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

2021年 3月 25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・劉 研究室

Tokyo Tech

ARTIC

AIコンピューティング研究ユニット: ArtIC

2019年4月
に発足

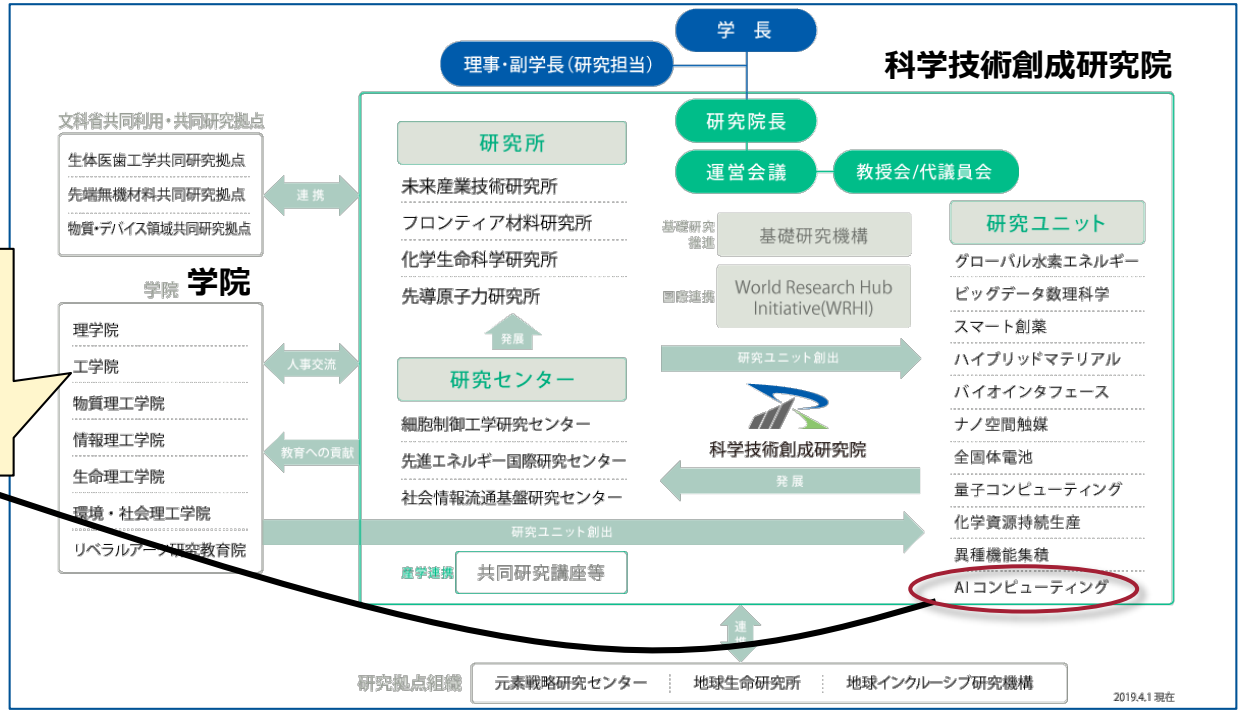
2020年4月からフル
