大阪大学 OSAKA UNIVERSITY
国立研究開発法人 科学技術振興機構 Japan Science and Technology Agency
Human-Machine Harmonious Collaboration
UEC TOKYO
iSYSLab Intelligent Systems Laboratory
Robot Learning Group
AIRC 人工知能先端研究センター Artificial Intelligence eXploration Research Center
EMOTION Emotion Model For Communication
ChiCaRo
COGNITIVE INTERACTION DESIGN

AI × Robotics:

Towards Human-like General Intelligence

2019.12.19 @JST-NSF-DATAIA Symposium

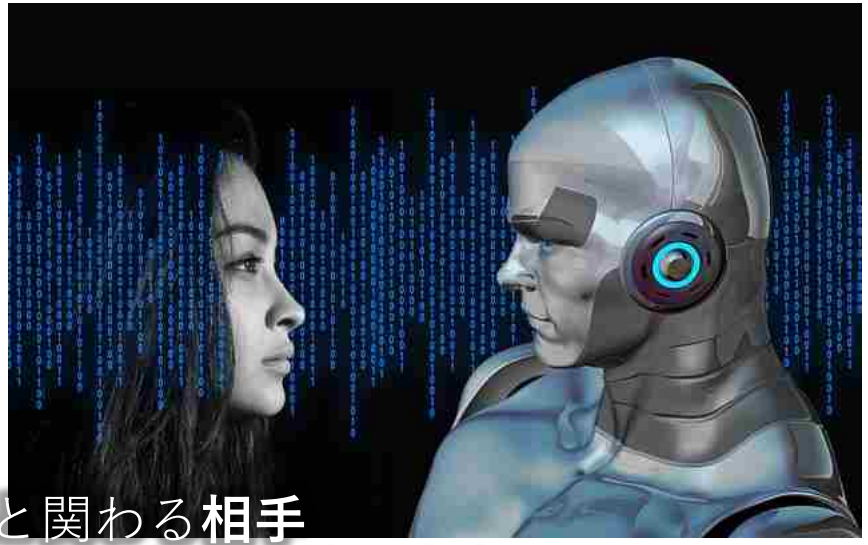
長井 隆行
大阪大学基礎工学研究科
電気通信大学AIX
Takayuki Nagai
Osaka University
AIX, University of Electro-Communications

Future of AI/Robots where are we heading?
AI/ロボットのミライ

- tools or partner
- 道具かパートナーか？

単なる道具
意味など理解する必要ない
理解するのは人間
大量のデータと計算パワーがあればOK

AI as a tool
No need to understand like humans
Humans understand the results
Big data and huge computational power

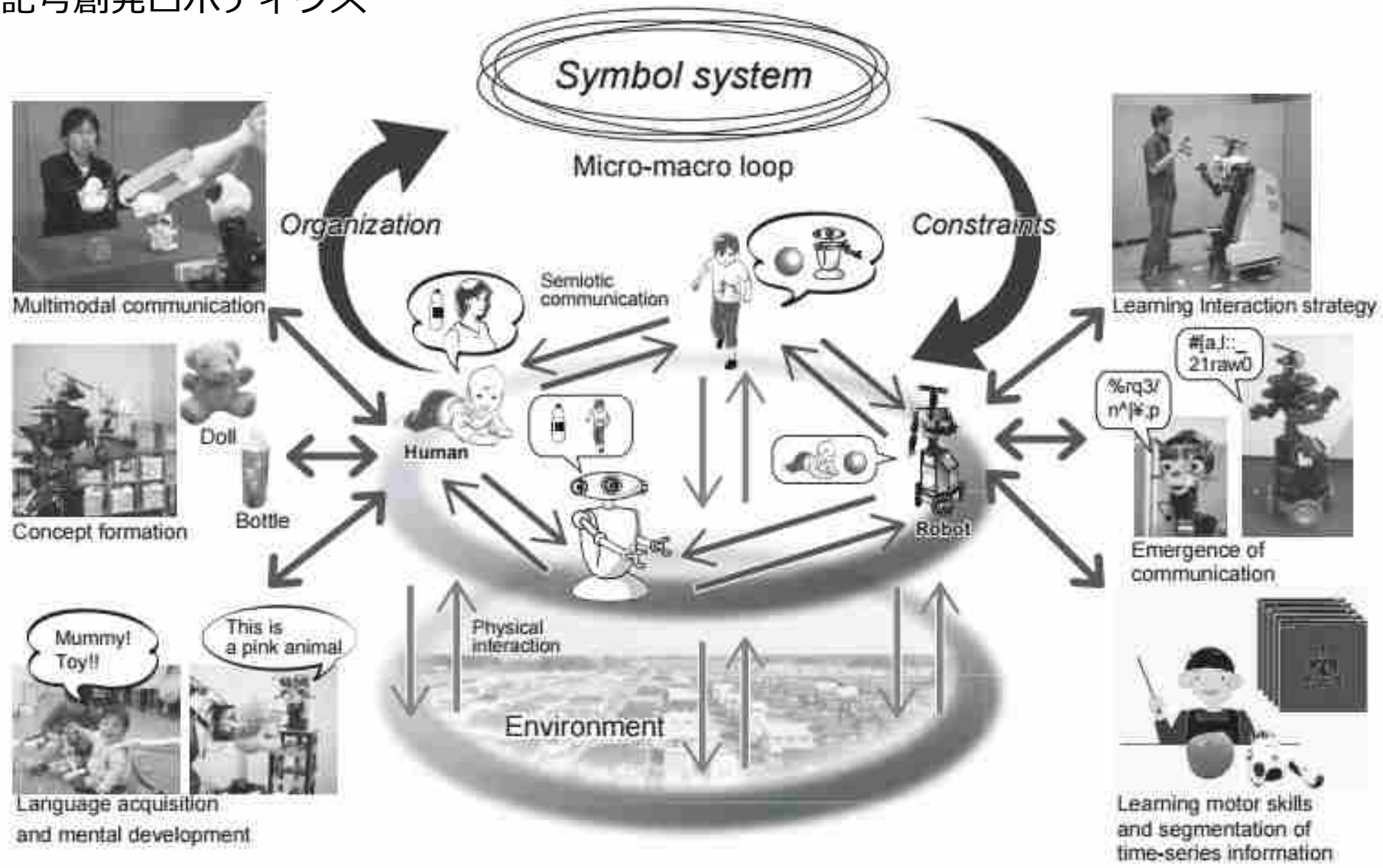


人と関わる**相手**
意味を**人間のように**理解する
大量のデータと計算パワーとは異なる問題
⇒意味の理解や相手の理解の本質は？

AI as a partner
Need to understand like humans
Different problem with data and comutations
⇒ What is understanding ?

