

全脳計算モデルの実現に向けた認知モデルの役割

# 全脳確率的生成モデル による統合認知モデルと ロボティクス応用

立命館大学  
情報理工学部  
講師 谷口 彰

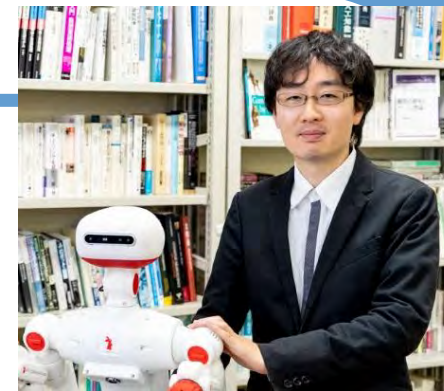


# 自己紹介

氏名： 谷口 彰

職位： 講師

所属： 立命館大学 情報理工学部 創発システム研究室

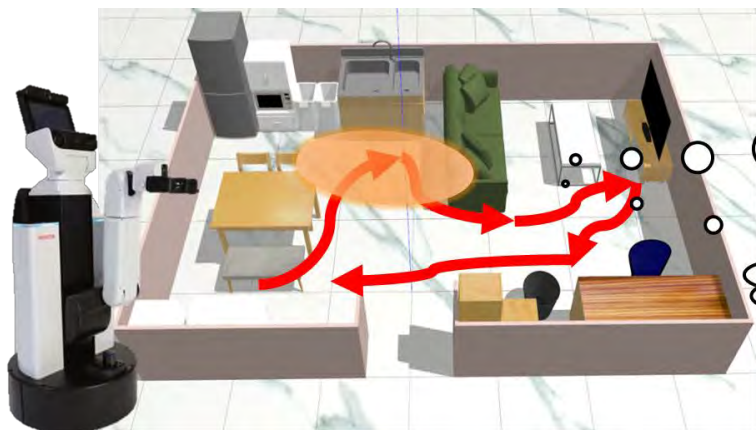


## 略歴

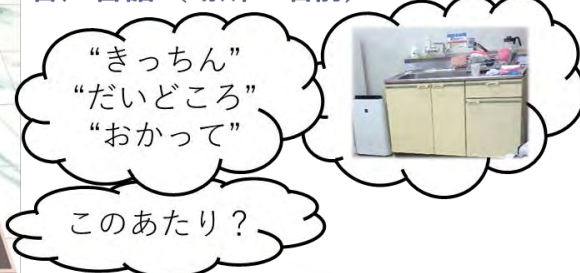
- ◆ 2018年 3月 立命館大学大学院 情報理工学研究科 博士後期課程 修了 博士（工学）
- ◆ 2017年 4月～2019年 3月 日本学術振興会 特別研究員 (DC2, PD)
- ◆ 2019年 4月～2022年 3月 立命館大学 情報理工学部 特任助教
- ◆ 2022年 4月～ 立命館大学 情報理工学部 講師

主な研究テーマ：「確率的生成モデルに基づくロボットによる場所概念・語彙の獲得と活用」

- 位置情報・言語情報・画像情報などのロボットの**マルチモーダル情報**からの**場所概念（場所のカテゴリ）の形成**や**空間的語彙の獲得**と、それらの**実世界タスクへの活用**を目指す研究。
- 記号創発ロボティクス, 知能ロボティクス, 人工知能, 機械学習, ロボット学習, 概念形成, 言語獲得



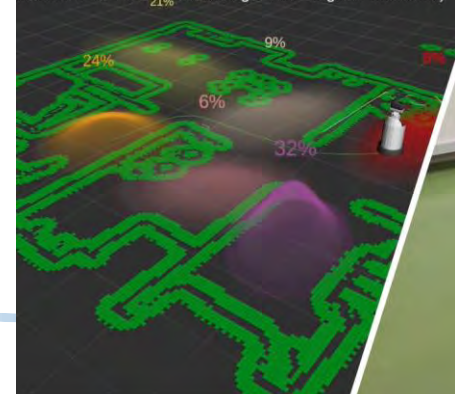
音声言語（場所の名前）



位置分布（空間的な広がり）

マルチモーダル情報に基づく場所概念

distribution of P(place | duck sponge)  
Combination of Onsite Knowledge and Logical Inference



# 今回の講演内容

## 概要

- 人の認知機能は、脳の全体的な働きと身体性による制約、環境との相互作用によって実現される。
- 多様な実世界における認知モデルの実装と実証のためには、身体を持つロボットの活用が重要となるだろう。
- 本講演では、人工知能学会誌の特集「生成AI時代における認知のモデリング」において執筆した記事「脳科学とロボティクスを結ぶ統合認知モデルの実践と展望」の内容をベースに紹介する。
- 主に、脳に学び、ロボットに実装可能な統合認知モデルとして**全脳確率的生成モデル (Whole Brain Probabilistic Generative Model : WB-PGM) のアプローチ**について概説する。
- また、WB-PGM実現に向けた個々の研究事例として、海馬体のナビゲーション機能や聴覚・言語野の二重分節化機能に着目したPGMを紹介する。
- その後、現在の課題と実現への道筋について考察する。

## 工学的な視点から事例紹介的な内容が中心

【特集】「生成AI時代における認知のモデリング」

## 脳科学とロボティクスを結ぶ統合認知モデルの実践と展望

Practice and Prospects of Integrated Cognitive Models Bridging Neuroscience and Robotics

谷口 彰  
Akira Taniguchi  
立命館大学  
Ritsumeikan University  
a.taniguchi@em.cs.ritsumei.ac.jp

山川 宏  
Hiroshi Yamakawa  
全脳アーキテクチャ・イニシアティブ、東京大学、理化学研究所、  
AI アライメント・ネットワーク  
The Whole Brain Architecture Initiative / The University of Tokyo / RIKEN / AI Alignment Network  
yamahiro@riken.jp

谷口 忠大  
Tadahiro Taniguchi  
立命館大学  
Ritsumeikan University  
taniguchi@em.cs.ritsumei.ac.jp

**Keywords:** brain reference architecture, embodiment, integrated cognitive architecture, probabilistic generative model, symbol emergence in robotics.

### 1. はじめに

人の認知機能は、脳の全体的な働きと身体性による制約、環境との相互作用によって実現される。このような認知機能を人工的に実現するためには、(1) 脳全体の機能と構造を参照した認知モデルの構築、(2) その認知モデルを駆動するために身体性をもつロボットへの実装、(3) 実世界において検証すること、が重要である。本稿では、(2)、(3) を見据えた (1) の実現について主に議論する。

昨今の人工知能の発展では、いわゆる生成AIや基礎モデルのように、大規模なデータセットで事前学習された大規模ニューラルネットワークの発展が目覚ましい。GATO, GPT-4V, Gemini を始めとして汎用人工知能 (AGI) の到来を予感させるような成果が今もなお次々と登場している [Google 23, OpenAI 23, Reed 22]。しかし、このような大規模なアプローチは、データ収集や学習にコストがかかる。加えて、生物のように身体性をもつ主観的な経験データを用いた学習や発達、および生物の認知機能や脳の構造との対照にはあまりフォーカスされていない。実世界との相互作用においてもまだ検討すべき課題がある。

人間の認知システムを包括的に模倣することで、よりリアルな環境で機能するロボットの認知システムを開発することは重要である。認知発達を達成するため

の計算モデルは、脳科学や認知科学にとっても優れた仮説や知見を提供することが期待される。そのため、脳に学び、ロボットに実装可能な統合モデルの開発フレームワークとして、**全脳確率的生成モデル (Whole Brain Probabilistic Generative Model : WB-PGM)** [Taniguchi 22d] の構築が必要であると考えられる。

このアプローチは、脳のアーキテクチャに基づいたAIロボットの開発と、確率的生成モデル (PGM) に基づく認知アーキテクチャの開発という二つの主要な要素に焦点を当てる。これらの要素はAI、ロボティクス、認知科学、神経科学の交差点に位置付けられ、それぞれがWB-PGMの発展に重要な役割を果たす。

WB-PGMのアプローチは、図1に示すように脳参照アーキテクチャ (Brain Reference Architecture : BRA) 駆動開発 [Yamakawa 21] と Symbol Emergence in Robotics Toolkit (SERKET) フレームワーク [Nakamura 18, Taniguchi 20b] の融合による脳全体の認知モデリングへの取り組みである (詳細は3章に記載)。WB-PGMは、Soar, ACT-R, Sigmaなどの従来の認知アーキテクチャ [Anderson 09, Laird 12, Rosenbloom 16] が記号的AIに基づいていたことは対照的に、ロボットの感覚運動情報に基づくシステムを通じて連続的に学習できる。

本稿では、全脳アーキテクチャ (Whole Brain Architecture : WBA) アプローチに触発され、統合的な認知システムとしてのWB-PGMの構築へ向けたAI

# 全脳確率的生成モデル(WB-PGM)



## Whole brain Probabilistic Generative Model

- PGMによる全脳アーキテクチャ構成のためのパースペクティブ論文 [Taniguchi et. al., 2021]を共同執筆

- 目的

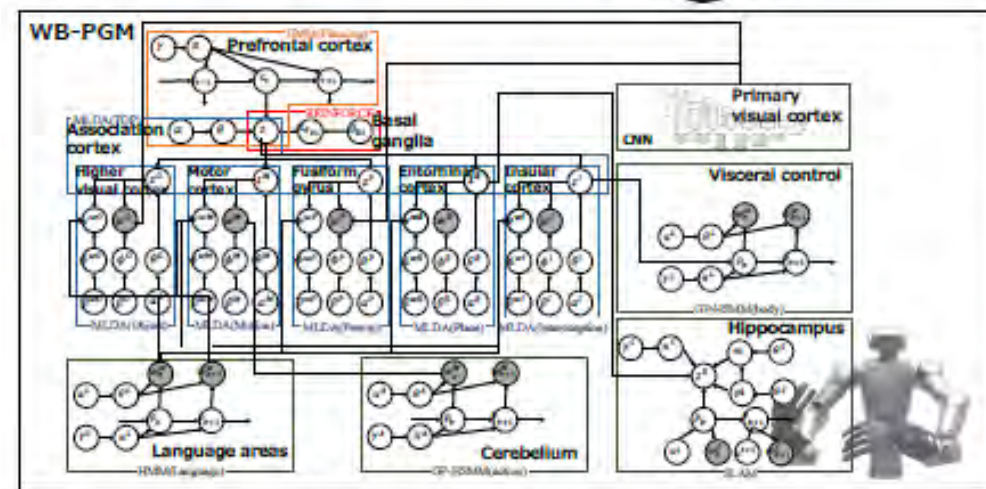
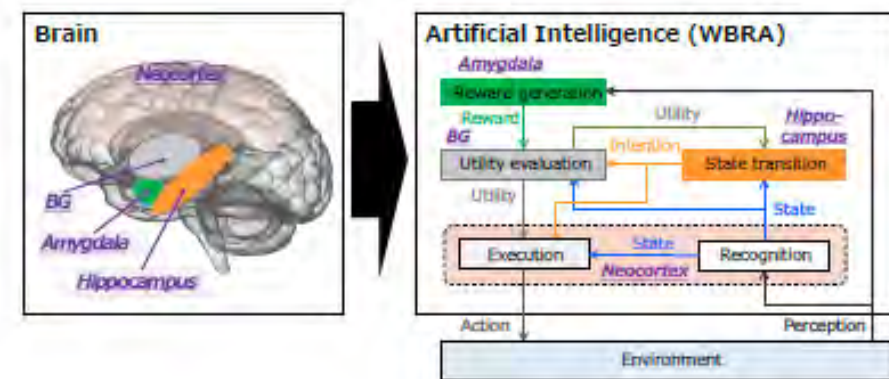
- 汎用人工知能 (AGI) の実現
- 人の脳に学び、脳を模した認知アーキテクチャの構築
- 確率的生成モデルを用いた統合的認知システムの構築

- Research Question

- どのような認知モジュールを実装すべきか?
- それらをどのように統合させ連携されればよいか?

- 以下の2つのアプローチを組み合わせることで、これらの疑問に答える。

- 脳にインスパイアされたAI**：人間レベルの知能を創り出すために、人間の脳アーキテクチャから学ぶ
- PGMベースの認知システム**：記号創発・認知発達ロボティクスにおける構成論的アプローチとしてPGMを統合した統合的認知システムを開発する



# 実世界での認知モデルの意義

- 人の認知機能は、**脳の全体的な働きと身体性による制約、環境との相互作用**によって実現される。
- このような認知機能を人工的に実現するためには
  - (1)脳全体の機能と構造を参照した認知モデルの構築
  - (2)その認知モデルを駆動するために身体性を持つロボットへの実装
  - (3)実世界において検証が重要である。
- 認知システムは、システム自身が観測するマルチモーダルな感覚運動情報を自己組織化し、様々なスキルを学習する必要がある。そのためには、**物理的および社会的環境**を能動的かつ自律的に探索する身体が必要であり、**実世界で行動できるロボットを使用**することが有効であると考えられる。

⇒ **記号創発ロボティクスにおける構成論的アプローチ**

# BRAとしての確率的生成モデル (PGM) 構築のための変換方法

# 確率的生成モデルとは？

## Probabilistic generative model (PGM)

- データが生成される過程をモデル化したもの
- 確率分布を仮定することでデータの統計的な性質をとらえる
- グラフィカルモデルによってモデル構造を可視化できる
- グラフィカルモデル
  - 変数間の依存関係をグラフ構造で表現
  - ノード：確率変数
  - エッジ：依存関係

