

山本研究室

(知能計算分野)

教授：山本 章博

助教：明石 望洋

研究室：総合研究7号館 324(教授室), 323, 325, 327

Web：<http://www.iip.ist.i.kyoto-u.ac.jp/>

Email：akihiro@i.kyoto-u.ac.jp

当研究室では、**機械学習理論**を中心にして人間の**高次推論機構**の性質を解明し、またそれらを用いて、与えられたデータから適切な情報を取出すための**計算機構**や**ソフトウェア**を構築することを目標に研究を行っている。この研究はデータ集合からの**知識発見**などへの応用、**数理論理学**や**計算数学**との関係の解明へと展開している。

第2次AIブームの成果を第3次AIブームに生かす

知一5とデー5

- 山本は令和8年3月で定年退職を迎えます
- 大学院情報学研究科
 - 知能情報学コース 知一5 については募集要項をよく読んでください
 - データ科学コース デー5 については, 昨年度よりデータ科学イノベーション教育研究センターの瀧川特定教授, フラナガン特定准教授に指導していただくことを前提に募集しています.
- 実は, 以前から4回生と修士両方が山本研であった学生は多くないです.

こだわり所

第2次AIブーム: 数理論理学を基盤

知識表現と証明(説明)

第3次AIブーム: ニューラルネットを基盤

大規模データからの機械学習

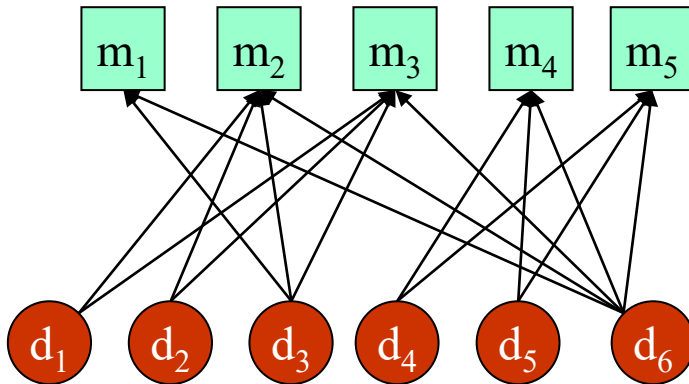
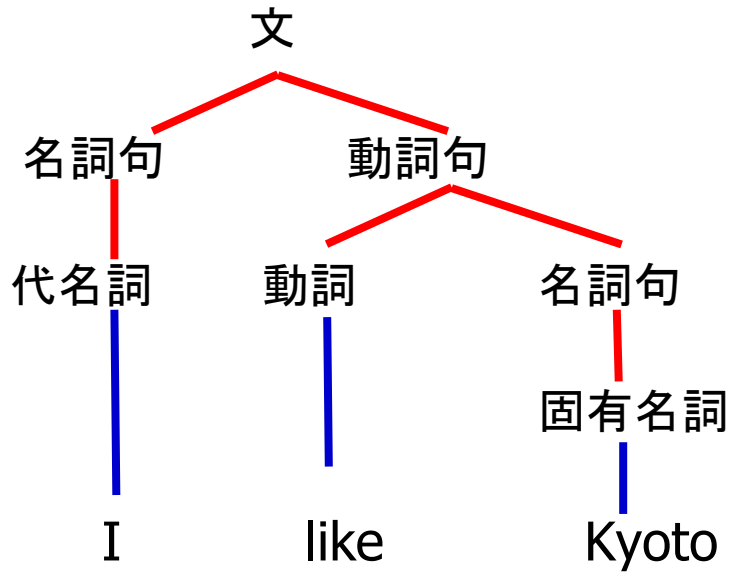
■ データ構造に着目した機械学習

- 木構造: ソースコードの構文解析木, HTMLのDOM-Tree
- 2部グラフ: トランザクション(商品と顧客)
- グラフ: 知識グラフ(Knowledge Graph)

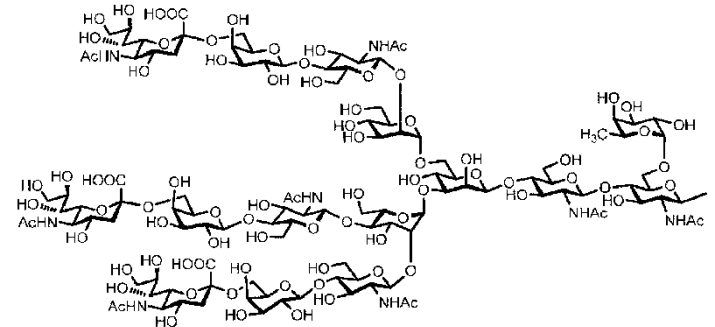
■ 説明可能な機械学習

- 数学・数理論理は「推論の表現と正しさ」を数学的に保証
- 機械学習でも「推論の表現と正しさ」を数学的に保証することが重要
機械学習アルゴリズムの正しさとは何か？
 - ・深層学習が有効なのは, 学習の結果に説明を求められない場合