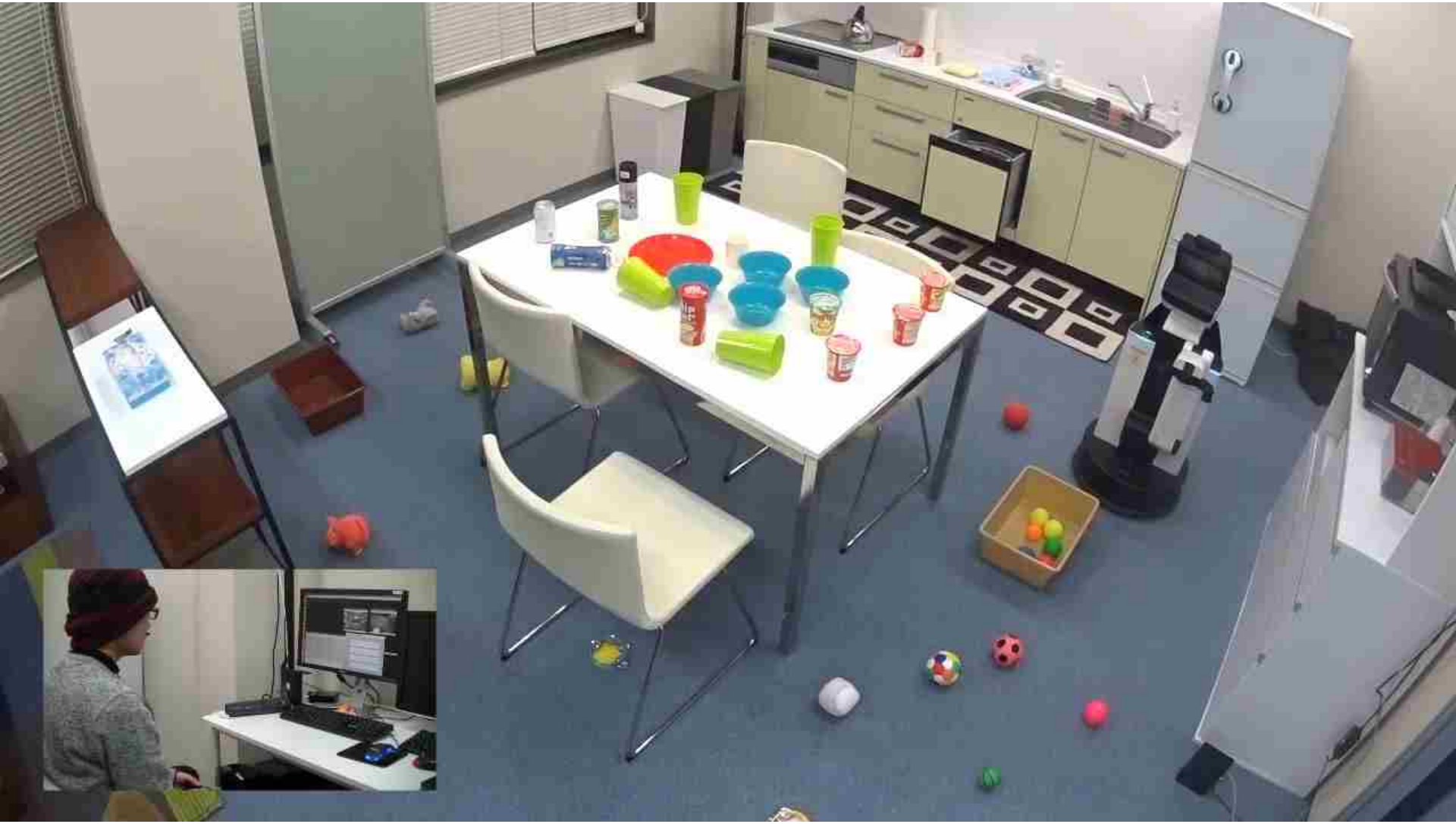
Symbol Emergence in Robotics

記号創発ロボティクス



http://arxiv.org/pdf/1509.08973v1.pdf

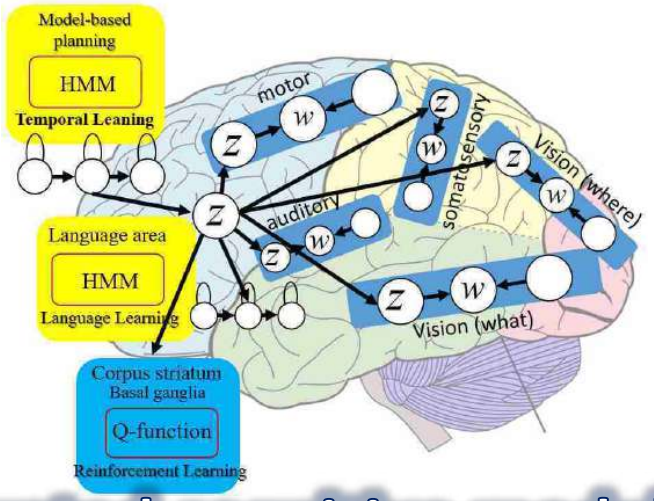
Example of home service robot
家庭用ロボットの例



Integrate as a whole!
全体を統合する！

身体
body
感情
affect
記号
symbol
論理推論
logical inference
プランニング
planning
反実仮想
counterfactual
意識
consciousness

身体制御(motor control)
言語(language)
推論(inference)
因果(causality)



Integrated cognitive model

統合認知モデル

- How can we integrate everything?

What is “Intelligence”?

知能とは何か？

Efforts

to see what cannot be seen

見えないものを見る努力

Prediction

予測

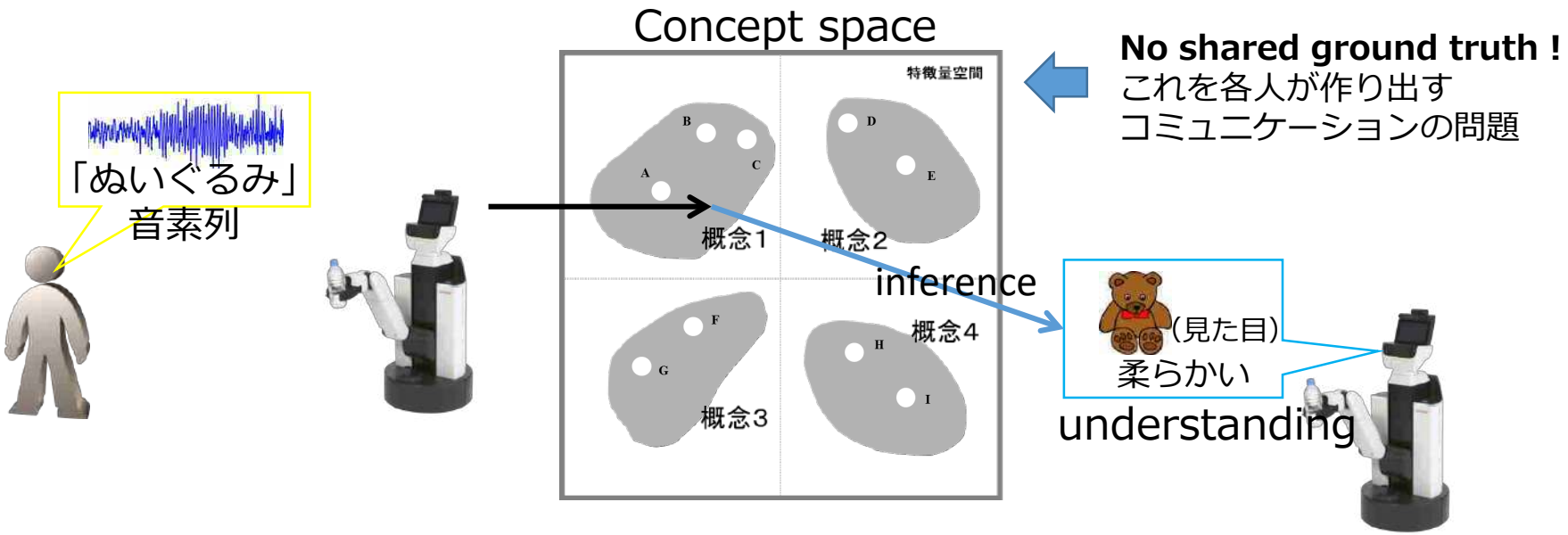
Language an aspect of intelligence

- ✓**ここにはないもの**について語ることができる
- ✓言葉の意味とは？
  - 正しい意味などどこにもない
- ✓言葉を理解するとは？
  - 予測という視点でとらえることができる
- ✓その言葉を聞いたときに予測したもの
  - それがその言葉の**その人**にとっての**その時**の意味
- **✓ロボットにできない理由はない！**