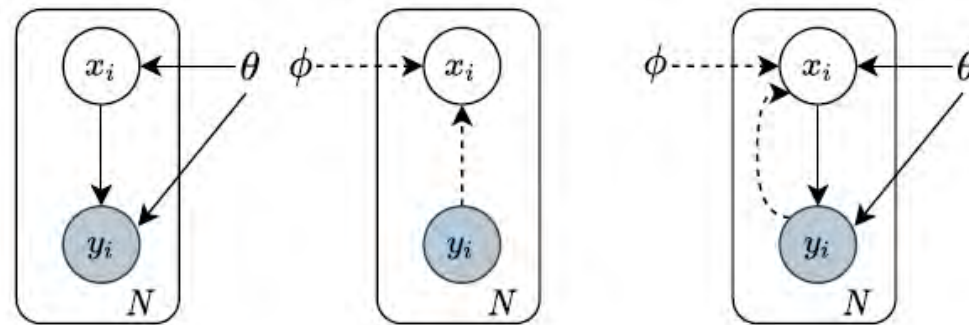
生成モデル

$$x_i \sim p(x_i|\theta) = p_\theta(x_i)$$

$$y_i \sim p(y_i|x_i, \theta) = p_\theta(y_i|x_i)$$

推論モデル

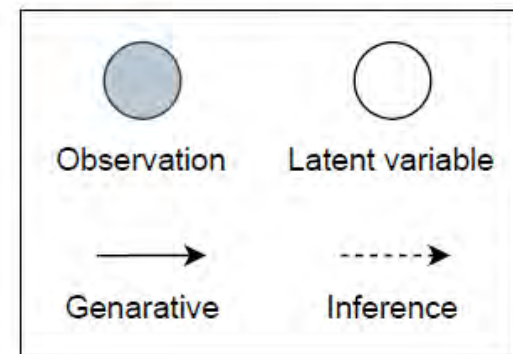
$$q(x_i|y_i, \phi) = q_\phi(x_i|y_i)$$



(a)  
生成モデル

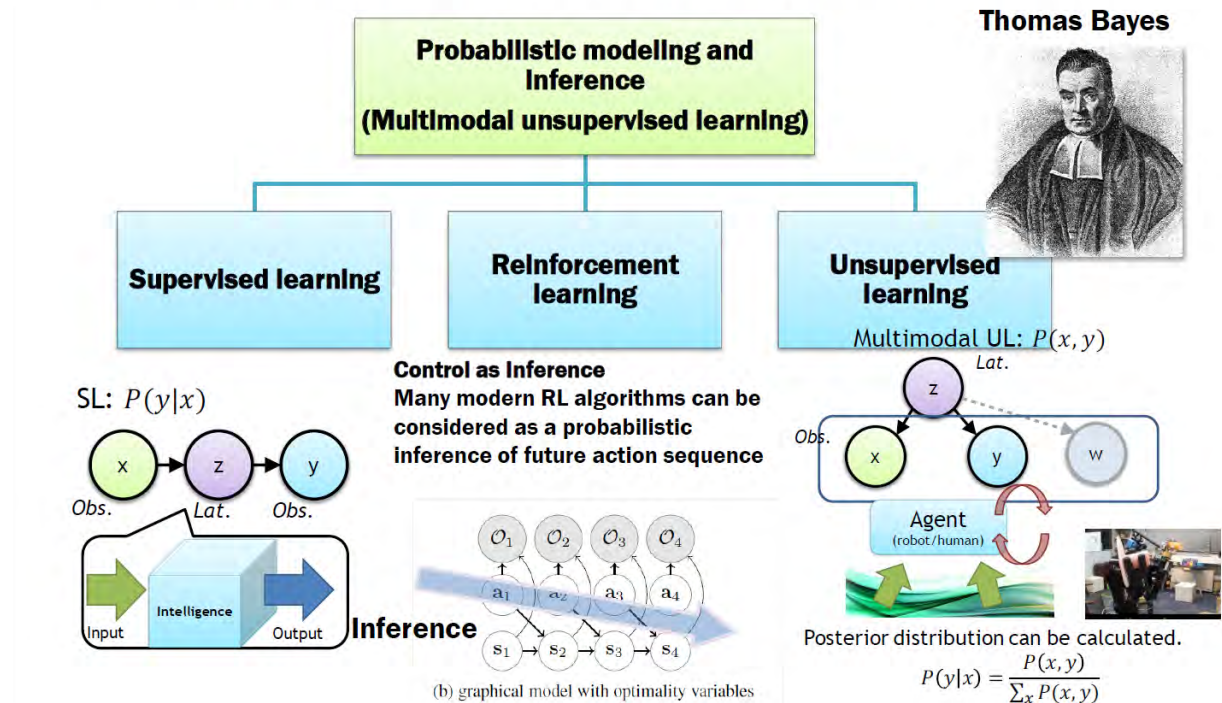
(b)  
推論モデル

(c)  
両方 (VAE)



# なぜ確率的生成モデル？

- グラフィカルモデルによって構造を容易に可視化できる
- PGM同士を統合しやすい  
(Neuro-SERKETフレームワーク [Taniguchi 2020])
  - 全脳PGMへ
- 教師あり学習、教師なし学習、強化学習をすべてPGMの上で議論できる
- 神経科学分野などでも議論・理論構築されている
  - ベイズ脳仮説 [Doya 2007]
  - 自由エネルギー原理/予測符号化仮説 [Friston 2012]
  - 世界モデル (World model)



↑ 谷口忠大先生スライドより

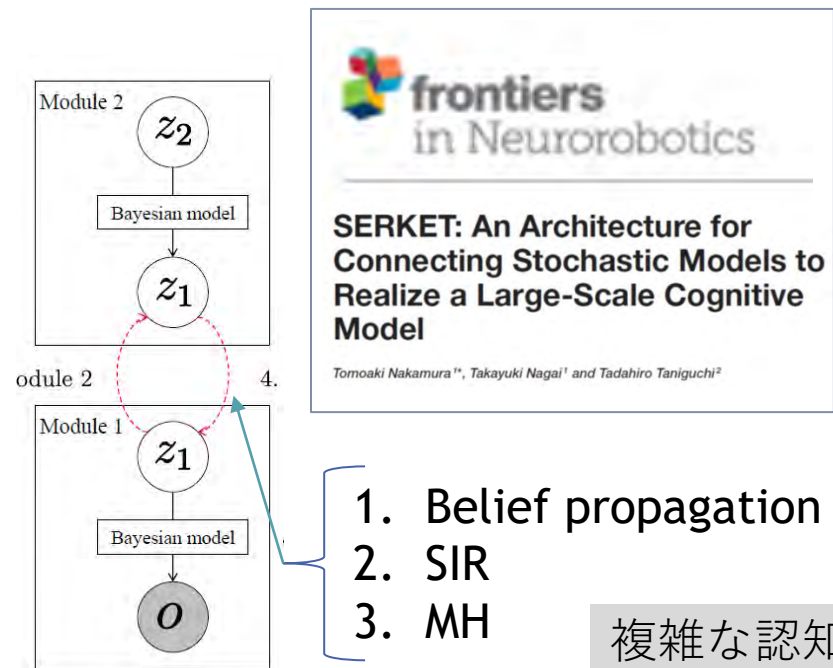
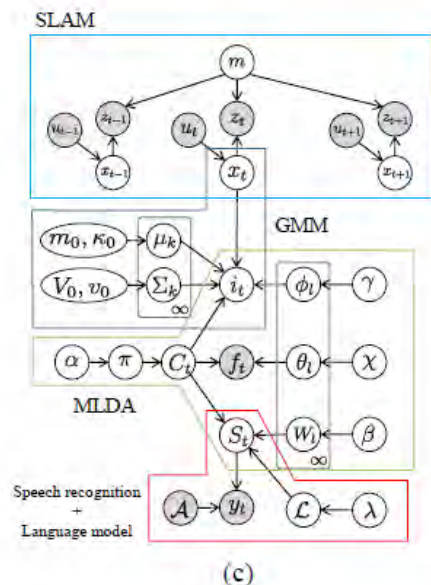
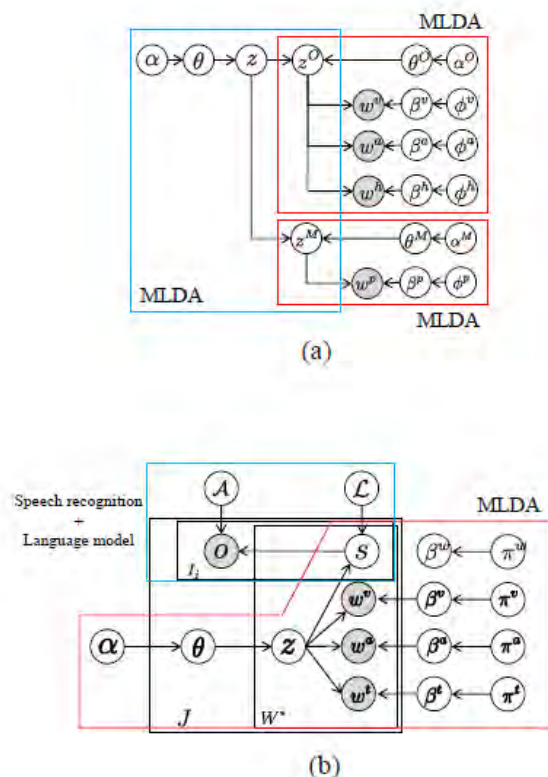
[Taniguchi 2020] Tadahiro Taniguchi, Tomoaki Nakamura, Masahiro Suzuki, Ryo Kuniyasu, Kaede Hayashi, Akira Taniguchi, Takato Horii, and Takayuki Nagai, "Neuro-SERKET: Development of Integrative Cognitive System through the Composition of Deep Probabilistic Generative Models", New Generation Computing, Jan. 2020.


[Doya 2007] Doya, K., Ishii, S., Pouget, A., Rao, R.P.N., Bayesian Brain: Probabilistic Approaches to Neural Coding. MIT Press. 2007.

[Friston 2012] Friston, Karl, and Ping Ao. "Free energy, value, and attractors." Computational and mathematical methods in medicine 2012.



# Neuro-SERKET: Development of Integrative Cognitive System Through the Composition of Deep Probabilistic Generative Models [Taniguchi+ 20]



複雑な認知システムを効率的に開発可能に 

- ❑ Connecting cognitive modules developed as probabilistic generative models and letting them work together as a single unsupervised learning system.
- ❑ Having inter-module communication of probabilistic information and guaranteeing theoretical consistency to some extent.
- ❑ Neuro-SERKET supports deep generative models, i.e., VAE, as well.

Nakamura T, Nagai T and Taniguchi T, SERKET: An Architecture for Connecting Stochastic Models to Realize a Large-Scale Cognitive Model. *Front. Neurobot.* 12:25. (2018) doi: 10.3389/fnbot.2018.00025

Taniguchi, T., Nakamura, T., Suzuki, M. et al. Neuro-SERKET: Development of Integrative Cognitive System Through the Composition of Deep Probabilistic Generative Models. *New Gener. Comput.* 38, 23–48 (2020). <https://doi.org/10.1007/s00354-019-00084-w>

# SERKET-SDE

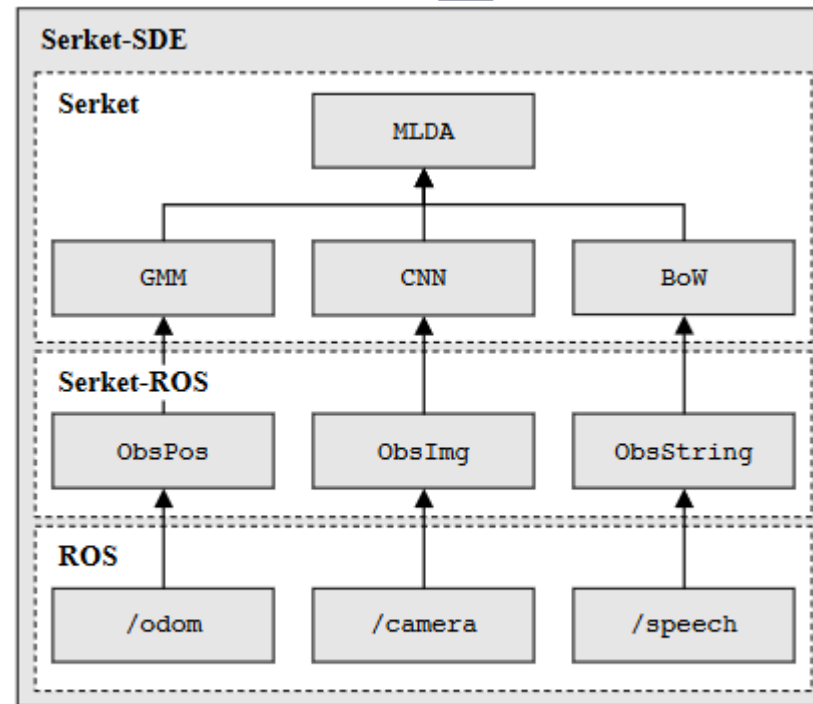
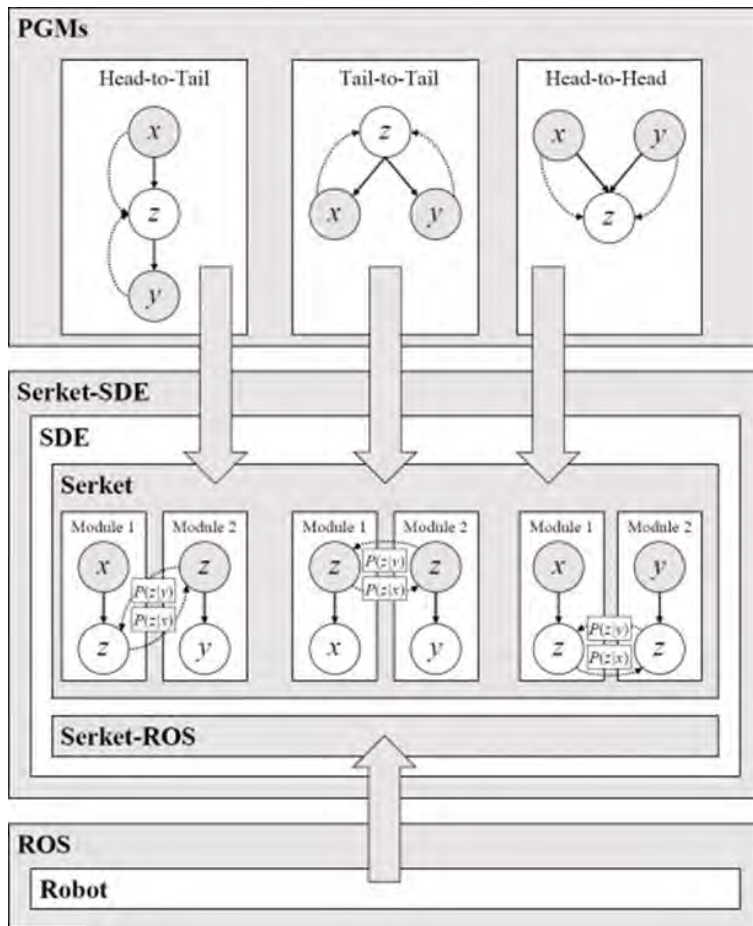


Fig. 4. Overview of the Serket architecture used for the multimodal place categorization experiments.