構造データの出どころ




```
無題 - メモ帳
ファイル(F) 編集(E) 書式(O) 表示(V) ヘルプ(H)
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "http://www.
<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="ja"
lang="ja">
<head>
<meta http-equiv="Content-Type"
content="text/html; charset=utf-8" />
<meta http-equiv="Content-Style-Type" content="text/css" />
<meta http-equiv="Content-Script-Type" content="text/javascript" />
<title>
ホーム
&dash;
京都大学
</title>
<base href="http://www.kyoto-u.ac.jp/ja/" />
<meta name="generator" content="Plone - http://plone.org" />
<meta content="京都大学, 京大, 京都大, Kyoto University, 大学, 日本の大学, 京
name="keywords" />
<meta content="京都大学のホームページです。京都大学の組織や機構の紹介を
name="description" />
<!-- Plone ECMAScripts -->
<script type="text/javascript"
src="http://www.kyoto-u.ac.jp/portal_javascripts/Pl
</script>
<script type="text/javascript"
src="http://www.kyoto-u.ac.jp/portal_javascripts/Pl
</script>
<script type="text/javascript"
src="http://www.kyoto-u.ac.jp/portal_javascripts/Pl
</script>
<script type="text/javascript"
src="http://www.kyoto-u.ac.jp/portal_javascripts/Pl
```



Tree-LSTMを利用した情報統合[修士]

- 自然言語データに対する深層学習モデルであるLSTMをベースに木構造向けモデルを設計
- 同一目的だがサイトによって表現形式が異なる情報の統合に応用
 - 大学のシラバスなど
京都大学のシラバスサイト



The image shows a screenshot of the Kyoto University OpenCourseWare website. The main content area displays course information for 'Introduction to Data Science' (データ科学のための数学入門). The course is listed as 'DAAB11 18006 L006' with 2 units. The instructor is listed as '山田 賢博 (山田 賢博)'. The course description is in Japanese, mentioning that it is a self-paced course for students interested in data science. Below the description, there is a table with columns for '科目名' (Course Name), '単位数' (Units), '担当教員' (Instructor), and '曜時限' (Days/Time). The table lists six courses: '数学入門' (2 units, 山田, 火1), '英語' (4 units, 鈴木, 水3), '統計 I' (2 units, 佐藤, 木4), '方言学' (2 units, 田中, 月2), 'グラフ理論' (2 units, 平山, 水6), and 'データ分析' (1 unit, 井上, 金1).

科目名	単位数	担当教員	曜時限
数学入門	2	山田	火1
英語	4	鈴木	水3
統計 I	2	佐藤	木4
方言学	2	田中	月2
グラフ理論	2	平山	水6
データ分析	1	井上	金1

東京大学のシラバスサイト

DFAとSATを用いた強化学習[修士]

- 報酬が決定性有限状態オートマトン(DFA)で表される強化学習
- エージェントはDFAを学習することにより方策を学習する
- 未知データに対してSATを用いて補完する



	ϵ	b
ϵ	0	0
a	0	1
ab	1	1
b	1	x_1
aa	0	x_2
aba	0	x_3
abb	0	x_4

SAT 補完

	ϵ	b
ϵ	0	0
a	0	1
ab	1	1
b	1	1
aa	0	1
aba	0	0
abb	0	1

形式概念解析における概念束の更新

目的 形式概念分析(FCA)において、**オブジェクト g に関する新たな関係**
(g, m)が追加された時の概念束の更新手法の効率化・簡略化

➤ 下図の例：薬2が病bにも有効であると分かった時の概念束更新

結論 文脈の変化が一部の時、既存手法より高速なアルゴリズムを開発

文脈		属性		
		病a	病b	病c
オブジェクト	薬1	×	×	
	薬2	×		
	薬3			×

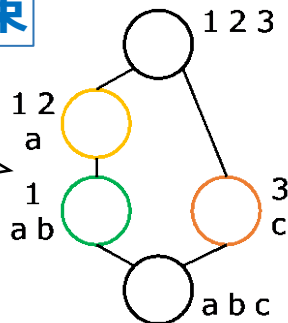
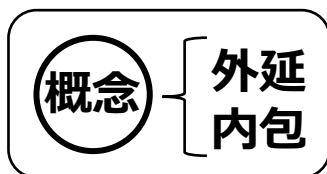
関係追加



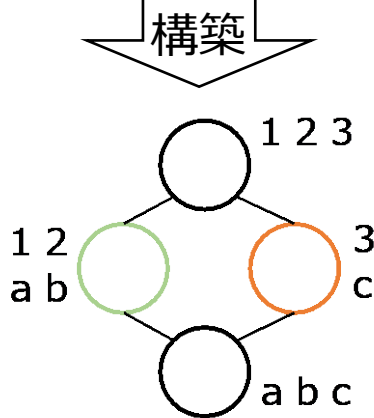
文脈		属性		
		病a	病b	病c
オブジェクト	薬1	×	×	
	薬2	×	×	
	薬3			×

