームが起こっているが、過去の2回と今回のブームは、脳科学、数学・数理科学とコンピュータ科学の間の緊密な連携の元に研究が進んでいる点で大きく異なっている。今回のブームの起爆剤となったのは、深層学習を用いたニューラルネットワークの飛躍的な発展である。深層学習の今日の隆盛の基盤になったのは、甘利俊一博士の確率降下学習法を用いた逆誤差伝搬法であり、ここでは数理モデルが重要な役割を果たしている。しかし、現在の深層学習は静的な数理モデルを用いたアルゴリズムを用いており、実際の脳のダイナミックな情報処理機構を十分反映しているとは言えない。例えば、脳は安静状態でもダイナミックに活動しているという特徴があり、これがノイズに強い要因とされているが、現状の深層学習は敵対的なノイズ攻撃に弱い。したがって、実際の脳の構造や活動に基づいた新しい数学・数理モデルの発展とそれに基づくアルゴリズムの開発は、コンピュータ科学の発展に大きなインパクトを持つと考えられる。

脳の活動解析をめざした数学モデルは、ニューロンのスパイク発火を記述する非線形の微分方程式群を中心とした力学系理論である。歴史的にはノーベル生理学・医学賞を受賞したHodgkin - Huxley方程式 [A. L. HODGKIN and A. F. HUXLEY: A quantitative description of membrane current and its application to conduction and excitation in nerve: The Journal of physiology, 117(4), 500-44 (1952).] から始まった分野である。その後、積分発火モデルなどより簡易なモデルや、発火頻度に着目したモデルなど様々なモデルが提案されており、さらにアナログ量で表現したモデル、しきい値関数、シグモイド関数など活性化関数のモデルも提案され、計算コストと生物学的妥当性を両立する数学モデルが研究されている。また、古典力学の粒子の集団運動は、運動粒子に統計的な性質を導入することによって統計力学という分野を切り開き、物性や相転移など多くの物理分野で成功を収めてきた。同様に脳の神経回路の時系列ダイナミクスに統計性を導入すると、脳全体のマクロな状態変化を統計力学的に扱うことができる（統計神経力学）。これをさらに発展させて、近年ではカオス理論を応用して、脳活動領域のクラスタリングのメカニズムが解明されるなど、シナプス発火のメカニズムを時間依存の微分方程式で記述・解析するのみならず、脳全体の活動をミクロなシナプス結合と結びつける新たな数学・数理モデルが発展しつつある。

一方、脳全体の情報ネットワークに着目する数学モデルは、ネットワーク理論やトポロジー・ホモロジー理論である。前者は、主に検索問題を効率的に解くことを目的として発展してきたが、最近では脳のモジュール構造に着目した理論の展開が進みつつある。後者は、例えば脳の神経系の発火データから脳全体のトポロジーを解析するアプローチであり、情報ネットワーク全体の時系列活動を大域的に研究するモデルである。こ

のモデルの解析から脳情報処理の堅牢性や信頼性の情報が得られると期待され、脳型AIの特徴を具体的なアルゴリズムやハードウェアに実装するときの活用も期待される。

この他にも、脳神経回路の環境適応について情報理論を用いてモデリングするなど、情報理論や情報幾何学の最先端との連携も進んでいる。さらに、ノイマン型のコンピュータのベースとなっている集合論を超えて、意識の因果的な側面や情報構造を圏論でとらえ直し、関係性のネットワークの更新による意識発現の研究 [土谷尚嗣、西郷甲矢人：圏論による意識の理解：Cognitive Studies, 26, 4, 462-179 (2019)] など、新たな数学分野からの脳科学、あるいは脳型AIの研究が進展しつつある。また、情報理論に計量を導入し情報多様体として取り扱う情報幾何学は、まず多層ニューラルネット関数空間上の多様体に幾何学構造を導入した機械学習への応用が目ざされてきたが、さらに脳の広範囲の情報統合によって意識が生じるという観点で意識を研究する情報統合理論への発展が期待される。この統合情報理論で意識をとらえ直す方向では、意識の因果的な側面や情報構造を圏論でとらえ直し、関係性のネットワークの更新による意識発現を研究するアプローチもされている。

今後も、脳型AIアクセラレータの開発を目指して、脳の構造・機能の新たな知見の獲得とその情報処理モデルを創出していくことで、数学・数理科学分野の一層の学術的な発展が期待できる。

(4) 高効率・低消費電力デバイス・回路技術への展開

脳が進化の過程で作りに上げた特異的な機能を物理的に実装するためには、デバイスの物理的・動力的性質を直接的に利用して、ニューロンやシナプスの機能をデバイス・材料レベルで模倣し再現することが求められる。脳型AIアクセラレータは、組み合わせ回路に基づく従来法とは全く異なる、新しい計算パラダイムとなるため、従来のCMOSデバイスを超える新動作原理のデバイスの研究開発を促進すると期待される。すでにデバイス・材料の研究者が多く集まる応用物理学会においては、2019年から「AIエレクトロニクス」という新たなセッションが設けられて活発な活動をしており、今後は回路・アーキテクチャ、数学・数理科学、アルゴリズム・ソフトウェア、情報処理、人工知能、脳科学などの分野の研究者も参加した大きなコミュニティを形成していくことが期待される。これにより、分野連携や融合が進み、デバイス・材料技術をコアにして研究領域が拡大していくものと考えられる。また、他の研究分野を含むデバイス・材料の技術シーズから、情報処理、人工知能、脳科学を含む様々な応用ニーズを含む形で研究開発が進められていくことにより、企業を含んだ形でデバイス・材料の研究開発が活性化し、材料技術、デバイス技術で強みを発揮してきた我が国が再びこの分野で主導権を握り、産業的な強化につながることを期待される。

具体的には、以下のようなデバイス・回路技術への展開が考えられる。シナプスは伝達効率を連続的に変化させることができ、それを長期記憶する機能を有するが、このシナプスの機能を模倣できるデバイスとして、不揮発性アナログ可変抵抗素子 (memristive device) の研究が活発化している。誘電体メモリ (Ferroelectric RAM: FeRAM)、磁気抵抗メモリ (Magnetoresistive RAM: MRAM)、相変化メモリ (Phase-change RAM: PRAM)、抵抗変化メモリ (Resistive switching RAM: ReRAM) など、ほぼ全ての不揮発性メモリが研究対象となり、材料探索も幅広く行われている。特に近年、アナログスピン軌道トルク (Spin-Orbit Torque: SOT) 磁気抵抗メモリで、ニューロンの応答特性であるリークを有する積分発火 (leaky integrate-and-fire) 特性と、シナプスの特性であるスパイクタイミング依存可塑性 (Spike-Timing Dependent Plasticity: STDP) が再現され [A. Kurenkov, S. DuttaGupta, C. Zhang, S. Fukami, Y. Horio, and H. Ohno: Adv. Mater. 31, 1900636 (2019)]、超小型で低消費電力のスパイクニューラルネットワーク (Spiking Neural Network: SNN) への応用の可能性が開かれつつある。このように脳型

AIアクセラレータの実現を目指す取り組みを通じて、今後も多種多様なデバイス・材料の研究開発が誘発されると期待される。

脳型AIアクセラレータの研究開発を推進することで、これまで見捨てられていたデバイスが再び注目される可能性も考えられる。従来のデジタルコンピュータでは、信頼性が高く高速なデバイスでの構成が要求され、ノイズや特性変動など、不確実な要素は徹底的に排除されてきた。一方、ニューロンの種類は多種多様であり、動作は低速かつ確率的でノイズが多く存在する。また、ニューロンの状態は常に変化し信頼性に乏しい。しかし、脳ではこれらニューロンの多様性や揺らぎをむしろ積極的に活用しているように見える。このように、脳型AIアクセラレータに適したデバイスが持つべき特性は従来のデバイスと大きく異なると考えられるため、従来のシステムで要求を満たせず忘れ去られてきた多くのデバイス・材料を従来の視点とは異なる視点でもう一度見直してその活用を考えていくことで、新たな方向のデバイス開発が進むことも期待される。

3 | 具体的な研究開発課題

今後取り組むべき研究開発課題を図3-1および図3-2を用いて以下に示す。図3-1は脳型AIアクセラレータ研究開発の方向性を示したものである。ここでは、①の矢印で示すように、脳科学における脳の構造や情報処理機能の知見から新たな数理モデルを構築し、それを脳型AIのソフトウェア・アルゴリズム、回路・アーキテクチャ、さらにはデバイス・材料の開発に繋げていく。また同時に、②の矢印で示すように、脳の分子レベル・細胞レベルなど低次構造の機能に類似あるいはヒントを得た特性を持つデバイス・材料に注目し、そこから新たな回路・アーキテクチャ、脳型AIのソフトウェア・アルゴリズムを創出していく。これら2つの方向の研究開発が相互に影響を与え、新たな発想や連携を促進していくことにより、新たな脳型AIアクセラレータの開発が加速していく。

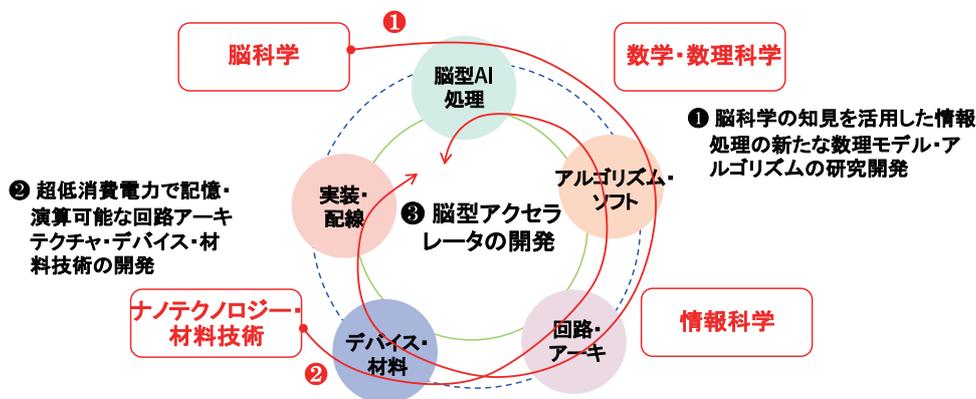


図3-1 脳型AIアクセラレータ研究開発の方向性

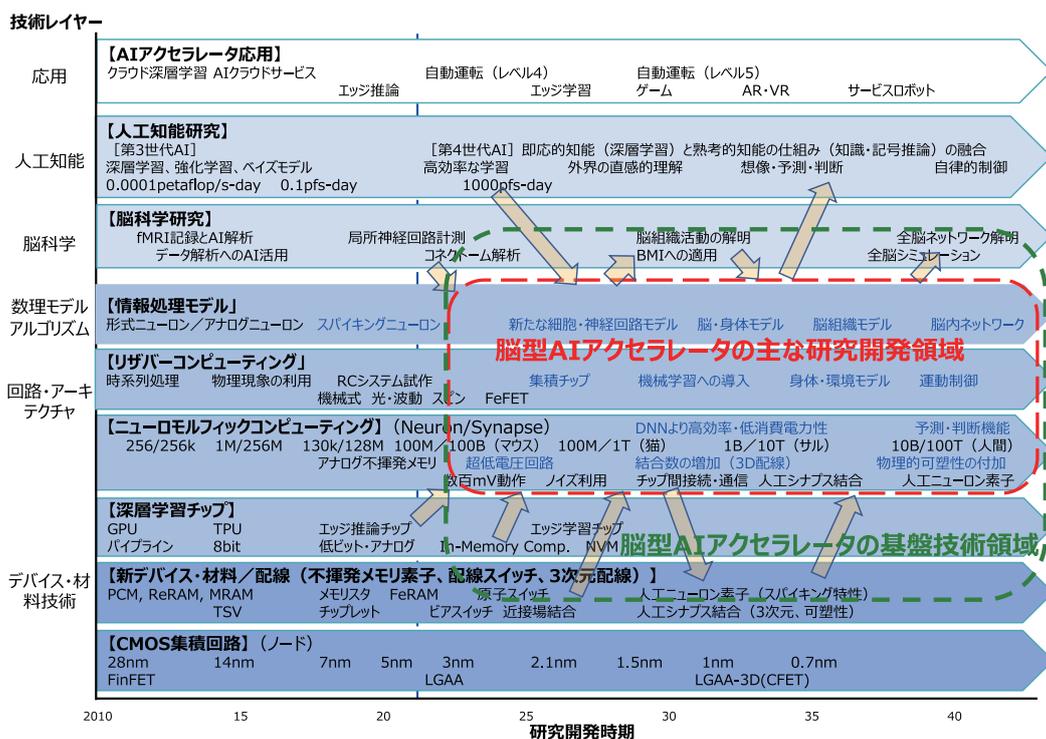


図3-2 脳型AIアクセラレータの研究開発ロードマップ

3 具体的な研究開発課題

研究の流れとしてはこの2つの方向があるが、今後取り組むべき具体的な研究開発課題としては、以下に示すように、脳科学の知見を活用した情報処理の新たな数理モデル・アルゴリズムの研究開発、超低消費電力で記憶・演算可能な回路・アーキテクチャ/デバイス・材料技術の開発、脳型AIアクセラレータの開発がある。取り組むべき研究開発課題の全体像を理解するために、図3-2にCRDSで作成した研究開発ロードマップを示す。ただし、このロードマップは今回の調査から見えてきたものであり、関連の学会やコミュニティでコンセンサスが得られているものではない。横軸は2040年までの長期的な研究開発時期を示し、縦軸は技術レイヤーを示しており、求められる研究開発内容、数値目標などを記載している。赤い点線で囲った部分が脳型AIアクセラレータの主な研究開発領域であり、緑の点線で囲った部分はそれを支える基盤領域である。図中の矢印はそれぞれの技術レイヤーの成果や知見を他の技術レイヤーに活用することを示している。

3.1 脳科学の知見を活用した情報処理の新たな数理モデル・アルゴリズムの研究開発

(1) 脳の各階層の構造・機能の知見からの情報処理モデル構築

脳の情報処理機能を参考・ヒントにした脳型AIアクセラレータの実現には、脳の構造や活動を細胞（ニューロン・シナプス）レベル、神経回路レベル、脳組織（視覚野、前頭前野、海馬など）レベル、脳ネットワーク全体など各階層ごとに理解し、それぞれを数理モデル化して情報処理アルゴリズムを作ることが重要である。そのためには、これまでの脳科学の様々な知見を情報処理の視点で見直して理解を深める必要がある。また、情報処理に活かすことを目的とした新たな脳の構造・活動の精密な計測・評価の最新動向、成果を知る必要があり、空間分解能および時間分解能の高いダイナミックな脳活動の計測を行っている研究者との議論が重要である。

