A Formal Account of Structuring Motor Actions With Sensory Prediction for a Naive Agent

Authors: Jean-Merwan Godon, Sylvain Argentieri and Bruno Gas
Presenter: Wenfan Jiang

ELECTRICAL ENGINEERING AND COMPUTER SCIENCE
UNIVERSITY OF MICHIGAN

Agenda

I. Notations & Concepts
II. Theoretical Knowledge
III. Practical Application
IV. Experiments & Results
V. Discussion

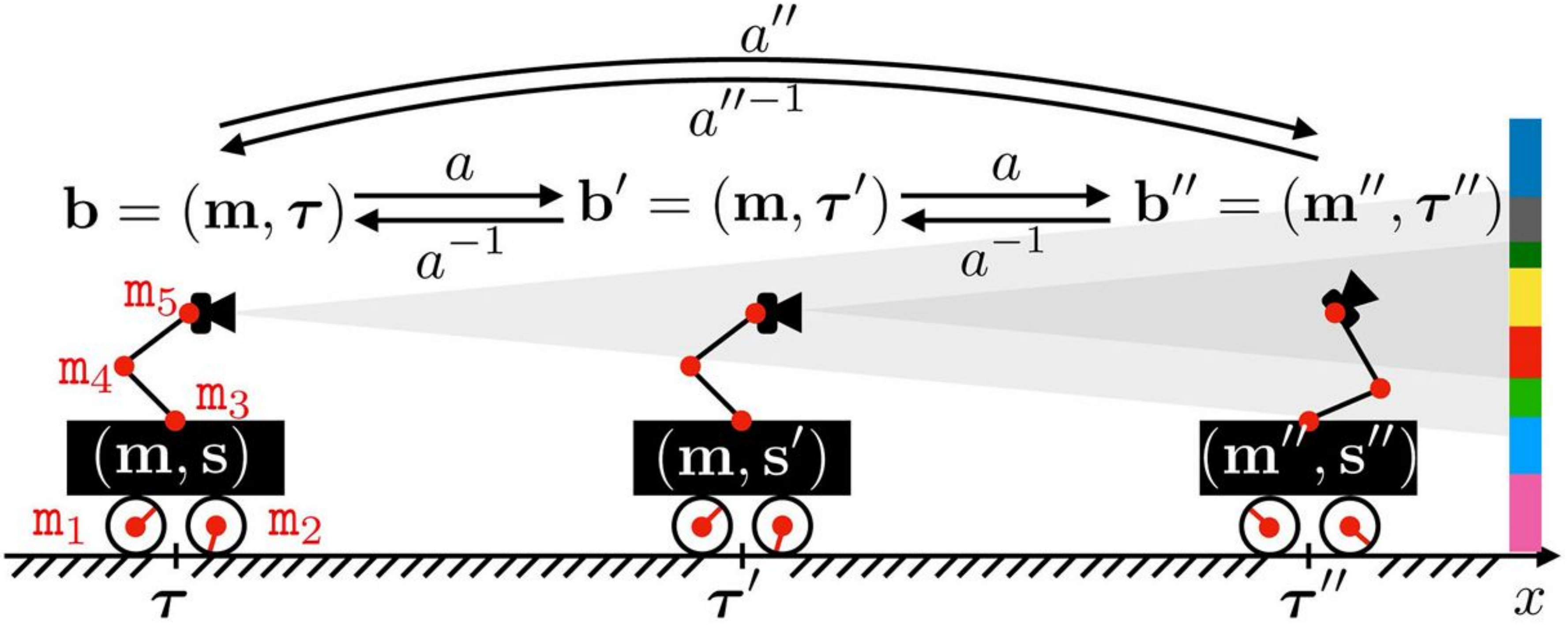
Notations & Concepts

Previous Work:

