

東工大・情報通信系 大学院説明会 本村・劉研究室 紹介スライド

2020年 4月 30日

東京工業大学 科学技術創成研究院
AIコンピューティング研究ユニット (ArtIC)
情報通信系 本村研究室 劉研究室

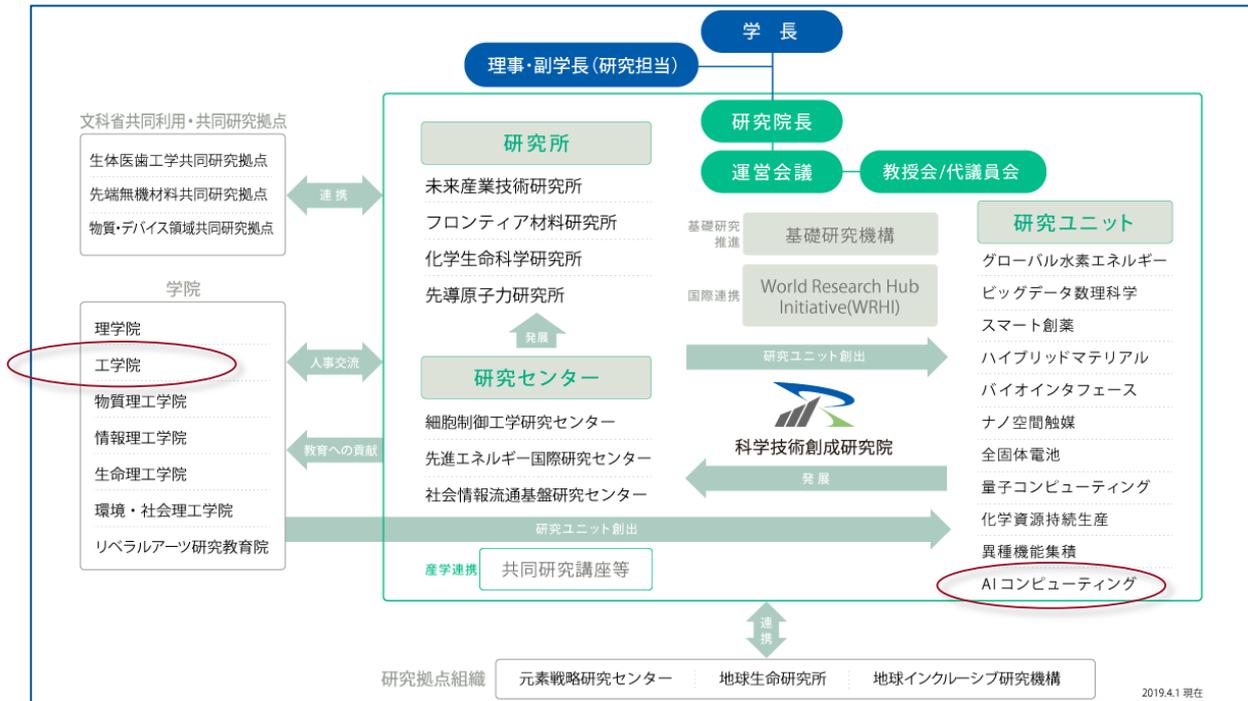


Tokyo Tech



ArtICの成り立ち

工学院と科学技術創成研究院



Artificially Intelligent Computing Research Unit

もう一つの意味: ArtなIC

=> 素敵なハードウェア

=> ソフトとハードの協調研究



2020年代: コンピューティングの新時代

AIコンピューティング

人工知能革命の急進

ムーアの法則の終焉

||

||

ポストノイマン時代

ポストムーア時代

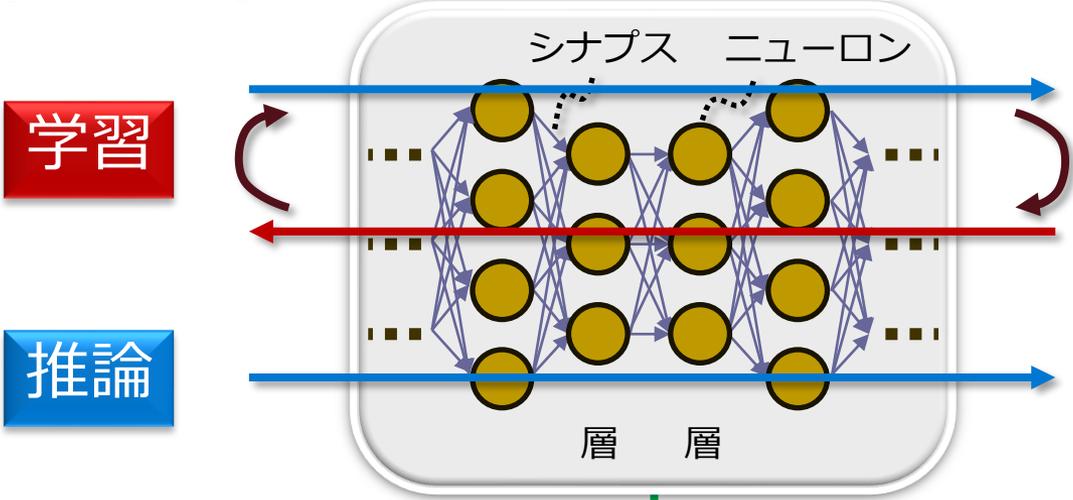
情報処理ハードウェアに
変化のチャンスが到来
- Pull -

情報処理ハードウェアが
変化せざるを得ない
- Push -

大きな成果を上げるチャンス
社会的に大きな意味のある研究

AIコンピューティングとは

深層(ディープ)ニューラルネットワーク (ディープラーニング)



- 既知データで**伝搬**と**逆伝搬**を反芻
- 所望の結果に向けてシナプス値調整

- **未知**データを**伝搬**させ結果を出力