細胞・神経回路レベルの低次機能のモデルとしては、すでにニューロンの機能を単純化した形式ニューロンモデルや、よりニューロンの活動に近いスパイクニューロンモデルがあるが、短期的には最近の空間分解能や時間分解能の高い計測技術により得られた実験結果を基にして、従来モデルの修正やより精緻なモデルの構築を行う必要がある。

また、長期的には脳組織として視覚野、前頭前野、海馬などの構造や機能に関する脳科学の研究成果から大胆な仮説を作ること、これらにおける情報処理モデルの構築も進めていく必要がある。例えば、海馬における短期記憶や場所記憶に対して場所情報処理、時間情報処理などのモデルをつくることや、大脳皮質の6層構造の特定の層に注目してモデル化し予測の情報処理に利用していくといった研究も進められており、このような研究の拡大が望まれる。さらには、人間の記憶に携わる海馬と大脳皮質との連携のモデル化、海馬と扁桃体と前頭前野からなる自律制御に関する統合機能モデルなどの研究開発の例もあり、脳組織の連携や脳全体のネットワークに関わる情報処理のモデル化についても進めていくことが重要である。

(2) 脳機能と身体性を考慮した情報処理モデル構築

脳における高度な判断ができる仕組みを理解しようとするときに、脳そのものの構造や機能を解析するだけでは不十分という考え方もある。人間の体には視覚、聴覚、触覚、嗅覚などのセンシング機能や手足のようなアクチュエーション機能が備わっており、脳の高度な機能の獲得にはこの身体性の利用や制約が大きく関

わっているというものである。脳は物理的に有限なリソースしかないにも関わらず、無限定で動的で不確定な世界に対峙するため、身体性を活用して非常に効率的な情報処理様式を進化的に獲得してきたとも考えられる。このため、身体性と脳機能との関係の研究や、知能の発達過程を研究することなどに対しても注目し、そこからヒントを得て新たな情報処理モデルを構築することが重要である。

特に、多数のセンサ情報から状況をリアルタイムに認識し、環境変化を予測して、次の行動を判断してアクチュエータを動かすといった、実空間上のエッジ応用に向けた脳型AIアクセラレータを実現するときには、大変重要になると考えられる。ヒトや動物の体に内包されているセンサは、現行の産業用ロボットについているものとは桁違いに膨大な数であり、触る、たたく、押すなどの行為に対するフィードバックを多数のセンサを介して得て、自らの存在する世界について学習している。このような身体経験をベースとした学習（身体知）が、実世界で行動するロボット等においても重要となり、長期的な研究になるが、この脳と身体性を含む情報処理モデルを構築する必要がある。この基盤となる研究としては、認知発達ロボティクスや計算論的神経科学における取り組みがあり、例えば深層予測学習により環境とのインタラクションをロボットが学習し、タオル畳みや粉体・液体の計量といった複雑な行動を実行するといった研究開発が行われている。これらの研究分野の知見の活用や、連携などが重要になる。

3

具体的な
研究開発課題

コラム 1

脳科学の新たな知見

2010年頃までは実験脳科学と理論脳科学の間には大きな分断があり、理論物理学と実験物理学が相互の成果に基づいて研究を発展させてきたような関係にはなかった。実験脳科学が得られる脳の情報が狭い範囲にとどまり、理論脳科学者が求める様な脳機能の情報が提供できなかった。この状況は脳計測技術の進展により大きく変貌した。今や小動物の数千以上の神経細胞の活動状態をモニターしつつ操作することができ、ヒトの脳機能が脳イメージング技術により高い時間・空間解像度で得られる。得られた脳領域や広域・局所神経回路の活動情報はそれぞれのレベルの回路モデルの構築に資するもので、これらに対応する様々なレベルの回路モデルが提起される時代になった。新たな回路モデルは実験脳科学により検証されさらに精緻になり、また、新たなAIのアルゴリズム創出の基盤となりうる。

細胞レベルから全脳レベルの各階層で脳計測技術の進展は近年著しい。脳ネットワークを構成するシナプス結合構造の全貌を電子顕微鏡レベルで網羅的に読み出す「コネクトーム」解析は、電子顕微鏡による自動撮像技術とその膨大な画像データからの人工知能による構造情報抽出技術の開発により軌道にのり、ゲノムプロジェクトに匹敵する規模のコネクトームプロジェクトが始動している。生きた脳深部の神経活動を細胞レベルで読み出す二光子顕微鏡技術や行動中の動物の脳深部

から長期間活動記録が可能な超小型内視鏡システムは蛍光プローブ技術とともに進展著しく、光学的あるいは化学的神経回路操作技術（光遺伝学・化学遺伝学）を組み合わせることで、マウスやマーモセットなどの小動物の脳回路を操作して活動を読み取る研究が可能となり、神経回路の動作原理の解明に革命をもたらしている。

全脳レベルの研究においては、fMRIデータのAIによる解析が成果を生み、また、ヒトの脳活動と深層学習の内部状態を比較することによる深層学習アルゴリズムの改良など、脳計測へのAIの寄与とともに脳計測のAI技術に対する寄与も見られるようになってきている。脳の視覚情報処理の階層構造に発想を得て1979年に福島邦彦（当時：NHK放送科学基礎研究所）が発表したネオコグニトロンに端を発した階層型AIは深層学習に結実した。深層学習は脳解析の強力なツールとしても用いられている。数千のトレーニング画像を特徴パターンのベクトルに分解し、ヒトの視覚皮質の反応と画像から計算して得た特徴ベクトルを深層学習に与えることによって、未知の画像を観たときの脳信号から画像の再構成ができることが示され、夢を見ているときのヒトのfMRI記録から夢の画像が再現されている [Horikawa T, Tamaki M, Miyawaki Y, et al. (2013) Neural decoding of visual imagery during sleep. *Science* 340:639–642]。さらに、深層学習と脳を相互作用させることによる深層学習のチューンアップが試みられている。同じ画像を動物と深層学習に入力して、得られた神経活動パターンを教師信号として深層学習を学習させると、脳の神経細胞が画像に対して活動するのと同じ様な反応が深層学習に見られるようになる。例えば1次視覚野での単純なパターンに応答する神経細胞の挙動あるいは高次視覚野における抽象的な視覚情報に反応する細胞活動と似た活動が深層学習で検出される。すなわち、視覚野のニューロン活動をコンピュータ上に再現した（写し込んだ）ともいえる。さらに深層学習のモデルにノイズ成分を加えたり後ろ向きの結合要素を加えたりすることにより、脳の反応により似た、誤判断の少ない挙動が得られることがわかってきた。

(3) 脳機能・ネットワークの新たな数理モデルの構築

脳科学の知見を脳型AIのアルゴリズムやハードウェアへ有効活用するためには、脳科学の知見を数学・数理モデルで記述し、アルゴリズム、回路へと結びつけていくことが必要であり、新たな数学・数理モデルの創出は重要な研究開発課題の一つである。以下、このモデル構築に関して今後重要となる研究開発の例について記載する。

a) 脳機能を解析する数学・数理モデル（力学系理論、情報理論）

ニューロンのスパイク発火を解析する数理モデルを扱う領域は、非線形の微分方程式を扱う力学系理論である。計算機応用を視野に入れ、計算コストに負担をかけず現象を精度良く記述するモデルとして

Izhikevichモデルなどが提案されてきたが、発火頻度など統計的な知見を入れるモデルや、発火過程を抽象化して方程式へ取り入れるモデルなど、脳科学の進歩に応じて、妥当な計算コストでより正確に現象を記述できる数理モデルの研究は引き続き重要である。脳活動のダイナミクスが数学により抽象化されると、同じような複雑系からなる様々の物理系との比較が可能となる。これにより物理系そのものをコンピューティング主体とするリザーブコンピューティング分野の発展や、脳型AIを構成するデバイス・材料に求められる特性・物性に対する有益な情報が得られ、脳型AIハードウェアの開発に資することになる。

さらに、脳活動のダイナミクスを各ノード全体の統計力学として扱い、脳全体のマクロな活動の時間発展を統計力学的に記述することも重要である。脳を構造化して、複数の異なった領域に相互作用として興奮性刺激と抑制性刺激を導入すると、より複雑なダイナミクスが検討できるようになる。このモデルを発展させると、系の平衡状態のみならず、振動や発振など回路的な描像を記述することができ、脳全体を回路と考える脳型AIのコンピューティングにも有益な情報が得られるようになると考えられる。さらに、複雑系の力学系理論の安定性を調べることのできるカオス理論は、脳活動領域のクラスタリングなど大域的な脳活動を解析するための有効な手法であり、これを活用することで、脳の低消費電力な情報処理の要因や、従来の万能型コンピューティングと異なる脳型の情報処理方式の本質へと迫る研究が進展する。

この方向での最近の発展としては、脳シナプスの活動は不均一であり、特定の少数シナプスが強く活動することで、刺激に対して閾値をもたず、弱い刺激でも活動し続けることがわかってきた。脳が極めて低電力で動作する秘密の一端が、特定のシナプスによる強い結びつきとその活動によっているのではないかという示唆を与える結果であり、このようなメカニズムを応用したアーキテクチャ・回路を研究開発することで、情報処理デバイスの超低消費電力化を実現する。

また、神経回路が環境に適応して進化するモデルを、情報理論の観点から、情報量の最大化に注目して説明する数理モデルの研究が始まっている。この研究に基づいたアルゴリズムで高効率に画像処理が行えるという報告もあり、脳全体を調整する役割のシナプスを活用するネットワークで高効率化や超低消費電力化が図れる可能性がある。このような情報ネットワークは小脳の中で実現していることが示唆されており、脳回路の数理モデル化によって、情報処理ハードウェアの高効率化や低消費電力化の道筋の解明が期待される。

b) ネットワーク理論・トポロジー

脳全体の情報ネットワークに着目する数学分野として重要な分野がネットワーク理論である。ネットワーク理論では、脳ネットワークをノードとリンクで構成して隣接行列で表現し、リンクに適切な仮定を与えて分析を行う。具体的には、ノード次数（各ノードの持つリンク数）、クラスター係数（各ノードの周りにある三角形数）、各ノード間の最短経路長などが計算でき、これらとfMRIの測定データとの比較により、脳のネットワーク的な活動の研究が行われている。今後重要になる課題としては、脳の階層モジュール構造の研究があげられる。ネットワーク理論を用いてそれらのモジュール同士がどのように協調し活動しているかを解析すると、脳型AIのトポロジー構築のベースとなる知見が得られる他、脳の持つロバスト性の解明などに関しても有益な情報が得られることが期待される。

脳科学にトポロジー・ホモロジーの手法を導入する研究は最近始まったばかりであるが、これは脳情報ネットワーク全体の時系列活動を大域的に研究する研究として、今後重要な研究課題となると予想される。特にネットワーク情報のトポロジー的な特徴を抽出してマップ化できるパーシステントホモロジーは、脳活動のネットワーク的な特徴を抽出する有益な方法であり、その時間発展が検討できる手法の開発が望まれる。またこれらのトポロジー的な手法により、脳情報処理に特有な堅牢性や信頼性の情報が得られると期待され、脳型

AIの特徴を具体的なアルゴリズムやハードウェアに落とし込むための指針を得る上でも重要である。

c) データ駆動型解析

ビッグデータの時代が到来し、計測されたデータから元々のシステムの構造や性質を解析するデータ駆動型解析が、機械学習の発展とともに極めて重要な研究分野となっている。この分野を支える解析法としてスパースモデリングが注目されている。スパースモデリングは、高次元データに普遍的に内在するスパース（疎）性を利用することで、計算量が次元数に対して指数関数的に爆発する状況でも、現実的な時間でデータから最大限の情報を効率よく抽出できる技術である。脳科学への応用では、一例として脳波情報などの複雑な時系列データを無限次元の線形作用素空間のモードに分解し、データ同士の距離（グラスマン距離）で分類することで、行動の予測や、脳ネットワークの解析を行う研究が進んでいる。脳活動のデータからその特徴量を取り出すための数学理論や数理モデルの研究は、脳科学の進展に重要な役割を果たすのみならず、確立された数理モデルをベースにした脳型AIの開発にも重要な手法になると考えられる。

(4) 脳の各階層での情報処理モデル化からの回路・デバイス・材料への機能要求

上で述べてきた、脳の細胞レベル、神経回路レベル、組織レベル、脳ネットワークレベルでの新たな知見を基に、新たな数理モデルを創出するとともに、そのモデルに従った動作を可能とする回路やデバイス・材料への特性要求を出していくことも重要である。例えば、弱い入力信号が時々入ってもその信号は記憶されず、弱い信号でも頻繁に入ったときや、一度でも強い信号が入ると、状態が大きく変わってその状態が保持されるといった機能を有するデバイスがほしい、といったものである。そのような特性を持つものが既存のデバイスにない場合には、望みの特性を持つ新たなデバイスを開発することも必要になる。脳レベルや神経回路レベルでは、考えやすいようにも思われるが、組織レベルや脳ネットワークレベルでもこのような要求が出てくることが望まれる。

汎用AIを作ろうとする動きや組織（例えば、「全脳アーキテクチャ・イニシアチブ」）も存在しており、それらの活動における研究開発内容についても注目しておく必要がある。脳の組織（モジュール）レベル、あるいは脳ネットワークレベルにおいて、新たな情報処理のモデルとして魅力ある提案が出されている場合には、その研究を進展させ、回路・デバイス・材料への機能要求として提示していくことも重要である。

3.2 超低消費電力で記憶・演算可能な回路・アーキテクチャ／デバイス・材料技術の開発

(1) 超低消費電力のデバイス・材料技術

脳の細胞レベルのニューロン・シナプスを模倣した代表的な演算回路として、スパイクニューラルネットワークがある。以前よりCMOS集積回路を用いて研究開発が進められているが、さらなる高効率化、低消費電力化、大規模化のためには、新たなデバイス・材料の開発および回路・アーキテクチャの開発が重要になる。特に、デバイス・材料技術としては、PCM、MRAM、ReRAM、FeRAMなどの不揮発メモリやメモリスター（memristor）を用いて、シナプスの接続状態を変更する機能や、一定時間保持する機能などが必要と考えられる。新材料を用いた不揮発メモリに関する研究は日本がリードしてきた領域であり、これらを深