メンバで活動中

すずかけ
台キャン
パス J3棟
17F

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

ArtICの成り立ち

工学院・情報通信系 と 科学技術創成研究院



情報通信系・コース
本村・劉 研究室

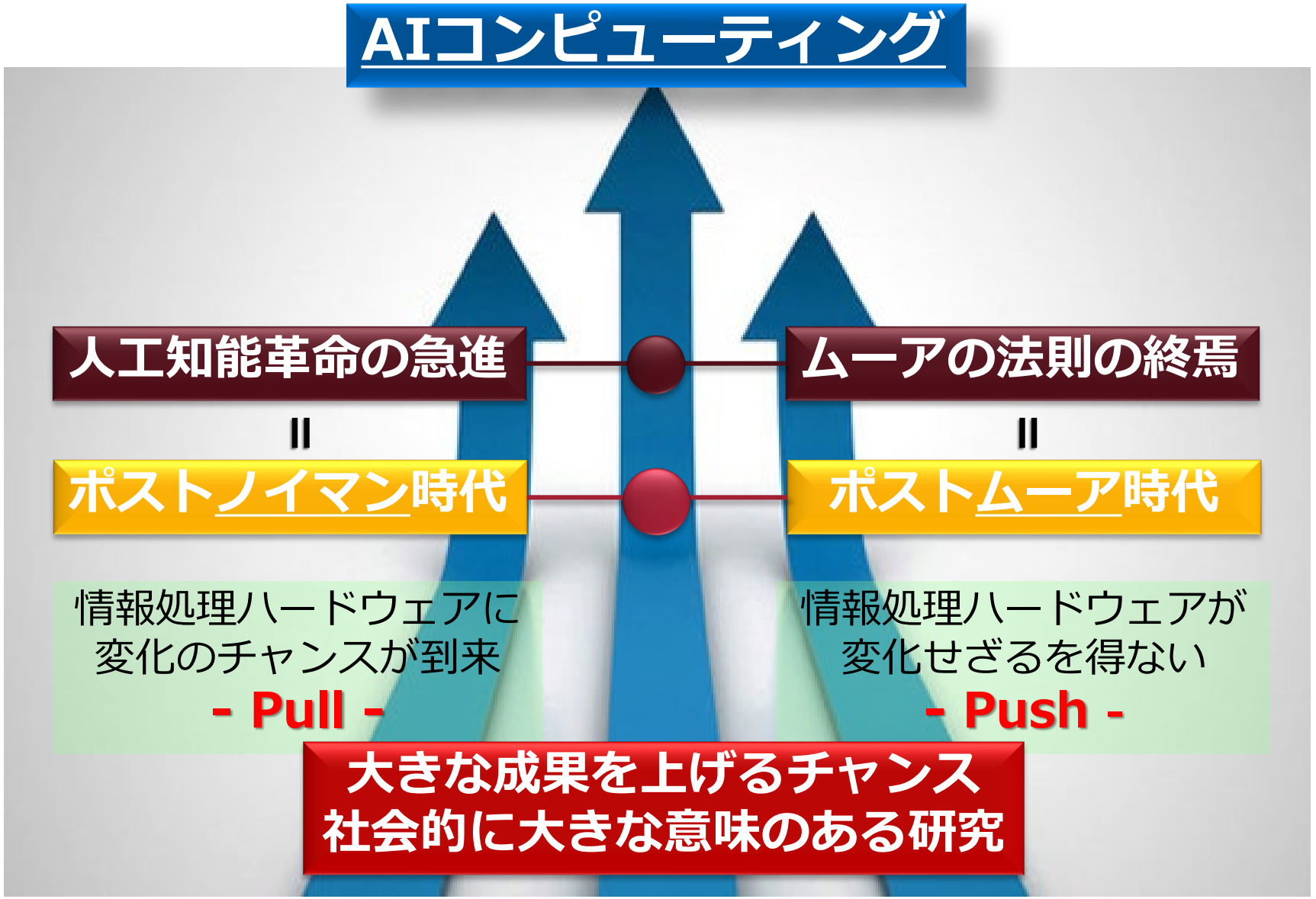
Artificially Intelligent Computing Research Unit