- SERKETとROSを接続するブリッジを提供
- 個別モジュールの独立した開発と統合モデルにおけるモジュール組み換えを容易に行うことができる

```
import serket as srk
import serket_ros as srkros

class PlaceCategorization():

    def __init__(self, category_nbr):

        # Define the Serket-ROS interfaces.
        self.obs1 = srkros.ObservationPos("/odom")
        self.obs2 = srkros.ObservationImg("/camera")
        self.obs3 = srkros.ObservationString("/speech")

        # Define the Serket modules.
        self.obs1_buf = Buffer()
        self.gmm1 = gmm.GMM(category_nbr)
        self.cnn1 = CNN.CNNFeatureExtractor()
        self.cnn1_buf = Buffer()
        self.bowl = bow.Bow()
        self.bowl_buf = Buffer()
        self.mlda1 = mlda.MLDA(category_nbr)

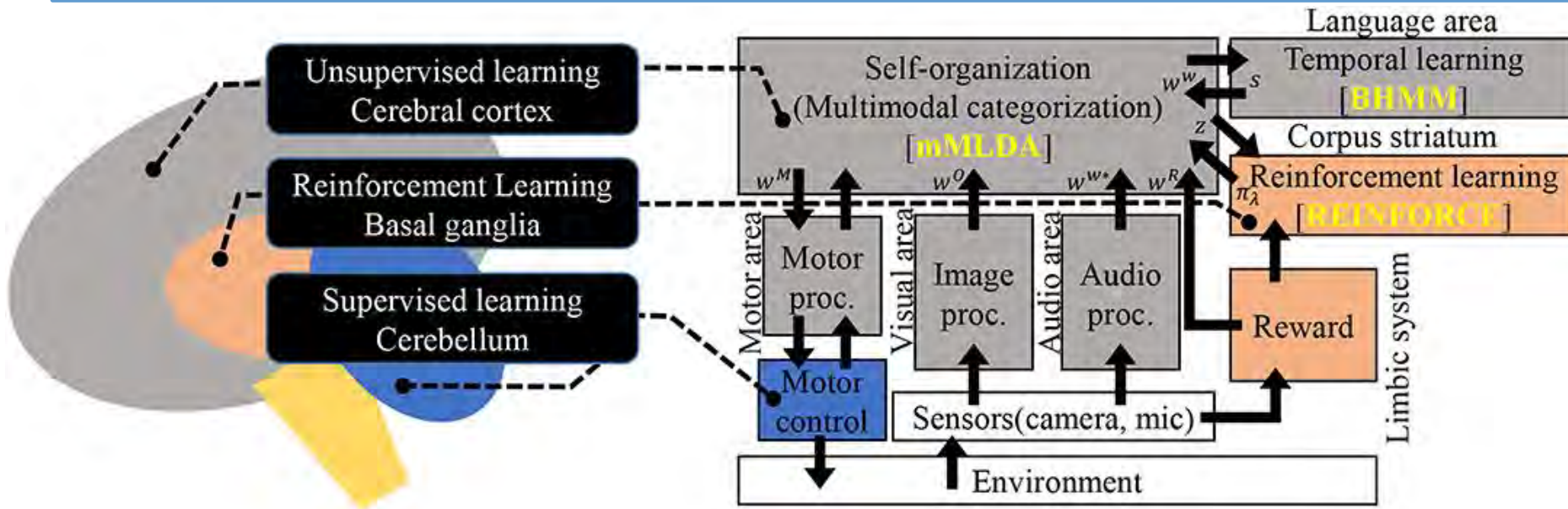
        # Define the Serket architecture.
        self.obs1_buf.connect(self.obs1)
        self.gmm1.connect(self.obs1_buf)
        self.cnn1.connect(self.obs2)
        self.cnn1_buf.connect(self.cnn1)
        self.bowl.connect(self.obs3)
        self.bowl_buf.connect(self.bowl)
        self.mlda1.connect(self.gmm1,
                           self.cnn1_buf, self.bowl_buf)

    def update(self):

        # Update the Serket architecture.
        self.obs1.update()
        self.obs1_buf.update()
        self.obs2.update()
        self.obs3.update()
        self.gmm1.update()
        self.cnn1.update()
        self.cnn1_buf.update()
        self.bowl.update()
        self.bowl_buf.update()
        self.mlda1.update()
```

Listing 1. Implementation of the PlaceCategorization class using Serket.

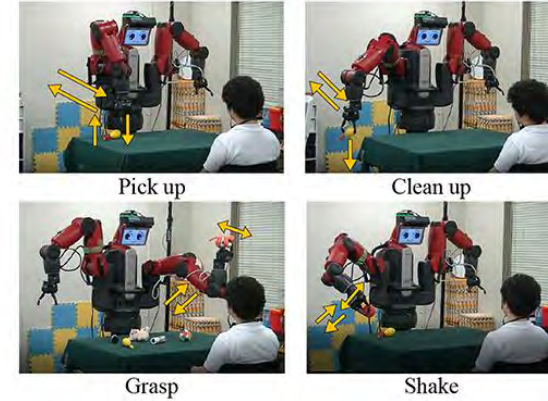
# プロトタイプモデル



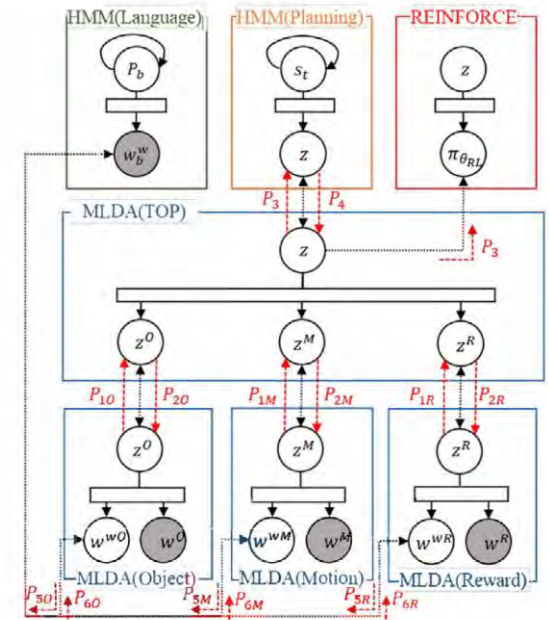
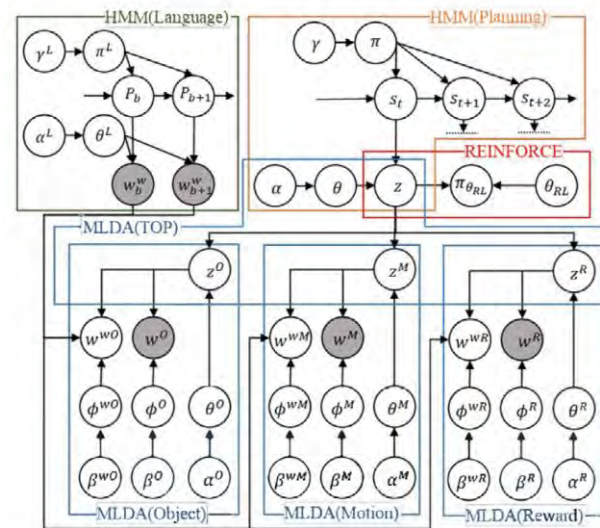
A



C



B



- 表現学習、言語学習、強化学習モジュールをNeuro-SERKETを使用して統合
- 脳の主要な機能と代表的な脳領域レベルの接続を参考に作られた
- 今後は、BRA駆動開発へと繋げることが重要
- Miyazawa, K., Aoki, T., Horii, T., & Nagai, T.: Integration of Multiple Generative Modules for Robot Learning. Workshop on Deep Probabilistic Generative Models for Cognitive Architecture in Robotics (in IROS2019). (2019)
- Miyazawa K, Horii T, Aoki T and Nagai T (2019) Integrated Cognitive Architecture for Robot Learning of Action and Language. Front. Robot. AI 6:131.

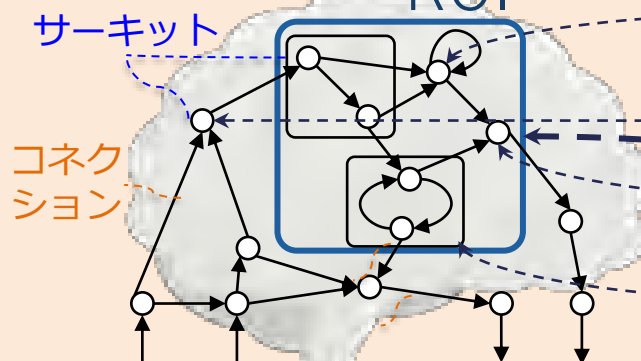
# 脳参照アーキテクチャ (BRA) と BRA 駆動開発

## 脳参照アーキテクチャ (BRA)

### 脳情報フロー (BIF)

メゾスコピック・レベルの解剖学的構造上の情報フロー

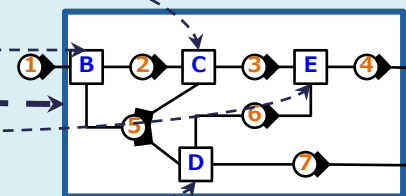
ROI



制約

### 仮説的コンポーネント図

ROIが担う最上位機能を解剖学的構造と一致するように機能を整理しようとする図



コンポーネントの機能: A, B, C, D, E  
信号の意味: 1, 2, 3, 4, 5, 6

開発

脳型ソフトウェア

脳型ソフトウェア

脳型ソフトウェア

脳の真実

構築

神経科学の論文・データ

構築

BRA設計

BRA活用

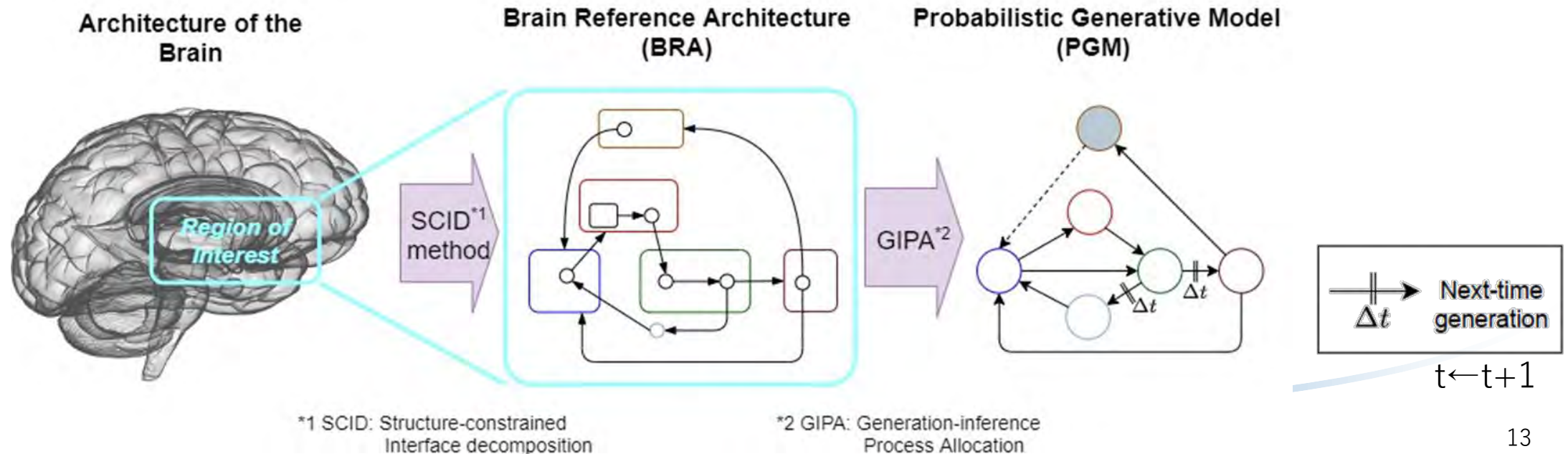
# GIPA: Generation-inference process allocation

- 脳部位の解剖学的構造を生成モデルの変数の依存関係へ変換
- PGMの制約
  - Directed acyclic graphsである必要性 (ループを作れない)
- 基本的なGIPAの手順
  1. エッジを生成モデルと推論モデルに分ける (Table 1を主な基準とする)
  2. 次時刻生成( $\Delta t$ )の表現を導入する

Table 1: Counter stream pathways in neocortex

| Pathways                                                      | Feedforward                                                              | Feedback                                                                 |
|---------------------------------------------------------------|--------------------------------------------------------------------------|--------------------------------------------------------------------------|
| Direction                                                     | From the outside world to the internal state (from lower to higher area) | From the internal state to the outside world (from higher to lower area) |
| Laminar on cortical microcircuits (Markov et al., 2014, 2013) | Layer 3, and Layer 4                                                     | Layer 2                                                                  |
| Meaning of signals (Yamakawa, 2020)                           | Observation                                                              | Prediction                                                               |
| Graphical model representation of PGM                         | inference process                                                        | generative process                                                       |

WB-PGM論文より



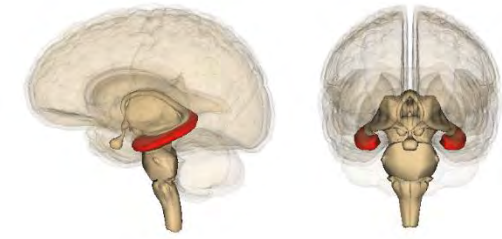
Hippocampal formation-inspired probabilistic generative model

# HF-PGM：海馬体の確率的生成モデル ～認知モジュール構築の事例として～

Akira Taniguchi, Ayako Fukawa, Hiroshi Yamakawa, "Hippocampal formation-inspired probabilistic generative model", Neural Networks, Vol.151, No. July, pp.317-335, 2022. DOI:  
10.1016/j.neunet.2022.04.001  
<https://www.sciencedirect.com/science/article/pii/S0893608022001332>

