HOKKAIDO UNIVERSITY

Propagation of Mind Through the Mechanism of Superposition

Center for Human nature, AI, and Neuroscience (CHAIN),
Hokkaido University, Japan

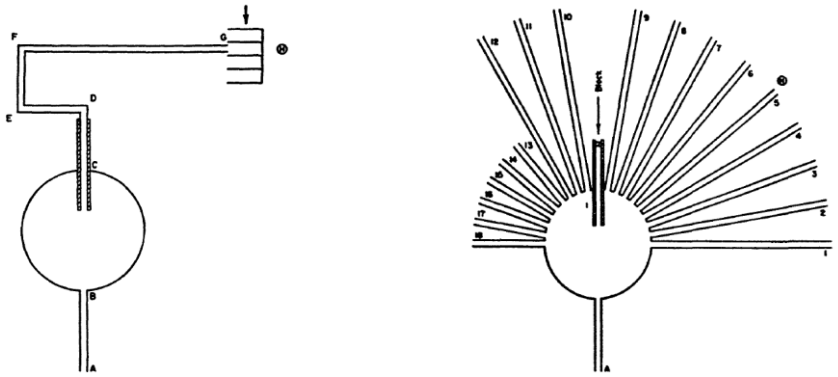
CHAIN
Center for Human Nature,
Artificial Intelligence, and Neuroscience

Hiroyuki Iizuka

iizuka@chain.hokudai.ac.jp

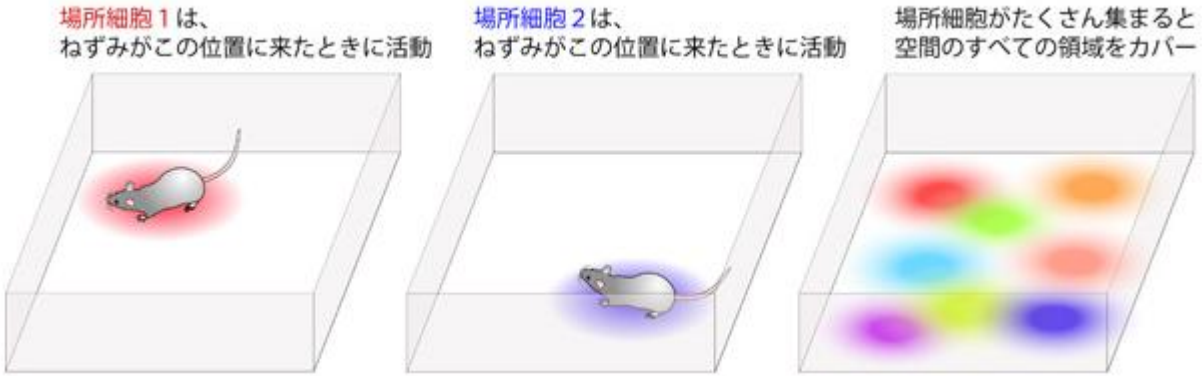
Spatial recognition and neural activation

Behavior



[Tolman, 1948]

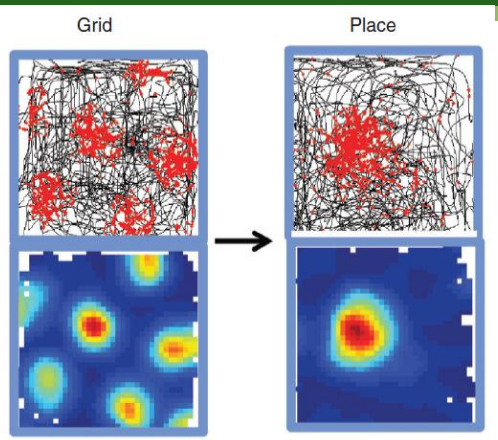
Neural activation



[日本神経科学会web]

Place Cell

- **Place cell**:
  neurons activated when animals enter in a particular place.
  [O'Keefe & Dostrovsky, 1971]
- **Grid cell**:
  neurons activated at an equal distances.
  [Fyhn et al., 2004, Hafting et al. 2005]



[Moser et al. 2015]

- Mathematical models have been proposed that retain spatial representation.
- Rats and infants are never told where they are.
- No model where spatial representation is built from only visuo-motor experiences.

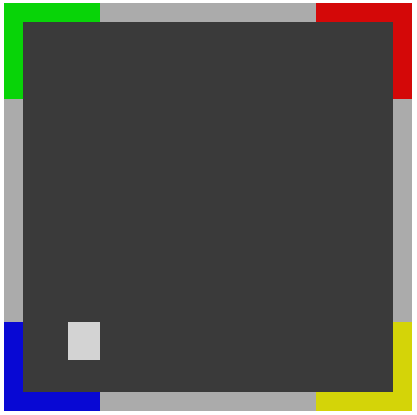
Simulation-based synthetic approach using deep neural networks

Noguchi, Iizuka, Yamamoto (2017), Cognitive map self-organization from subjective visuomotor experiences in a hierarchical recurrent neural network, Adaptive behavior 25(3), 129-146.

Predictive learning

- **Simulation environment**

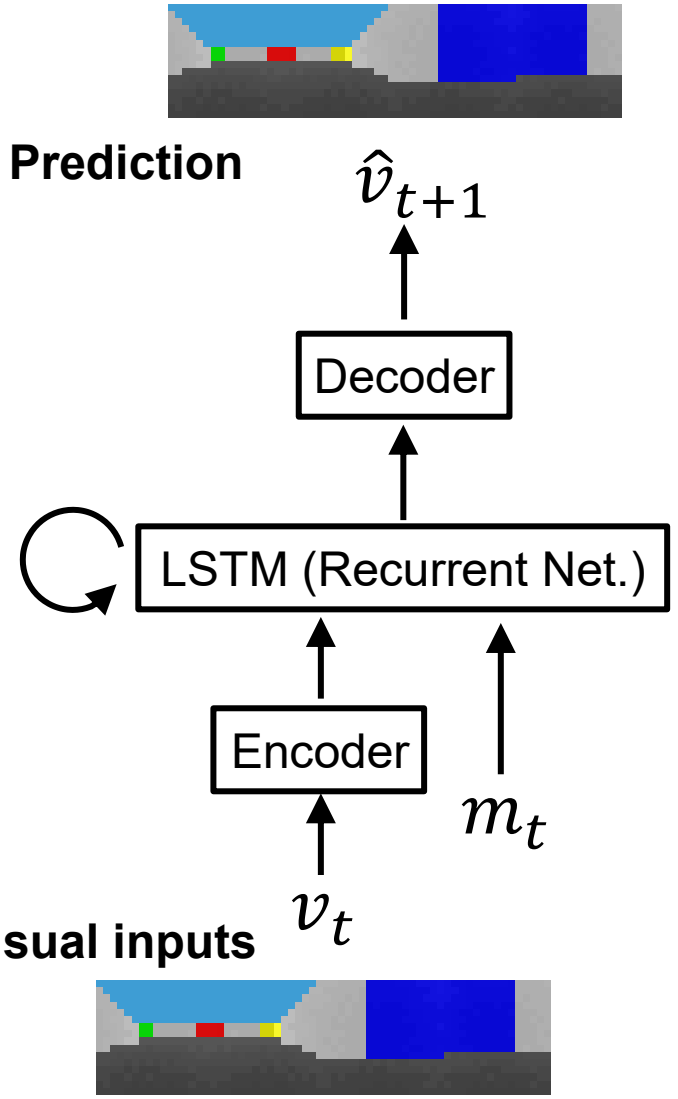
Agent : a robot moving around 2-D space



➢ **Four landmarks at the corners**

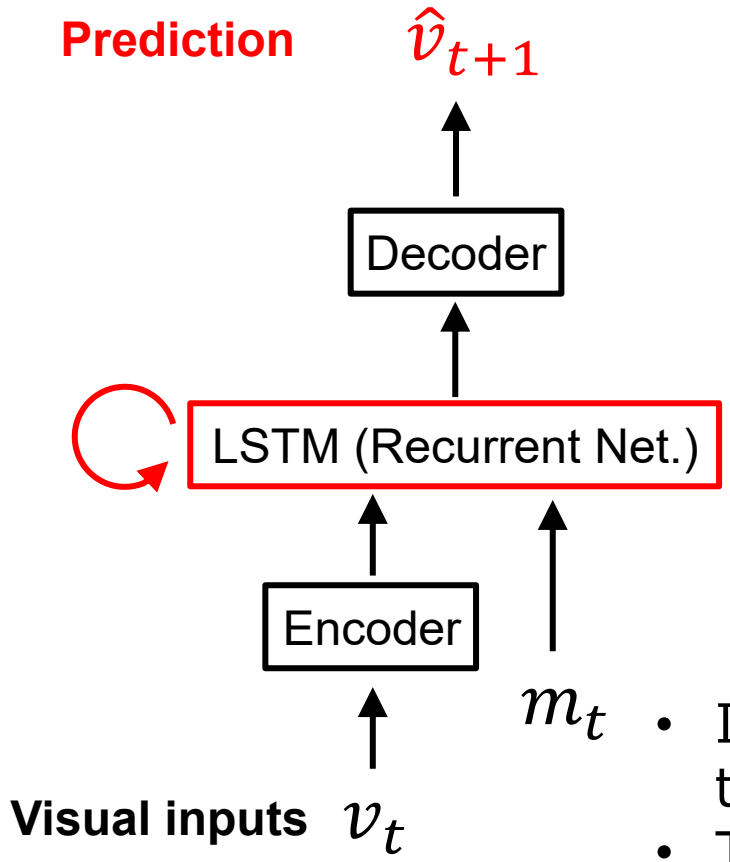
The neural network is trained to predict future visual inputs.

- **Neural network model**



Visualizing internal states

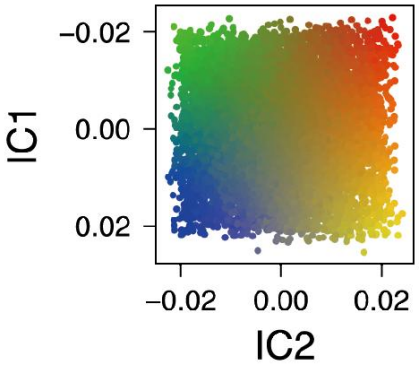
Prediction performance

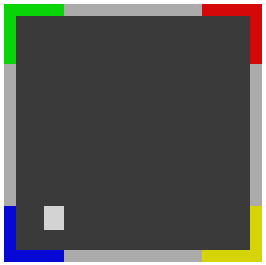


Ground Truth

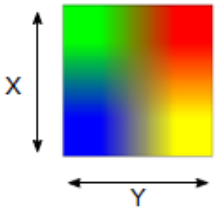
Prediction

Visualized internal states





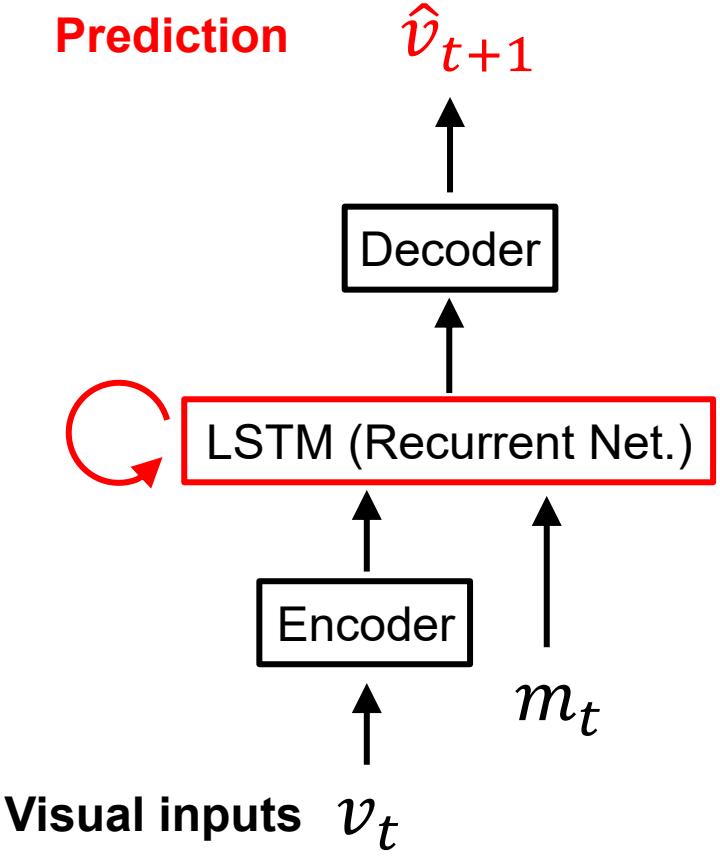
Color map



- Internal states are stored in such a way that they correspond to the locations.
- Two dimensions enough same as spatial dimensions.

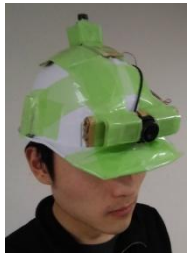
Real-world experiments

Prediction performance



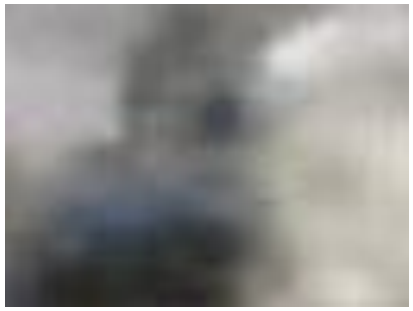
Ground Truth

acc.