正帰還ループにより爆発的に発展

高速な計算ハードウェア

大量のデータ

我々は深層ニューラルネットなどのAIコンピューティングを加速するハードウェアとソフトウェアの技術を中心に研究を進めています

オープン環境によるアルゴリズム開発

TensorFlow Spark Caffe Chainer cuDNN

arXiv.org

AIコンピューティング研究ユニット: ArtIC

2019年4月
に発足
2020年4月からフル
メンバで活動

すずかけ
台キャン
パス J3棟
17F



<http://www.artic.iir.titech.ac.jp>

ArtICのミッション

- 

新たな AI プラットフォームの創出
Creation of Novel Platform for AI Computing
- 

共通基盤コンピューティングアーキテクチャの構築
Construction of Common-basis Computing Architecture
- 

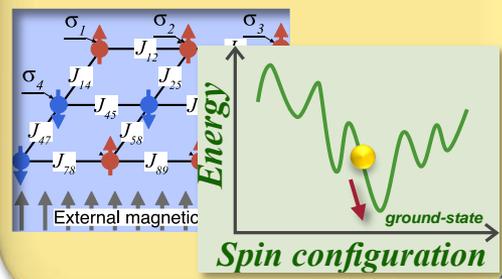
自律性・安全性・エネルギー効率・コスト効率の高いハードウェア基盤の実現
Realization of Hardware Foundation with Good Autonomy, Safety, Energy, and Cost Efficiency