層学習の重みづけの部分に使う研究も活発であり、日本の強みを利用しさらに新たな技術を創出する必要がある。また、これまでの2次元的な接続だけでなく、脳のように3次元的な接続を可能とするような配線技術も必要である。さらに、神経回路の可塑性を模倣する場合には、回路の接続を自由に変更できるような機能を持つデバイスも重要になる。3次元的な接続を可能にする配線技術としては、現在の3次元集積化技術のさらなる細粒度化、無線や光を使ったインターコネクション技術、液体や固体の電解質を用いて電界・電流印加により金属細線を伸長させる技術（例えば、原子スイッチ）、などの研究開発が必要となる。原子スイッチは日本発のデバイスであり、すでにFPGA応用のチップが製造できる技術レベルになっており、不揮発メモリ機能と回路接続を自由に変更できる機能、3次元配線機能を合わせ持つデバイスとしても注目され、日本の強みとして研究開発を進めることが期待される。

深層学習よりも簡単な構造で時系列データを取り扱い易いものとして、リザーブコンピューティングが注目されるようになってきている。その中でも非線形の物理現象を用いた物理リザーブコンピューティングでは、望まれる物理特性に合ったデバイスや材料の開発が重要になる。特に、チップ化に適している物理特性を持つデバイス材料として、スピンドバイス、FeFET、フォトニックデバイスなどの研究が重要になる。

新たにつくられる脳のモデルからデバイス・材料に要求される特性によっては、現状のデバイス・材料で使われている物理現象では実現できないものもあると考えられ、新たな物理現象や、ノイズ、ゆらぎの積極的な利用も期待される。特に、ノイズやゆらぎの利用については、超低消費電力化の回路という視点でも興味深い。人間の脳が使用できるエネルギーは電力換算で20W程度と言われ、理化学研究所と富士通が開発した「富岳」の消費電力は30～40MWであり、実に100万倍以上の開きがあるが、その要因の一つとして挙げられるのが、ノイズやゆらぎの扱い方の違いである。現在のデジタルコンピュータでは、多段の論理ゲートを正確に動作させるため、信号のノイズが演算の都度積み重ならないよう、論理ゲートを通過する度に信号が電源電圧に規格化され、徹底的にノイズが排除される。一方、脳を構成するニューロンの状態は常に変化し動作も確率的であるが、これらニューロンの多様性や揺らぎを脳は積極的に活用しているように見える。超低消費電力の脳型AIアクセラレータを実現するためには、こうしたノイズやゆらぎをうまく利用している脳の機能を理解して模倣し、それに適するデバイスおよび回路・アーキテクチャの開発が重要と考えられる。

(2) アナログメモリ

大規模脳型ハードウェアを実現するには、ニューロン部の非線形特性よりもシナプス部の方が重要性が高いと考えられており、シナプス機能を実現する不揮発性アナログ可変抵抗素子の研究開発が重要となる。不揮発性アナログ可変抵抗素子として、FeRAM、MRAM、PRAM、ReRAMなどがあり、ほぼ全ての不揮発性メモリが研究対象となる。近年、アナログスピン軌道トルク磁気抵抗メモリで、スパイクタイミング依存可塑性が再現され注目を集めている。不揮発性メモリは、車載用マイコンのロジック埋め込み不揮発性メモリや、インメモリコンピューティングへの応用に向けて現在最も盛んに研究開発が行われている領域であり、脳型AIアクセラレータの研究開発にも多大な影響を及ぼすと考えられ、多ビット化、アナログ記憶、特性安定性・再現性の向上など、脳型AIアクセラレータに向けた利用形態の多様化、制御性の向上に関する研究開発を進めていく必要がある。

(3) インメモリコンピューティング

超低消費電力化、つまり単位電力あたりの演算量を高める方法の一つとしてインメモリコンピューティング手法がある。従来のノイマン型の構造ではデータが演算部とメモリ部の間をアドレスとデータバスなどを介し

て行き来している。バス間をデータが行き来するだけで大きな消費電力を要する。この手法はその消費電力を削減するために演算部とその演算に必要なメモリ部を近くに配置することでその消費電力を削減するものであり、深層学習用の積和演算を行う回路として研究開発が盛んになっている。脳においても演算部と記憶部が一体化した構造になっており、このようなインメモリコンピューティングの研究は重要と考えられ、演算部とメモリ部をうまくアナログ回路化することでさらなる低消費電力化に向けた研究開発を進めていくことも必要である。

(4) 積和演算におけるアナログ回路技術

深層学習の各層の結合のウェイトを決める演算の基本は積和演算である。その実現方法としては、例えば信号を電流値として和を電流加算、積はオームの法則 ($V=I \times R$) を利用し R を係数とすればよく、積和演算の出力となる V は信号 I と係数 R の積となる。スパイク回路ではその積和の値 V を閾値と比較するためのコンパレータが必要になる。ここで電力を消費する主な要素は電流などの信号を表現伝達する部分、およびコンパレータであり、その実現方法に多くの工夫が求められる。複数のトランジスタの組み合わせで構成される CMOS コンパレータでは出力が変化しない場合でも電力を消費するのが一般的なため、低消費電力化・小型化には適さない。このため、ゼロアイドル電流の特性を持つ小型のコンパレータ素子の実現が求められる。またメモリスタなど重みを表現する接続や書き込みが容易な小型素子の開発や、その低消費電力化へ向けた回路技術の開発も進める必要がある。

(5) 配線再構成技術

ニューロンは数多くのシナプス結合で他のニューロンとつながっているが、現在の CMOS 回路の配線技術では多くの結合を効率良く配線することができない。また配線自体は事前に決定する必要があり一旦製造されれば柔軟に接続を組み替える事は不可能である。実際の脳に近い構造や機能を実現するには、多くの信号を集め結合する方法を開発する必要がある。更に理想的にはその結合が柔軟に変えられる方法が実現できれば、回路の設計手法に革命をもたらす可能性がある。先に述べた原子スイッチのようなデバイスをうまく活用できる回路・アーキテクチャの開発が期待される。

3.3 脳型AIアクセラレータの開発

(1) 新たな脳の数理モデル・情報処理技術と超低消費電力回路・デバイス・材料技術の融合

上記の 3-1 および 3-2 で記載した技術を融合させることにより、脳型AIアクセラレータの開発を行う。具体的には、短期的にアクセラレータの有力な候補としてこれまで研究開発が進められてきたニューロモルフィックコンピューティングや、リザーブコンピューティングの高度化を進めるとともに、長期的には脳の高次機能による新たな数理モデル・アルゴリズムに対応したコンピューティング（高次脳型コンピューティング）のチップ設計に関する研究開発を行う。ニューロモルフィックコンピューティングや、リザーブコンピューティングの高効率化・低消費電力化を進めることで、従来の深層学習チップよりも効率的な学習や低消費電力化が可能なることを示すことが必要であり、研究開発の詳細については以下の (2) および (3) に示す。

脳の高次機能を模倣したアクセラレータの設計・開発については、身体・環境の把握と運動制御、予測・

判断をリアルタイムに行えることなどを目標にする。具体的にどのような数理モデルが提案されるかはまだわからない状況ではあるが、身体性を考慮した数理モデルや、脳の大きな特徴である記憶と演算が共存・一体化していることを記述する数理モデルの構築が望まれる。特に、身体性を考慮した数理モデルに基づく脳型AIアクセラレータについては、実空間上のエッジ応用に向けた新たなAI技術として、多くのセンサからの情報収集とアクチュエータの制御とも密接に関係してくるので、大きな期待がある。

これらのアクセラレータ開発に関する短期的および長期的な研究開発課題に加えて、応用に向けた研究開発、アクセラレータへの最適なタスクの割り当てを可能とするソフトウェアなどの研究開発も必要であり、これらについて以下の(4)および(5)に示す。

(2) ニューロモルフィックコンピューティング

脳の基本構成要素であるニューロンやシナプスの動作を物理デバイスにより再現するニューロモルフィックコンピューティングの開発が、脳型AIアクセラレータの研究開発の中心的な課題の一つであり、図3-2に示すロードマップに沿った研究開発が望まれる。特に、ニューロンの膜電位のスパイク状のパルス信号を考慮したニューロンのモデル「スパイクニューロン」、およびそれを用いて構成したネットワーク「スパイクニューラルネットワーク (Spiking Neural Network: SNN)」に基づく高集積・低消費電力のチップの開発が重要である。

1998年に、前段ニューロンと後段ニューロン間のシナプス結合加重の変化がそれぞれの出力スパイクのタイミング差に依存するという、スパイクタイミング依存シナプス可塑性 (Spike-Timing Dependent Synaptic Plasticity: STDP) [G.Q. Bi and M.-M. Poo: J. Neurosci. 18, 10464 (1998).]が発見され、これを契機としてスパイクタイミングを考慮した積分発火型SNNの研究が世界的に活発化している。最近では学習アルゴリズムの改良によりSNNで高い学習性能が得られるようになり、消費電力が低くノイズに強いSNNの特長が活かせるようになってきている。2014年にIBMが発表したSNNのチップ「TrueNorth」では、ニューロン1個に対するシナプス数が256個にとどまり、カエルの脳程度の規模しか実現できなかったが、2017年にIntelが1ニューロン当たりのシナプス数を約1000個に増加させたチップ「Loihi」では、スパースモデリングのタスクを、同社のマイクロプロセッサ「Core i7」と比べて演算速度で1万倍、推論1回あたりの消費電力は100万分の1で行うことができ、圧倒的な性能を示した。さらに、2020年3月にはIntelからLoihiを768枚結合させたシステムでネズミの脳並みの数のニューロンとシナプスを持つSNN演算システム「Pohoiki Springs」が発表されている。SNNの真価は、人間並みあるいはそれを超える能力を発揮することであり、さらに高密度化・高接続数や超低消費電力化を可能とする必要があり、この実現に向けた研究開発が必須である。

(3) リザバーコンピューティング

リザバーコンピューティング (RC) についても、図3-2のロードマップに沿った研究開発が望まれる。特に、アクセラレータとして従来のマイクロプロセッサと一緒に使うことを想定すると、素子を集積してチップ化できる物理リザバーコンピューティングの開発が重要となる。

80年代から、現在の深層学習のベースである順伝搬型・階層構造のニューラルネットワークに加えて、深層学習にはないフィードバックを有する再帰型ニューラルネットワーク (RNN) の研究が始まり、今世紀初頭にその一つとしてRNNで出力のみの重みを学習するRCが提唱された。このRCのモデルは、大脳皮質の微小カラム構造のモデルとして提案されたLiquid State Machineに端を発する [M. Lukosevicius, H.

Jaeger, “Reservoir computing approaches to recurrent neural network training,” *Computer Science Review* 3, 127 (2009)]. 整然とした階層型ニューラルネットワークを用いる深層学習と異なり、入力と出力の中間の層（リザーバー）は非線形な入出力特性さえ示せばブラックボックスでよく、システムに必要とされる複雑な非線形性を持つ様々な物理・化学システムをリザーバーとして動作させることができる。RCは時系列のパターン認識が重要な役割をしている応用に有望であると考えられ、2015年頃からNNを構成するネットワークに実際の物理系をリザーバーとして用いるRCが盛んに研究されるようになっていく。

物理リザーバーの種類としては、①非線形のネットワーク構造を持つもの（メモリスト、FeFET、バネ、細胞など）、②フィードバックを持つ単一の非線形システム（光デバイスなど）、③連続媒体（液体、磁性体、弾性体など）がある。物理リザーバーを用いたRCでは、システムがエコステートプロパティ（応答が過去の変数系列の関数）であることが必要条件であるため、使用するシステムが数学的にこの要請を満足するかをまず検討する必要がある。また、物理リザーバーの善し悪しを判断する基準は、リザーバーを構成する物理系がどれだけリッチで複雑な非線形システムであるか、また物理系そのものに多くの情報が記憶可能であるかなどであり、これらの点を検討することが重要になってくる。また、現在主流の深層学習は層の数や計算精度、データ量などの計算コストに依存して学習・認識性能が向上していくという特徴があるが、RCにおける性能はタスクと物理リザーバーの種類によって大きく変わってくると考えられ、タスクと性能の関連を物理系に即して数理モデルを用いて定量的に解析することと、タスクに即した最適な物理リザーバーを探索・設計していくことは重要な研究開発課題となる。

RCはその単純な構成にもかかわらず、複雑系のカオス的なダイナミクスによる豊かな情報表現と、多様で分散した情報の統合とを同時に実現する可能性のあるシステムとして、脳機能の一部を実現する可能性を秘めている有望なモデルと言える。このため、デバイスの物理的・動力学の性質を直接的に利用するRCの実現は、デバイス・材料分野から脳型AIアクセラレータ構築への大きな礎になると期待される。

一方で応用面から見ると、RCは現在画像認識等で大きな市場を獲得している深層学習と比較してどのような応用がキラーアプリとなるかがはっきりしていない。一般的には、リアルタイム性や少量のデータ処理が必要であるエッジAIとしての適用が想定されているが、用途によって必要なハードウェアの性能、大きさ等が決まってくる。応用に応じた仕様の策定、仕様に合わせたアルゴリズムの開発や物理リザーバーの非線形性を制御する周辺回路の開発なども重要である。エッジ応用では周辺回路のオーバーヘッドを極力低減したいので、その観点が社会実装へのチャレンジを成功させる鍵となる。

(4) 応用に向けた研究開発

実際のアクセラレータ開発に当たっては、デバイス・回路設計だけでなく、大規模化のための配線・実装技術についても研究開発しておくことが重要である。特に、シナプスのように3次元的に大量の接続を行うようなことを考えると、これまで半導体プロセスで用いられている3次元配線とは桁違いの接続が可能な技術を検討する必要がある。例えば、近接場結合、電解質中の金属細線形成などの新たな3次元配線・実装技術の開発が望まれる。また、新材料による不揮発メモリ機能を利用する場合には、トランジスタなどの演算デバイスと同じ層にするのか、配線と同じ層にするのかといったレイアウト上、プロセス上の課題もあり、デバイス・回路設計とのつながりも考慮したチップ全体の設計を考えていく必要がある。

想定する応用先で求められる機能を想定した設計（ドメインスペシフィックな設計）が重要になる。特に、脳型AIアクセラレータの重要な応用領域として考えられる実空間におけるサービスロボット、スマートロボットなどのロボティクス分野においては、様々な多くのセンサおよびアクチュエータを搭載しており、多くのセン