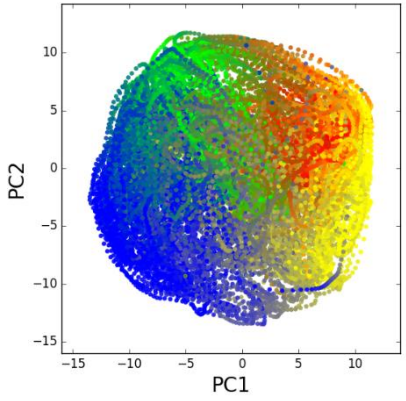


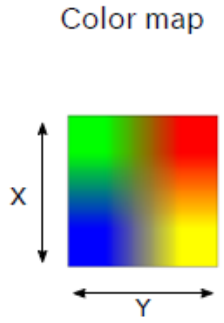
camera

Prediction



Visualized internal states





- **Distinction of self and others through superposition mechanism**
- **Propagating a mind from other to self via superposition mechanism**

How do we predict and explain other's behavior?

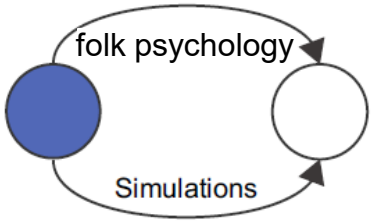
- Theory theory [Gopnik, 1994 …]
  - Using implicit general knowledge called folk psychology
  - Knowledge-based inference using beliefs and desires

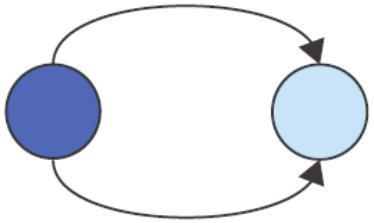
- Simulation theory [Gordon, 1986 …]
  - Using its own model to simulate other's mental states.
  - Model-based inference

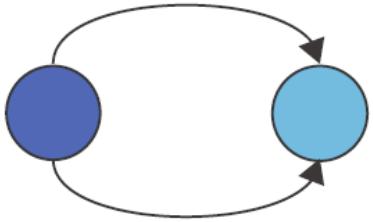
- Interaction theory [Gallagher, 2008 …]
  - Criticizing the view of unilateral observation of others and making inferences about others' mental states
  - Ongoing interaction itself causes direct perception of others

Illustration of those approaches

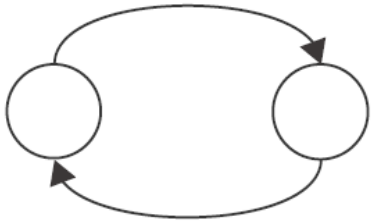
Theory theory
Simulation approach

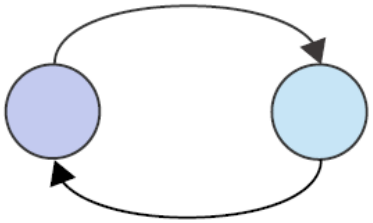


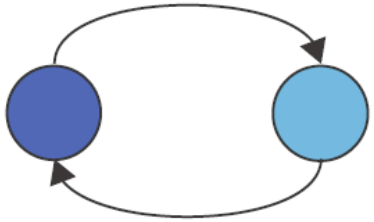




Interaction theory

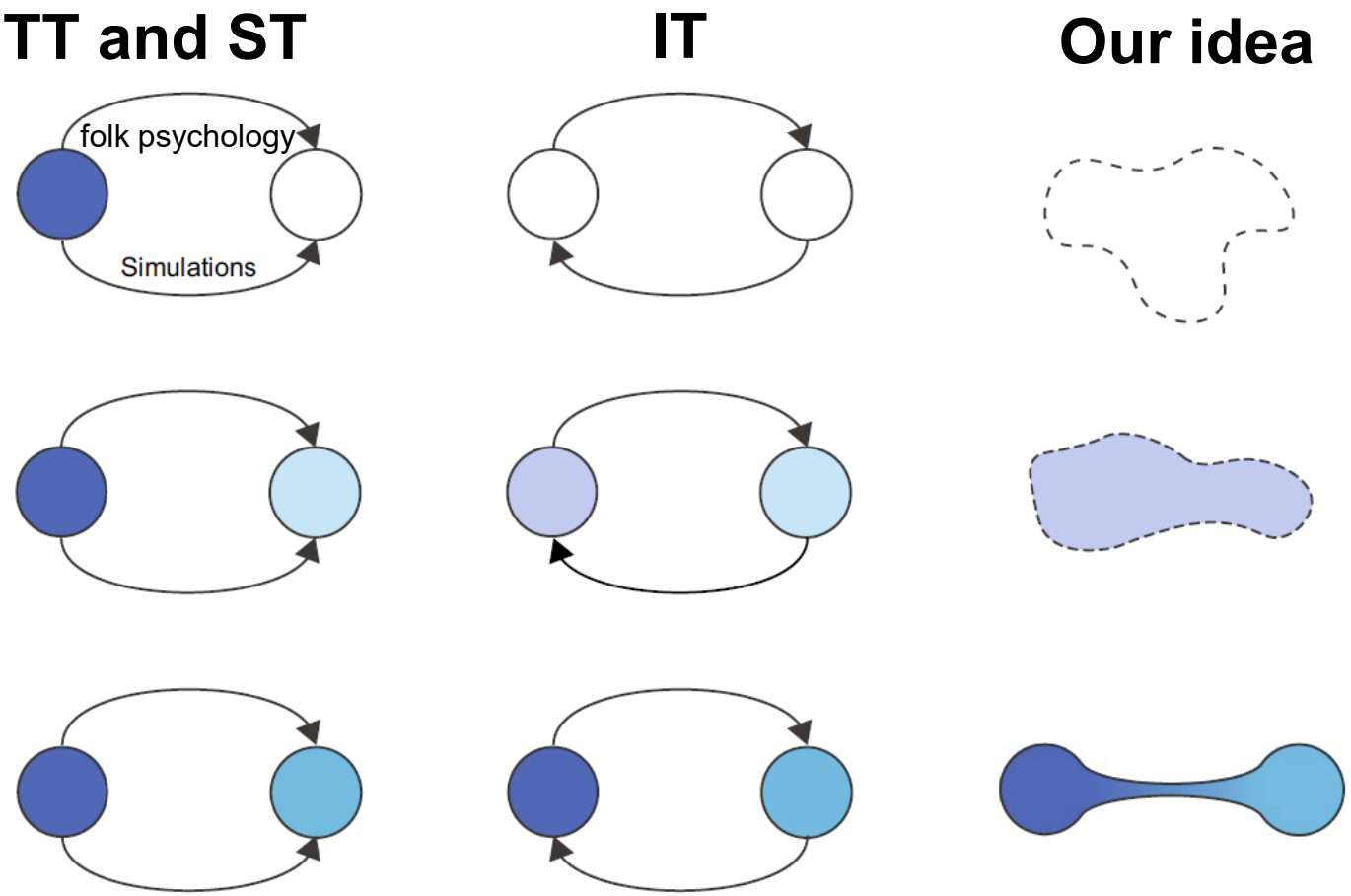






How can the pre-existing frameworks of self and others be formed?

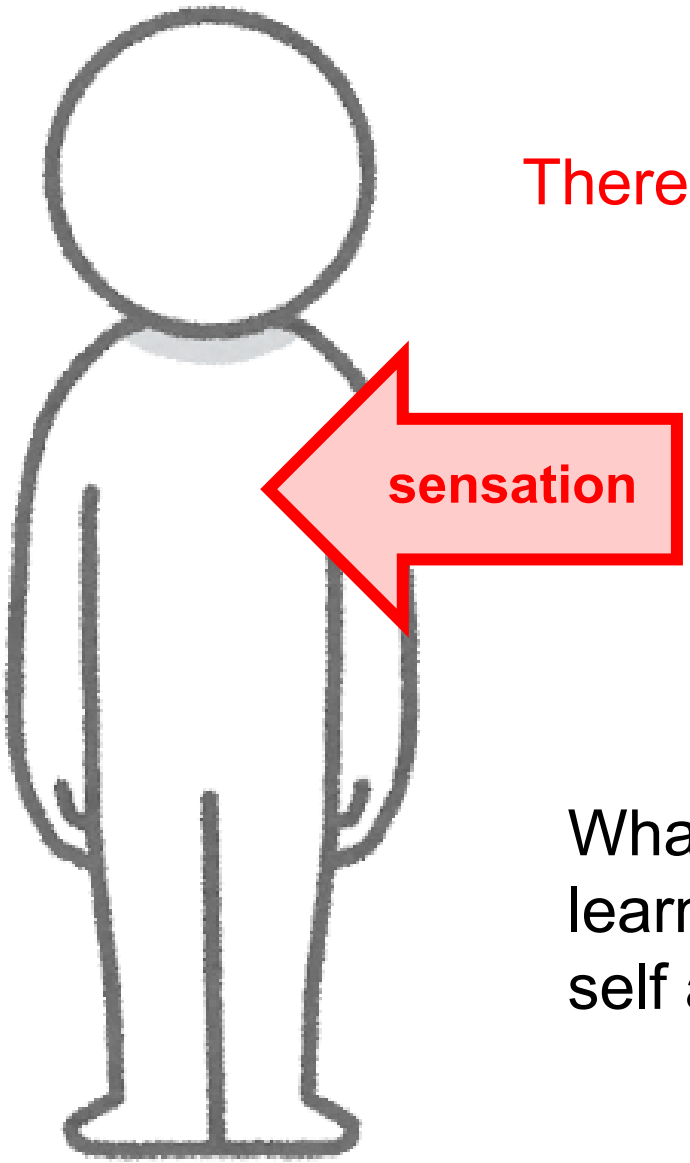
No pre-existing frames of self and other



We assume

- there are no pre-existing frames of self and other.
- these concepts themselves are acquired through learning.
- starting from a situation where there is neither self nor other.

How we understand self and others?



There is no distinction between self and other.

Sensations are nothing more than external stimuli, even if they are considered belonging to a particular body from the third-person perspective.

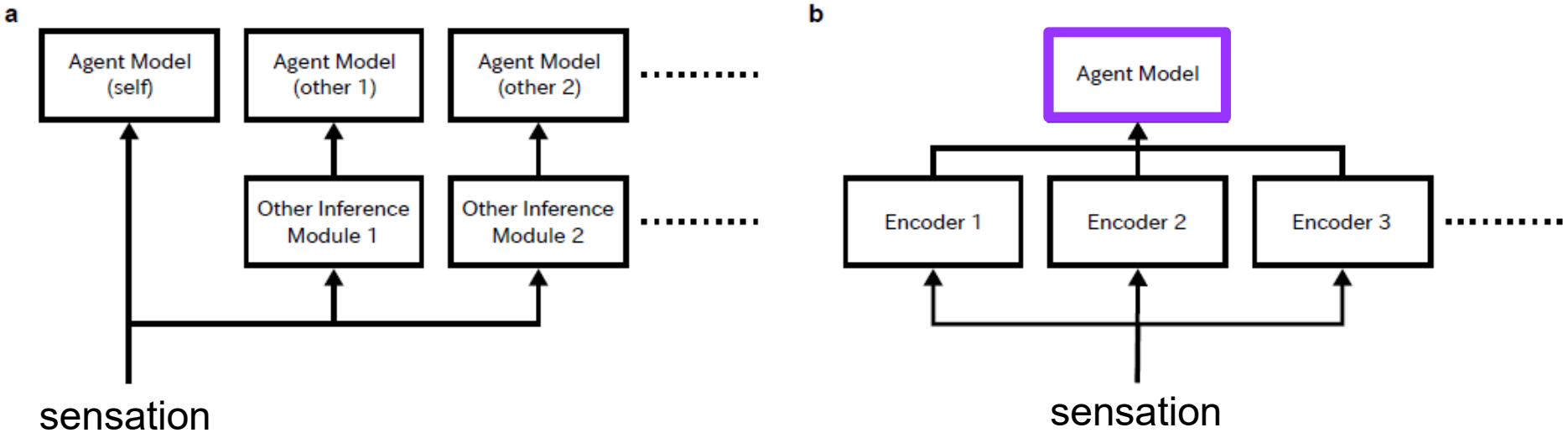
What mechanism makes it possible to learn and obtain the representations of self and other?

Anyone model that doesn't differentiate between self and others

Instead of creating models of others, we assume a single model for anyone without distinction between self and others.