構築

概念束



更新



多肢選択問題での解答選択

- 多肢選択問題に対して、説明可能な過程で解答を行う。MCQA (multiple choice question answering)
 1. 空欄に単語を当てはめ、4つの文を用意
 2. 各文を命題と捉え、真らしさを指標化
 3. 最も真らしい文を与える選択肢を正解として選択
- **知識グラフ**を用いる

知識グラフ (KG)

- 検証用の関係グラフ
- Wikipediaを変換して作成



REBEL

KG



命題グラフ (PG)

- 検証すべき命題を表すグラフ
- 問題文 $mcq(c_i)$ を変換して作成

Shohei Ohtani is a Japanese baseball player, playing for the Los Angeles Dodgers in MLB.

REBEL

PG



Wikipedia (https://en.wikipedia.org/wiki/Shohei_Ohtani) より作成

問題文 $mcq(x)$

___ is a Japanese baseball player, playing for the Los Angeles Dodgers in MLB.

選択肢 $C = \{c_0, c_1, c_2, c_3\}$

1. Shohei Ohtani
2. Munetaka Murakami
3. Ichiro Suzuki
4. Mike Trout

機械学習(AI)から見た多肢選択問題

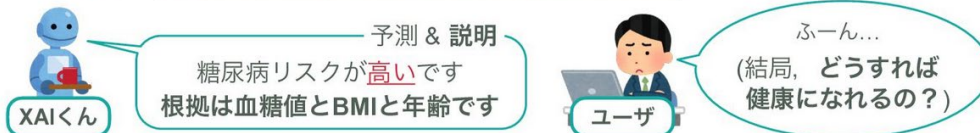
選択肢には説明可能性を持つ必要がある

- 説明可能AI(XAI) : AIが導き出した答えに対して, 人間が納得できる根拠・理解しやすい根拠を示す技術
- 反実仮想説明 : 予測結果が覆る(反対になる)特徴量の摂動例を提示する説明
 - 反実仮想(Counterfactual) : 事実と反対のことを想定すること

反実仮想説明法 (Counterfactual Explanation, CE)

所望の予測結果を得るための“アクション”を説明として提示

- 従来の局所説明法 (例: **LIME** [Ribeiro+ 16])
 - モデルの予測結果の根拠となった特徴量を提示



- **反実仮想説明法 (CE)** [Wachter+ 18]
 - モデルから**所望の予測結果を得るための特徴量の変更方法を提示**



研究目標

正答⇒予測結果

誤答⇒反実仮想, 敵対的摂動

理解しやすい根拠

⇒言葉 or 視覚的表現

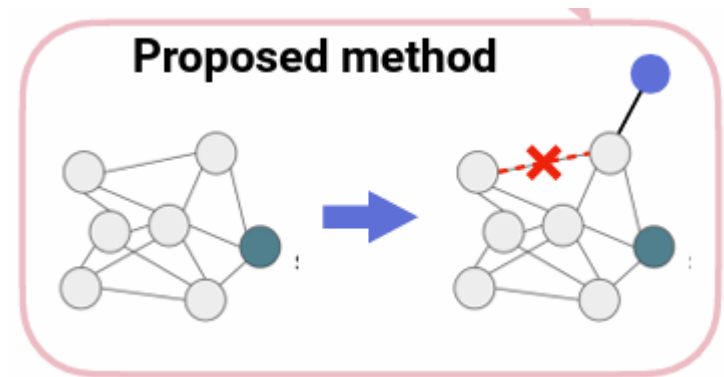
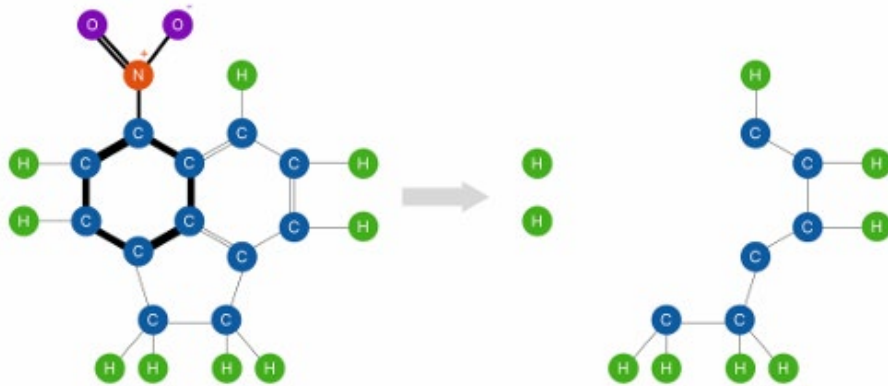
- 当初はAIは用いない
- 単純な探索問題とする

[金森氏(北大, 現富士通)の
スライドより]

GNNに対する反実仮想説明[修士]

