東北大学
TOHOKU UNIVERSITY

Press Release

2023 年 12 月 13 日

国立大学法人東北大学
国立研究開発法人科学技術振興機構（JST）

AI 処理を高速・超低電力で行う新技術を開発
～現行 AI の計算方式に対応した スピントロニクス『P』コンピュータの動作を実証～

【発表のポイント】

- 高速・超低電力での演算が可能なスピントロニクス(注 1)技術を用いた確率論的（『P』）コンピュータ(注 2)で人工知能（AI）処理を行う新技術を開発
- 現行 AI で利用されている「順伝播型ニューラルネットワーク」の動作を実証
- 病気の原因の推定、気象予測などを超低電力で行う技術への応用が期待

【概要】

人工知能（AI）やデジタル社会の進展に伴い、コンピュータで処理するタスクは複雑かつ多様化しています。この要請に応えるため、各用途に特化した新概念コンピュータの研究開発が活性化しています。

スピントロニクス確率論的（『P』）コンピュータは確率性を伴う複雑な問題を省電力で超高速に処理できると期待される新概念コンピュータの一種です。東北大学とカリフォルニア大学サンタバーバラ校（アメリカ）のチームは以前から共同研究を行っており、昨年実験結果に基づき汎用的なコンピュータと比べて演算速度を約 5 桁向上、消費電力を約 1 桁低減できることを示していました。一方で昨今急成長する AI 技術への応用に向けては、現行 AI の大部分が採用する計算方式である順伝播型ニューラルネットワーク[図 1]に整合した技術の開発が求められていました。

今回研究チームはスピントロニクス P コンピュータにて順伝播型ニューラルネットワークに基づく計算を行うための新技術を開発し、行動履歴や生活習慣と病気の発症の因果関係を確率的に解析する AI 計算のデモ実験などに成功しました。また併せて、これまでの P コンピュータの動作速度を 3 桁向上する新素子技術を開発しました。

AI の更なる発展に向けてはコンピュータの演算能力の向上と省エネ化の両立が喫緊の課題となっています。スピントロニクス P コンピュータはまさにこの要請に応えるものであり、本研究にて現行 AI と高い整合性を有した技術が確立されたことから、今後社会実装に向けた研究開発がより一層進展するものと期待されます。

本成果は、2023 年 12 月 9-13 日（米国時間）に米サンフランシスコで開催される学術会議「International Electron Devices Meeting: IEDM」で発表されます。

【詳細な説明】

研究の背景

AI（人工知能）は言うまでもなく現代の社会基盤の一つであり、直近では生成 AI の出現が社会に様々な革新をもたらそうとしています。一方で AI の発展はコンピュータの演算能力に過剰な要求を突き付けてもいます。コンピュータの演算能力は 1960 年頃から約 50 年にわたって、ほぼ 1.5 年ごとに 2 倍のペースで向上してきました（ムーアの法則）が、2010 年以降はこれを 5 倍以上上回るペース（1 年で約 10 倍）で AI 計算量が増加しています。このような AI 計算量の増大はデータセンターでの電力消費を 3 倍増加させているとも言われており、持続可能な社会の実現を目指す以上、看過できない懸案でもあります。この問題に対処するためには、AI 計算を超低電力で行える革新的なコンピューティング技術の導入が不可欠です。特にコンピュータに求められるタスクが多様化する昨今、全てに対処し得る万能コンピュータの実現を期待することは非現実的であり、各用途に特化した様々なコンピュータ（Domain-specific computer）を開発して使い分けていくことが有効と考えられています。

複雑な問題（例えば気象予測など）を処理するソフトウェア技術としてしばしば乱数アルゴリズムが用いられています。これは特に確率性を伴う事象を扱うのに適していますが、現行の決定論的に動作するコンピュータとは本質的には相性が悪く、計算に多大な電力が費やされています。この問題を解決する手法の一つとして、1981 年にアメリカの物理学者リチャード・ファインマン博士は、ハードウェアのレベルで確率論的に振る舞うコンピュータ（確率論的（Probabilistic）コンピュータ = P コンピュータ）の可能性に言及しました。しかしそれから約 40 年の間、P コンピュータを実現するための有効策は明らかではありませんでした。

