# Symbol Emergence in Robotics: Pursuing Integrative Cognitive Architecture using Probabilistic Generative Models for Real-world Language Acquisition

Tadahiro Taniguchi

1) Professor, College of Information Science & Engineering, Ritsumeikan University

2) Visiting General Chief Scientist, Technology Division, Panasonic

International Symposium on Artificial Intelligence and Brain Science

11th October, 2020, online, Japan

# Tadahiro Taniguchi @tanichu


- Professor, Emergent System Laboratory, College of Information Science and Engineering, Ritsumeikan University, Japan
  - 2003-2006: PhD student, Kyoto University
  - 2005-2008: JSPS research fellow, Kyoto University
  - 2008: Assistant professor, Ritsumeikan University
  - 2010: Associate professor, Ritsumeikan University
  - 2015-2016: Visiting Associate Professor, Imperial College London
  - **2017: Professor, Ritsumeikan University**
  - **2017: Visiting General Chief Scientist, Panasonic Corporation**
- **Research Topics**
  - Machine learning, Cognitive & Developmental robotics, Symbol emergence in robotics, Language acquisition

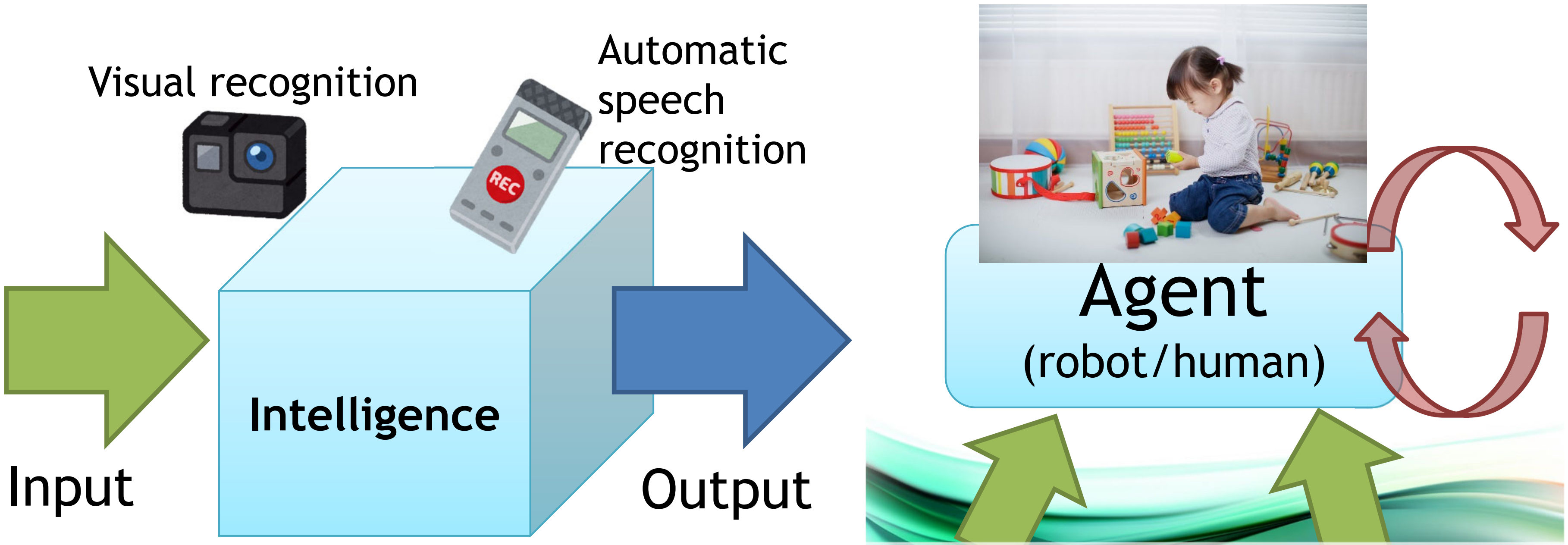
# Computational Understanding of Mental Development
## From Behavioral Learning to Language Acquisition

- Human children acquire many physical skills, concepts, and knowledge, including **language**, through physical and social interaction with their environment.
- We'd like to obtain an understanding of the **computational process** of mental development and language acquisition.

**Constructive approach**
Develop robotic and computational models to better understand the original

## Developmental Robotics /Symbol Emergence in Robotics

# Intelligence as a **cognitive dynamics**




- ☐ Single-purpose
- ☐ Task-oriented
- ☐ Intelligence as a procedure
- ☐ No autonomy

- ☐ Not single task-oriented
- ☐ Self-organization of sensorimotor information flow
- ☐ Autonomy

Intelligence is an existence emerging through self-organization of sensorimotor information flow.

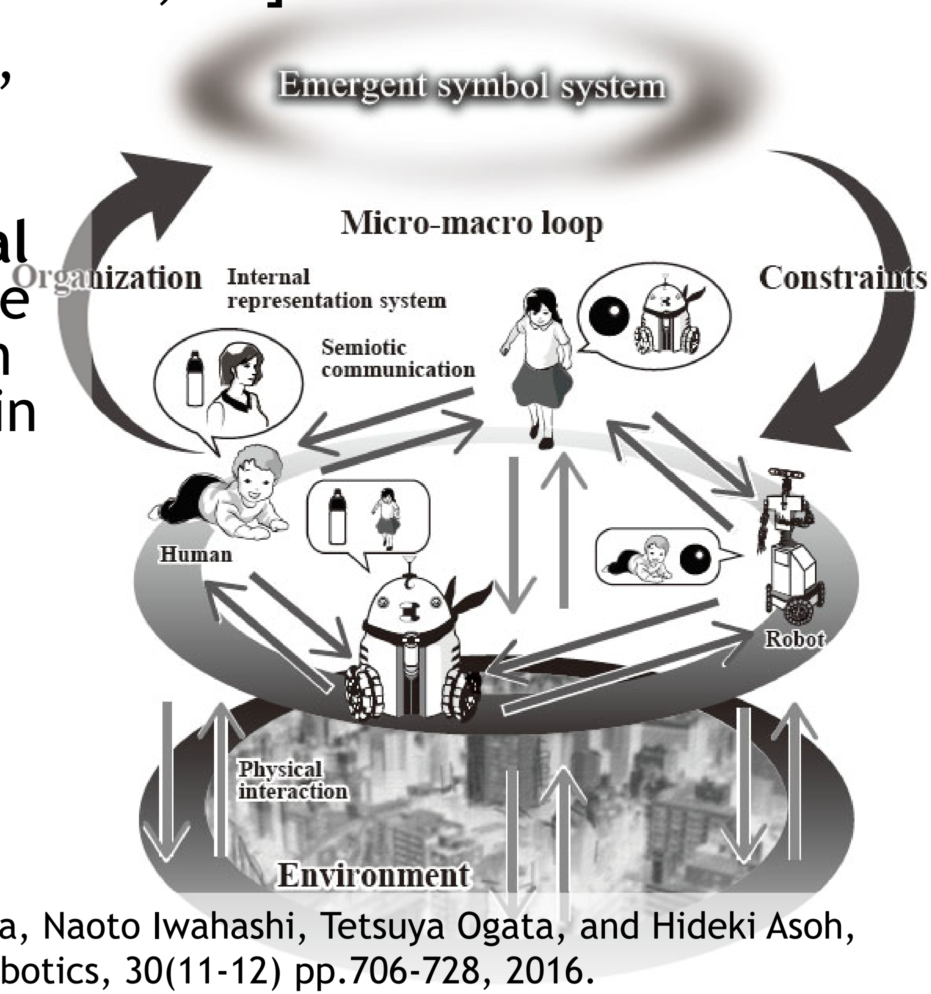
Related keywords

*Predictive coding, world model, unsupervised learning and latent variable models*

# Symbol emergence in robotics

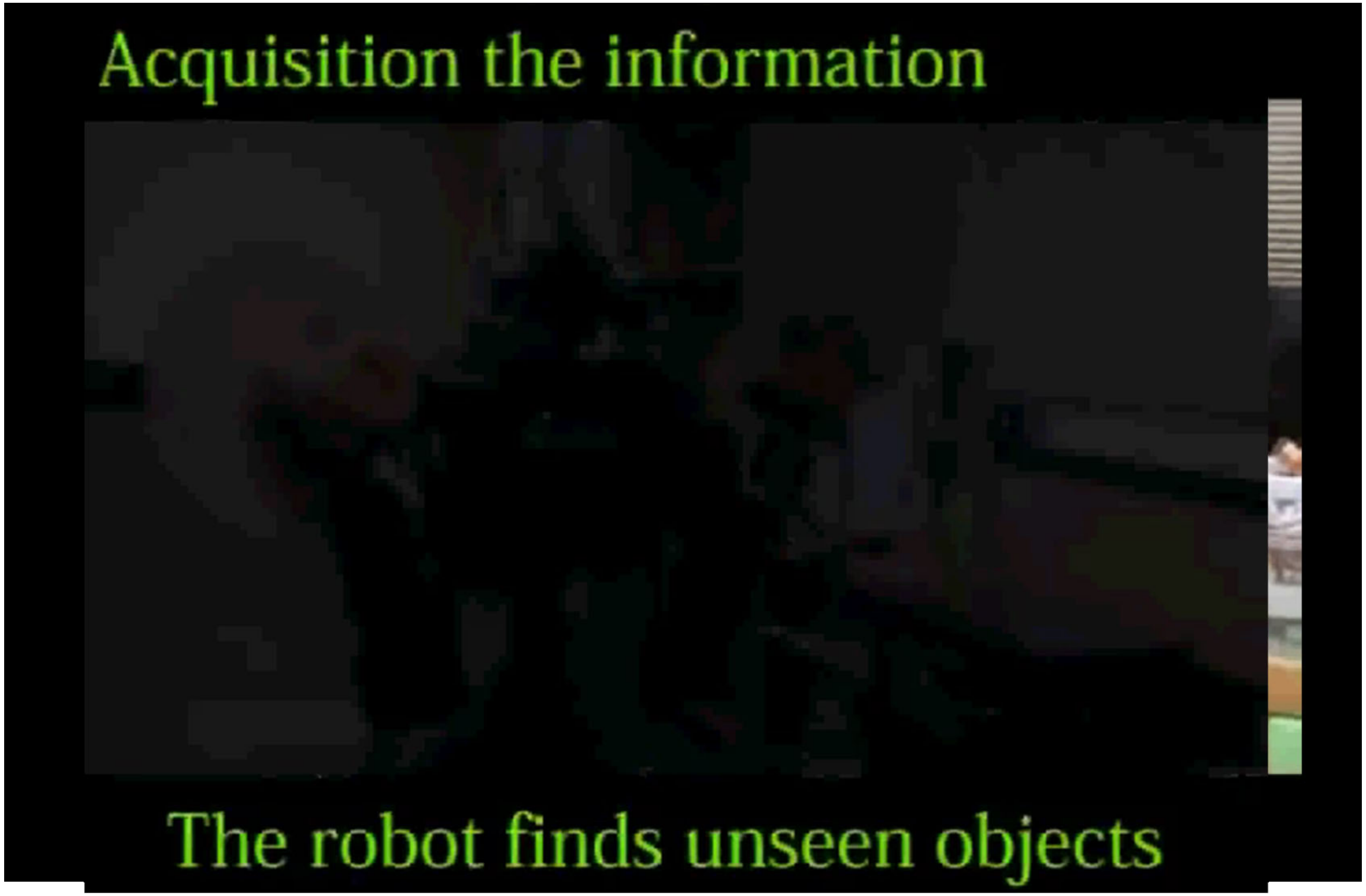
[Taniguchi+ 16, 19]

- Symbol systems, e.g., language, in our society is formed in a bottom-up manner.
- Bottom-up formation of **internal representations** in our cognitive systems and **symbol systems** in our society should be modeled in a computational/constructive manner.
- The phenomena should be reproduced in a **real-world environment.**




