

CRDS-FY2017-SP-02

戦略プロポーザル 革新的コンピューティング

～計算ドメイン志向による基盤技術の創出～

STRATEGIC PROPOSAL

Disruptive Computing

- A Computing-Domain-Oriented Approach -



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

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っ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。

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

エグゼクティブサマリー

「革新的コンピューティング ～計算ドメイン志向による基盤技術の創出～」とは、物理的な世界（フィジカル空間）とコンピュータやネットワーク上のサイバー空間とが融合したサイバーフィジカルシステム（Cyber Physical System : CPS）の実現に必要な高度情報処理が可能なコンピューティング技術の研究開発課題と推進方法に関するものである。

CPSに必要な高度な情報処理システムの実現には、コンピューティングのこれまで以上の高速化・低消費電力化とともに、ネットワークの末端（エッジ）側の機器におけるリアルタイム性やロバスト性、認識・判断といった人工知能（AI）技術、クラウド側に価値ある情報を安全に送るデータ圧縮・暗号化技術などの高度な情報処理技術が必要になる。

一方、現在の情報処理性能向上の担い手である半導体集積回路は微細化の限界に直面しており、従来の CMOS デジタル回路によるノイマン型のコンピューティング技術だけではさらなる性能向上は困難になっている。このため、従来のコンピューティングのアルゴリズムやアーキテクチャにとらわれない新たなコンピューティング技術が望まれる。

近年はネットワーク・Web の拡大、スマートフォンなど携帯機器の普及、コンピュータの飛躍的な性能向上などにより、ビッグデータ処理による購買分析や材料探索、深層学習（Deep Learning）を活用した画像認識や病気診断など、新たなサービス創出の動きが活発である。これらはクラウド側で行われているが、今後は IoT（Internet of Things）の普及により、エッジ側におけるセンサからの生データの解析と価値あるデータへの変換、安全なデータ転送（プライバシー保護、セキュリティ確保）、認識・判断（画像認識、音声認識、危険回避など）などが重要になってくる。このようなエッジ側における情報処理は、超低消費電力性、リアルタイム性、ロバスト性などクラウド側とは異なる特性が要求されるため、これらを実現する新たなコンピューティング技術に対する期待は大きい。

海外では、米国 DARPA の SyNAPSE プログラムによるニューロモルフィックチップ開発、欧州の Human Brain Project による脳研究と脳を模倣したコンピューティング技術開発、新たなコンピューティング技術の研究開発を目指す IEEE の Rebooting Computing Initiative の設立、Google、Microsoft、IBM など企業による深層学習のアクセラレータや量子コンピュータの研究開発など、新たなコンピューティング技術の活動が活発になっている。このため、我が国でもこのコンピューティング技術の新たな流れをチャンスとしてとらえ、革新的コンピューティングの研究開発を推進すべきである。

今後取り組むべきコンピューティング技術の研究開発課題は、アルゴリズム・ソフト、回路・アーキテクチャ、デバイス・材料の全ての技術レイヤーに関わる。しかし、これらの技術を網羅的に行うのではなく、重要な応用に向けて、時間軸を考慮した計算ドメイン志向により各技術レイヤーから適切な技術を選択し、垂直統合的な技術開発を行うことが社会実装や効率性の点から重要である。また、新たに開発した技術や知識を日本の強みとして蓄積していくことが望まれる。これらに対し、以下の2つの研究開発領域を提案する。

（1）ソフト・ハードの垂直統合技術開発と性能検証

特定の技術レイヤーの要素技術開発だけでは、情報処理全体の性能向上・最適化につながらない場合もあり、コンピューティングの全ての技術レイヤーを垂直統合的に見た技術

開発が重要である。特にエッジ側においては、リアルタイムの画像認識、自律制御などの重要な計算ドメインがあり、各計算機能に着目しそれぞれに最適なコンピューティング技術を開発することが重要である。このような計算ドメイン志向により、超低消費電力性やリアルタイム性、ロバスト性などに注目し、ハードウェア技術やシステム化などの日本の強みを活かした統合技術を開発し、性能優位性を検証することが求められる。また、クラウド側においても計算ドメイン志向の技術開発を適用し、データ検索、深層学習、経路最適化などに対して超高速・高効率に行うアクセラレータ技術の開発が求められる。

(2) 新たな共通基盤技術の体系化と各技術レイヤーの強化

上記の計算ドメイン志向により新たに開発された各技術レイヤーのコア技術は、特定の計算ドメインだけでなく、他の計算ドメインにも活用できると考えられる。これらを新たな共通基盤技術として体系化し、各技術レイヤーを強化することが重要である。

アルゴリズム・ソフトにおいては、エッジ側やクラウド側のワークロードに対し、近似計算、ニューラルネット、ニューロモルフィック、量子計算など非ノイマン型の新たなアルゴリズムの適用範囲を明確化し、知識を体系化することが求められる。

回路・アーキテクチャにおいては、新たなアルゴリズムの実行に適するアーキテクチャ技術を蓄積し、他の計算ドメインへの展開に向けた技術のブラッシュアップを行って体系化する。この技術レイヤーはソフトとハードを繋ぐ重要な研究領域であり新たな実装・配線技術やデバイスの活用方法の整理も求められる。また、アルゴリズムやソフトのオープンソースに接続可能で、使い易い設計ツールも開発しておくことが重要である。

デバイス・材料においては、新たなデバイス・材料の性能・機能・信頼性などを整理し、その適用範囲を正しく理解することや、それらの設計指針を明確にしておくことが重要である。この技術領域は日本が強く、コンピューティングの視点からの強化も期待される。

本研究開発の推進に当たっては、計算ドメインごとの研究開発を垂直統合的に行う体制構築が必要である。これには、技術レイヤーの異なる研究者・研究組織をまとめていくプロジェクト・マネージャー（ドメインスペシフィック・アーキテクト）の役割が重要であり、最適なチーム編成を行い、トップダウン的に技術レイヤー間の連携を深めて一体的な研究開発を進めることが求められる。また、産学官の研究者・技術者が手軽にソフト開発、機能設計、回路設計、ツール開発、ハードウェア試作ができる研究環境の整備が必要である。このような研究環境の利用により、効率的な研究開発ができるとともに、産学官連携の促進、様々な技術レイヤーの研究開発に関心を持つ人材の育成なども期待できる。

このような研究開発の推進や研究環境の整備については、新たなコンピューティング技術の革新性・リスク、技術成熟度（Technology Readiness Level：TRL）、実現可能時期などを考慮して、府省横断の戦略的なファンディングを実施することが重要である。関連府省が今後取り組むべき研究開発のポートフォリオを描き、戦略を共有し、効果的な役割分担を行う必要がある。また、産業界の協力を得て技術の知財化、共有化、オープン化、国際標準化などを戦略的に取り組むなど、産学官連携のエコシステムの構築が望まれる。

新たなコンピューティング技術の研究開発は大きな技術トレンドであるため、世界の知識や人材を活用し、アルゴリズム・ソフト、回路・アーキテクチャ、デバイス・材料関連学会の連携による新たな研究分野の創成と新たなコミュニティの形成が望まれる。

Executive Summary

This proposal “Disruptive Computing: A Computing-Domain-Oriented Approach” addresses challenges in R&D on disruptive computing technologies capable of performing the advanced data processing required for cyber-physical systems (CPSs), which closely integrate the physical world (physical space) with cyber space, and the method of implementing such R&D.

Realizing the advanced data processing systems needed for CPSs will require faster speeds and lower power consumption than what current computing can achieve, artificial intelligence (AI) technologies for real-time and robust recognition and decision-making in edge devices, and advanced data processing technologies such as data compression and encryption for safely transferring information of value to the cloud.

In the meantime, semiconductor integrated circuits, which are responsible for today’s advances in data processing performance, are up against limits in miniaturization. Any further improvements in performance will be difficult with a von Neumann computing technology that employs conventional CMOS digital circuits. Consequently, it is desirable to conduct R&D on new computing technologies free from conventional computing algorithms and architectures.

Owing to the expansion of networks and websites, the popularization of smartphones and other mobile devices, and dramatic improvements in the performance of computers in recent years, there is a growing movement to create new services related to purchase analyses and material searches enabled by big data processing and to image recognition, disease diagnosis, and other areas that utilize AI technologies such as deep learning. While the current standard for these services is to perform data processing in the cloud, the increasing prevalence of the Internet of Things (IoT) will make it more important to perform such data processing on the edge, including analysis of raw data from sensors and other devices, conversion of this raw data to data with value, safe data transfers (protecting privacy and ensuring security), and recognition and decision-making (image recognition, speech recognition, danger avoidance, etc.). Since such data processing on the edge requires different attributes from processing in the cloud, including a need for ultra-low power consumption, real-time processing, and robustness, there is much anticipation for new computing technologies capable of implementing these qualities.

Overseas there has been a growing interest in and increased R&D activity on developing new computing technologies, including the development of neuromorphic chips in America’s DARPA SyNAPSE Program, research on the brain and the development of new computing technologies simulating the brain through Europe’s

Human Brain Project, the establishment of the IEEE Rebooting Computing initiative to conduct R&D on new computing technologies, and R&D carried out by such corporations as Google, Microsoft, and IBM on accelerators for deep learning and quantum computing. Japan must take this opportunity to chart a new course for computing technologies by promoting R&D on disruptive computing.

Future R&D challenges in computing technologies that must be addressed involve all technology layers, including algorithms/software, circuits/architecture, and devices/materials. However, rather than including all technologies comprehensively, it will be important from the perspective of social implementation and efficiency to select only technologies suitable for important applications from each technology layer specific to the computing domain and with consideration for the timeline, and to develop technologies through a vertical integration approach. New knowledge and newly developed technologies should be accumulated to eventually serve as Japan's assets. For these purposes, the following two areas of R&D are proposed.

(1) Technological development and performance verification through vertical integration of software and hardware

The development of component technologies in specific technology layers does not always lead to performance improvements or optimization of data processing as a whole. Therefore, technological development must be conducted with a view toward vertically integrating all technology layers of computing. Computing domains such as real-time image recognition and autonomous control are particularly important on the edge, and it is important to focus on each computing function to develop an optimal computing technology for these domains. This type of computing-domain-oriented approach must be utilized to emphasize ultra-low power consumption, real-time processing, and robustness; to develop integrated technologies leveraging Japan's skill in hardware technology and system creation; and to verify performance advantages. Computing-domain-oriented technology development must also be applied in the cloud in order to develop accelerator technologies for achieving ultrafast speeds and high-efficiency in data searches, deep learning, and route optimization.

(2) Organizing new common fundamental technologies and strengthening each technology layer

New core technologies in each technology layer developed through a computing-domain-oriented approach are likely applicable to more than just one specific computing domain. It is essential that these core technologies be organized as new common fundamental technologies to strengthen each technology layer.

Within algorithms and software, it will be necessary to define the scope of application for new algorithms in non-von-Neumann computing, such as approximate computing, neural nets, neuromorphic computing, and quantum computing, to

workloads on the edge and in the cloud, and to codify this knowledge.

For circuits and architecture, it will be important to accumulate architecture-related technologies suitable for implementing new algorithms, and to refine and organize these technologies for use in other computing domains. This technology layer is an important research area for linking software to hardware. As such, it will be necessary to organize methods of utilizing new packaging and interconnect technologies and new devices. User-friendly design tools must also be simultaneously developed with accessible algorithms and open-source software.

In devices and materials, it is essential to organize the performance, functions, and reliability of new devices and materials, accurately learn the range of their applications, and clarify guidelines for their design. Japan is strong in this technological area and is only expected to become stronger from the perspective of computing.

Conducting this R&D will require construction of a vertically integrated system of R&D for each computing domain. Here, project managers (domain-specific architects) that assemble researchers and research organizations from different technology layers will play an important role in ensuring the best team composition and facilitating integrated, top-down R&D with close cooperation among researchers from each technology layer. Establishing a research environment that facilitates researchers and engineers from industry, academia, and government in developing software, designing functions and circuits, developing tools, and manufacturing hardware prototypes will also be important. Use of such a research environment is hoped to improved efficiency in R&D, promote collaboration among industry, academia, and government, and foster human resources with an interest in R&D in various technology layers.

Promoting this style of R&D and setting up a suitable research environment will require strategic cross-ministerial funding with considerations for the innovativeness and risk in new computing technologies, the technology readiness level (TRL), and feasible time frames. The relevant ministries and agencies must draw up a complete portfolio covering short-to-long term plans for R&D and what topics are to be addressed, necessitating open sharing of strategies and an effective division of roles. With the cooperation of industry, an ecosystem for industry, academia, and government collaboration should be constructed to implement strategic efforts on acquiring intellectual property rights, sharing and making available technologies, and establishing international standards.

Since R&D on new computing technologies is expected to continue along its upward trend, it will be desirable to utilize knowledge and human resources from around the world to create new fields of research and to form new communities through collaboration with societies involved in algorithms/software, circuits/architectures, and materials/devices.

目 次

エグゼクティブサマリー

Executive Summary

1. 研究開発の内容	1
2. 提案を実施する意義	5
2-1. 現状認識および問題点	5
2-2. 社会・経済的効果	8
2-3. 科学技術上の効果	10
3. 具体的な研究開発課題	14
3-1. ソフト・ハードの垂直統合技術開発と性能検証	14
3-2. 新たな共通基盤技術の体系化と各技術レイヤーの強化	20
4. 研究開発の推進方法および時間軸	25
4-1. 技術レイヤー間に跨がる統合的研究推進	25
4-2. 産学官連携のエコシステムの形成	26
4-3. 人材育成	27
4-4. 時間軸	28
付録1. 検討経緯	30
付録2. 国内外の状況	35
付録3. コンピューティング技術が重要な応用領域	39
付録4. 専門用語解説	43

1. 研究開発の内容

これからの社会の課題として、少子高齢化に伴う労働力不足の解消、高齢者・障がい者の介護・自立支援、安全で便利な移動手段の確保、ものづくりの効率化、社会インフラの効率的な保全、セキュリティ強化、プライバシーの確保などが顕在化している。これらに対し、スマートロボット、自動運転、IoT（Internet of Things）などの物理的な世界としてのフィジカル空間とコンピュータやネットワーク上におけるデータ、情報などからなるサイバー空間とが高度に融合したサイバーフィジカルシステム（Cyber Physical System：CPS）により、人間の能力の強化や活動を支援し、安全で快適な生活や活力ある社会を目指す Society 5.0 の実現が期待されている。このような高度な情報処理システムを実現するためには、これを支えるコア技術の一つであるコンピューティング技術において、これまで以上の高速化・低消費電力化とともに、リアルタイム性やロバスト性、認識・判断といった人工知能（AI）技術のような高度な情報処理が必要になる。

一方、現在の情報処理性能向上の担い手である集積回路は微細化の限界に直面しており、従来の CMOS デジタル回路によるノイマン型のコンピューティング技術だけではさらなる性能向上は困難になってきている。このため、Society 5.0 の実現に向けて、新たなコンピューティング技術の開発が望まれる。本提案は、このような高度な情報処理を高速・低消費電力・高効率で行う革新的なコンピューティング技術の開発を目指し、ソフトウェアからハードウェアまでの垂直統合的な技術開発を計算ドメイン志向により行い、新たに創出される共通基盤技術の体系化と各技術レイヤーの強化に関する研究戦略を示すものである。

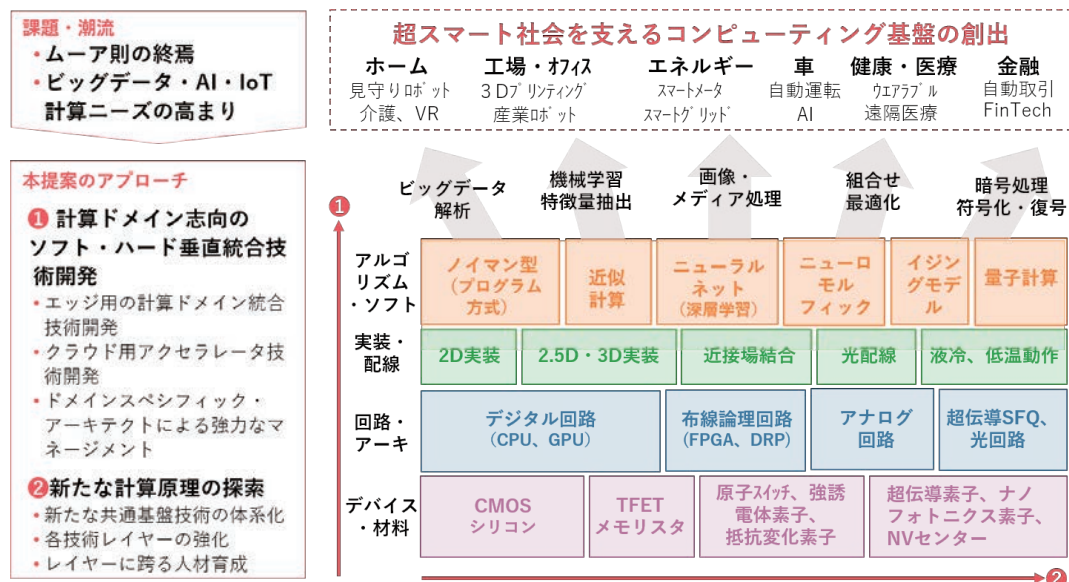


図 1-1 計算ドメイン志向のコンピューティング技術の研究開発

近年はネットワーク・Web の拡大、スマートフォンの普及、コンピュータの飛躍的な性能向上などにより、ビッグデータ処理による購買分析、材料探索や、深層学習（Deep Learning）などの人工知能技術を活用した画像認識や病気診断などにおいて、新たなサービスを創出する動きが活発になっている。これらはクラウド側での情報処理が基本になっ

ているが、今後は IoT の普及により、それらエッジ側における、センサなどからの生データの高効率（高速かつ低消費電力）な解析と価値のあるデータへの変換（画像処理など）、安全なデータ転送（プライバシー保護、セキュリティ確保）、リアルタイムの認識・判断（画像認識、音声認識、危険回避など）などが重要になってくると考えられる。このようなエッジ側における情報処理は、低消費電力性、リアルタイム性、ロバスト性などクラウド側の要求とは異なる特性が要求されるため、これらを実現できる新たなコンピューティング技術に対する期待は大きい。

海外では、米国 DARPA の SyNAPSE プログラムによるニューロモルフィックチップ開発、欧州の Human Brain Project による脳研究と脳を模倣した新たなコンピューティング技術開発、新たなコンピューティング技術の研究開発を目指す IEEE の Rebooting Computing Initiative の設立、Google の深層学習のアクセラレータ TPU (Tensor Processing Unit) の開発と利用、Google、Microsoft、IBM、Intel などによる量子コンピュータの研究開発と利用環境の提供など、新たなコンピューティング技術の開発に対する動きが活発になっている。このため、我が国においても、コンピューティング技術の新たな流れが生まれてきている今をチャンスの時期ととらえ、革新的コンピューティングの研究開発を推進すべきである。