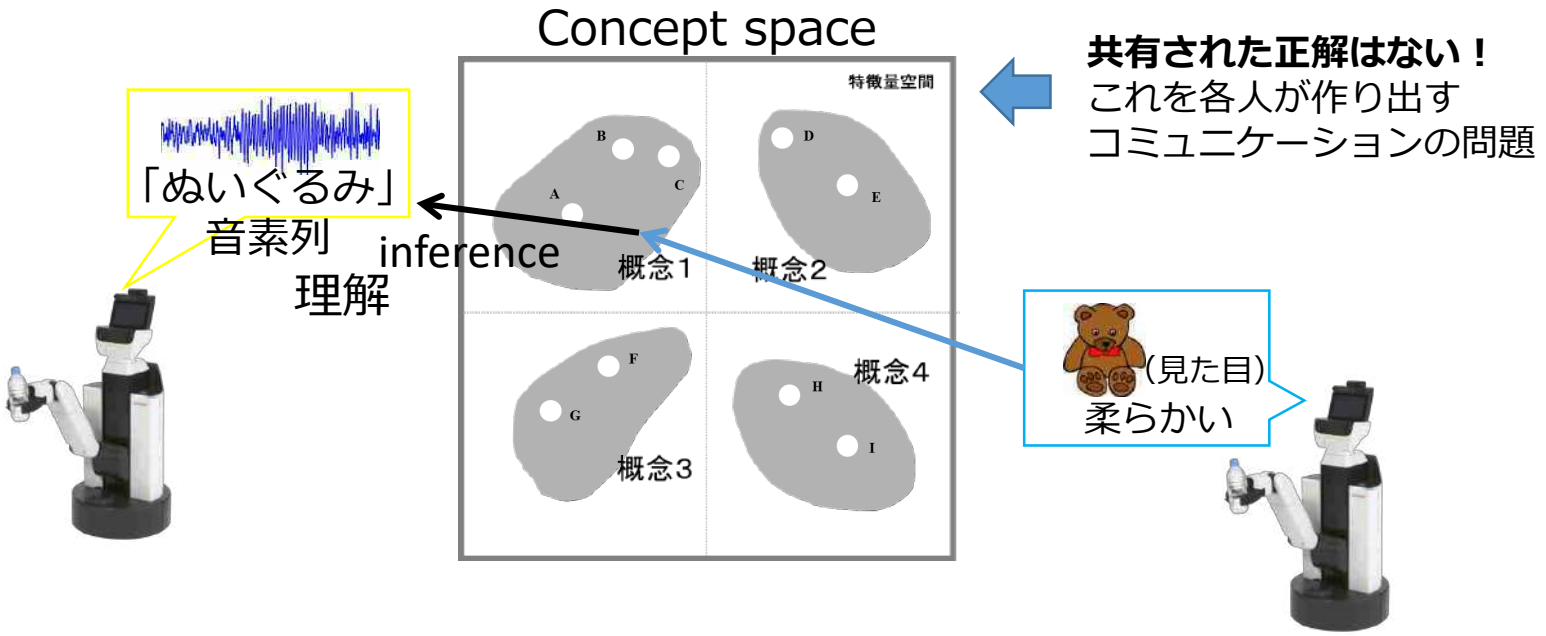
We can talk about something even doesn't exist
What is the correct meaning?
There must be "NO" correct meaning
What is understanding?
"Prediction" can answer this question
Something that comes to mind
That is what it "means" for the person at the moment
I'm sure robots can do this!

What is “understanding”? Naïve view

- ロボットによる実世界理解
  - **“理解”**：概念を通した未観測情報の予測
  - **“意味”**：予測した内容
  - **“概念”**：経験（マルチモーダルデータ）の
    カテゴリ分類によって形成される
- Symbol grounding problem