シング情報の中から重要な情報を選択して認識・判断し、判断した結果を複数のアクチュエータに効率的に伝えるなど、これまでのAI処理になかった実空間対応の機能の開発が要求される。例えば、スマートロボットには、オープンな環境である程度自律性を持って稼働することが求められ、(1) 自己と環境の的確な状況認識、(2) オープンな環境変化への対応、(3) リアルタイムの制御などが求められる（戦略プロポーザル「ナノ・IT・メカ統合によるロボット基盤技術の革新 ～人に寄り添うスマートロボットを目指して～」：CRDS-FY2015-SP-03参照）。このような実空間対応の機能を実現するためのシステム設計が重要になる。このような多くのセンサ入力を瞬時に分別・処理する技術は非常に難度が高く挑戦的な課題であるが、逆にそれが可能になると技術的にも産業的にも大きな優位性をもつことになると考えられる。

さらに、脳型AIアクセラレータの性能や潜在的な能力を社会に対してアピールできるように、特定の応用を想定したフラグシップチップを開発し、アクセラレータの進捗具合や実使用への段階が分かるようにしていくことも重要である。例えば、サービスロボットが周囲の環境変化を把握して、障害物を取り除いたり避けたりしながら目的の場所まで移動するという場面を想定すると、複数のイメージセンサ、音響センサ（マイク）、温度センサ、GPS、加速度センサ、接触センサ、匂いセンサなどからの情報を総合的に把握して、今後の環境変化を予測し、次の行動を判断し、腕や足のアクチュエータに指令を与えることができるロボットAI制御チップが必要になる。この状況判断や予測、制御の正確性や処理の速さを進捗の指標にして、技術レベルを可視化していくことが重要になる。

(5) アクセラレータチップを効率的・適切に利用するためのソフトウェアなどの研究開発

脳型AIアクセラレータのチップを作製しても、それを効率的に動かすためのソフトウェアがないと、全体のコンピュータシステムの効率化・低消費電力化が図れない。このため、ホストとなる汎用コンピュータ上に、アクセラレータへのタスクを割り当て、実行させるソフトウェアの開発も重要である。ここでは、ソフトウェア開発の視点から全体の効率化のためにアクセラレータのハードウェアの仕様に対する要求も出てくることが考えられるため、アクセラレータの開発と利用ソフトウェアの開発は密に連携しながら行う必要がある。

また、脳型AIアクセラレータの誤動作やノイズによるエラーが発生しても、全体のコンピュータシステムとして正常に機能し続けるフォールトトレランスについても検討しておくことが重要である。人間の脳の場合には、突発的に誤った判断をしそうになった時に、社会倫理、過去の経験、将来予測などと照らし合わせて誤りを訂正することができるが、同様な機能についても検討しておく必要がある。このようなフォールトトレランスの機能が脳のどの階層で行われているかまだ解明されていないので、脳型AIアクセラレータ自身に実装できるかどうかはわからない。しかし、コンピュータシステムはCPUとアクセラレータとの協調で動くので、アクセラレータからおかしな結果が返されてきたときには、少なくともソフトウェアにより特異な出力結果を破棄したり、アクセラレータにやり直しを指示したりすることは可能と考えられる。

脳型AIアクセラレータに求められる機能の一つにリアルタイムの判断がある。いくつかの判断の候補を出して、その中から状況に最適なものを選別していくことが考えられるが、その候補があまりにも多くなってしまうと、それぞれの候補を出すための時間や、それらと比較する時間が長くなってしまい、リアルタイム性が失われる。このため、候補の数を少なくするためにある程度バイアスを掛けて候補を出す必要があり、それには脳の思考にバイアスがかかる特徴を積極的に活かすことも重要と考えられる。

コラム 2

物理リザバー AI ブーム

近年人工知能や機械学習による情報処理が飛躍的な発展をとげている。それを支えているのが画像認識をはじめとする多様な応用で高い性能を発揮するディープラーニング（深層学習）で、これには階層的ニューラルネットワークが利用されている。一方、再帰的ニューラルネットワークは、前の時刻のニューロンの状態が、他のニューロンにフィードバック入力されるような構造であり、時系列情報の機械学習に適している。リザバーコンピューティングはその特殊なモデルを基礎として発展してきた。

リザバーコンピューティングのシステムでは、リザバーと呼ばれる非線形な応答や情報記憶を司るデバイスに、複数の入力と複数の出力が接続されている。最近のアルゴリズムでは、出力層の結合のみを学習する方式が開発され、高速な学習が可能となった。初期のリザバーは、非線形活性化関数を有する人工ニューロンから構成される Echo State Network (ESN)、あるいはスパイクニューロンの結合からなる Liquid State Machine (LSM) のように脳のシナプス結合を模した再帰ニューラルネットワークであった。ESN とはリザバーの中で過去の入力情報を反響 (Echo) させて、時系列入力のパターンを学習するアルゴリズムで、各時刻での入力の影響がリザバー内で時間とともに徐々に消失していく性質 (Echo state property (ESP)) を持つように設計する。他方 LSM は、脳の皮質における時間依存情報のリアルタイム処理を模した構造で、非線形な連立微分方程式で記述される。いずれも、リザバーが適切な「記憶」が可能で豊かな非線形性を持つことが要請されている。

このようなリザバーに求められる性質は、数学的には入力を非線形変換により高次元の特徴空間に射影することであり、必ずしも再帰的ニューラルネットワークを用いる必要はない。そこで最近、非線形性を有する様々な物理系をリザバーとして用いる試みが極めて盛んになってきた。これまで報告されている物理系は、力学系モデル (遅延力学系、セルオートマトン、結合振動子)、電気回路 (ASIC、FPGA、VLSI、アナログ回路、ニューロモルフィック回路、メモリスタ)、光学系 (光増幅器、レーザ共振器)、スピン系 (スピントルク発振器、スピン波)、機械系 (ばね系、テンセグリティロボット、ソフトロボット)、生物系 (脳部位、培養細胞) など百花繚乱状態である。これらの物理系のリザバーとしての性能は、ESP や記憶容量 (Memory capacity) を調べることでベンチマークされる。また、その性質は非線形力学系理論、情報・統計理論を用いて解析され、リザバーのダイナミクスと全

体の計算性能の間の関係を明らかにしようとする研究も活発に行われ、数学・数理の研究者の間にも流行の波が広がっている。

以上のように、リザーバーコンピューティングは、従来の機械学習が得意でない時系列パターン処理を行うのに大変便利なシステムで、今後ますます実用化に向けたハードウェアデバイスの開発が期待される。他方、リザーバーコンピューティングの数理モデル化の研究が進展することで、似たような非線形性を有する脳の情報処理機構の理解が進むと予想される。この分野は、ハードとアルゴリズムが相互に絡み合って発展するという意味でも、たいへん興味深い研究分野と言えるのではないだろうか。

(参考文献)

- ・ 田中剛平：リザーバーコンピューティングの概念と最近の動向：電子情報通信学会誌 Vol. 102, No. 2, 108 (2019).
- ・ Kohei Nakajima: Physical reservoir computing—an introductory perspective: Japanese Journal of Applied Physics 59, 060501 (2020).

4 | 研究開発の推進方法および時間軸

上記の研究開発を効率的に進めていくためには、長期的な視野に立ち、強いリーダーシップにより脳科学、数学・数理科学、情報科学、ナノテク・材料技術の異分野の連携を促進することや、産業界を巻き込んで魅力的な応用領域（例えばロボティクス）を決めて我が国がリードできるシナリオを描き、時間軸とテクノロジーレベルを考慮した研究開発の推進が重要になる。そのためには、5年程度の短期的な視点、10年程度の長期的な視点に関わらず、脳科学の知見を新たな情報処理技術、次世代の人工知能技術、チップ化技術に活用することを強く意識し、コンピュータやプロセッサの研究開発に携わってきたリーダーが必要である。その下で明確な目標を掲げ脳型AIアクセラレータの開発を目指すことが重要である。また、具体的な研究開発課題のところでも述べたが、脳科学の知見から出発して数理モデル、アルゴリズムへと進む流れと、逆にデバイス材料から出発して回路・アーキテクチャ、アルゴリズムへと進む2つの流れで、それらの融合を図る体制を構築することが重要である。一方で、これらの体制の中で得られた知見や技術を蓄積していき、国内外との連携によるチップ化や応用領域の探索などのエコシステムの構築を進めていく必要がある。さらに日本国内において裾野を広げるため、脳科学、数学・数理科学、情報科学、ナノテク・材料など異分野の学会間の連携を進め、他分野に興味を持つ人材育成を促進することで、新たなコミュニティの形成に結びつけていくことが重要である。以下、これらについて図4-1を参照して説明する。

- ・脳科学、数学・数理科学、情報科学、ナノテクノロジー・材料分野の強固な連携による脳型AIアクセラレータの研究開発の推進
- ・次世代人工知能用のアクセラレータチップ作製を強く意識したリーダーによる推進

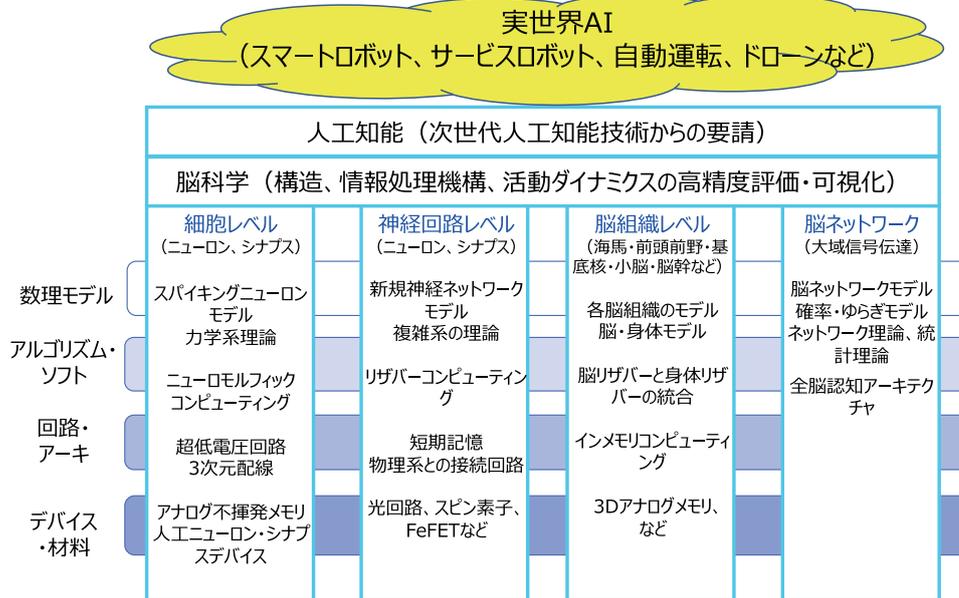


図4-1 脳型AIアクセラレータ研究開発の推進方法

4.1 脳科学、情報科学、数学・数理科学、ナノテク・材料との密な連携

脳型AIアクセラレータの研究開発には、脳科学、情報科学、数学・数理科学、ナノテク・材料との密な連

携が必要であるが、特に情報科学、数学、ナノテク・材料分野の研究者が新たな情報処理技術・AI技術を創出することを目的に、脳科学の研究者の領域へ積極的に踏み込んでいくことが重要である。これまでの脳型情報処理に関する国内外の大型のプロジェクトにおいては、脳の理解や情報処理機構の解明を最上位の目的とし、それに付随する形で脳型の計算原理・コンピューティングの研究が行われてきた。このような研究体制では、脳の情報処理機構を忠実に再現することを目指す研究に偏ってしまい、脳の特異な構造や信号伝達・信号処理からヒントを得た情報処理に関する飛躍的な発想は生まれにくい。現状の人工知能技術では不可能な高度な情報処理機能や消費電力を大幅に低減できる新たな技術の獲得を目指す際には、脳の情報処理機能の再現という目的はかえって技術開発の妨げになりかねない。したがって、脳の知見を工学的に利用する視点で、脳にヒントを得たAIハードウェア技術の獲得を第1の目的にした連携体制の構築が必要である。また、脳科学の研究者も、脳の忠実な再現ではない目標に対する理解と、新たな情報処理のモデル構築、ハードウェア開発に対する積極的な取り組みが重要である。

その要になるのが、この研究を推進するプロジェクトのリーダーである。リーダーとしては将来のコンピュータシステムを見据えて、新たな情報処理モデルの構築とチップ化をやりとげる強い意志を持ち、ハードウェアとソフトウェアの両方を把握できることが必要なため、これまでにコンピュータシステム、プロセッサ開発、新たな回路・アーキテクチャ開発に携わってきた人材が望まれる。出身の研究機関としては大学、国研、企業が考えられるが、研究開発のフェーズや技術開発レベル（TRL）を考慮して人選することが必要であろう。例えば、機能検証段階では大学や国研の研究者をリーダーにし、性能実証段階では国研あるいは企業の研究者をリーダーにすることも考えられる。

脳科学、情報科学、数学・数理科学、ナノテク・材料の異分野の研究者が日常的に活発に交流できる場所を提供し、脳の情報処理モデルに対する熱い議論を通して相互作用を起こすような研究環境の整備も重要である。このような研究環境を実現するためには、異分野の研究者・技術者が集まれる研究拠点の構築が望まれる。例えば、すでにWPIとして脳科学、数学・数理科学、情報科学などの連携によるIRCNが設立されており、このような形の基礎研究を主な目的とした研究拠点の形成が考えられる。また、もう少し応用や実用化を意識し国のプロジェクトを効率的に遂行することを目的に、中心となる研究機関において産官学のコンソーシアムのような形も考えられる。さらに、研究拠点は一つである必要はなく、例えば図4-1に示すように脳の情報処理モデルごとにいくつかの研究拠点を形成してもよい。複数の研究拠点を設ける際には、リーダーあるいはそれぞれの拠点でのサブリーダーの意思やマネージメントが全体に行き届くような体制の構築が必要である。これらの研究拠点において、脳科学、数理モデル、アルゴリズム・ソフトウェア、回路・アーキテクチャ、デバイス・材料など異分野の研究者・技術者が集まり議論することになるが、それぞれの興味や使う専門用語が異なっているため、円滑なコミュニケーションが難しい。このため、それぞれの立場・研究の興味の理解や、言葉の理解（同じ言葉でも専門領域で意味が異なる）など、日々の研究活動や生活の中で相互理解が進むような取り組みや場の設定が重要である。