図 1-1 に計算ドメイン志向で今後取り組むべき研究開発領域とその推進方法を示す。図の縦軸には、アルゴリズム・ソフト、実装・配線、回路・アーキテクチャ、デバイス・材料といったコンピューティングに関わる技術レイヤーを示しており、横軸は新たな計算原理に関わる各種技術の革新性や技術開発の困難さを考慮したものとなっている。各技術レイヤーの左端は上から順に現在のコンピューティングの主流であるノイマン型のアルゴリズム、二次元的な実装・配線、デジタル回路、シリコンの CMOS デバイス技術となっており、右に行くほど技術的には難しくなるが、性能向上や消費電力に対するインパクトが大きくなる。

今後取り組むべき研究開発領域はこの図に示した全ての技術に関わるが、全ての技術を網羅的に行うのではなく、社会的課題の解決や重要な応用に向けて、時間軸を考慮して適切な技術を選び技術レイヤー間をまたいで垂直統合的な技術開発を行うことが重要である。ここでは、エッジ側やクラウド側におけるコンピューティングのワークロード（稼働中の IT システムにかかる負荷やその度合い）を明確化することや、日本の強い技術を活かした研究開発を行うといった視点が求められる。このため、以下の 2 つのアプローチで研究開発を行うことが必要である。

(1) ソフト・ハードの垂直統合技術開発と性能検証

コンピューティング技術はアルゴリズム・ソフトから回路・アーキテクチャ、デバイス・材料に跨っており、特定の技術レイヤーの要素技術だけの開発では、全体としての性能向上・最適化につながらない場合もあるため、全ての技術レイヤーを垂直統合的に見た技術開発が重要である。特にエッジ側においては、リアルタイムの画像処理、画像認識、音声認識、自律制御などの重要な計算ドメインがあり、各計算機能に着目してそれぞれに最適なコンピューティング技術を開発することが重要と考えられる。このような計算ドメイン

志向により、超低消費電力性やリアルタイム性、ロバスト性などに注目し、ハードウェア技術やシステム化といった日本の強みを活かした統合技術を開発し、性能優位性を検証することが求められる。

また、このような計算ドメイン志向の技術開発はエッジ側だけでなくクラウド側にも適用できる。クラウド側にもデータ検索、深層学習、経路最適化などある特定の計算処理を超高速・高効率に行うアクセラレータが重要になってきており、ここでも計算ドメイン志向による技術開発が求められる。先に述べた日本の強みを活かして開発したエッジ用の基盤技術をこのアクセラレータ開発に展開することも期待される。

(2) 新たな共通基盤技術の体系化と各技術レイヤーの強化

上記の計算ドメイン志向により新たに開発された各技術レイヤーのコア技術は、特定の計算ドメインだけでなく、他の計算ドメインにも活用できると考えられる。これらを新たな共通基盤技術として体系化し、各技術レイヤーを強化することが重要である。

アルゴリズム・ソフトにおいては、エッジ側やクラウド側のワークロードに対し、近似計算、ニューラルネット、ニューロモルフィック、量子計算など非ノイマン型の新たなアルゴリズムの適用範囲を明確化し、知識を体系化することが求められる。

回路・アーキテクチャにおいては、新たなアルゴリズムとソフトの実行に適する新たなアーキテクチャ技術を蓄積し、他の計算ドメインへの展開に向けて技術ブラッシュアップや組み合わせ方などを体系化する。この技術レイヤーはソフトとハードを繋ぐ重要な研究領域であり、広く使われているデジタル回路に加え、アナログ回路、超電導回路、光回路などの新たな利用方法や、フィールドプログラマブルゲートアレイ（FPGA）や動的再構成プロセッサ（DRP）などのプログラマブルロジックデバイス（PLD）の効率的な活用、さらには実装・配線の技術レイヤーの3次元実装や光配線、近接場結合などの効果的な利用、不揮発メモリ、メモリスタ、原子スイッチなどの新たなデバイスの活用に対する知識の整理が求められる。また、新たに開発する基盤技術を広く活用していくためには、アルゴリズムやソフトのオープンソースに接続可能で、使いやすい設計ツールも同時に開発しておくことが重要である。

デバイス・材料においては、特定の計算ドメインに適する性能・機能のものを開発することが重要であるが、それだけではなく、他の計算ドメインにも適用できるように、新たなデバイス・材料の性能・機能・信頼性などを整理しその適用範囲を正しく理解することや、それらの設計指針を明確にしておくことが重要である。デバイス・材料技術は特に日本の強い領域であり、コンピューティング技術という視点からもさらに強化しておくことが期待される。

本研究開発の推進に当たっては、計算ドメインごとに適したアルゴリズム、アーキテクチャ、ハードウェアの研究開発を垂直統合的に行う体制を構築することが肝要である。これには、技術レイヤーの異なる研究者・研究組織をまとめていくプロジェクト・マネージャー（ドメインスペシフィック・アーキテクト）の役割が重要であり、最適なチーム編成を行い、トップダウン的に技術レイヤー間のギャップを埋め連携を深めて一体的な研究開発を進めることが求められる。

また、具体的なワークロードに対する新たな計算原理・計算手法の基本的な性能の検証

(Proof of Concept : PoC) を行うプロトタイプを集積回路を実現するためのソフト開発、ツール開発、ハードウェア試作ができる研究環境（研究施設、設計ツール、試作装置など）を整備することが必要である。このような研究環境を利用することで、ソフトからハードまでの垂直統合的な研究開発を進めて今後重要となるコア技術を見極めるとともに、オープン化、デファクト化を見据えた戦略を構築することも可能になる。

垂直統合的な研究開発や研究環境の整備については、新たなコンピューティング技術の革新性・リスク、技術成熟度（Technology Readiness Level : TRL）、実現可能時期などを考慮して、府省横断の戦略的なファンディングを実施することが求められる。また、実施側においても産学官がそれぞれの特徴を活かした研究開発のエコシステムを構築することが重要である。産業界の様々なステークホルダにおける将来の製品化・産業化のニーズと、大学・国研のシーズ技術を、時間軸や TRL を考慮してうまくマッチングさせることが望まれる。

研究開発プロジェクトの遂行に当たっては、新たな研究開発に挑戦していることを海外に示すために、主要国際学会での発表や論文作成を積極的に行ってプレゼンスを向上させ、海外との共同研究の推進や優秀な人材の確保により、研究開発を加速することが重要である。また、プロジェクトを進める中で、アルゴリズム・ソフト、回路・アーキテクチャ、デバイス・材料といった各技術レイヤーに跨る新たなコミュニティの形成を積極的に行い、関連学会の連携や、新たな研究領域の創設を図るなど、コンピューティングに関わる研究者・技術者を増やしていくことも重要である。その中でも、技術レイヤー全体に興味を持ち、殻を破って自分の専門外の技術レイヤーに出て行ける人材を育成することが求められる。

2. 提案を実施する意義

2-1. 現状認識および問題点

コンピューティング技術は単なる数字の四則演算や計算機シミュレーションの数値計算だけでなく、画像処理、検索、音声認識、暗号化、高速デジタル通信など様々な情報処理に使われるものであり、我々の生活の中で欠かせなくなったパソコン、携帯電話、スマートフォン、ゲーム機、デジタル家電、クラウドサーバーなどの中核技術になっている。一方、これからの社会の課題として、少子高齢化に伴う労働力不足の解消、高齢者・障がい者の介護・自立支援、安全で便利な移動手段の確保、ものづくりの効率化、社会インフラの効率的な保全、セキュリティ強化、プライバシーの確保などが顕在化している。これらに対し、スマートロボット、自動運転、IoTなどによる物理的な世界としてのフィジカル空間と、コンピュータやネットワーク上における様々なデータや情報からなるサイバー空間とが高度に融合した CPS により、人間の能力の強化や活動を支援し、安全で快適な生活や活力ある社会を目指す Society 5.0 の実現が期待されている。

第5期科学技術基本計画（平成28年1月閣議決定）では、「超スマート社会」（Society 5.0）の実現に向けて、速やかな強化を図る必要のある基盤技術として、安全な情報通信を支える「サイバーセキュリティ技術」、IoTやビッグデータ解析、高度なコミュニケーションを支える「AI技術」、大規模データの高速・リアルタイム処理を低消費電力で実現するための「デバイス技術」、IoTの高度化に必要な現場システムでのリアルタイム処理の高速化や多様化を実現する「エッジ・コンピューティング」などが挙げられている。

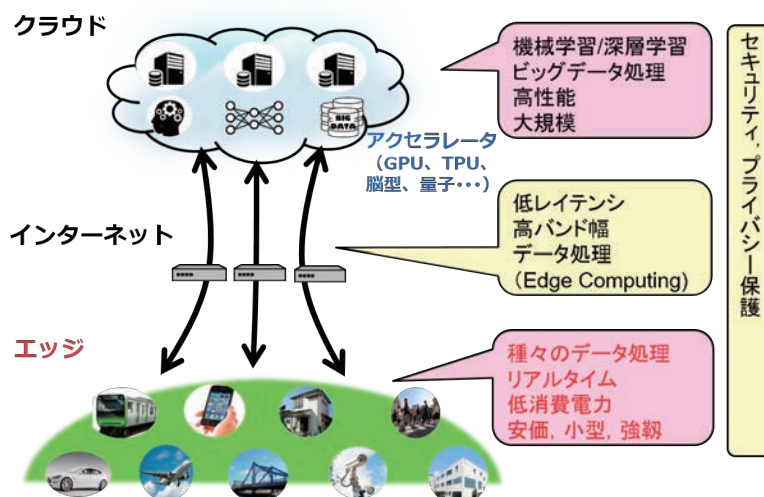


図 2-1 Society 5.0 実現に向けた情報処理システム

図 2-1 にこのような高度な情報処理システムを示している。この実現には、コア技術の一つとなるコンピューティング技術において、ネットワークの末端（エッジ）側の機器のこれまで以上の高速化・低消費電力化とともに、そこでのリアルタイム性やロバスト性、認識・判断といった AI 技術のような高度な情報処理などが必要になる。

フィジカル空間に散りばめられたセンサーデバイスが収集する膨大な量のデータをクラ

ウドに転送してからマルチモーダル情報処理をするのでは、莫大な通信コストがかかってしまう。また、インターネットを経由する際の遅延の大きさがネックとなり、リアルタイム性が求められる領域で CPS/IoT の効果が十分に発揮されない。例えば、リアルタイム高精度画像解析・認証システムを構築するには、画像が撮影されて解析・認証されるまでの待ち時間を実用レベルまで小さくすることが必要であり、そうした情報処理をクラウド上だけで行うことは難しい。また、仮想現実（Virtual Reality：VR）において、VR ヘッドセットを装着した頭の動きに追従して目に見える映像も動かす際に、その映像が実際の頭の動きより遅くなれば、脳内での視覚系と前庭系の間にミスマッチ（感覚のずれ）が生じ、乗り物酔いに似た映像酔い（VR 酔い）が起こる可能性がある。このような遅延の問題を回避するためには、クラウドにデータを転送するのではなく、センサデバイス、すなわちエッジ側において中心的な情報処理を高速に行う必要がある。これがエッジ・コンピューティングという計算パラダイムであり、従来行われている知的情報処理におけるクラウド集約型からエッジ分散型へのパラダイムシフトを意味する。しかも、エッジ側では電力や大きさ等の制限が非常に厳しくなってくるため、これまで以上に小型かつ低消費電力で高度な知的情報処理能力（機能）をもった多様なデバイスが必要となる。

IoT の高度化が望まれている多様な現場のシステム（例えば、少子高齢化に伴う労働力不足を補い人と一緒に作業する産業用ロボット、高齢者・障がい者の介護・自立支援や遠隔医療や患者のケアを行う医療・介護ロボット、安全で便利な移動手段としての自動運転車、ものづくりの効率化のためのファクトリーオートメーションシステム、社会インフラの効率的な保全システム、IoT セキュリティ強化とプライバシー保護を万全にするインテリジェント IoT システムなど）には、どのような環境や状況下にあっても、周囲の環境さらにはシステム自身の状態まで含めて瞬時に判断できる能力が要求される。そのためには、例えば、ネットワークが途切れた状況でも末端のデバイス上で直接的に確実に適切な認識・判断ができる高度な機能が必要となってくる。このため、コンピューティング技術に対してはこれまでとは桁違いの高速性、低消費電力性、ロバスト性が求められる。

一方、現在のコンピューティング性能向上や情報処理性能向上の担い手である集積回路は微細化の限界に直面しており、従来の CMOS ロジック回路によるノイマン型のコンピューティング技術だけではさらなる性能向上は困難になってきている。

これまでの電子機器の発展において半導体、特に CMOS の微細加工技術の進展が果たしてきた役割は非常に大きい。特にインテル社創業者の一人であるゴードン・ムーア氏提唱の「ムーアの法則」をよりどころに、IBM 社のロバート・デナード氏による「スケーリング則」の設計指針もあり、3 年もしくは 2 年で 4 倍の集積度と 20%～30% の動作速度向上もしくは低消費電力化を達成してきた。このように高性能化と製造コスト低減を同時に達成する電子機器は大変魅力的であり、その対象を大型計算機からパーソナルコンピュータあるいはデジタル家電へと変えつつも半導体メーカーはその技術革新に注力してきた。

1998 年からは国際半導体技術ロードマップ（International Technology Roadmap for Semiconductors, ITRS）が発行されるようになり、それはムーアの法則を堅持するための課題を国際的に共有する活動に発展し、国際学会でもその課題解決の構造や材料の提案が相次いだ。2017 年の IEEE International Electron Devices Meeting (IEDM) で

はインテル社の 10nm CMOS 技術や AMD 社の 7nm CMOS 技術の発表があり、現在も CMOS 微細加工技術の進展が続いている。なお、ここでの 10nm、7nm という数字は実際の寸法ではなく、加工プロセスの技術ノードを表しており、実際のデバイス寸法はもっと大きいことには注意が必要である。

5nm 以降はばらつきによる物理限界が近づいているとの意見がある。また、それ以前に消費電力低減のメリットが薄れていることや、製造装置の導入に莫大な投資が必要であることから、10nm 以降の微細 CMOS プロセス技術を構築しているのはわずか数社に留まっている。もちろん、EUV (極端紫外線リソグラフィ) やナノインプリントリソグラフィなどの新たな露光技術や 3 次元化集積技術などの導入により、今後も微細化・集積化はもう少し続くと思われるが、集積回路の性能向上・消費電力低減の傾向は確実に鈍化していく。

IoT の進展やビッグデータなど巨大で複雑なデータ処理の要求は大きいことから CMOS 微細加工によらない新たな展開で性能向上を目指す動きが活発化してきている。すでにメモリでは 3 次元に積層することでコストスケールを実現するデバイスが量産され、ロジックデバイスでもチップを積層する技術が導入されてきている。今後はシステムレベルで新たなアーキテクチャが展開されることが重要となっている。

このように CMOS 集積回路微細化によるコンピューティング性能向上の限界が見えてきたことから、Society 5.0 の実現に向けて、今後はシステム、アルゴリズム、アーキテクチャ、デバイス・材料などの技術レイヤー全体を含む新たなコンピューティング技術の開発が望まれる。

米国では、IEEE「Rebooting Computing Initiative」が主導し、ポストムーア時代の新たなコンピューティングを模索する動きが加速している。ここでの中心的な技術テーマは DARPA「SyNAPSE」プログラム、IBM「TrueNorth」の流れを汲むニューロモーフィックチップの開発と機械学習の実装、NASA や Los Alamos National Laboratory (LANL) と D-Wave Systems 社が進める量子アニーリングマシンを含む量子コンピューティングである。

コンピュータアーキテクチャの国際会議 (ISCA、HPCA、MICRO など) や、半導体回路・デバイスの国際会議 (ISSCC、VLSI、IEDM など) では、ディープニューラルネットワークのハードウェア実装事例が多数報告されブームを迎えている。また、GPU アーキテクチャに関する研究も加速している。

産業界の動きはとりわけ活発化している。NVIDIA は深層学習向けの次世代 GPU アーキテクチャ「Volta」の開発で AI 分野での躍進を目指し、Intel は AMD と共同で CPU/GPU 混載チップの開発を発表するなど、競争が過熱している。国内でも、日立の CMOS アニーリングチップ、富士通の革新的基本最適化回路の開発など、IT ベンダーの研究開発が進められている。また、Google による深層学習専用プロセッサ「TPU」の独自開発・利用や、デンソーによる自動運転用プロセッサ開発の新子会社「エヌエスアイテクス (NSITEXE)」設立など、従来のプレイヤー以外の参入も特徴である。FPGA についても、Intel による FPGA 技術会社 Altera の買収、Qualcomm と Xilinx による戦略的提携など動きが目立つ。また、PC、スマートフォン、IoT の融合も進んでいる。これまで Microsoft 社は IoT 機器への組み込み用途で Windows 10 IoT を x86 と ARM の 2 種類の命

令セットアーキテクチャに対応させてきたが、Qualcomm 社と共同で Snapdragon835 プロセッサで動作する ARM 版 Windows 10 を発表するなど、x86 以外の命令セットアーキテクチャの探索も加速している。

以上のように、世界的にも注目される新たなコンピューティング技術に関して、我が国も積極的に取り組むべきと考えられるが、研究開発を効率的・効果的に行い社会実装が確実になされることが求められる。これには、過去に行われた様々なプロジェクトの進め方について振り返り、良いところや反省点を生かすことが重要である。

例えば、これまでもコンピューティングのシーズ技術や技術レイヤー毎のプロジェクト（新デバイス、新材料・プロセスなど）が行われてきたが、応用分野のシステム構築に必要な様々な技術レイヤーの技術の統合ができず、新たな産業創出に結びついていない場合も見受けられる。このような反省から、ユーザーの視点で、現状あるいは近い将来の問題点・ボトルネックを明確にし、それを解決するための、アルゴリズム・ソフト、設計ツール、回路・アーキテクチャ、デバイス・材料の最適な組合せまで検討するようなことが望まれる。

また、各府省のプロジェクトはそれぞれ独自の目標を持って行われ、日本全体が効率的な研究開発を進めるための戦略的な連携には至っていない場合もある。世界の研究開発の状況を把握し、それに対する日本全体の取り組み姿勢、研究開発の戦略を策定し、各府省が戦略を共有してそれぞれの分担の施策を実施することが望まれる。特に、新たなコンピューティングに関する技術はソフトからハードまでのいくつもの技術レイヤーがあり、非ノイマン型のアルゴリズムとして、深層学習、近似計算、ニューロモルフィック、イジングモデル、量子アニーリング、ゲート型量子計算など多岐にわたり、シリコントランジスタとは異なる機能を持つ不揮発メモリやメモrista など新たなデバイスの提案もあり、時間軸や技術成熟度（TRL）を考慮しておく必要がある。

さらに、研究開発を効率的・効果的に行い確実に社会実装に繋げていくための仕組みや人材育成も検討しておく必要がある。日本は各技術レイヤーの技術レベルは高く、各学会も各技術レイヤーの横の連携活動もできているが、技術レイヤーを跨がる新たなコンセプトの製品やサービスを生み出すという点では産業界を含めて弱い。欧米ではこの縦を繋げるところが強く、見通しを立てられる人がいることが強みであるが、日本にはこのような人材はほとんどいない。このため、産業界だけに頼るのではなく、国全体としてあえて技術レイヤーを縦に繋げる仕組みを作り、この仕組みの中でリーダーシップを取って研究開発を引っ張っていく人材を育成する必要がある。