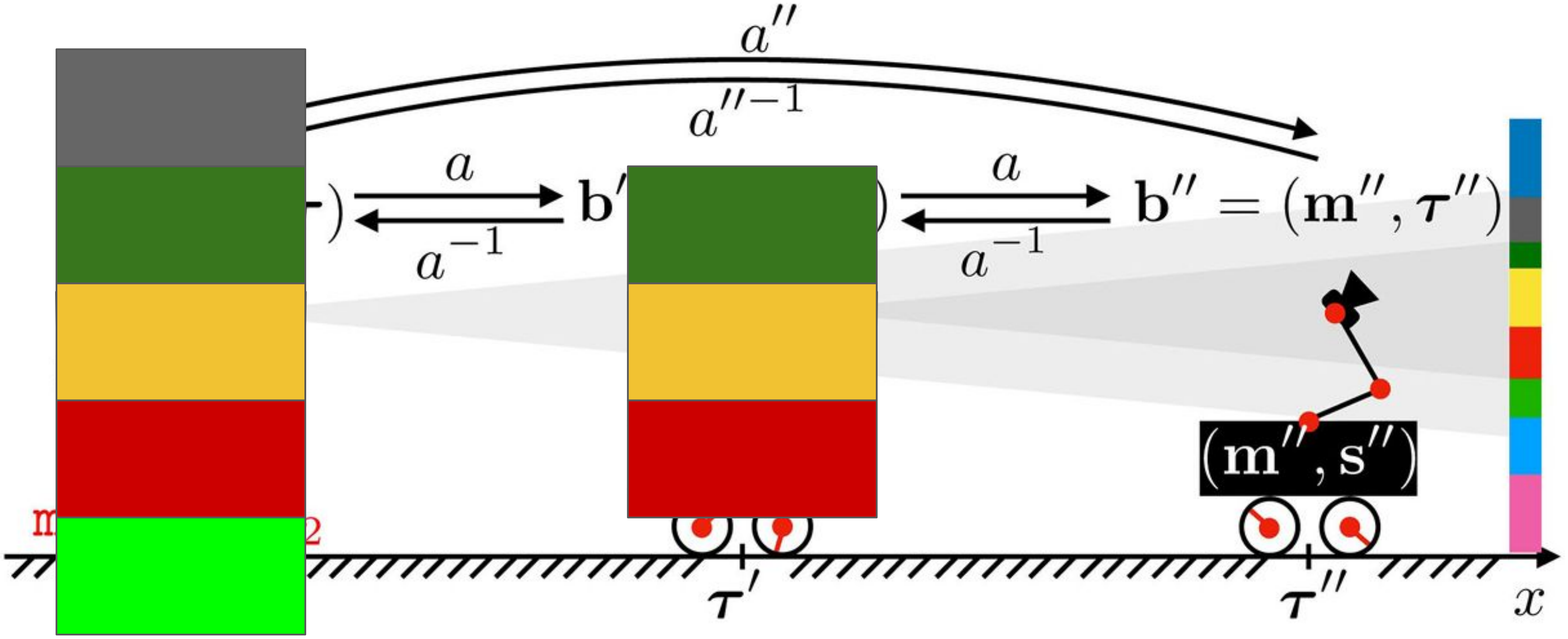
Sensorimotor interaction = Internal motor configuration + sensory configuration

$$\psi : \underset{\uparrow \atop \mathbf{m}}{\mathcal{M}} \times \underset{\uparrow \atop \epsilon}{\mathcal{E}} \rightarrow \mathcal{S}$$

Notations & Concepts



Notations & Concepts



Notations & Concepts

Previous Work:
Sensorimotor interaction = Internal motor configuration + sensory configuration

This paper:
Sensorimotor interaction = (Internal motor configuration, position) + sensory config

$$\psi : \mathscr{B} \times \mathscr{E} \rightarrow \mathcal{S}$$

$$(\mathbf{m}, \boldsymbol{\tau}) \qquad \epsilon$$

Notations & Concepts

Action: a function mapping one of this new configuration to another

$$\begin{tikzcd} \mathbf{b} \arrow[rr, "a''"] \arrow[dr, "a"'] & & \mathbf{b}'' \\ & \mathbf{b}' \arrow[ur, "a'"'] & \end{tikzcd}$$

Closure

Inverse

$$\begin{tikzcd} \mathbf{b} \arrow[r, shift left, "a"] & \mathbf{b}', \arrow[l, shift left, "a^{-1}"] \end{tikzcd}$$

$$\begin{tikzcd} \mathbf{b} \arrow[rrrr, "a"] \arrow[dr, "a_{i_1}"] & & & & \mathbf{b}' \\ & \mathbf{b}_1 \arrow[r, "a_{i_2}"] & \cdots \arrow[r, "a_{i_{n-1}}"] & \mathbf{b}_{n-1} \arrow[ur, "a_{i_n}"] & \end{tikzcd}$$

Associativity

Also have Identities

Notations & Concepts

Ambient geometrical space: Our 3D physical world

$$\epsilon : \mathcal{X} \rightarrow \mathcal{P}$$

Environmental states: Temperature, luminance, humidity…

Theoretical Knowledge - Spatial Information

$$\forall \epsilon_1, \epsilon_2 \in \mathscr{E}, \forall \mathbf{b} \in \mathscr{B},$$

$$\epsilon_{1|F_{\mathscr{C}'}(\mathbf{b})} = \epsilon_{2|F_{\mathscr{C}'}(\mathbf{b})} \Rightarrow \psi_{\mathscr{C}'}(\mathbf{b}, \epsilon_1) = \psi_{\mathscr{C}'}(\mathbf{b}, \epsilon_2).$$

“No matter what motor you're doing and what position you're at, the environment can only give one sensorimotor mapping result, which is the perception.”