The other way round

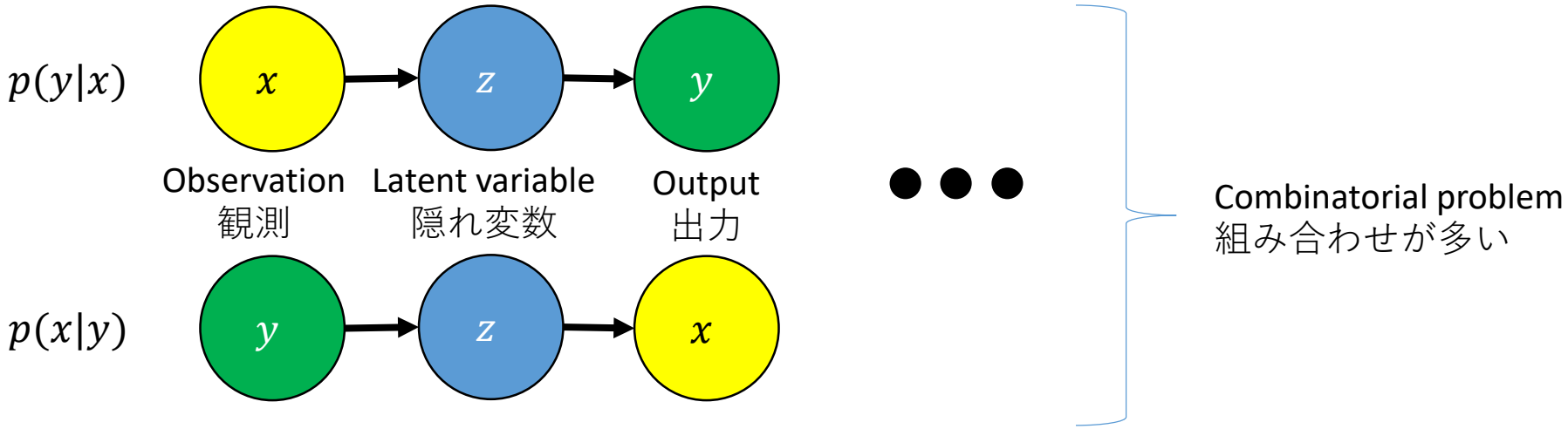


- 単一の識別タスク $p(y|x)$をモデル化するだけではダメ
- 同時分布をモデル化する

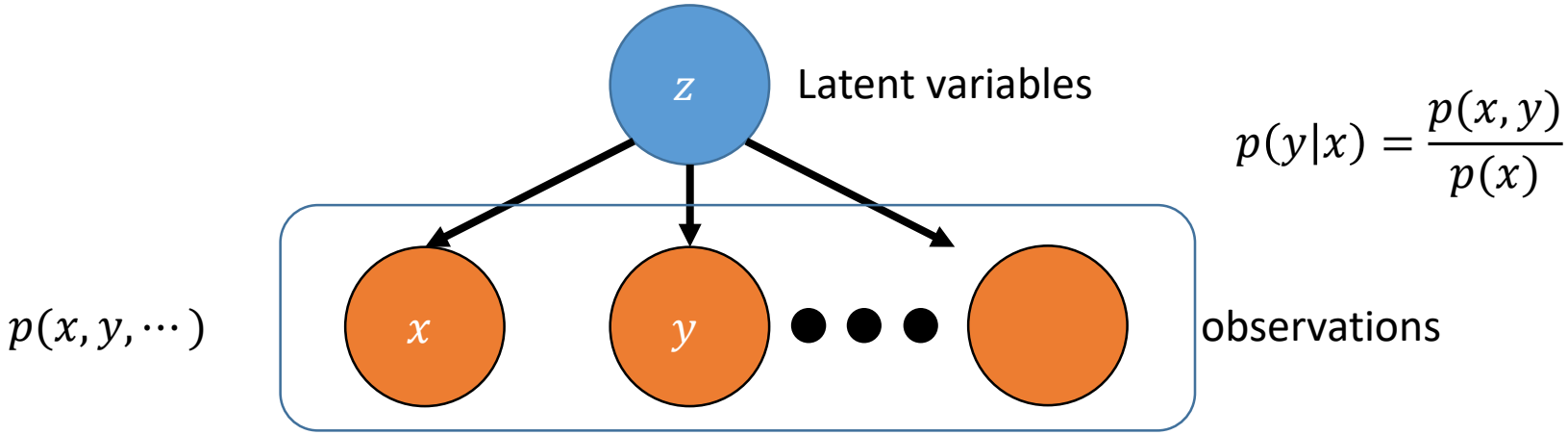
$$p(x, y, \cdots) = \sum_{z} p(x, y, \cdots | z) p(z)$$

Multimodal Generative Model (unsupervised learning)

Multimodal supervised learning（discriminative model)

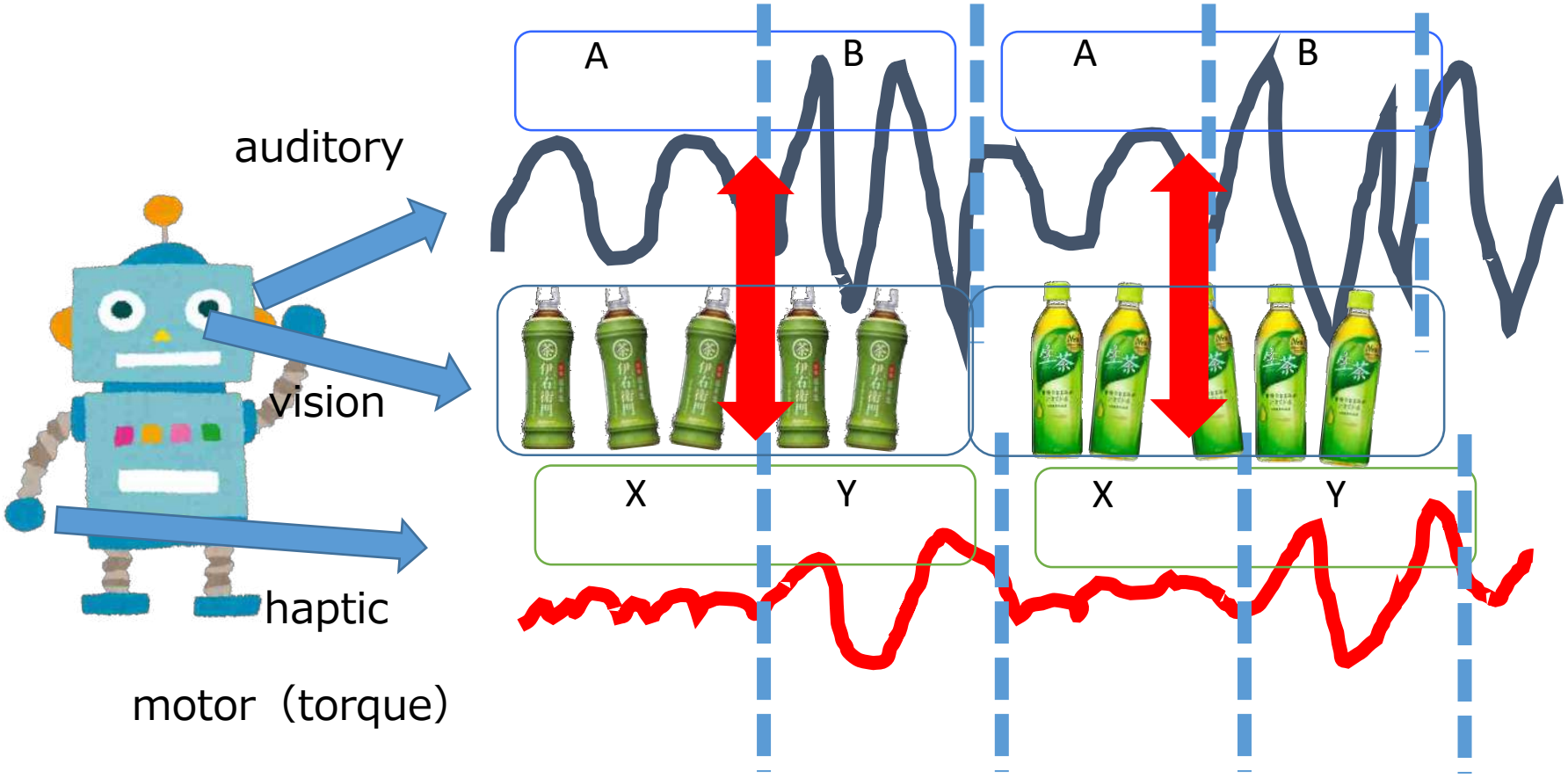


Multimodal unsupervised learning（generative model)



Idea

- Segmentation and categorization
- Gaussian process and hidden Markov models

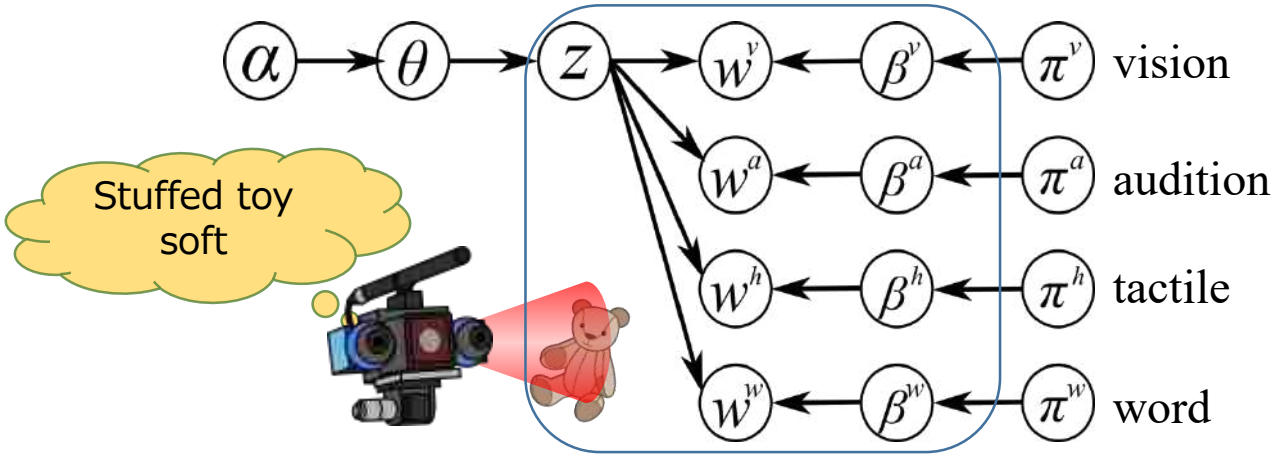


Multimodal categorization

Categorization of multimodal data

- **マルチモーダルカテゴリゼーション**
- **Multimodal Latent Dirichlet Allocation**（MLDA, MHDP, … ）