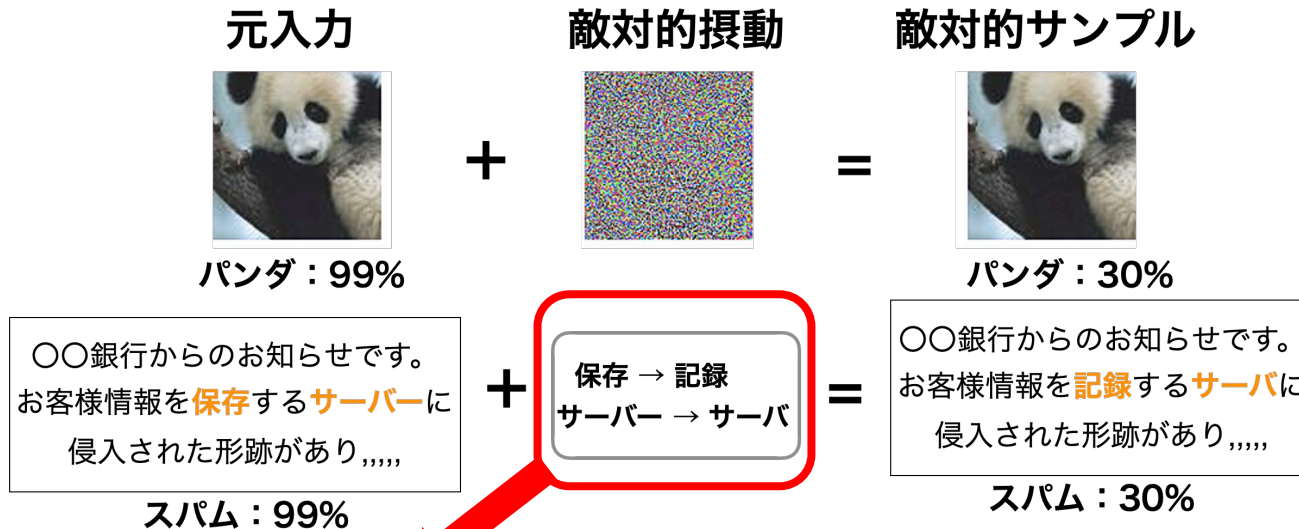
- **グラフ**・ニューラルネットワークに対する**反実仮想説明**を構成する
- 入力の部分グラフだけでなく、**非部分グラフ**も反実説明に用いることができる手法を考案

従来手法



数理最適化を用いた言語モデルの頑健性検証 [修士]

- 最小単語置換の敵対的サンプルを求める問題を整数線形計画問題として定式化



最小単語置換問題

$$R = \min_{e \in \mathcal{S}} |e|$$

subject to $f(x + e) \neq y$

x : 入力テキスト e : 単語置換パターン
 y : 正解ラベル f : NNモデル
 $|e|$: 単語置換パターン e の単語置換数

どの程度誤差を加えると間違えてしまうのかを評価

整数線形計画問題(提案法)の解は必ず最適解(最小の単語置換)

LLMを用いたAPRの信頼性検証[修士]

- 大規模言語モデル(LLM)によるプログラム自動修正(APR)の信頼性を検証する方法を提案
 - プログラムに摂動を与える
 - 記述の詳細さの影響を測る
 - 変数名の可読性の影響を測る

Write a function called "find_sum_of_multiples" that takes one argument, a positive integer, and returns the sum of the first **51** positive multiples of the given integer.

○ (a) $p = 51$. Correct answer.

```
def find_sum_of_multiples(n):  
    if n <= 0:  
        return 0  
    sum = 0  
    for i in range(1, 52):  
        sum += i * n  
    return sum
```

Write a function called "find_sum_of_multiples" that takes one argument, a positive integer, and returns the sum of the first **56** positive multiples of the given integer.

✗ (b) $p = 56$. Wrong answer.

```
def find_sum_of_multiples(n):  
    if n <= 0:  
        raise ValueError(*n must be a positive integer*)  
    sum = 0  
    for i in range(1, 57):  
        if i % n == 0:  
            sum += i  
    return sum
```

[Hanarvar+, 23]

正しいプログラム

```
def search(x, seq):  
    for i, e in enumerate(seq):  
        if x >= e:  
            return i  
    return len(seq)
```

演算子
算術, 比較, ブール演算子の
1か所のみの置換を総当たり

```
def search(x, seq):  
    for i, e in enumerate(seq):  
        if x == e:  
            return i  
    return len(seq)
```