2-2. 社会・経済的效果

コンピューティング技術は、情報処理の中核をなす技術の一つであることから、コンピューティング技術の進化が経済、産業、生活など社会に及ぼす影響は非常に大きいと考えられる。今後、社会への影響が大きく、高度な情報処理が要求される産業の例として以下のものが挙げられ、それぞれ1年間で1兆円を超える大きな市場規模が見込まれている。

○クラウドサービス [2019年の世界市場規模：約27兆円（IHS Technology）]

- ・ビッグデータ解析の高速・高効率化
- ・個人に合ったサービスの提供（購買行動分析、行動予測）
- ロボット [2021 年の世界市場規模：約 25 兆円（約 2,307 億ドル）（IDC¹）]
 - ・産業用ロボット（人との安全な協働作業、多様な形状・材質の扱い）
 - ・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）
- AR/VR [2021 年の世界市場規模：約 24 兆円（約 2,150 億ドル）（IDC²）]
 - ・高速な物理モデルシミュレーション、画像・映像処理
- 自動運転車 [2030 年の世界市場規模：約 14 兆円（シード・プランニング³）]
 - ・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）
- ファクトリーオートメーション機器・システム [2022 年の世界市場規模：約 7 兆円（富士経済⁴）]
 - ・安定した生産の継続（生産最適化、装置やプロセスの早期異常検出、性能劣化予測）
- ホームセキュリティ [2022 年の世界市場規模：約 5.7 兆円（約 515 億ドル）（Lucintel⁵）]
 - ・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）
- ヘルスケア・ウェアラブルデバイス [2021 年の世界市場規模：約 2 兆円（約 178 億ドル）（Tractica⁶）]
 - ・健康管理（日々の生体データの取得・分析、生活のアドバイス）

上記以外にも、例えば社会インフラの保守のように、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）や修理・建て替えの順番最適化（組み合わせ最適化問題）などで高度なコンピューティング技術の要求があると考えられる。また、半導体集積回路の市場においても、新たなコンピューティング技術を組み込んだ半導体の市場が立ち上がってくると考えられる。現状の半導体世界市場（約 44 兆円：約 4000 億ドル）の 4 割程度をロジック・マイクロ関係の半導体チップが占めることを考えると、大きな規模となることが期待できる。例えば、ニューロモルフィックチップについては Future Market Insights⁷ が 2026 年での世界市場規模を約 1.2 兆円（約 108 億ドル）と推測している。

このように、上記の応用領域やその波及効果も考えると、新たなコンピューティング技術の経済的効果は相当に大きなものになると言える。

このような新たなコンピューティング技術の開発による情報処理の高度化は、社会的に

1 IDC Japan, プレスリリース（2017 年 8 月 2 日）<https://www.idcjapan.co.jp/Press/Current/20170802Apr.html>

2 IDC Japan, プレスリリース（2017 年 8 月 15 日）<https://www.idcjapan.co.jp/Press/Current/20170815Apr.html>

3 シード・プランニング, 市場調査レポート「2017 年 世界の自動運転関連企業の戦略とロードマップ」<http://store.seedplanning.co.jp/item/9498.html>

4 富士経済, プレスリリース（2016 年 8 月 10 日）http://www.group.fuji-keizai.co.jp/press/pdf/160810_16067.pdf

5 Lucintel, “Growth Opportunities in the Global Home Security Market” (May 2017)
<http://www.lucintel.com/home-security-market-2017-2022.aspx>

6 Tractica, “Healthcare Wearable Device Shipments to Reach 98 Million Units Annually by 2021” (April 14, 2016)
<https://www.tractica.com/newsroom/press-releases/healthcare-wearable-device-shipments-to-reach-98-million-units-annually-by-2021/>

7 Future Market Insights, “Neuromorphic Chip Market: U.S. Neuromorphic Chip Market Projected to Register High Y-o-Y Growth During the Forecast Period: Global Industry Analysis and Opportunity Assessment, 2016-2026”
<https://www.futuremarketinsights.com/reports/neuromorphic-chip-market>

も大きなインパクトを与えられとされる。図 1-1 に示しているように、本プロポーザルで着目している計算ドメイン（自律制御、リアルタイム画像認識、診断・予測など）は超スマート社会（Society 5.0）の応用領域（ホーム、工場・オフィス、自動車等）で活用できるコンピューティングの機能群と見ることもできるため、一つの応用領域で有効性が示されたものは、他の応用領域にも急速に広がるとされる。例えば、リアルタイム画像認識技術が自動運転に有効となった場合には、この技術は産業用ロボットや VR などの応用領域にも波及していくと考えられる。また、多くのセンサと繋がった IoT 機器は 2015 年で 154 億個だったものが 2020 年には 304 億個と 2 倍に増加する (IHS Technology の推定) など今後も加速度的に増えていくことが期待されており、IoT 機器からのデータを分析して診断・予測する技術がスマート工場で有効となれば、健康医療や社会インフラ保守にも利用されとされる。

このように、計算ドメイン志向の新たなコンピューティング技術の開発は、多くの応用領域に展開されるようになり、産業の効率化だけでなく、例えば以下の示すような社会的課題の解決や安全・安心な生活にも大きく貢献するとされる。

- ・ 人間と同じように部品・物体の認識、状態の判断、作業計画策定ができるようになり、工場や農作業などでの労働量不足が解消し人間の負担も軽減。
- ・ 知的なロボット機能が一般的になり、高齢者、障害者等が人による介護や助けなしに自立して生活。
- ・ 自動運転が可能になることで、だれでも安全にどこへでも行ける便利な移動手段が確保。
- ・ 工場内に各種のセンサを張り巡らし、各種のセンサの情報を統合して全体の稼働状況をリアルタイムに把握することで、機器の故障回避や計画的な保守、生産状況の把握と作業計画の修正を容易に行うものづくりの効率化。
- ・ 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、異常な時にだけ情報をクラウド側に送ることによる社会インフラの効率的な保全。
- ・ 公共の場に、人間や物体の認識と自律的に状況判断や行動予測が可能なスマート監視カメラの導入で、プライバシーの確保とセキュリティの強化を両立。
- ・ 実生活における高度なサービス、多様な情報・アドバイスを容易に入手できるようになり、仕事の効率化、医療・健康への対応等の生活上の安心感が増加。

2-3. 科学技術上の効果

本提案の計算ドメイン志向により革新的コンピューティングの基盤技術創出に関する研究開発を推進することにより、アルゴリズム・ソフト、回路・アーキテクチャ、デバイス・材料といった異なる技術レイヤー間の垣根を超えた活動が生まれ、技術レイヤー間を跨ぎ、計算ドメインに沿って垂直統合的に考える新たな学術領域や技術領域が形成されると期待される。

2-3-1. 学術の発展

(1) エッジ用の計算ドメイン志向を支える基礎科学・基礎工学

エッジにおいて重要な計算ドメインとしては後述するように、自律制御、リアルタイム画像認識、診断・予測、分散学習、特徴抽出、暗号化、データ圧縮などがあり、これらを

高速、超低消費電力、ロバストに行う基礎科学や基礎工学が進展する。アルゴリズムレベルでは、必要最小限の計算を行って桁違いの高速・低消費電力化を図る近似計算、画像から迅速なものの認識を行う深層学習（推論）、膨大な候補の中から最適なものの選択に強力なイジングモデル、超低消費電力で認識や判断が可能なニューロモルフィックなどの新たなコンピューティング手法に対する知識が蓄積されると期待される。回路・アーキテクチャレベルでは、デジタル回路に加え布線論理回路、アナログ回路などに関する知識の蓄積が期待される。

(2) クラウド用の計算アクセラレータ技術を支える基礎科学・基礎工学：

クラウドにおいて重要な計算ドメインとしては、深層学習アクセラレータ、近似計算アクセラレータ、組み合わせ最適化ソルバー、大規模グラフ処理などがあり、これらを超高速、高効率に実行することを可能とする基礎科学や基礎工学が進展する。アルゴリズムレベルでは、エッジ側の基礎科学・基礎工学に加え、深層学習（学習）、量子コンピューティング（量子アニーリング、ゲート型）などの知識が蓄積される。回路・アーキテクチャレベルではプリフェッチ・分岐予測、インメモリ計算、データフロー計算、超電導回路、光回路などの知識の蓄積が期待される。

さらには、IoT/AI 時代に最適なエッジ・コンピューティングとクラウド・コンピューティングの有機的連携を実現するための、融合的な研究領域も創発的に誕生することが期待される。また、挑戦的な新たなコンピューティング技術を研究開発する中で、物理現象や生物の行動などにヒントを得て、新たな計算原理が生み出されることも期待される。イジングモデルは、結晶を構成する原子のスピンの向きを解析する簡易なモデルであり、ニューロモルフィックは脳の機構をまねている。他にも参考になるものがあるはずであり、このような視点での新計算原理の創出が望まれる。最近注目されている量子コンピュータの一つの方式である量子アニーリングは、1998年に東工大の西森秀稔教授により提案されたものであり、これに続くような提案が出てくることを期待したい。

2-3-2. 技術の発展

技術の発展としては、ノイマン型の従来のプロセッサより高いエネルギー効率のコンピューティング技術（非ノイマン型）の進展が期待でき、各種の近似計算手法を活用したコンピューティング技術、ニューラルネット・深層学習によるコンピューティング、ニューロモルフィック・コンピューティング、イジングモデルによるコンピューティング、汎用量子コンピューティング等の革新的なコンピューティングの実現に向けた基盤技術の研究開発が進展すると期待できる。

エッジ側での実際の応用の課題に対応した高速性・低消費電力性・リアルタイム性・ロバスト性に優れた統合技術として、深層学習の推論、分散学習、マルチメディア処理、経路最適化、暗号化、データ圧縮などが進展する。また、クラウド側では、データ分析用途のアクセラレータ技術として、ビッグデータ解析、大規模学習、最適化問題、大規模シミュレーションなどが進展する。

また、計算ドメイン志向で開発する計算モジュールやIP コアはその組み合わせで様々な応用領域に適用可能になるとともに、そこで得られる様々な知見は他の計算ドメインにも良い影響を与えられと考えられるため、コンピューティング技術全般にわたる質の向上や、

これに携わる研究者・技術者の知識および技術レベルの向上につながる。

特に、クラウド側とは桁違いの低消費電力・高効率化やリアルタイム性、ロバスト性が求められるエッジ・コンピューティングについては、日本の強いデバイス・材料技術やシステム化技術を生かして、世界をリードする技術の創出が期待される。例えば、MRAM、ReRAM など不揮発性メモリ機能を有するデバイスや、高性能で面積効率の高い FPGA が可能な原子スイッチ、動的再構成が可能な回路技術などの新たなコンピューティング技術への活用が望まれる。このような技術の進展により、さらに日本の強いデバイス・材料技術やシステム化技術を進化させるとともに、アルゴリズムやアーキテクチャに関するコンピューティング技術の強化が期待される。

2-3-3. 人材育成

本提言の実践により、計算複雑性・計算原理、デバイス・材料、回路・アーキテクチャ、アルゴリズム・ソフト全体に興味を持ち、自分の殻を破って他の技術レイヤーに出て行ける、多様な知識を有する学際的な IT アーキテクトの育成が期待される。計算ドメイン志向コンピューティングの研究開発においては、上記の各技術レイヤーの総合的な知識が必要になるとともに、多様な研究領域、多彩な応用分野、さらにはサービス科学分野なども関わってくる。これらの研究領域に携わる研究者・技術者・事業戦略立案者などが共通の研究開発拠点やネットワーク構築により密に協力・連携して、垂直統合的な研究開発を全員で総力を挙げて進めていくことができれば、相互に研究目的、内容、考え方などを理解し、また知識や技術を獲得することで、創発的な成長が期待できる。その結果、革新的コンピューティングに関する研究開発全体の中での自分の置かれている立場を明確に認識し、様々な学問領域、理論と実験、基礎研究から応用研究、さらには社会への実装までの広い範囲に興味を持てる人材の育成が可能となる。その結果、情報科学分野に留まらない新たな融合研究領域を切り拓く人材や、社会システムデザイン、Fintech、環境・エネルギー、ライフサイエンス・臨床医学等の分野の課題解決に応用していく人材が出てくることも期待される。

コラム 1 多彩な分野で超人的な「数学的定式化」力を発揮した 天才数学者 John von Neumann

John von Neumann (1903 年～1957 年) ほど、あらゆる分野に数学で切り込み、超人的な業績を上げた数学者はいない。彼は、「Neumann 型と呼ばれる現代コンピュータの父」でもあり、「ゲーム理論」の創始者でもある。また、「量子力学の数学的定式化」に成功したのも彼であり、「数値気象学」の創始者でもある。彼は、ユダヤ系物理学者の Eugene P. Wigner (1902 年生まれ：1963 年ノーベル物理学賞者) や Edward Teller (1908 年生まれ：水爆開発者) と同じハンガリーの首都ブダペストに生まれ、ギムナジウムで英才教育を受けたユダヤ系の神童であった。彼らは、後に、ナチスのユダヤ人迫害を逃れ 1930 年代初頭に米国に亡命し、第 2 次世界大戦中にはナチス原爆開発に対抗するための「マンハッタン計画」(原爆開発・製造のために科学者・技術者を総動員した計画) に携わっている。以下では、彼のことを、米国

的な愛称風に Johnny と呼ぶことにする。

量子力学（Heisenberg の行列力学や Schrödinger の波動力学）が誕生した 1926 年頃、若干 22 歳の Johnny は、「現代数学の父」と呼ばれる David Hilbert に師事し、Hilbert 学派の旗手として純粋数学界に新風を吹き込んでいた。Johnny は、また、相容れなかった Heisenberg の思想（行列力学）と Schrödinger の発想（波動力学）が混迷していた「量子論」の分野に切り込み、ヒルベルト空間の概念を使って、数学者にも受け入れられる形に量子力学を再構成することに成功し、数学の力で混迷を解きほぐしたのである。

Johnny の頭脳の凄さを物語る米国ロスアラモス国立研究所でのエピソードはあまりにも有名である。Johnny が「マンハッタン計画」に携わっていた時の、Enrico Fermi（1901 年生まれ：1938 年ノーベル物理学賞者）と Richard P. Feynman（1918 年生まれ：1965 年ノーベル物理学賞者）と 3 人で、核融合爆弾としての水爆の効率概算を競ったエピソードである。Feynman は卓上計算機で、Fermi は計算尺で、そして Johnny は天井を向いて暗算したが、驚くべきことに Johnny の答えが最も速く正確だったのである。

2 月 8 日は、Neumann 式と呼ばれるプログラム内蔵型の計算機（Neumann 型コンピュータ）について数学的に定式化した Johnny の命日である。彼は、1955 年にイエール大学から依頼された連続講義の準備をしている最中に、左肩に骨肉腫が発症し、「脳と計算機」というチャレンジングな話題に関する未完の原稿を残して、1957 年 2 月 8 日に闘病の末この世を去った。彼自身、未来を見据えて、“Beyond von Neumann” を目指していたに違いない。驚くべき計算効率と電光石火の計算能力を有する革新的な新計算原理コンピュータを、我々が総力を結集して世に送り出せたとき、初めて Johnny を超えることができると言っても過言ではない。

（参考文献）

- ・ノーマン・マクレイ（著）、渡辺正（訳）、芦田みどり（訳）、「フォン・ノイマンの生涯」（朝日新聞，1998）。
- ・Robert Jungk, Brighter than a Thousand Suns: A Personal History of the Atomic Scientists (Mariner Books, 1970).
- ・ジョン・フォン・ノイマン（著）、柴田裕之（訳）、「計算機と脳」（筑摩書房，2011）。

3. 具体的な研究開発課題

今後取り組むべき研究開発領域は図 1-1 に示した全ての技術に関わるが、全ての技術を網羅的に行うのではなく、社会的課題の解決や重要な応用に向けて、時間軸を考慮して適切な技術を選び、技術レイヤー間を跨いで垂直統合的な技術開発を行うことが重要である。ここでは、エッジやクラウドにおけるコンピューティングのワークロードを明確化することや、日本の強い技術を活かした研究開発を行うといった視点が求められる。このため、計算ドメイン志向により、①ソフト・ハードの垂直統合技術開発と性能検証、②新たな共通基盤技術の体系化と各技術レイヤーの強化、に取り組む必要がある。以下、この2つのアプローチでの研究開発課題について述べる。

3-1. ソフト・ハードの垂直統合技術開発と性能検証

3-1-1. エッジ用の計算ドメイン志向の技術開発

図 3-1 にエッジにおける応用領域と重要な計算ドメインとの関係を示す。コンピューティング技術が重要な応用領域の例については付録 3 で説明している。エッジで重要な計算ドメインとしては、自律制御、リアルタイム画像認識、診断・予測、分散学習、特徴抽出、暗号化、データ圧縮などがあり、それぞれの処理の特徴を反映した効率的なコンピューティング技術の開発が必要である。

応用領域 計算 ドメイン	自動運 転	ロボット	警備・ 防犯	スマート 工場	健康・ 医療	社会イン フラ 保守	VR	自動 翻訳
自律制御	○	○						
リアルタイム 画像認識	○	○	○				○	
診断・予測				○	○	○		
分散学習	○	○					○	○
特徴抽出	○	○	○			○		
暗号化	○		○	○	○			
データ圧縮			○			○	○	

図 3-1 エッジにおける応用領域と計算ドメインとの関係

自律制御については、自動車の自動運転レベル 4（高度自動運転）やレベル 5（完全自動運転）で不可欠となり、また、作業内容や作業手順を自分で判断して動く産業用ロボットや、人間に危害を与えず置かれた状況・環境に対して的確に動作するスマートロボット

などの応用領域で重要になる。ここでは、クラウド側との接続（通信）が途切れても、カメラや各種センサからの情報（距離、速度など）から端末側（自動車、ロボット）で状況を総合的に判断し、行動計画を立て、機器（アクチュエータ）を制御することが必要になる。ここでは、AI 的な計算処理が必須であり、脳の機能を模倣した深層学習、ニューロモルフィック・コンピューティング技術などが重要になる。

この自律制御の研究開発の例を図 3-2 に示す。環境の変化に対して瞬時に対応できるような自律制御を実現するために、必要最小限の精度で情報処理を行うとする。現在の利用可能な技術としてシリコンの CMOS デバイスとデジタル回路を選択し、少し挑戦的な技術として論理回路とメモリやセンサなどを三次元的に集積してバンド幅の拡大や回路遅延を縮小できる 3D 実装や、状況に合わせて計算の精度を倍精度から単精度、8 ビット、3 ビット、2 ビットなどに変化させられる近似計算を選択するとしたものである。このような研究開発は、5 年程度の短中期的に実現可能と考えられるが、10 年程度の中長期的な研究開発としてさらなる高速化・低消費電力化を目指して、ニューロモルフィックのアルゴリズムとアナログ回路を用いるものなども考えられる。