[Nakamura+ 09]



α : Dirichlet prior

θ : multinomial parameters

z : categories

w^*: multimodal infromation

β^*: multinomial parameters

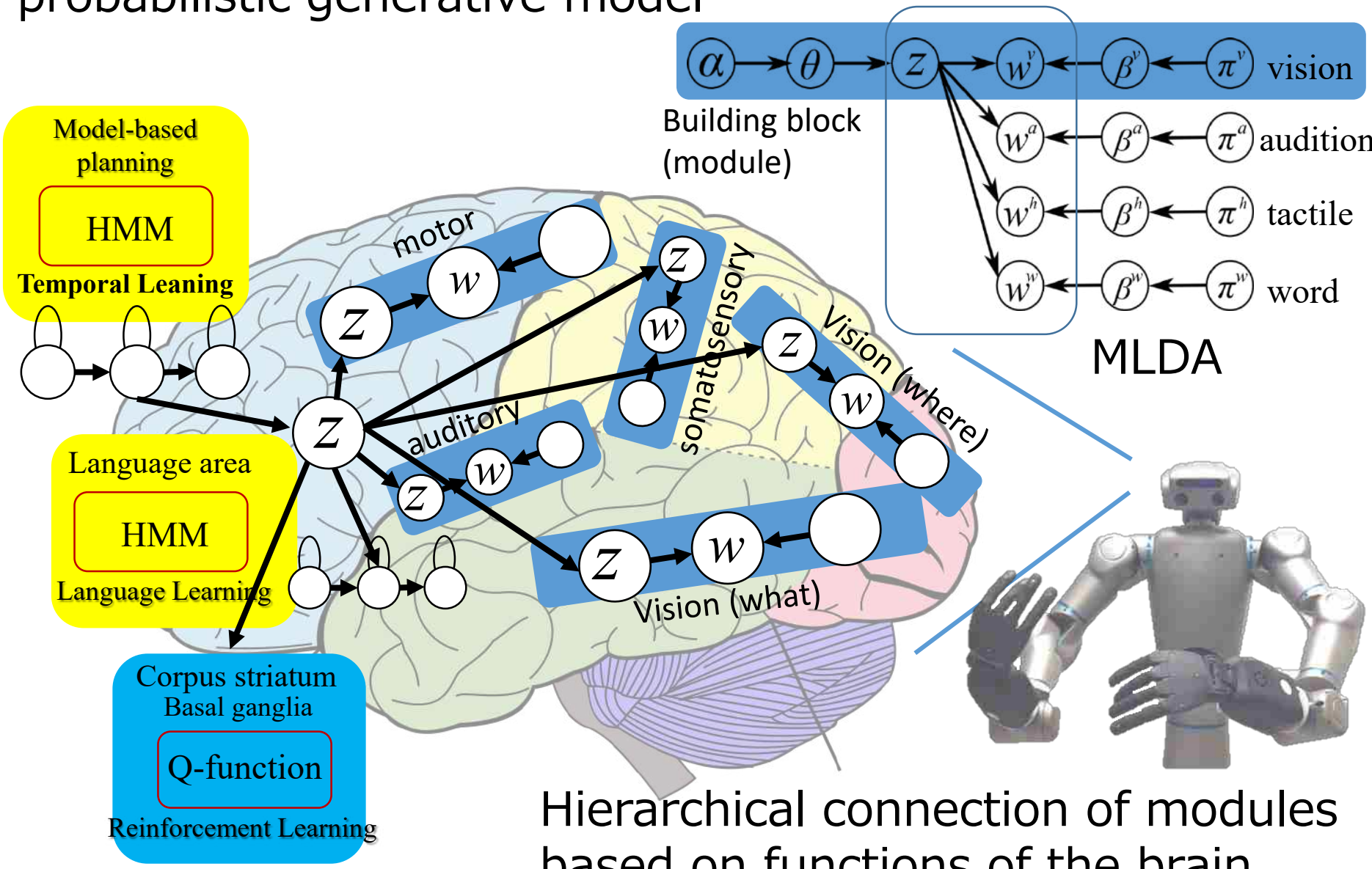
π^*: Dirichlet prior

Inference of the parameters β^* and θ by Gibbs Sampling

[Nakamura + 09] Nakamura,T. *et al.*, Grounding of word meanings in multimodal concepts using LDA, in Proc. IROS2009, pp.3943–3948, 2009

Integrated cognitive model

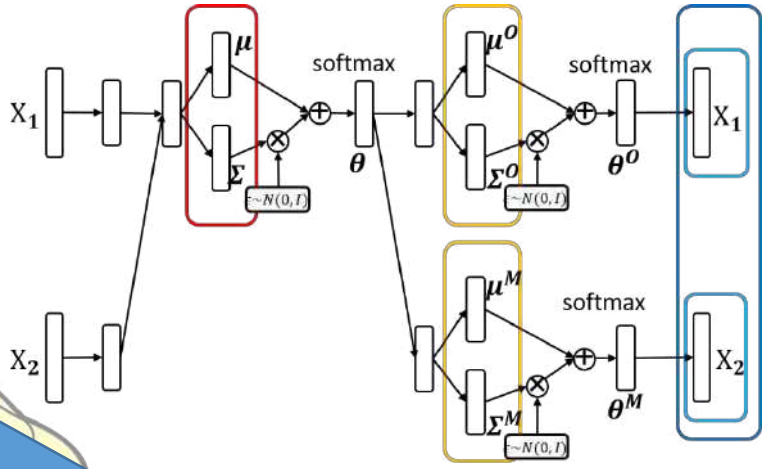
probabilistic generative model



Hierarchical connection of modules based on functions of the brain

Integrated cognitive model

deep generative model



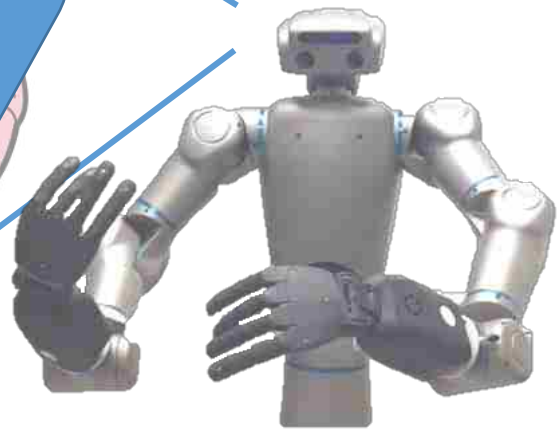
Deep mMLDA

LSTM
(temporal learning)