Theoretical Knowledge - Spatial Information

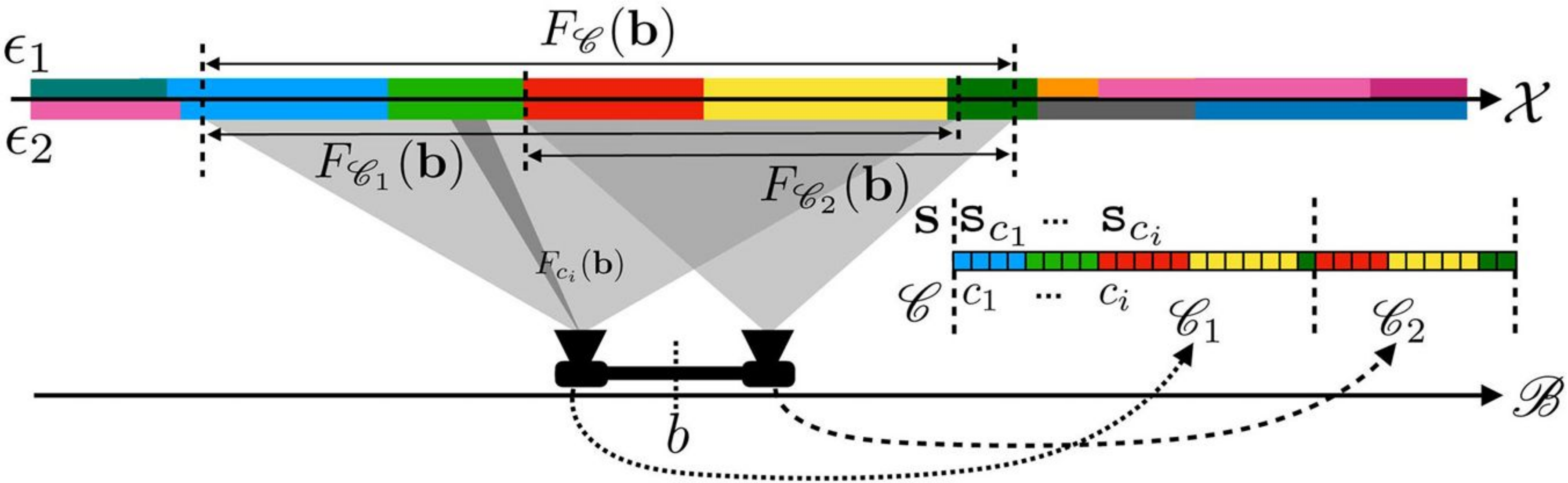


Figure 2

Theoretical Knowledge - Spatial Information

$$\forall \mathbf{b} \in \mathcal{B}, \forall \epsilon \in \mathcal{E}, \psi_c(\mathbf{b}, \epsilon) = f_c(\epsilon_{|F_c(\mathbf{b})}),$$

f() is the sensitivity function that converts physical properties into a sensory output.
F(b) is minimal region of space which entirely determines the output.

This equation represents the sensorimotor dynamics, and the observer can now speak of sensels that look at the same region of space.

Theoretical Knowledge - Spatial Information

$$\forall \mathbf{b} \in \mathscr{B},\ F_{c_i}(a\mathbf{b}) = F_{c_j}(\mathbf{b}).$$

This relation describes the situation that “after an action, sensel i samples the same point as sensel j sampled before this action.”

Theoretical Knowledge - Spatial Information

$$\forall \mathbf{b} \in \mathscr{B}, \qquad \forall \epsilon \in \mathscr{E},$$

$$\begin{aligned} \psi_{c_i}(a\mathbf{b}, \epsilon) &= f_{c_i}(\epsilon_{|F_{c_i}(a\mathbf{b})}) \\ &= f_{c_j}(\epsilon_{|F_{c_j}(\mathbf{b})}) = \psi_{c_j}(\mathbf{b}, \epsilon). \end{aligned}$$

“After an action, the output of the sensels i and that of j before this action, which are represented by the sensitivity functions, are also equal.”