リアルタイム画像認識については、自動運転、ロボット、警備・防犯、VR の応用領域で非常に重要である。静止している構造体（道路、建物など）や動いている物体（自動車、人など）が何であるかを画像の特徴から瞬時に認識することが必要になる。この機能は次につながる判断や予測、自律制御に先立って処理を行う必要があるため、リアルタイム性（低レイテンシー）が非常に重要であり、深層学習の高速推論処理、近似計算による最小限のコンピューティング処理を行う技術などが重要になる。このリアルタイム画像認識の

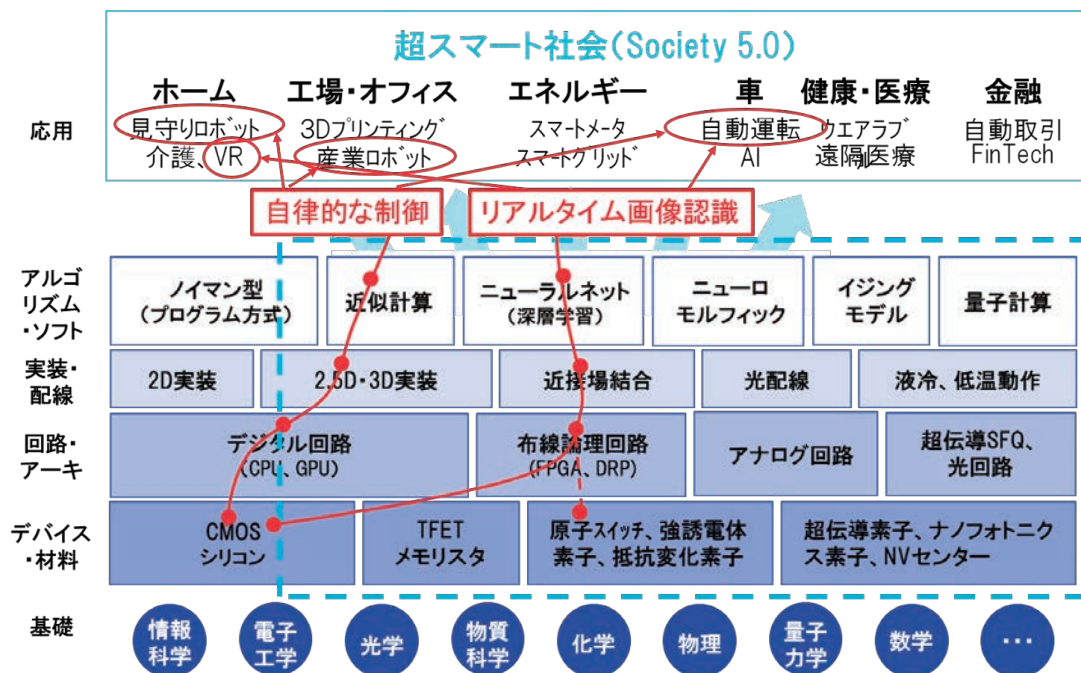


図 3-2 計算ドメイン志向の研究開発の例

ここでは「自律的な制御」と「リアルタイム画像認識」の2つの事例を示しているが、赤線は技術の選び方の一例を示しただけであり、技術が必要な時期や各技術のTRLにより、自由な選択が可能である。

研究開発の例について図 3-2 に示している。リアルタイム画像認識を深層学習で実現することにしたときに、現在の利用可能な技術としてシリコンの CMOS デバイスと FPGA を用いた布線論理回路を選択し、少し挑戦的な技術として論理回路とメモリとのバンド幅を大幅に拡大できる近接場結合を利用するとしたものである。このような研究開発は、5 年程度の短中期的に実現可能と考えられるが、10 年程度の中長期的な研究開発としては、さらに桁違いの高速性や低消費電力性を目指し、深層学習＋近似計算を選択し、布線論理回路の代わりにアナログ回路を選択することや、CMOS の代わりにトンネル電界効果トランジスタ (Tunnel Field Effect Transistor : TFET)、不揮発性の機能を持つ抵抗変化素子や原子スイッチなどの新規デバイスを利用するなど、様々な技術レイヤーの組み合わせが考えられる。

診断・予測については、スマート工場、健康・医療、社会インフラ保守などの応用領域で重要である。それまでに蓄積された各種のデータと、現在の状態を示すデータとの比較により、異常がないかどうかの判定を行うことや、状態の変化から長期的な変化や故障に至る期間などを見積もることが必要である。

分散学習については、自動運転、ロボット、VR、自動翻訳などで重要である。物の識別や顔の認識などだれにでも共通する学習は、クラウド上で行うことが効率的であるが、特定の人の好みや、置かれた特殊な環境などに対する学習は個別に行う方が効率的であり、それにはエッジ側での学習が必要になる。このようなエッジ側で学習を行うためには、ユーザーの好みの変化、身体的能力の変化、置かれた環境の変化などに対応して継続的に学習していく必要があり、時系列での学習が重要になる。

特徴抽出については、自動運転、ロボット、警備・防犯、社会インフラ保守などで重要である。イメージセンサからの画像情報から認識に繋がる特徴抽出を行ったり、時間的な変化から動きの方向や速さを導出したりすることが求められる。この処理は、情報処理の次のステップである認識、判断、診断につながるものであるため、リアルタイム性が求められる。

暗号化については、自動運転、警備・防犯、スマート工場、健康・医療などで重要である。エッジ側で取得したデータを安全にクラウド側へ送る場合や、秘密やプライバシーに係わるデータを保存しておく場合に、データを暗号化しておくことが不可欠である。第三者に暗号を解読されないようにすることが必須であるが、エッジ側で暗号化することから、暗号化の容易性や低消費電力性も求められる。

データ圧縮については、警備・防犯、社会インフラ保守、VRなどで重要である。IoT のセンサ端末などでは、使える電力や通信容量・通信速度が制限される場合が多く、有用なデータだけを選別して取り出したり、価値あるデータ形式に変換したりすることは非常に重要である。データ圧縮は暗号化とも関わるため、通信の制約を考慮して、全体として最適となるような技術開発が必要である。

3-1-2. クラウド用のアクセラレータ技術開発

クラウドにおける計算では、ビッグデータ解析、AI 処理のニーズが急速に高まる一方、それを実現するハードウェアは、微細化による低コスト・高性能化は望めなくなりつつある。クラウドの計算基盤であるデータセンターでは、FPGA や GPU の導入に加え、専用設計チップを導入する動きや、計算要求に応じて適材適所に処理を移すという手法もとら

れている。

また、従来、スパコンが対象としてきた科学技術計算についても、観測データとシミュレーションのデータ同化、機械学習によるシミュレーションのバイパス、大量データからの深層学習など、データセンター技術との融合や外部データとの接続を重視する動きが見られる。そのため、計算結果の質を担保しつつ計算量を削減する方法や、新しい計算原理に基づく特定計算のアクセラレーションへの期待は強い。クラウド側での研究開発に軸足をもつ研究開発課題の代表例は下記が挙げられる。

(1) 深層学習アクセラレータ

深層学習の演算を加速させる「深層学習アクセラレータ」を目指す流れが加速している。Google は機械学習フレームワーク「TensorFlow」向けのニューラルネットワーク演算専用プロセッサ「TPU」を独自開発し、自社データセンターで使用したときの性能評価を CPU や GPU と比較して 10 倍の電力性能比であることを発表した。ニューラルネットワークを用いた推論演算で頻出する行列積（積和演算）の高速化がキーとなることが知られ、その高速化・低消費電力化を狙ったアーキテクチャ研究がすでに ISCA（International Symposium on Computer Architecture）や ISSCC（International Solid-State Circuits Conference）などの国際学会で多数報告されている。

精度を必要最小限に取れば、データ量が減り、回路も小さくてすみ、全体として消費電力を抑えることにつながる。先述の Google の TPU の場合には、計算途中のパラメータは 8 ビット精度である。しかし、どの程度まで精度を落としても推論結果の精度に影響しないのかはケースバイケースであり、ハードウェアによる推論の高速化・低消費電力化は研究開発の余地が十分にある。GPU による推論の加速は深層学習に特化した GPU アーキテクチャと各種の深層学習フレームワークに対応した統合開発環境「CUDA」により NVIDIA が先導している。

学習の高速化は推論に比べ挑戦的な研究開発課題である。深層学習の精度向上には、より大きい学習データと、より多くのパラメータを使うモデルが必要と考えられており、そのための計算時間は膨大なものとなる。また、学習ではパラメータの精度を推論時ほどには落とせない場合が多い。現状では大学の研究室で利用できる GPU リソースでは 1 回の学習に数週間かかる場合もある。学習時間を短縮できれば、新しいアイデアの検証・試行錯誤を短時間で何度も行えるようになるため、ニューラルネットワークを用いた機械学習の研究開発基盤として、学習の高速化は極めて重要な課題である。

単純に多数の GPU を並列化しても GPU 間の通信オーバーヘッドにより精度向上や学習速度の向上は限定的である。したがって、分散深層学習をマルチノードで効率的に行うアルゴリズム開発とその性能検証が求められる。近年、Tesla P100 GPU を 1024 基搭載した深層学習向けスパコンを Preferred Networks 社が稼働させ、ResNet-50 の学習を 15 分で完了するなど有望な成果が出始めている。深層学習に適した GPU アーキテクチャ研究、ライブラリの整備、フレームワークとのコデザイン、サービスとしての深層学習（Deep Learning as a Service : DLaaS）を構築するためのクラウドソフトウェア技術も重要な研究開発要素である。また、深層学習の教師なし学習、転移学習、強化学習の方向への進展と合わせ、それぞれの計算に適した計算機アーキテクチャの研究もこれから重要な方向性である。

(2) Approximate Computing (近似計算)

ビッグデータ・AI 処理の計算要求が増大すると予想されており、絶対量としてコア数やメモリバンド幅を確保して対応することが必要である。しかし、要求計算量の伸びが、提供可能なハードウェア性能の伸びを上回る見込みである。したがって、計算結果の質は保ちつつ、実効的な計算量を大幅に減らすことで、対電力性能を大幅向上させる手法が不可欠となる。この方向の有望なアイデアとして、計算結果の近似や誤差をある程度許すことによって高速化・低消費電力化する近似計算 (Approximate Computing : AC) が注目されている。

回路レベルでは、安定動作しないほどの低い供給電圧で駆動させるなど、ハードウェアの不安定な挙動を積極的に利用したり、ストレージレベルでは、SRAM のリフレッシュレートを下げたり、SSD でエラーコレクションを行わないなどの方法が取られる。メモリアクセスパターンなどからデータの重要度を推定し、メモリコントローラによって重要なものを高信頼度のメモリに、そうでないものを低信頼度のメモリに自動的に振り分ける技術は、エッジ用途だけでなく、サーバーに用いられるストレージクラスメモリでも電力性能比を向上させる有効な手段と考えられる。ソフトウェアレベルでは、サブルーチン呼び出しの結果を再利用するメモイゼーション (Memoization) やループ処理の一部をスキップするループ・パーフォレーション (Loop perforation) などが知られている。モンテカルロ法や乱択アルゴリズムも広義の近似コンピューティングに含まれる。

重要な研究開発要素としては、画像処理など誤りがある程度許容できる応用について、結果の精度と計算コストのトレードオフの統一的な理解と、それに基づくアーキテクチャ設計指針の構築が挙げられる。また、状況に応じた演算精度の自動チューニングもクラウド・スパコンで重要な研究開発テーマである。

(3) 組み合わせ最適化ソルバー

組み合わせ最適化問題を固体物理でよく知られるイジング模型として定式化し、その最小エネルギー状態を求める別の問題として解くアプローチが注目されている。近年の最適化問題ソルバー高速化のブームの発端は、D-Wave Systems 社が進めている量子アニーリングによる組合せ最適化であろう。ここでは、2次元イジング模型に横方向から磁場がかかる場合の最小エネルギー準位を、量子アニーリングと呼ばれる手法で高速に求める。

実行できる問題のサイズや種類はハードウェアの実装形態に依存する。D-wave マシンのイジングスピン (ハードウェアとしては超伝導回路で実装される) は 2000 ビット程度まで実装が進んでいるが、ビット同士は特殊な疎結合のグラフ (キメラグラフ) になっており、巡回セールスマン問題では 47 ノード程度の問題サイズまでしか解けない。組合せ最適化問題をイジング模型として解くこの方向は、日立製作所が進めている CMOS イジングマシン、富士通のデジタルアニーラー、NTT 等が進めているコヒーレントイジングマシンなど、様々なアプローチで研究が進められている。それぞれ、有効イジング模型が異なり、イジングスピンの実装方法、スピン間の結合ネットワークの構造・重み付けの自由度、最低エネルギー状態を探索するアルゴリズムも異なる。

したがって、多ノード・密結合のイジングネットワークをいかに実装するかは技術課題であるが、いずれの実装方法とアルゴリズムを用いるにせよ、既存のソルバーと比較した場合の優位性の検証は必須である。特に、実際に計算にかかる所用時間そのもののほか、

問題サイズを大きくした時に計算所用時間がどのように変化するか（スケーリング）が重要である。巡回セールスマン問題など、産業応用上重要な組み合わせ最適化問題で問題サイズが大きい問題のソルバー高速化は、アクセラレータ技術として極めて重要なテーマといえる。現在注目されている GPU が画像処理専用チップからその歴史が始まったことを鑑み、たとえ特定の問題のみに有効な手法であったとしても、応用上重要な最適化問題であれば開発意義はある。このため、解くべき問題を研究開発の初期から明確に想定し、数学的な理論裏付けとともにハードウェア開発を行うことが肝要である。このような意味からは、ゲート方式の量子コンピューティングは特定の計算で従来計算機より高速なアルゴリズムが理論的に保証されているが、明確になっていない方式もある。キラーアプリを明確化し、その特定用途での従来計算機と比べて高速化するという理論保証（ないしは実験検証）によって、研究の健全化を図る必要がある。

(4) 大規模グラフ処理

大規模グラフの計算では、メモリアクセスパターンが不規則となるなど並列性が無く、GPU には不向きな計算が大量に必要な場合が多い。とりわけ、データ移動に伴う電力消費が課題であり、データをできる限り動かさずに計算する方法が求められる。大規模グラフ処理に限らず多くの場合でメモリバンド幅が計算性能を律速することは多く、メモリ内処理（Processing in Memory）、ニアメモリ処理（Near Memory Processing）、SSD コントローラでの計算（Processing in SSD）などがあるが、汎用 CPU との使い分け、従来のメモリシステムとの一貫性の確保、活用するためのソフトウェア技術の開発などが重要である。また、クラウドのサーバーで FPGA を使用する動きもあり、ハードウェア構成を C 言語や Python などから自動生成する高位合成も重要な研究開発要素となる。

(5) 息の長いテーマの再考

プリフェッチ、分岐予測、インメモリ計算、データフロー、アナログなど、アーキテクチャ研究の流れの中で何回かブームがあり、ブームが去った研究テーマを、今一度、IoT・ビッグデータ・AI の文脈で再考することも有意義である。ブームとなった当時の文脈や用途では低性能・低評価であった技術が新たな見方で改めて評価される場合、従来は実装技術が限定的であったアイデアがデバイスの高性能化により実験検証可能となる場合、荒唐無稽なアイデアであり大企業が行うにはリスクが高すぎた研究などに注意しておくことは重要である。特に、なぜその研究がスタックまたは主流とはならなかったかの考察が極めて重要である。例えば、アナログがデジタルに勝てない理由の一つは精度とスケーラビリティであろう。これらの問いに対して、深層学習の推論計算では精度は必要ないのでアナログが有利となるなどの一定の理由付けは必要である。

データセンターにおける全体的なコンピューティング性能や消費電力を大きく左右するチップ間・ボード間・ラック間通信なども含めた技術開発も重要である。今後チップ間の光配線などで用いられるシリコンフォトニクスデバイスを用いて、計算アルゴリズムを光で実行する光コンピューティングや、データセンター内でのラック間のレイテンシをマイクロ秒レベルに下げるようなストレージクラスメモリ技術なども重要な研究課題である。

3-2. 新たな共通基盤技術の体系化と各技術レイヤーの強化

図3-2に示したような計算ドメイン志向による垂直統合技術開発では、各技術レイヤーにおいても新たなコア技術を創出する。このようなコア技術は、特定の計算ドメインだけでなく、他の計算ドメインにも活用できると考えられ、これを新たな共通基盤技術として体系化していき、知的基盤の強化と各技術レイヤーの国際競争力の強化につなげていくことが重要である。

3-2-1. 新たなコンピューティングアルゴリズムとソフト

アルゴリズム・ソフトウェアの技術レイヤーについては、現在、ノイマン型コンピュータに基づいたものが主流である。ノイマン型コンピュータは、処理するデータおよびプログラムをあらかじめ主記憶装置に記憶するプログラム内蔵方式であることと、CPUが主記憶からプログラムの命令やデータを順に呼び出して実行する逐次制御方式であることを特徴とする。様々な問題に適用できる汎用性の高い計算モデルであるが、問題の性質によっては、計算量や消費電力等で多大なリソースを必要とするケースが生じており、ノイマン型での改良や非ノイマン型の新しい計算モデルが必要になっている。

具体的なケースとしては、昨今、AI技術の応用が広がり、大規模なデータ（ビッグデータ）を用いた機械学習（特に深層学習）や組み合わせ最適化等の計算において、計算量や消費電力等で多大なリソースを必要とする問題が顕在化している。このようなタイプの問題に有効と考えられる、非ノイマン型の新しい計算モデルとして、近年注目されているのは脳型コンピューティングと量子コンピューティングである。

脳型コンピューティングは、脳神経回路網を参考にした計算モデルで、ニューラルネットワークやニューロモルフィックが注目される。ニューラルネットワークは、生体ニューロンの動作を簡易化した人工ニューロンをノードとして、ノード間の結合強度を学習によって変化させる計算モデルである。ノードを非常に多く重ねた多層ニューラルネットワークを用いる深層学習によって、パターン認識の精度が大幅に向上したことから、機械学習の高精度化・高性能化のために大きく注目されている。ニューラルネットワークの計算モデルは、ノイマン型コンピュータ上にも実装可能であるが、専用のチップやハードウェアを用いた並列処理実装によって、より高い性能が得られるため、利用できるハードウェアも含めてニューラルネットの高速で効率的なアルゴリズムを開発していくことが必要である。

一方、ニューロモルフィックは、生体ニューロンのスパイク発火の仕組みを含め、ニューラルネットワークよりもさらに脳神経回路網を模した計算モデルとなっている。脳神経回路網を模することで、人間（脳の消費電力：約20W）のように大幅な低消費電力化も可能になると期待されているが、脳の機能や構造がまだ解明されているわけでもないことから、脳の機能性発現に関する理論的研究などと連携したアルゴリズム研究が望まれる。