また、これらの研究拠点においては、研究者・技術者の入れ替わりが頻繁に起こると考えられるが、そこでの研究成果や知識が研究拠点に蓄積されていくような仕組みの構築も重要である。これについては、以下の4-3において述べる。

このような研究拠点の構築は時間軸を考慮して検討する必要がある。脳科学、数学・数理科学、情報科学、ナノテクノロジー・材料の分野の研究者を集めて、脳型AIアクセラレータ開発の研究拠点をすぐに構築することも考えられるが、魅力ある応用先やコアになる技術がまだ明確になっていない状況では、産業界の参加が難しい。このため、最初はすでに存在する研究拠点（例えばIRCNなど）と連携・協力して、プロジェクト

の下でネットワーク的な研究拠点を形成し、魅力的な応用先を見つけながら、コア技術を創出していくのがよいと考えられる。その後に産業界も含めて異分野の研究者が集まる実際の研究拠点を構築して、明確な応用を目指した産学連携の強固な研究体制をつくっていくことが必要であろう。

4.2 脳科学・数理モデルからと材料・デバイス・回路から攻める体制

これまでの脳型AI研究においては、脳科学の知見から出発して数理モデル、アルゴリズムと進む流れが主であった（図3-1の①の矢印）。しかし、背景のところで述べたように、ムーアの法則が破綻して微細化による半導体の高性能化・低消費電力化が困難になることから、現状のコンピュータ技術やAI技術に対して危機感を持っているのは材料・デバイス・回路の研究者であり、脳型AIアクセラレータへの期待も大きい。このため、脳科学・数理モデルから出発し、アルゴリズム、回路・アーキテクチャ、デバイス・材料へと波及していく研究の方向性ととも、デバイス・材料から出発して回路・アーキテクチャ、アルゴリズムへと進む流れを構築することも重要である（図3-1の②の矢印）。ただし、ここで注意すべきことは、デバイス・材料の研究者が独りよがりにならないようにすることである。新たなデバイス・材料を新たな情報処理モデルに繋げていくために、回路・アーキテクチャ、アルゴリズム、数理科学、アプリケーションなどの技術レイヤーの研究者たちと常に議論し、立ち位置や研究の方向性を明確にしていくことが重要である。

ハードウェアに関する研究開発においてはこれまで、デバイス・材料、集積回路・アーキテクチャ、モデル・アルゴリズム、システム・応用という、各技術階層の研究者による分業体制で行われてきた。しかし、脳型AIアクセラレータの研究では図1-1に示したように、脳科学、数学・数理科学といった分野の異なる新たな技術レイヤーの研究者が加わり、これらが絡み合いながら研究を進めていくことが必要になる。異なる分野・階層の技術者間ではバックグラウンドとなる知識も大きく異なり、議論がかみ合わないことも少なくない。脳科学、数学・数理科学、デバイス・材料、回路・アーキテクチャ、実装・配線、アルゴリズム・ソフト、人工知能、応用までの広い学術分野、技術階層間の連携を効果的に進めるためには、複数の技術階層に精通するインターディシプリナリな人材の育成が重要となる。例えば、材料やデバイスの研究者は、物性物理や材料科学、電子工学の知識だけでなく、集積回路のアーキテクチャ、ソフトウェアに関する専門知識も併せ持ち、応用側の視点から材料技術、デバイス技術の新しい可能性を引き出す能力が求められる。

4.3 技術・知識の蓄積とチップ化のエコシステムの構築

研究拠点において研究開発を進めていくときに、それぞれのプロジェクトで得られる研究成果やノウハウを蓄積していき、それを研究拠点全体で活かすとともに、日本の強みとして活用していくことが重要である。国のプロジェクトを実施している間は、公的研究機関の研究拠点に最先端の装置や産官学から優秀な研究者・技術者が集まり、新たな研究成果や装置を生み出したが、プロジェクト終了後に、人がそれぞれの研究機関に戻り、先端の技術やノウハウ、装置が散逸して、研究拠点となっていた公的研究機関に技術や知識が残らなかった例が多く見られる。このようなことを避け、研究拠点あるいは中心となる研究機関に技術や知識が蓄

積され活用できるようにしていく仕組みを検討しておく必要がある。例えば、図4-1において、脳の各階層におけるそれぞれのプロジェクトを遂行しているときに、プロジェクト間で情報を共有することに加え、横ぐしとなる技術レイヤーの研究者がそれぞれのプロジェクトの知見について議論し、それを学術的な視点で一般化し共有化していくことなども考えられる。これにより、あるプロジェクトの人が入れ替わっても、他のプロジェクトの人からそれまでの知識を受け継ぐことが可能となり、中心となる研究機関にも知識が蓄積されると考えられる。

脳型AIアクセラレータの研究開発のようなハードウェアの研究開発体制を強化するためには、新しいアイデアを速やかに具現化できるよう、ファウンドリーでの試作を仲介する設計サービス機能を強化する必要がある。現状、日本国内ではファウンドリービジネス自体が育っておらず、半導体製造プロセスの知識や技術を蓄積する体制すら脆弱である。このため、先端のプロセス技術を有する海外のファウンドリーを使えるような体制を構築することが重要である。すでに大学の中には海外のファウンドリーと提携して先端プロセスによる試作サービスを受けられるようにしているところもあるが、このような海外ファウンドリーとの協力体制を独自に構築するか、すでにこのような仕組みを持つところと連携を図る必要がある。また、脳型AIアクセラレータでは、従来のSi-CMOSデバイスだけではなく新動作原理のデバイス・材料の開発が求められるため、これまでの標準的なシリコン集積化プロセスだけではなく、新たな材料を取り扱えるプロセスの開発も必要になる。そのため、画一的なファウンドリーとは異なるフレキシブル・マニュファクチャリングのサービスも必要になる。これには、TIAのような様々な材料を扱える集積プロセス技術を持つ機関との連携も重要になる。また、将来的に多様なセンサと脳型AIアクセラレータとの集積化・一体化を考えていくには、産業技術総合研究所が開発しているミニマルファブのような新たな製造方法の検討も必要になる。ミニマルファブでは、それぞれの装置を小型化してクリーンルームを不要とすることで、大幅な省エネ化と低コスト化、短期間の製造を実現し、多品種少量といった潜在的な半導体生産ニーズにフレキシブルに適應できると期待されており、脳型AIアクセラレータのように様々なアイデアをすぐにチップ化できるファウンドリーサービスの新たな形としても注目される。このような国内のファウンドリーの機能を活用することで、そこに新たな材料のプロセスなどの知識やノウハウが蓄積するような形になれば、日本の技術力強化につながることができると考えられる。

さらに、職業研究者にとどまらず、アマチュアが個人的趣味でデバイスや回路のプロトタイプを製造できる時代の到来も予想される。Chris Andersonが提唱したMaker Movement [クリス・アンダーソン著”Makers” NHK出版(2012)]では、3DプリンタやArduinoなどのオープンソースハードウェアの登場で、個人がプロトタイプ製品を製造できるようになり、ものづくりを趣味とするアマチュア集団による新産業革命が起こるという未来予測がなされており、こうした動きがナノエレクトロニクス分野に波及する可能性も十分考えられる。特に脳型AIアクセラレータは、ロボットなどエッジにおける応用が期待されるものの、魅力的な製品イメージや市場がまだ明確になっておらず、すぐに大規模な研究開発体制を敷ける状況にはない。むしろアマチュア集団の趣向的活動の中からキラアプリが登場し、研究開発をけん引する可能性も考えられる。従来の常識にとらわれない、新発想のコミュニティやプラットフォームを形成することが、次の時代のイノベーションの醸成につながると考えられる。

4.4 産業応用を見据えた新たなコミュニティの形成、産学連携、国際連携

脳型AIアクセラレータを実用化し魅力的なマーケットや社会に実装していくためには、アカデミアと産業界との連携が必須である。明確な市場がまだ見えていない中長期的な研究開発に対して、産業界がすぐに人材と資金を投入することは難しいが、技術的な魅力と将来産業・市場へのポテンシャルをアカデミアと産業界と一緒に議論し将来のターゲットを共有していくことで、徐々に連携やプロジェクトへの参加が増えていくと考えられる。例えば、スマートロボットは今後の社会的なニーズも大きく、これまでの日本の強い技術力を生かせる領域であるので、脳科学、数学・数理科学、情報処理、デバイス・材料関係の学会を跨ぐ形で、スマートロボットあるいはスマートロボット用のAI技術について議論するコミュニティを形成していくことが望まれる。2019年から応用物理学会にAI技術を議論する「AIエレクトロニクス」のセッションが立ち上がっており、このような活動の発展を期待したい。

このような学会レベルの活動だけではなく、先に述べたプロジェクトや研究拠点においても産学連携の仕組みを構築していくことが重要である。個々の大学と企業との共同研究もよいが、共通の問題意識を持つ多くの人の情報交換・交流の中から多くの新たなアイデアが出てくるので、プロジェクトや研究拠点において産業界が入ってきやすい目標や仕組みを作ることが望まれる。例えば、マルチセンシングに対応した介護ロボット用の脳型AIアクセラレータを作るという目標を立てると、チップメーカー、IT企業だけでなく、センサ、ロボット、介護サービスなどの企業も参加しやすくなると思われる。

このような産学の連携を進めていく中で、4-3で述べた技術・知識の蓄積とともに、技術の知財化、IPコア化、国際標準化などの戦略的な取り組みを行っていくことが重要である。特に、情報処理・人工知能技術における日本の産業競争力の強化に向けて、知財、IPコアの取得は必須であり、脳科学の知見から新たな数理モデルを創出した段階や、新たな回路・アーキテクチャ、新デバイス・材料を考えた段階での基本特許の出願、その後の周辺技術の継続的な出願を行い、コア技術に対する強い特許群を形成していく必要がある。そのためには、産業界だけでなく大学・国研の研究者も特許意識を高めていくことが必要であり、産学での日々の議論ができる体制・仕組みの中での知財化の活動が望まれる。さらに、コア技術を基にした国際標準化の活動を積極的に進めていくことも期待される。人工知能関係の国際標準化としては、2017年に国際標準化機構（ISO）と国際電気標準会議（IEC）がAIに関する合同技術小委員会（ISO/IEC合同技術委員会1、AIに関する小委員会42）が設置され、人工知能に関する概念や用語の定義、システムのフレームワークの整理がなされている。この動きに注目しておくとともに、ISO、IECにおけるデバイス・材料、回路・アーキテクチャ関係の活動に提案できるかを検討し、可能性がある場合には積極的に提案して世界での産業化の促進に貢献し、日本の存在感を高めていくことが望まれる。

全てのことを国内だけで実施しようとする、日本がまだ取り組んでいない研究領域や弱い技術領域に対しては遅れがちになってしまうので、海外との連携についても積極的に行う必要がある。例えば、先に挙げた先端プロセスを用いたチップ作製に関しては、台湾の研究機関や企業との連携が考えられ、また欧州のHuman Brain Projectや米国のBRAIN Initiativeに参加している研究機関との連携も考えられる。例えば、ReRAMやFeFET、原子スイッチなどのデバイス・材料開発や動作解析では日本の研究レベルが高いが、これらとCMOS回路との集積プロセス技術では台湾の研究機関やファウンドリー企業の技術力が高くなっており、互いの技術を補完する形の国際共同研究の推進が考えられる。また、2013年から脳の理解や情報処理機構の解明を最上位の目的として進められてきた欧州のHuman Brain Projectでは、脳科学・神経科学からの知見とスーパーコンピュータによる脳機能のシミュレーションによる多くの知見が得られていると考えられ、これらの参加機関との情報交換や人材交流を進めていくことが考えられる。脳型のハードウェア開発だけでなく、汎用プロセッサとつなぐソフトウェア、設計のためのツールなどの開発も行われてきており、参考になるこ

とが多い。日本は先に述べた新たなデバイス・材料の技術や、応用分野としてのロボット及びその制御技術に優れているので、今後の脳型AIアクセラレータの実用化に向けて、国際連携が構築できると期待される。このような国際連携活動により、世界の様々な分野の研究者との交流を深めて日本の研究力・技術力を発信し、海外からの優秀な研究者の招聘に繋げていくことが期待される。また、魅力的な応用領域を探すために、ハッカソンやコンテストを共同で企画して、新たなアイデアや魅力的な使い方を見出していくような活動も期待したい。

海外との連携においてオープンイノベーションの基本的な方針は維持するものの、開発レベルに応じてオープン・クローズ戦略を持つことも必要である。プレコンペティティブな状況ではほとんどの情報をオープンにしてもよいが、企業が参加して製品化が見えてきた段階では、キーになる技術についてはクローズにして、優位性を確保しておくことも必要である。

4.5 人材育成

脳型AIアクセラレータの研究開発には脳科学、数学・数理科学、情報科学、ナノテクノロジー材料分野の人材が必要になるが、国内では十分な数の研究者がいるわけではなく、その育成が重要である。例えば、脳の機能解明や病気の治療には興味があっても、情報処理への応用を考える研究者は少なく、ハードウェアの開発まで考える人工知能の研究者も少ない。また、アクセラレータ開発の中心となるべき回路・アーキテクチャの人材はもともと多くない。一方で、デバイス材料の研究者は、今後の集積回路の微細化の限界が迫っていることから、危機感を持って他の技術領域に発展させていこうという機運が高まっている。また、最近の深層学習は脳科学やデバイス材料の研究にも活用されるようになり、デバイス材料研究者の情報処理分野への障壁は低くなりつつある。このような状況を踏まえると、脳を生理的に研究し、かつ数理科学・情報科学にも興味を持ち親和性ある研究者の育成や、回路・アーキテクチャ人材の育成、デバイス材料研究者の回路・アーキテクチャ、人工知能など情報処理分野への進出を促進していくことが必要である。

また、次世代のプロジェクトリーダーを育てていく仕組みも必要と考えられる。リーダーとして望まれる経歴としては、コンピュータシステム、プロセッサ開発、回路・アーキテクチャなどの経験があることを述べたが、日本におけるコンピュータ産業や半導体産業が小さくなっている状況では、そのような人材はかなり少ないと思われる。脳型AIアクセラレータの実現を目指すようなプロジェクトや研究開発体制、研究拠点は、このような経験を積む機会として非常に重要と考えられる。プロジェクトや研究拠点の中でチップ開発に直接携われる機会などを通して、今後のコンピュータシステムまで考えられる人材が育成されることが期待される。