Building a model of self and others separately

Building a single model that doesn't differentiate between self and others

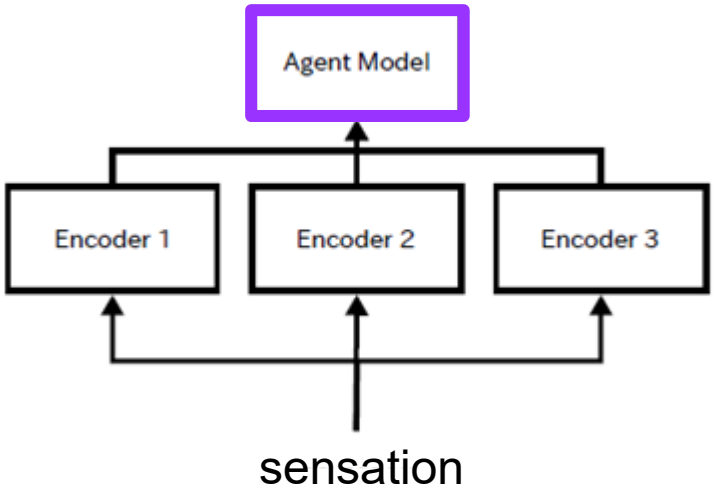


c.f. agent-neutral coding by Gordon (originator of ST)

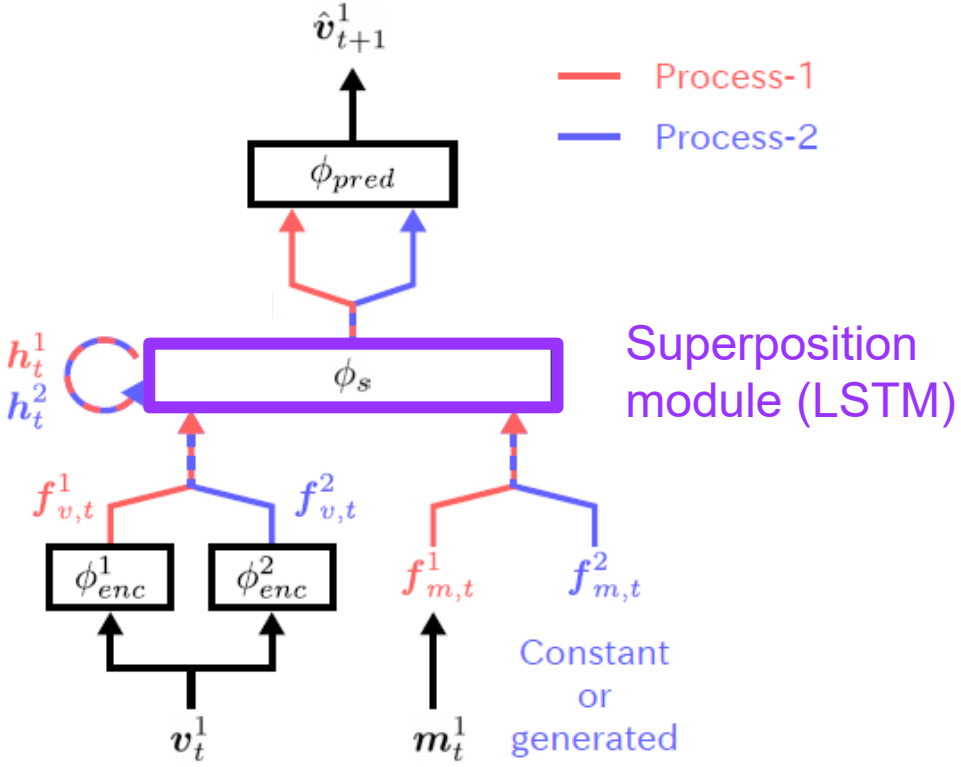
Gordon, R. M. (2021). Simulation, predictive coding, and the shared world. In *The Neural Basis of Mentalizing* (pp. 237-255).

Superposition network model

- Receiving own sensory (visual and somatosensory) inputs
- Multi-stream processing, process-1 and process-2
- Processed by a superposition module
- Integrating outputs of both processes to predict sensory inputs

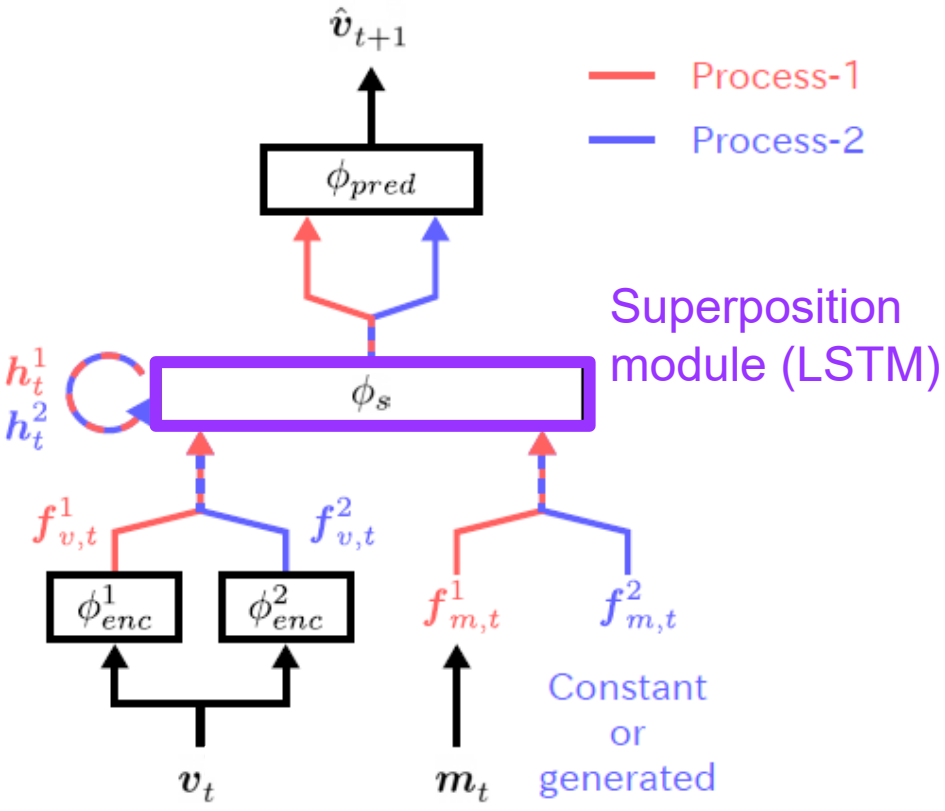


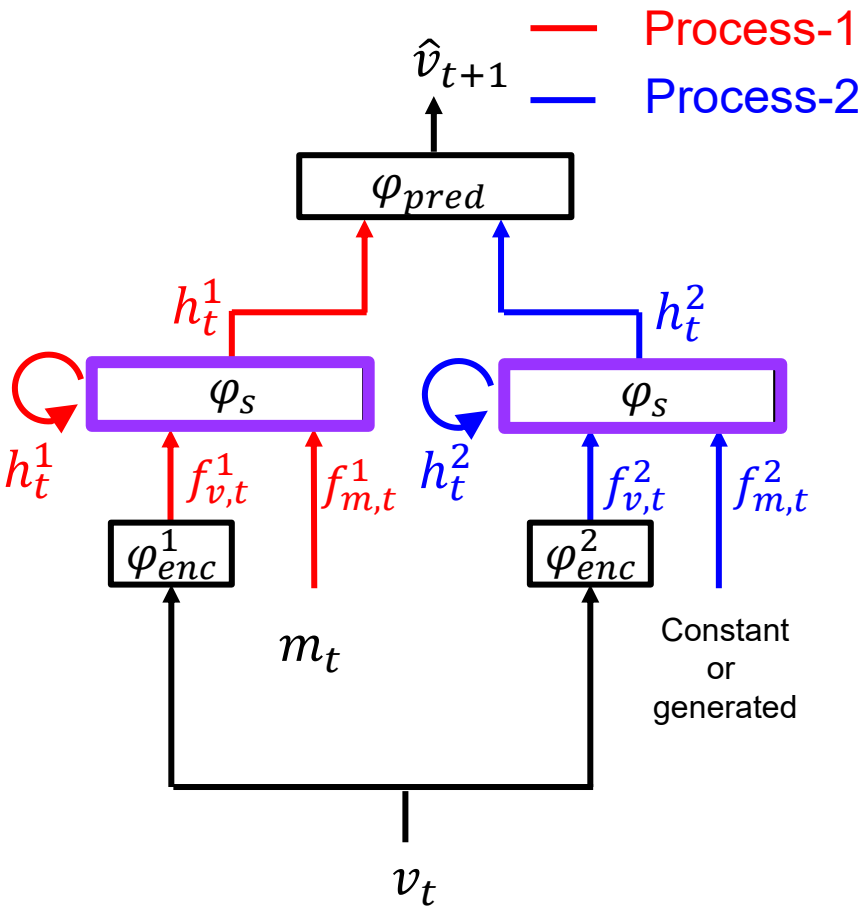
Anyone model (concept)