量子コンピューティングは、量子力学的な状態の重ね合わせを利用して、超並列的な計算を可能にする計算モデルであり、従来のコンピュータの論理ゲートを量子ゲートに置き換える量子ゲート方式と、イジングモデルのシミュレートにより組み合わせ最適化問題を解く量子イジングモデル方式が知られている。すでに、それぞれの方式のハードウェアが試作され、ネットワーク経由で実際に試すことも可能になってきている。今後は、量子コ

コンピューティングをどのように使いこなすかが重要になると考えられ、組み合わせ最適化などの具体的な問題をいかに量子コンピューティングの処理に帰着させるかなど、ソフトウェアの研究開発が重要になる。

また、ノイマン型であっても、計算に用いる数値の有効ビット数を減らし、多少の誤差を許容した近似計算を行う近似コンピューティングが、計算量や消費電力を削減するのに有効である。特に、機械学習への適用においては、精度をあまり低下させることなく、計算量や消費電力の削減が可能であることがわかってきた。デバイスの動作も含めて、どの程度の近似計算をどこに入れば誤差を少なくして利用できるか、アルゴリズムやソフトウェアの研究開発が必要である。

以上で取り上げたような新しい計算モデルや改良アルゴリズムは、必ずしも汎用的なものではなく、機械学習や組み合わせ最適化等の限定的なケースで効果がある。したがって、それぞれの計算モデルやアルゴリズムについて、どのようなケース・条件で高い効果が得られるのかを明らかにし、ターゲットとする問題の性質に応じて、適切に使い分け、組み合わせられるように、使いこなしの知識やソフトウェア群としての体系化が重要になる。それぞれについて、より性能や実用性を高めるための研究開発が必要であるとともに、体系化への取り組みが求められる。

3-2-2. 新たなアーキテクチャとそれを実現する実装技術

新たなアルゴリズムやそれを具体的なハードウェアに帰着させるアーキテクチャの研究開発が進展すれば、より高速かつ低消費電力な大規模システムへの展開が期待されるが、実際の電子機器を実現するためには大規模システムを実装する技術が重要である。すでにシリコン CMOS LSI によるデジタル回路では複数のチップを 2 次元で実装する技術が実用化されており、今後一部を積層する 2.5 次元実装や 3 次元実装まで発展させる動きがある。このような実装技術において課題となるのがチップ間の接続である。特に、演算部と大規模なメモリとの間のデータのやりとりをするときに、大きな遅延の発生や電力消費が行われる。新たなコンピューティング技術においては、このようなボトルネックを解消しようとするものもあるが、根本的な問題として残る問題でもある。従来の電氣的接続では処理能力が限界に来ており、3 次元化で配線距離を短くするだけでなく、新たな結合・配線技術として、チップ間近接場結合技術や光配線技術などの研究開発が必要である。

アーキテクチャ技術としては、デジタル回路の発展として深層学習を目的に布線論理回路を FPGA に展開する動きや、計算結果にあまり影響を与えないような範囲で演算のビット数を減少させて高速化・低消費電力化を図る近似計算などが非常に活発になっている。ニューラルネットの様々なアルゴリズムが提案される中で、実際に動作する集積回路を実現するためには、アーキテクチャレベルの研究開発が非常に重要である。日本の現状として、このアーキテクチャ技術においては世界に比べて強いとは言えないが、今後注目される IoT、エッジ側の応用に向けての技術開発は世界でも端緒についたばかりである。日本からもエッジ側の応用を目指してオリジナリティの高い論文・発表が出てきており、世界と戦える技術を開発し蓄積していくことが重要である。

また、非ノイマン型コンピューティングの一つとして注目される、ニューロモルフィックにおいては、デジタル回路よりもさらに低消費電力化が可能ということで、アナログ回路を部分的に導入する動きがある。アナログ回路で均一な特性や再現性が得られるかと

いった疑問もあるが、深層学習の推論のようにある程度の誤差を許容できる応用分野においては魅力的であり、アーキテクチャ研究としても重要である。ただし、アナログ回路をデジタル回路と同一チップに混載するか3次元実装とするかはアプリケーションによるところもあり、実装面での新たな課題でもある。

シリコンデバイスとは異なる機能や特徴を持つ新たなデバイスを利用するためのアーキテクチャや実装技術に関する研究開発も重要である。例えば、超伝導デバイスは高速性と低消費電力性から1980年代にジョセフソンコンピュータとして盛んに研究開発されたが、4K程度の低温動作や集積化の難しさなどから、CMOS集積回路を凌駕することはできず実用化には至らなかった。しかし、現在では量子コンピュータ用の素子として利用される段階まで来ている。また、クライオジェニックコンピューティングとして再び注目されるようになっており、従来のシリアルな処理ではなくパラレル処理の新たな検討など、アーキテクチャ技術の蓄積が必要と考えられる。また、光コンピューティングも国際学会ICRC (IEEE International Conference on Rebooting Computing) などで注目されるようになっており、シリコンフォトリソグラフィのようなデバイスだけでなく、光の特性を活用したアーキテクチャ技術も蓄積しておくことが重要である。このような新たなデバイスの活用に向けて、その利用方法や必要実装技術を含めて、知識の体系化にむけた準備を推進すべきである。

一方、システム構築で常に問題となるのが、どの機能をハードウェアで実現すべきで、どこをソフトウェアで組むべきかの議論となる、ハードウェアとソフトウェアの協調設計である。このようなハード／ソフトの境界に関する研究は今後ますます重要になり、その境界に位置するアルゴリズム研究がリードしていくことを期待したい。

以上のようなアルゴリズム、回路、実装の技術蓄積や利用において重要になるのが、コンパイラなどの設計ツールの整備である。技術だけが開発され体系化されても、それを様々な応用で利用できるようにしなければ意味がない。このため、やりたい計算処理内容から、適切なアルゴリズム、回路・アーキテクチャの選択、デバイスレイアウト、動作検証などを自動的に行える設計ツールの整備が不可欠である。既存の標準化されたツールを有効に活用しながらも、独自の技術に対するツールを開発して従来ツールにアドオンすることも、同時に行う必要がある。

3-2-3. 新たなコンピューティングの実行に適するデバイス・材料

新たなコンピューティング技術の開発には、従来のトランジスタやダイオードとは異なる機能を有するデバイス・材料の特性向上、高信頼化などを進めることも必要になる。デバイス・材料技術は、もともと日本が強い分野であり、これらの強みを生かして、新たなコンピューティング技術を開発することが重要である。新たなコンピューティング技術に不可欠な要素は、超高性能、超低消費電力、超小型、などである。デバイス (LSI) 技術開発の潮流の一つは「新規デバイスの開発」であり、もう一つは「LSIの3次元化」である。

新規デバイスは多くの候補が研究されているが、その中でも、Steep-Slopeの「トンネルFET (TFET)」、電荷の有無ではなく金属架橋の有無によってスイッチを実現する「原子スイッチ素子」、第4の回路素子とも呼ばれ、通過した電荷を記憶しそれに伴って抵抗が変化する受動素子「メモリスト」、スピンを利用した「スピントロニクス素子」、粒子と波の性質を兼ね備えた量子を利用する「量子デバイス」などが有力な候補である。これら

のデバイスの先には、いわゆる本格的な「beyond CMOS」と呼ばれるものが控えており、エキシトンなど電荷以外の状態変数を用いたデバイス、熱平衡以外で動作するデバイス、非ブール関数で動作するデバイスなどが提案されているが、それらの本格的な実現はかなり先のことになると考えられる。

一方、新規材料の開発も将来のコンピュータ実現に不可欠である。炭素の1次元物質であるカーボンナノチューブ（CNT）や2次元物質であるグラフェン、h-BN（ヘキサゴナルボロンナイトライド）、MoS₂やWS₂に代表される遷移金属ダイカルコゲナイド、トポロジカル絶縁体などが盛んに研究されている。

どのようなデバイス・材料を選択するかは、応用分野や計算ドメインにより異なるため、必要な計算処理とそれに適するデバイス・材料との関係を整理しておくことが重要である。また、デバイス・材料の研究開発には10年以上かかるものが多いため、技術開発レベル（TRL）を明確にしておくことが必要である。計算ドメイン志向で研究開発を進めていくときに、時間軸およびTRLを考慮して、上位の技術レイヤーとの統合化を図ることが重要である。

コラム2 科学技術の「カンブリア爆発」的進化とエッジ AI 生存競争

英国 ARM 社（売上高約 1800 億円の半導体設計企業）がソフトバンク社により約 3 兆 3000 億円で買収される発表（2016 年 7 月）は、世間に衝撃を与えた。「1 兆個を超える IoT デバイスで人類の進化を加速する」ことを目指す孫正義・ソフトバンク社長が、2016 年 10 月に開催された ARM カンファレンスの基調講演において、5 億 4000 万年前の生物学的な「カンブリア爆発」から話を始めて、超知性誕生としてのシンギュラリティ（技術的特異点）に向けた衝撃的な一手を語ったことは、科学技術の「カンブリア爆発」を我々の脳裏に刻み込むことになった。

生物学的な「カンブリア爆発」は生命進化のビッグバンとも呼ばれ、カンブリア紀（寒武紀：約 5 億 4000 万年前～約 4 億 8000 万年前）の最初の 1000 万年の間に、今日見られる生物の分類学的な「門」のほとんど全てが爆発的に一気に出現した現象を指す。このような生物の複雑多様な進化の謎は、Charles Darwin 流に「突然変異」と「自然淘汰」のみでは語ることはできず、「自己組織化」が決定的な役割を担っていると理論生物学者の Stuart Kauffman は考えている。「カンブリア爆発」の要因には諸説があるが、生物が「眼」を持ったことで進化のスピードが急激にアップし「カンブリア爆発」に結びついていると古生物学者の Andrew Parker は唱えている。これは、いわゆる「光スイッチ説」と呼ばれ、AI 研究者やロボット研究者たちが、深層学習の進化により「眼をもった機械」が誕生することで機械・ロボットの世界での「カンブリア爆発」が起こると考える論拠になっている。

カンブリア紀において、多種多様な生物による激しい生存競争が繰り広げられたとされており、同じような熾烈な生存競争が AI チップの世界でも起こり始めている。Huawei 社は、2017 年 9 月、エレクトロニクスショー（IFA2017）において、「On-Device AI + Cloud AI = Mobile AI」というモバイル環境における AI コンセプトを示した上で、NPU（Neural-network Processing Unit）と呼ばれる世界初の AI 処理機能

を搭載した SoC「Kirin 970」を発表し、NVIDIA 社の GPU (Graphics Processing Unit) に対する優位性を強調した。このチップは Huawei スマホのフラグシップモデル Mate 10 Pro に搭載され出荷されているが、IP ライセンス提供された中国のベンチャー企業 Cambricon Technologies (寒武紀科技) の技術が使われている。このカンブリア紀 (寒武紀) の名を冠した Cambricon Technologies とは、中国科学院計算技術研究所の 2 人の若き陳教授 (兄弟、Yunji Chen : 35 歳、Tianshi Chen : 33 歳) が創立した AI 特化型チップを手がけるスタートアップ企業であり、2017 年 8 月時点で、中国政府系ファンドの主導による初期投資で 1 億ドルを調達している。また、韓国サムスン電子は、IPU (Intelligent Processing Unit) を開発する英国 AI スタートアップ企業 Graphcore に投資しており、現在のハードウェアの 10 倍から 100 倍高速であることが予備テストで示されたとしている。一方、NVIDIA 社は、少量のデータでも訓練できるように精度を損なうことなく AI アルゴリズムを小型化する新手法 (ネットワーク・プルーニング) に関する研究を進めながら、機械学習の「パラダイムを覆す」ことを狙っている。また、ARM 社も、メモリへのデータの入出力が速い低消費電力 AI 専用チップを開発したことを 2018 年 2 月 13 日に発表し、覇権争いが激化してきた。

NVIDIA 社を中心とした産業生態系を一変させるような「一を聞いて十を知り、そして百を語る」ことを可能とする高効率超省エネ AI チップは何か？ 驚くべきゲームチェンジテクノロジーを有した AI チップが創発的に誕生した時、シンギュラリティの到来が近いと言えよう。

(参考文献)

- ・中田敦, “孫社長、「1 兆の IoT が人類の進化を加速」”, 日経ビジネスオンライン (2016 年 10 月 31 日). <http://business.nikkeibp.co.jp/atcl/report/15/061700004/102600155/>
- ・スチュアート・カウフマン (著), 米沢富美子 (監訳), 「自己組織化と進化の論理 宇宙を貫く複雑系の法則」(日本経済新聞, 1999; 筑摩書房, 2008).
- ・アンドリュー・パーカー (著), 渡辺政隆 (訳), 今西康子 (訳), 「眼の誕生 カンブリア紀大進化の謎を解く」(草思社, 2006).
- ・Gill A. Pratt, “Is a Cambrian Explosion Coming for Robotics?” Journal of Economic Perspectives, Vol.29, No.3, pp.51-60 (2015).
- ・松尾豊, 「企業の投資戦略研究会ーイノベーションに向けてー」研究成果報告 3 章 “人工知能 (AI) 搭載の「眼のある機械」の社会実装に向けて人材投資を加速させることが必要”, 財務省広報誌「ファイナンス」平成 29 年 4 月号 pp.52-58 (2017).
- ・劉堯, “Kirin 970 の AI 専用プロセッサは中国独自開発の「寒武期」”, インプレス PC Watch (2017 年 10 月 16 日) <https://pc.watch.impress.co.jp/docs/news/yajiuma/1086229.html>
- ・Yiting Sun, “More efficient machine learning could upend the AI paradigm,” MIT Technology Review (February 2, 2018).
- ・Jamie Condliffe, “The company that made smartphones smart now wants to give them built-in AI,” MIT Technology Review (February 13, 2018).

4. 研究開発の推進方法および時間軸

4-1. 技術レイヤー間に跨がる統合的研究推進

本研究開発の推進にあたっては、計算ドメインごとに適したアルゴリズム、アーキテクチャ、ハードウェアの研究開発を垂直統合的に行う体制を構築することが肝要である。ドメインスペシフィック・アーキテクトをプロジェクト・マネージャーとするチーム編成を行い、トップダウン型の連携を進める。また、計算ドメインを跨がった複数のアーキテクトの連携により、共通化可能部分の共通化を進める必要もある。とりわけ、異なるレイヤー間ではテクニカルチームや研究文化が異なるため、研究開始にあたっては、相互理解を深めるレイヤー間コミュニケーターとしての役割が重要となる。ドメインスペシフィック・アーキテクトになる人材としては、応用領域にも知見があり、上位レイヤー（アルゴリズム・ソフトや回路・アーキテクチャ）に詳しく、技術レイヤー全体を見渡せることが必要なため、企業の研究開発部門の技術に責任を持つマネージャークラスの人や、企業との共同研究などで実用化の実績のある大学教授などが望まれる。

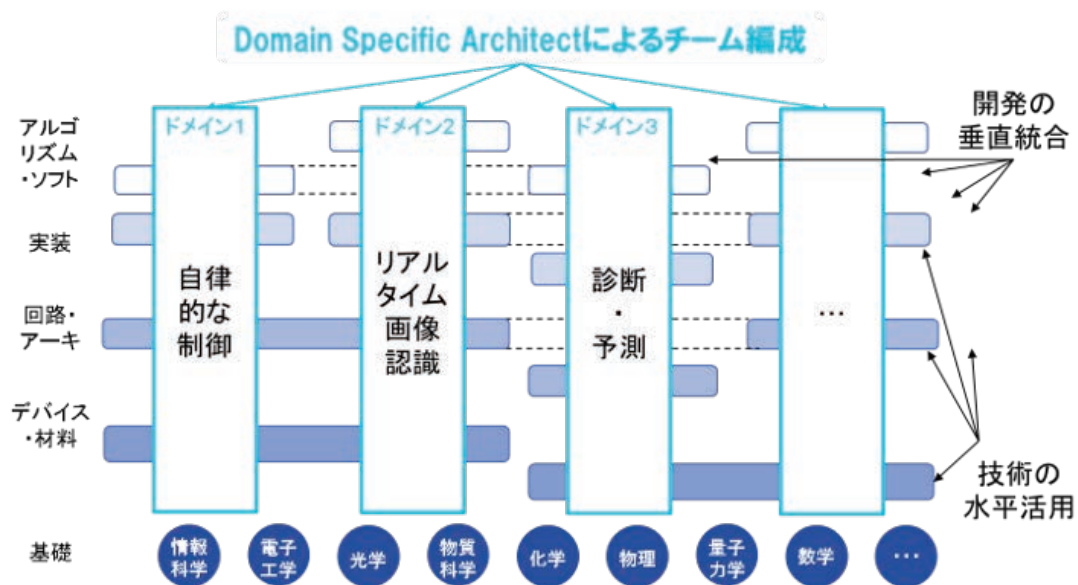


図 4-1 技術レイヤー間に跨る統合的な研究の推進方法

体制構築の例としては、産学官コンソーシアムを設立し、研究開発、人材育成、知財管理、広報などを主体的に行うことが考えられる。半導体の産官学コンソーシアムまたは技術研究組合はこれまで複数のプラクティスが知られており、運営にはそれらの知見を活用することが求められる。特に、競争・非競争領域を定義すると同時に、試作インフラ、計算・データインフラなどに関する共同利用拠点として産業界の参画を促す仕組みづくりが重要である。そのためには、いわゆる箱物施設だけでなく、使いやすい設計ツールやアプリ開発基盤の整備・公開、データ取得・活用の場の提供、ハード設計 IP コアの流通など、ソフトウェアやサービスへの投資が極めて重要である。

また、公開は必ずしも無制限・無料である必要はなく、ユーザーの属性や利用形態に応じた課金も必要である。産業界、とくに、元来半導体が主たる事業ではない企業やスター

トアップに、いかに間口の広い設計環境やハードウェア試作環境を提供できるかが肝である。設計ツールのライセンス料は一般に高価であり、教育目的の場合のみ料金優遇が受けられるのが慣習であった。しかし、中長期的に見てツール利用の間口を広げることは国益であり、企業規模に応じた利用補助金政策などは一考に値する。実際には、ツールベンダとの粘り強いライセンス交渉は不可避だろう。しかし、新しいアイデアの原理実証（Proof of Concept : PoC）を行うための、ソフト開発、ツール開発、ハードウェア試作の高速化は本研究開発のキーであり、共同利用拠点の整備は政府および公的研究機関がイニシアティブを取り重点的に進めるべき施策の一つである。内閣府、文部科学省、経済産業省など府省横断での役割分担と協調が重要である。

4-2. 産学官連携のエコシステムの形成

研究開発プロジェクトや産官学コンソーシアムの成功を判断する指標は、多くは論文数や特許出願数などであろう。本戦略プロポーザルでは、「エコシステムの形成」を成功判断の新規指標として提案したい。

産学官がそれぞれのメリットを感じ協力して研究開発を進めるためには、非競争領域の研究開発テーマの設定が重要である。競争領域のテーマでは同業他社と一緒に取り組むことができず、基礎的な研究テーマでは産業界や国の研究機関は産業化や実用化の時期を予測することが困難で、参加をためらうことになる。非競争領域のテーマ設定は容易ではないが、新たなコンピューティング技術の研究開発においては、複数の応用分野に共通して利用可能な計算ドメインを設定することが有効と考えられる。ある程度の産業利用が見えているが、実際の応用には企業の最適化が必要なレベルであれば、国の研究機関や同業他社も参加できると考えられる。