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

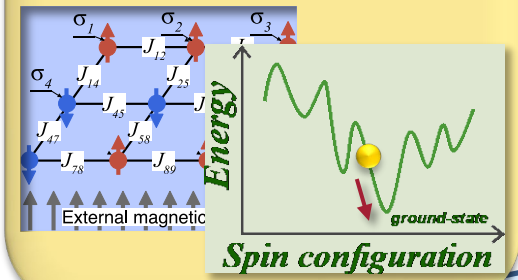
2020年代: コンピューティングの新時代



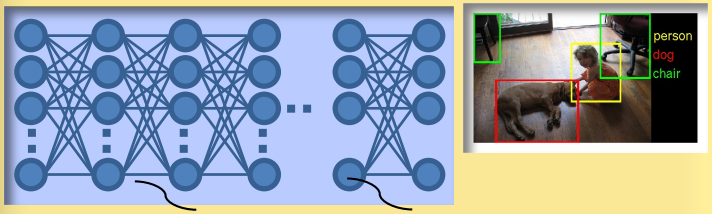
ArtIC: 研究ターゲット

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

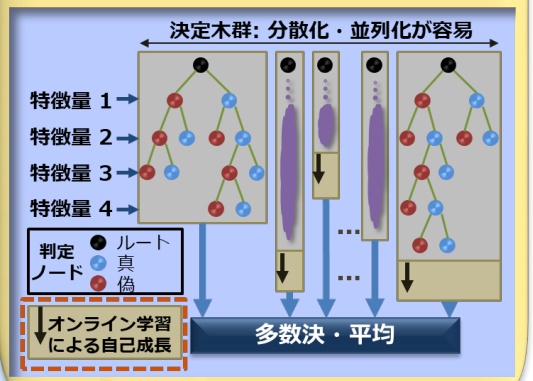


大量データの学習 → 強力な推論・識別・予測能力



抽象シナプス　抽象ニューロン

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）

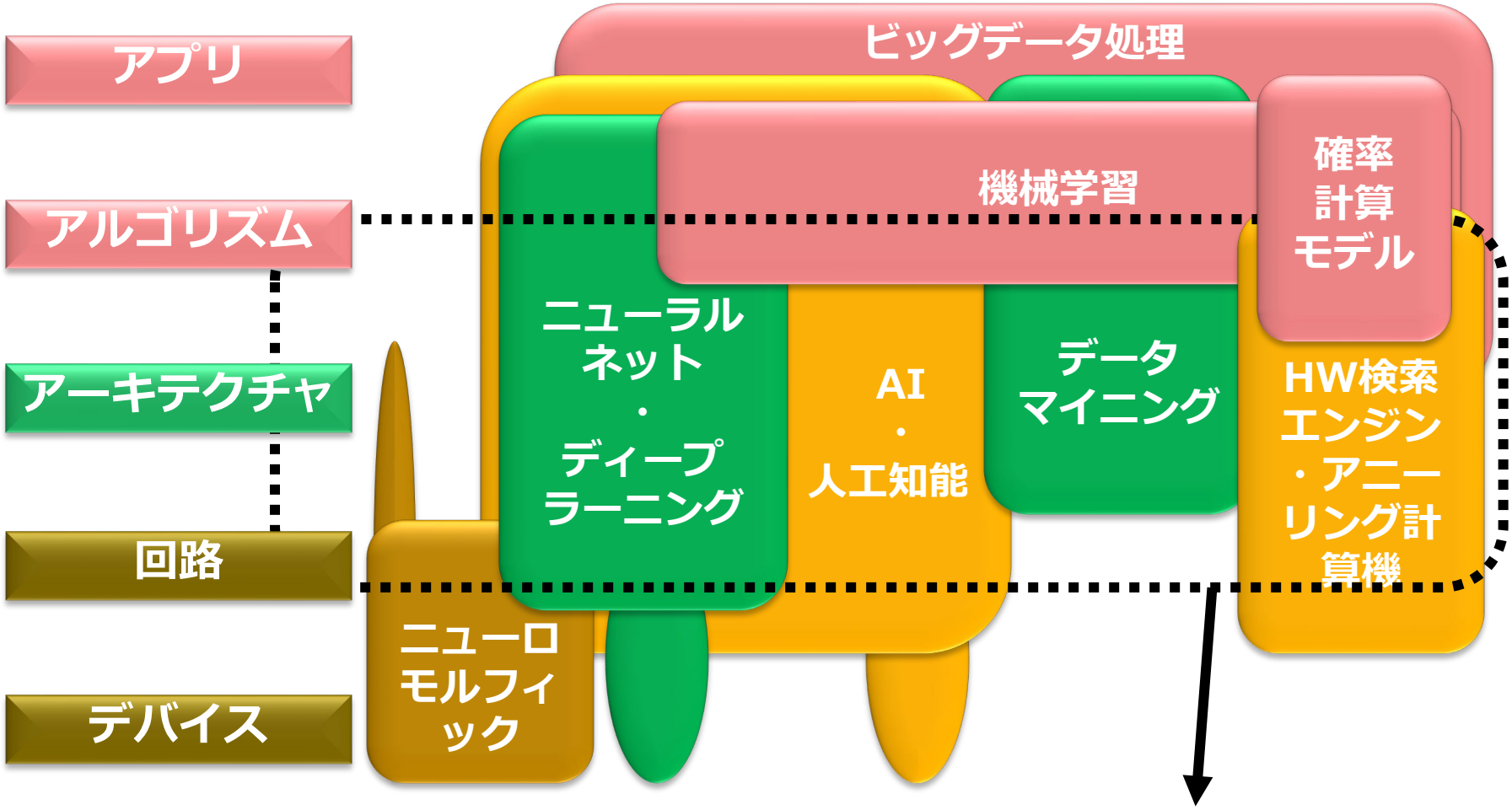
統計的機械学習（アンサンブル学習等）

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
アーキテクチャの研究を推進

研究ユニット教員の紹介

本村


日経新聞 17/9/18
日朝刊に掲載

AI支える テクノロジー ㊤専用半導体

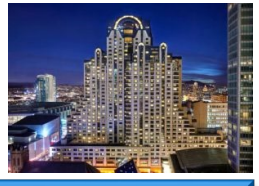
高速処理・省エネ 両立探る

AIの普及には新しい半導体が必要

現在	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普及期	AIに特化した半導体	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