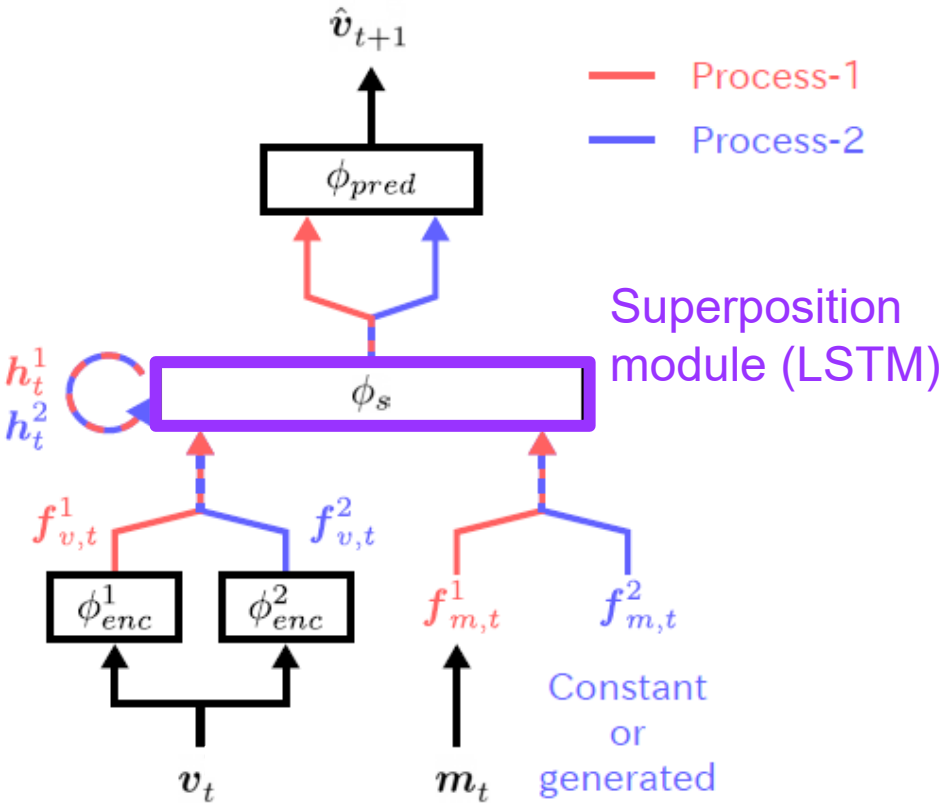
Superposition network model

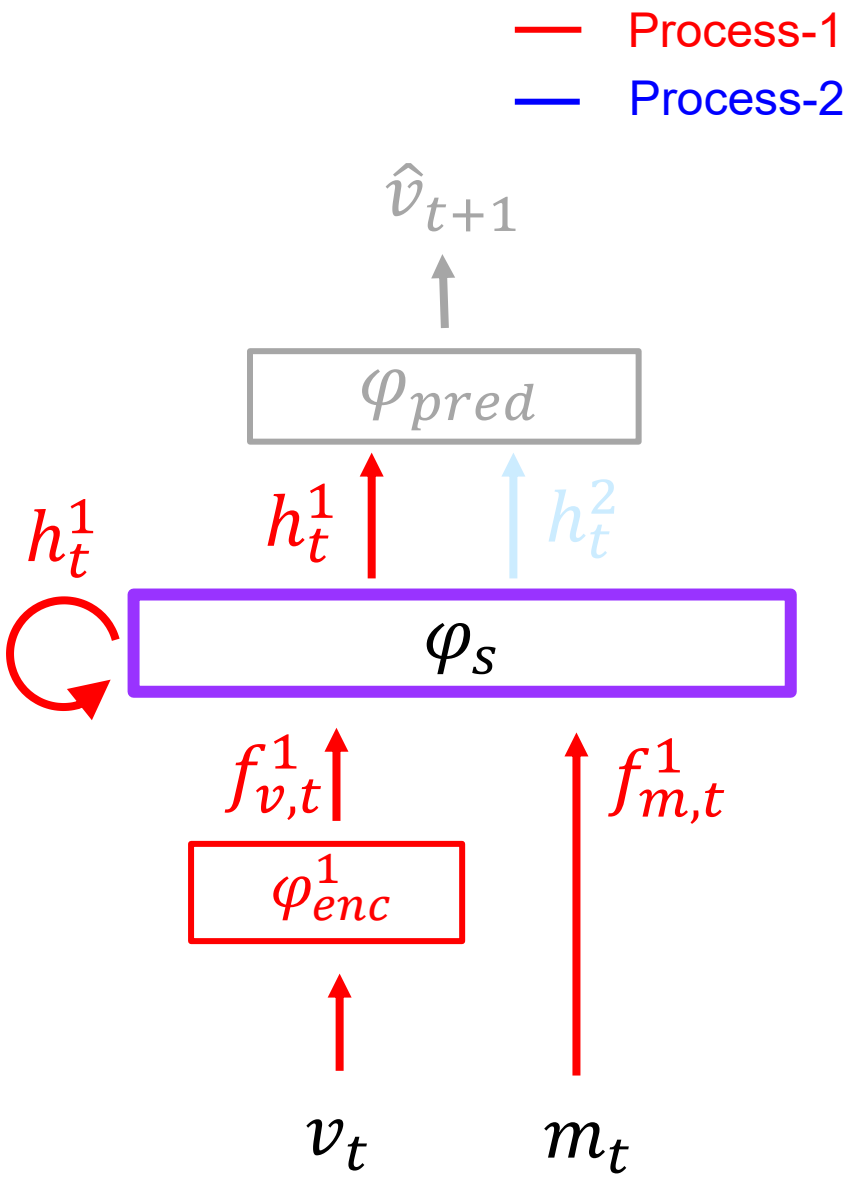
Another illustration form of our model



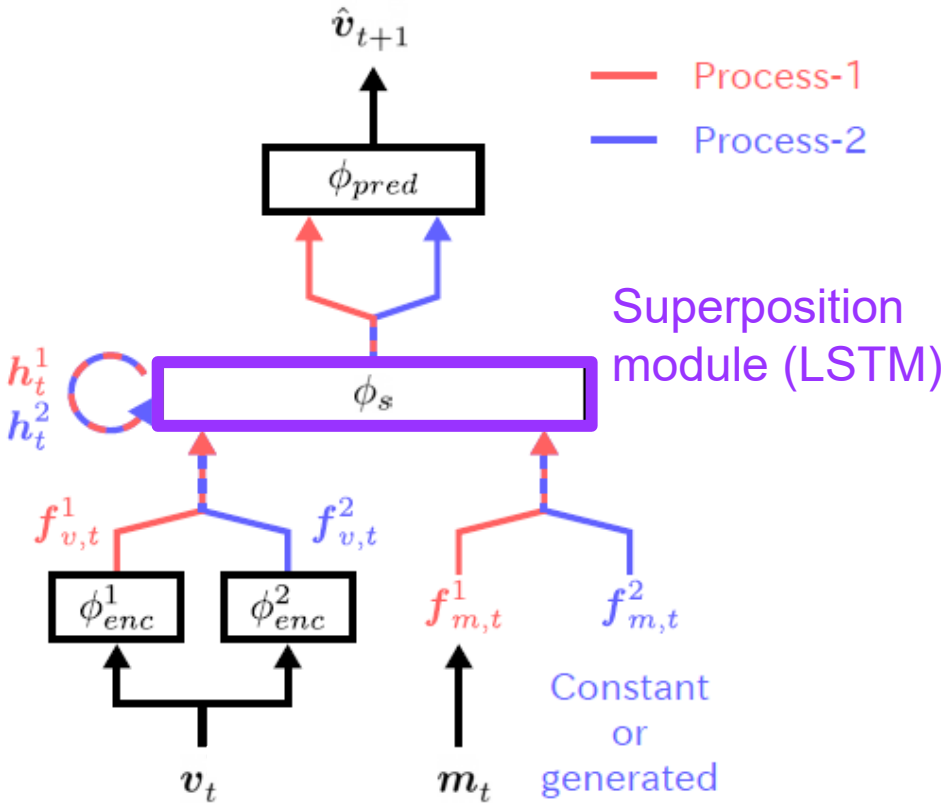


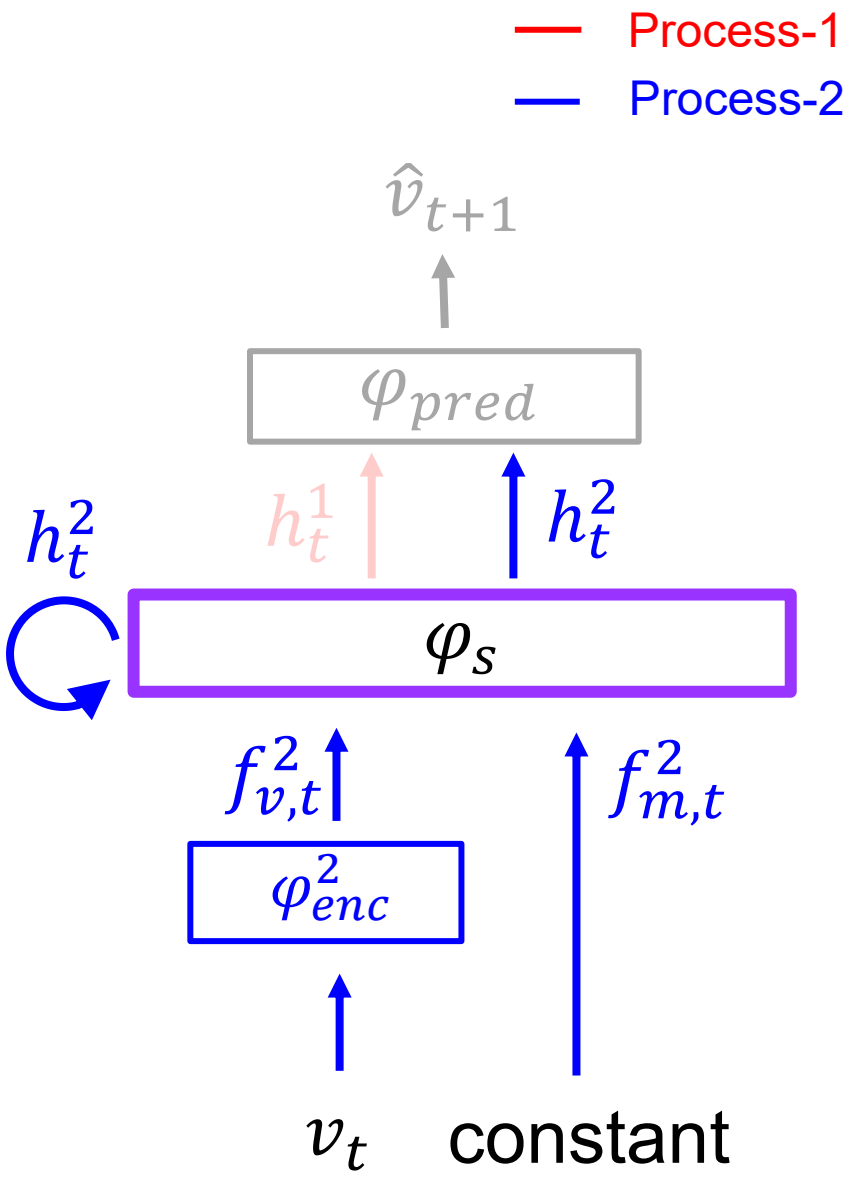
Another illustration form of our model



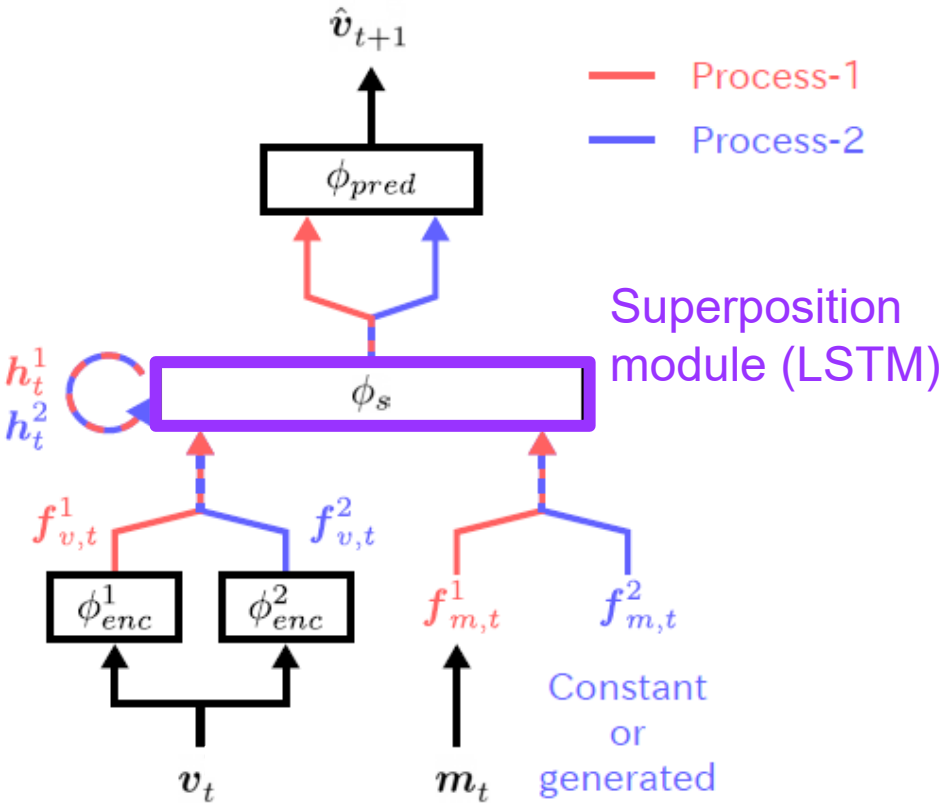


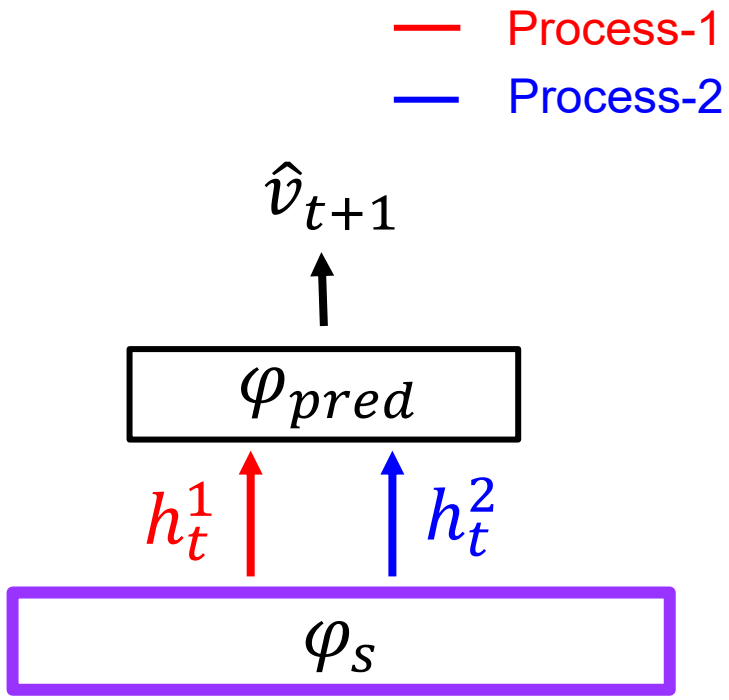
Another illustration form of our model





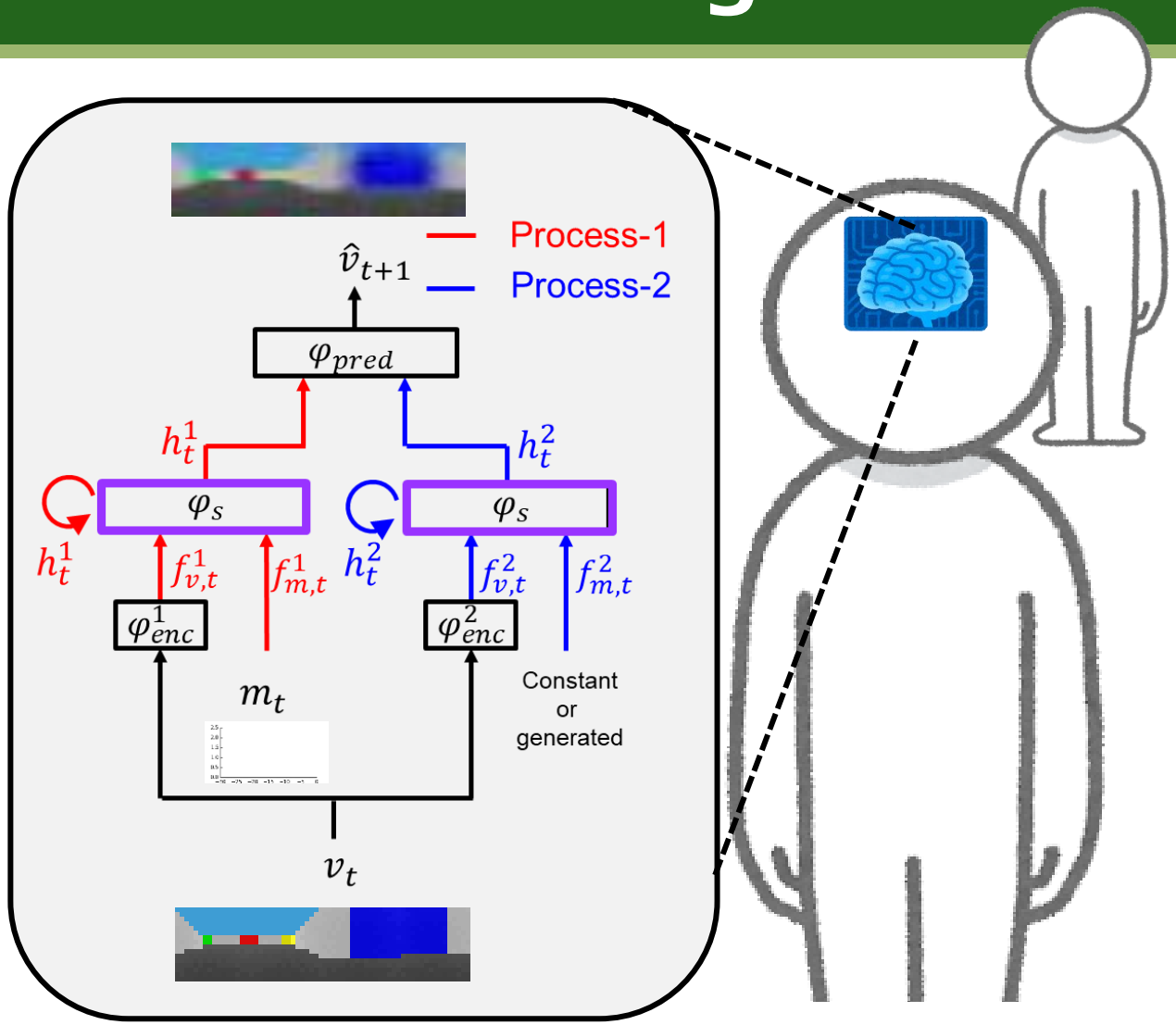
Another illustration form of our model





Predictive learning

The network is trained to minimize the prediction error for vision.