2019 年に東北大学とパデュー大学（アメリカ）の共同研究にて、自然の熱で確率的に状態が更新されるスピントロニクス素子を用いた P コンピュータのデモシステムが構築され、複雑性の高い問題の典型例である組合せ最適化の原理実証が行われました（東北大学プレスリリース『室温動作スピントロニクス素子を用いて量子アニーリングマシンの機能を実現』(注 3)）。また、2022 年には東北大学とカリフォルニア大学サンタバーバラ校（アメリカ）の共同研究にて、実験結果に基づいた P コンピュータの性能予測から、従来型の決定論的に動作する汎用コンピュータと比べて、P コンピュータは乱数アルゴリズムを用いた計算の速度を約 5 桁、消費電力を 1 桁改善できることが示されていました（東北大学プレスリリース『確率動作スピン素子を用いた高性能・省電力「P」コンピューターを実証』(注 4)）。

一方で、スピントロニクス P コンピュータを社会で広く用いられている AI 技術に適用するためには一つの障壁がありました。それは AI 計算とこれまでのスピントロニクス P コンピュータの間での数学的な計算モデル（ニューラルネットワークの種類）の違いに関係します。図 1 に代表的なニューラルネットワークである再帰型ニューラルネットワーク（Recurrent Neural Network）と順伝播型ニューラルネットワーク（Feed-

forward Neural Network）の構造を示しました。再帰型ニューラルネットワーク（図 1(a)）では、ノード間での情報の行き来の方向が定まっていない（例えばノード A はノード B に影響を与え、同時にノード B はノード A に影響を与える）のに対して、順伝播型ニューラルネットワーク（図 1(b)）はノードからノードへと情報が流れる方向が定まっています。順伝播型ニューラルネットワークは現行の逐次的に演算を行うコンピュータとの相性が良くほぼ全ての AI で用いられているのに対して、これまでに開発された P コンピュータの演算では再帰型ニューラルネットワークが用いられていました。すなわち、順伝播型ニューラルネットワークに対応したスピントロニクス P コンピュータを実現できれば、現在汎用コンピュータを用いて行われている AI 処理をそのままスピントロニクス P コンピュータにてより高速にかつ低い電力で実行でき、それを契機にその後は P コンピュータの特徴を活かした様々なアプリケーションが開拓されていくものと期待されます。

今回の取り組み

今回、東北大学の金子遥南氏（大学院生・工学研究科）、金井駿准教授（電気通信研究所）、大野英男教授（現総長）、深見俊輔教授（電気通信研究所）らは、カリフォルニア大学サンタバーバラ校の Kerem Camsari 博士らのチームと共同で、スピントロニクス P コンピュータを順伝播型ニューラルネットワークに適応させる新技術を開発し、その基本動作を実証しました。

図 2 に構築したスピントロニクス P コンピュータの原理実証システムの写真が示されています。システムは 4 つの確率動作スピントロニクス素子を搭載した確率ビット（Probabilistic bit: P ビット）[注 2]のユニット（左側）と、プログラマブル半導体（Field-Programmable Gate Array: FPGA）[注 5]（右側）から構成されています。FPGA はどのスピントロニクス P ビットを動作させるか指定し、指定されたスピントロニクス P ビットは物理乱数を生成して FPGA を駆動する、という手順を繰り返すことで演算が行われるよう設計されています。

このシステムを用い、順伝播型ニューラルネットワークの一種であるベイジアンネットワーク[注 6]の動作実証を行いました。図 3 にその結果の一例が示されています。これは 1980 年代に提唱された「アジアネットワーク」と呼ばれるベイジアンネットワークの教科書的な例題であり（S. L. Lauritzen and D. J. Spiegelhalter, J. R. Stat. Soc., B: Stat. 50, 157-224 (1988)）、アジアへの旅行歴や喫煙習慣の有無と X 線検査での異常や呼吸器系の障害の有無の間の因果関係を表したものです。図 3(a)に示されるように、ベイジアンネットワークを構成する 4 層を 4 つのスピントロニクス P ビットで順次更新する手順を 10,000 回繰り返し、各事象のセットの起こりやすさを計算しました。図 3(b)はその結果です。実験結果（青の棒グラフ）はベイズの定理から予測される理想分布（赤の棒グラフ）と酷似した分布となっており、これはスピントロニクス P コンピュータが期待通りの動作をしていることを意味しています。