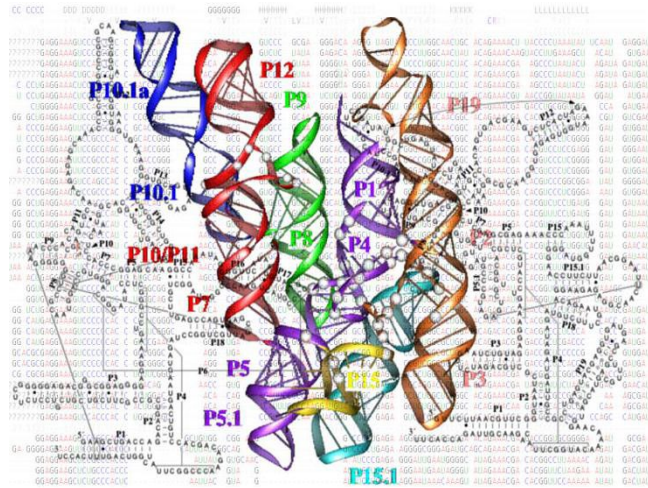
変数
各変数を **1か所のみ** ランダム
な他の変数に置換

```
def search(x, seq):  
    for i, e in enumerate(seq):  
        if x >= i:  
            return i  
    return len(seq)
```

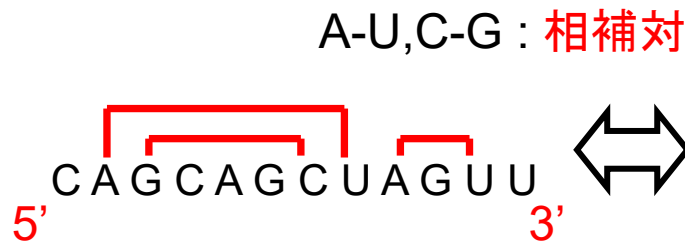
行
全ての行で, **1行** 削除した
コードの穴埋め問題

```
def search(x, seq):  
    for i, e in enumerate(seq):  
        >>> [ INFILL ] <<<  
    return i  
    return len(seq)
```

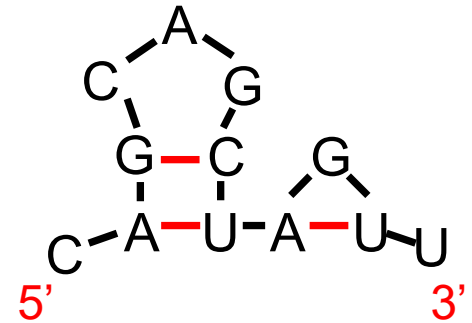
Transformerを用いたRNAの2次構造予測



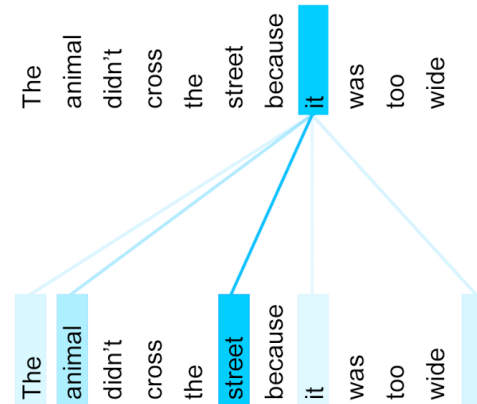
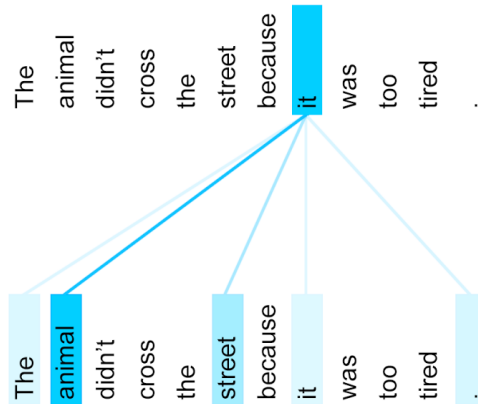
[Wikipedia]



RNA配列



RNA配列の二次構造



Uszkoreit, J.: Transformer: A Novel Neural Network Architecture for Language Understanding
<https://ai.googleblog.com/2017/08/transformer-novel-neural-network.html>

本年度の研究に向けて

■ Transformerの形式言語理論・数理論理学からの解析

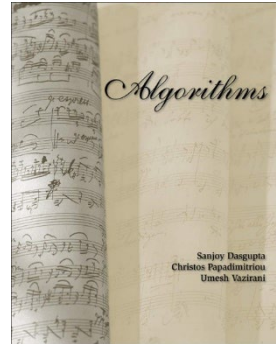
先行研究例

Yao, S., Peng, B., Papadimitriou, C., and Narasimhan, K.: Self-Attention Networks Can Process Bounded Hierarchical Languages (2021)

定理 任意の k と D に対して、 k 種類の括弧と構文木の高さが高々 D であるような形式言語 $Dyck_{k,D}$ を受理するような $D+1$ レイヤの attentionネットワークを構築することができる.