Loss function

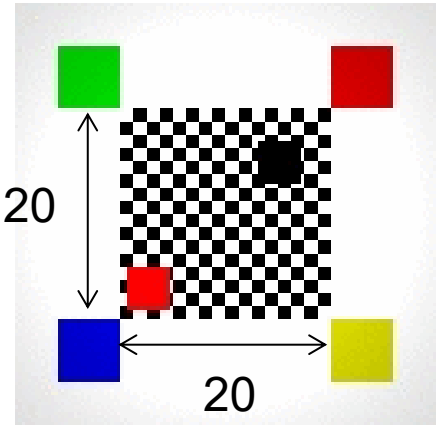
$$\mathscr{L}_{pred} = \sum_{t=1}^{T} \left\{ \lambda_{\text{MSE}} \text{MSE}(\hat{v}_{t+1}, v_{t+1}) + \lambda_{\text{MAE}} \text{MAE}(\hat{v}_{t+1}, v_{t+1}) \right\},$$

Simulation environment

■ : Agent-1 ■ : Agent-2

Agent-1 and agent-2 exist

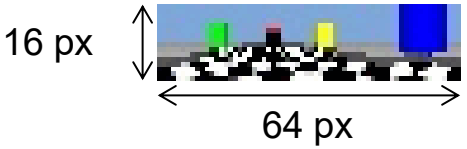
- Agent-1 moves around in the arena.
- Agent-2 does not move but change its locations every trial.



Agent-1:

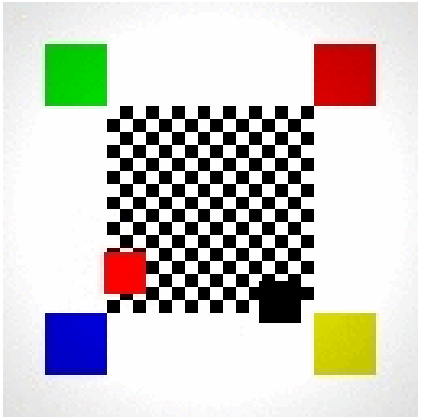
- Equipped with the superposition network
- Receiving visual images and proprioception
- Predicting future visual images