Tadahiro Taniguchi, Takayuki Nagai, Tomoaki Nakamura, Naoto Iwahashi, Tetsuya Ogata, and Hideki Asoh, Symbol Emergence in Robotics: A Survey, Advanced Robotics, 30(11-12) pp.706-728, 2016. DOI:10.1080/01691864.2016.1164622
Tadahiro Taniguchi, Emre Ugur, Matej Hoffmann, Lorenzo Jamone, Takayuki Nagai, Benjamin Rosman, Toshihiko Matsuka, Naoto Iwahashi, Erhan Oztop, Justus Piater, Florentin Wörgötter, Symbol Emergence in Cognitive Developmental Systems: A Survey, IEEE Transactions on Cognitive and Developmental Systems, 11(4), pp.494-516, 2019. DOI: 10.1109/TCDS.2018.2867772

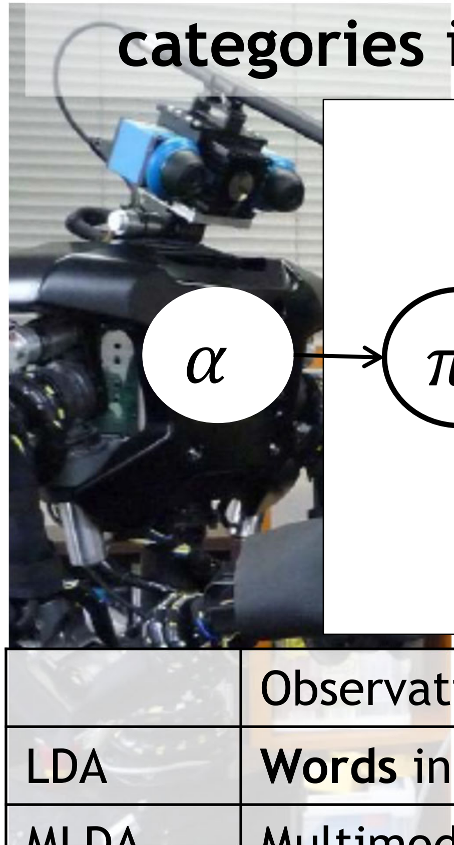
# **Multimodal Categorization** and Lexical Acquisition by an Autonomous Robot [Nakamura+ 2009-]

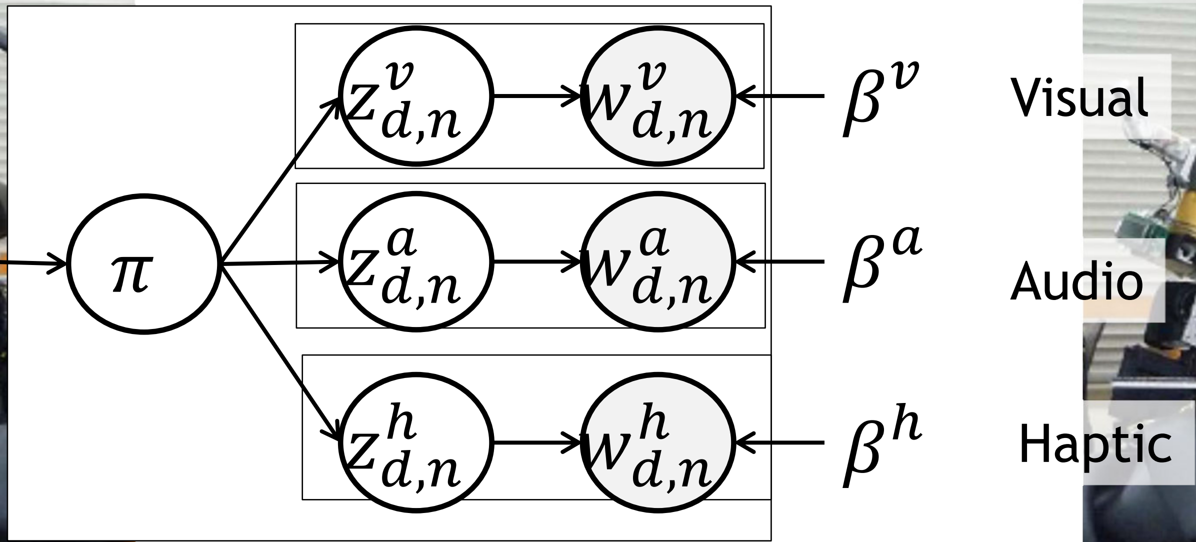


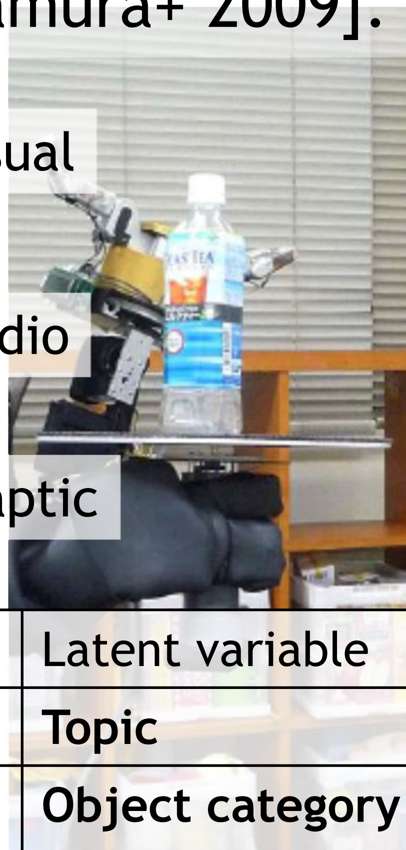

Takaya Araki, Tomoaki Nakamura, Takayuki Nagai, Shogo Nagasaka, Tadahiro Taniguchi, Naoto Iwahashi. Online Learning of Concepts and Words Using Multimodal LDA and Hierarchical Pitman-Yor Language Model. IEEE/RSJ International Conference on Intelligent Robots and Systems 2012 (IROS 2012), 1623-1630 .(2012)

# Multimodal latent Dirichlet allocation(MLDA)

- The **MLDA** is a multimodal categorization method that is an extension of the LDA [Blei+ 2004].
- The MLDA was proposed for making a **robot** form **object categories** in an unsupervised manner [Nakamura+ 2009].



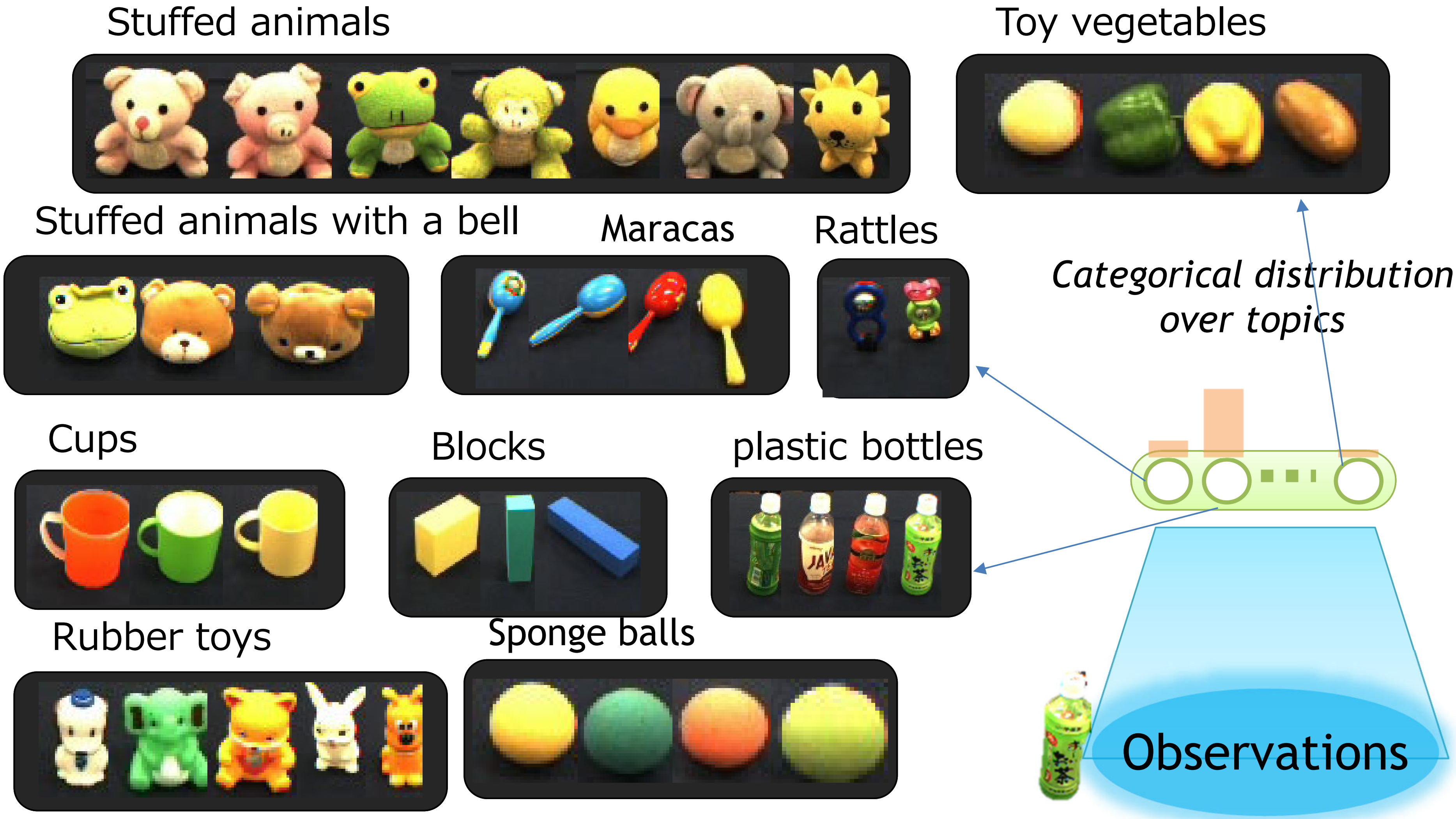


| | Observations | Latent variable |
|---|---|---|
| LDA | **Words** in a **document** (i.e., Bag of words) | **Topic** |
| MLDA | Multimodal (**visual**, **auditory**, and **haptic**) **features** obtained from an **object** (i.e., Bag of features) | **Object category** |