また今回の実験では、スピントロニクス P ビットの素子にも工夫が施されています。一般に P コンピュータの性能は P ビットの乱数生成の速度で決定され、速ければ速いほど高速での演算が可能となります。図 4(a)にはこれまでの研究と今回の研究で用いられた確率動作スピントロニクス素子の膜構成が示されています。今回の研究では、これまでの東北大学の研究成果に基づき、コバルト鉄ホウ素（CoFeB）とルテニウム（Ru）が積層された構造（CoFeB/Ru/CoFeB）が用いられました。これによって外乱磁界に対する頑強さと、高速での乱数生成が実現されています。従来の構成ではミリ秒の時間スケールで乱数生成が可能であったのに対し、図 4(b)に示されているように本研究で用いた P ビットはマイクロ秒（0.001 ミリ秒）で物理乱数を生成しています。なお東北大学のチームはさらに 2 桁以上速い 10 ナノ秒（0.01 マイクロ秒）以下で乱数を生成する素子技術も確立しており（東北大学プレスリリース『スピントロニクス疑似量子ビットを従来比 100 倍超に高速化』(注 7)）、今後の更なる高速化の余力を残しています。

今後の展開

本研究により、これまで再帰型ニューラルネットワークの動作実証に限られていたスピントロニクス P コンピュータが、順伝播型ニューラルネットワークにも対応できることが明らかになりました。先述の通り、順伝播型ニューラルネットワークは、現行の AI の大部分が採用する計算モデルです。特に今回動作実証されたベイジアンネットワークは、複雑な事象が絡み合う一連のプロセスを解析するのに適しており、気象予測や病気の原因の推定、故障診断、マーケティングなどで広く利用されています。現在これらの計算はスーパーコンピュータ（スパコン）などの汎用性の高いハードウェアで比較的大きな電力を用いて実行されていますが、本研究によりスピントロニクス P コンピュータを用いてこれらの計算を超低電力で行う道筋が開かれました。すなわち今回開発された技術はスピントロニクス P コンピュータの社会実装の扉を開く鍵となり、それを皮切りにして P コンピュータの特徴を活かした様々なアプリケーションが開拓されていくことが期待されます。

今後の AI の発展に向けては、コンピュータの演算能力の更なる向上と省電力化の両立が求められています。一方、スピントロニクス P コンピュータは従来型の決定論的かつ逐次的な動作を基本とするコンピュータと比べて演算能力と省エネ性を劇的に向上しうるポテンシャルを有しています。すなわち本研究成果はまさに社会的要請に応えるものであると言え、今後大規模化、集積化に向けた研究開発が進展し、利便性とエネルギー効率の高い社会の実現に向けて利用されていくことが期待されます。

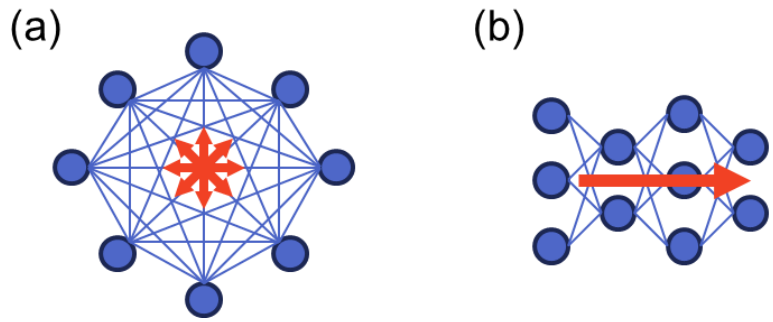

図 1. 再帰型ニューラルネットワーク（Recurrent Neural Network）(a) と、順伝播型ニューラルネットワーク（Feed-forward Neural Network）(b) の模式図。再帰型ニューラルネットワークは情報が双方向に流れるのに対して、順伝播型ニューラルネットワークは情報が流れる方向が規定されている。順伝播型ニューラルネットワークは現行の決定論的かつ逐次的に動作するコンピュータと相性が良く、大部分の AI 処理において利用されている。