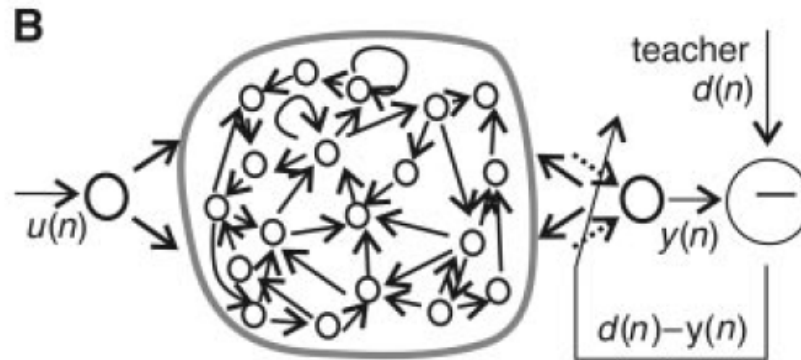
J. Pérez, P. Barceló, and J. Marinkovic :Attention is Turing Complete (2021)

定理 任意のTuring機械 M に対して、 M をシミュレートするような Transformer $T(M)$ を構成することができる.

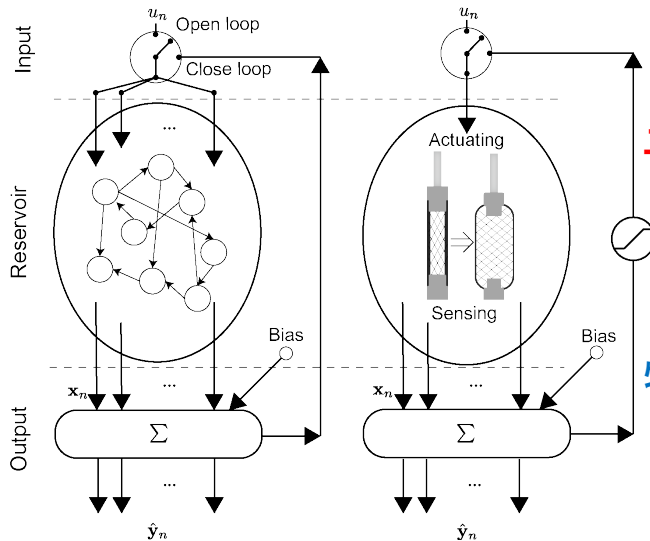


機械学習とダイナミクスの横断研究

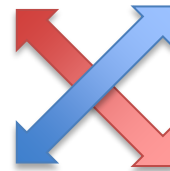
- 機械学習と動力学解析を組み合わせた知能計算の解明と開発



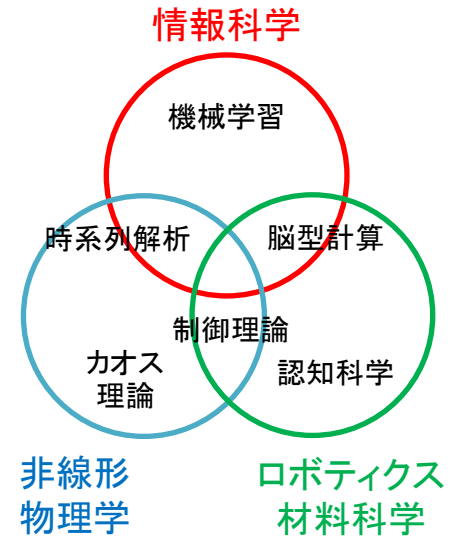
リザーバ計算[1]: 再帰型ニューラルネットの学習手法



ニューラルネットを動学的に解析



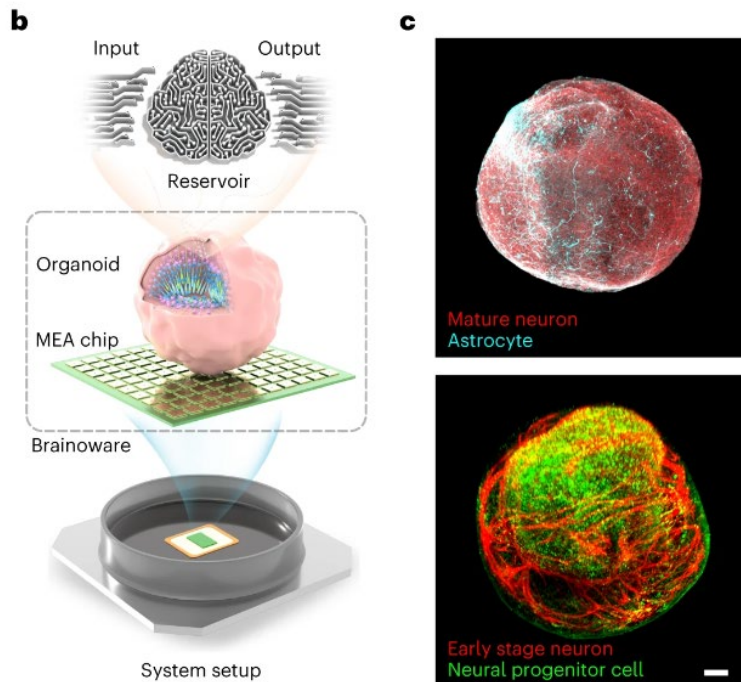
物理ダイナミクスを情報処理へ活用



物理リザーバ計算:
物理ダイナミクスをニューラルネットとして活用

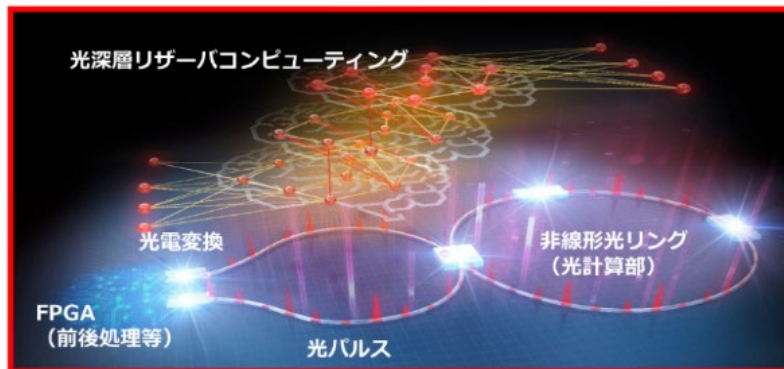
機械学習とダイナミクスの横断研究

Extremely Active Research Topics!



脳オルガノイドによる音声認識

H. Cai, Z. Ao, C. Tian, Z. Wu, H. Liu, J. Tchieu, M. Gu, K. Mackie, and F. Guo, *Brain Organoid Reservoir Computing for Artificial Intelligence*, Nat Electron 6:, 1032 2023.

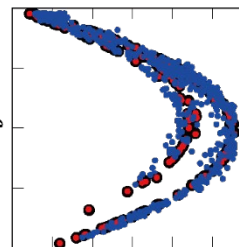
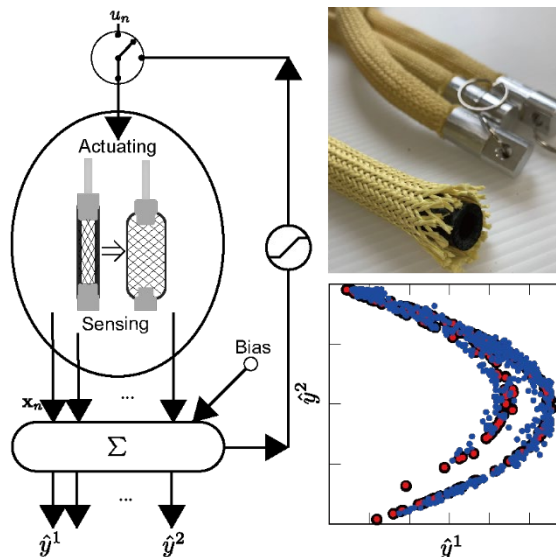


本成果②:
提案法による光NNの実装

- ・提案法を光電実装
- ・NTTの光深層リザーバコンピュータに適用
- ・推論と学習の両方を光計算でアシスト

次世代の光コンピュータの計算基盤