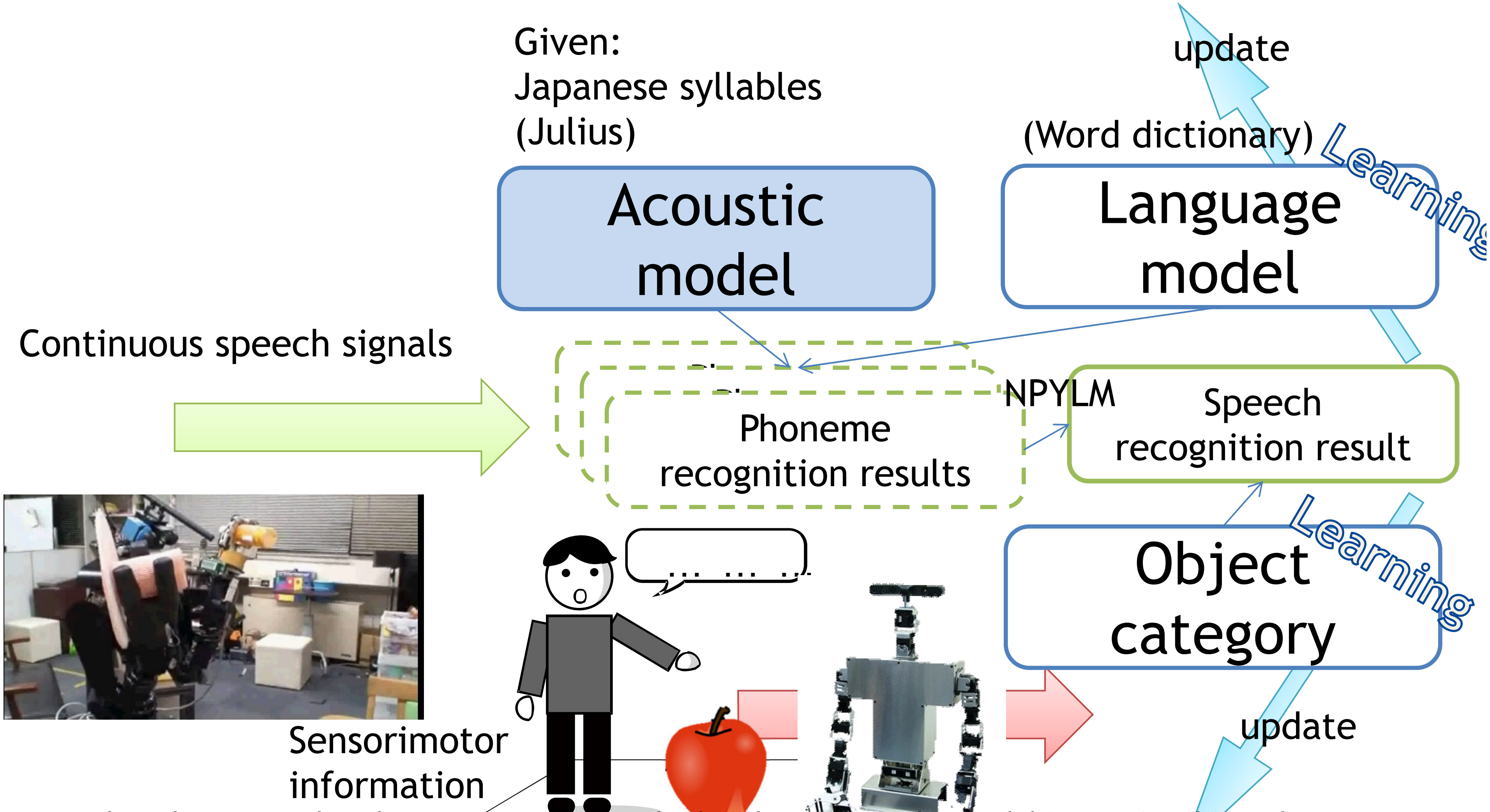
Nakamura, T., Nagai, T., & Iwahashi, N. (2009). Grounding of word meanings in multimodal concepts using LDA. *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 3943-3948.

# Categorization result based on real-world multimodal sensorimotor information



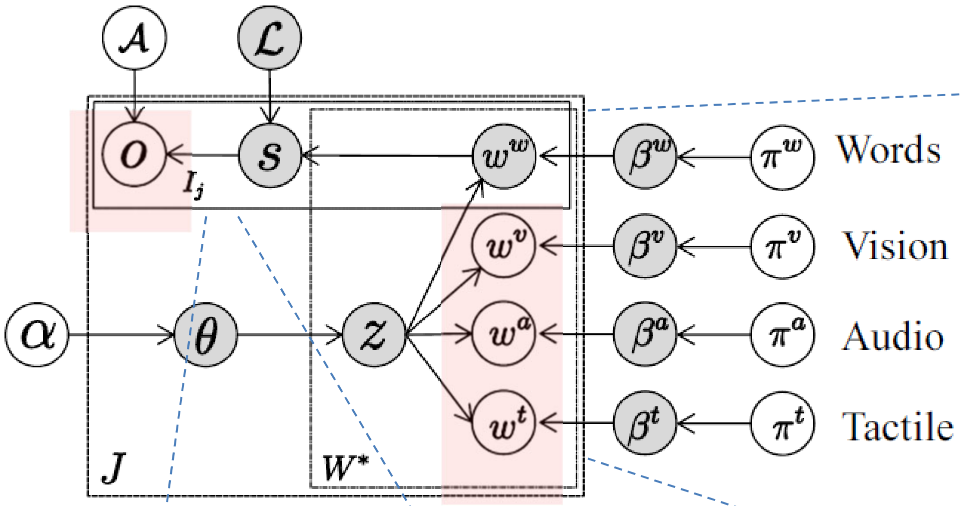

By integrating multimodal information, the robot formed categories represented by latent variables that were similar to most of the human participants.

# Simultaneous acquisition of word units and multimodal object categories [Nakamura+ 2014]



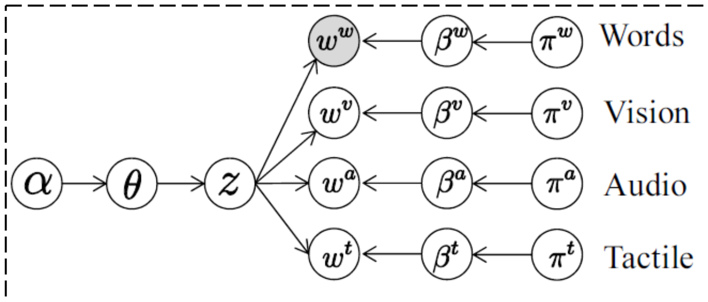

Tomoaki Nakamura, Takayuki Nagai, Kotaro Funakoshi, Shogo Nagasaka, **Tadahiro Taniguchi**, and Naoto Iwahashi, Mutual Learning of an Object Concept and Language Model Based on MLDA and NPYLM, 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'14), 600 - 607 .(2014)

# Probabilistic generative model for simultaneous acquisition of word units and multimodal object categories

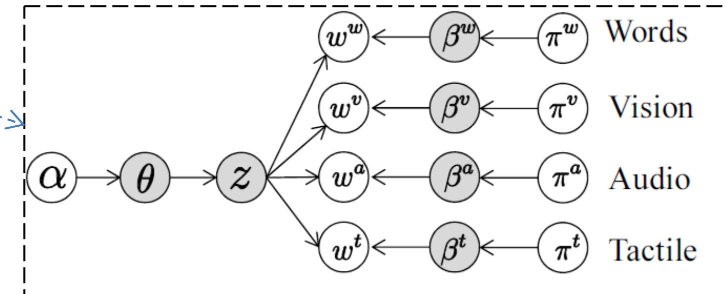



observations

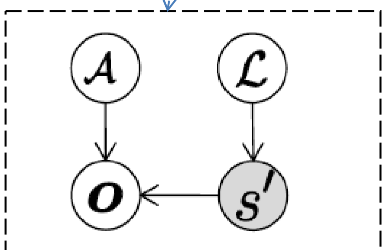
by MLDA




(c) Word candidates selection

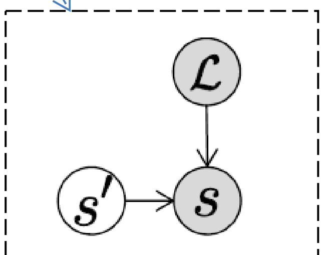



(d) Object concept



(a) Speech recognition

by Julius



(b) Language model

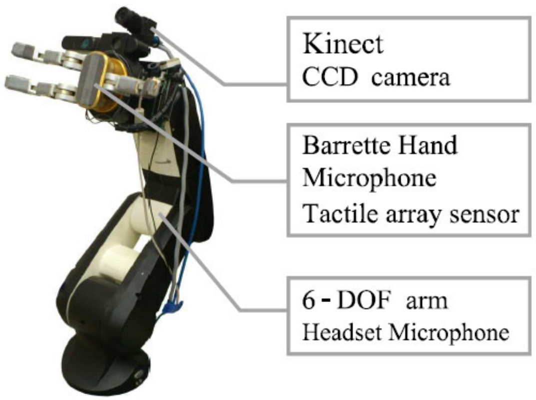
by NPYLM

Each component is updated using approximate MAP estimation iteratively.

-> **Approximate inference procedure**

Tomoaki Nakamura, Takayuki Nagai, Kotaro Funakoshi, Shogo Nagasaka, **Tadahiro Taniguchi**, and Naoto Iwahashi, Mutual Learning of an Object Concept and Language Model Based on MLDA and NPYLM, 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'14), 600 - 607 .(2014)

# Overview of experiment and results



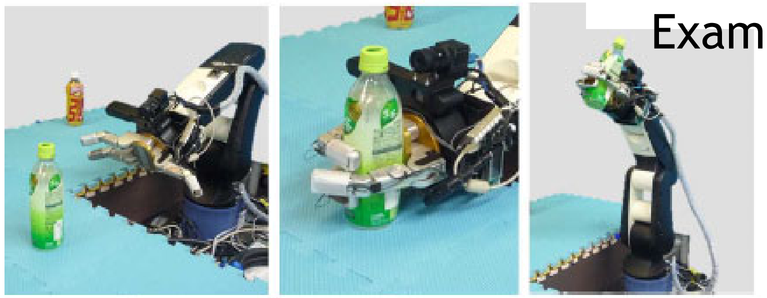

Robot

This is a red spray can. (ko re wa a ka i su pu re e ka N)
This makes a sound when shaken. (ko re wa o to ga shi ma su)
This is made of metal and is hard. (ko re wa ki N zo ku de de ki te i te ka ta i)

A green plushie of a frog. (mi do ri no ka e ru no nu i gu ru mi)
This is soft. (ko re wa ya wa ra ka i)
This is an animal. (ko re wa do u bu tsu)

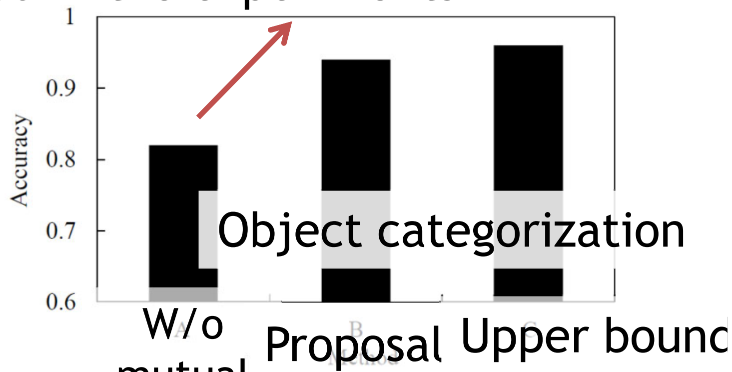
A green plastic bottle. (mi do ri no pe tto bo to ru)
This is green tea. (ko re wa ryo ku cha)

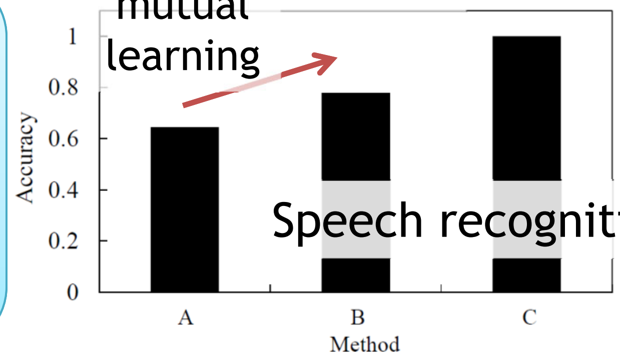
Example sentences used in the experiments



(a) (b) (c)