Agent-1's vision



Prediction performance

Example of agent-1's trajectory

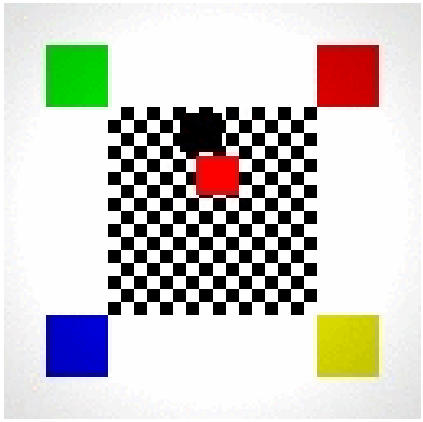


visual prediction

Truth

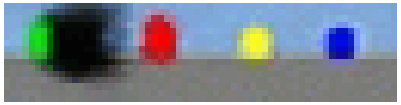
predict

Another example



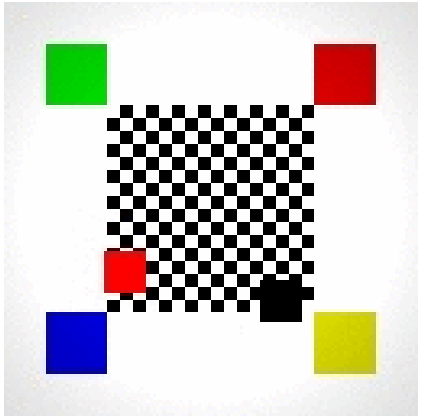
Truth

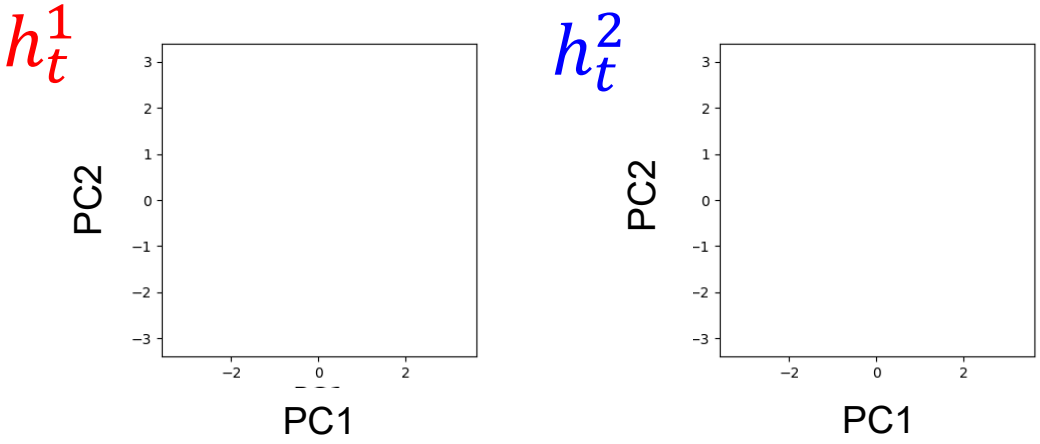
predict

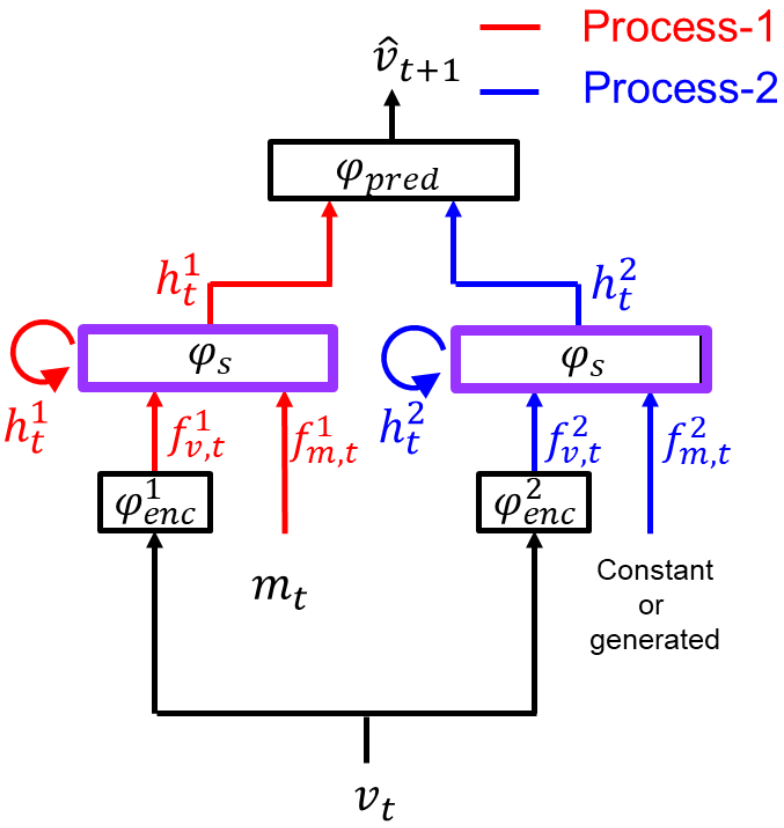


Visualizing internal states of SM

Example of agent-1's trajectory



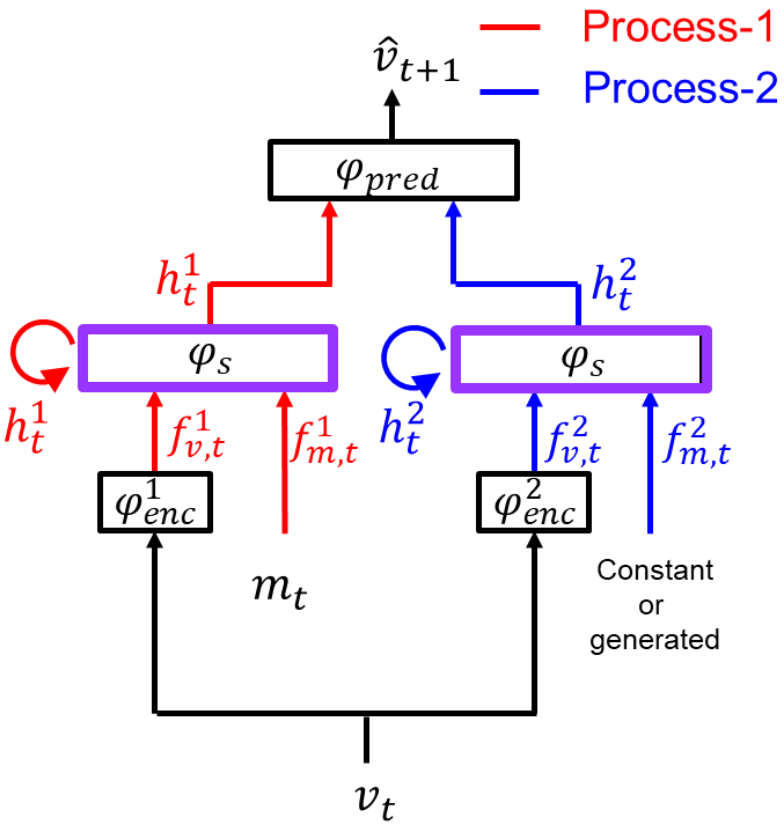




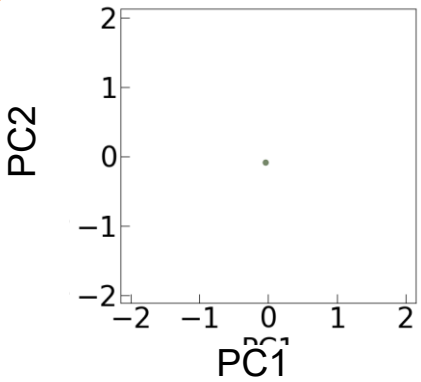
Coloring : agent-1's positions

agent-2's positions

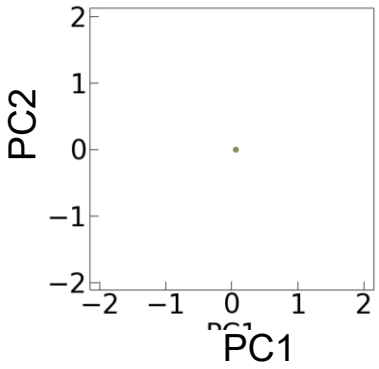
Visualizing internal states of SM



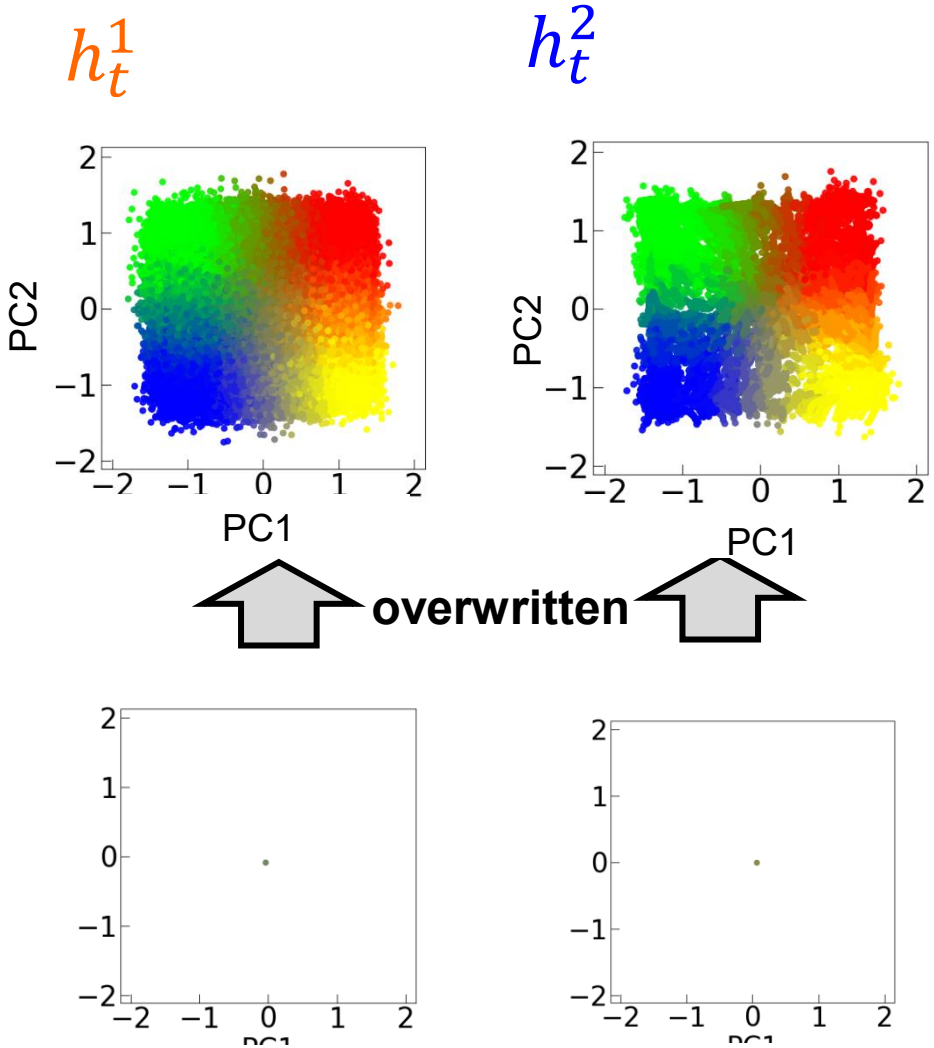
h_t^1

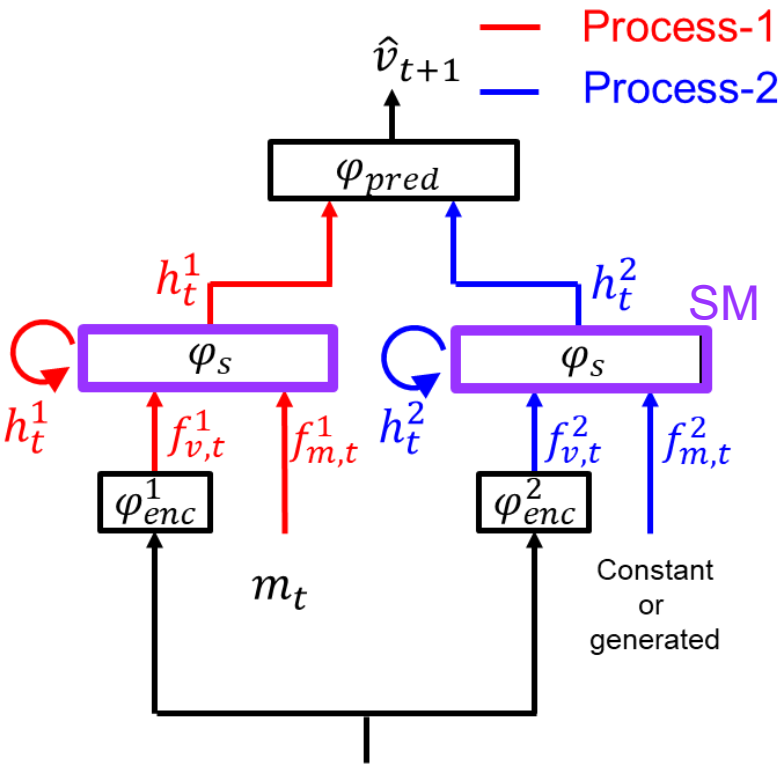


h_t^2



Visualizing internal states of SM





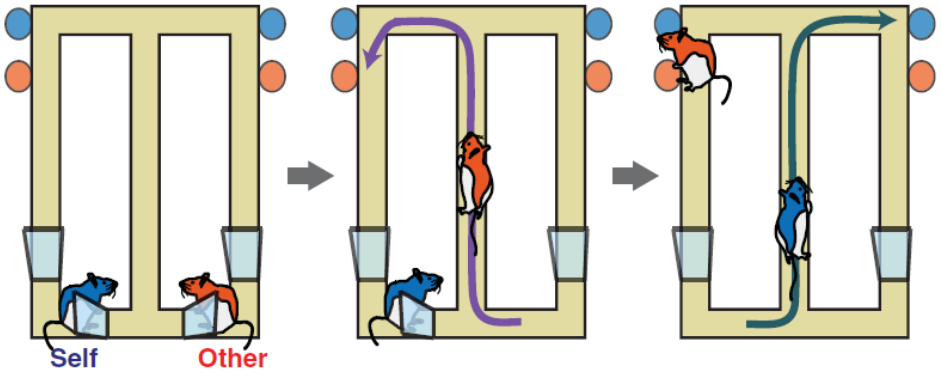
- The internal states of SM are activated in response to A-1 and A-2's positions.
- Similar to the social place cell found in neuroscience.

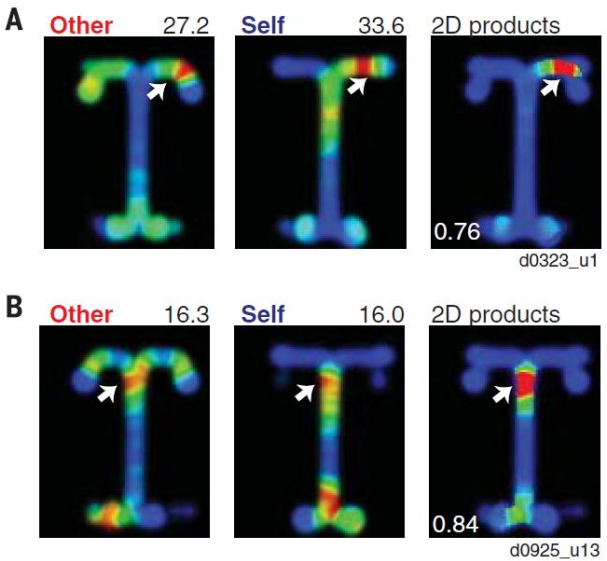
[Danjo, et. al, 2018, Omer, et. al, 2018]

Discovery of Social Place Cells

Spatial representations of self and other in the hippocampus

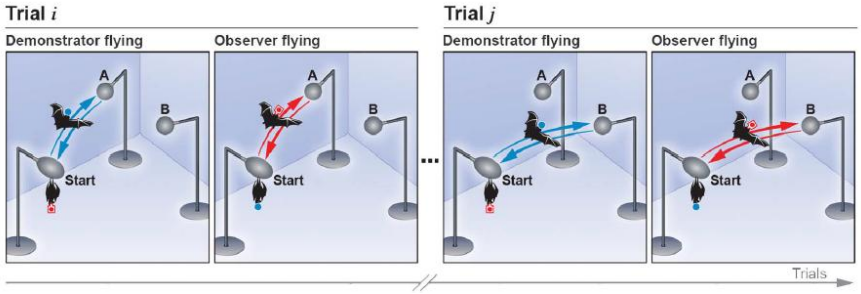
[Danjo et al., 2018]

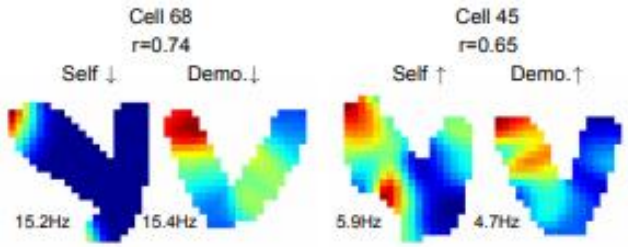


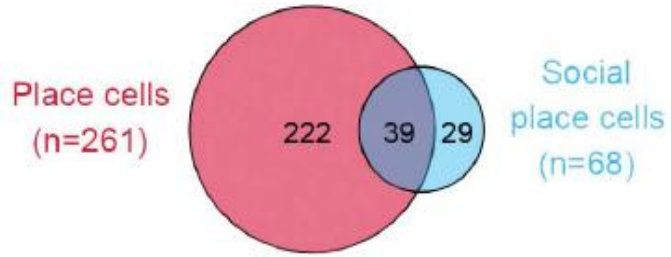


Social place-cells in the bat hippocampus

[Omer et al. 2018]

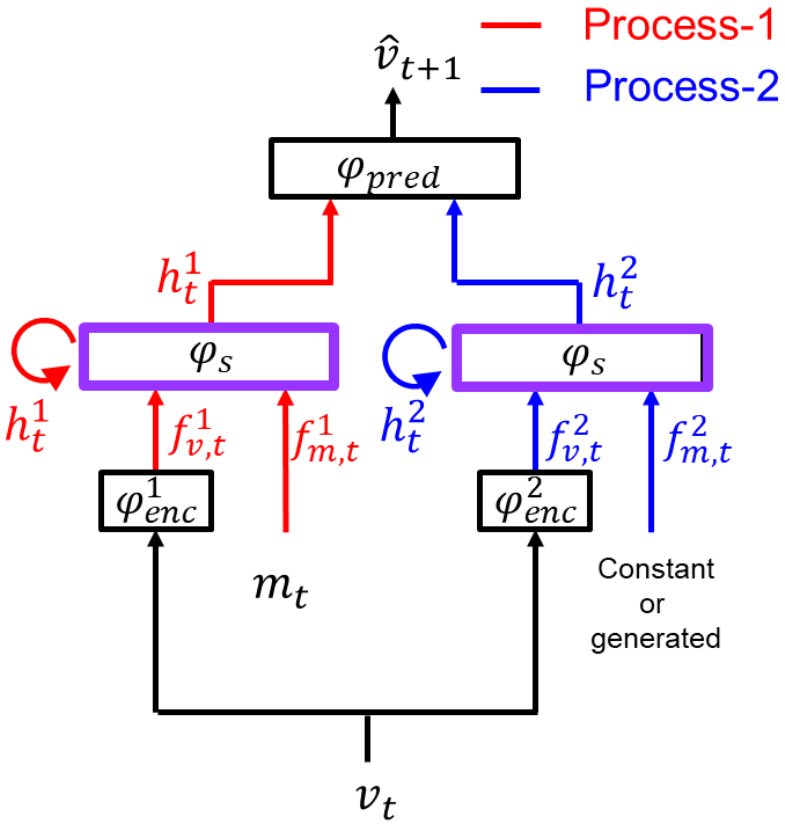


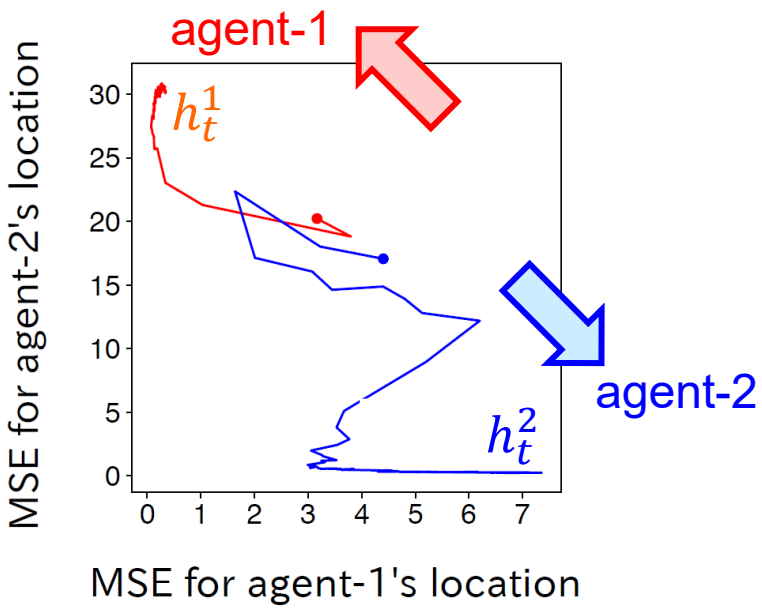




Self / other separation

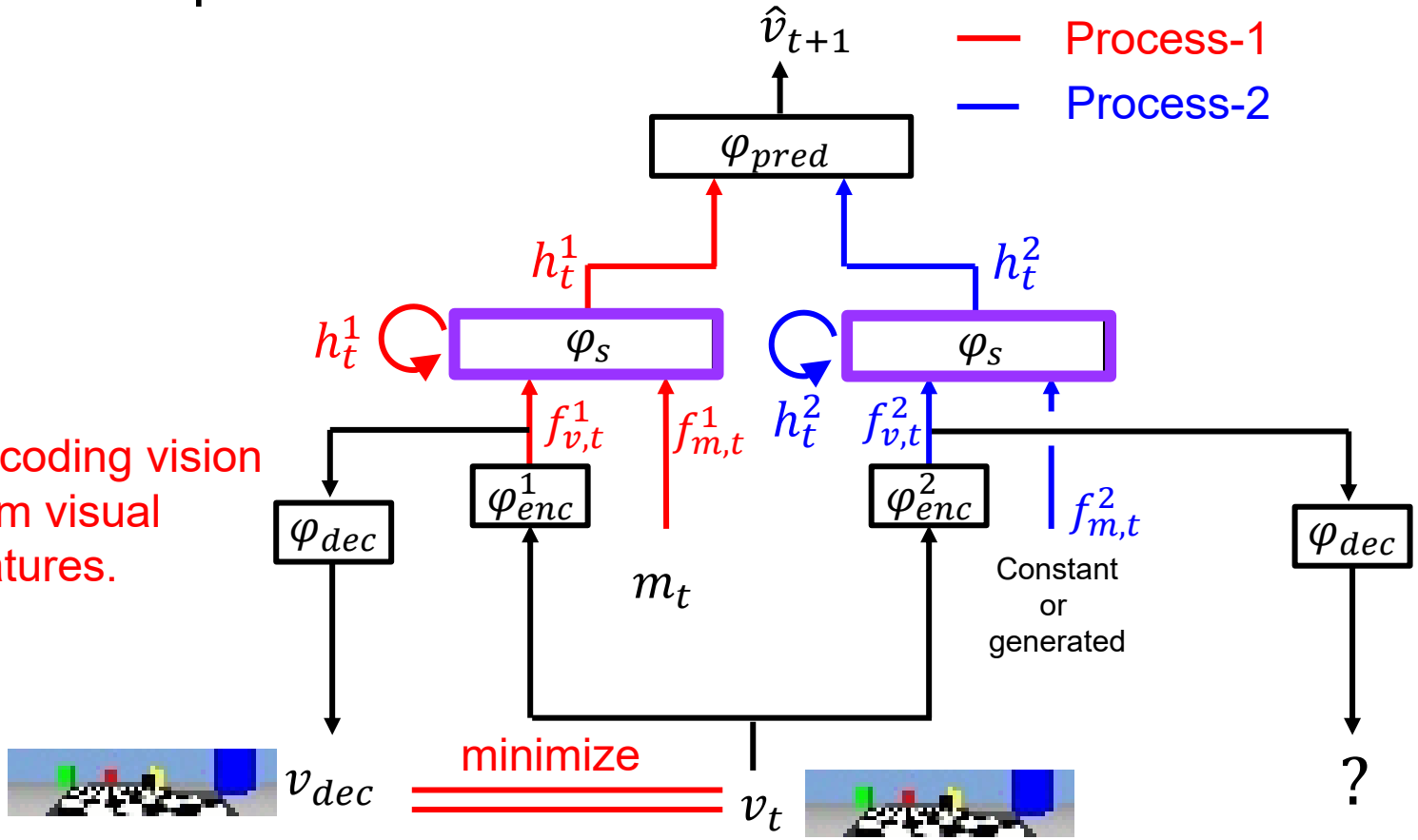
Process-1 and -2 are gradually specialized to self and other, respectively. Process-1 handles information about self and process-2 handles information about other.