# なぜ海馬体なのか

**海馬体：脊椎動物の記憶を司る脳部位**



空間学習や空間探索といった空間記憶に深く関わる

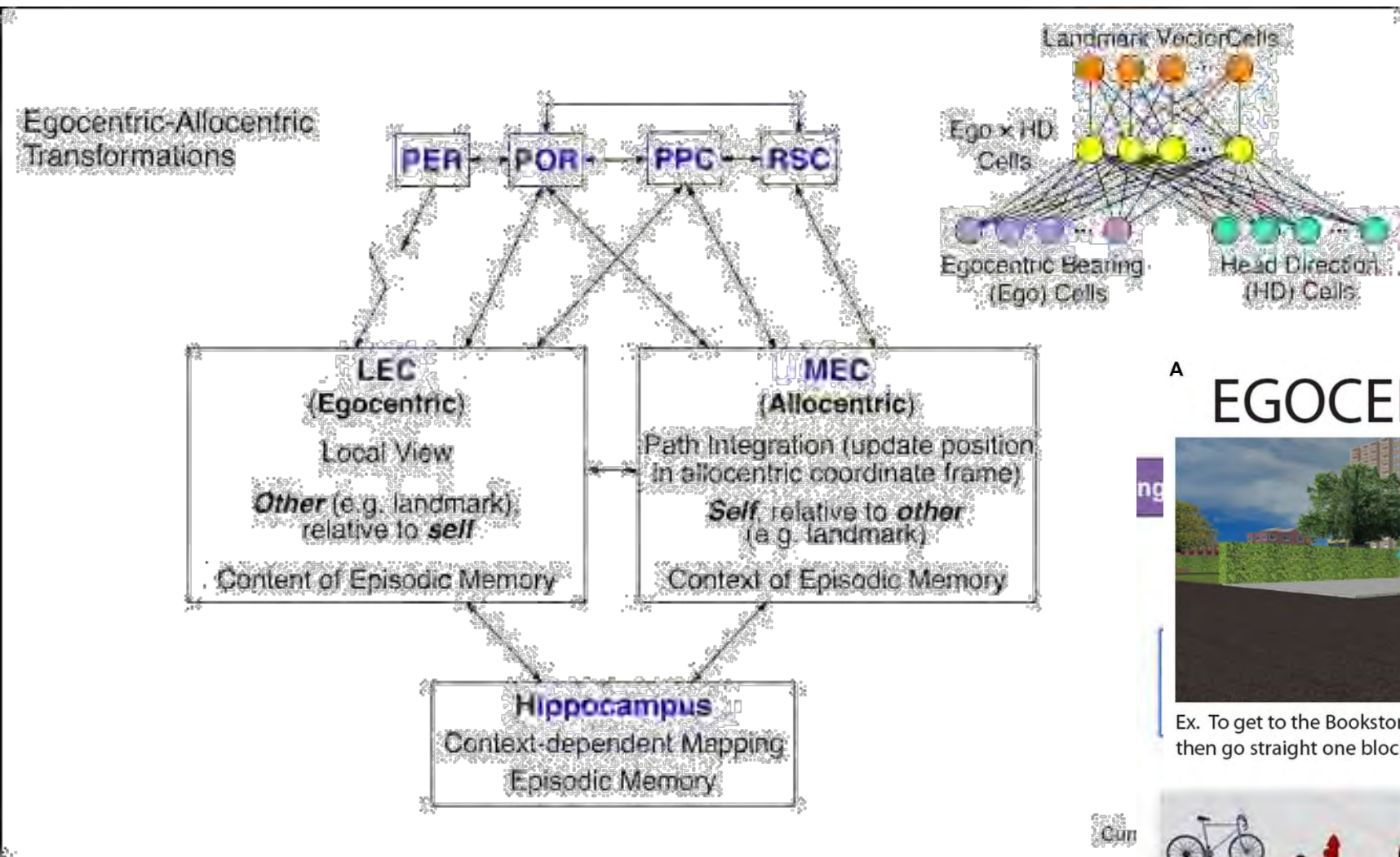
認知的処理を担うコア領域

脳型の汎用人工知能の構築に  
不可欠の構成要素

海馬体に関する  
数多くの神経科学知見が  
蓄積されている

# Egocentric and allocentric representations of space in the rodent brain

LECがエゴセントリック, MECがアロセントリック Current Opinion in Neurobiology 2020



## A EGOCENTRIC B ALLOCENTRIC



Ex. To get to the Bookstore I need to turn right then go straight one block and turn left.



Ex. The Bookstore is 30 meters north of the Gym and 20 meters east of the Camera Store.



perirhinal cortex (PER) and postrhinal cortex (POR)  
posterior parietal cortex (PPC) and retrosplenial cortex (RSC)

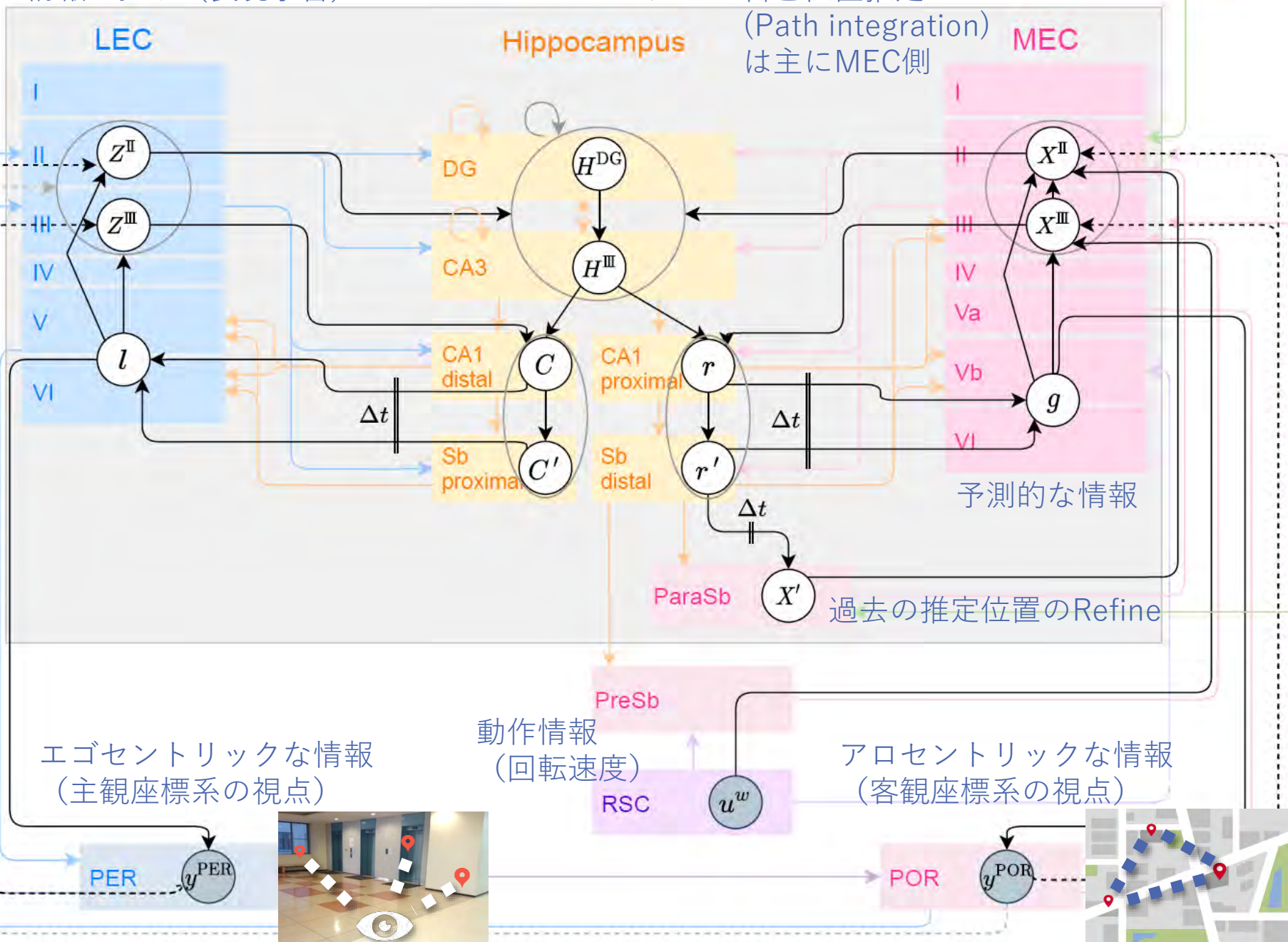


LEC側は場所の概念的な情報を担う (表現学習)

LECとMECからの2つの情報を統合

自己位置推定 (Path integration) は主にMEC側

Medial septum



| Symbol    | Function of components on HCD                                                                                                  |
|-----------|--------------------------------------------------------------------------------------------------------------------------------|
| $C$       | Place category (internal representation of visual spatial information)                                                         |
| $r$       | Position distribution (cluster information regarding positions)                                                                |
| $C'$      | Place category at the previous time                                                                                            |
| $r'$      | Position distribution at the previous time                                                                                     |
| $H^{III}$ | Integrated semantic memory and episodic memory of information from $X$ and $Z$                                                 |
| $H^{DG}$  | Integrated semantic memory                                                                                                     |
| $X^{II}$  | Self-position information, predictive distribution ( $\{x_t\}$ )                                                               |
| $X^{III}$ | Self-posture information (position and orientation), observation likelihood ( $\{x_t\}$ )                                      |
| $g$       | Predictive representation of $X$ (Prediction at future time regarding movement/speed amount or posture)                        |
| $Z^{II}$  | Abstraction of information from $y^{PER}$ (transmission of prediction, generation of prediction signal)                        |
| $Z^{III}$ | Abstraction of information from $y^{PER}$ (Observation transmission)                                                           |
| $l$       | Predictive representation of $Z$ (Prediction at future time from the difference between $C'$ and $C$ )                         |
| $X'$      | Self-posture information ( $\{x_{t-1}\}$ )                                                                                     |
| $y^{POR}$ | Allocentric visual information (distal distance/landmarks, absolute object positions)                                          |
| $y^{PER}$ | Egocentric visual information (proximal distance/landmarks, relative object positions, object category, landscape information) |
| $u^w$     | Rotational speed movement ( $\{u_t\}$ )                                                                                        |

エゴセントリックな情報 (主観座標系の視点)

動作情報 (回転速度)

アロセントリックな情報 (客観座標系の視点)

予測的な情報

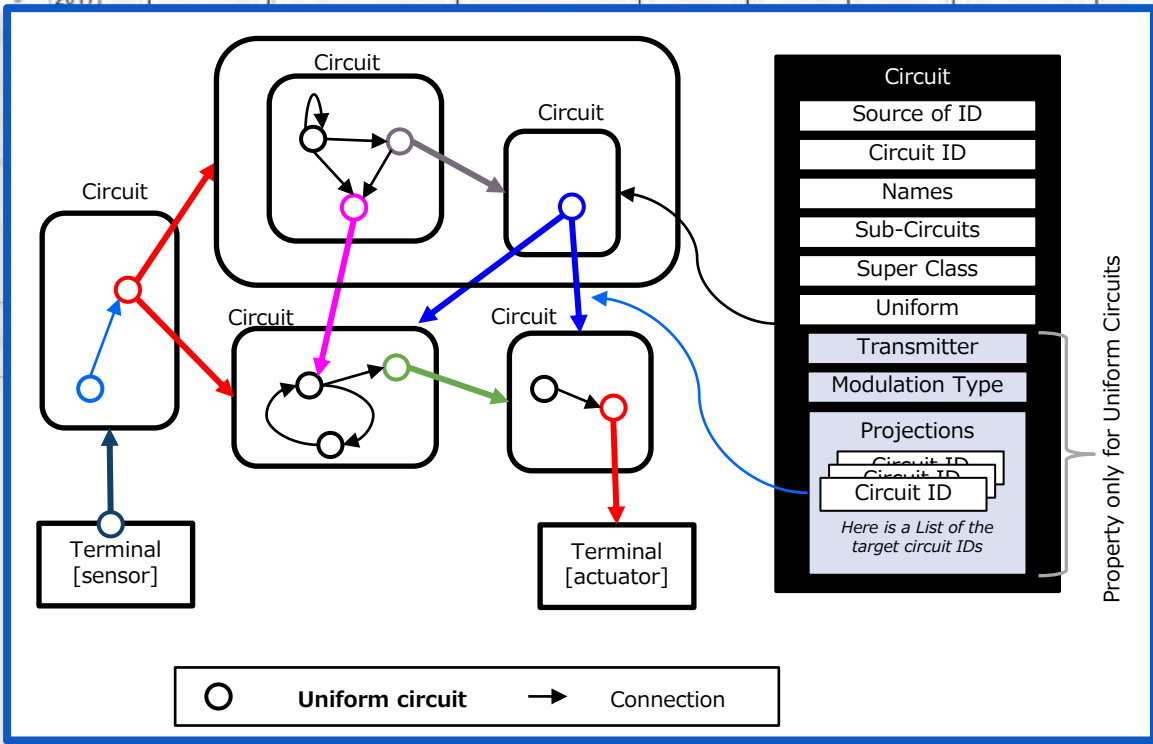
過去の推定位置のRefine

# 海馬体のBRA形式データ (HF.bra)

| Source of ID | Circuit ID | Names         | Sub-Circuits    | Super Class | Uniform | Transmitter | Modulation Type | Size | Projections                                                                                                                                                                                                             | Interpretation( $\phi$ & $f$ )                                                                                                                                                                                                                                                             | Label(1)             | Function(1)                                                                                                                        | Projections in Use(1) |
|--------------|------------|---------------|-----------------|-------------|---------|-------------|-----------------|------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------|------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| collection   | HPF        | Hippocampal   | Hipp: LEC: MEC: |             | FALSE   |             |                 |      | PERI (origin LEC_deep_uc1);                                                                                                                                                                                             |                                                                                                                                                                                                                                                                                            | HPF                  | Top Level Functions for HPF: (1) Self-localization, (2) Place categorization by integrating allocentric and egocentric information | PERI (GIPA:G)         |
| DHBA         | Hipp       | Hippocampus   |                 |             |         |             |                 |      |                                                                                                                                                                                                                         |                                                                                                                                                                                                                                                                                            |                      |                                                                                                                                    |                       |
| DHBA         | CA1        | cornu Ammonii |                 |             |         |             |                 |      | MEC_III [Moser, 2010]; MEC_V [Moser, 2010]; Medial septum [Rolls, 2013]; Nucleus accumbens [Rolls, 2013]; PFC.L2 [Verwer, 1997]; PFC.L3 [Verwer, 1997]; S [Llorens-Martin, 2014];                                       | Ensemble activity in CA3, but not in CA1, is sensitive to the content of a scene or place in the environment. CA3 supports the retrieval of memories about where an object is in a scene [Robertson, 1998] [Rolls, 1998] or where a reward could be found in the environment [Brun, 2002]. |                      |                                                                                                                                    |                       |
| [Bubb, 2017] | CA1_distal | CA1, distal   |                 |             | TRUE    |             | Excitatory      |      | S_proximal [Knierim, 2014]; LEC_deep [Knierim, 2014]; MEC_III [Moser, 2010]; S_distal [Knierim, 2014]; MEC_deep [Knierim, 2014]; CA1 [Llorens-Martin, 2014]; CA3 [Llorens-Martin, 2014]; MEC_II [Llorens-Martin, 2014]; | Non-spatial semantic memory [author, year]                                                                                                                                                                                                                                                 | C                    | Place category (internal representation of visual spatial information)                                                             | S_proximal (GIPA:G)   |
|              |            |               |                 |             |         |             |                 |      | CA1 [Llorens-Martin, 2014]; CA2 [Llorens-Martin, 2014][Knierim, 2015]; CA3 [Knierim, 2015]; DG [Llorens-Martin, 2014];                                                                                                  | Place cells [Brun, 2002] [O'Keefe, 1976]                                                                                                                                                                                                                                                   | r                    | Position distribution (cluster information regarding positions)                                                                    | S_distal (GIPA:G)     |
|              |            |               |                 |             |         |             |                 |      | CA1_distal [Knierim, 2015]; CA1_proximal [Knierim, 2015]; CA2 [Llorens-Martin, 2014][Knierim, 2015]; CA3 [Knierim, 2015]; DG [Llorens-Martin, 2014];                                                                    | Pattern completion, information integration [Kesner, 2007] [[Rolls, 2007]                                                                                                                                                                                                                  | H <sup>2</sup> (III) | Integrated semantic memory and episodic memory of information from X and Z                                                         | CA1_distal (GIPA:G)   |
|              |            |               |                 |             |         |             |                 |      | CA2 [Llorens-Martin, 2014]; CA3 [Llorens-Martin, 2014]; DG [Knierim, 2015];                                                                                                                                             | Pattern separation [Kesner, 2007] [[Rolls, 2007]                                                                                                                                                                                                                                           | H <sup>2</sup> (DG)  | Integrated semantic memory                                                                                                         | CA3 (GIPA:G)          |