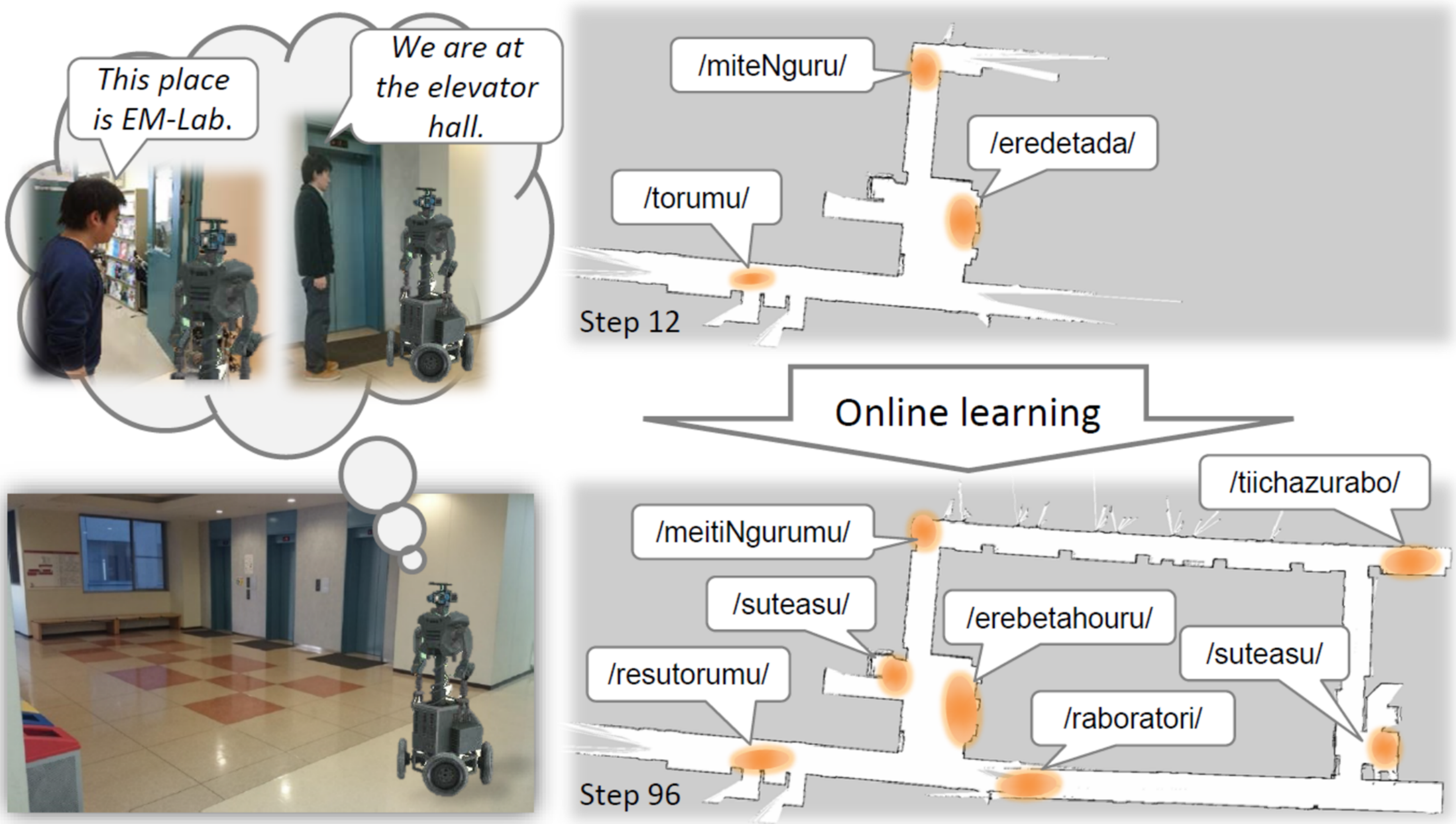
Obtaining multimodal sensory information




- ✓ Unsupervised lexical acquisition was performed successfully.
- ✓ Both **object categorization** and **speech recognition performances** increased using co-occurrence cues.

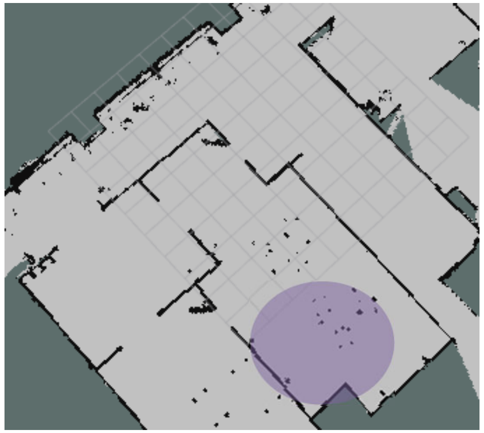
# Online spatial concept acquisition method
# **SpCoSLAM** [Taniguchi+ 2017]
# (including word discovery task)




Akira Taniguchi, Yoshinobu Hagiwara, Tadahiro Taniguchi and Tetsunari Inamura, Online Spatial Concept and Lexical Acquisition with Simultaneous Localization and Mapping, IEEE IROS 2017 p. 811-818 .(2017) oral

# Spatial concept is multimodal

Position



Where is the space?

Word

"This is *the third table*"
"*A meeting space*"
"*Under the air conditioner*"

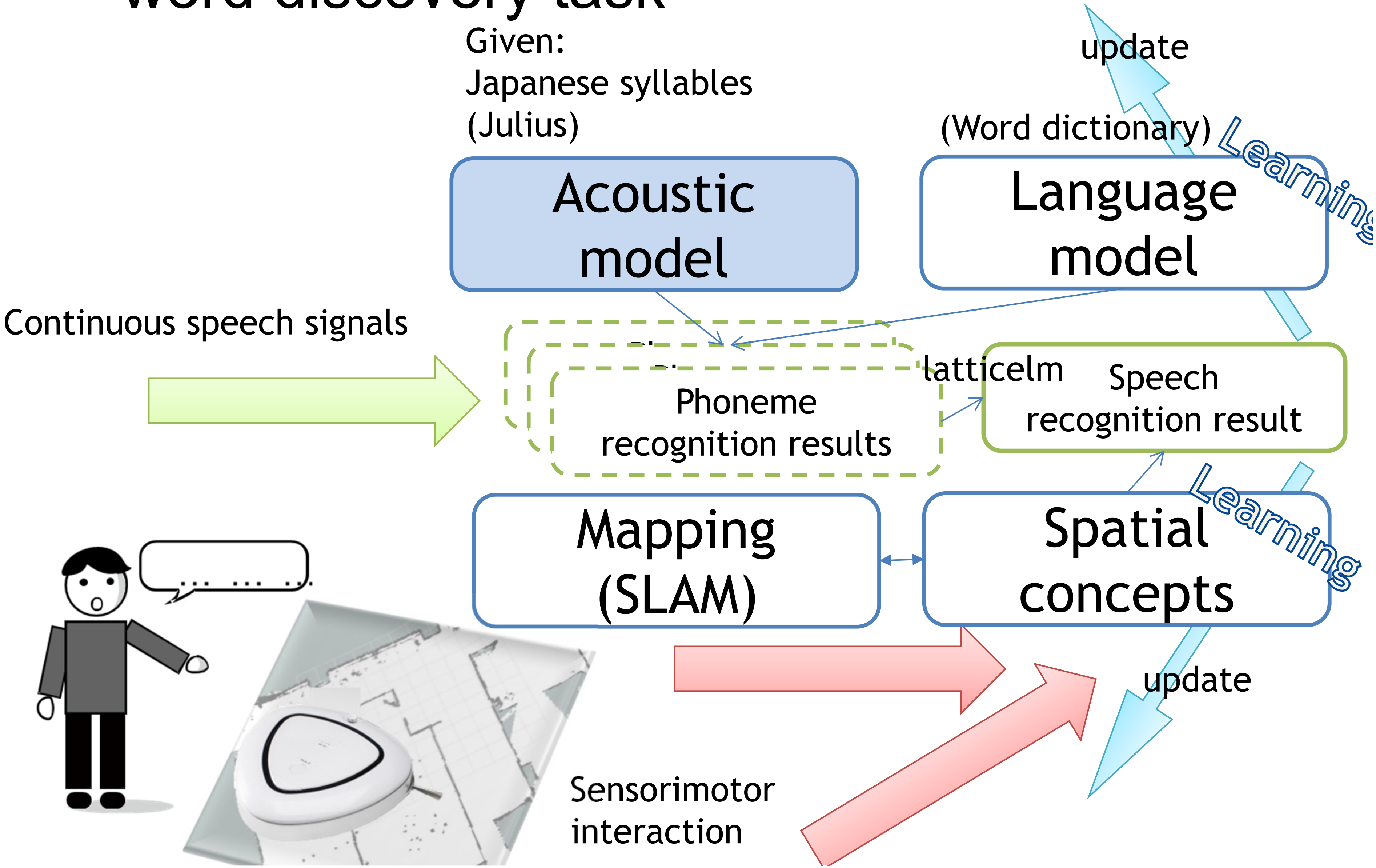
How do they call the space?