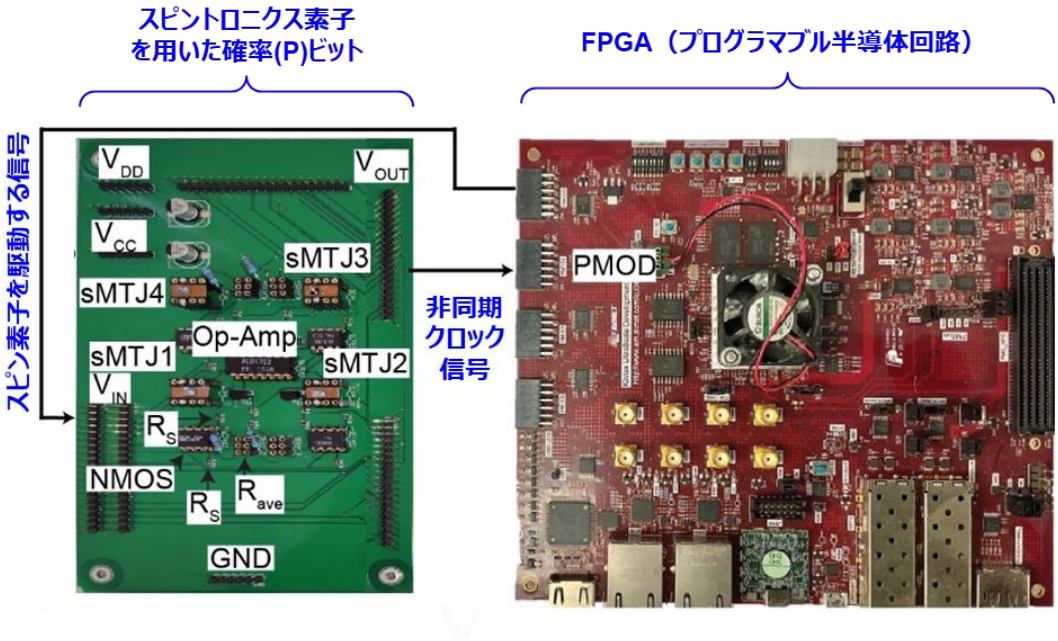

図 2. 作製したスピントロニクス P コンピュータの原理実証システムの写真。左側は確率動作スピントロニクス素子からなる P ビットのユニットであり、右側はプログラマブル半導体回路（FPGA）。