脳科学、情報科学から材料分野など広範囲にまたがる関連分野すべてを知悉する専門家の育成は不可能でも、研究者が各々の専門分野に加えてこれら広範囲の関連分野の知識を持っていれば、脳型AIアクセラレータ研究開発のような学際的研究を推進する上で決定的なポテンシャルとなる。日本では高等教育における科目選択の硬直性が問題になる。大学院以降は専門教育・研究を最優先にするとしても、学部教育に柔軟性を持たせて、脳科学と数学のダブルディグリー、脳科学系、情報科学系、工学系科目の幅広い選択の可能性を与える方策が必要と考えられる。

コラム3

脳の構造・機能と精神的な働き

300年以上前、ライプニッツは彼の代表的著作「モノドロジー」の中で、精神的な働きを脳の機械的構造や機能で説明することはできない、という主張を、次のようなわかりやすい例えで述べている。

「ものを考えたり、感じたり、知覚したりできる仕掛けの機械があるとする。その機械全体をおなじ割合で拡大し、風車小屋のなかにも入るように、そのなかにはいってみたりする。だがその場合、機械の内部を探って、目に映るものといえは、部分がたがいに動かしあっている姿だけで、表象について説明するに足るものは、決して発見できはしない。とすると、表象のありかは、複合体や機械のなかではなく、単一実体（モノド）のなかでなくてはならなくなる。」

本稿で述べてきた脳型AIアクセラレータは、必ずしも人間の精神活動までも再現することを目的とするものではないが、人々が「これは脳に似ている」と認める高度な情報処理機能の表象はまさしくライプニッツの言うモノドであり、ライプニッツの論によれば、それは回路の構造や素子の機能では説明できないことになる。

「モノドは、発生も終焉も、かならず一挙に行われる」とライプニッツは主張する。開発しているものが精神的な働きも含めて脳型AIと呼べるものなのかどうかも、それが具現化し動作したとき初めて判明することなのかもしれない。

(参考文献)

- ・「ライプニッツ モノドロジー／形而上学叙説」清水富雄、武田篤司、飯塚勝久 訳、中央公論社

4.6 時間軸

脳型AIアクセラレータの実現を目指す取り組みは、長期的に継続される必要がある。例えば、深層学習が注目されるようになったのは、コンピュータによる物体認識の精度を競う国際コンテスト「ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2012」で、トロント大学のSuperVisionチームが格段に高い認識率を達成して優勝した時からと考えられるが、そこで使われている基本的なアルゴリズムの大部分は1990年代に完成していた。また、誤差逆伝搬法は1990年代に研究され、福島邦彦らのネオコグニトロン

[K. Fukushima et al., Pattern Recognition, 15, 455, (1982)] にルーツを持つ畳み込みニューラルネットワークも1990年代終わりには完成していた[Y. LeCun et al., Proceedings of IEEE, 86, 2278 (1998)]。深層ニューラルネットワークの真価が発揮されるのに時間を要したのは、事前学習に必要な一定以上の規模のデータが手に入るインフラが必要だったこと、またそれら処理できるGPUやマルチコアCPUなどのハードウェアが登場する必要があったためと考えられている。脳型AIアクセラレータも、深層学習と似たような形で世に登場する可能性がある。基礎的な研究を今から長期的かつ着実に継続し、脳型AIアクセラレータが幕を跳ね上げ登場できる時代に備えておくことが重要である。

一方で、大きな変革をもたらす脳型AIアクセラレータの早期登場のためには、その技術の発展性や革新性、社会的な波及効果を期待させる成果を出し続け、周辺技術の底上げや研究開発人材の育成、産業界の参加促進により、研究開発を加速させていくことが必要である。例えば、深層学習がこのように発展する前に、将棋の世界で強いソフトが次々と開発・改良されていき、ついには名人を打ち負かすようになり、世界的にも注目されるようになったことなどがある。このような形で脳型AIアクセラレータの研究開発を進めていくためには、これまでに得られている脳機能の数理モデルとデバイス材料の最適な組み合わせを考えていくことや、現状の深層学習プロセスの一部を新たなハードウェアで実行して、高速化や低消費電力化を示していくことが重要である。

今後取り組んでいく研究開発の短期的（5年程度）な目標としては、エッジ向けに高効率・低消費電力の学習やリアルタイムの認識・判断ができることを目指し、スパイクニューロンモデルを基にしたニューロモルフィックチップや、物理リザーバーによる機械学習チップの開発が考えられる。現状でもエッジ・IoT分野の応用を中心に、デジタル回路のスパイクニューラルネットワークが研究開発されているが、今後、不揮発メモリやハードウェア的に可塑性を持つデバイス材料を用いたアナログ回路により、ニューロン・シナプスを模倣した大規模なニューロモルフィックチップが登場し、消費電力および規模の上でこれまでのデジタル回路に対する優位性が示されると期待される。また、FeFETやスピン、光などを用いた物理リザーバーコンピューティングの研究が現在活発化してきており、これらの技術を高度化することで、深層学習の一部の層を置き換えて高速化・低電力化を図ることも考えられる。そして、このような例題を基に、脳型AIアクセラレータ技術が現状の深層学習、人工知能技術が抱える問題の一部を解決していくことを示していくことと、脳科学、数学・数理科学、情報科学、ナノテクノロジー材料分野にまたがる異分野連携のコミュニティや体制を構築していくことが重要である。

長期的（5年～10年）な目標としては、エッジおよびクラウドでの高効率な学習、転移学習、外挿的な予測、環境変化に合わせた判断、科学的発見など高度な機能の実現を目指して、脳細胞レベルや神経ネットワークレベルだけではなく、脳組織レベル、グローバルなネットワークといった高次の脳機能の知見を活用した新たな情報処理モデルの構築、それによるニューロモルフィックアクセラレータ、物理リザーバーコンピューティングアクセラレータ、記憶と演算が一体化した脳型記憶回路、などの開発が考えられる。また、多様なセンサからの情報を収集・選別してリアルタイムに判断し、アクチュエータに適切な指示を出し、一連の動作からリアルタイムの学習ができるチップの開発なども期待される。これらのためには、脳全体での柔軟な判断、低エネルギー消費の機構に基づく新たな情報処理モデルが必要になるとともに、大規模な3次元結合、配線・結合制御、不揮発演算回路、などの新たなデバイス技術、回路技術などが必要になると考えられる。このような技術が開発されていくことで、人間並みの高度な情報処理が可能なクラウドレベルでの汎用人工知能への展開も期待される。

これらの活動を後押しして、日本における脳型AIアクセラレータ関連の知識や技術力を高めていくためには、

ファイディングの充実が望まれる。付録2の図S2-1に示すように、これまで脳科学と情報科学の間を繋げてきた2つの新学術領域は終了することから、これまでの成果や知識、人脈を活かす形での継続的な基礎研究の推進が必要と考えられる。また、プロジェクトの一部に脳型AIのハードウェアに関する研究が含まれるCREST・さきがけ「革新的コンピューティング」はすでに公募が終了している。NEDOの「高効率・高速処理を可能とするAIチップ・次世代コンピューティングの技術開発」は2027年度まで継続されるが、現状ではAIエッジコンピューティングや量子コンピュータなどの割合が多く、脳型AIに関するテーマは少ない。このような状況にあるため、脳型AIアクセラレータに関する新たな目的基礎研究のプロジェクトの推進や、NEDOプロジェクトの中で脳型AIに関するテーマを拡充（2023年度に事業見直し予定）していくことが望まれる。

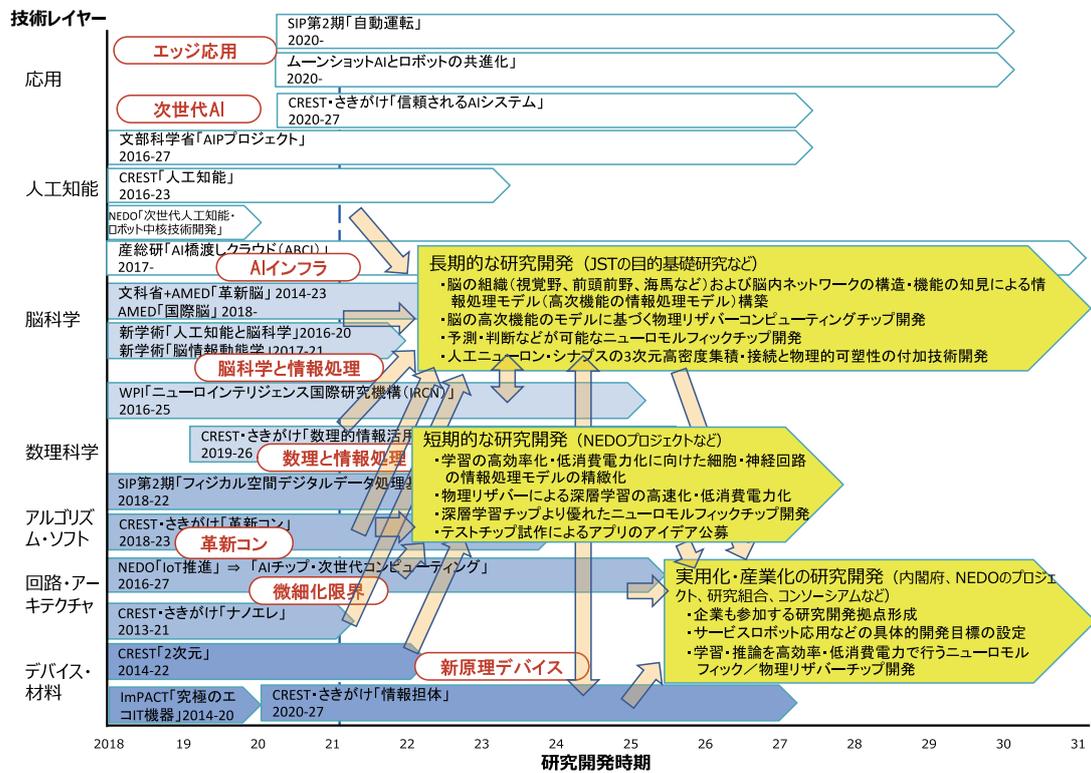


図4-2 脳型AIアクセラレータ関連のプロジェクト推進案

このような状況を踏まえ、図4-2に日本における脳型AIアクセラレータに関するプロジェクトの推進案を示す。これまでの関連するプロジェクトの成果や知見を活かす形（黄色の矢印）で、新たなプロジェクトおよびファイディングが必要と考え、黄色の3つのブロック内に記述するように、長期的な研究開発、短期的な研究開発、および実用化・産業化の研究開発について時間軸に沿った推進案を示している。長期的な研究開発については、JSTの目的基礎研究のプロジェクトが適当と考えられる。このプロジェクトでは脳の組織（視覚野、前頭前野、海馬など）および脳内ネットワークの構造・機能の知見による情報処理モデル（高次機能の情報処理モデル）の構築、脳の高次機能のモデルに基づく物理リザバーコンピューティングチップの開発、予測・判断などが可能なニューロモルフィックチップの開発、人工ニューロン・シナプスの3次元高密度集積・接続と

物理的可塑性の付加技術開発、などを行う。ここでは、これまで行われてきた2つの新学術領域やCREST・さきがけ「革新的コンピューティング」、WPIなどの研究成果や知見などの活用が期待される。この目的基礎研究のプロジェクトにおいては具体的な応用を目指してこれらのいくつかの研究開発課題を含む長期的な目標を掲げて、例えば図4-1の脳の各階層（レベル）に対応して複数の技術レイヤーを跨るグループを構成すること、さらに個別にこれらの研究開発課題の中で優れた脳型AIアクセラレータ開発につながる挑戦的なテーマを実施することが考えられる。ここでは、グループ間での知識・成果の共有に加え、各グループの同一技術レイヤーの研究者やさきがけの研究者が連携して、各技術レイヤーでの新たな学術的な流れを作っていくことも期待される。

短期的な研究開発としては、NEDOのプロジェクトなどが適当と考えられる。学習の高効率化・低消費電力化に向けた細胞・神経回路の情報処理モデルの精緻化、物理リザバーによる深層学習の高速化・低消費電力化、深層学習チップより優れたニューロモルフィックチップ開発、テストチップ試作によるアプリのアイデア公募、などの研究開発課題と取り組みを行う。ここでは、これまで行われてきたCREST・さきがけ「革新的コンピューティング」、NEDO「AIチップ・次世代コンピューティング」などの研究成果や知見などの活用が期待される。NEDO「AIチップ・次世代コンピューティング」はまだ後半が残されており、新規プロジェクトでなくとも、見直し段階でここに挙げた研究開発課題を取り込んでいくことも考えられる。

実用化・産業化の研究開発においては、応用を明確化して企業を含む体制での大型のファンディングが必要であり、内閣府のプロジェクト、NEDOプロジェクトでの推進や、研究組合を含むコンソーシアムでの推進が適当と考えられる。企業も参加する研究開発拠点の形成、サービスロボット応用などの具体的開発目標の設定、学習・推論を高効率・低消費電力で行うニューロモルフィック／物理リザバーアクセラレータの開発などを行う。ここでは、上記の長期的な研究開発、短期的な研究開発、およびそれまでの様々な研究開発プロジェクトの成果の中で企業の関心が高く実用化に近いものを選択し、大きな予算で多くのチップ試作を行いながら従来チップよりも優れた性能を実証し、応用システムでの有効性確認や新たな応用領域の開拓などを進めていくことが重要である。

このようなファンディング、プロジェクトの実施に当たっては、内閣府、文科省、経産省など関係府省や、NEDO、JSTなどのファンディング機関との間で、脳型AIアクセラレータに関するプロジェクトの研究開発の状況を共有するガバニングボードや推進協議会などの枠組みを考えておくことも今後必要と考えられる。

4

付録1 検討の経緯

- ・JST 研究開発戦略センター（CRDS）では、令和2年度に戦略プロポーザルを作成すべきチームの候補をCRDS戦略スコープ2020検討委員会を経て指定し、令和2年5月に検討チームを発足させた。その後、検討チームにおいて提言作成へ向けた調査・分析・検討を重ねた。
- ・チームの活動では、調査によって国内外の研究開発動向・技術水準を明らかにしながらスコープの焦点を絞り、その過程において提言の方向性を検討するため、以下の有識者へのインタビュー・意見交換、およびセミナー講演を実施した。
- ・その上で、脳型AIアクセラレータに関してCRDSが構築した仮説を検証する目的で、科学技術未来戦略ワークショップを開催した。ワークショップの結果は報告書として、令和3年2月にCRDSより発行している（CRDS-FY2020-WR-07）。
- ・CRDSでは以上の調査・分析の結果と、ワークショップにおける議論などを踏まえて、令和3年3月に本戦略プロポーザルを発行するに至った。