’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のオリンピック 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

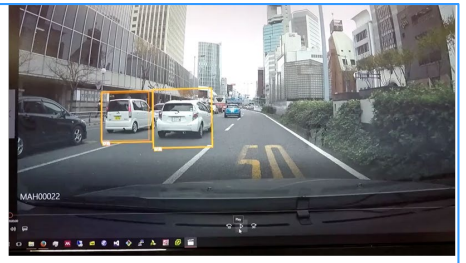
世界初: 二値化ニューラルネットチップ
日本初: 深層ニューラルネットチップ

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

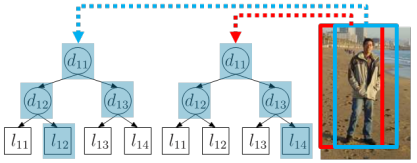
汎用物体認識システム(FPGA利用)



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

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

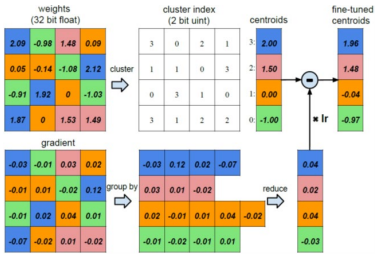
リソース使用率

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

識別速度
- 歩行者識別: Full HD 350 fps

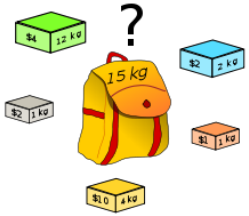
BDT: ブースティング決定木

近似計算によるディープラーニング



最新の研究成果: アニーリングプロセッサ

- 組合せ最適化問題は社会のあらゆる場面に存在
  - ✓ 物流、創薬、工場、広告、交通、機械学習、集積回路設計…
- 選択肢が指数関数的に増大(組合せ爆発)
- 精度の高い**ベター**な解を**高速**に得る計算技術のニーズ



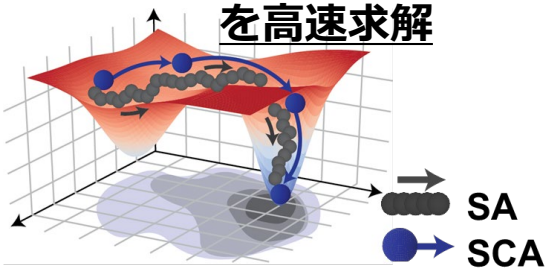
ハミルトニアン

$$H(\boldsymbol{\sigma}) = -\sum_{(i,j)\in E} J_{ij}\,\sigma_i\sigma_j - \sum_{i\in V} h_i\sigma_i$$

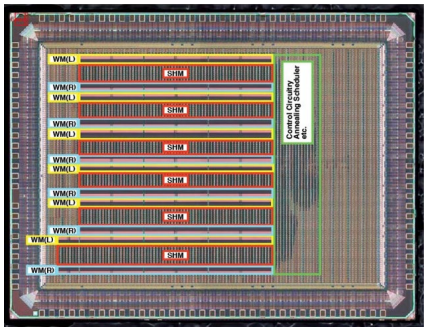
二値変数の二次関数

LSIのオリンピックISSCC2020で発表

組合せ最適化問題を高速求解



STATICA CHIP



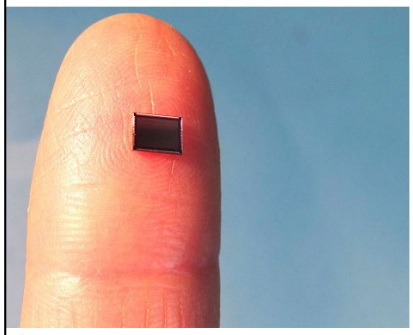
IEEE Spectrum

14 Apr 2020 | 17:00 GMT

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

Tokyo Tech engineers say their CMOS processor bests ologies in solving the traveling salesman drum and other complex puzzles