このような産学官の連携を進めていく中で、それぞれの計算ドメインに関する知識や技術を蓄積し皆が共有して使えるような体制を構築するとともに、日本の産業競争力の強化に向けて、技術の知財化、IP コア化、オープン化、国際標準化などの戦略的な取り組みを行うことが重要である。例えば、強い IP コアの一部をオープン化することで海外のパートナーを呼び込むことや、世界に先駆けて開発した計算ドメイン技術を早期に国際標準化する活動などが考えられる。

前節で述べた共同利用拠点や開発環境は、整備しただけでは機能しない。多数のステークホルダによるエコシステムが形成されて初めて機能する。そのため、各種ツール・ライブラリ・アプリ開発基盤・ハード設計 IP の整備・利用促進は必須であり、いずれも使いやすいインターフェイスを備え幅広いユーザー層に訴求することが重要である。NVIDIA の GPU がこれほどまで急速に並列計算に使われるようになった背景に、統合開発環境 CUDA の存在があったことを参照したい。

本研究開発領域で創出される新しいハードウェアは計算ドメイン志向の専用機となる場合が多いが、プログラマ視点ではそれだけでは使いにくい。機械学習分野で近年利用が進む言語は Python であり、インタープリタ、ライブラリ、コンパイラなどソフトウェア技術の整備はエコシステムの形成に極めて重要である。また、深層学習の場合には Python 上での学習フレームワークである TensorFlow、Theano、Caffe、Chainer などの利用をハー

ドウェア側からも視野に入れておく必要がある。

また、産業用ロボット、建設機械、自動車、医療機器などの応用分野が有望視されているが、技術的・人的に本プロポーザルが主たる対象とする計算機科学・半導体分野からはまだやや距離がある。各種のセンサやアクチュエータの開発・高度化は進められているが、それと比べて制御・計算は遅れている。制御マイコンや組み込み OS と、革新的コンピューティング技術の接続は課題である。IoT の重要課題と認識される情報技術と制御技術との融合（IT/OT 融合）は技術レベルだけではなく、それを開発・利用する人レベルでの融合が極めて重要であろう。研究開発プロジェクトやコンソーシアムを通じて、それらの人の間のオフィシャル・カジュアルな交流が進み、エコシステムを形成していくことが重要である。クラウド上で FPGA のプログラミングを可能とするサービスも進められており、プラットフォーム整備の一形態として参考になろう。

アカデミックなサイロ間の交流を促進する意味では、学会との密な連携も重要である。本研究開発領域はデバイス・材料から、回路・アーキテクチャ、アルゴリズム・ソフトウェアに至る異分野融合領域であり、関連する複数学会の連携で、新規分野創出を促す議論の活発化が必要である。具体的には、情報処理学会、電子情報通信学会、応用物理学会などが連携して新たな研究分野を創出することや、新たなコミュニティを形成することが求められる。

4-3. 人材育成

自動運転などの重要な応用が見えているものの、ハードウェアの微細化による大幅な性能向上が望めなくなっている現状においては、アーキテクチャでの工夫が面白い時代になっている。このアーキテクチャの重要性をエコシステムの形成を通して他の技術レイヤーの研究者や学生にもアピールしていくことが重要である。とりわけ、若手人材の確保は喫緊の課題である。米国ではすでに電子工学科は計算機科学科へと大きくウェイトを移しており、我が国の大学においても計算機科学科の重点化など、工学部の見直しを図る必要がある。国内の大学でアーキテクト人材育成が不足しており、このままではアルゴリズムやアーキテクチャのアイデアをチップに落とし込める人がいなくなる。システムまで見られる専門性を持ったアーキテクトの養成が重要となる。

また、応用領域のニーズを把握し、技術レイヤー全体に対する総合的な知識を持ち、研究開発を指導していくことができる将来のドメインスペシフィック・アーキテクトの育成についても考慮していくべきである。大学の中でもこの基本的な教育を期待したいところであるが、それに加え、計算ドメイン志向の産学官コンソーシアムやプロジェクトの中で、実践的に他の技術レイヤーの研究者や企業の人と交流して知識や興味を広げていく育成体制の構築が望まれる。

国内の人材育成に限らず、海外から優秀な研究者や学生を誘引することも重要である。海外から見てオリジナリティが高く魅力的なコンピューティングの研究開発を進めることはもとより、海外のトレンドの取り入れた多様な研究開発を進めることで、グローバルな先頭集団からの参加者を誘引して、人材育成の場としての国際競争力を高めることも重要である。

機械学習・データサイエンス分野での優秀な若手研究者は Google や Microsoft などの米国 IT 企業との引き合いとなるため、大学や公的研究機関においても、大幅な裁量を認めるような人事制度の確保、柔軟な制度づくりが必要となる。民間企業では、これまで以上に、コンピュータサイエンスの専門性を持つ博士人材の積極的活用と人事制度改革が必要となる。

4-4. 時間軸

計算ドメインや技術レイヤーという空間軸での水平・垂直での役割分担・連携に加えて、時間軸についての検討も重要である。本研究開発領域は、図 4-2 に示すような 3～5 年程度の短・中期的な目標のものと、10 年以上にわたる長期的な視野で行うべきものの両者を含み、政府による公的資金で行うべき部分と、ベンチャーキャピタルや民間企業を通して市場から資金を獲得する部分、民間企業が内部投資で行う部分が混在する。資金に関わる政府機関としては、内閣府、文部科学省、経済産業省、JST、NEDO、JSPS などが想定される。また、実際の研究開発では、国立大学や民間企業のほか、理化学研究所や産業技術総合研究所などの公的研究機関の参画も期待される。

府省や資金配分機関においては、その支援対象によって施策の棲み分けを行うのではなく、政策目標として想定する成果の時間軸に従って緩やかに役割分担をするのが妥当である。また、その際には Society 5.0 の実現という大きなアンブレラの視点から、重複箇所、不足箇所などに俯瞰的な目配せをして、我が国全体としてのポートフォリオの意識が不可欠である。このような司令塔機能は内閣府に期待されるところであり、内閣府がポートフォリオを作成して全体像を描き、研究開発戦略を関係府省と共有し、各府省が効果的な役割分担をして、それぞれに実施する施策の連携をとることで、成果を最大化することが必要である。また、内閣府、文部科学省、経済産業省で関連する施策を実施していく場合には、効率的な研究開発に向けて、研究者の連携推進や、それぞれのプロジェクトの成果の共有と他のプロジェクトへ展開などを図るガバニングボードなどの設置も望まれる。

より長期的な視点で本分野全体の発展を目指し、エコシステムを形成していくためには、研究コミュニティによる新学術領域などの大型科研費の活動を活発化させていくことや、研究機関として世界トップレベル研究拠点プログラム（WPI）に代表される拠点形成事業を積極的に利用して研究環境と研究水準の向上を図ることが期待される。これらにより、全体の研究レベルの底上げをし、自由なアイデアにより多くの失敗も許容できる研究・教育システムの醸成ができると考えられる。特に、大企業がハイリスクすぎて実施しない荒唐無稽とも言えるアイデアが、突如として実現され、工学的応用が開かれるケースが頻出している。従来の性能評価軸で見ると既存技術に比べて見劣りし、かつ観察されるインクリメンタルな進歩の外挿から応用・実用化までかなり距離があるように判断される技術であっても、静観すべきではない。

ソフトウェア開発においては、クラウド上で学習させた識別器をそのままアプリケーションプログラミングインタフェース（Application Programming Interface : API）として公開してサービスインしたり、研究開発レベルであってもあえてベータ版として公開してユーザーのエコシステムを形成したり、その中で実証研究として改良を進めるケースが多い。ハードウェアにおいても、クラウドサービスとして研究開発段階からユーザーと

ともに開発を進めることや、外部のプロセスを利用することでソフトウェアレベルのユーザーによるハードウェアの開発が可能となってきた。その意味では、時間軸での棲み分け境界を随時「越境」することを許容し、基礎研究段階から一足飛びに実用化を目指す際に機動的に資金提供できるような柔軟なファンディングスキームが鍵となろう。

また、図 4-2 に記載されていない新たな計算原理の創出にも積極的に取り組むべきである。イジングモデルや、深層学習、ニューロモルフィックなどは物理現象や脳の構造にヒントを得たものであり、他にも様々な物理現象や生物の行動など参考にすべきものは多いと考えられる。長期的な視点でインパクトの大きさも考慮して、このような挑戦的な研究にも取り組んでももらいたい。

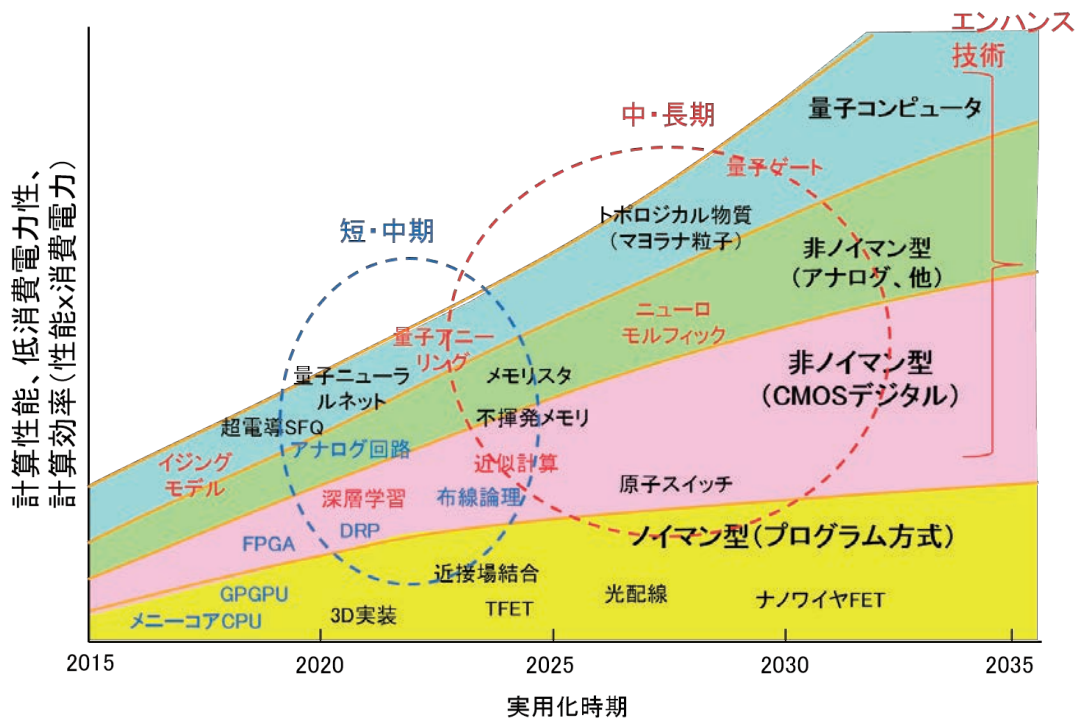


図 4-2 コンピューティング技術の研究開発予測

付録 1. 検討経緯

- ・ CRDS では、平成 29 年度に作成すべき戦略プロポーザルの複数のテーマを戦略スコープ 2017 検討委員会等において議論し選定した。本テーマはその一つであり、本戦略プロポーザル作成チームは平成 29 年 4 月に発足した。
- ・ チームでは、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こんでいったが、その過程では提言の方向性を検討するため、下記の有識者とのインタビューやミニセミナーにおいて意見交換を行ってきた。
- ・ また、本テーマの研究開発に関してチームが構築した仮説を検証する目的で、平成 29 年 7 月 26 日に研究開発戦略検討会および平成 29 年 11 月 29 日に科学技術未来戦略ワークショップを開催した。後者のワークショップについては報告書（CRDS-FY2017-WR-08）を平成 30 年 2 月に発行している。
- ・ 以上の調査・分析等を踏まえ、平成 30 年 3 月に本戦略プロポーザルを発行するに至った。

インタビュー、ミニセミナーにおいて意見交換を行った有識者(所属・役職は実施時のもの)

氏 名	所 属	役 職
インタビュー		
天 野 英 晴	慶應義塾大学理工学部	教授
井 崎 武 士	エヌビディア合同会社エンタープライズ事業部	事業部長
石 川 正 俊	東京大学大学院情報理工学系研究科	教授
井 上 淳 樹	株式会社富士通研究所コンピュータシステム研究所	特任研究員
井 上 弘 士	九州大学大学院システム情報科学研究院	教授
岩 田 太 地	日本電気株式会社事業イノベーション戦略本部 FinTech 事業開発室	室長
内 山 邦 男	株式会社日立製作所研究開発グループ	技術顧問
岡 田 隆 三	株式会社東芝研究開発センター	研究主幹
織 田 和 夫	アジア航測株式会社総合研究所	総合研究所長
金 山 敏 彦	産業技術総合研究所	フェロー
金 子 俊 一	北海道大学大学院情報科学研究科	教授
加 藤 千 幸	東京大学生産技術研究所	教授
清 川 清	奈良先端科学技術大学院大学情報科学研究科	教授
小 出 俊 夫	日本電気株式会社セキュリティ研究所	主任研究員
小 松 崎 常 夫	セコム株式会社	顧問
佐 藤 智 和	奈良先端科学技術大学院大学情報科学研究科	准教授
杉 本 英 樹	株式会社デンソー電子基盤技術統括部 IP 開発室	室長
杉 山 太 香 典	東京大学先端科学技術研究センター	特任助教
堂 前 幸 康	三菱電機株式会社先端技術総合研究所	主席研究員
中 川 章	株式会社富士通研究所コンピュータシステム研究所	准フェロー
中 島 研 吾	東京大学情報基盤センター	教授
中 島 康 彦	奈良先端科学技術大学院大学情報科学研究科	教授

氏 名	所 属	役 職
永 田 聡 美	エヌビディア合同会社ディープラーニング部	ビジネスディベロップメントマネージャ
中 野 恵 一	オリンパス株式会社人事本部教育統括部	部長代理
中 村 泰 信	東京大学先端科学技術研究センター	教授
羽 生 貴 弘	東北大学電気通信研究所	教授
林 喜 宏	ルネサスエレクトロニクス株式会社インダストリアルソリューション事業本部ホームソリューション事業部	シニアスペシャリスト
平 木 敬	東京大学大学院理学系研究科	特任研究員
増 田 宏	電気通信大学大学院情報理工学研究科	教授
松 岡 聡	東京工業大学学術国際情報センター	教授
丸 山 宏	株式会社 Preferred Networks	最高戦略責任者
森 祐 司	株式会社デンソー電子基盤技術統括部 IP 開発室	担当次長
本 村 真 人	北海道大学大学院情報科学研究科	教授
森 川 博 之	東京大学大学院工学系研究科	教授
山 名 早 人	早稲田大学理工学術院	教授
横 矢 直 和	奈良先端科学技術大学院大学	学長
ミニセミナー		
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 彰 二 郎	リガク株式会社	特別顧問
浅 井 哲 也	北海道大学大学院情報科学研究科	教授
浅 野 種 正	九州大学大学院システム情報科学研究院	教授
勝 満 徳	株式会社 TRL	最高経営責任者
黒 田 忠 広	慶應義塾大学理工学部	教授
阪 本 利 司	日本電気株式会社システムプラットフォーム研究所	主任研究員
佐 藤 正 幸	株式会社 TRL 新事業推進室	室長
竹 内 健	中央大学理工学部	教授
田 中 宗	早稲田大学高等研究所	准教授
津 田 健 二	株式会社セミコンダクタポータル	編集長
戸 井 崇 雄	ルネサスエレクトロニクス株式会社第二ソリューション事業本部インダストリ・ソリューション事業部	課長
中 原 啓 貴	東京工業大学工学院	准教授
橋 本 昌 宜	大阪大学大学院情報科学研究科	准教授
益 一 哉	東京工業大学科学技術創成研究院	教授
松 谷 宏 紀	慶應義塾大学理工学部	准教授
三 河 巧	パナソニックセミコンダクターソリューションズ株式会社	主幹
本 村 真 人	北海道大学大学院情報科学研究科	教授
安 田 哲 二	産業技術総合研究所ナノエレクトロニクス研究部門	研究部門長

研究開発戦略検討会「革新的コンピューティング」プログラム

(敬称略)

日時：2017年7月26日（水）13:30～18:00

場所：JST 東京本部別館 2 階会議室 A-2

司会：馬場 寿夫（JST-CRDS）

13:30～13:35 開会挨拶

木村 康則（JST-CRDS）

13:35～14:10 検討会の開催趣旨と仮説の説明

馬場 寿夫（JST-CRDS）

（話題提供 1）革新的コンピューティングに対する産業界からの期待

14:10～14:40 新たなコンピューティングへの期待（Strict から Best Effort へ）

井上 淳樹（富士通研究所）

14:40～15:10 変曲点を迎えた AI 時代のプロセッサ

～布線論理型アーキテクチャの提案～

戸井 崇雄（ルネサスエレクトロニクス）

（話題提供 2）今後取り組むべき研究開発領域

15:30～16:00 Deep Learning に向けた Approximate Computing とシストリックアレイ

アーキテクチャ

中島 康彦（奈良先端科学技術大学院大学）

16:00～16:30 新奇デバイスを用いた革新的コンピューティング・プラットフォームの開発

～ポストムーア時代を見据えた技術レイヤー縦断的アプローチ～

井上 弘士（九州大学）

16:40～17:55 総合討論

ファシリテータ：嶋田 義皓（JST-CRDS）

1. 社会的・経済的・科学技術的效果（日本のターゲットとすべき重要なワークロードなど）
2. 重要な研究開発課題（アルゴリズム、アーキテクチャ、デバイス、ツール開発など）
3. 研究開発体制（レイヤー間融合、研究拠点、産学官連携など）
4. その他（学会の役割、人材育成など）

17:55～18:00 閉会挨拶

木村 康則（JST-CRDS）

招聘識者（所属・役職は開催時のもの）

氏 名	所 属	役 職
話題提供者		
井上 淳 樹	株式会社富士通研究所コンピュータシステム研究所	特任研究員
井上 弘 士	九州大学大学院システム情報科学研究院	教授
戸井 崇 雄	ルネサスエレクトロニクス株式会社第二ソリューション事業本部インダストリ・ソリューション事業部	課長
中島 康 彦	奈良先端科学技術大学院大学情報科学研究科	教授
コメンテータ		
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
平木 敬	東京大学大学院理学系研究科	特任研究員

科学技術未来戦略ワークショップ

「ドメインスペシフィック・コンピューティング ～新たなコンピューティングの進化の方向性～」
プログラム

(敬称略)

日時：2017年11月29日(水) 10:00～18:00

場所：TKP 市ヶ谷カンファレンスセンター バンケットホール 9A

司会：馬場 寿夫 (JST-CRDS)

10:00～10:05 開会挨拶 木村 康則 (JST-CRDS)