脳情報フロー (BIF)

仮説的コンポーネント図 (HCD)



## 参考資料:

- [脳参照アーキテクチャ\(BRA\)データの記述マニュアルおよび具体例の公開](#)
- [第1回 WBAレクチャー～認知機能の脳構造に沿った分解手法—脳機能の体系的理解を目指して](#)

# 研究アプローチ

海馬体の概要

## LEC (外側嗅内皮質)

- ✓ 場所の概念的な表現学習
- ✓ エゴセントリックな視点



## 海馬

## MEC (内側嗅内皮質)

- ✓ 環境のメトリックな処理
- ✓ アロセントリックな視点



本研究

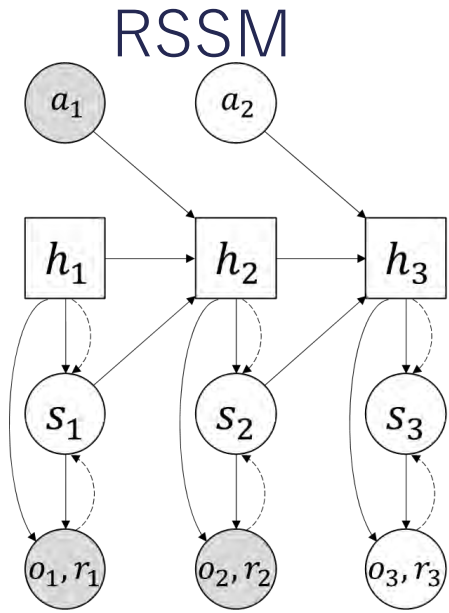
世界モデル  
(RSSM/Multimodal RSSM)

SLAM  
(地図生成と自己位置の同時推定)

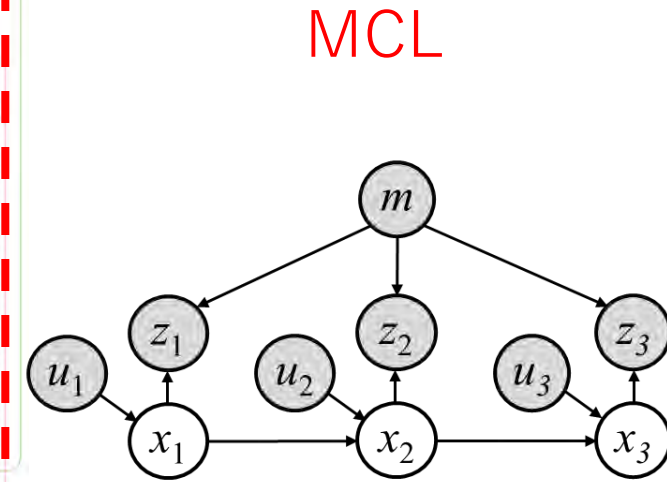
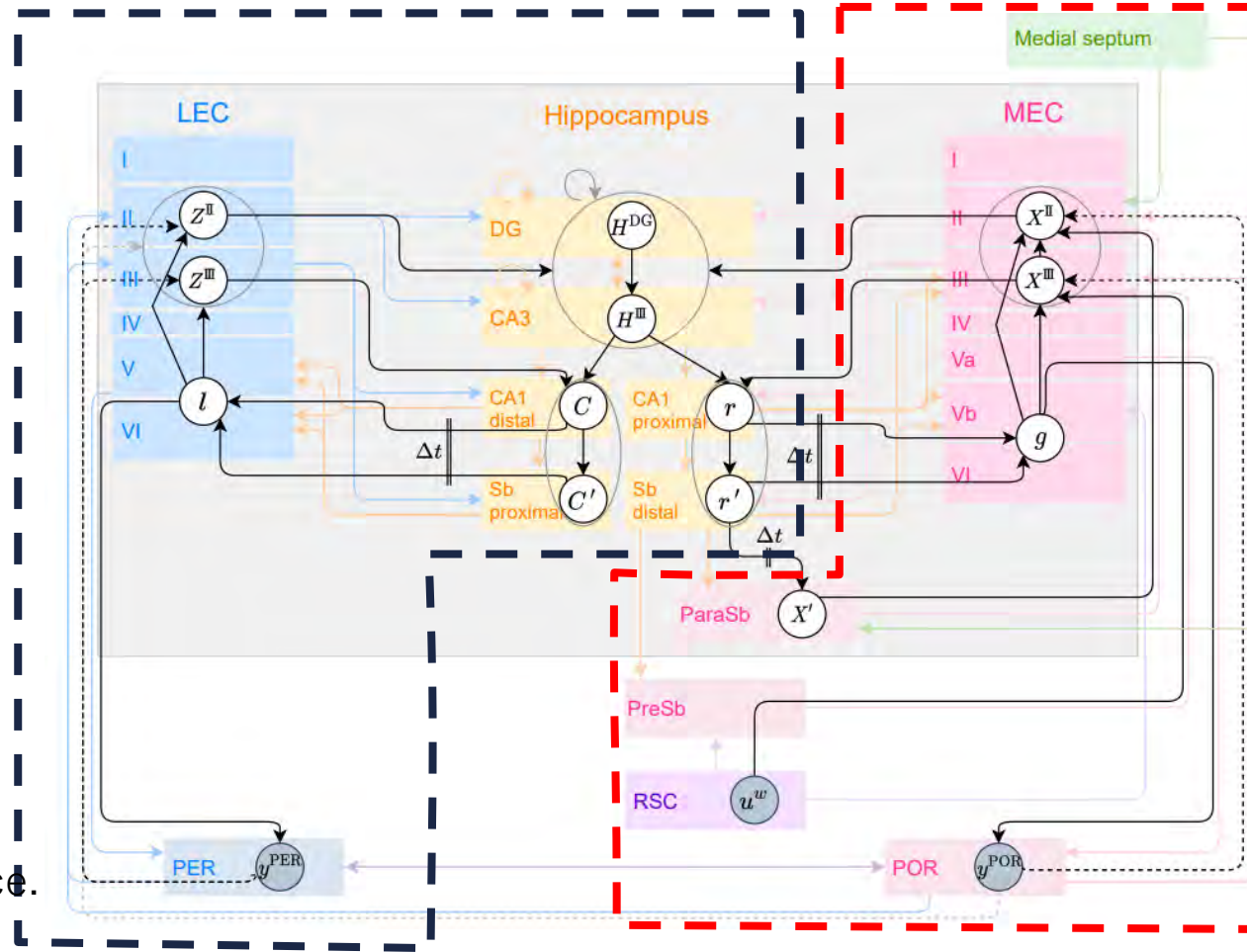
○大竹 俊輔<sup>1</sup>、前山 功伊<sup>1</sup>、長谷川 翔一<sup>1</sup>、中島 毅士<sup>1</sup>、谷口 彰<sup>1</sup>、谷口 忠大<sup>1</sup>、山川 宏<sup>2,3</sup> (1. 立命館大学、2. 全脳アーキテクチャ・イニシアティブ、3. 東京大学)、海馬体に学んだ確率的生成モデルの実装と有効性検証、人工知能学会全国大会, Vol. 37, No. 1Q3-OS-7a-05, 熊本, 2023年6月.

# Proposed method

Combine RSSM and MCL to create a model that **approximately** follows the HF-PGM.



Extract representation of place.



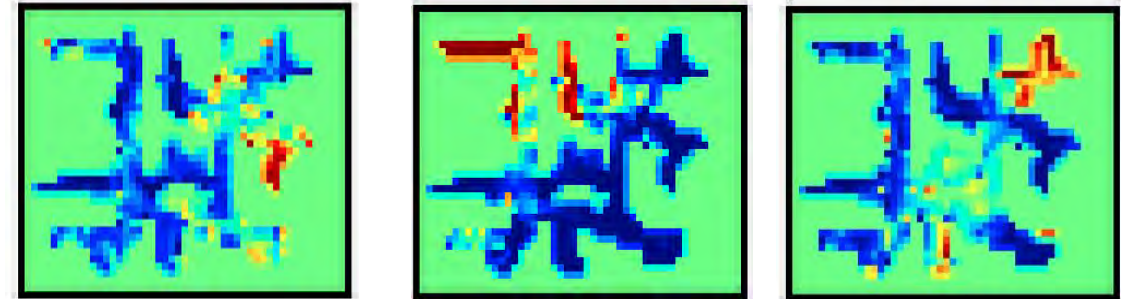
Estimate self-location.

Two different merging methods. (Model1 , Model2)

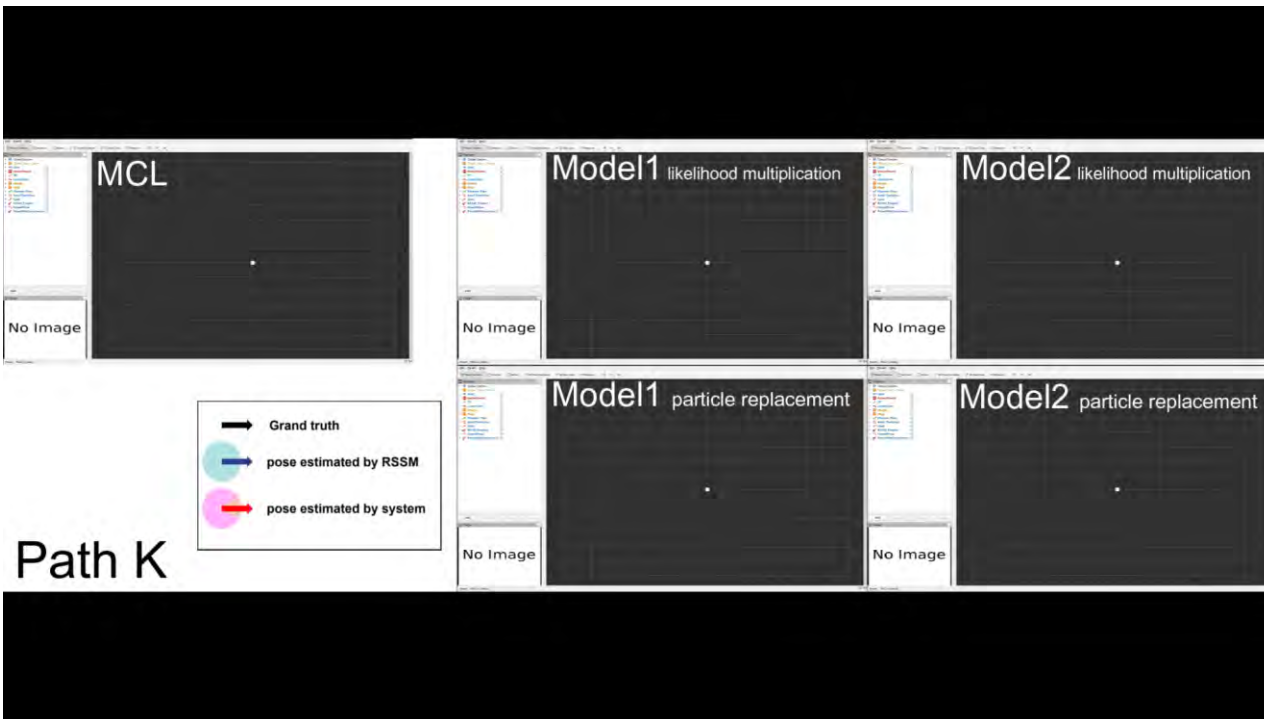
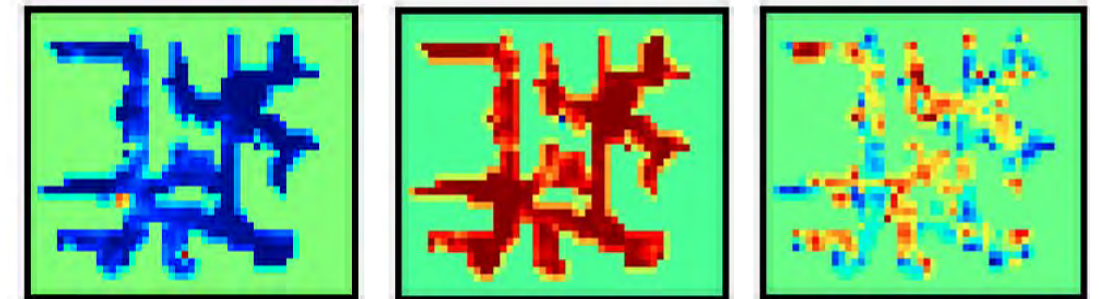
# RSSM + MCLによる大域的自己位置推定、誘拐ロボット問題への対応

## ■ Visualize latent variables as a cell

### latent variables like a place cell



### latent variables difficult to interpret



- ✓ 提案モデル2のパーティクル置換によって統合したモデルがMCL (ベースライン) より優位
- ✓ モデルの統合により、**パーティクル枯渇問題・誘拐ロボット問題を回避**する効果を得た