Theoretical Knowledge - Spatial Information

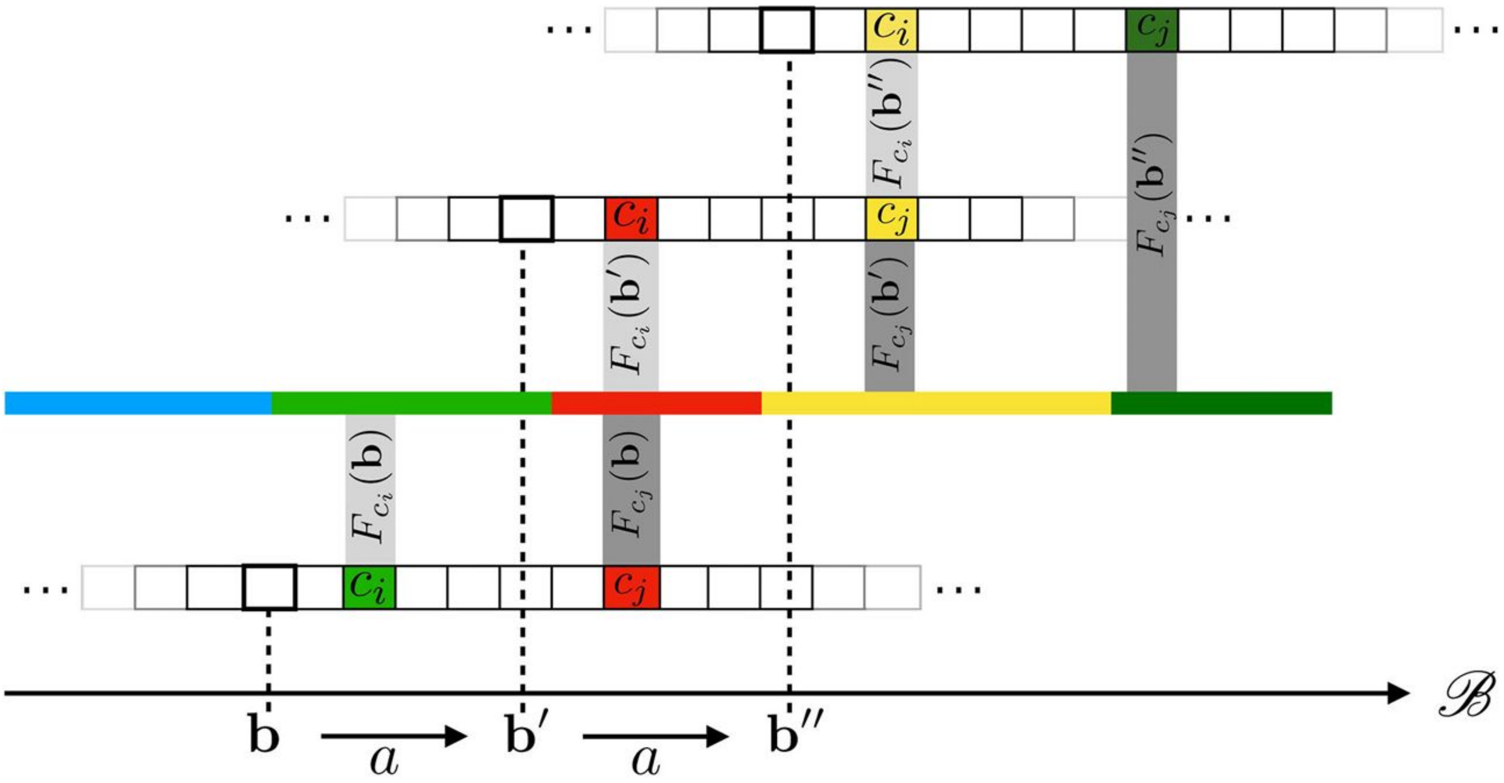


Figure 3

Theoretical Knowledge - Spatial Conservation

An action is *conservative* if

$$\forall c \in \mathscr{C}, \exists c' \in \mathscr{C} \text{ such that } \forall \mathbf{b} \in \mathscr{B}, F_c(a\mathbf{b}) = F_{c'}(\mathbf{b}),$$

“After this action, no matter what motor you’re doing or which location you’re at, sensels sample a new permutation in general.”

Theoretical Knowledge - Spatial Conservation

$$\Pi : \mathcal{A}_{\mathscr{C}} \to \mathrm{Bij}(\mathcal{S})$$
$$a \mapsto \Pi_a$$

$$\mathcal{A}_{\mathscr{C}} \cong \Pi(\mathcal{A}_{\mathscr{C}}).$$

Theoretical Knowledge - Spatial Conservation

Can directly use permutation of the matrix of actions!

$$\mathcal{A}_{\mathscr{C}} \cong \Pi(\mathcal{A}_{\mathscr{C}}).$$

Experiment - Setup

- Ambient Space
- Physical properties (environmental states)
- Configuration of motors and positions
- Grayscale sensory output
- 7 actions (translations along x and y coordinates, rotations in the plane)