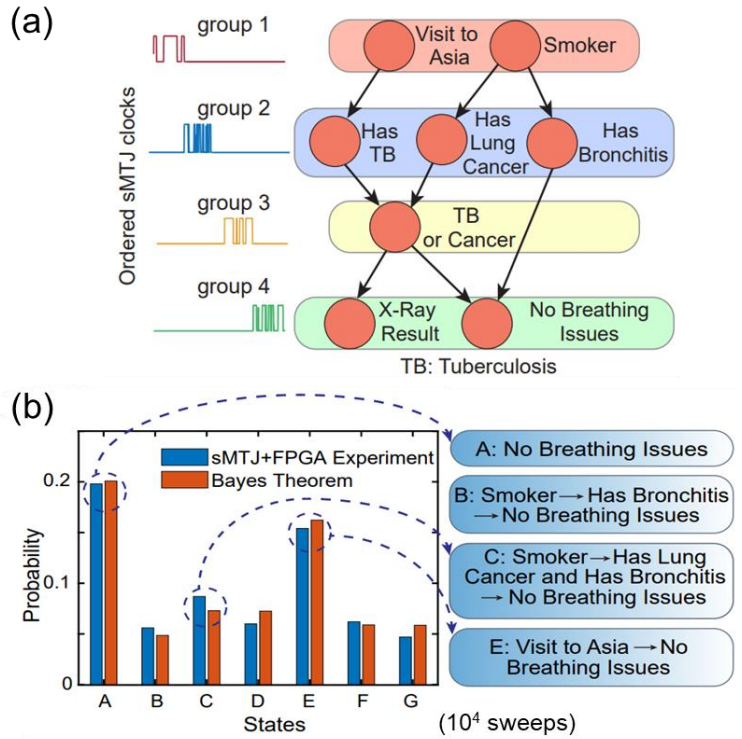

図 3. ベイジアンネットワークの構成、スピントロニクス P ビットからの出力信号(a)と、実験結果 (b)。

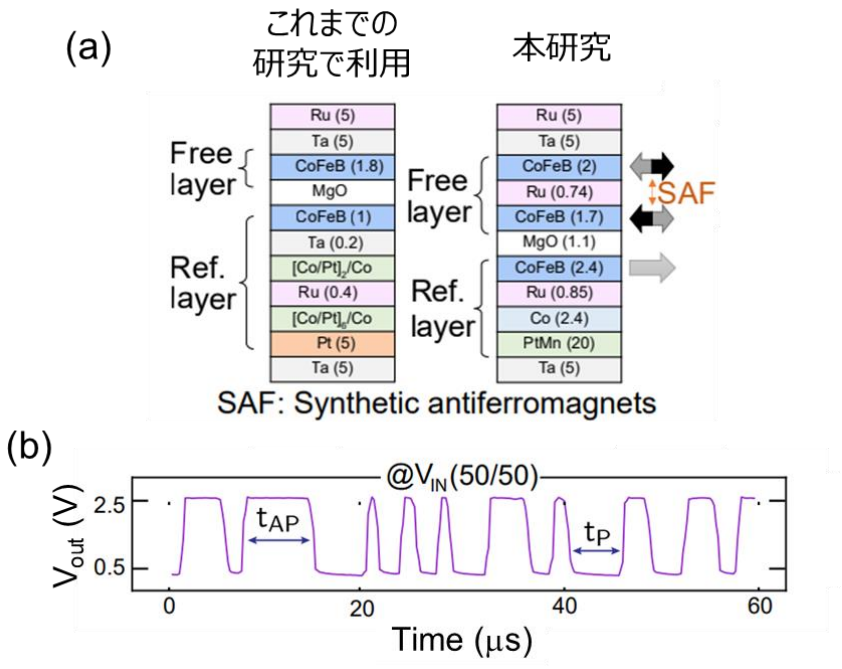

図 4. 従来研究を3桁上回る高速乱数生成動作が可能なスピントロニクス素子の膜構成（a の右側）と、今回作製した P ビットが生成するランダムテレグラフノイズの測定結果（b）。

【謝辞】
本研究は、科学技術振興機構（JST）世界のトップ研究者ネットワーク参画のための国際研究協力プログラム（AdCORP）「スピントロニクス確率論的コンピュータの大規模化に向けた材料・素子・回路・アルゴリズム融合研究」（研究代表者：深見 俊輔）JPMJKB2305、戦略的創造研究推進事業 CREST「スピンエッジコンピューティングハードウェア基盤」（研究代表者：佐藤 茂雄）JPMJCR19K3、および同事業さきがけ「不確定性スピントロニクス素子」（研究代表者：金井 駿）JPMJPR21B2、などの支援を受けて行われたものです。

【用語説明】

注1. スピントロニクス

電子の持つ電気的性質（電荷）と磁気的性質（スピン）を同時に利用することで発現する物理現象を明らかにし、工学的に利用することを目指す学術分野。例えば従来は不可能であった磁気的性質や磁化方向の電気的な検出や制御（スピントルク磁化反転）、電気伝導特性の磁場や磁化による制御（磁気抵抗効果）などが可能となり、現在も様々な現象が発見され続けている。

注2. 確率論的（『P』）コンピュータ、確率ビット（P ビット）

確率ビット（P ビット）とは、短時間で出力信号が 0 と 1 の間で確率的に変化し、かつ各ビットを電気的に相関させられる情報処理の基本単位。確率論的コンピュータ（P コンピュータ）は P ビットを用いて演算を行うコンピュータ。