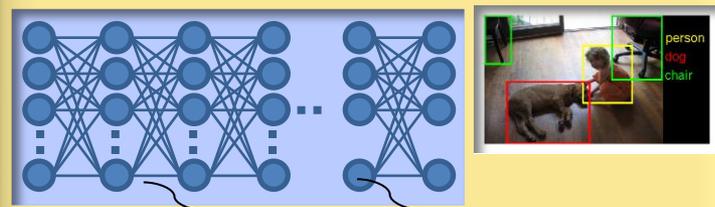
ArtIC: 研究ターゲット

人工知能(AI)応用の急速な拡大
「**コントロール**駆動から**データ**駆動へ」
計算機アーキテクチャの革命

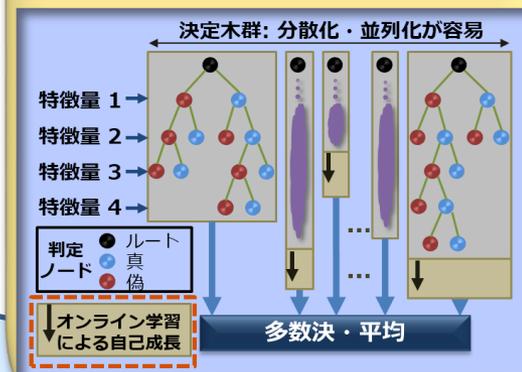
組合せ最適化問題 →
スピン格子のエネルギー
最小化問題



大量データの学習 →
強力な推論・識別・予測能力



説明性・制御性の高さ
と低学習負荷の両立



深層ニューラルネット
・ディープラーニング

アニーリング計算機
(非量子)

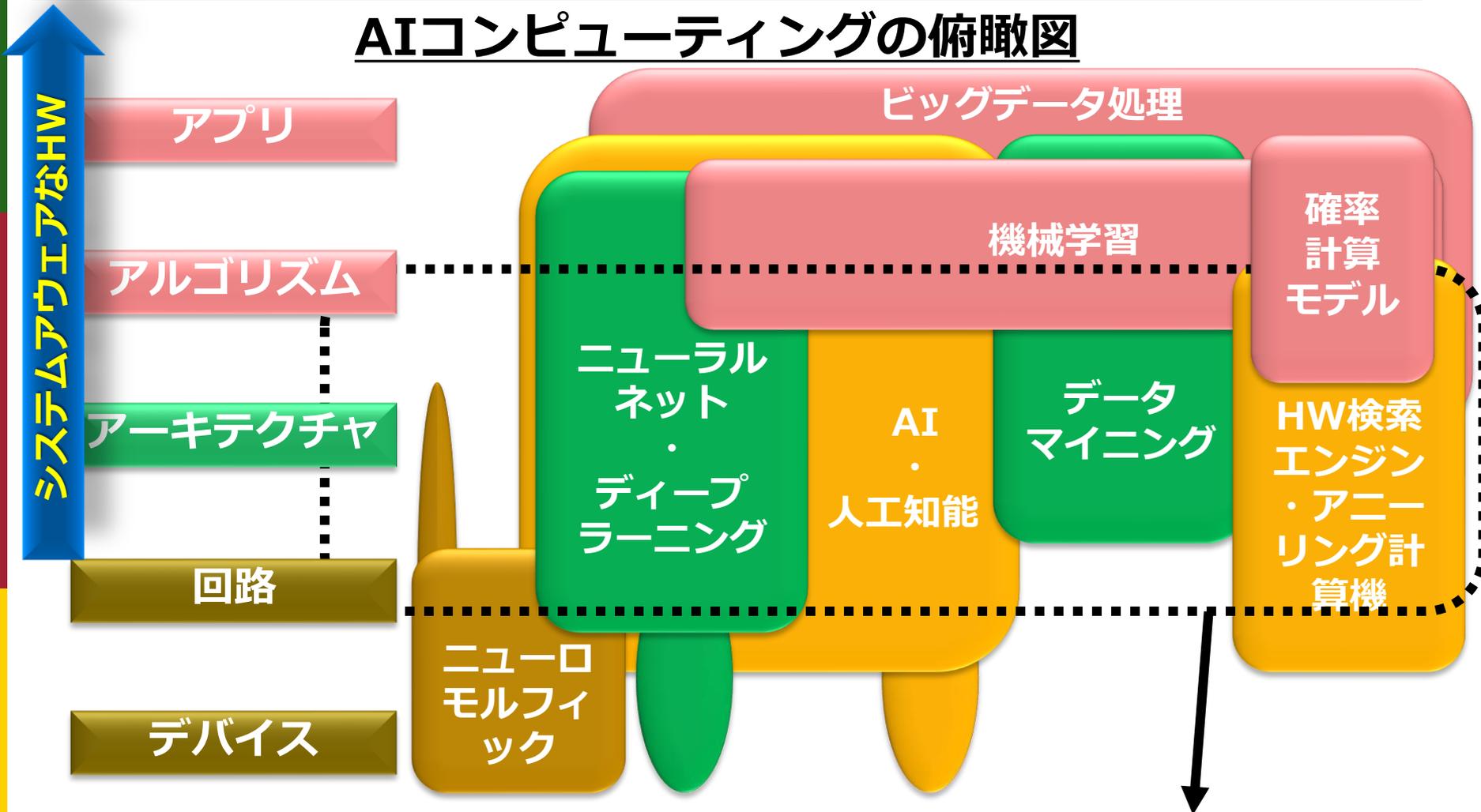
統計的機械学習
(アンサンブル学習等)