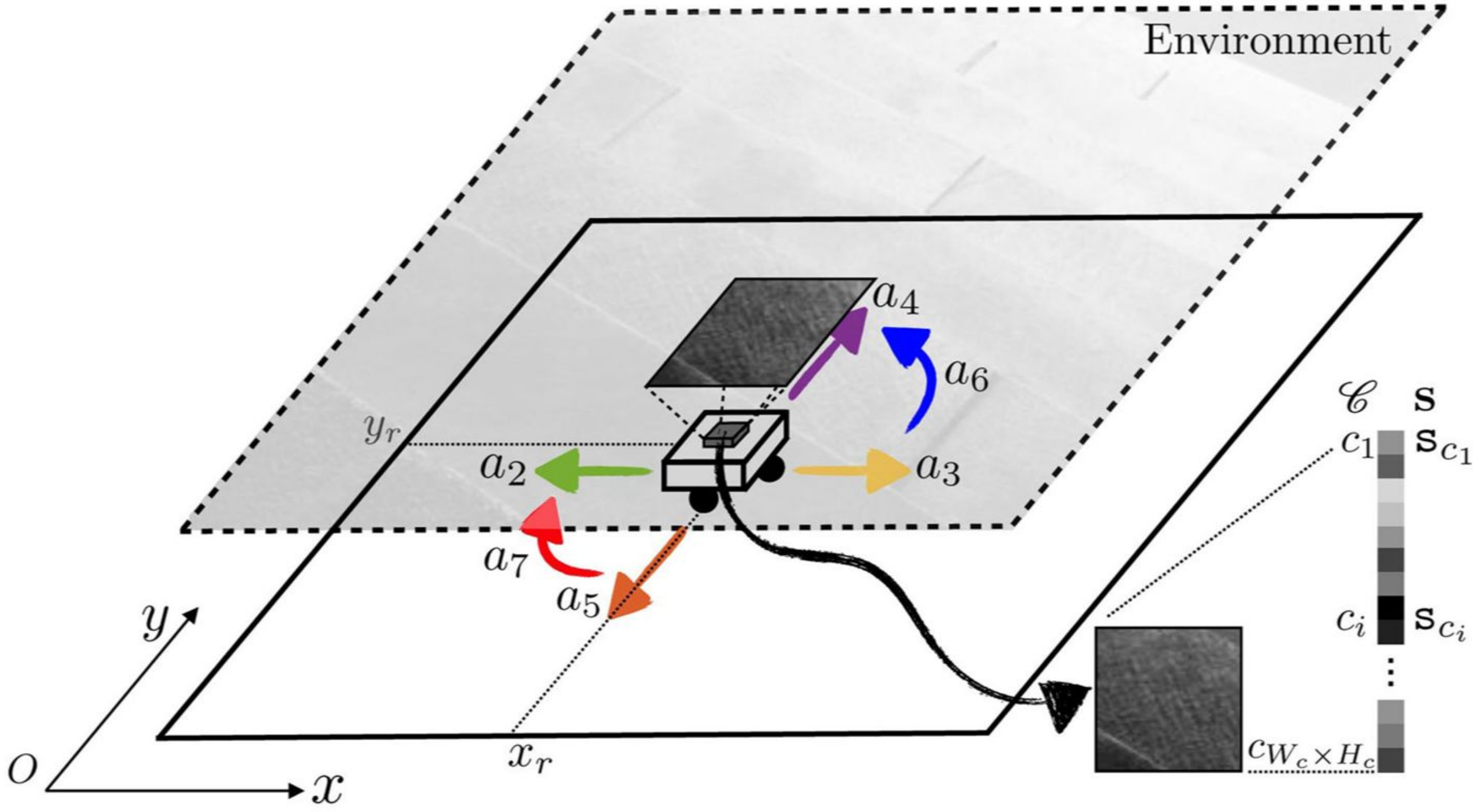
1. Agent effectively runs available actions and find out conservative ones by computing the sensel permutation matrices.

2. Environmental configuration is changed via changing the grayscale image.

3. Agent gets the resulting set of permutation matrices.

Experiment - Setup

Figure 4



Experiment Details

1. Compute the permutation matrix associated to each of its available motor actions. Conservative motor actions will converge to a certain permutation matrix, while non-conservative ones will converge to null matrix, so the agent can distinguish between conservative and non-conservative actions.

2. The agent uses the prediction functions it discovered for elementary conservative moves to infer how combinations of these moves relate to each other, where theoretically, combination of their permutation matrices results the same as the permutation matrix of the combination of two conservative actions.