P ビットは 0 と 1 の重ね合わせ状態を持ち、かつビット間でもつれあい（相関状態）を形成できる量子ビット（Q ビット）とは本質的に異なるが一定の類似性があることから、確率論的コンピュータは量子コンピュータと並んで新概念コンピュータの一つとして注目されている。1981 年にリチャード・ファインマンが行った講演において、量子コンピュータと並んで、確率的な現象を効率的に計算する仕組みとして紹介されている。

注3. 東北大学 2019 年 9 月 18 日プレスリリース
『室温動作スピントロニクス素子を用いて量子アニーリングマシンの機能を実現』
https://www.tohoku.ac.jp/japanese/2019/09/press20190918-01-spin.html

注4. 東北大学 2022 年 12 月 7 日プレスリリース
『確率動作スピン素子を用いた高性能・省電力「P」コンピューターを実証』
https://www.tohoku.ac.jp/japanese/2022/12/press20221207-02-spin.html

注5.　プログラマブル半導体（Field Programmable Gate Array; FPGA）

ユーザーが現場（Field）で論理回路の機能をプログラムできる論理集積回路。パソコンの頭脳である CPU（Central Processing Unit）と比べると、汎用性では劣るものの、論理回路の構成を変えられることから、ユーザーがプログラムした計算を行う速度は速くなる。一方、ASIC（Application Specific Integrated Circuit）と比べると、速度では劣るものの、ユーザーが機能を書き換えられ汎用性が高いという特徴がある。

注6.　ベイジアンネットワーク

ある事象が起こった際に次の事象が起こる確率を意味する「条件付き確率」（例えば、雲が出ているという条件下で、雨が降る確率が 30%、降らない確率が 70%）をもとに、確率的現象の集合をグラフィカルに整理する計算モデル。ある初期状態から期待される最終状態を予測する、あるいはある観測結果をもとにその原因を推定する際などに用いられる。後者は「ベイズ推定」と呼ばれる。

18 世紀の英国の数学者、哲学者であるトーマス・ベイズにより考案された。

注7.　東北大学 2021 年 3 月 18 日プレスリリース
『スピントロニクス疑似量子ビットを従来比 100 倍超に高速化』
https://www.tohoku.ac.jp/japanese/2021/03/press20210318-02-bit.html

【論文情報】

タイトル：“Hardware Demonstration of Feedforward Stochastic Neural Networks with Fast MTJ-based p-bits”（高速磁気トンネル磁気トンネル接合からなる確率ビットを用いた順伝播型確率的ニューラルネットワークのハードウェア実証）

著者：Nihal Sanjay Singh, Shaila Niazi, Shuvro Chowdhury, Kemal Selcuk, Haruna Kaneko, Keito Kobayashi, Shun Kanai, Hideo Ohno, Shunsuke Fukami and Kerem Y. Camsari

*責任著者：カリフォルニア大学サンタバーバラ校 Kerem Camsari

掲載誌：69th Annual IEEE International Electron Devices Meeting (IEDM 2023)

DOI：1 月以降に付与

URL: 1 月以降に決定

【問い合わせ先】
（研究に関すること）
東北大学電気通信研究所
教授　深見 俊輔
TEL: 022-217-5555
E-mail: s-fukami[at]tohoku.ac.jp
（兼）東北大学先端スピントロニクス研究開発センター (CSIS)
（兼）東北大学国際集積エレクトロニクス研究開発センター (CIES)
（兼）東北大学材料科学高等研究所 (WPI-AIMR)
（兼）稲盛科学研究機構 (InaRIS)

（報道に関すること）
東北大学電気通信研究所 総務係
TEL: 022-217-5420
E-mail: riec-somu[at]grp.tohoku.ac.jp

科学技術振興機構 広報課
TEL: 03-5214-8404
E-mail: jstkoho[at]jst.go.jp

（JST 事業に関すること）
科学技術振興機構 国際部
佐藤 正樹
TEL: 03-5214-7375
E-mail: adcorp[at]jst.go.jp