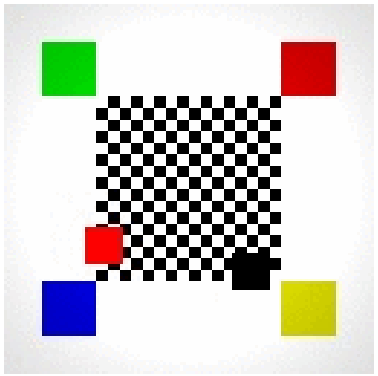
Visualizing what agent-1 sees

- Building a decoder to reconstruct the visual image from (own) visual features on process-1.
- Applying the decoder to the visual features on process-2 and visualize what it sees.



Visual perspective taking

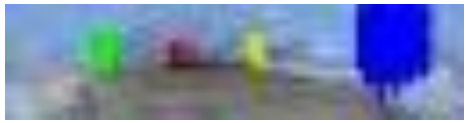
Example 1



Process-1

A-1's vision

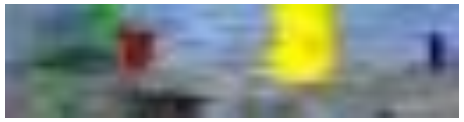
Decode

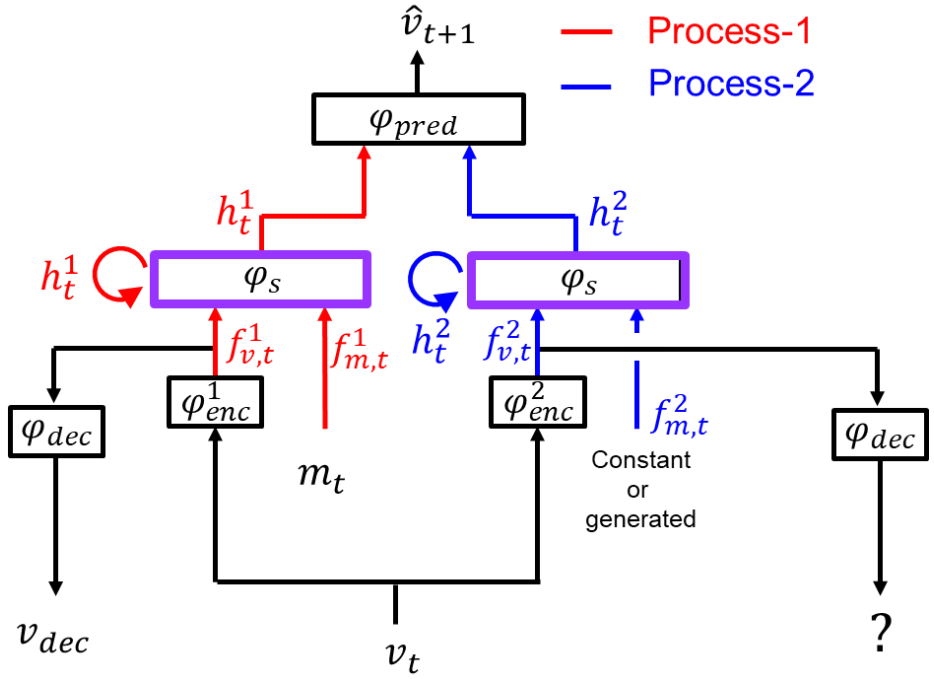


Process-2

A-2's vision

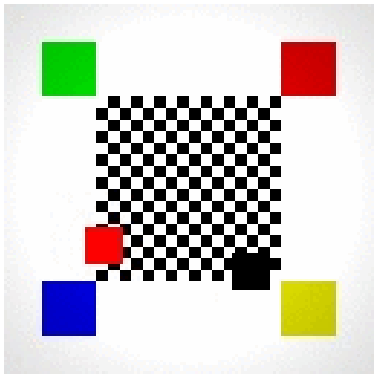
Decode





Visual perspective taking

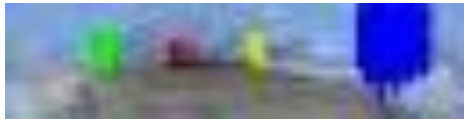
Example 1



Process-1

A-1's vision

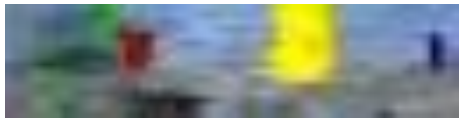
Decode



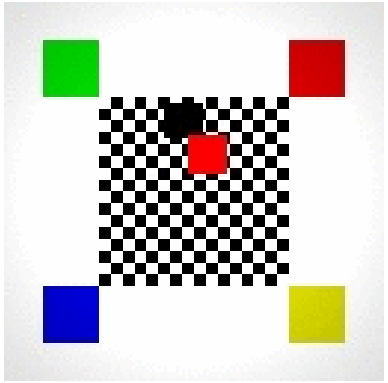
Process-2

A-2's vision

Decode



Example 2



Process-1

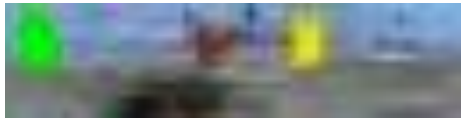
A-1's vision

Decode

Process-2

A-2's vision

Decode



Just by using the superposition module, the agent can naturally obtain other's perspective.

Training motion generator

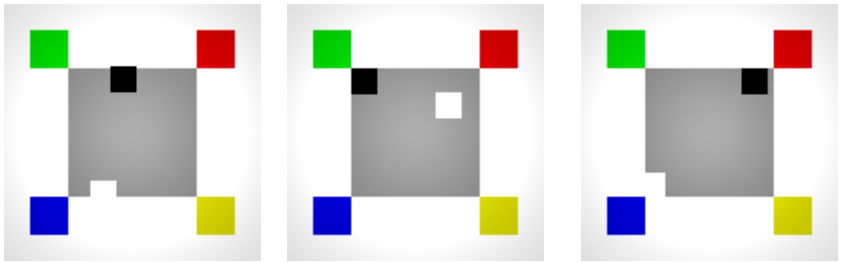
■ Agent-2 has intentions and begins to move

(So far, only agent-1 was movable)

- Agent-2 moving around
- Adding a motion generator (MG(LSTM))
- Training the MG via predictive learning

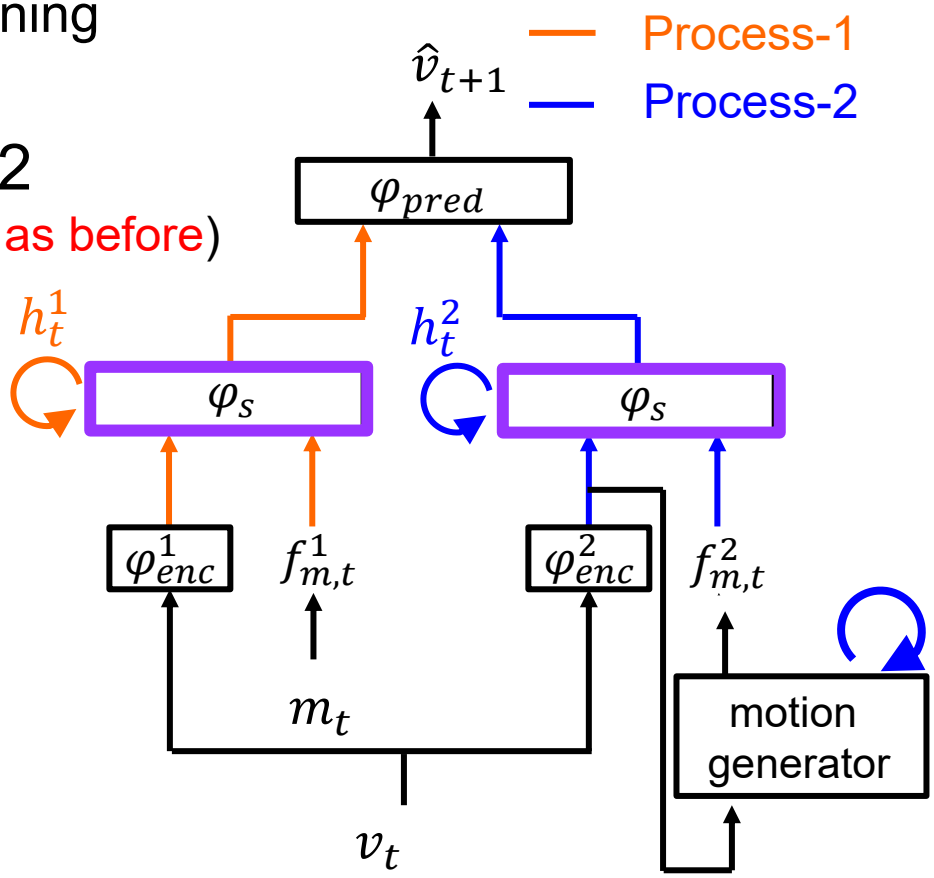
■ Three motion patterns of Agent-2

(Agent-1 : Moving in a random way same as before)



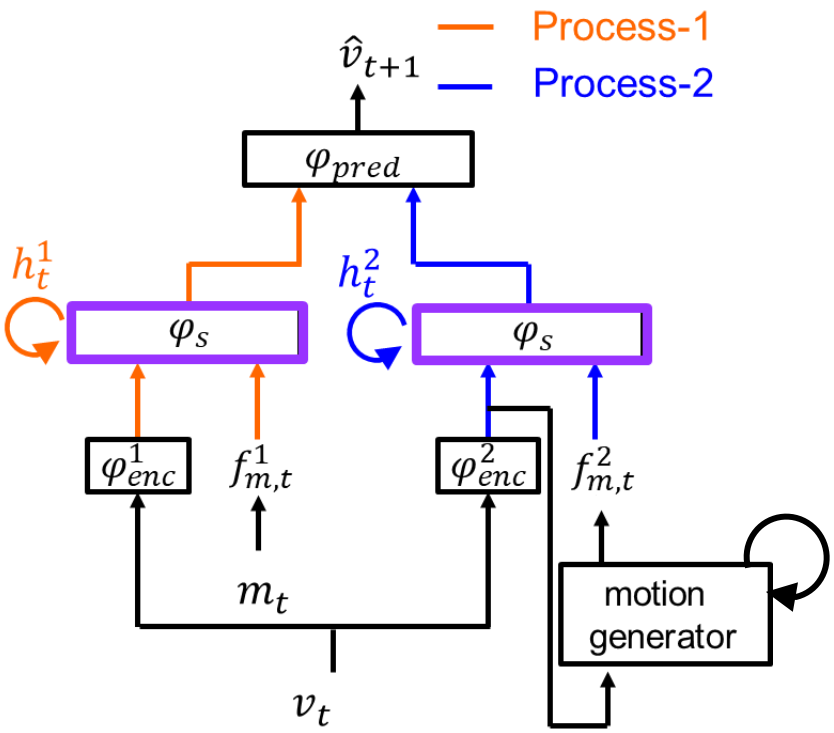
Stopping CW CCW

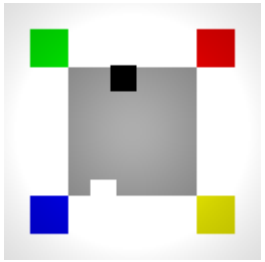
White : Agent-1 Black : Agent-2

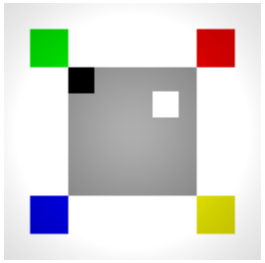


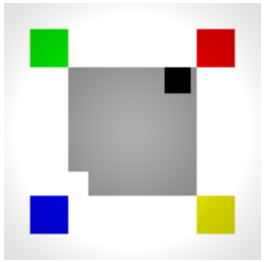
Mirror neuron-like activations

Motion commands that create A-2's movements as if A-1 were A-2.