Experimental Results

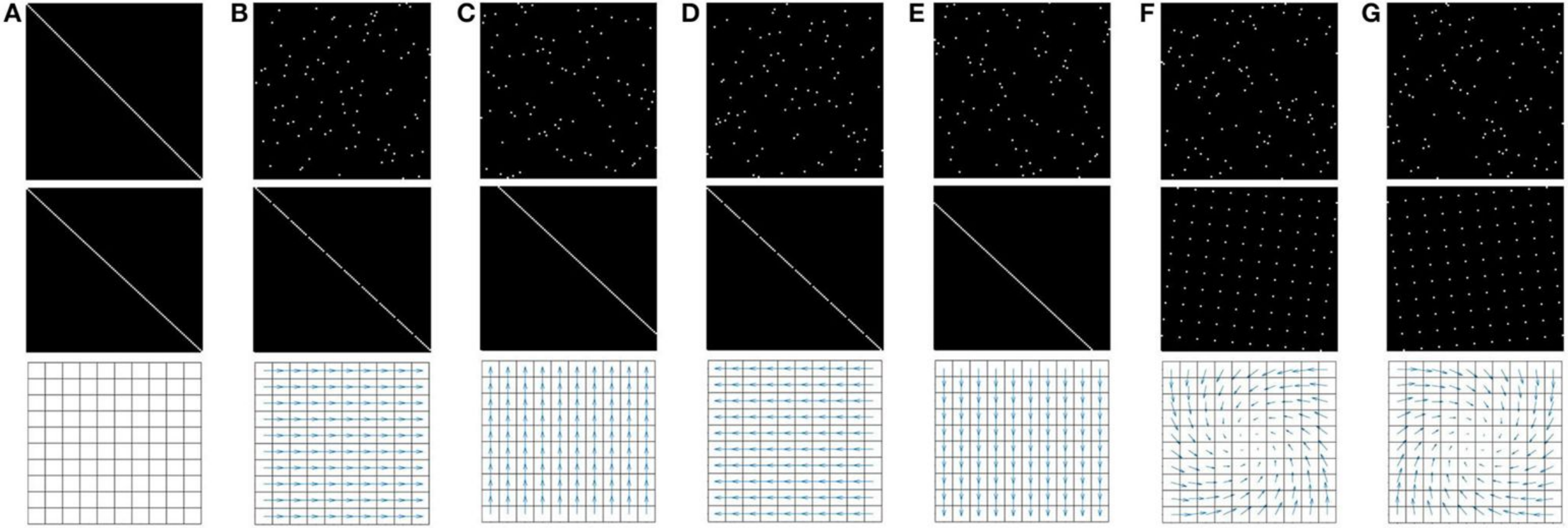


Figure 5

Experimental Results

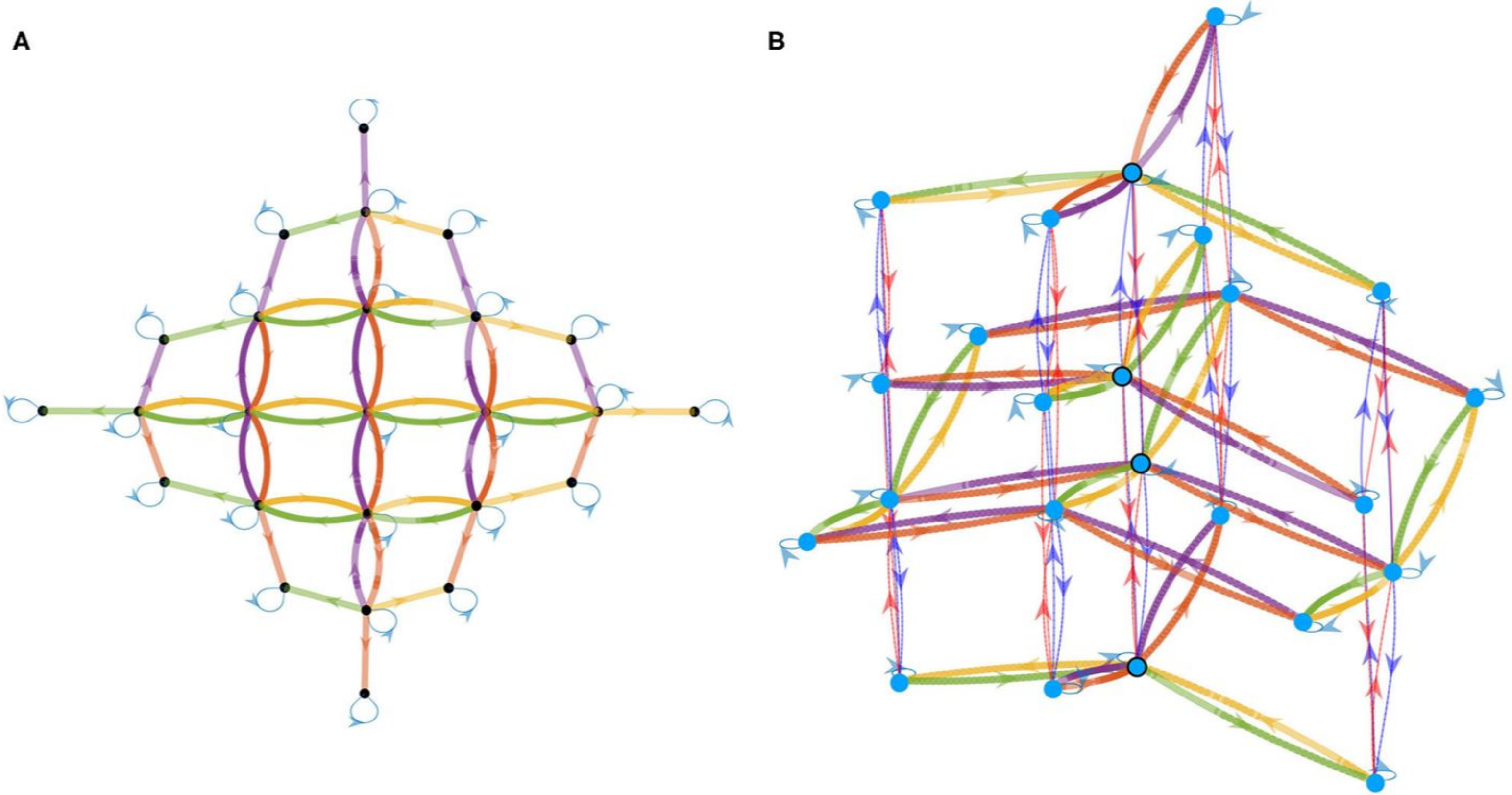


Figure 8