構造型情報処理アーキテクチャ
として共通基盤化

アルゴリズム理解 => アーキテクチャ研究 => ハードウェア実現

ArtIC: 研究分野・研究レイヤ

AIコンピューティングの俯瞰図



これら広範囲のAIコンピューティング群を加速する
アーキテクチャプラットフォームの研究を推進する

研究ユニット教員の紹介

本村

'87 京大理学部修士
 '96 京大工学博士
 '87-'11 NEC研究所
 '11-'18 北大
 '19- 東工大

LSIのオリピック
 ISSCC2018で、量子化ニューラルネットチップの発表



4千人参加の最高峰会議

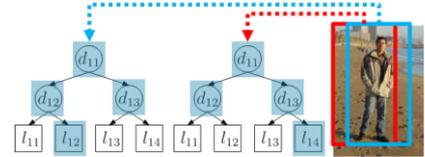
多くのシンポジウムでAIハードウェアの招待講演

劉

'07 京大情報学修士
 '13 阪大情報科学博士
 '13-19 阪大
 '19/10- 東工大

ブースティング決定木の並列アクセラレータ

物体検出における並列処理の場合



FPGA 実装

- Xilinx ZC706 評価ボード
- 1,024並列: M=8, 2次元方向8x16
- 6.6倍x128並列=845倍高速化

リソース使用率

Slice	32Kb BRAM	DSP
18,904 (35%)	186 (34%)	0 (0%)

識別速度

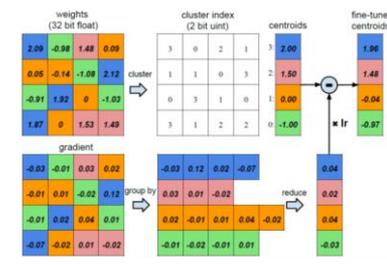
- 歩行者識別: Full HD 350 fps
- BDT: ブースティング決定木

汎用物体認識システム(FPGA利用)



実装	ミス率*	フレームレート	処理検出窓数
既存 SVM	46%	Full HD 60 fps	6,284k/秒
既存 ACF	17%	VGA 30 fps	1,972k/秒
案 ACF	17%	Full HD 170 fps	112,501k/秒