Representation Learning
(State Space)

Latent Variables

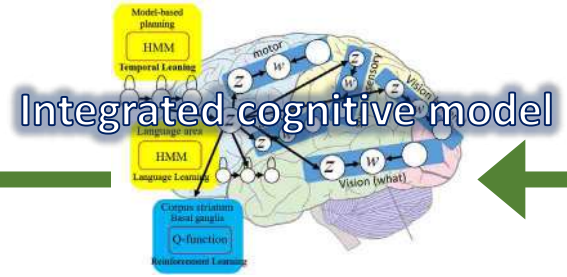
(Deep) Q Network
RL



Hierarchical connection of modules
based on functions of the brain

Implementation on a real robot

ROS



Integrated cognitive model

organization

decision-making

sensors

actuators

camera

microphone

arm

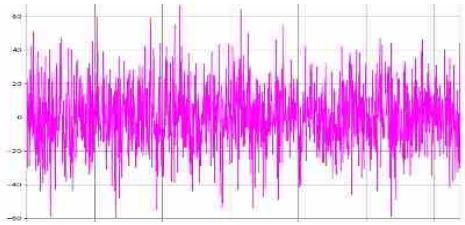
encoders

hand

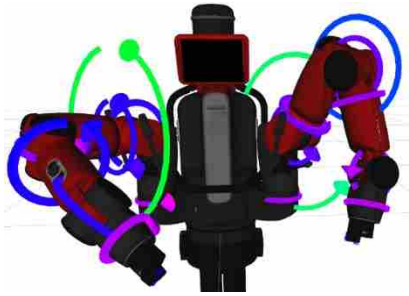
Haptic sensor

data

image



sound



proprioception

Tactile sensor

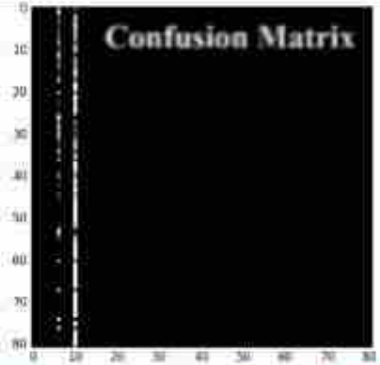
Data from various sensors

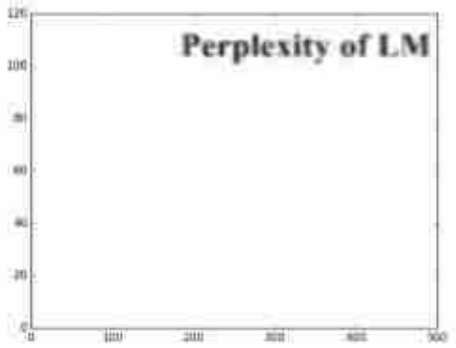
Long-term learning experiment [Nishihara+ 16]

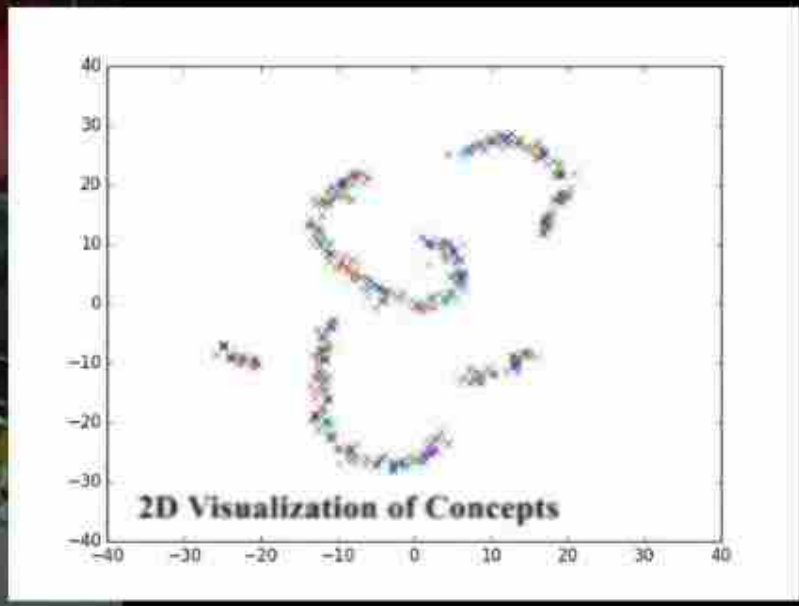
- 3 ~ 5 時間/日 × 1か月 (100 時間以上)
- 500 物体 (81 カテゴリ) を使用
- Over 100 hours of interaction
- Using 500 objects (81 categories)

Long-term learning experiment [Nishihara+ 16]