Visual information

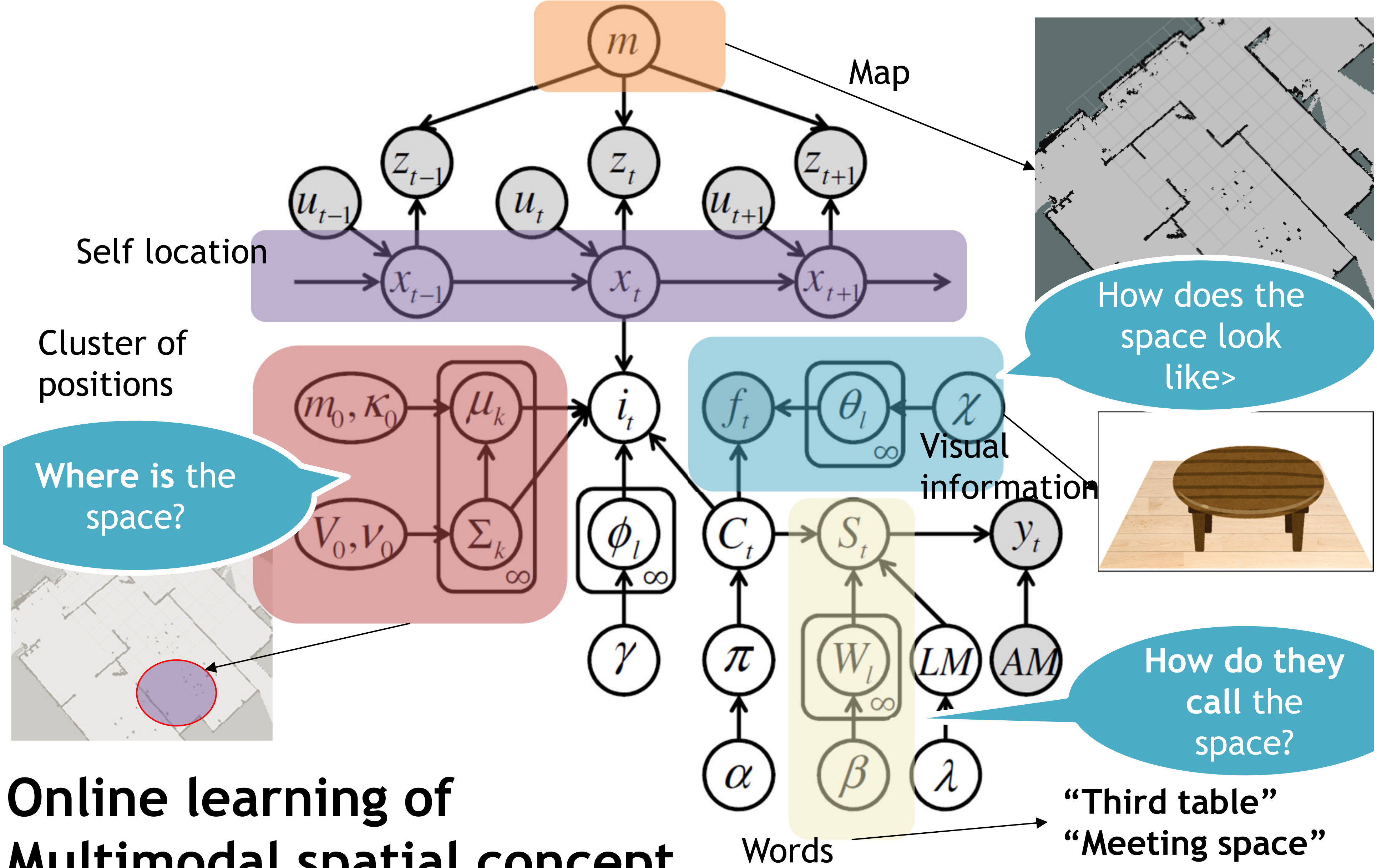
How does the space look like>

Sound, smell, ,,,

# Online spatial concept acquisition with word discovery task

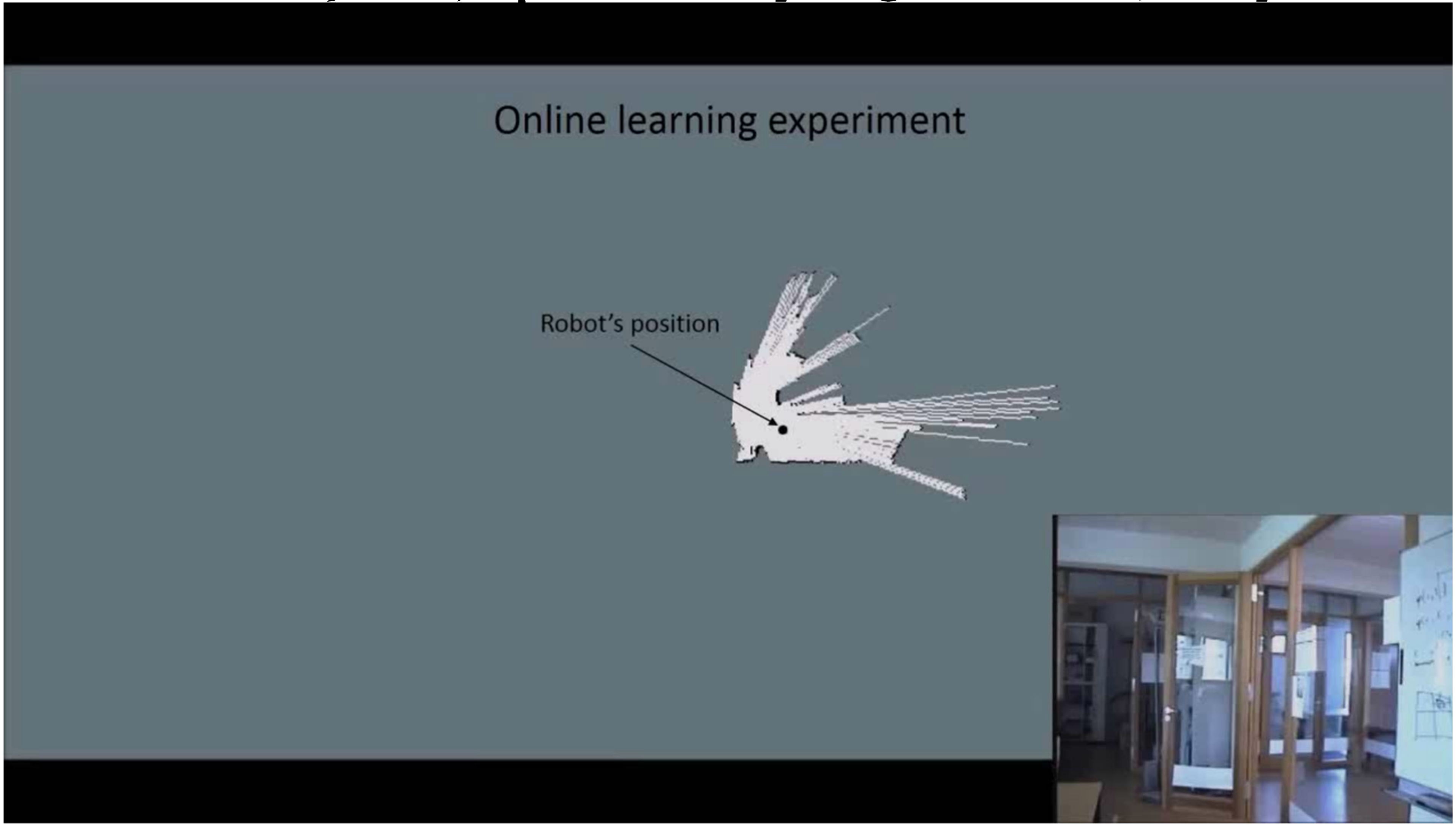


# Graphical model of SpCoSLAM



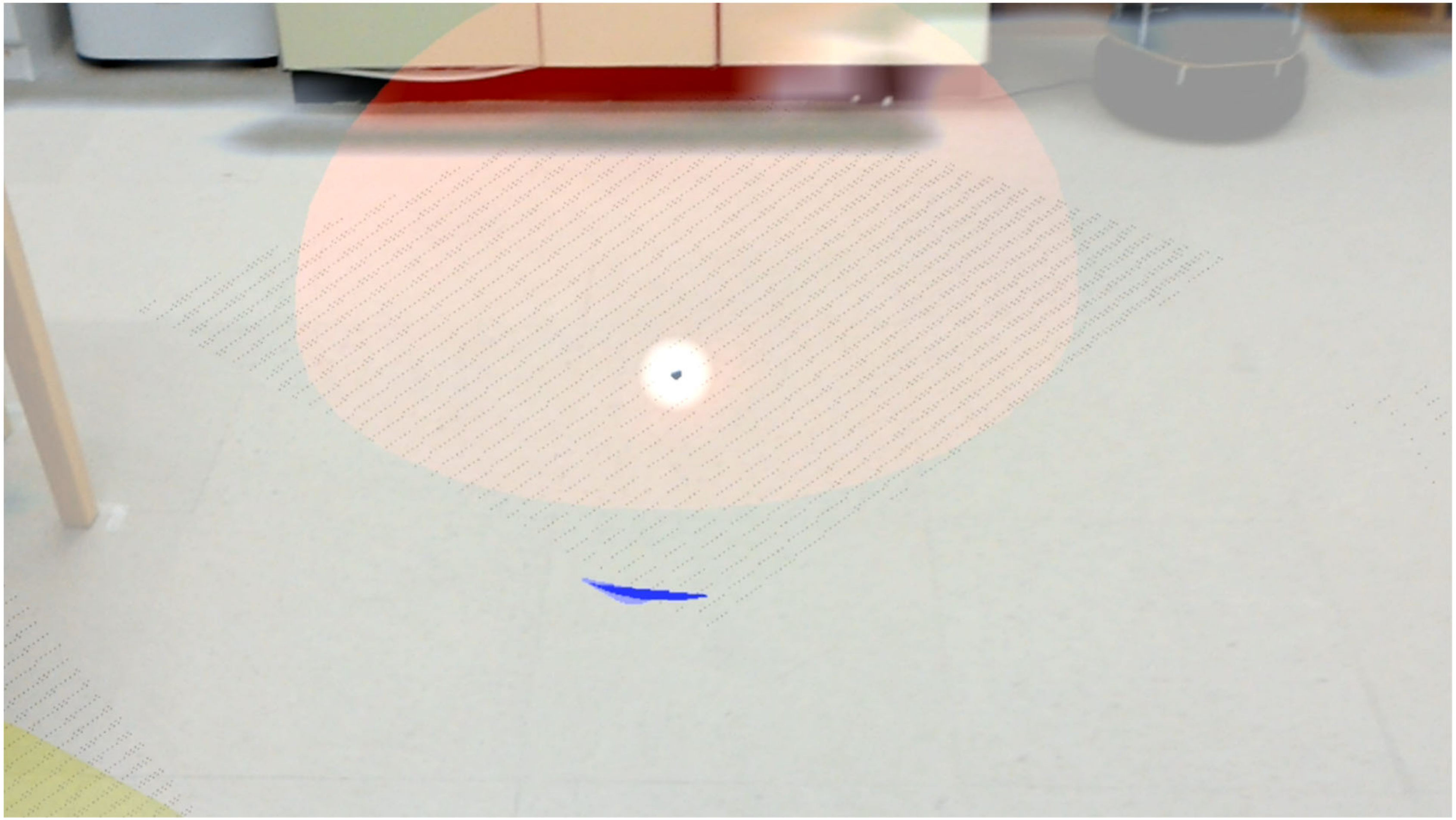

**Online learning of**
**Multimodal spatial concept**

# Online spatial concept acquisition method (including word discovery task) **SpCoSLAM** [Taniguchi+ 2017,2020]



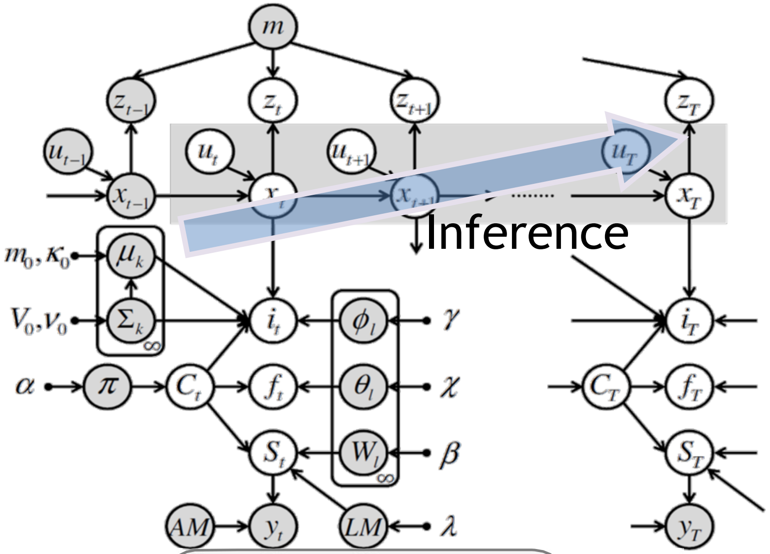

Akira Taniguchi, Yoshinobu Hagiwara, Tadahiro Taniguchi and Tetsunari Inamura, Online Spatial Concept and Lexical Acquisition with Simultaneous Localization and Mapping, IEEE IROS 2017
Akira Taniguchi, Yoshinobu Hagiwara, Tadahiro Taniguchi, Tetsunari Inamura, Improved and scalable online learning of spatial concepts and language models with mapping, Autonomous Robots, 44(6), pp.927-946, 2020. DOI: 10.1007/s10514-020-09905-0

# Visualizing robot's memory and perception

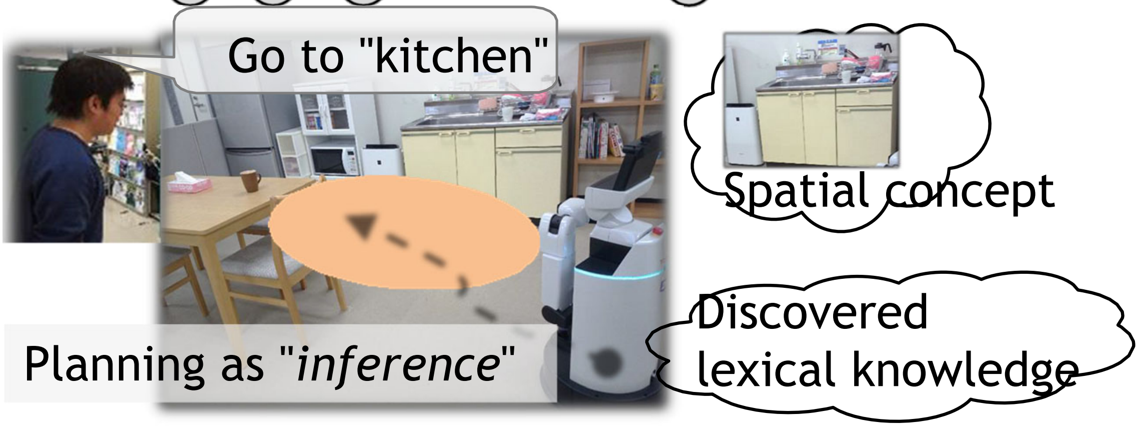


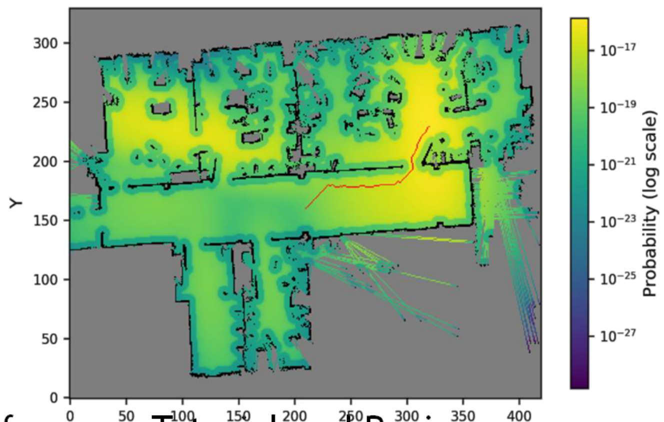
L. El Hafi, S. Isobe, Y. Tabuchi, Y. Katsumata, H. Nakamura, T. Fukui, T. Matsuo,G.A. Garcia Ricardez, M. Yamamoto, A. Taniguchi, Y. Hagiwara, and T. Taniguchi, System for augmented human-robot interaction through mixed reality and robot training by non-experts in customer service environments, Advanced Robotics, 34(3-4), pp.157-172, 2020. DOI: 10.1080/01691864.2019.1694068

# Spatial Concept-Based Navigation with Human Speech Instructions via Probabilistic Inference [Taniguchi+ 20]




- Control as Probabilistic Inference (CoI)
  - We can reformulate reinforcement learning as Bayesian inference on (PO)MDP.[Levince 18, Okada 20]
- SpCoNavi [Taniguchi+ 20]
  - In the same way as CoI, Bayesian inference on the PGM of SoCoSLAM can perform path planning/navigation.