10:05～10:45 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)

(話題提供1) 革新的コンピューティングに関する国内外の活動状況

10:45～11:10 国際学会、論文誌における最新の動向 平木 敬 (東京大学)

11:10～11:35 ポストムーア/ポストノイマン世代に向けた国際ロードマップ活動の現状
林 喜宏 (ルネサスエレクトロニクス)11:35～12:00 日本工学アカデミーの「次世代コンピューティング」活動
金山 敏彦 (産業技術総合研究所)

(話題提供2) ユーザー側からの革新的コンピューティングに対する要望

13:00～13:25 自動運転時代に必要なコンピューティング特性
杉本 英樹 (デンソー)13:25～13:50 ロボット：製造・物流を中心とした最新動向と課題
堂前 幸康 (三菱電機)13:50～14:15 医療応用領域からの期待 ～内視鏡を例に～
中野 恵一 (オリンパス)14:15～14:40 スパコン／科学技術計算 (シミュレーション)
齊藤 元章 (PEZY Computing)

(話題提供3) 重要な研究開発領域とその取り組み方

15:00～15:25 革新的コンピューティングを支えるソフトウェア基盤
中島 研吾 (東京大学)15:25～15:50 知的情報処理時代を支える情報処理アーキテクチャの革新
本村 真人 (北海道大学)

15:50～16:15 ハードウェア側からの提案 竹内 健 (中央大学)

16:25～17:55 総合討論 ファシリテータ：嶋田 義皓 (JST-CRDS)

1. 日本のターゲットとすべき応用領域と重要なドメイン
2. 重要なドメインの具体的な技術統合 (アルゴリズム、アーキテクチャ、デバイス、ツール開発など)
3. 効果的な研究開発の仕組み (技術レイヤー統合ができる研究拠点、設計ツールの共有化、ソフト・ハードの標準化、実効的な府省連携など)
4. その他 (海外との連携、アーキテクチャ研究強化など)

17:55～18:00 閉会挨拶 木村 康則 (JST-CRDS)

招聘識者（所属・役職は開催時のもの）

氏 名	所 属	役 職
話題提供者		
金 山 敏 彦	産業技術総合研究所	フェロー
齊 藤 元 章	株式会社 PEZY Computing	代表取締役社長
杉 本 英 樹	株式会社デンソー電子基盤技術統括部 IP 開発室	室長
竹 内 健	中央大学理工学部	教授
堂 前 幸 康	三菱電機株式会社先端技術総合研究所	主席研究員
中 島 研 吾	東京大学情報基盤センター	教授
中 野 恵 一	オリンパス株式会社人事本部教育統括部	部長代理
林 喜 宏	ルネサスエレクトロニクス株式会社インダストリアルソリューション事業本部ホームソリューション事業部	シニアスペシャリスト
平 木 敬	東京大学大学院理学系研究科	特任研究員
本 村 真 人	北海道大学大学院情報科学研究科	教授
コメンテータ		
天 野 英 晴	慶應義塾大学理工学部	教授
内 山 邦 男	株式会社日立製作所研究開発グループ	技術顧問
黒 田 忠 広	慶應義塾大学理工学部	教授
本 村 真 人	北海道大学大学院情報科学研究科	教授

付録2. 国内外の状況

◆研究開発の代表例

欧米、中国、韓国におけるニューロモーフィックコンピューティングおよび深層学習の研究開発の代表例を挙げる。

(1) ニューロモーフィック

(ア) アナログ回路

- ① **cxQuad** (チューリッヒ大、ETH チューリッヒ): 画像認識用に設計されたチップ。アナログニューロン (Adaptive-Exponential Integrate-and-fire 方式) により、プーリング処理と畳み込み演算を行う。
- ② **ROLLS** (チューリッヒ大、ETH チューリッヒ): 3種類のシナプスを持つ画像認識用チップ。アナログニューロンは上記 **cxQuad** と同様。主に画像分類を行う。

(イ) デジタル回路

- ① **SpiNNaker** (マンチェスター大学): ARM コアベースの超並列処理システム。Python ライブラリ **sPyNNaker** を使用しニューラルネットワークを ARM コアに乗るようグラフ分割し、シミュレーションモデルとして実行する。
- ② **TrueNorth** (IBM): 推論用ニューロモーフィックチップ。コアあたり 256 個のニューロンが実装され、シナプスは符号付き 9 ビットをサポートしている。MNIST データセットで、64 コア構成で 99.42% 精度の画像分類を $121\mu\text{J}$ で処理可能と発表されている。
- ③ **BrainScales** (ハイデルベルグ大学):
- ④ **Tianji** (清華大学): 時分割方式で処理するニューロンを持つニューロモーフィックチップ。ツールチェーン **NEUTRAMS** も平行して開発され、ニューラルネットワーク表記から実際のプラットフォームへのマッピングを行える。

(2) 深層学習

(ア) アナログ回路

- ① **PRIME** (UC サンタバーバラ): メモリ機能とアクセラレータ機能を動的に切り替え可能な **ReRAM** アレイによる **Processing-in-memory (PIM)** モデルの実装。
- ② **ISAAC** (ユタ大学): **DaDianNao** (後述) をベースとする、畳み込みニューラルネットワーク計算用のチップ。積極的パイプライン化により高スループット化を狙う。

(イ) ASIC

- ① **DianNao, DaDianNao, PuDianNao** (中国科学院): ニューラルネットワーク処理用のアーキテクチャ **DianNao**。データ再利用法によりメモリアクセストラフィックを削減する。**DaDianNao** は **DianNao** をノードとする分散型アーキテクチャ (推論・学習用)。**PuDianNao** は DNN 以外にも、サポートベクターマシン (SVM) や k 近傍法などの機械学習モデルもサポートする。
- ② **Cambricon-ACC, Cambricon-X** (中国科学院): **Cambricon-ACC** はニューラルネットワーク特化型の命令セットアーキテクチャ (ISA) である **Cambricon ISA** を一般的 RISC アーキテクチャに組み込んだアクセラレータである。また、行列のスパース性を利用したアクセラレータとして **Cambricon-X** も平行して開発されている。

- ③ K-Brain (KAIST (韓国科学技術院)) : 動的再構成可能なシストリックアレイから構成される画像認識用プロセッサ。
 - ④ Deep CNN Processor for IoE (KAIST (韓国科学技術院)) : 畳み込み演算の並列化や冗長なメモリアクセス削減などの特徴をもつ CNN 計算用プロセッサ。
 - ⑤ Tensor Processing Unit (Google) : 深層学習に用いられる TensorFlow API 処理用のアクセラレータ。数値表現を 8 ビットにすることで高いスループットを実現、シストリックアレイ方式の行列積ユニットにより行列演算を行う。TPU をインストールしたシステムが運用され、NVIDIA K80 GPU 比 30 倍の電力効率と発表されている。
 - ⑥ Ecient Inference Engine (EIE) (スタンフォード大学) : 画像認識用アクセラレータ。ニューラルネットワークの無駄な結合の剪定や重み共有、パラメータ量子化などの手法によりモデルとパラメータを圧縮する。
 - ⑦ Eyeriss (マサチューセッツ工科大学) : 画像認識用アクセラレータ。シストリックアレイ上での積極的なデータ再利用が特徴。
 - ⑧ Cnvlutin (トロント大学) : CNN 処理用アーキテクチャ。DaDianNao (前述) をベースに、積算におけるゼロとの演算を削減することで処理の高速・低消費電力化を狙う。
 - ⑨ Minerva (ハーバード大学) : アルゴリズム・アーキテクチャ・回路の設計環境。ネットワークモデルからアクセラレータのレジスタ転送レベル (RTL) コードを自動生成。
- (ウ) FPGA
- ① Catapult (Microsoft Research) : データセンター用途のサーバーに Altera の FPGA (Stratix-V D5) を実装した、画像認識処理のアクセラレータ。
 - ② FINN (Xilinx Research Labs) : 1 ビット表現のパラメータと活性化関数を持つニューラルネットワークモデルの FPGA 実装。
 - ③ FPGA-based Accelerator (北京大学) : ルーフラインモデルによって、CNN の FPGA への実装の最適化を行う。FPGA は Xilinx の Virtex7 を使用。単精度浮動小数点による演算。
 - ④ Embedded FPGA Platform (清華大学) : CPU と FPGA からなる画像認識用のアクセラレータ。メモリアクセスを管理する CPU 上の Processing system と、FPGA 上に実装される Convolver complex からなる。FPGA として Xilinx Zynq706 を使用。
 - ⑤ DNNWeaver (ジョージア工科大学) : DNN の仕様から FPGA 向けのコードを自動生成するフレームワーク。プログラマに高レベルでの深層学習モデル開発環境を提供。
 - ⑥ FPGA Acceleration of RNNLM (ピッツバーグ大学) : Xilinx Virtex6 LX760 を使用した再帰型ニューラルネットワーク (RNN) による言語モデルの FPGA 実装。隠れ層は 1 層。
 - ⑦ NeuFlow (ニューヨーク大学) : データフローアーキテクチャによる画像認識用プロセッサ。専用のコンパイラ LuaFlow により、flow-graph 表現を機械語にコンパイル。

◆日本の状況

基礎研究・応用研究とも産官学で行われている。深層学習での演算能力を重視した「人工知能処理向け大規模・省電力クラウド基盤（AI Bridging Cloud Infrastructure: ABCI）」の開発がスタートした。富士通のAIエンジン「Deep Learning Unit」や日立の「CMOS アニールリング」などクラウド向けの専用チップ・アーキテクチャ研究開発が企業中心に進んでいる。FPGA や ASIC によるニューラルネットワークの実装やニューロモルフィックチップに関する研究開発は量・質ともに米国と比べると劣る。

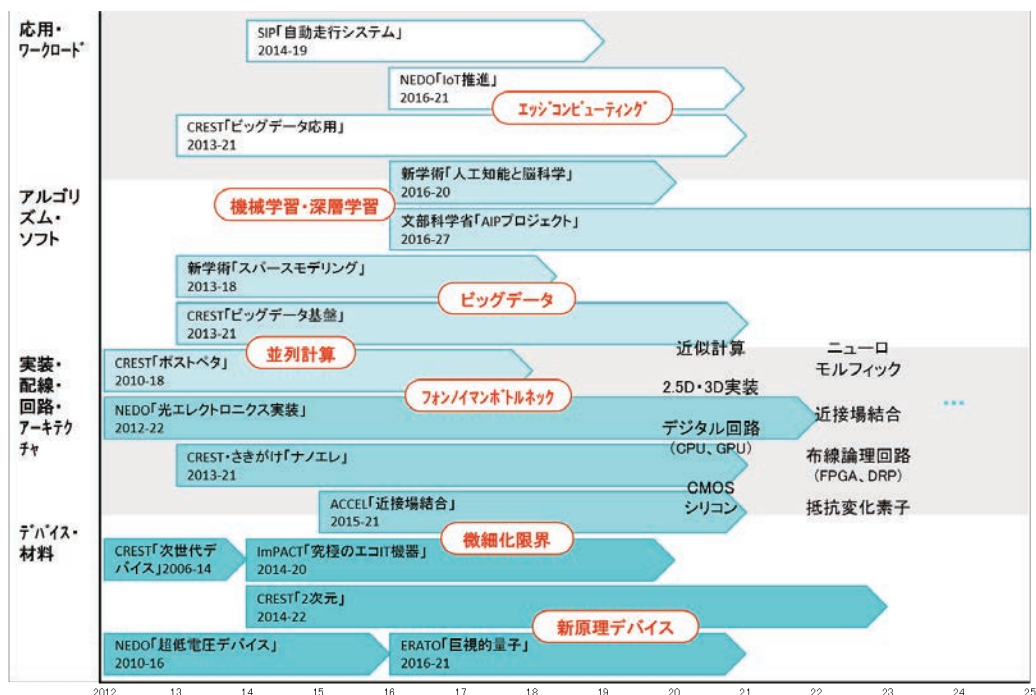


図 S2-1 革新的コンピューティング関連のプロジェクト

◆米国の状況

新たなコンピューティング技術の分野で最も活発なのが米国である。脳型・量子コンピューティング研究者の層が厚く、様々な基礎研究が行われている。DARPA がニューロモルフィックチップ開発プログラム「SyNAPSE」を開始（2008年）、IBM がニューロシナプティックコンピュータチップ「TrueNorth」を発表した（2016年3月）。DARPA はグラフ解析に特化した研究プログラム「HIVE」も開始している（2017年）。NSF では「Expedition in Computing」プログラムを実施、量子コンピュータアーキテクチャ、DNA コンピューティング、分子コンピューティング、視覚野ニューロモルフィック計算などの5年間プロジェクトに最大年200万ドルのグラントを行っている。HPC 分野では、米国はTop500において圧倒的シェアをもち、アプリケーション開発、システムソフトウェア開発、ハードウェア開発技術も最先端を走る。機械学習のアクセラレータとしては、Google が専用チップ「Tensor Processing Unit (TPU)」を開発し、実際のデータセンターでの運用性能が発表されるなど耳目が集まっている。量子コンピュータについても Google、Microsoft、IBM など主要企業が開発に挑戦している。IEEE では Rebooting Computing Initiative が設立され、脳型、量子、光計算などによるコンピューティングの革新が議論されている。

◆論文の推移

当該分野においてはニューラルネットワーク、深層学習、ニューロモルフィック、イジングマシン、量子コンピューティング、Approximate Computing など、新たなコンピューティング技術が提案され、アルゴリズムから回路・アーキテクチャ、新たなデバイス・材料を含めたソフトからハードウェアまでの研究が活発化している。

図 S2-2 に示すように、これら新たなコンピューティングのアルゴリズム関係論文数は全体的な傾向として増加傾向にある。上位 5 カ国までの論文数の動向をみると、米国と中国が突出しており、それ以降の国は日本も含めて横並びで微増傾向にある。(米国は 1990 年から増加し、中国は 2000 年から急激に増加)

分野別にみると、近年はニューラルネットワーク、深層学習、Approximate Computing、量子コンピューティング、ニューロモルフィックなどの分野の増加が顕著になっている。

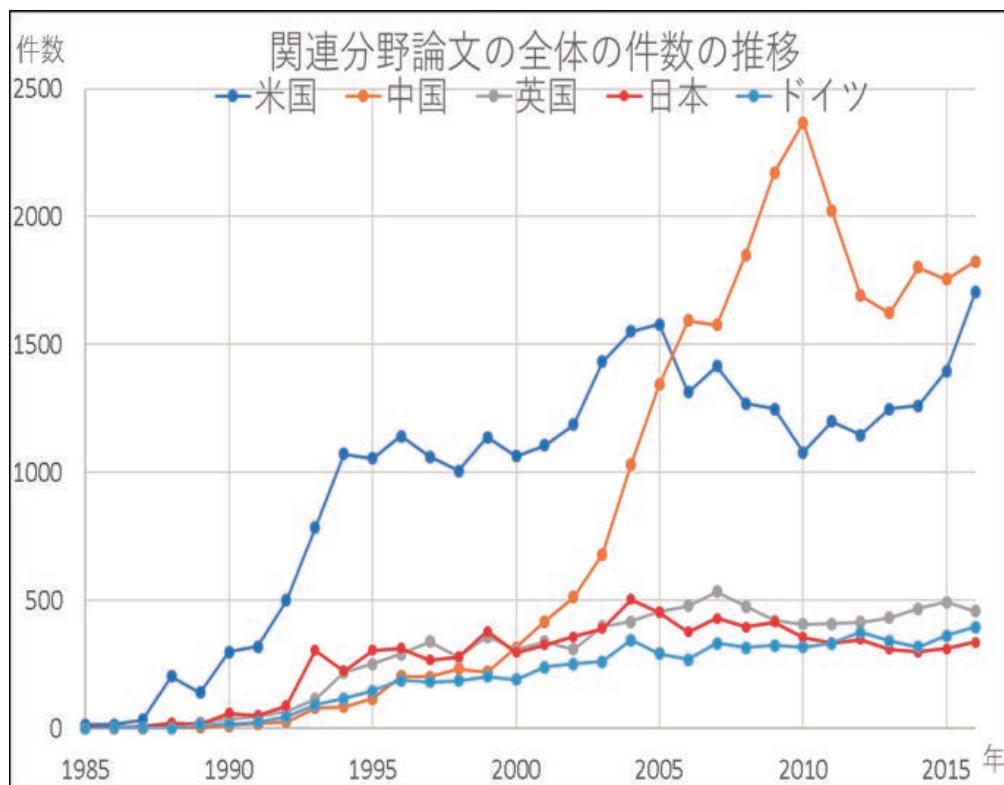


図 S2-2 革新的コンピューティングに関する国別論文数の推移

アルゴリズムに関する以下のキーワードを選び、エルゼビアの SCOPUS で検索

- (1) ニューラルネットワーク (“neural network” AND comput*),
- (2) 深層学習 (“deep learning” AND comput*),
- (3) Approximate Computing (approximate comput*),
- (4) 量子コンピューティング (“quantum comput*),
- (5) ニューロモルフィック (“neuromorphic comput*),
- (6) イジングモデル (“ising model” AND comput*)

付録3. コンピューティング技術が重要な応用領域

(1) 自動運転、ロボット

①自動運転

自動運転技術には、クラウド側とエッジ側の役割分担が必要であり、同時にその間の通信技術がキーとなる。この性能によってクラウド側とエッジ側の必要能力が異なってくるが、データ転送に伴うデータ圧縮や暗号化などの新規技術の開発が必要になる。また、これら統合化技術と全体を概観したシステム設計が重要になる。

自動運転を実現させるためには、自己位置と環境を同時に認識しながら、車載間（前方、後方、左右）でお互いの位置・速度を認識しながらの連携、協調、分散動作を行うことが必須である。これには他の車の行動予測を行うだけでなく、さらに突発的な事象への対応が求められ、かなり高度なリアルタイム予測、判断機能が必要である。これをレベル3（条件付き自動運転）で行うよりは、レベル4（高度自動運転）として、人間の運転（運転技術は個人により異なるため予測が難しい）を排除し、全車が協調して行うほうがはるかに容易である。

以下自動運転での各部の役割例を記す。

【クラウド側】各自動車から取得される画像情報を統合して地図作成、ランドマーク作成等を行い、エッジ側での画像認識のための様々なランドマークを学習させる役目をもつ。日々道路の周辺環境が変化する中で常に更新し続け、学習していく仕組みが必要である。各車載に取り付けられたカメラ画像を回収することで、莫大な量の学習データを収集できる。この処理にリアルタイム性は必要ないものの、自動車に常に更新した地図を送信配布する必要がある。その際に画像は膨大な量となるため、通信速度、容量の制限から圧縮技術が必要になる。