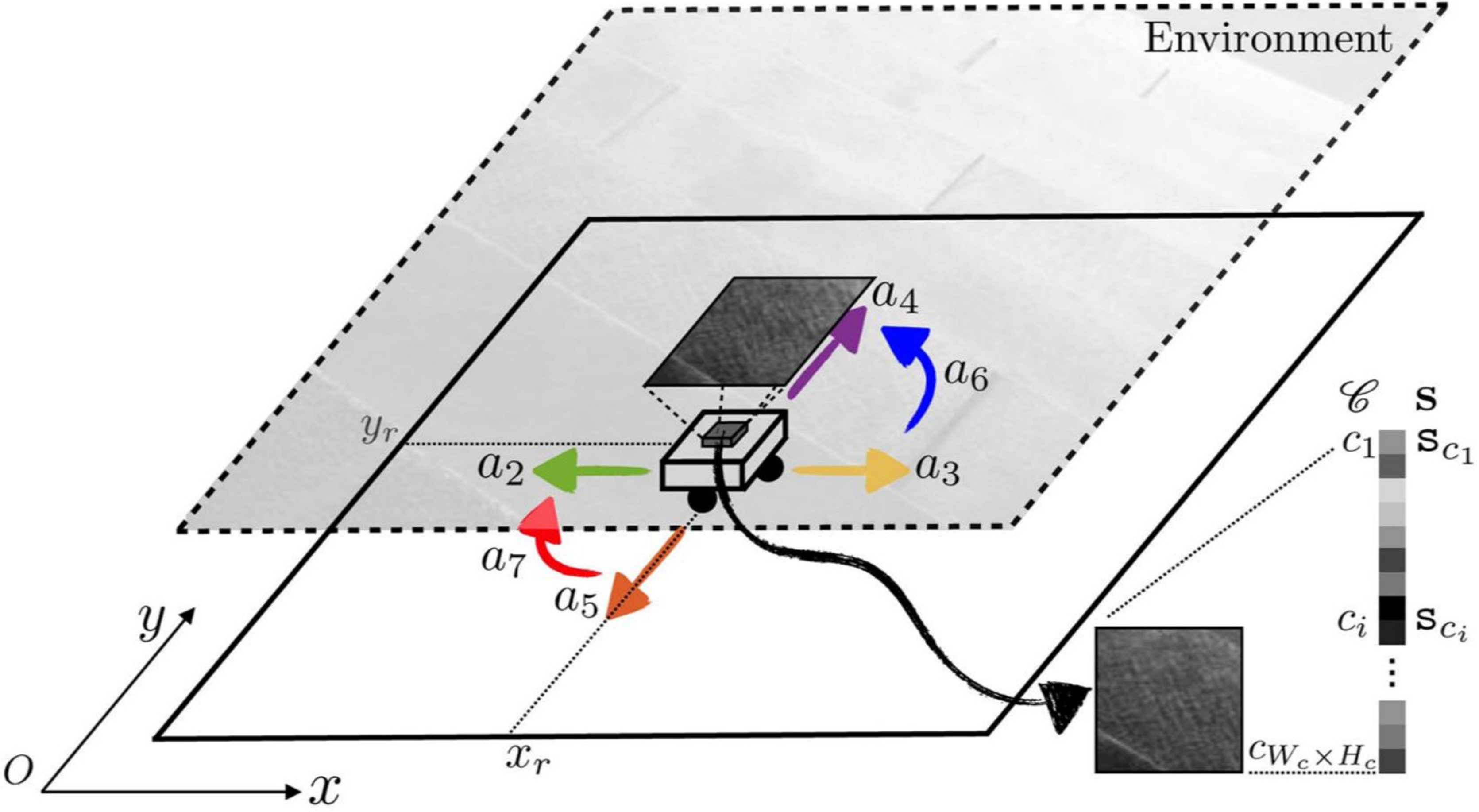

Environment
a_4
a_6
a_2
a_3
a_7
a_5
y_r
x_r
y
O
x
$\mathscr{C}$
$\mathbf{s}$
c_1
$\mathbf{s}_{c_1}$
c_i
$\mathbf{s}_{c_i}$
$c_{W_c \times H_c}$