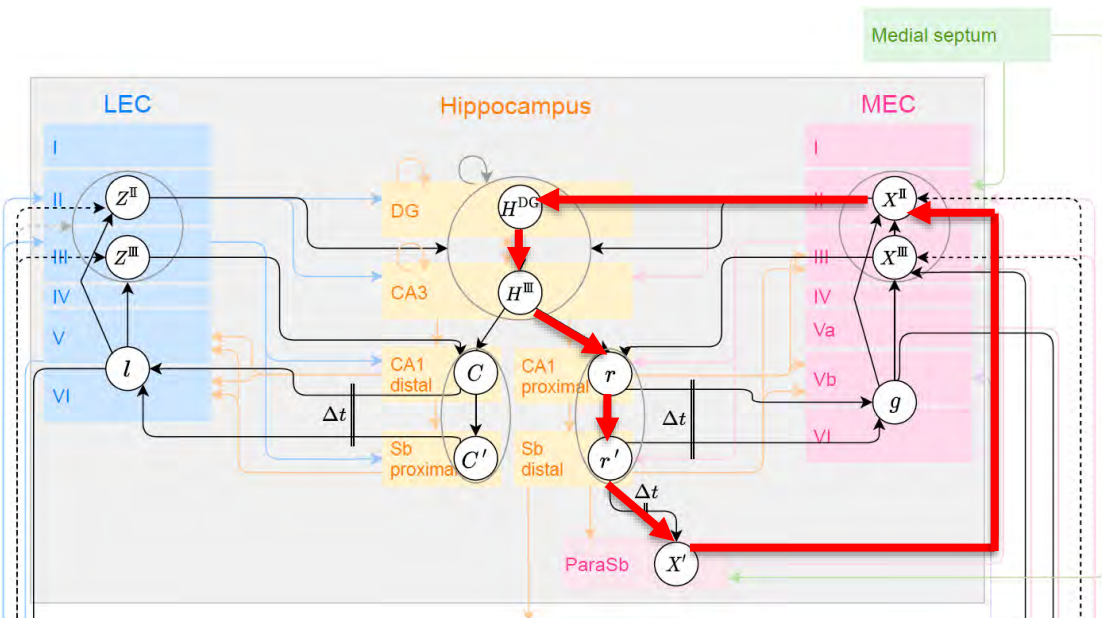
# DEQ-MCL: Discrete-Event Queue-based Monte-Carlo Localization

This research outcomes will be presented at the following international conference:

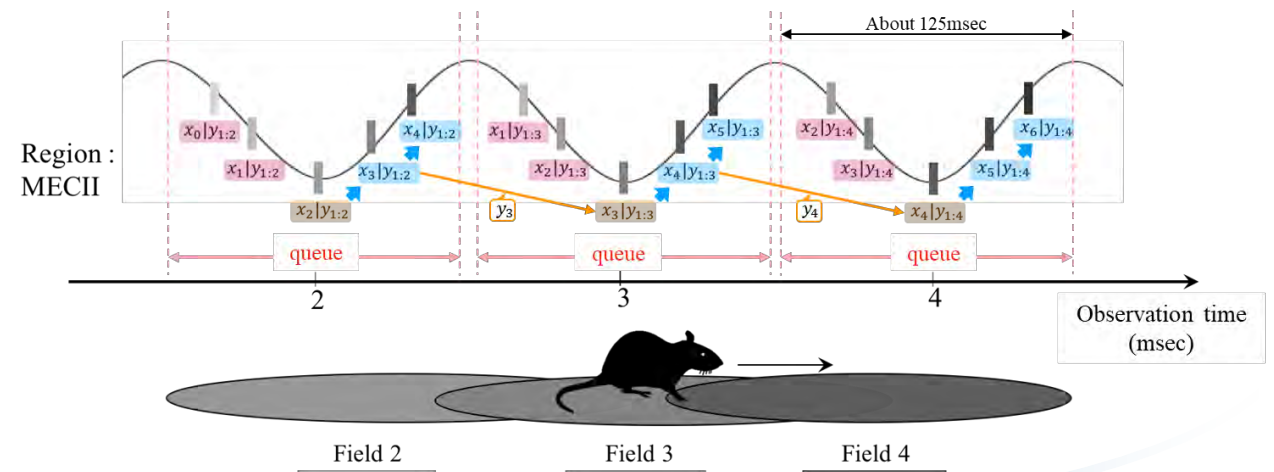
•Akira Taniguchi, Ayako Fukawa, Hiroshi Yamakawa, "DEQ-MCL: Discrete-Event Queue-based Monte-Carlo Localization", AROB-ISBC-SWARM 2024. (Accepted)

# Abstraction as discrete-event queue

- シータ波位相歳差によって経験を離散化し、時間的に圧縮して処理することが知られている
- 現在、過去、未来の事象が位相内に符号化されている**と考えられている [Terada et al., 2017]
- 位相歳差キュー仮説 (Phase precession queue assumption)
  - Pentasyntacticループ回路上を循環するシータ波の1周期内に圧縮された観測状態は、離散的な時間間隔でサンプリングされた状態を表すキュー (**discrete-event queue**) とみなすことができる。
  - 海馬とMECの間のpentasyntacticループ回路
    - MECII-DG-CA3-CA1-Sb-ParaSb-MECII.

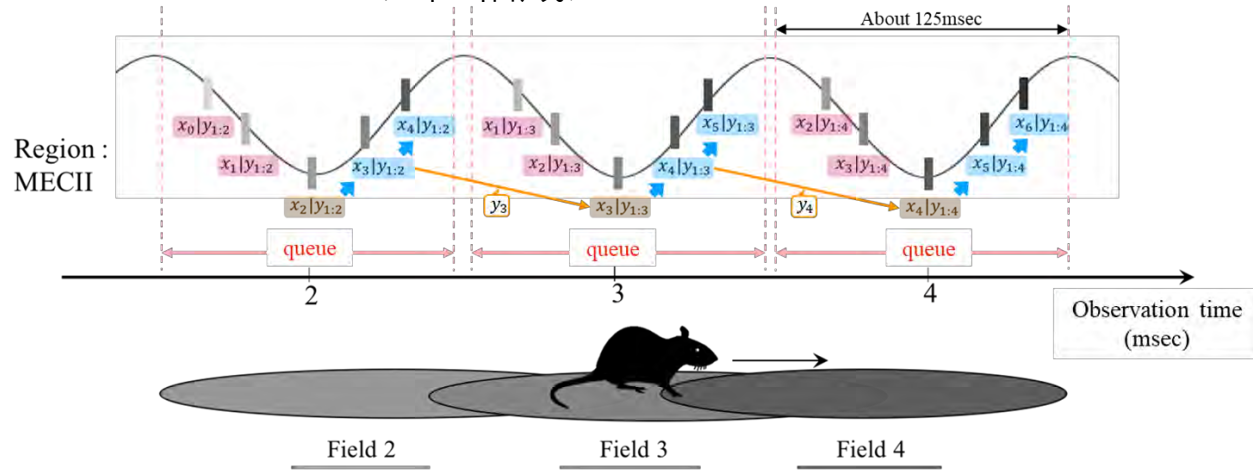


(a) Phase precession



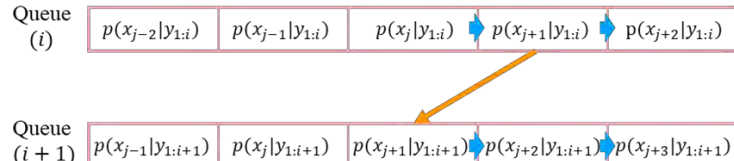
# Discrete-event queue-based Monte-Carlo localization (DEQ-MCL)

(a) Phase precession シータ位相歳差

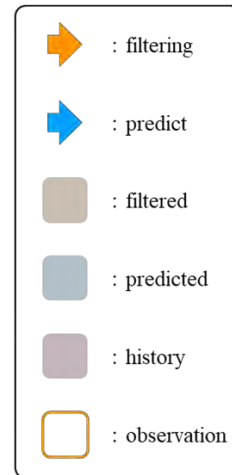


(b) Discrete-event queue 離散イベントキュー

$p(x_j | y_{1:i})$   
 $x_j$  : state  
 $y_{1:i}$  : observation  
 $i$  : observation time  
 $j$  : state time



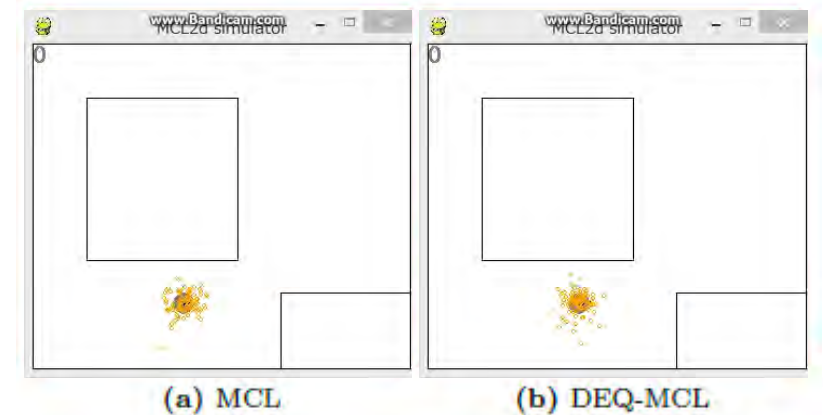
|                      |   | State time (j)   |                  |                  |                  |                  |
|----------------------|---|------------------|------------------|------------------|------------------|------------------|
|                      |   | 1                | 2                | 3                | 4                | 5                |
| Observation time (i) | 1 | $p(x_1 y_{1:1})$ | $p(x_2 y_{1:1})$ | $p(x_3 y_{1:1})$ |                  |                  |
|                      | 2 | $p(x_1 y_{1:2})$ | $p(x_2 y_{1:2})$ | $p(x_3 y_{1:2})$ | $p(x_4 y_{1:2})$ |                  |
|                      | 3 | $p(x_1 y_{1:3})$ | $p(x_2 y_{1:3})$ | $p(x_3 y_{1:3})$ | $p(x_4 y_{1:3})$ | $p(x_5 y_{1:3})$ |
|                      | 4 |                  | $p(x_2 y_{1:4})$ | $p(x_3 y_{1:4})$ | $p(x_4 y_{1:4})$ | $p(x_5 y_{1:4})$ |
|                      | 5 |                  |                  | $p(x_3 y_{1:5})$ | $p(x_4 y_{1:5})$ | $p(x_5 y_{1:5})$ |



## DEQ-MCLの定式化

$$\begin{aligned} \text{Queue}(t) &= p(x_{t-2:t+2} | y_{1:t}), \quad j \in \{i-2 \leq i \leq i+2\}, \quad i = t, \\ &= \underbrace{\eta p(x_{t+2} | x_{t+1})}_{\text{Prediction}} \underbrace{p(y_t | x_t)}_{\text{Filtering}} \underbrace{p(x_{t-2:t+1} | y_{1:t-1})}_{\int \text{Queue}(t-1) dx_{t-3}}, \quad t \geq 3. \end{aligned}$$

- 過去、現在、未来を含めた状態の同時事後分布をキューとして表現
- ベイズフィルタの拡張として表現できる
  - フィルタリングに平滑化と予測を組み合わせた逐次更新アルゴリズム



Akira Taniguchi, Ayako Fukawa, Hiroshi Yamakawa, "DEQ-MCL: Discrete-Event Queue-based Monte-Carlo Localization", AROB-ISBC-SWARM 2024. (Accepted)

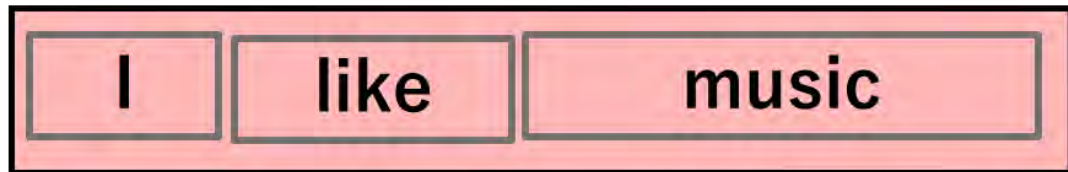


# 海馬体に学んだモデルのまとめ

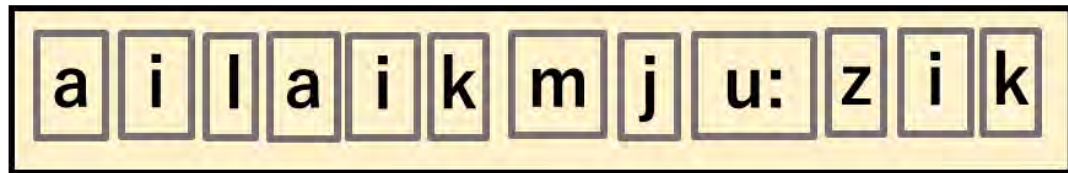
- 共に空間認知の機能を担う**SLAMと海馬体の知見の橋渡し**を目指した
- PGMベースのSLAMと海馬体の脳科学的知見をまとめた
- **海馬体の機能や解剖学的構造と整合性のとれたPGM**を構築した
  - 布川らが提案した海馬体のBIFをベースにLECを加えたBIFに発展させ議論
  - PGMベースのSLAMやSpCoSLAMを海馬体BIFの解剖学的構造に合うように拡張
- HPF-PGMの従来のSLAMとの主な違い
  - MECのみならずLECも統合してモデル化
  - Discrete-event queueの導入、定式化
- 今後の方向性
  - HF-PGMをWB-PGMのサブモジュールとして統合
  - 生成モデル・推論モデルの確率分布やアーキテクチャの詳細設定
  - 移動ロボットへの実装・導入