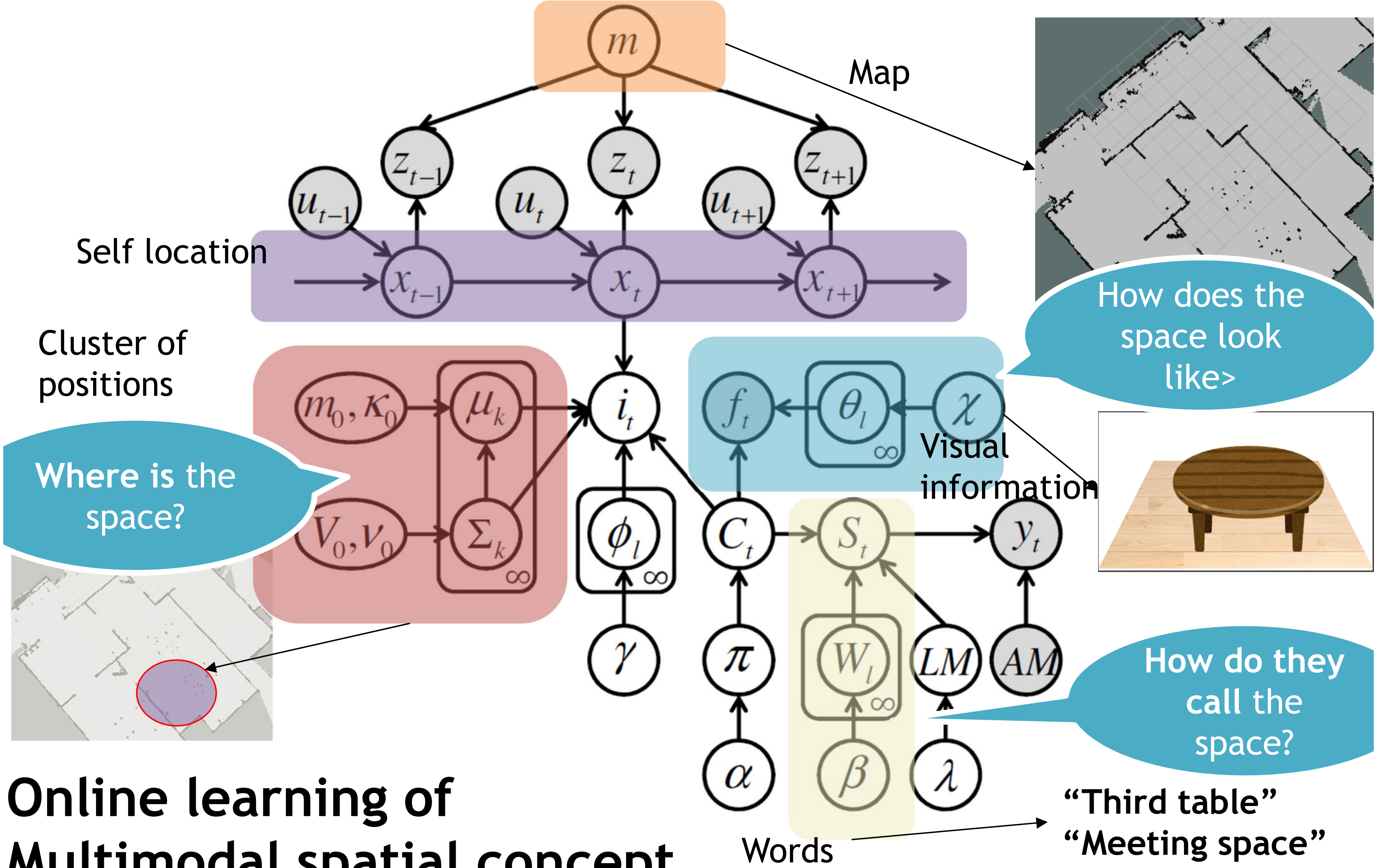



Sergey Levine, Reinforcement Learning and Control as Probabilistic Inference: Tutorial and Review, arXiv:1805.00909 [cs.LG], 2 May, (2018) https://arxiv.org/abs/1805.00909
Masashi Okada, Norio Kosaka, Tadahiro Taniguchi, PlaNet of the Bayesians: Reconsidering and Improving Deep Planning Network by Incorporating Bayesian Inference, IROS2020
Akira Taniguchi, Yoshinobu Hagiwara, Tadahiro Taniguchi, Tetsunari Inamura, Spatial Concept-Based Navigation with Human Speech Instructions via Probabilistic Inference on Bayesian Generative Model, Advanced Robotics, 34(19), pp.1213-1228, 2020. DOI: 10.1080/01691864.2020.1817777

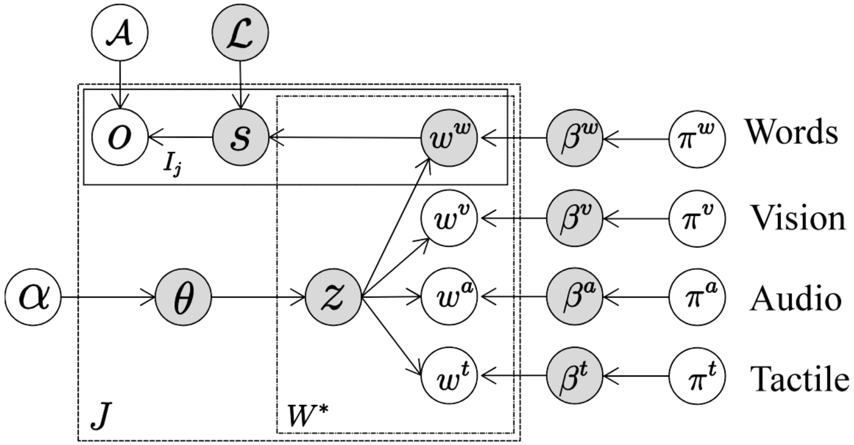
# World Robot Summit 2018
# METI/NEDO, Japan

# Graphical model of SpCoSLAM

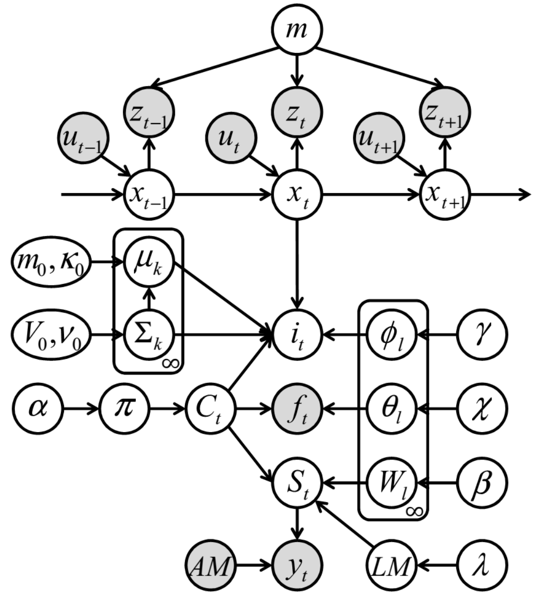


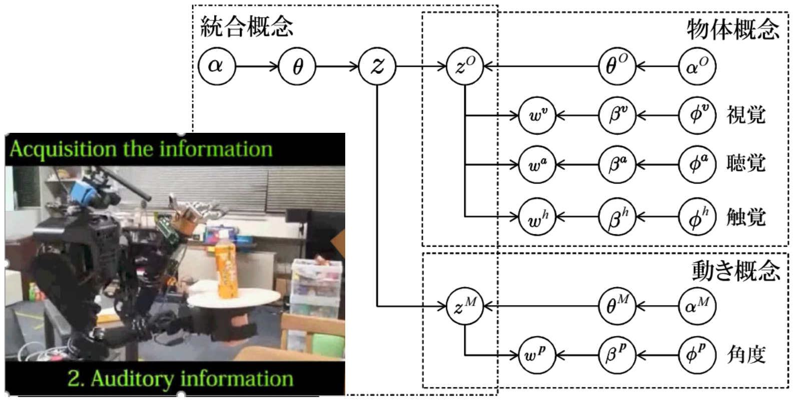

**Online learning of Multimodal spatial concept**

# Probabilistic generative models for integrative cognitive systems

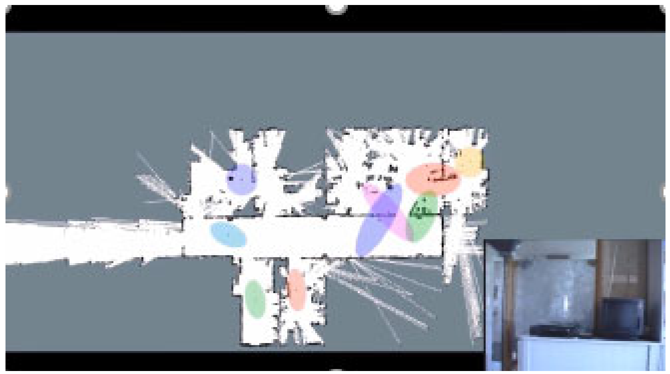


Learning object and language [Nakamura+ 15]


Place concept formation [Taniguchi+ 17]

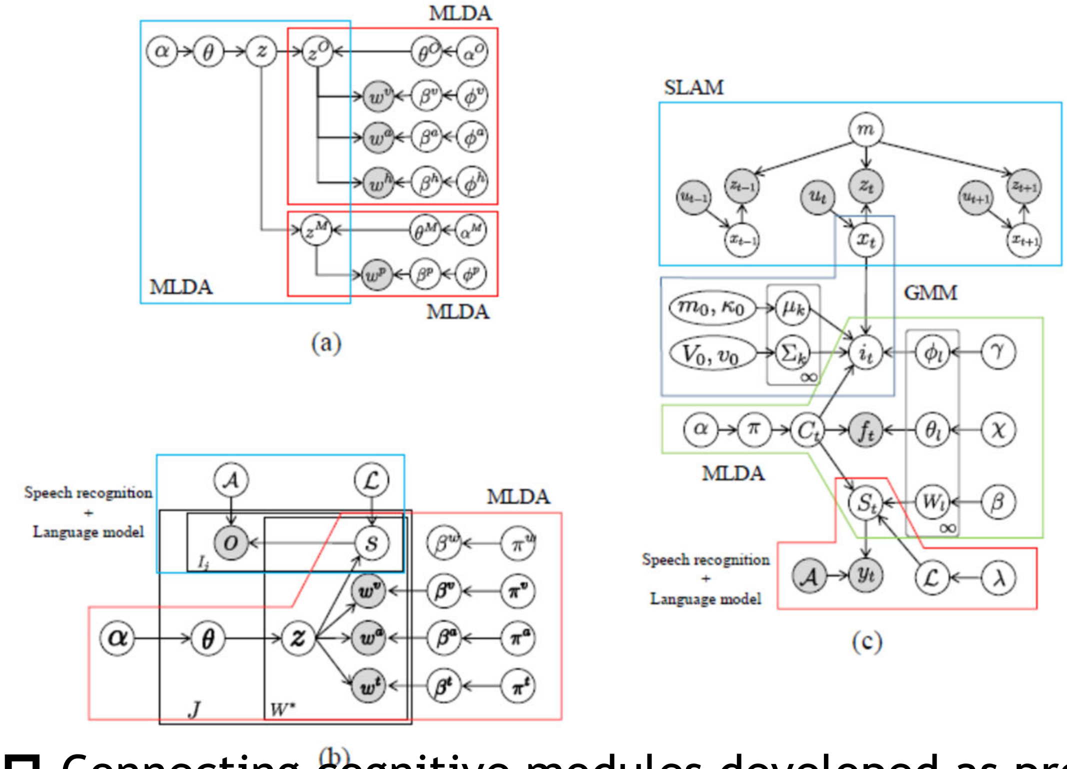


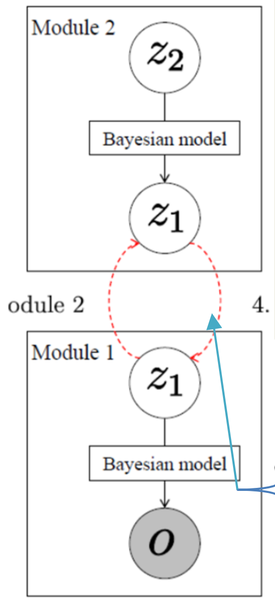
Formation of integrated concepts with object and motion [Attamimi+ 14]