- 3000以上の発話（300ユニーク単語）100単語程度獲得
- 約7割程度の正解率

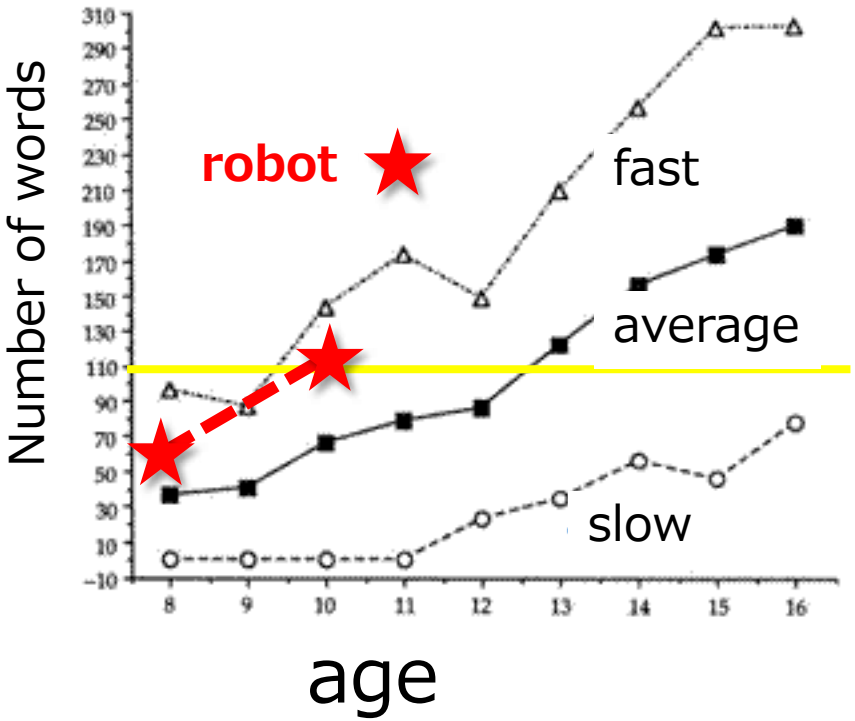






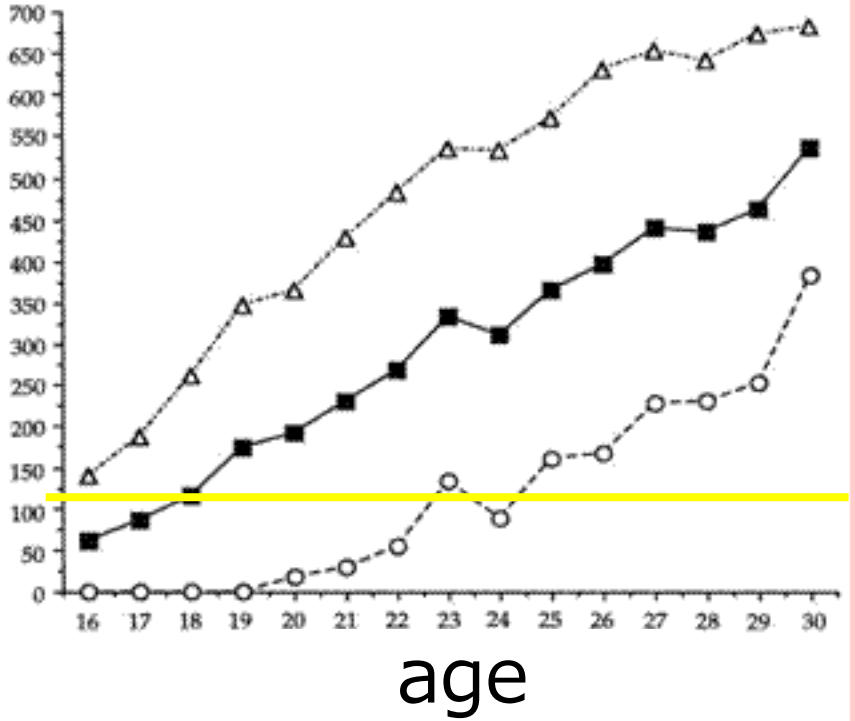
Vocabulary development

understanding



About 180words in 16-month

production



About 60words in 16-month

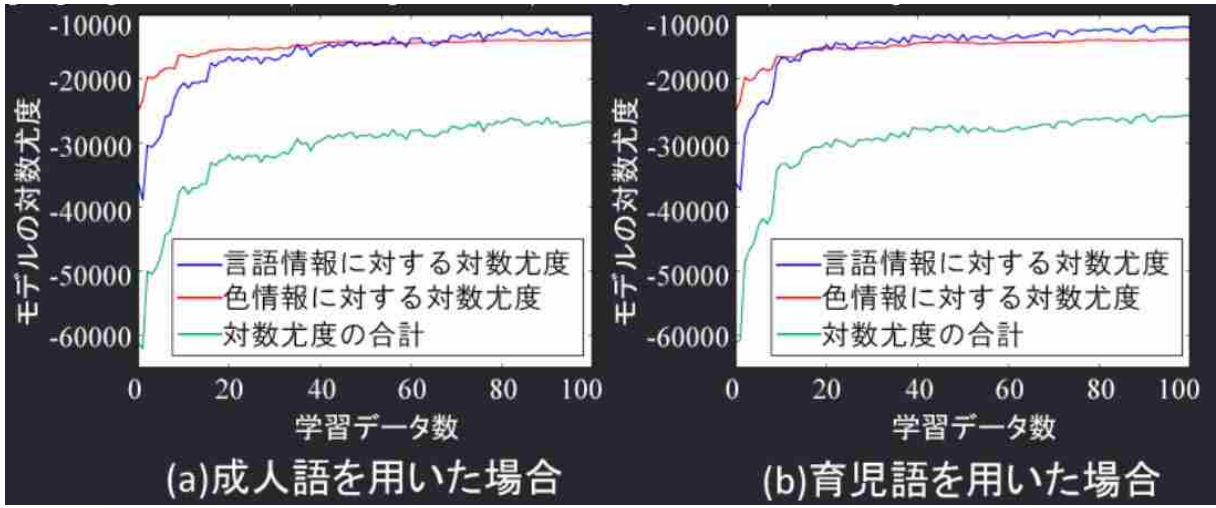
About 1000words in 3 years old

Bates et al, 1995

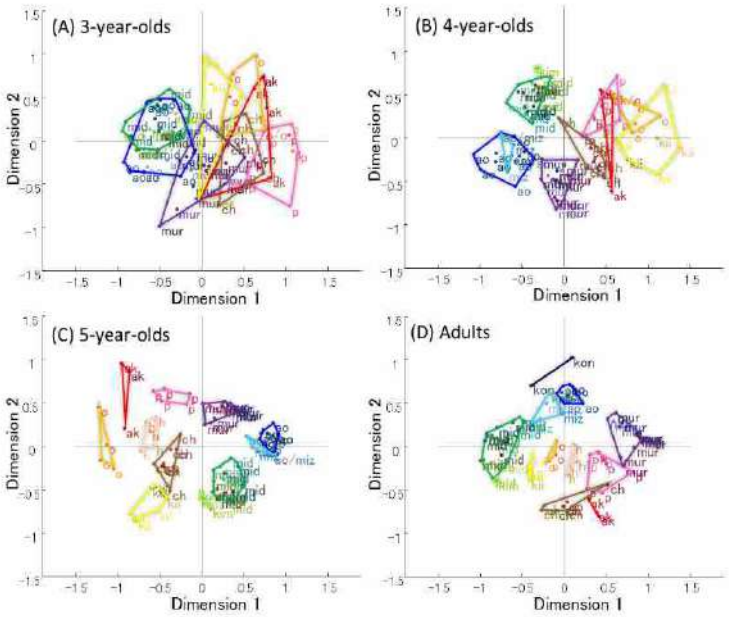
Findings

Suggestions obtained by analyzing robot learning

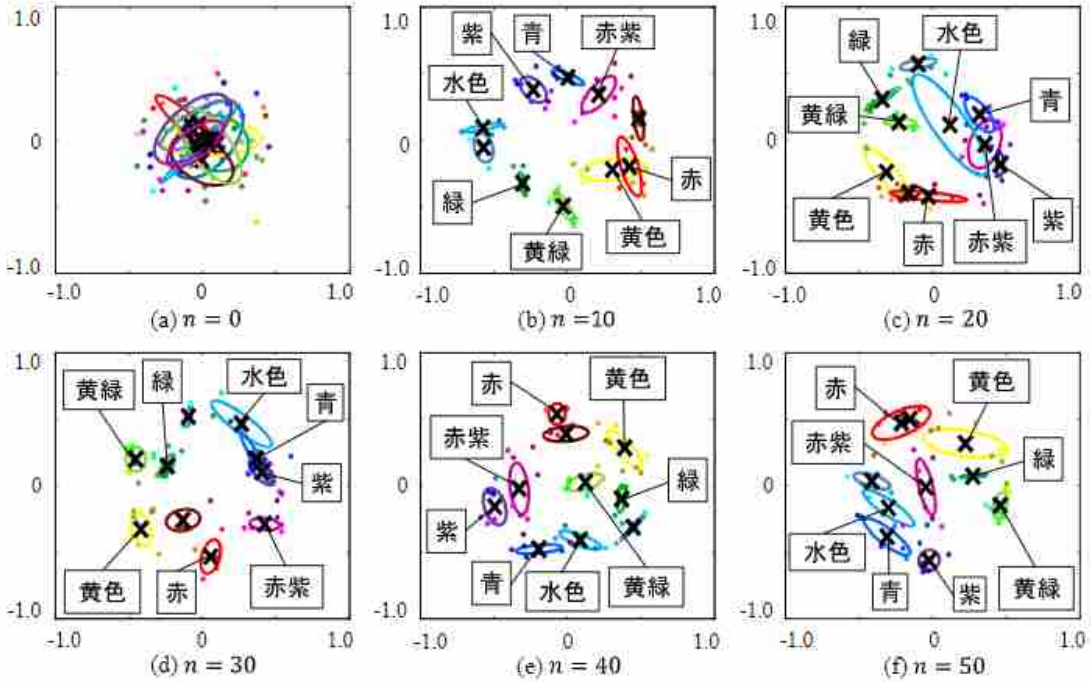
- ✓ The role of child directed speech[Funada+ 17]
- ✓ Learning of color [Funada+ 16]