近似計算によるディープラーニング



日経新聞
 17/9/18日朝刊

AI支えるテクノロジー ④専用半導体

高速処理・省エネ両立探る

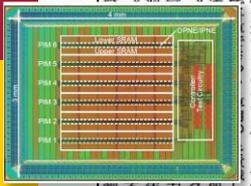
AIの普及には新しい半導体が必要

現在	未来
CPU (中央演算処理装置)	あらゆる計算をこなせる。AI向けの計算は苦手
GPU (画像処理半導体)	AI技術「深層学習」向けの応用に強み

AIに特化した高速処理と省エネを両立

▽主な研究例

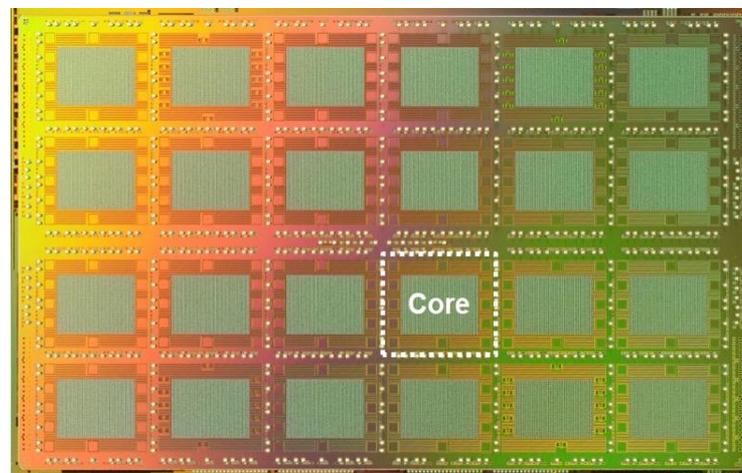
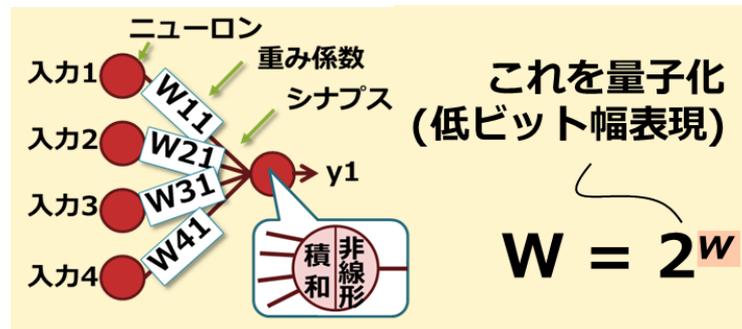
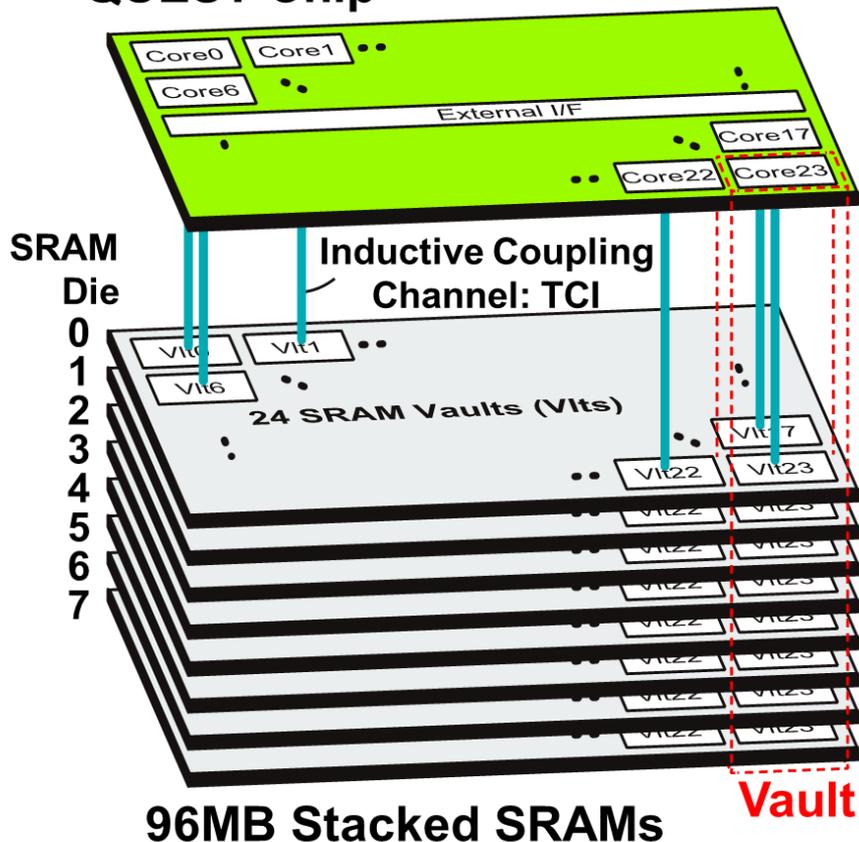
- 富士通 「学習」を担う半導体
- 北海道大 「推論」向けの技術
- 米グーグル 学習、推論を兼ね備える半導体



世界初: 二値化ニューラルネットチップ
 日本初: 深層ニューラルネットチップ

補足: 北大での研究実績 (2) QUEST

QUEST Chip



ISSCCで発表 (18年2月)

世界初の対数量子化・3次元積層DNNチップ。Silkroad Award受賞。日経新聞等掲載

- **対数量子化**近似DNN推論エンジン
- **枝刈りによる不規則 & スパース NNがターゲット**
- 7.49TOPS(バイナリ時)