# 二重分節解析機能に関する領域と対応する機能

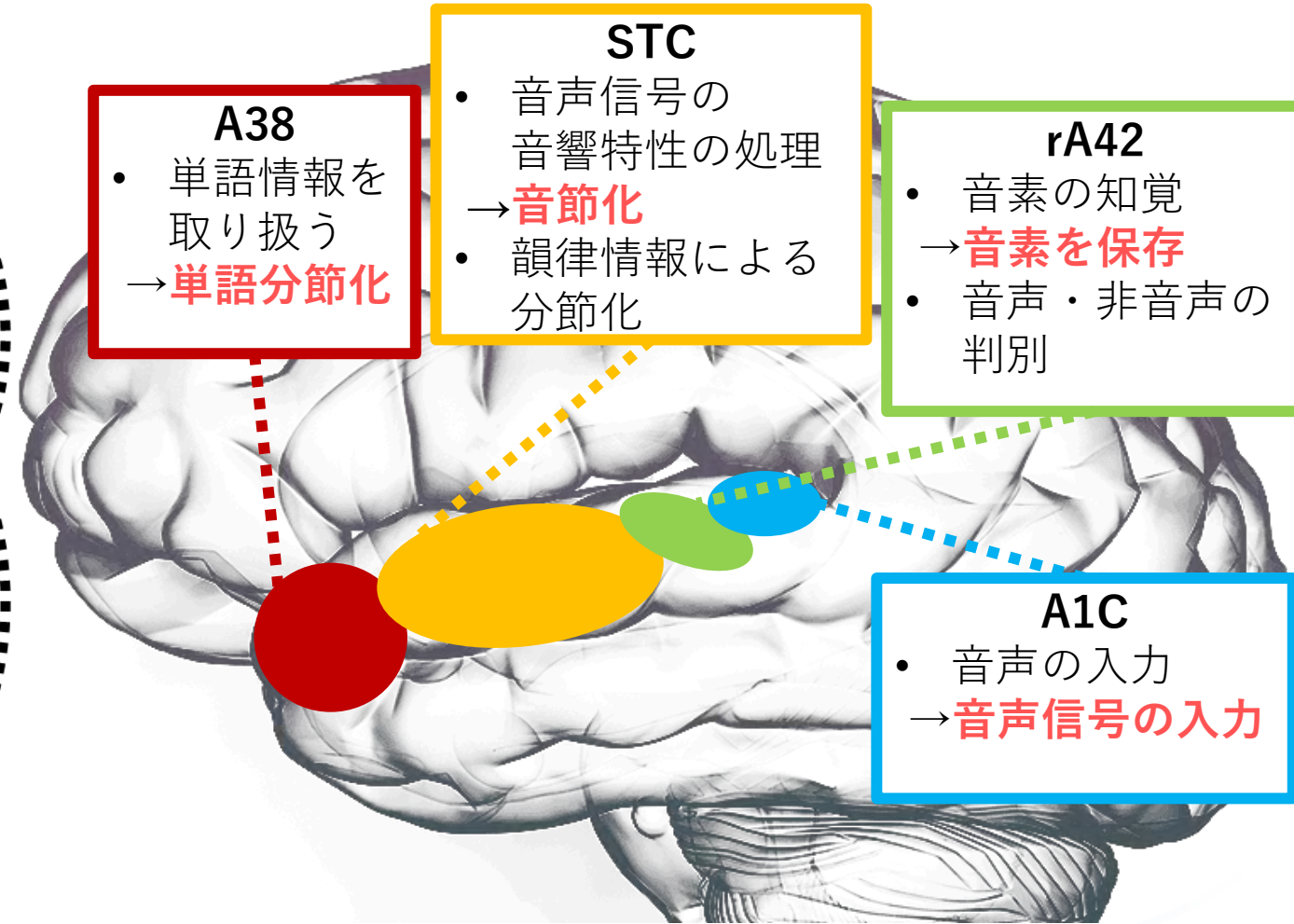
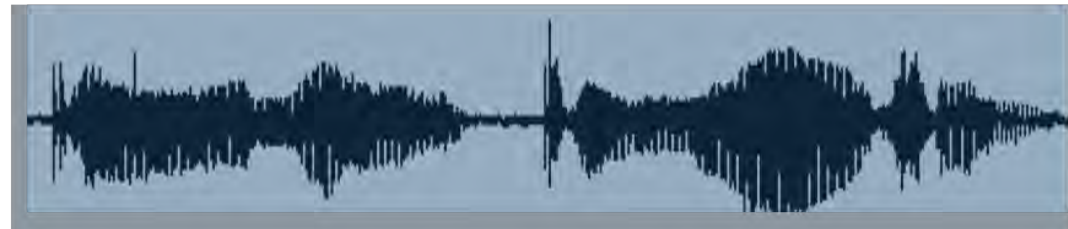
単語列



音素列



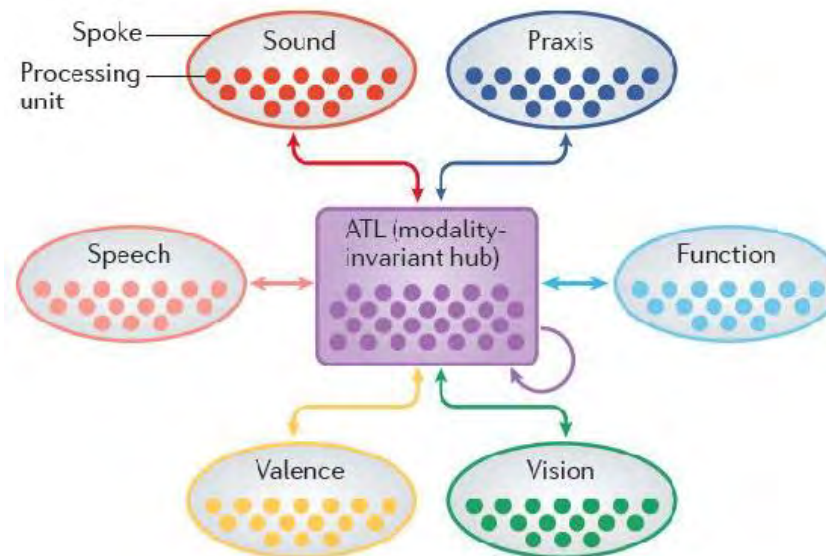
音声信号



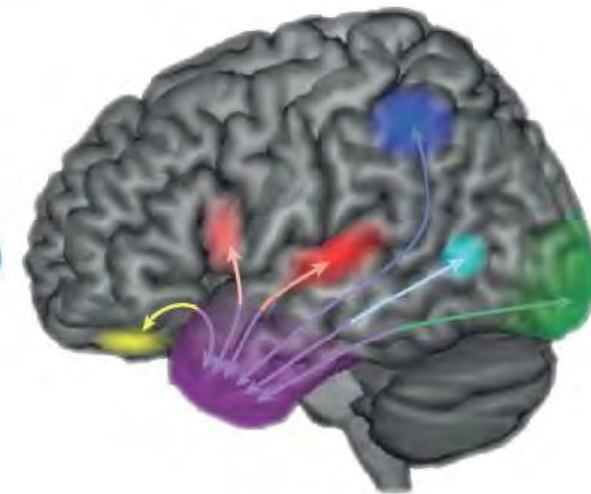
# モダリティの意味情報のハブ [6-16]

- ハブ&スポークモデル：  
言語および非言語的な経験による  
**概念を構築するための中核的な材料 (material)** となり、  
これらの情報源は**脳全体に分布している特異的な皮質 (=スポーク)** に  
対応していると仮定したモデル
- 全モダリティ固有な情報源の  
クロスモーダルな相互作用は、  
**the anterior temporal lobes (ATLs)** の両側に位置する  
**単一のトランスモーダルハブ**によって媒介されると提唱している

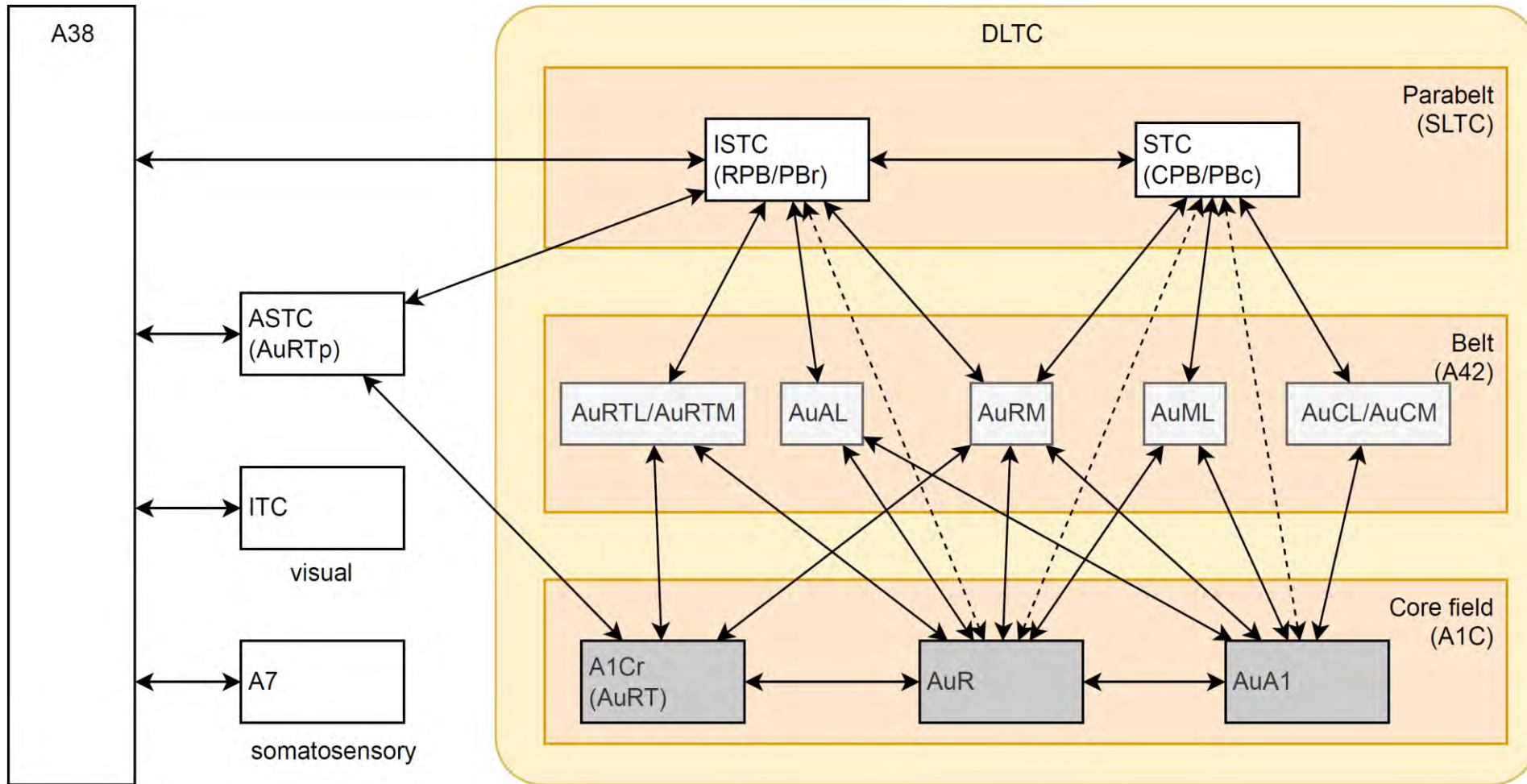
a Computational framework



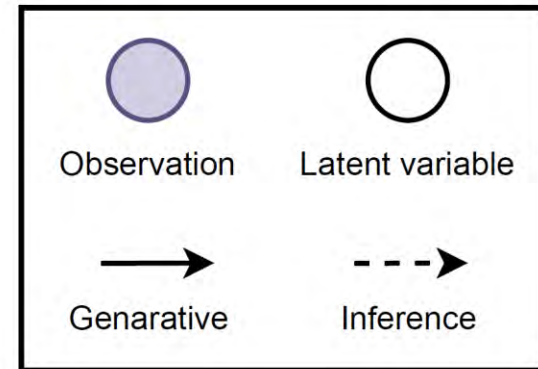
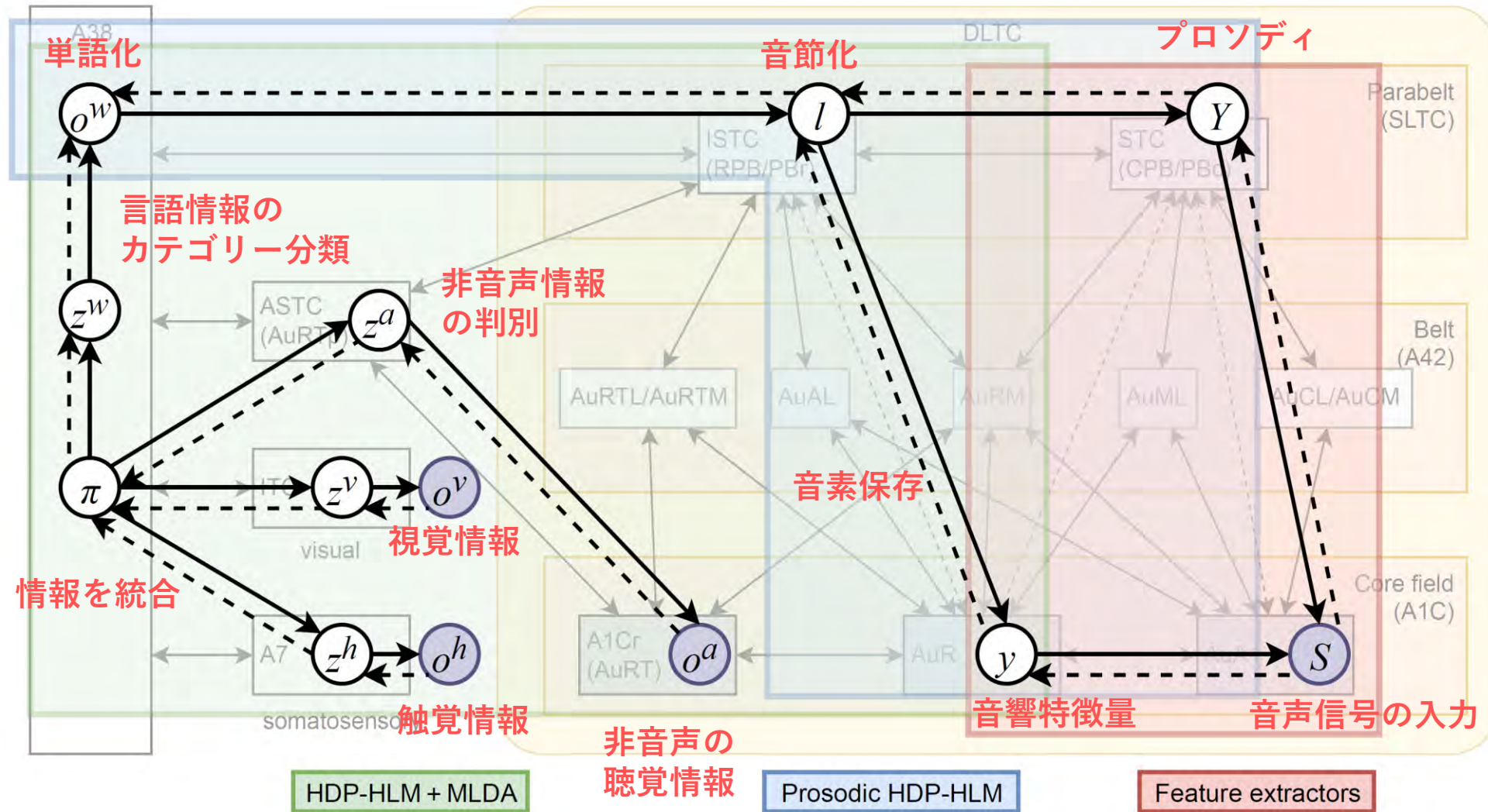
b Neuroanatomical sketch