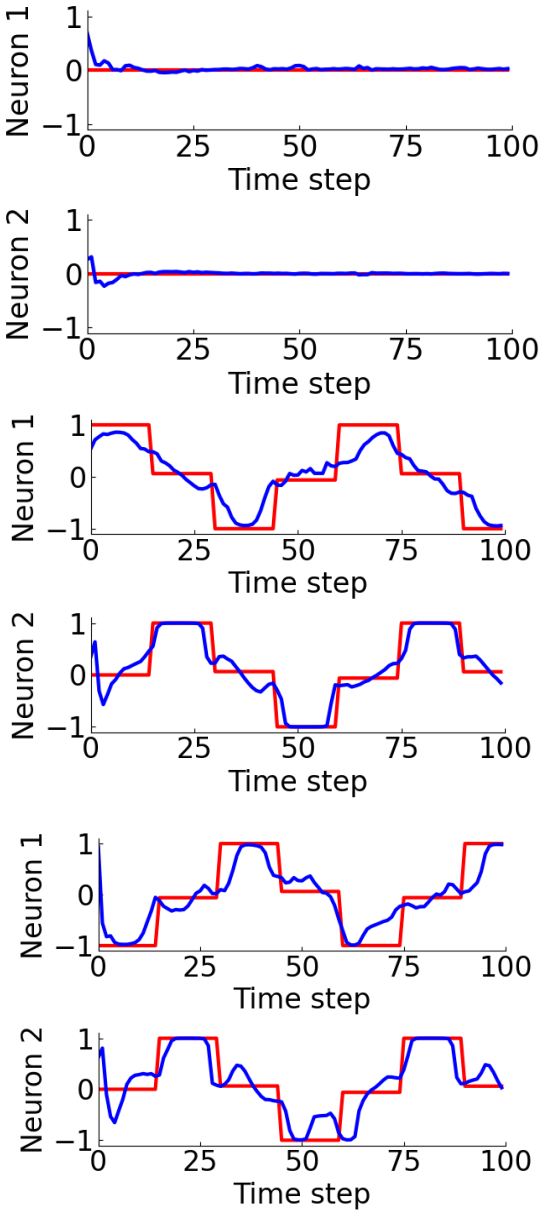








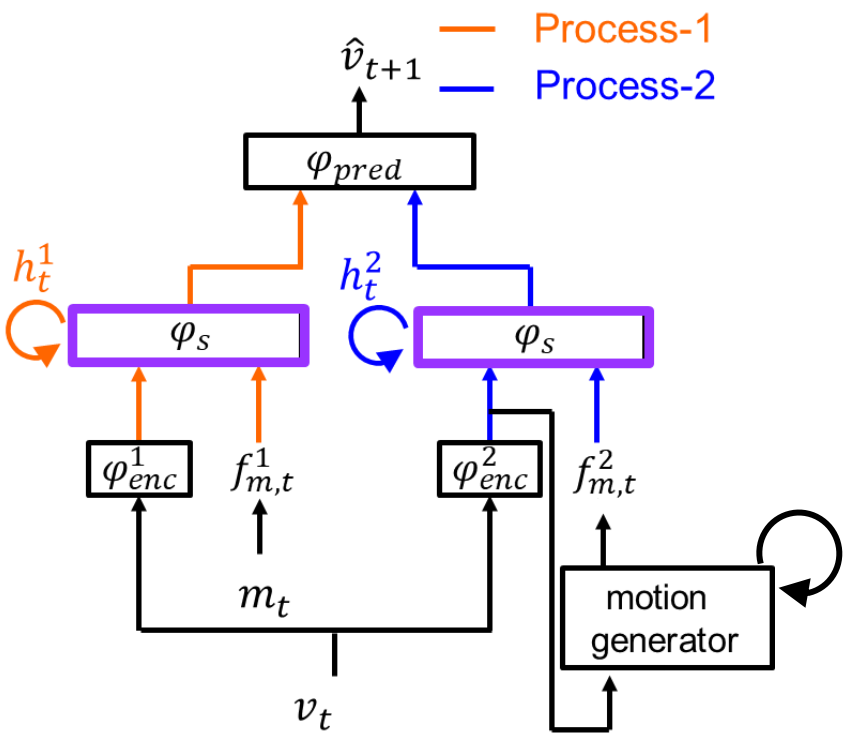
White : Agent-1
Black : Agent-2

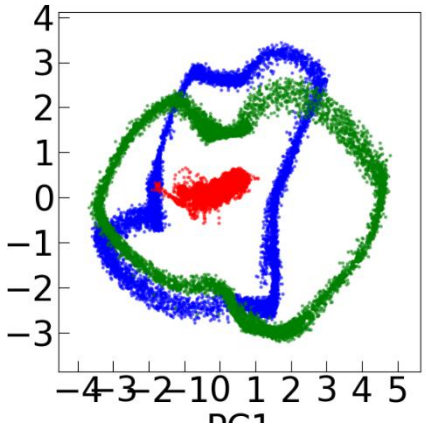


- When observing agent-2, motions to generate agent-2's movements are internally created within agent-1.

Internal states of motion generator as intentions of the other

- Visualizing the internal states of the motion generator by color-coding them for each action of Agent-2.



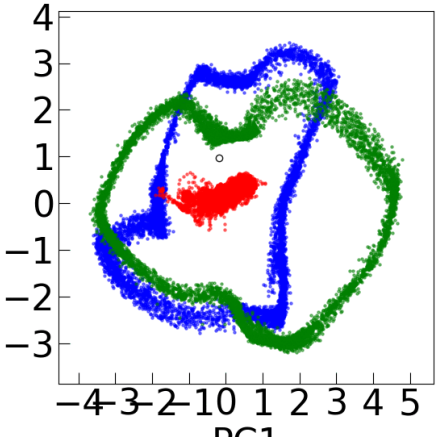


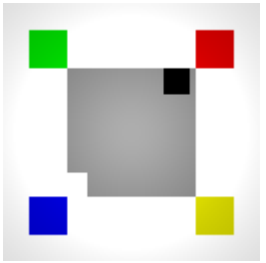
Stopping

CW

CCW

CCW

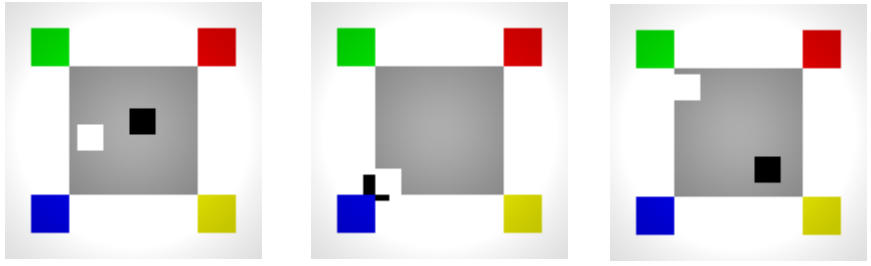




As a result, the attractors corresponding to the behavior patterns were formed.

Using MG of the other as self's

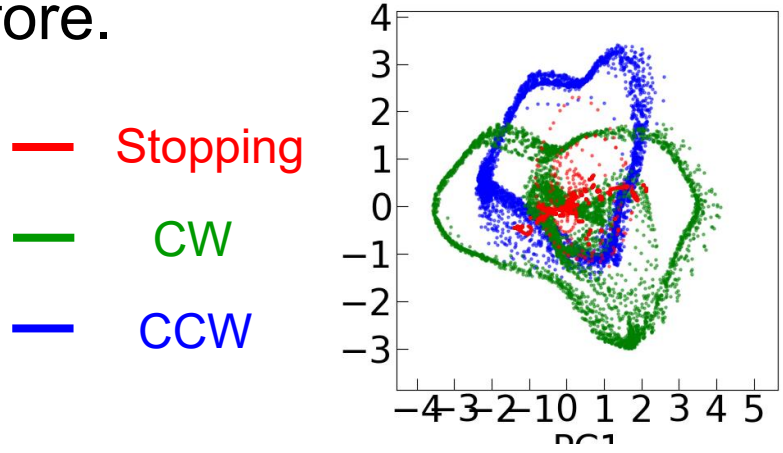
- Although these are patterns agent-1 has never moved in before, agent-1 tries moving in the three patterns

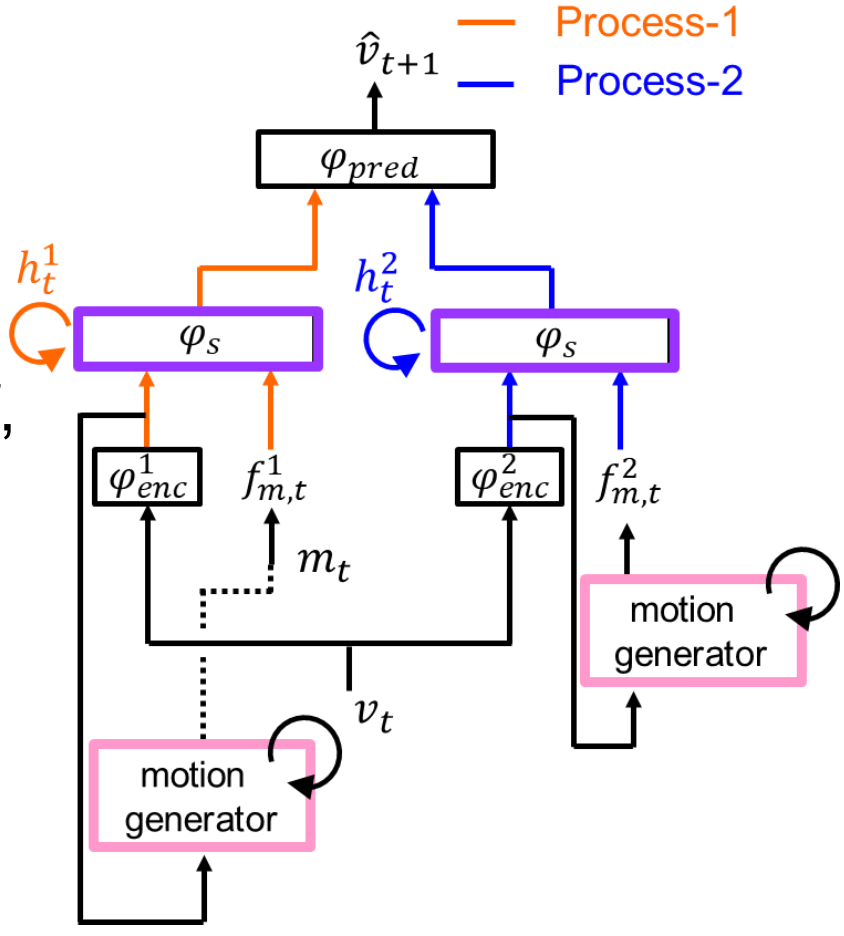


Stopping CW CCW

White : Agent-1 Black : Agent-2

- Applying the MG of the other to self, visualizing the internal states same as before.





Superposition mechanism (shared module network)

Multisensory integration model using superposition

- Wataru Noguchi, Hiroyuki Iizuka & Masahito Yamamoto (2022) Multi-modal shared module that enables the bottom-up formation of map representation and top-down map reading, Advanced Robotics, 36:1-2, 85-99
- Wataru Noguchi, Hiroyuki Iizuka & Masahito Yamamoto (2019). Navigation behavior based on self-organized spatial representation in hierarchical recurrent neural network, Advanced Robotics, 33:11, 539-549

Symbol grounding problem

- Yasuhiro Shimada, Wataru Noguchi, Hiroyuki Iizuka, Masahito Yamamoto, Bottom-up formation of number representation and top-down understanding of symbolic manipulation, Proceedings of the ALIFE 2022, pp. 337-344 (2022)

Tool-body assimilation

- Haruna Kasashima, Wataru Noguchi, Yasumasa Tamura, Masahito Yamamoto, Hiroyuki Iizuka (2024). A Deep Learning Approach to Understanding the Integration of Tools into Body Schema , ALIFE 2024: Proceedings of the 2024 Artificial Life Conference

Thank you !

Article | Open Access | Published: 21 February 2022

Superposition mechanism as a neural basis for understanding others

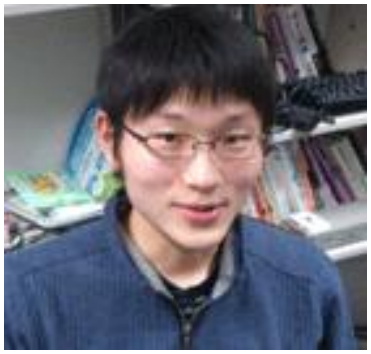
Wataru Noguchi ✉, Hiroyuki Iizuka, Masahito Yamamoto & Shigeru Taguchi

Scientific Reports **12**, Article number: 2859 (2022) | Cite this article

1887 Accesses | **30** Altmetric | Metrics

scientific reports

Research movie



Wataru Noguchi

A.I.

Hiroyuki Iizuka

Artificial Life

Shigeru Taguchi

Phenomenology