最新の研究成果: アニーリングプロセッサ

STATICA: A 512-Spin 0.25M-Weight Full-Digital Annealing Processor with a Near-Memory All-Spin-Update-at-Once Architecture for Combinatorial Optimization with Complete Spin-Spin Interactions

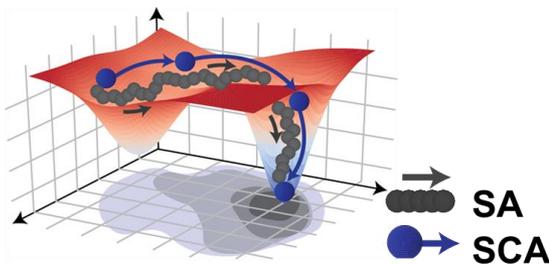
Kasho Yamamoto^{1,2}, Kota Ando¹, Normann Mertig³, Takashi Takemoto³, Masanao Yamaoka³, Hiroshi Teramoto², Akira Sakai², Shinya Takamaeda-Yamazaki⁴, Masato Motomura¹

¹Tokyo Institute of Technology, Yokohama, Japan, ²Hokkaido University, Sapporo, Japan ³Hitachi, Sapporo, Japan, ⁴University of Tokyo, Tokyo, Japan

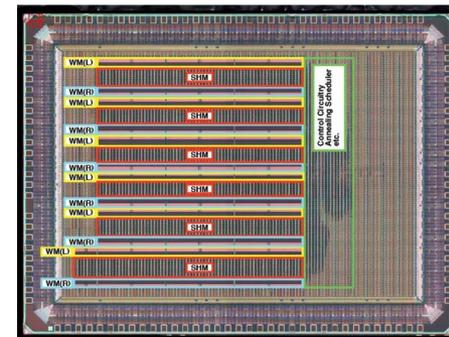


© 2020 IEEE International Solid State Circuits Conference 1 of 22

組合せ最適化問題を高速求解



STATICA CHIP



疑似量子計算チップ、東工大など開発 渋滞解消・創薬に応用

2020/2/17付 [有料会員限定]

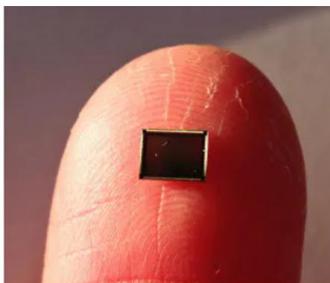
保存 共有 印刷 複製 共有 他

東京工業大学や北海道大学、日立製作所、東京大学などは共同で、量子コンピューターの計算を疑似的に再現して、組み合わせ問題を高速で解くことのできる半導体チップを開発した。計算を並列で処理できる理論を考案し、数ミリ角のチップを試作した。従来法よりも計算が約4倍速く、消費電力は約60分の1になった。量子コンピューターよりも先に、渋滞の解消や創薬、材料開発などで応用できるとみている。

成果は米サンフランシスコで開催される半導体の国際会議「ISSCC」で発表する。

既存のコンピューターを超える計算能力を持つ次世代計算機として量子コンピューターが注目される。現状では極低温まで冷やしたり複雑な配線が必要だったりするため、装置が大がかりで計算も安定しない。

既存のコンピューターを使い、量子コンピューターの計算方法をまねる技術が注目を集める。様々な組み合わせの中から最適解を探す「組み合わせ最適化問題」の計算を得意とする。従来コンピューターでは計算量が多すぎて効率よく計算するのは難しい。装置の小型化や安定した計算、より大規模な計算に対応できると期待されている。



画像の拡大

数ミリメートル角の半導体チップで量子コンピューターを模した計算が高速で処理できる=東工大提供

日経新聞

IEEE Spectrum

Novel Annealing Processor Is the Best Ever at Solving Combinatorial Optimization Problems

Tokyo Tech engineers say their CMOS processor bests current technologies in solving the traveling salesman conundrum and other complex puzzles