By John Boyd



nology

o years, *IEEE Spectrum* has spotlighted several new ng combinatorial optimization problems,

日経新聞

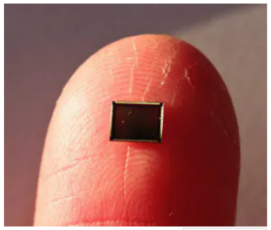
疑似量子計算チップ、東工大など開発 渋滞解消・創薬に応用

2020/2/17付 [有料会員限定]

保存 共有 印刷 その他

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。



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

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。従来の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。

最新の研究成果: 機械学習

- 深層ニューラルネットワーク・決定木アンサンブルは様々な応用分野で実用
  - 画像認識，音声認識，自然言語処理，意思決定，…
- 学習・推論にかかる計算量・消費電力の問題
- 高い精度の予測を**低計算量**，**低電力**，かつ，**高速に**行う技術への要求

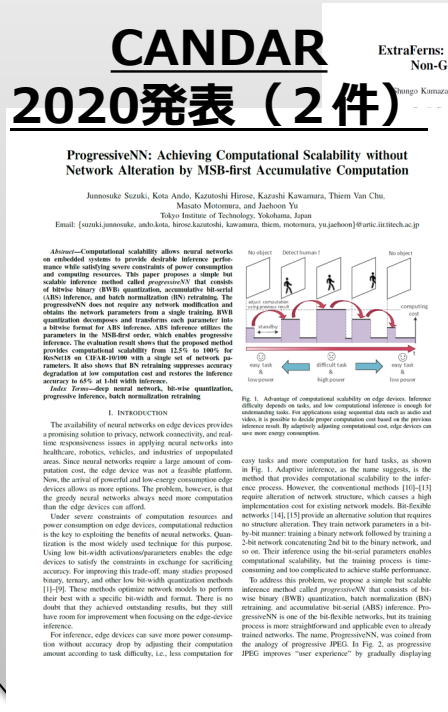
DNNアクセラレータの研究

VLSI2021
投稿中

ISSCC

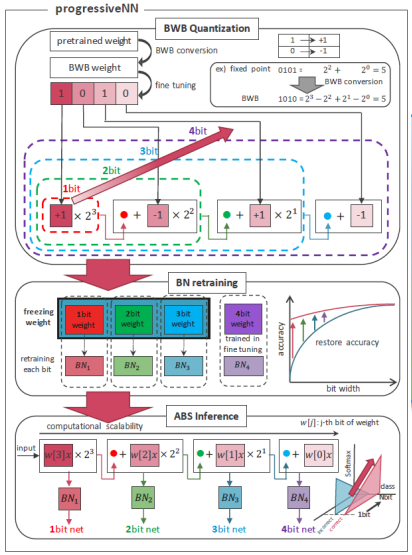
機械学習アルゴリズム

CANDAR
2020発表（2件）


ExtraFerns: Fully Parallel Ensemble Learning Technique with Non-Greedy yet Minimal Memory Access Training

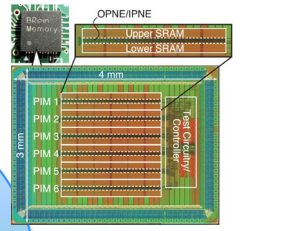
ProgressiveNN: Achieving Computational Scalability without Network Alteration by MSB-first Accumulative Computation