【エッジ側】実際に走行する自動車の内部でリアルタイム処理を行う必要がある。その際に必要なのは、地図上の自己位置認識機能と環境の三次元再構築機能、ランドマーク検出・認識機能、環境認識機能、である。これらの機能の実現にセンサは必須で、複数台のカメラ、ライダー（LIDAR）、GPS、IMU（慣性航法装置：GPSが受信できないときの自己位置検出を行う装置）、超音波センサ、車輪の回転数などのモニタリング等をリアルタイムで統合処理し、自己位置を検出、補正しながら地図上を走行することが必要になる。この実現には、リアルタイムに画像から特徴抽出を行い、複数のセンサ情報を統合処理する必要がある。エッジ側では、リアルタイム処理だけでなく、超低消費電力化、超小型化も必要である。また、自動運転に限らないが、エッジ側のオンライン学習が可能となれば、各地域やローカルでの特徴を反映した学習が可能となり、サービスもきめ細かく個々に提供できるようになる。オンライン学習が可能となれば、データの爆発やデータ転送ネックなども解消されることが予想される。

②ロボット

ロボットの場合も利用される技術は、クラウド側もエッジ側も自動運転と同様である。ロボット分野の場合は、ソフトやハードのオープン化が進み、さらに DARPA のロボットチャレンジや、アマゾンピッキングチャレンジなどに代表される競技会等で

著しく進展している。その原動力は、ROS (Robot Operating System : ソフトウェア開発者のロボット・アプリケーション作成を支援するライブラリとツール) がフリーで、ライブラリ化されており、研究者であれば誰でもが比較的簡単にロボットを動かすことができる環境になっている。ロボットのハードウェアもオープン化が進みつつあり、これに openCV (Open Source Computer Vision Library) などのコンピュータビジョンのフリーの画像認識ライブラリを組み合わせることで、ロボットに画像認識機能を付加することができる。産業用のロボットメーカーとは異なり、大学などを中心とするオープン化を軸にした動きは今後の動向が注目される。ロボットは複数のセンサによるリアルタイム認識機能を備え、予測や判断により命令を即時にアクチュエータに伝え動作させる必要があり、リアルタイム性の高いコンピューティング技術が期待される。

(2) 警備・防犯

この分野では主に、監視カメラなどの用途が主流になるため、画像／映像／音声データ圧縮などの技術や、認識技術が必須になる。街や道路、店舗、ビルなどいたるところに監視カメラが配置され、科学捜査に代表されるように犯罪や事故などの防止に利用されている。ここでの問題はプライバシー保護や暗号化である。自由主義圏では、画像の取得等はプライバシー保護の観点から、大量の画像データの安易な利用はできない（許可が必要）。しかし、共産圏など（特に中国）では、政府の政策により個人情報をも自由に研究に利用できることもあり、顔認識などの技術が急ピッチで進歩している。

(3) スマート工場

ドイツ主導で進められているインダストリー 4.0 は日本では第 4 次産業革命とも言われている。第 4 次産業革命に対応した工場のことをスマート工場（賢い工場）と呼ぶ。その特徴は 3 つあり、①顧客の要求 1 個にも応じるカスタマイズ性、②注文即生産のリアルタイム性、③低価格、である。そしてそれを実現するための 3 つのポイントは、①センサーネットワークによる自動化、②ビッグデータと AI による品質向上、③データの標準化とサプライチェーンの統合、である。（http://www.oki.com/jp/iot/doc/2016/16vol_12.html）

今や大量生産の時代は終わり、個々人のカスタマイズ化へと市場の流れが変わりつつある。それゆえ、生産工程の最適化や工程管理において、従来の大量データではなく、多品種少量に適した工程の自動化が検討されている。例えば、注文された商品を自動でとりにいくロボットや、多品種少量に合わせたプログラムレス・学習型のロボットなどが研究され利用され始めている。また植物工場では、各種センサを配置し、工場全体の管理、監視を行い、植物の成長を適正に行えるように温度や水や照明、CO₂ 等の管理、制御を行うなどの流れがある。

生産現場でのロボットに特に求められているのは、熟練作業者の経験と勘をいかに自動化するかで、すでに大量のセンサデータ等があるところは、AI によりそれらを熟練者がいなくともできるように学習解析している。その代表的なサービスは IBM の Watson の生産現場における活用である。

(4) 健康・医療

病理画像診断では、CT や MRI、OCT (Optical Coherence Tomography) 等の画像処理技術が進み、細胞内を断層像として 3D 化し、細胞の中などの見えない部分の可視化による診断が画像解析技術として機器開発とともに進んでいる。また、内視鏡などのカメラ技術により人間の体の中を覗き、その場で腫瘍の切除などをリアルタイムで行えるようになってきている。ここでのキーは、ハードウェアの進歩もあるが、そのリアルタイムでの画像診断技術である。医者を支援する技術は、特に深層学習に代表される画像認識技術の発展が著しく、データの蓄積により医療のノウハウが蓄積され進化していく。

健康、リハビリテーションに関しては、人体の動きを高速度カメラで計測し、その動きを解析するとともに、人の失った機能や不自由となった手や足を、ロボットで介護、支援する技術がでてきている。健康人の動作をロボットがアシストし、訓練させることで、リハビリテーションや失った機能の回復にも役立っている。また、事故などで失われた機能を指令するチップを脳内に埋め込む等の研究もおこなわれている。

失った機能のロボットによる代替、自律移動介護車、義足・義手、スマートスーツなど、エッジ・コンピューティングが活躍する現場は数多いと予想される。これら機能を実現させる機器やチップの超小型化、超低消費電力化はこの分野では必須であり、まさにエッジ・コンピューティングの活躍の場であり、日本での大きな進展が期待されるところでもある。

(5) 社会インフラ保守

老朽化してきた社会インフラ（道路、橋、トンネル、ビルなど）の診断には大量のデータが必要であり、現状ではデータの取得と解析の双方に問題がある。エッジ側のデータの取得には、自動制御機能や画像認識機能、位置認識機能を搭載した自走車やドローンが必要である。ここでも超低消費電力、超小型、超軽量の自律移動ロボットが必要になる。

解析にはその大量の画像データ等を解析する必要があるため、現状ではクラウド側での処理が必要となっている。取得した大量の画像等のデータを通信・転送する有効な手段がないため、この分野は解析自体がデータ取得と相まってあまり発達していない。エッジ側のデータ圧縮技術、通信技術の進展により、大量のデータ取得が容易に行われ、データ転送のネックが解消されれば、解析技術も飛躍的に進歩すると考えられる。

(6) VR

VR、特に Mixed Reality 技術の進展は現在海外で著しい。この技術がアミューズメントから生産現場やメンテナンス、修理などの分野に波及すると大きな市場となることが予測される。人が実際の環境を観察する際に、コンピュータで形成された VR 映像を同時に映し出し重畳することで、実空間上にアノテーション（看板や注意、標識、マニュアル、シミュレーション結果、リアルタイムでの指示等）を人間に提示することが可能となり、建物の案内や展示物の注釈だけでなく、修理作業のガイドや、仮想展示品の表示、あるいは、クレイモデルの 3D 表示など、設計や製作、その他様々な現場での活用が考えられる。バーチャル空間を生み出すことで、実際のものがなくとも、触ることや見ることができ移動する必要がなくなる。そのため、この分野は爆発的に発展する可能

性がある。本技術のキーはエッジ・コンピューティングで、人が装着する場合は超小型軽量かつ低消費電力で人間に負担のない表示装置が必須である。現状では重さと人間への負担が問題であり普及していない。

また、この技術による遠隔操作・操縦技術も進展しており、地球の裏側でも遠隔操作により作業が可能となることから可能性は図り知れない。特に災害現場や宇宙等、人間の入り込めない空間や行くことのできない場所でのロボットの作業は重要である。この場合のキーは、ロボット側の視覚だけでなく聴覚や触覚、臭覚など五感に相当するものをいかに人間に不自然なく伝え、制御と動作をロボットと人間が相互にコミュニケーションし一体化できるかである。

(7) 自動翻訳

Google や Microsoft、Apple 等の深層学習等による音声認識技術の発展は顕著で、個人の音声レベルではすでに十分実用化がなされている。さらに自動翻訳がリアルタイムで可能となれば、世界中の言語の壁が取り払われる。コミュニケーションが容易になることで、世界中の人との交流が促進され、相互理解が深まることが予想される。このようなデバイスはスマートフォンや時計等のウェアラブルデバイス、さらには人体にも組み込まれ、誰でもがどこの国の人とも自動翻訳でコミュニケーションできる未来が想像される。

これには、様々な環境でもノイズに埋もれた音声を抽出する技術と、言語や意味を理解する技術の進展がキーとなる。これらの技術が進めば、従来のキーボードだけではなく、自動翻訳も同時にできる音声タイピングも実現できる。

付録4. 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、コラム 1 を参照) 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。

計算ドメイン

任意の特定目的に対応した一連の技術による計算機能を、計算ドメインと呼ぶ。例えば、リアルタイムの画像処理、画像認識、音声認識、自律制御などが計算ドメインの種類である。特定の計算領域（ドメイン）向けにコンピューティング技術を最適化することで、高い性能・機能を発揮することができる。

CMOS

Complementary Metal Oxide Semiconductor (相補型金属酸化膜半導体)。電子が多い n 型と正孔が多い p 型の MOS トランジスタの組み合わせで構成される半導体回路の一種。消費電力が少ない、小型化・高集積化に適するため製造コストを安くできるといった利点から、現在のコンピュータに使用されている CPU のほとんどが CMOS を採用している。

ロバスト性

外乱に対する強さを意味する語で、日本語では堅牢性や強靱性。システムや機械が持つ、外的要因による変化を内部で阻止する仕組みや性質を意味する表現。ロバストネスともいう。外乱を受けても挙動が安定していたり、外部からの影響を最小限に抑えたりする仕組みのこと。

アルゴリズム

コンピュータで計算を行う際の、問題を解くための処理手順を定式化した形で表現したもの。「計算方法」のこと。問題を解く手続きを一般化するものであり、コンピュータ処理するための具体的な手順を記述したものがプログラムである。

アーキテクチャ

コンピュータにおけるハードウェア、OS、ネットワーク、アプリケーションソフトなどの基本設計や設計思想などを意味する。「方式」という語が使われることもある。ソフトウェアアーキテクチャ、システムアーキテクチャ、のように用いられる。情報システムにおいて、アーキテクチャの選択や個々のアーキテクチャ間の整合性などが重要になる。元々は建築分野における設計思想や建築様式のことを意味していたが、1964年、

IBM が最初の汎用コンピュータの設計思想を表現するためにアーキテクチャの語を用いたことに由来する。

ワークロード

稼働中の IT システムにかかる負荷またはその度合いを表す。CPU やメモリの占有率などで測る。計算機資源がその最大容量に対してどのくらいの割合が発揮・利用されているかを表す。例えば、「浮動小数点演算」「データベース処理」「画像処理」など、アプリケーションごとに重要となる様々な処理のタイプを指して用いる。

エッジ

コンピュータネットワークの末端部をエッジと呼ぶ。ネットワークを介してサーバーの集約化を図るのがクラウド・コンピューティングであることに对比して用いられるのが、エッジ・コンピューティングである。コンピュータネットワーク上で、利用者の端末に程近い場所にサーバーを分散配置し、クラウドでは難しいとされる負荷の分散や通信の低遅延化を図ることができる。リアルタイム処理を必要とする MtoM や IoT への対応が可能となる。

アナログ回路

連続的に変化する信号を取り扱う電子回路のこと。これに対するデジタル回路は、通常は 0 と 1 の 2 値しか持たない信号を扱う。アナログ回路では、回路における物理量の連続的な変化を利用して情報処理をする。一方、デジタル回路では、ある物理量（閾値）より大か小かで 0 か 1 かを判定する。安価で大量複製が可能という観点からデジタル化が進んでいるが、現実世界との接点では必ずアナログ量の入出力が必要となるため、いかにアナログ量の情報を失わずにデジタル量に変換するか、またはいかにデジタル量の情報を失わずに目的のアナログ量に変換するかが、アナログ回路には求められ、IoT を支えるキー技術の一つとされる。

アクセラレータ

コンピュータの処理能力を高めるために、追加して利用するハードウェアやソフトウェアの総称。アクセラレータを組み込むことによってコンピュータの処理能力を高めることができる。より高性能な CPU に置き換えて処理能力を高める CPU アクセラレータや、グラフィック表示を高速化するグラフィックアクセラレータ、暗号通信を高速化する SSL (Secure Socket Layer) アクセラレータ、web サーバーとの接続を高速化するインターネットアクセラレータなどが知られる。

ディープニューラルネットワーク

計算がおこなわれる層が多数積み重なったニューラルネットワーク (Neural Network) による機械学習手法のこと。一般に深層学習 (Deep Learning) として表現される。ニューラルネットワークは、人間の脳神経細胞を模倣したアルゴリズムで、パターンが認識できるように設計されている。データを効率的にクラスタリングや分類するのに役立つ。

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) と上述の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である。

GPU

グラフィックス プロセッシング ユニット (Graphics Processing Unit)。元々は 3D 画像処理に特化した演算プロセッサであったが、CPU よりも並列演算性能に優れていることから、現在はより一般的な計算に活用する GPGPU (General-purpose computing on graphics processing units; GPU による汎用計算) がさかんに行われるようになってきている。並列性が高い深層学習の処理に多用されており、映像出力端子を持たない専用製品も多く現れている。

FPGA

FPGA (field-programmable gate array) は、製造後に購入者や設計者が現場で論理回路の構成を設定・再構成できる集積回路である。現場でプログラム可能な回路アレイであることから、このように呼ばれている。通常の LSI では、設計者が作成した回路図が製造時に固定的に論理回路を形成し、製造後に変更することはできないが、FPGA は製造後に外部から設計情報を送り込んで任意の論理回路を構成することができる。また、何度も書き換えて動作を変更することができる。

Python

汎用のプログラミング言語であり、フリー・オープンソフト。コードがシンプルで扱いやすく設計されているとされ、C 言語などに比べて、さまざまなプログラムを少ないコード行数で書くことができるといった特徴がある。組み込み開発、WEB アプリケーションなどで広く利用されており、人工知能や機械学習の分野では、Python が使われているケースが多くなっている。

ストレージクラスメモリ

メモリとストレージの両方の特徴を併せ持つ技術として、NAND 型フラッシュメモリより高速で、低消費電力かつ長寿命で容量単価が DRAM と同等以下の不揮発性メモリのこと。NAND 型フラッシュメモリは HDD より速いが、DRAM と比べるとアクセス速度にギャップがあり、データ処理のボトルネックになっている。このギャップを埋める高速な次世代メモリの開発が活発化しており、このような次世代メモリを指す。ストレージクラスメモリには、MRAM、ReRAM、PRAM、FeRAM、STT-MRAM などがある。

イジング模型 (Ising Model)

結晶を構成する原子のスピン向きを計算する簡易的なモデルである。実際の磁性体

を単純化して解析した統計力学上のモデル。D-wave systems 社の量子アニーリングマシンはこの発想にもとづいており、量子化された Ising Model を計算するための専用のプロセッサを搭載している。多数のスピンの状態を並列にシミュレーションすることで、より高速かつ正確に Ising Model の最小値問題を解くことができる。シミュレーションの際、高温状態から出発して、徐々に温度を下げながら計算を進めていくことで、本当の最小値に到達できるようにする。

大規模グラフ処理

大規模グラフとは、ビッグデータの持つデータ構造の一つであり、人・モノ・場所といった多様な情報のつながりを表現するデータ構造を処理することを指す。大規模グラフは、データおよびデータどうしの関連性をそれぞれ「ノード」と「エッジ」で表現するデータ構造を持ち、たとえばこのデータ構造は Web ページのリンク関係や SNS の交友関係、商品の購買関係など様々なものがある。近年の web や SNS の広まりに伴い、グラフとして取り扱うデータサイズが劇的に大きくなったことから、大規模グラフをいかに短時間で効率的に処理・最適化するかが課題となっている。

レイテンシ

データ転送における指標のひとつで、遅延時間のことをいう。CPU やハードディスクなどのデータの処理待ち時間や、通信におけるスイッチング遅延などがある。例えばメモリやハードディスクなどの記憶装置からデータを読み出す際に、命令を送ってから記憶装置におけるデータが記録された箇所までアクセスする時間と、記憶装置から読み出されたデータが命令の結果として返ってくるまで、遅延時間が生じる。また、コンピュータの CPU とメモリとの間ではクロック周波数が異なるため、処理にレイテンシが生じやすい。通信のレイテンシは、即応性が求められるサービスでは特に重視されている指標である。

コンパイラ

人間が理解できる言語や数式で記述されたプログラムを、コンピュータが解釈・実行できる機械語の形式に変換するプログラムのこと。プログラミング言語は人間向けの文法を持った言語であるため、そのままではコンピュータには理解できない。そこで、プログラミング言語を用いたソースコードをコンピュータが理解できるように翻訳する必要がある。この翻訳のことを「コンパイル」といい、コンパイルに用いるプログラムを「コンパイラ」という。

■戦略プロポーザル作成メンバー■

総括責任者：木村 康則	上席フェロー	(システム・情報科学技術ユニット)
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：河村誠一郎	フェロー	(ナノテクノロジー・材料ユニット)
嶋田 義皓	フェロー	(システム・情報科学技術ユニット)
永野 智己	フェロー	(ナノテクノロジー・材料ユニット)
福島 俊一	フェロー	(システム・情報科学技術ユニット)
藤井新一郎	フェロー	(システム・情報科学技術ユニット)
的場 正憲	フェロー	(システム・情報科学技術ユニット)
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)
高地 伸夫	主任調査員	(戦略研究推進部 ICT グループ)

※お問い合わせは、下記ユニットまでお願い致します。

CRDS-FY2017-SP-02

戦略プロポーザル

革新的コンピューティング

～計算ドメイン志向による基盤技術の創出～

STRATEGIC PROPOSAL Disruptive Computing

- A Computing-Domain-Oriented Approach -

平成 30 年 3 月 March, 2018

ISBN 978-4-88890-582-4

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 システム・情報科学技術ユニット

Systems and Information Science and Technology Unit,
Center for Research and Development Strategy, Japan Science and Technology Agency

〒102-0076 東京都千代田区五番町 7

電 話 03-5214-7481

ファックス 03-5214-7385

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

©2018 JST/CRDS

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。

No part of this publication may be reproduced, copied, transmitted or translated without written permission. Application should be sent to crds@jst.go.jp. Any quotations must be appropriately acknowledged.

ATTAATC A AAGA C CTAAC T CTCAGACC
CT CTCGCC AATTAATA
TAA TAATC
TTGCAATTGGA CCCC
AATTCC AAAA GGCCTTAA CCTAC
ATAAGA CTCTAACT CTCGCC
AA TAATC

ISBN 978-4-88890-582-4

AAT A TCTATAAGA CTCTAACT CTAAT A TCTAT
CTCGCC AATTAATA
ATTAATC A AAGA C CTAAC T CTCAGACC
AAT A TCTATAAGA CTCTAACT
CTCGCC AATTAATA
TTAATC A AAGA C CTAAC T CTCAGACC
AAT A TCTATAAGA CTCTAACT
ATTAATC A AAGA C CT
GA C CTAAC T CTCAGACC

0011 1110 000
00 11 001010 1
0011 1110 000
0100 11100 11100 101010000111
001100 110010
0001 0011 11110 000101