Figure 4

Experimental Results

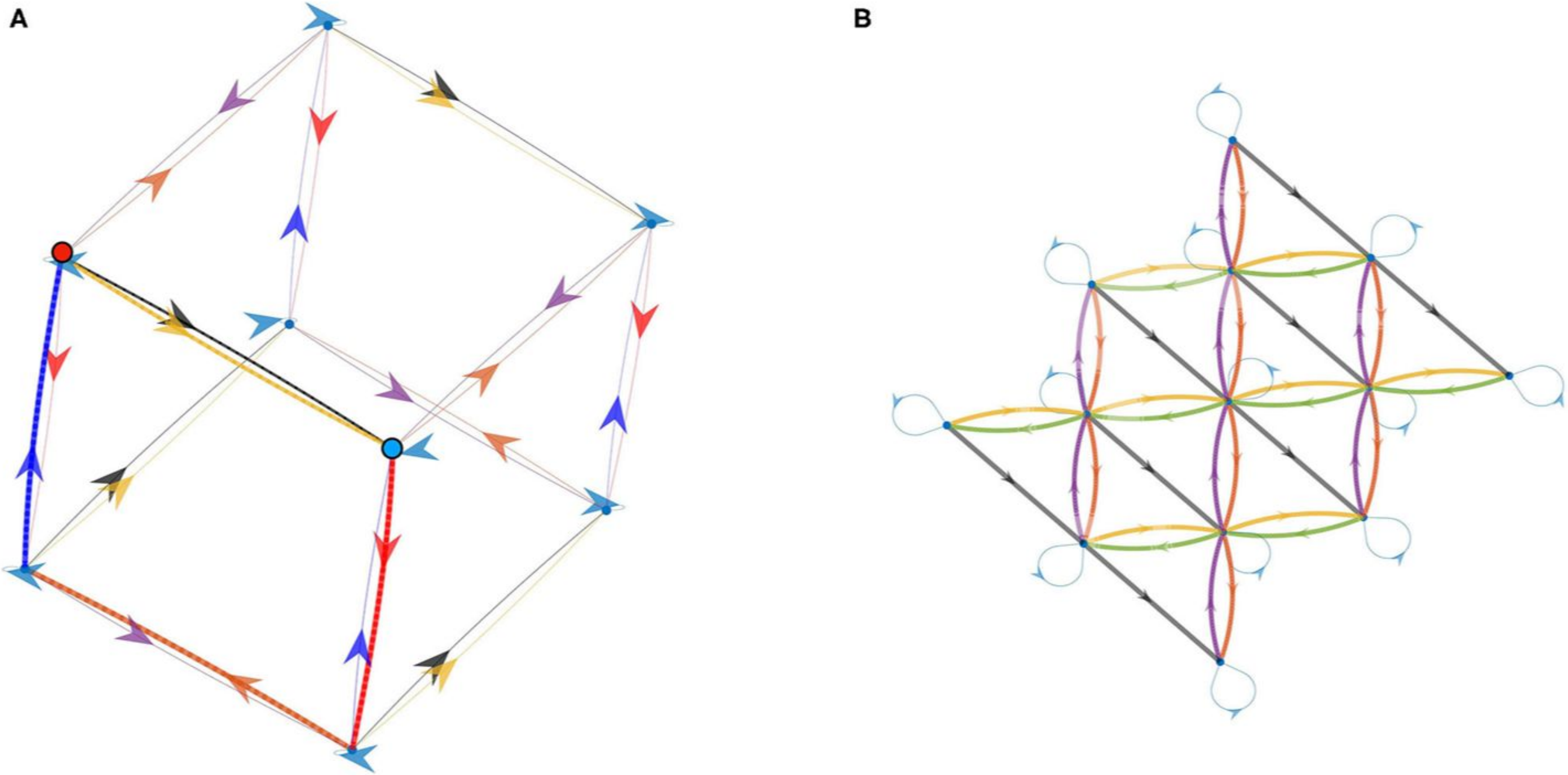


Figure 9

Discussion: Piazza @53_f1

“Conservative actions are rare in real life, so it is still hard to apply this formalism. Is there a way to map non-conservative actions to conservative ones?"

We can try to use a deep network to map the non-conservative actions to a conservative prototype, we may investigate a way to mathematically compute it in a conservative form, or we can extract the conservative portion and do the computation on that part only, which might give useful results, too. Same thing applies to noisy actions: we can try to denoise them or extract the known part.

Discussion: Piazza @53_f4

“The experiment only deals with pixels of an image. Can we incorporate other types of sensory information such as touch, taste, or auditory inputs?"

Yes, auditory inputs is promising, where people have already used it in different learning applications, and we can measure them in either the frequency or the volume even if it might be hard to do a permutation. It’s also possible to combine the auditory inputs with the current approach of vision, too. However, the other sensory information seems to not be quantifiable, which means it is hard to be used in any mathematical rules provided in this paper, while they might still be considerable in a deep learning networks.

Discussion: Piazza

In addition to the previous two questions, people were also discussing about the meaning of “noisy action”, which can be inferred as a stochasticity in action, because in most other related work, action tends to be seen as a distribution.

Besides, we were talking about the further exploration of generalizing this method to non-conservative actions, and there was also agreement about “deep learning network can be useful in this direction.” especially when the author also mentioned that the sensory prediction in this paper learns in the same way as neural network does.

In the end, we also got the interest from the authors’ comparison and relations between the “external viewpoint” and the “internal viewpoint”.