IJNC
投稿中（2件）

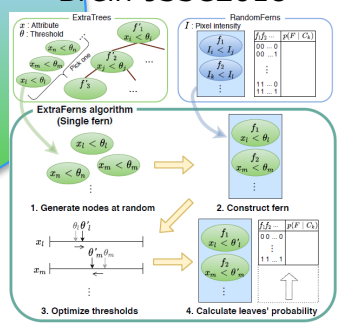


スケーラブルネットワーク
Progressive-CANDAR2020

HW
協調
Algo



Brein-JSSC2018

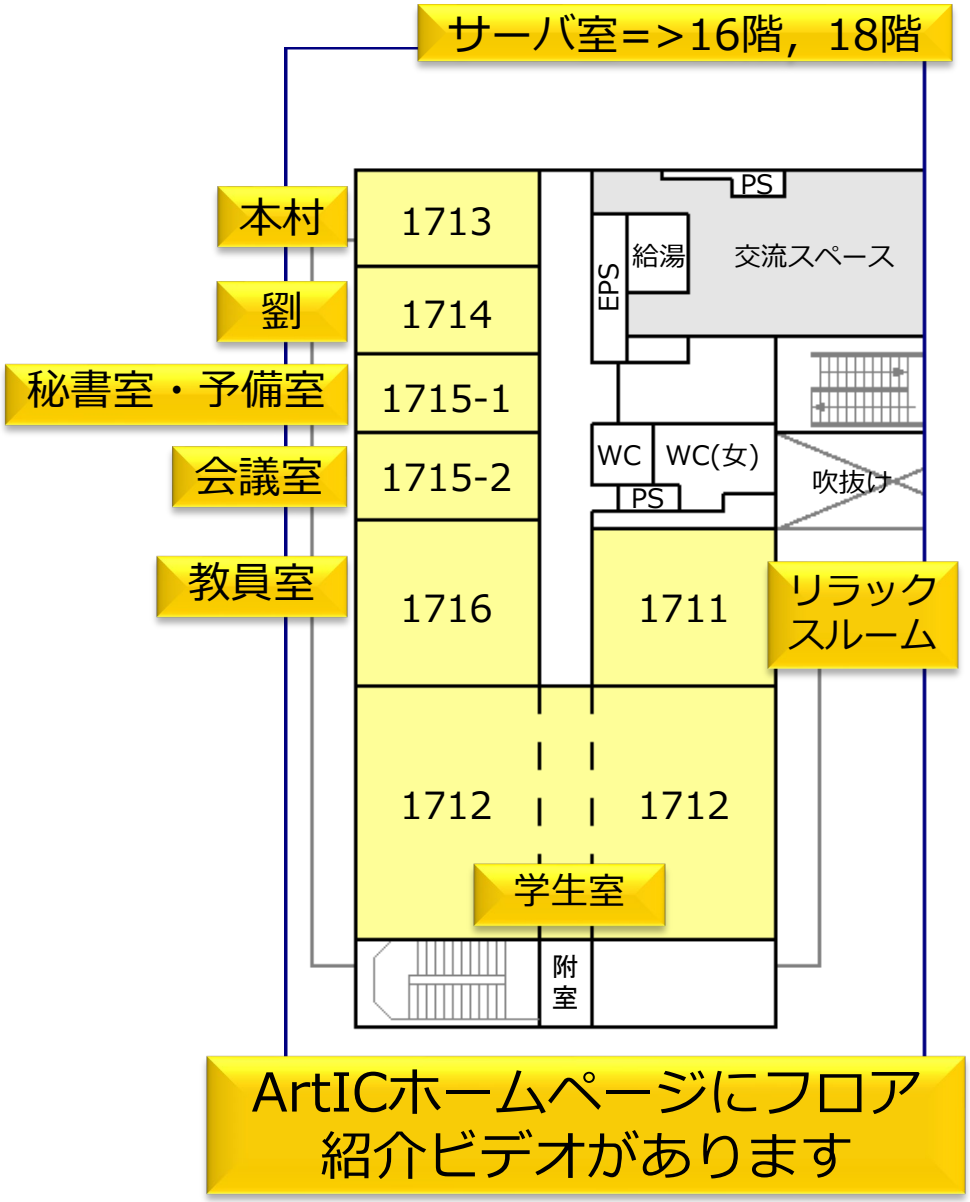


決定木アンサンブル
ExtraFerns-CANDAR2020

ArtIC人員・居室構成 (21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，安藤	
スタッフ	技術支援員	(募集中)	
	秘書	2名	
学生	D3	1名	
	D1	1名	
	M2	3名	
	M1	5名	
	B4	3名	

ユニット
一体運営



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）

- 機械学習アルゴリズム
- 画像処理アルゴリズム
- ディープラーニング(人工知能)
- 学習環境構築・GPU活用
- 組み込みシステム設計
- プロセッサアーキテクチャ
- 非線形・近似計算理論・応用
- 柔らかいHW(動的再構成LSI)
- 信号処理
- デジタル回路・Verilog設計
- FPGA設計・利用
- LSI設計

最後に…

- 本資料は，研究室HPの入学志望者向け情報のページに載せる予定です
- ユーチューブチャネル (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