M. Nakajima, K. Inoue, K. Tanaka, Y. Kuniyoshi, T. Hashimoto, K. Nakajima, Physical deep learning with biologically inspired training method: gradient-free approach for physical hardware, *Nature Communications* 13: 7847, 2022.



ロボットの身体による計算

[N. Akashi](#), Y. Kuniyoshi, T. Jo, M. Nishida, R. Sakurai, Y. Wakao, K. Nakajima, Embedding bifurcations into pneumatic artificial muscle, *Advanced Science*, (Accepted) (preparing press)

機械学習とダイナミクスの横断研究

■ 機械学習と動力学解析を組み合わせた知能計算の解明と開発

これまで/現在の研究例

■ スピントロニクスによる脳型計算

- [N. Akashi](#), Y. Kuniyoshi, S. Tsunegi, T. Taniguchi, M. Nishida, R. Sakurai, Y. Wakao, K. Kawashima, K. Nakajima, A Coupled Spintronics Neuromorphic Approach for High-Performance Reservoir Computing, *Advanced Intelligent Systems* 4: 2200123, 2022.
- [N. Akashi](#), T. Yamaguchi, S. Tsunegi, T. Taniguchi, M. Nishida, R. Sakurai, Y. Wakao, K. Nakajima, Input-driven bifurcations and information processing capacity in spintronics reservoirs, *Physical Review Research* 2: 043303, 2020.