# Brain information flow (BIF) for DAA



# BIFの上に配置した音声言語の二重分節解析のグラフィカルモデル

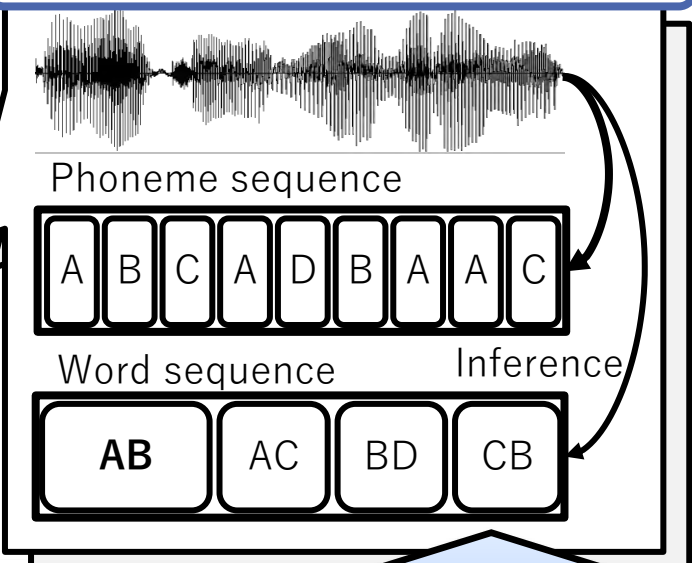


# ロボットのマルチモーダルな観測情報を活用した単語分割の実例

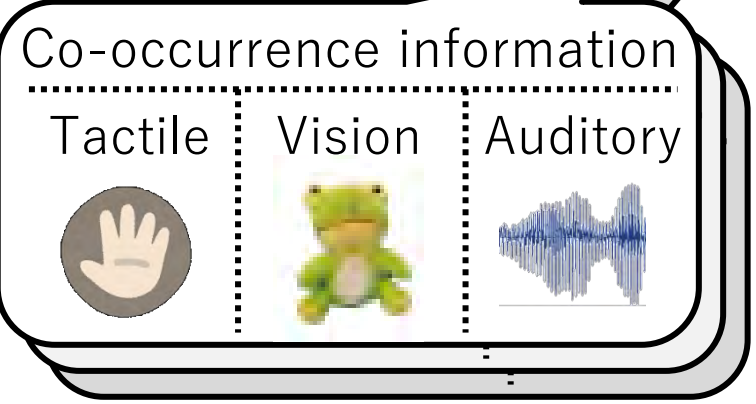
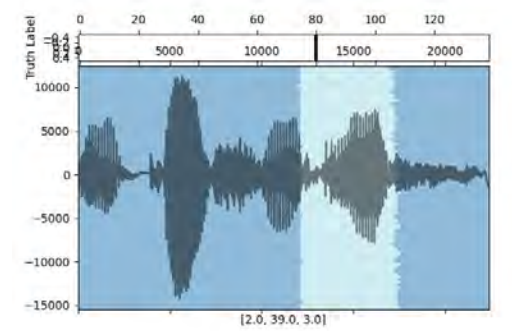
Speech regarding each object



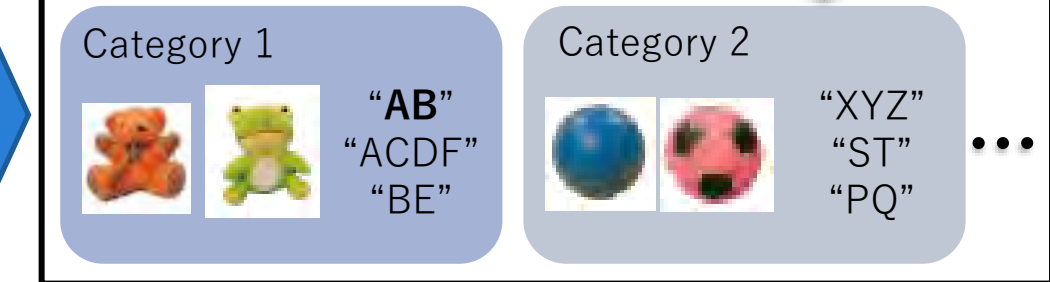
## Phoneme/word discovery



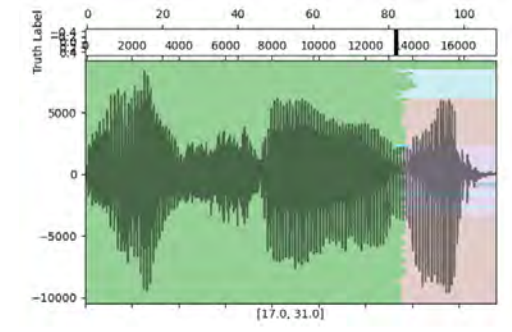
Ex.1: “おかし|です”



## Multimodal object categorization



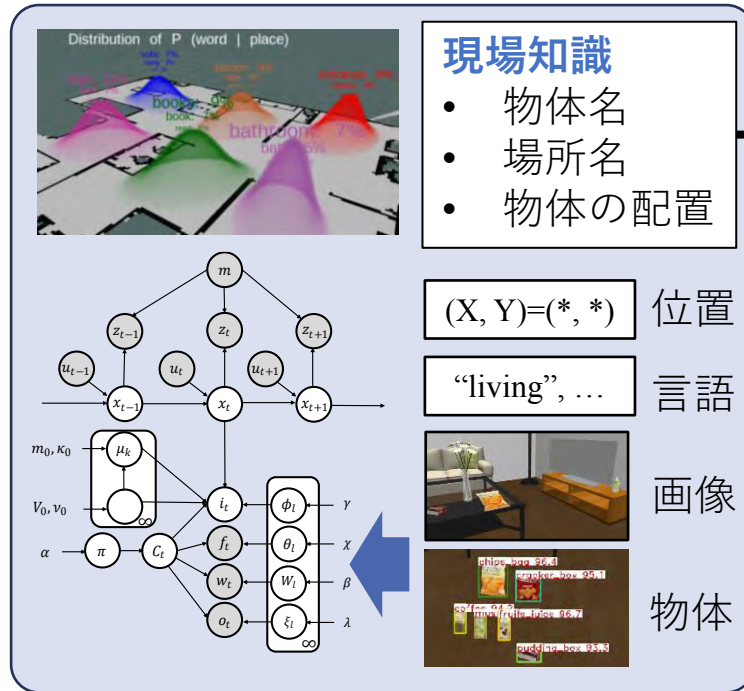
Ex.2: “おいしい|ね”



# 現在の課題と実現への道筋

# 現場知識を用いたGPTによるロボットの行動計画

## 場所概念モデルに基づく 現場知識の獲得と記述



- スキルセット**
- Navigation
  - Object Detection
  - Pick
  - Place

**言語命令**  
Bring a cup to the kitchen.

## GPTによる行動計画

```

You are an excellent household robot.

Please generate a series of behaviors to accomplish the user's
instructions, referring to the information below. Let's think step-
by-step.

List of location_name:
[living_room, kitchen, toy_room]

List of object_name:
[bottle, cup, stuffed_toy, banana, apple, orange, bowl]

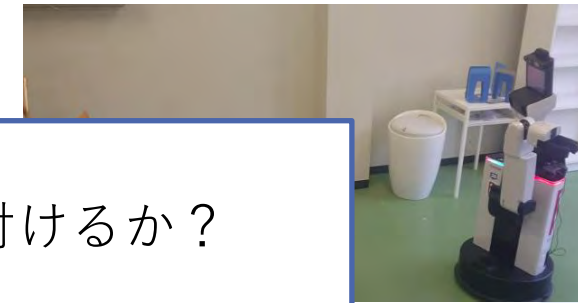
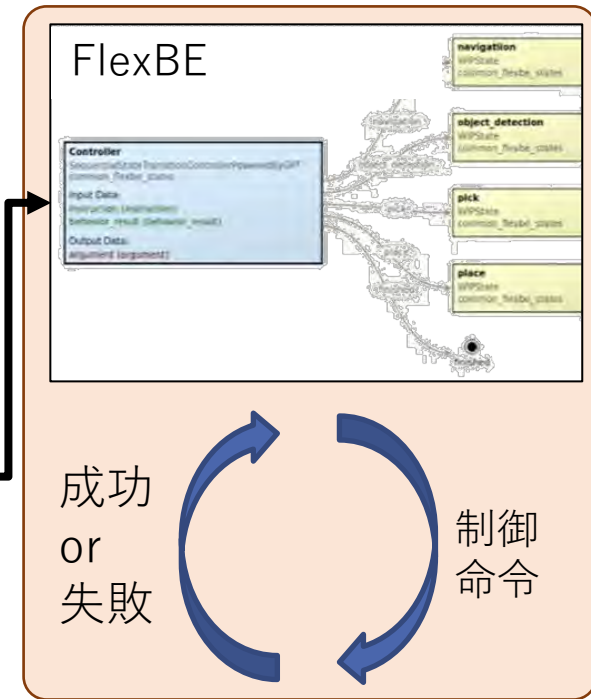
List of probabilities that an object exists at [living_room, kitchen,
toy_room]:
bottle: [0.8, 0.15, 0.05]
cup: [0.7, 0.2, 0.1]
stuffed_toy: [0.1, 0.05, 0.85]
banana: [0.15, 0.8, 0.05]
apple: [0.15, 0.8, 0.05]
orange: [0.15, 0.8, 0.05]
bowl: [0.15, 0.8, 0.05]

List of behaviors:
1. navigation (location_name): move to location_name
2. object_detection (object_name): detect an object_name and its
position from a captured image
3. pick (object_name): pick up an object_name
4. place (location_name): place an object to the location_name
These behaviors return "succeeded" or "failed". If "failed" is
returned, try the same or another behavior again.

USER : bring the cup to the kitchen
ASSISTANT : navigation (living_room)
USER : succeeded
ASSISTANT : object_detection (cup)
USER : succeeded
ASSISTANT : pick (cup)
USER : failed
ASSISTANT : pick (cup)
USER : succeeded
ASSISTANT : navigation (kitchen)
    
```

ASSISTANT: navigation (kitchen)  
ASSISTANT: place (kitchen)

## 逐次的な行動実行



LLMや基盤モデルをどのように脳と対応付けるか？

実世界における  
生活支援タスクの達成



# 自由エネルギー原理、能動的推論のロボット応用



- ロボットが「意味的に未知」な場所に移動しユーザに能動的に質問することで場所概念を獲得するSpCoAEを提案[Taniguchi 2022]
- 理論的には**期待自由エネルギー最小化における能動的推論**と接続性がある

$$a^* = \underset{a}{\operatorname{argmax}} (\operatorname{IG}(Z; X_a | X_{n_0}) - \eta \operatorname{TravelCost}(a))$$

$$G(\pi, \tau) = - \underbrace{E_{q(X_\tau | \pi)} [D_{\text{KL}} [q(Z_\tau | X_\tau, \pi) || q(Z_\tau | \pi)]]}_{\text{information gain}} - \underbrace{E_{q(X_\tau | \pi)} [\log p(X_\tau)]}_{\text{expected log-evidence}}$$

## Active SpCoSLAM (右図：実行例)

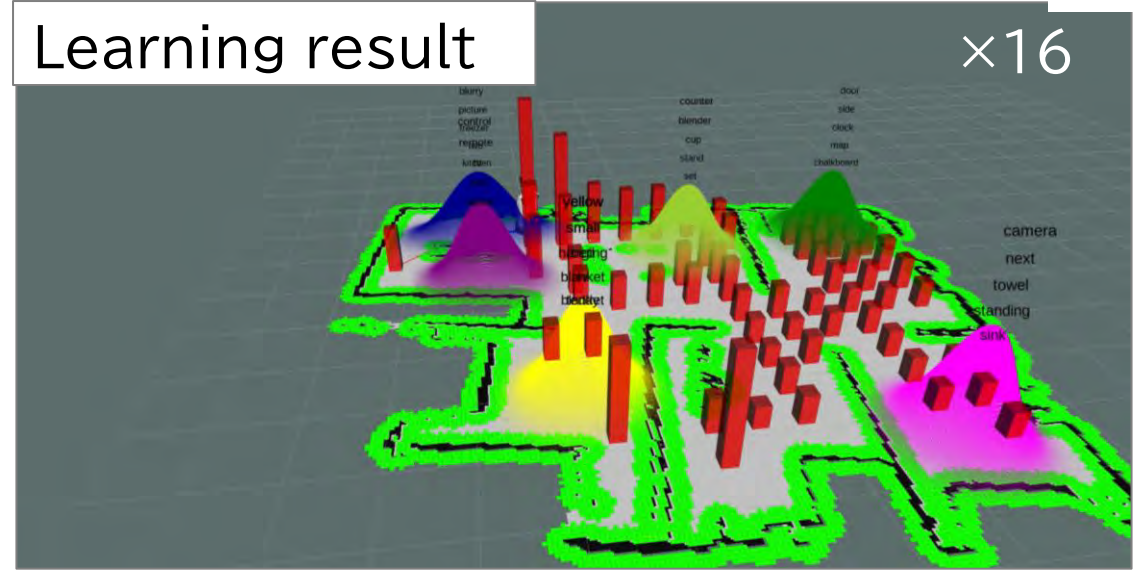
- SpCoSLAMをActive SLAMを融合し地図を作成しながら環境理解を進めるActive Semantic mapping手法

神経科学的な知見をモデルのみならず推論アルゴリズムとして表現する方向性も検討したい

能動的な探索による学習

Learning result

×16



Camera



Taniguchi, A., Tabuchi, Y., Ishikawa, T., Hafi, L. E., Hagiwara, Y., & Taniguchi, T. (2022). Active Exploration based on Information Gain by Particle Filter for Efficient Spatial Concept Formation. *arXiv preprint arXiv:2211.10934*.

Tomochika Ishikawa, Akira Taniguchi, Yoshinobu Hagiwara, Tadahiro Taniguchi, "Active Semantic Mapping for Household Robots: Rapid Indoor Adaptation and Reduced User Burden", IEEE International Conference on Systems, Man, and Cybernetics (SMC), Oct. 2023.

# 現在の課題と実現への道筋

- 個別モジュールの刷新
  - 深層生成モデルやTransformerベースのモデルを採用していく
- SERKETの理論的发展とSERKET-SDEの拡張
  - 確率的生成モデル以外の多様なモデルも取り込む
  - 対照学習や変分推論などの多様な推論アルゴリズムへの対応
  - センサデータやモジュール間の計算の同期・非同期
  - 各モジュールを容易に組み替えながら、より適したモデルを比較検討できるようにする
  - モデルとROSがリアルタイムに通信し逐次的な学習や推論を可能にする
- WBAIのロードマップとのリンク
  - 構築される全脳参照アーキテクチャを参考に実世界でのロボットに実装・検証へとつなげる