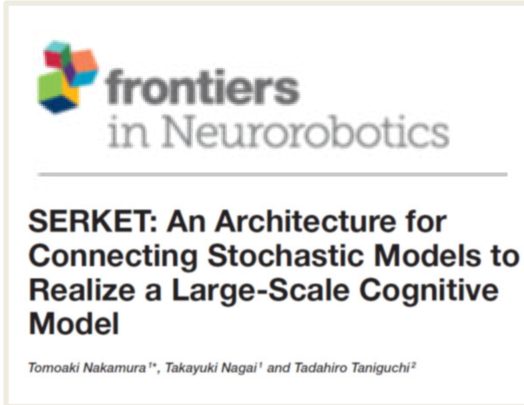


Developing integrative cognitive systems from full-scratch requires huge cost, i.e., labor force.

# SERKET: An Architecture for Connecting Stochastic Models to Realize a Large-Scale Cognitive Model [Nakamura+ 18]



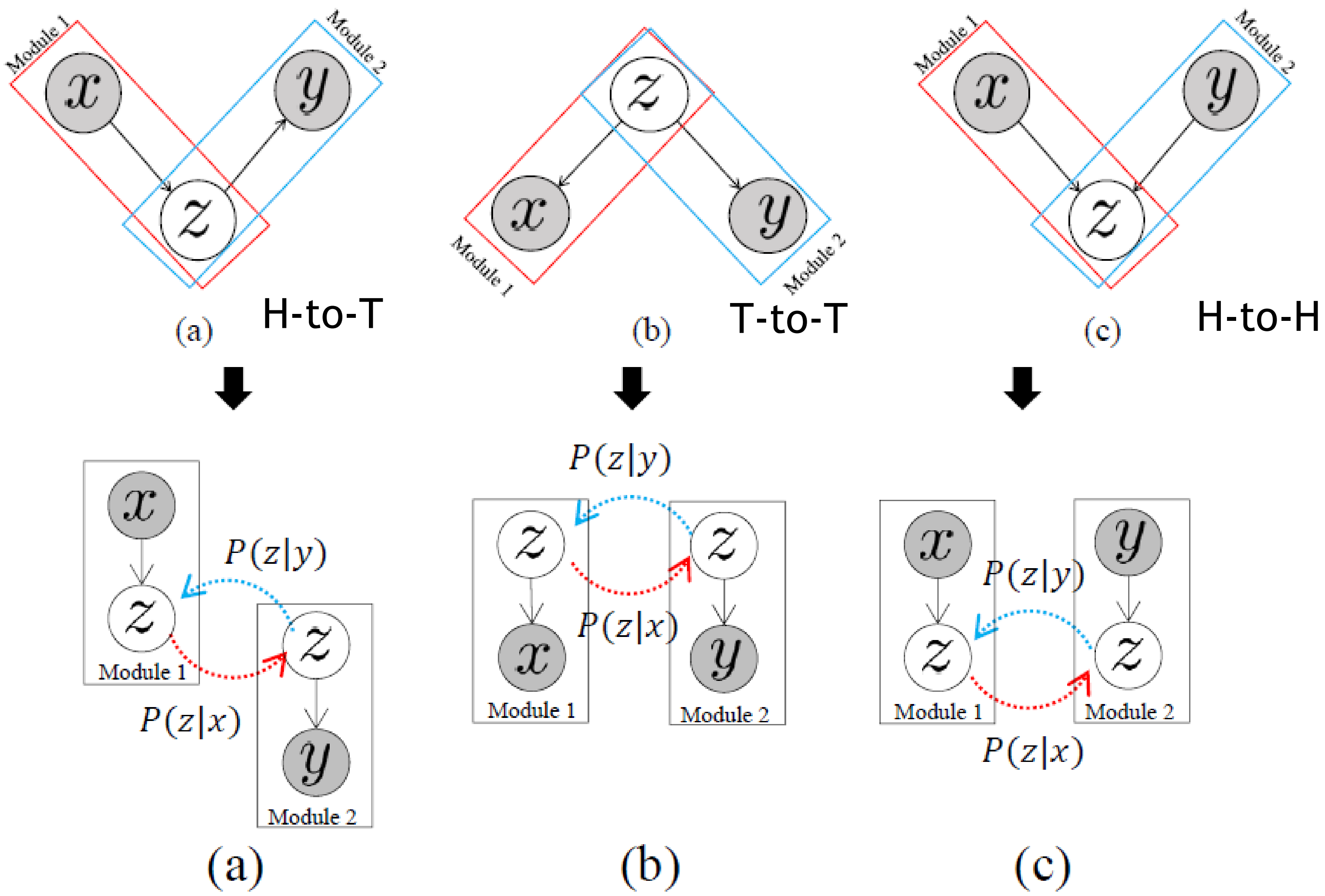

1. Belief propagation
2. SIR
3. MH

- 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. Neurorobot. 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

**Generation:**
Decomposition of Complex Graphical Model




**Inference:**
Composition of Complex Graphical Model

# Example: unsupervised categorization of image and speech [Taniguchi+ 2020]

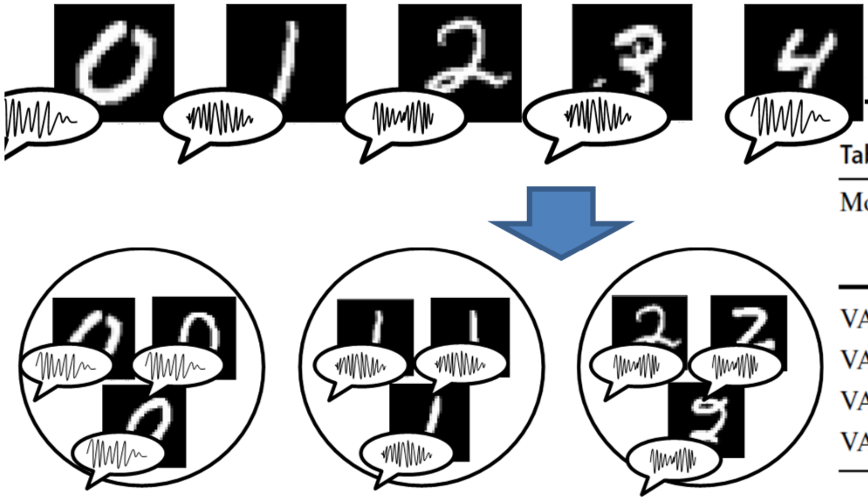


Table 4 Classification accuracy in the GMM and LDA modules

| Model | Accuracy (%) | | Features introduced in Neuro-SERKET | | |
|---|---|---|---|---|---|
| | GMM | LDA | Head-to-head | Tail-to-tail | Neura |
| VAE GMM LDA ASR | 62.0 | 27.4 | | | |
| VAE GMM LDA + ASR | 62.0 | 91.8 | ✓ | | |
| VAE + GMM LDA + ASR | 63.7 | 91.8 | ✓ | | ✓ |
| VAE + GMM + LDA + ASR | **91.0** | **93.7** | ✓ | ✓ | ✓ |

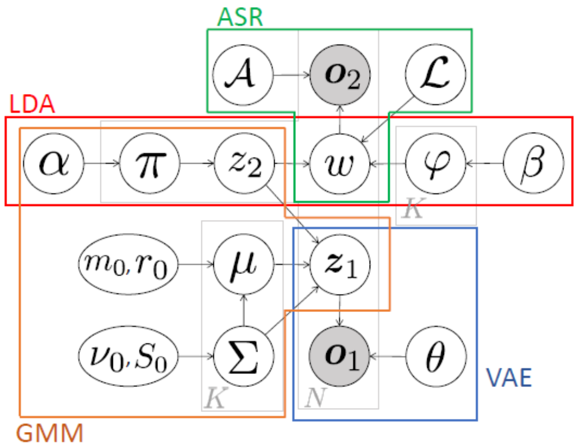



Fig. 5 The original graphical model of the integrative PGM (VAE+GMM+LDA+ASR). Each block shows each module.

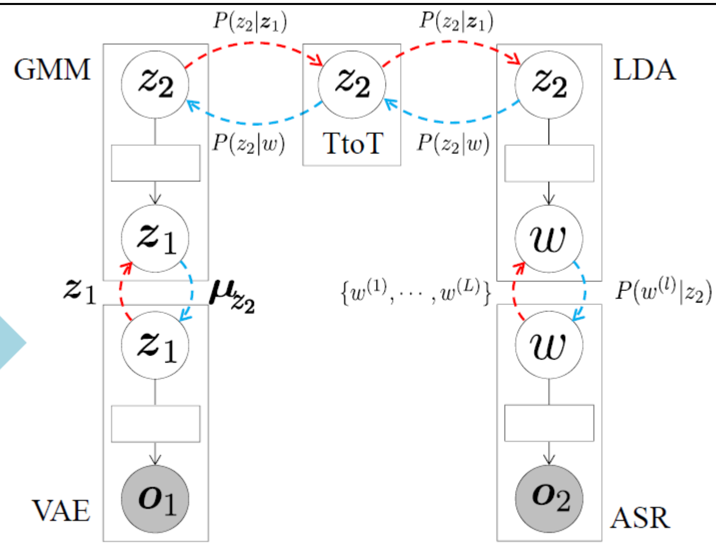
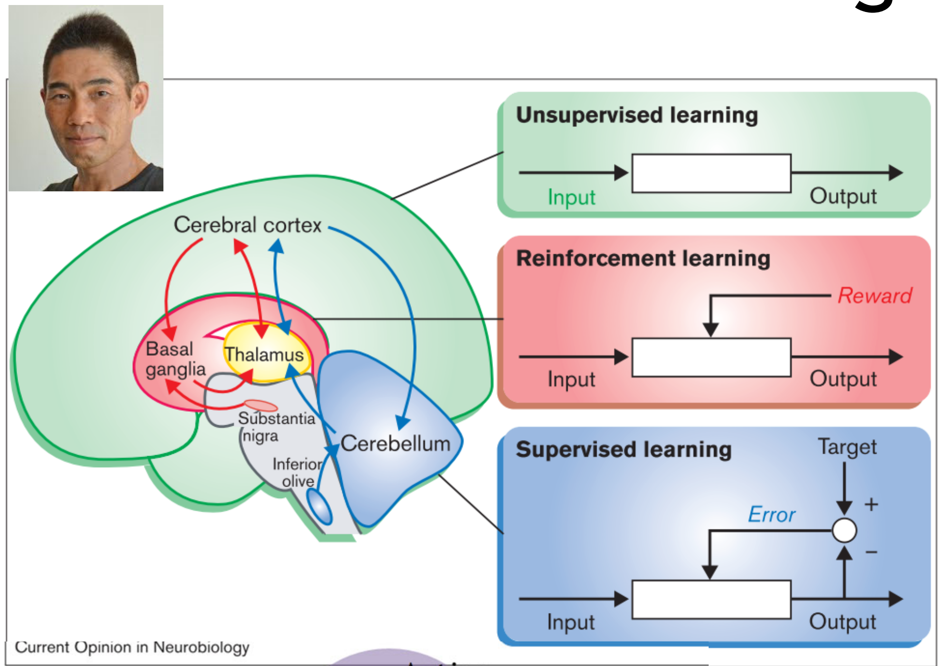