(1) インタビューとセミナー

■インタビュー・意見交換を実施した識者（50音順、敬称略、所属・役職は実施時点）

合原 一幸	東京大学 特別教授
秋永 広幸	産業技術総合研究所 ナノエレクトロニクス研究部門 総括研究主幹
浅井 哲也	北海道大学大学院情報科学研究科 教授
井上 淳樹	元富士通研究所 特任研究員
大木 研一	東京大学大学院医学系研究科 機能生物学専攻統合生理学 教授
岡田 隆三	東芝研究開発センター 研究主幹
尾形 哲也	早稲田大学 基幹理工学部 教授
小川 宏高	産業技術総合研究所 人工知能研究センター チーム長
葛西 誠也	北海道大学 量子集積エレクトロニクスセンター 教授
加納 敏行	日本電気株式会社 データサイエンス研究所 主席技術主幹
神谷 之康	京都大学大学院情報学研究科 教授
久保田 均	産業技術総合研究所 エレクトロニクス・製造領域 総括研究主幹
河野 崇	東京大学 生産技術研究所 情報・エレクトロニクス系部門 教授
西郷 甲矢人	長浜バイオ大学 教授
杉本 英樹	デンソー電子基盤技術統括部 室長
高木 信一	東京大学大学院工学系研究科 教授
田中 剛平	東京大学大学院工学系研究科 電気系工学専攻 特任准教授
常木 澄人	産総研スピントロニクス研究センター 主任研究員
銅谷 賢治	沖縄科学技術大学院大学 教授
トープラサートボン カシディット	東京大学大学院工学系研究科 助教
豊泉 太郎	理化学研究所 脳神経科学研究センター チームリーダー
中嶋 浩平	東京大学大学院情報理工学系研究科 特任准教授
中野 大樹	日本IBM 東京基礎研究所
西 義史	東芝研究開発センター
尾藤 晴彦	東京大学 医学系研究科 脳神経医学専攻 教授

- 平岡 裕章 京都大学 理学部数学科 高等研究センター長・教授
 平木 敬 (株)Preferred Networks シニアリサーチャー、東京大学 名誉教授
 深見 俊輔 東北大学 電気通信研究所 教授
 福島 俊一 JST CRDS フェロー
 福水 健次 統計数理研究所 統計的機械学習センター センター長
 堀尾 喜彦 東北大学電気通信研究所 教授
 丸山 宏 (株)Preferred Networks フェロー
 三宅 陽一郎 日本デジタルゲーム学会 理事
 森江 隆 九州工業大学大学院生命体工学研究科 教授
 山川 宏 全脳アーキテクチャ・イニシアティブ 代表
 山根 敏志 日本IBM 東京基礎研究所 主任研究員
 古田 雅則 東芝研究開発センター

■セミナー講演を依頼した識者（50音順、敬称略、所属・役職は実施時点）

- 浅井 哲也 北海道大学大学院情報科学研究科 教授
 河野 崇 東京大学 生産技術研究所 教授
 銅谷 賢治 沖縄科学技術大学院大学神経計算ユニット 教授
 森江 隆 九州工業大学大学院生命体工学研究科 教授
 山川 宏 全脳アーキテクチャ・イニシアティブ 代表

(2) ワークショップ

科学技術未来戦略ワークショップ「脳型AIアクセラレータ ～柔軟な高度情報処理と超低消費電力化の両立～」

開催日時：2020年11月28日（土）9:30～17:00

開催会場：TKP市ヶ谷カンファレンスセンター バンケットホール8Bおよびオンライン会議

(敬称略)

- | | | |
|-------------|--------------------------|------------------|
| 9:30～9:35 | 開会挨拶 | 曾根 純一 (JST-CRDS) |
| 9:35～9:55 | ワークショップの開催趣旨と骨子案の説明 | 馬場 寿夫 (JST-CRDS) |
| 9:55～11:10 | (話題提供1) 高効率・高度情報処理に対する要求 | |
| 9:55～10:20 | 人工知能技術開発のための高速計算基盤「ABCI」 | 小川 宏高 (AIST) |
| 10:20～10:45 | リザーブ・コンピューティングとエッジAIへの応用 | 山根 敏志 (日本IBM) |
| 10:45～11:10 | 予測誤差最小化モデルを基盤としたロボット動作学習 | 尾形 哲也 (早稲田大学) |

11:20～12:35 (話題提供2) 高度情報処理に向けた脳科学、人工知能、数理科学の研究動向

11:20～11:45 情報処理の活用に向けた脳科学分野の進展

豊泉 太郎 (理化学研究所)

11:45～12:10 脳のふるまいに倣う人工知能技術”ゆらぎ学習”研究開発への取り組み

加納 敏行 (NEC)

12:10～12:35 高次元ビッグデータ活用のための数理研究

合原 一幸 (東京大学)

14:00～15:15 (話題提供3) AIチップ開発

14:00～14:25 ブレインモルフィックコンピューティング

堀尾 喜彦 (東北大学)

14:25～14:50 新デバイス材料のAIチップへの応用

秋永 広幸 (AIST)

14:50～15:15 強誘電体FeFETを利用したリザーバコンピューティング

高木 信一 (東京大学)

15:25～16:55 総合討論

ファシリテーター：東 良太 (JST-CRDS)

1. 我が国が注力すべき脳型AIの研究領域 (短期、中長期)
2. 脳科学、数学・数理科学、情報科学、ナノテク・材料技術の異分野を連携させる仕組み (脳科学と情報科学との密な連携、研究開発フェーズ、時間のずれへの対応など)
3. アカデミアと産業界との連携の仕組み
4. その他 (コミュニティ形成、人材育成、海外連携)

16:55～17:00 閉会挨拶

曾根 純一 (JST-CRDS)

招聘識者

(敬称略)

(発表者)

- ・小川 宏高 産業技術総合研究所 人工知能研究センター チーム長
- ・山根 敏志 日本IBM 東京基礎研究所 主任研究員
- ・尾形 哲也 早稲田大学 基幹理工学部 教授、産業技術総合研究所
人工知能研究センター 特定フェロー
- ・豊泉 太郎 理化学研究所 脳神経科学研究センター チームリーダー
- ・加納 敏行 日本電気株式会社 データサイエンス研究所 主席技術主幹、
大阪大学 NEC Brain Inspired Computing 協働研究所 副所長
- ・合原 一幸 東京大学 国際高等研究所 IRCN 特別教授
- ・堀尾 喜彦 東北大学 電気通信研究所 教授
- ・秋永 広幸 産業技術総合研究所 デバイス技術研究部門 総括研究主幹
- ・高木 信一 東京大学 大学院工学系研究科 教授

(コメンテータ)

- ・平木 敬 (株)Preferred Networks シニアリサーチャー、東京大学 名誉教授
- ・森江 隆 九州工業大学 大学院生命体工学研究科 教授
- ・浅井 哲也 北海道大学 大学院情報科学研究院 教授
- ・銅谷 賢治 沖縄科学技術大学院大学 教授

付録2 国内外の状況

(1) 日本の状況

日本の主なプロジェクトと研究拠点の状況について図S2-1に示し、それぞれの活動状況について簡単に記載する。

統合イノベーション戦略推進会議「AI戦略2019」(2019年)

AI戦略2019における研究開発戦略は、世界をリードできる次世代AI基盤技術の確立を目的に、研究開発重点項目として「基礎理論」、「高品質かつ信頼できるAI」、「AIのシステム・コンポーネント」とともに、「AIのためのデバイス・アーキテクチャ」の研究開発の必要性が記載されており、この中で「脳を模倣した情報処理を実用的な時間で実現するアーキテクチャの開発」、「情報処理に係る消費電力性能を従来比10倍以上に向上させる革新的AIチップ技術の確立」が挙げられている。

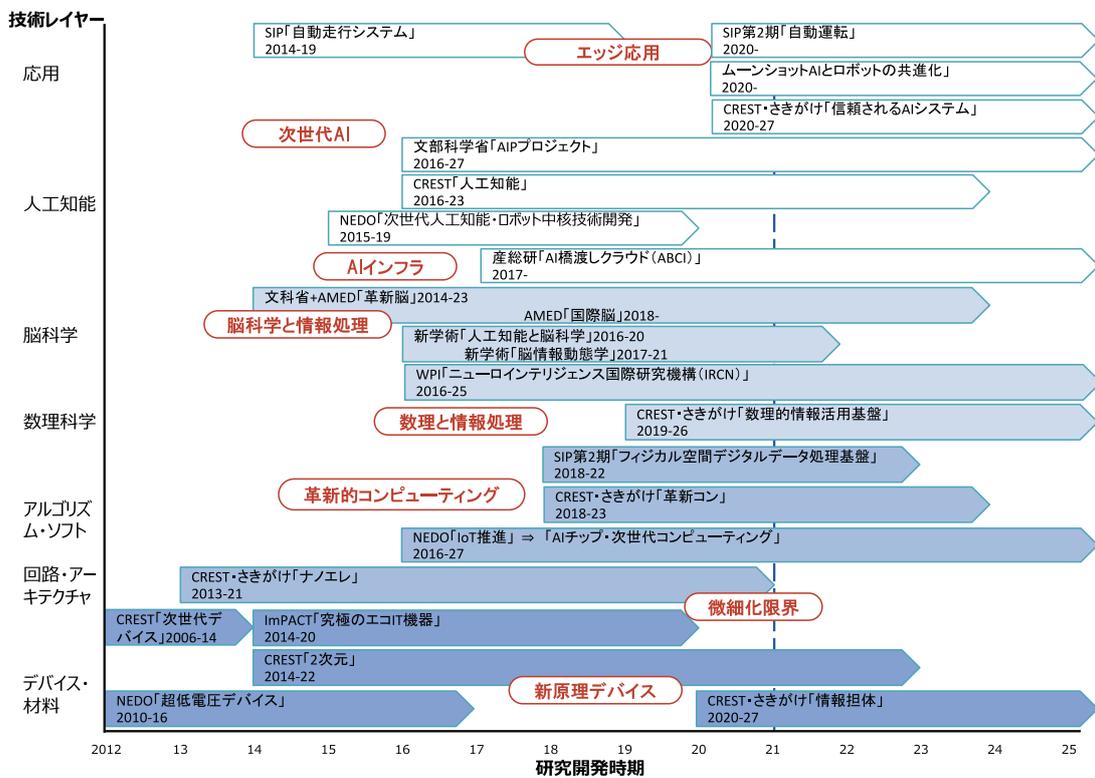


図 S2-1 脳型 AI アクセラレータ関連の日本の主なプロジェクト

新学術研究領域「人工知能と脳科学の対照と融合（人工知能と脳科学）」(2016～2020年度)

感覚運動情報の背後にある構造を捉える表現学習、それらの変化を予測する内部モデル学習、さらに予測結果の評価による探索学習について、それぞれを確実に効率よく実現するアルゴリズムとその脳での実現を探るとともに、それらをつないだ全脳アーキテクチャにならった柔軟な人工知能システムの実現を目指している。

新学術研究領域「脳情報動態を規定する多領域連関と並列処理（脳情報動態学）」(2017～2021年度)

脳機能の多次元活動動態を先端的計測操作技術により解明・再現し、記憶・予測・判断に基づく行動原

理を明らかにすることを目指している。

文科省+AMED 革新的技術による脳機能ネットワークの全容解明（革新脳：2014～2023、400億円／10年）

霊長類（マーモセット）の高次脳機能を担う神経回路の全容をニューロンレベルで解明することにより、ヒトの精神・神経疾患の克服や情報処理技術の高度化に貢献する。姉妹プログラムである国際脳とも連携している。

AMED 戦略的国際脳科学研究推進プログラム（国際脳）

AMEDで推進されてきた脳科学研究事業や国際的な脳プロジェクト（Brain Initiative, Human Brain Projectなど）との連携を強化しつつ、健常から疾患に至る脳画像等の総合的解析、AIによる脳科学技術開発、ヒトと非ヒト霊長類動物との神経回路比較研究を推進。次世代AIに関する基盤技術開発として「非線形動力学に基づく次世代AIと基盤技術に関する研究開発」が推進されている。

全脳アーキテクチャイニシアティブ（WBAI）

脳全体のアーキテクチャに学ぶことにより人間のような知的能力を持つ汎用人工知能の実現に向けて、研究開発活動への支援等を行っているNPO法人。脳の各器官を機械学習モジュールとして開発し、それら複数の機械学習モジュールを脳型の認知アーキテクチャ上で統合するという2つのアプローチにより2030年頃の汎用人工知能の構築を目指している。

東京大学 ニューロインテリジェンス国際研究機構（IRCN）

世界トップレベル研究拠点プログラム（WPI）の1つとして2017年に設立。神経回路発達の基礎研究、精神疾患の病態研究、神経の動作原理に基づく革新的な人工知能（AI）の開発を目指している。

北海道大学 人間知×脳×AI研究教育センター（CHAIN）

人文社会科学、脳科学（神経科学）、AI研究の分離融合型学際研究と大学院教育を目的として2019年に設立。

(2) 米国の状況

Brain Initiative（2013～、\$4.5B[2016年～2025年]）

神経疾患や精神疾患の克服に向けて、革新的ツールの開発と応用を進めることにより、脳の個々の細胞と複雑な神経回路が時間的、空間的にどのように相互作用するのかを理解し、さらに、脳が情報の記録、処理、利用、保存をどのように可能にするか解明することを目指している。

例えば、MITは加齢にフォーカスして、記憶の仕組みを直接ニューロンにアプローチして調べ、老化した脳の謎を解明する。神経科学、基礎生物学と遺伝学、調査医学、工学とコンピューターサイエンス、経済学、化学、都市計画、人工知能にまたがり、包括的なシステムアプローチで研究している。

DARPA Lifelong Learning Machines (L2M) プログラム（2018—2021年）

コンピュータ科学者と生物学者とが連携した分野融合的なアプローチにより、過去学習したことを忘れるこ

となく、システムが継続的に新しい状況に適応していけるような、まったく新しい機械学習手法の実現を目指している。例えば、海馬と皮質の二重記憶構造の研究から入力を既存の記憶と比較することによって潜在的な結果を予測することができる機械学習システム、サラマンダーのような動物の再生メカニズムを調べることで環境の変化に適応するために構造と機能を変更することができる柔軟なロボット、といった研究が実施されている。