By John Boyd

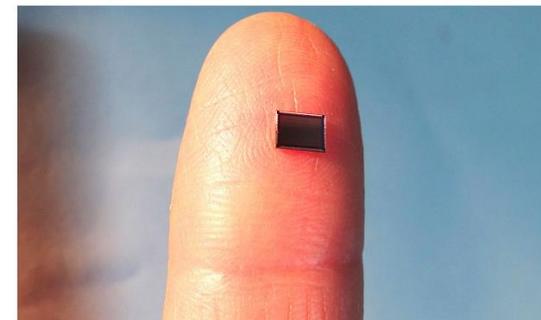


Photo: Tokyo Institute of Technology

During the past two years, IEEE Spectrum has spotlighted several new approaches to solving combinatorial optimization problems,

ArtIC人員・居室構成 (20年4月)

本村研 劉研

教員	教授	本村	
	准教授		劉
	助教	ティエム	
	特任助教	川村	(募集中)
スタッフ	技術支援員	(募集中)	
	秘書	2名	
学生	D3	1名	
	D2	1名	
	M1	3名	
	B4	4名	



ArtICの研究レイヤ

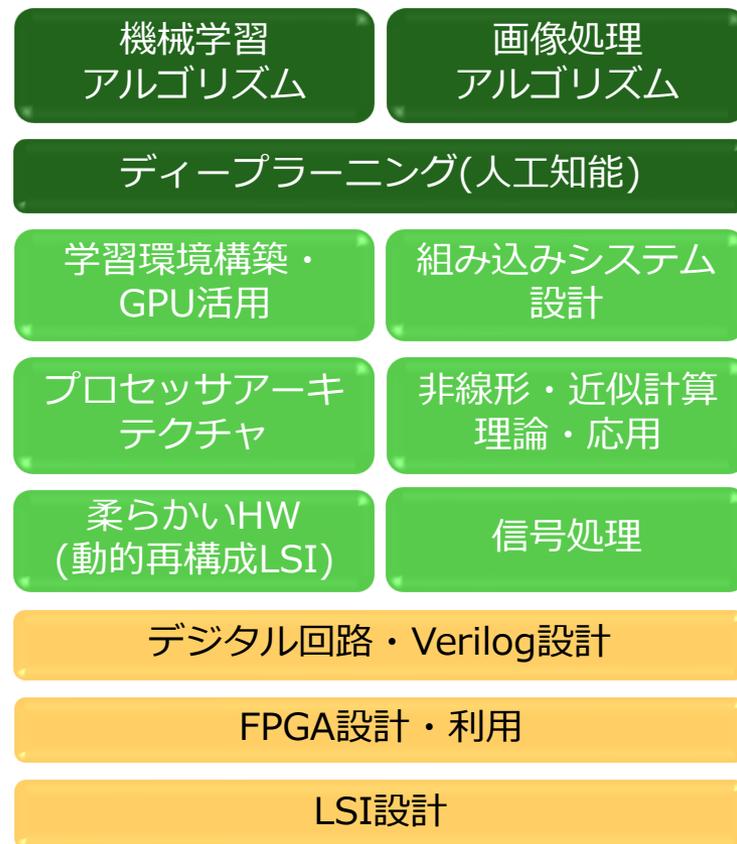
技術レイヤ

「論理」を相手にする世界



「物理」を相手にする世界

身に付くスキル・知識 (例)



最後に…

- ユーチューブチャンネル (研究室HPからリンクあり/Artic 東工大で検索可能)
 - <https://www.youtube.com/channel/UCJY897-DXhrnfMWC4gYIwFw>
- LSI設計やFPGAを用いたハードウェア設計に興味を持つ皆さんのArtICへの参加を歓迎します
 - 特別な知識は求めません。この分野の研究に対する意欲を期待します
 - 基礎知識が身に付くよう、輪講や研修を行います
- 二研究室で共同運営しており、研究ユニット内に垣根はありません
 - 居心地よい環境づくりを心掛けています
- 一線級の国際会議で発表できるグループを目指しています
- 近未来の社会ニーズに即した、実戦的な研究活動を主体としています
- 従って、実社会で役立つスキル・知識を身に着けることができます
- 産学連携、大学間連携、国家プロジェクト参画を活発に進めています