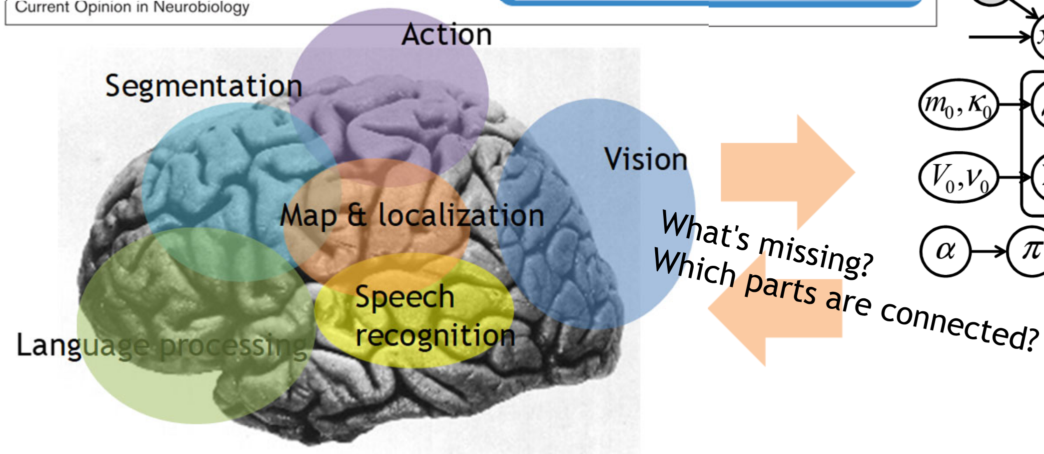

Fig. 6 Decomposed modules and communication between them following the Neuro-SERKET framework.

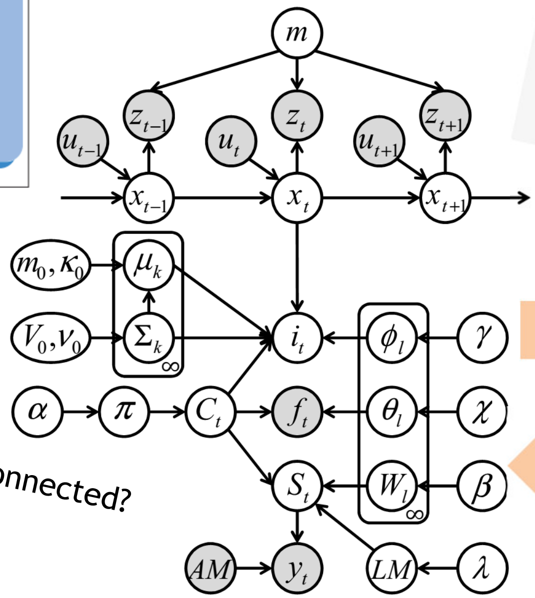
**Taniguchi, T.,** Nakamura, T., Suzuki, M. et al. Neuro-SERKET: Development of Integrative Cognitive System Through the Composition of Deep Probabilistic Generative Models. New Gener. Comput. 38, 23–48 (2020). https://doi.org/10.1007/s00354-019-00084-w

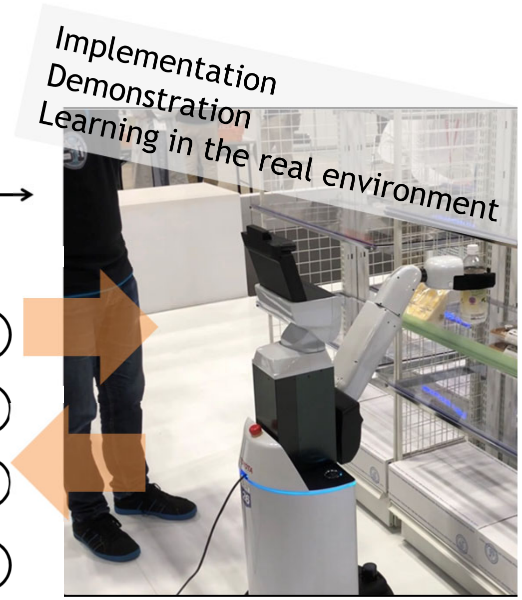
# "What can we further learn from the brain for next-generation AI?"




- ✓ It was hypothesized that cerebral cortex is for unsupervised learning.
- ✓ Unsupervised learning can perform a wide range of tasks using PGMs.
- ✓ The brain is an excellent reference to build an integrative cognitive architecture using PGMs.






Doya, Kenji. "Complementary roles of basal ganglia and cerebellum in learning and motor control." Current opinion in neurobiology 10.6 (2000): 732-739.
Doya, Kenji. "What are the computations of the cerebellum, the basal ganglia and the cerebral cortex?." Neural networks 12.7-8 (1999): 961-974.

ADVANCED ROBOTICS
2019, VOL. 33, NOS. 15–16, 700–730
https://doi.org/10.1080/01691864.2019.1632223

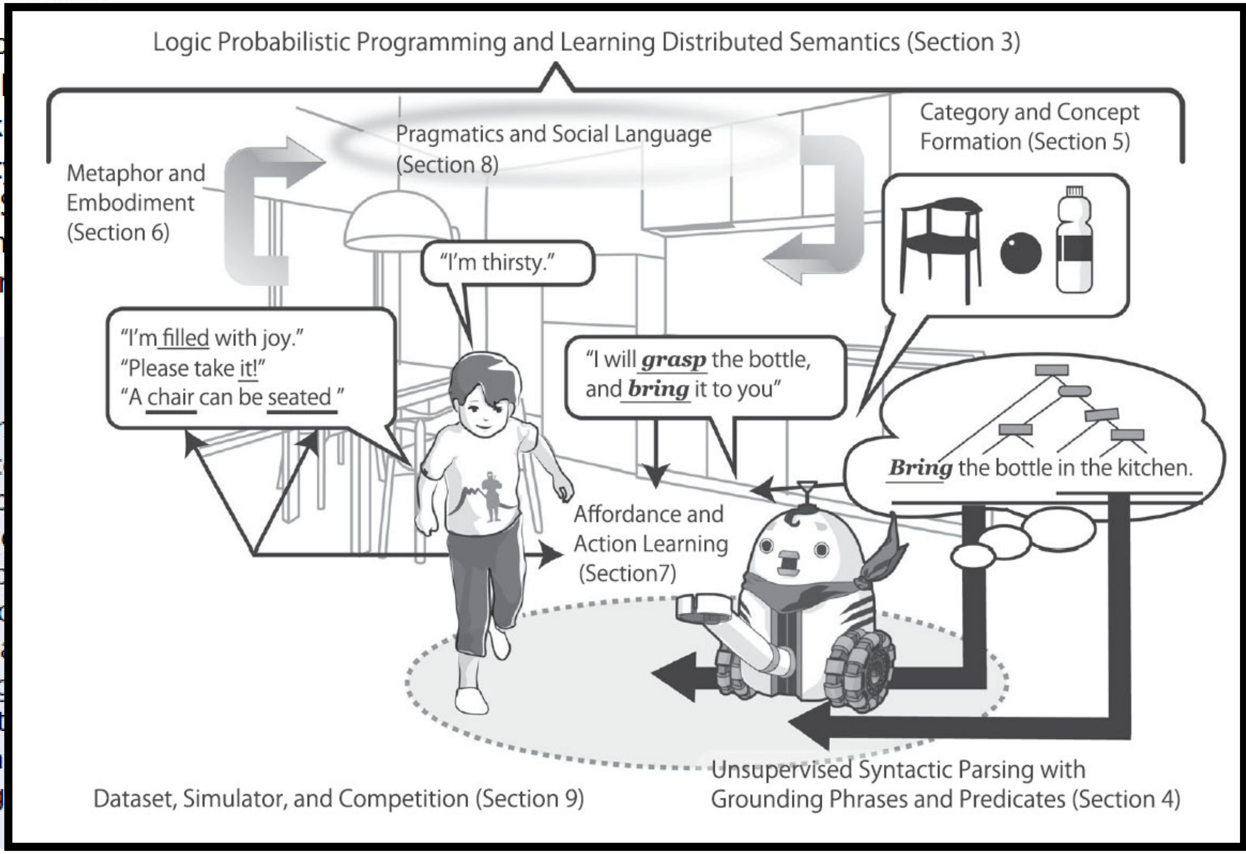
SURVEY PAPER

OPEN ACCESS Check for updates

# Survey on frontiers of language and robotics

T. Tangiuchi[a], D. Mochihashi[b,c], T. Nagai[d], S. Uchida[e], N. Inoue[f,g], I. Kobayashi[h], T. Nakamura[i], Y. Hagiwara[a], N. Iwahashi[j] and T. Inamura[c,k]

[a]Department of Informatic … thematics,
Tachikawa, Japan; [c]SOKEN … g Science, Osaka
University, Toyonaka, Osak … l of Information
Sciences, Tohoku Universit … nces, Graduate
School of Humanities and … gent Systems, The
University of Electro-Comm … a Prefectural
University, Okayama, Japa




**ABSTRACT**
The understanding ar
future robotics servic
ing on the problem fo
despite significant pr
models) during the p
and organized, as mo
This study conducts
and robotics, ranging
ate language underst
language from intera
robots to learn a lang

… HISTORY
… 6 March 2019
… 9 May 2019
… 2 June 2019

… RDS
… botics; cognitive
… symbol emergence;
… nguage processing;
… understanding;
… acquisition

T. Tangiuchi, D. Mochihashi, T. Nagai, S. Uchida, N. Inoue, I. Kobayashi, T. Nakamura, Y. Hagiwara, N. Iwahashi & T. Inamura, Survey on frontiers of language and robotics, Advanced Robotics, 33(15-16), 700-730, 2019. DOI: 10.1080/01691864.2019.1632223

# Summary

- ✓ Symbol emergence in robotics is a field in which we explore computational and robotic model that can form internal representations and learn symbolic systems, e.g., language, in a real environment.
- ✓ Probabilistic generative models, including DPGMs, allow us to build integrative cognitive systems that can learn many kinds of concepts from real sensorimotor experience in an unsupervised manner.
- ✓ SERKET framework is introduced for further development of PGM-based cognitive systems.

# Future challenges

- ✓ Learning wide range of concepts and language, e.g., abstract concepts and syntactic knowledge, through real-world interactions
- ✓ Developing service robots that can be installed into a real service environment using learning-based approach.
- ✓ Developing a whole-brain cognitive architecture integrating multimodal sensorimotor information and high and low-level cognitive processes.

# Information

Funding information

MEXT Grand-in-Aid for Scientific Research on Innovative Areas
Correspondence and Fusion of
Artificial Intelligence and Brain Science
Leader: Prof. Kenji Doya

JST-CREST
**Symbol Emergence in Robotics for Future Human-Machine Collaboration**
Leader: Prof. Takayuki Nagai

Ritsumeikan Global Innovation Research Organization
International and Interdisciplinary Research Center of
Next-generation Artificial Intelligence and Semiotics
Leader: Prof. Tadahiro Taniguchi

Email: taniguchi@ci.ristumei.ac.jp
Twitter: @tanichu/ Facebook: Tadahiro Taniguchi