(3) 欧州の状況

Human Brain Project (2013-2023年、\$1.3B/10年)

6つの情報通信技術（ICT）基盤研究プラットフォームを中心として、ICTを用いた脳の理解を目指す。脳機能と脳疾患の網羅的データベースの構築と人工脳の結合や脳の基礎理論とモデルの構築、脳に着想を得たコンピュータやロボット工学など広範な領域をカバーしている。ニューロモルフィックのコンピューティングシステムとして、アナログ型で高速なハイデルベルグ大学（ドイツ）のBrainScaleS、デジタル型でリアルタイム性の高いマンチェスター大学（イギリス）のSpiNNakerが提供されている。

(4) 中国の状況

次世代人工知能発展計画（2017—2030年）

2017年に国務院が発表したAI技術の開発計画で、2030年までに理論・技術・応用を世界トップ水準に引き上げることなどが目標に掲げられている。この計画に基づき、百度（バイドゥ）が自動運転、阿里雲（アリババ・クラウド）が公共サービスシステム、騰訊（テンセント）が医療画像、科大訊飛（アイフライテック）が音声識別について、それぞれ中心となってオープンイノベーションプラットフォームの構築を行うこととなっている。

科学技術イノベーション2030・重大プロジェクト「脳科学・脳関連研究」

今後2030年まで中国が取り組む分野として、2016年に発表した16プロジェクトの一つで、中国脳計画（China Brain Project）とも呼ばれる。2017年に実施計画の策定準備が開始されているが、2020年12月現在、開始は確認できていない。脳の認知原理の研究を中心とし、脳型コンピュータや脳型AI開発と脳疾患の診断・治療法開発の2つを目的とする研究の展開を計画している。2018年に北京と上海にそれぞれ設立された脳科学及び脳関連研究センター（北京脳科学と類脳研究中心、上海脳科学と類脳研究中心）が中核を担うこととなっている。

脳型知能技術・応用国家技術実験室

2017年合肥市に中国科学技術大学が中心となって、復旦大学、中国科学院、バイドゥ（百度）などと共同設立。脳型知能技術応用・研究プラットフォームを共同設立し、脳型認知・神経計算、脳型マルチモーダル知覚・信号処理、脳型チップ・システム、脳型コンピューティングシステム、量子人工知能、脳型知能ロボットの6つの方向について研究開発と実用化を推進する。

清華大学 脳型コンピュータ研究センター

2014年、清華大学の7学部（脳科学、電子機器、マイクロ電子機器、コンピュータ、自動化、材料、および精密機器）により設立。基礎理論から、チップ、ソフトウェア、システム、およびアプリケーションまで

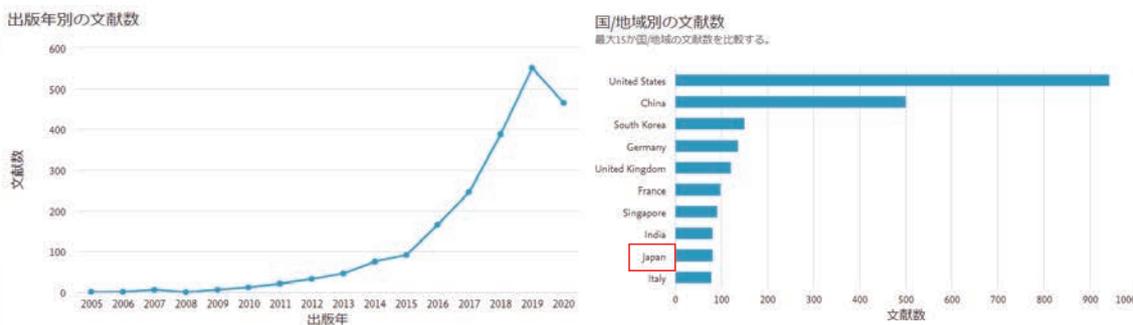
研究開発を実施。センター長の施路平のチームは、2019年脳型コンピューターチップ「天機芯（Tianjicチップ）」を発表。人工ニューラルネットワークとパルスニューラルネットワークのいずれにも対応可能な世界初のハイブリッドチップとして、Natureの表紙に採用された。

(4) 論文の推移

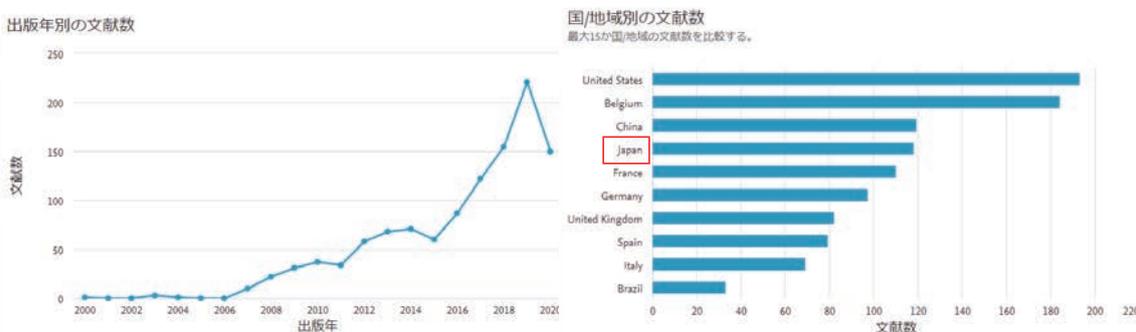
図S2-2に脳型AIアクセラレータの研究として中心的なものとなっているニューロモルフィックコンピューティングとリザーバーコンピューティングに関する論文の2000年から2020年までの出版年別と国別の論文数のデータを示す。

出版年別の論文数の推移をみると、両技術とも2010年頃から増加しており、これらの研究が活発になっていることがわかる。また、国別で見ると、ニューロモルフィックコンピューティングでは、米国、中国がリードしており、日本は9位となっている。リザーバーコンピューティングでは、米国の1位は変わらないが、ベルギーが2位、中国が3位、日本が4位になっており、まだ国別の有意な差がついていない状況にある。

検索式：“neuromorphic comput*”、2000～2020年（2107件）



検索式：“reservoir comput*”、2000～2020年（1131件）



図S2-2 脳型AIアクセラレータに関する論文数の推移と国別文献数

以下のキーワードを選び、2000～2020年についてエルゼビアのSCOPUSで検索

- (1) ニューロモルフィックコンピューティング (“neuromorphic comput*”)
- (2) リザーバーコンピューティング (“reservoir comput*”)

付録3 専門用語解説

■ノイマン型（コンピュータ）

数学者ジョン・フォン・ノイマン (John von Neumann) によって提唱されたコンピュータの基本構成のこと。ノイマン型コンピュータは、演算装置、制御装置、メモリ、入力装置、出力装置とこれらの間を接続するデータ伝送機構から構成されている。今日の一般的なコンピュータシステムのほとんどが、このノイマン型である。また、ノイマン型コンピュータの構造以外の構造をとるコンピュータアーキテクチャは、非ノイマン型コンピュータと総称される。

■アクセラレータ

特定用途向けの計算・情報処理を行う専用のプロセッサ。汎用プロセッサでは計算負荷の大きな処理をアクセラレータに渡すことで、アクセラレータが高速・低消費電力にその処理を実行し、結果を汎用プロセッサに返す。

■メモリスタ

通過した電荷を記憶して抵抗が変化する素子であり、次世代不揮発性メモリとして期待されている。

■機械学習

データの背後に潜む規則性や特異性を発見することにより、人間と同程度あるいはそれ以上の学習能力をコンピュータで実現しようとする技術である。教師あり学習、教師なし学習、強化学習といったタイプがあり、各タイプについて、さまざまな方式（深層学習、サポートベクトルマシン、決定木、回帰分析、他）が開発されている。

■深層学習

多層ニューラルネットワーク（Deep Neural Network）を用いた機械学習技術である。

■インメモリコンピューティング

外部記憶装置を用いず全てのデータをメモリ上に持つことで、高速処理を実現する技術の総称である。

■ニューロモルフィックコンピューティング

機械学習ハードウェアには大きく二つ、ニューロモルフィックコンピューティング（Neuromorphic Computing）とニューラルネットワークがある。ニューロモルフィックコンピューティングは脳のシナプスの機能と発火の仕組みを模倣したアプローチである。脳神経系を構成する神経ネットワークの神経活動をリアルタイムあるいはそれ以上の速度で再現する回路システムである。

■リザバーコンピューティング

リザバーコンピューティングは再帰的ニューラルネットワークの特殊なモデルを一般化した概念で、時系列情報処理に適した機械学習の枠組みの一つである。入力信号をリザバーで変換した後、読み出し部のみを簡便なアルゴリズムで訓練することで高速な学習を可能となる。

■スパイクングニューラルネットワーク

ニューロンとシナプスで構成される脳の神経ネットワークの挙動を模倣するモデルである。情報はニューロンの膜電位の変化とそれに伴う発火パターンで表現される。

■ストカスティック演算回路

近似計算法の一つであり、デジタル回路で確率的な計算を行う。耐故障性や小面積・低電力性から再注目されており、画像処理や機械学習などでの利用が期待されている。

■畳み込みニューラルネットワーク (CNN)

ニューラルネットワークの中間層に畳み込み層とプーリング層を組み込んだものである。画像分析のための学習手法の一つで、一部が見えにくくなっているような画像でも解析することが可能である。

■ムーアの法則

米インテル社の創業者のひとりであるゴードン・ムーアが1965年に論文で発表したのでこの名がある。経験則にもとづく将来予測であり、1975年には次の10年を見据えて集積回路の部品数が2年ごとに2倍になるという予測をした。

■フォン・ノイマン・ボトルネック問題

コンピュータアーキテクチャの一つであるノイマン型に存在する性能上のボトルネックを指す。ノイマン型では、記憶装置に命令を格納するので、プロセッサが命令を実行するには必ずバスを通して記憶装置にアクセスしなければならず、プロセッサと記憶装置のアクセス速度が遅ければコンピュータ全体のボトルネックになる。

■シグモイド関数

生物の神経細胞が持つ性質をモデル化した関数である。シグモイド曲線がギリシャ文字のシグマに似た形なのでこの名がある。

■圏論

数学的構造とその間の関係を抽象的に扱う数学理論の1つである。考えている種類の「構造」を持った対象とその構造を反映するような対象間の射の集まりからなる圏が基本的な考察の対象になる。

■Izhikevich モデル

神経細胞の挙動を模倣するモデルの一つで、計算負荷の軽減と挙動の多様性において理想に近い数理モデルと言われている。2003年にユージン・イジケヴィッチ (Eugene M. Izhikevich) が提唱した。

■IPコア

LSIを構成するための部分的な回路情報で、機能単位でまとめられているものである。LSIが大規模複雑化する中、開発手法としてハードウェア記述言語による開発が盛んになり、開発効率の向上が求められた。そこで、既存開発製品の回路を、機能ブロック単位で再利用可能な形にまとめ、他の製品でも利用可能な部分はそれを流用する方法が用いられた。

■誤差逆伝搬法

バックプロパゲーションとも言われ、機械学習においてニューラルネットワークを学習させる際に用いられるアルゴリズムである。

■ネオコグニトロン

1979年に福島邦彦（当時：NHK放送科学基礎研究所）によって提唱された階層的、多層化された人工ニューラルネットワークである。手書き文字認識やその他のパターン認識に用いられており、畳み込みニューラルネットワークの元となった。

■大脳皮質

大脳の表面に広がる、神経細胞の薄い層であり、知覚、言語、思考、推理、記憶などの脳の高次機能を司る組織であり、分子層、外顆粒層、外錐体細胞層、内顆粒層、内錐体細胞層、多型細胞層の6層構造を持つ。

■海馬

脳の内部にある古い大脳皮質の部分であり、短期・長期記憶に関係し、様々な感覚入力に応じて時間・空間情報の認知・学習に関わる。

■扁桃核

側頭葉内側の奥に存在し、恐怖、不安、悲しみ、喜び、など情動の処理、価値判断、記憶において主要な役割を持つ。

■前頭前野

前頭葉の前側の領域にあり、記憶や感情の制御、行動の抑制、推論などの認知・実行機能、高次な情動・動機づけ機能とそれに基づく意思決定過程などを担っている。

作成メンバー

総括責任者	曾根 純一	上席フェロー	(ナノテクノロジー・材料ユニット)
リーダー	馬場 寿夫	フェロー	(ナノテクノロジー・材料ユニット)
メンバー	渡邊 孝信	フェロー	(ナノテクノロジー・材料ユニット)
	東 良太	フェロー	(システム・情報科学技術ユニット)
	井上 眞梨	フェロー	(システム・情報科学技術ユニット)
	山原 恵子	フェロー	(ライフサイエンス・臨床医学ユニット)
	井上 貴文	特任フェロー	(ライフサイエンス・臨床医学ユニット)
	伊藤 顕知	主任調査員	(未来創造研究開発推進部)
	勝又 康弘	主任調査員	(戦略研究推進部 グリーンイノベーショングループ)
	上田 幸弘	主任調査員	(戦略研究推進部 先進融合研究グループ)

戦略プロポーザル

CRDS-FY2020-SP-04

脳型AIアクセラレータ

～柔軟な高度情報処理と超低消費電力化の両立～

STRATEGIC PROPOSAL

Brain-Inspired AI Accelerator

Brain-Inspired AI Accelerator - Realizing flexible advanced information processing with ultra-low power consumption -

令和3年3月 March, 2021

ISBN 978-4-88890-726-2

国立研究開発法人科学技術振興機構 研究開発戦略センター

Center for Research and Development Strategy, Japan Science and Technology Agency

〒102-0076 東京都千代田区五番町7 K's 五番町

電話 03-5214-7481

E-mail crds@jst.go.jp

<https://www.jst.go.jp/crds/>

本書は著作権法等によって著作権が保護された著作物です。

著作権法で認められた場合を除き、本書の全部又は一部を許可無く複写・複製することを禁じます。

引用を行う際は、必ず出典を記述願います。

This publication is protected by copyright law and international treaties.

No part of this publication may be copied or reproduced in any form or by any means without permission of JST, except to the extent permitted by applicable law.

Any quotations must be appropriately acknowledged.

If you wish to copy, reproduce, display or otherwise use this publication, please contact crds@jst.go.jp.

FOR THE FUTURE OF
SCIENCE AND
SOCIETY



<https://www.jst.go.jp/crds/>