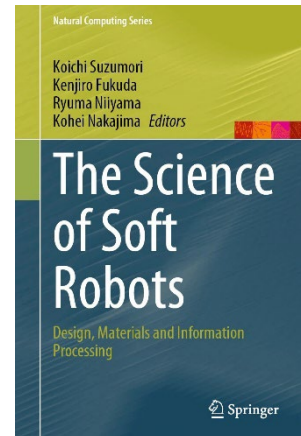
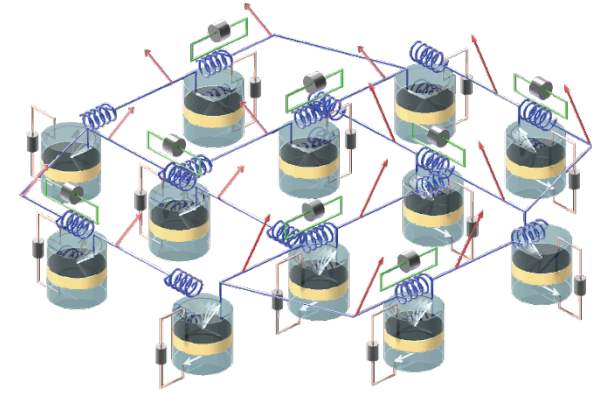
■ ロボットの身体による物理リザーバー計算

- [N. Akashi](#), Y. Kuniyoshi, T. Jo, M. Nishida, R. Sakurai, Y. Wakao, K. Nakajima, Embedding bifurcations into pneumatic artificial muscle, *Advanced Science*, (Accepted)

■ 物理知識に基づくニューラルネットワーク

- W. Sun, [N. Akashi](#), Y. Kuniyoshi, K. Nakajima, Physics-informed recurrent neural networks for soft pneumatic actuators, *IEEE Robotics and Automation Letters*, 7 (3) pp. 6862-6869, 2022.

■ Transformerの記憶の解析.



R5年度卒論題目

- 形式概念分析におけるオブジェクトの修正に伴う概念束の更新手法
- 大規模言語モデルを活用した推理小説読み解きのための知識獲得と含意論理式の生成
- 知識グラフを利用した多肢選択問題での解答選択手法

R4年度卒論題目

- ディープラーニングと自由エネルギーを用いたシュードノットを含むRNAの二次構造予測
- 埋め込み理論を用いた新型コロナウイルス新規陽性者数のモデル化

R3年度卒論題目

- モルフォロジー変換による前処理を用いた染めの型紙データへの自動タグ付け
- 推理小説と辞書における共起を用いた含意論理式の生成手法
- 最短非分離パス問題のパラメータ化計算複雑度に関する研究

R2年度卒論題目

- WordNetを用いたデータ拡張による知識グラフに対する質問応答精度の向上
- データマイニングツール群を利用した相関規則分析手法の評価
- 多段階頂点被覆集合問題とその一般化に対する固定パラメータ容易性の研究

R1年度卒論題目

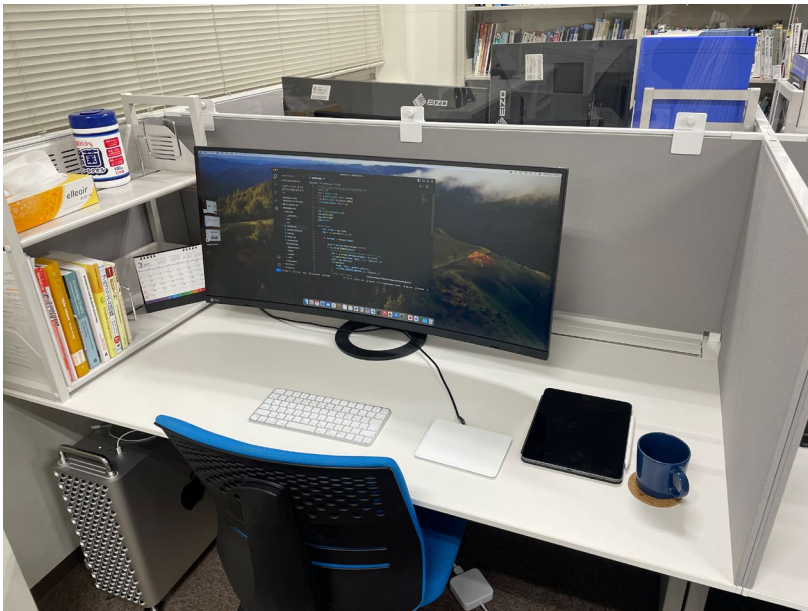
- 抽象構文木の関係に基づく条件付き確率場を用いた変数名予測
- 最大平衡連結部分グラフ問題に対するアルゴリズムと計算複雑性
- 出現頻度に対する定義を緩和した飽和アイテム集合とその列挙アルゴリズム

歓迎

- システム構築, 基礎理論いずれに興味があっても歓迎
 - 議論を通じて自ら追究する
 - 簡単な事柄でも, 動的計画法などの数学的手法に抵抗がない
 - 理論だけでなく実装による成果の確認を行える
 - 実装の数学的に意味を自分の言葉で説明する

研究室の環境

- 学生全員がMacbook/Mac Pro使用できます！
 - Windows/Linuxも使えます！
 - 欲しい本やデバイスなども！
- (他の研究室と比べていませんが)机広いです！



質問や相談があればメールで

教授: 山本章博 akihiro@i.kyoto-u.ac.jp

※ メールで質問の際は件名を「研究室配属2024」とし, 氏名を忘れずに!