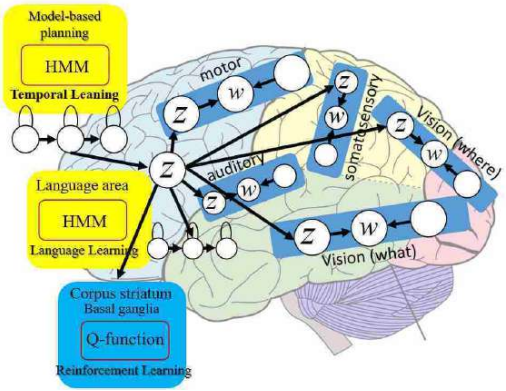
M.Imai, *et.al* "Meanings of symbols emerge within a system" JSAI 2015

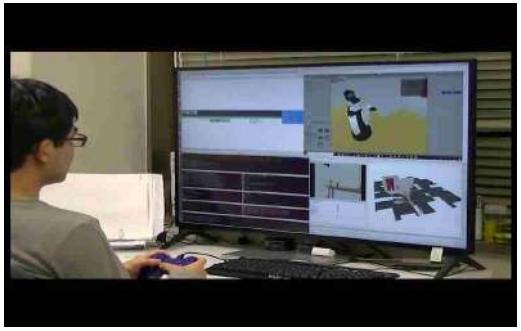


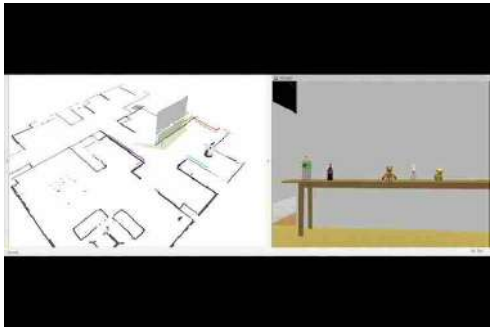
ak = 赤, ao = 青, ch =茶色, h = 肌色, kii = 黄色, kim = 黄緑 mid =緑, miz =水色, mur =紫, o =オレンジ色, p =ピンク色



Learning from teleoperation for domestic service robot







Iwata *et al.* "Learning and generation of actions from teleoperation for domestic service robots," IROS 2018

Miyazawa *et al.* "Integrated cognitive architecture for robot learning of action and language," Frontiers in Robotics and AI, 2019

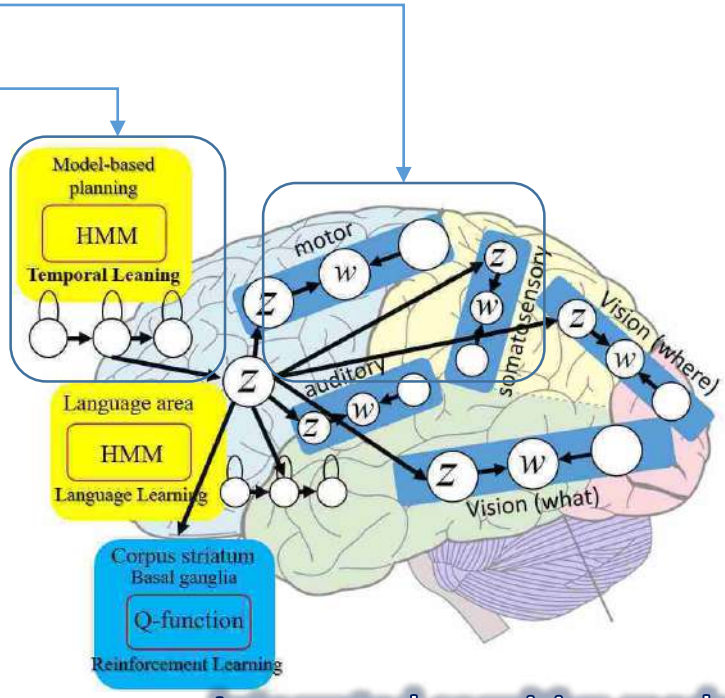
Currently working on experiments @ three different real home environments

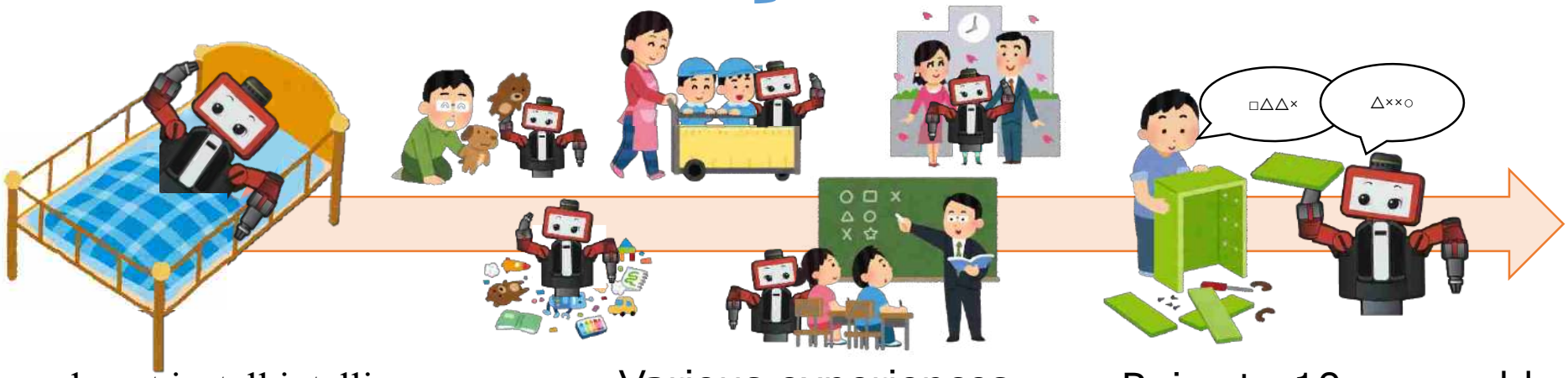
Three different real home environments

After learning (autonomous task execution)

Challenges

- Communication（sociality）
  - Self/other discrimination
  - Mirror neuron system
- Emotion/affection
  - Information on body（**interoception**）
  - Rewards
- Higher level cognitive functions
  - Induction, deduction, abduction
  - Causality
  - Counterfactual
  - Explainability
- Consciousness





> sudo apt install intelligence

Various experiences

Raise to 10 years old