SMA group meeting

Hierarchical Sequential Learning with Constructivist Paradigm

Jianyong XUE

Université de Lyon, LIRIS CNRS UMR5205

Université Claude Bernard Lyon 1

10.06.2020

Learning from interactions - why?

- We cannot design **all** interaction situations to the robots.
- Agent's interaction with the environment could make up de **deficiency** in the artificial designing.
- We need a way let the agent to **learn by itself** with requiring little or no manual intervention
- Also, this learning process need to be effective and efficient. (learning fast and with as much as complete knowledge)

Self-motivated Self-adaptive

Learning from interactions - how?

- Action selection with a **probabilistic model** to select one most likely action.
- A **predictive model** to select the actions to generate the maximum reward.
- A sequence of **specific** actions for completing a **specific** task.
- **Structured behaviors** with hierarchical sequential patterns for accelerating behaviors construction process.

Current methods

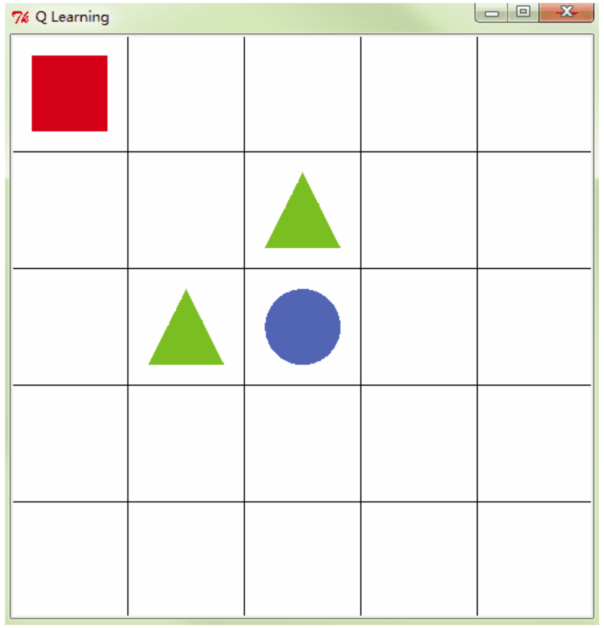


Fig. 1. Q-learning with ε-exploration in grid environment.

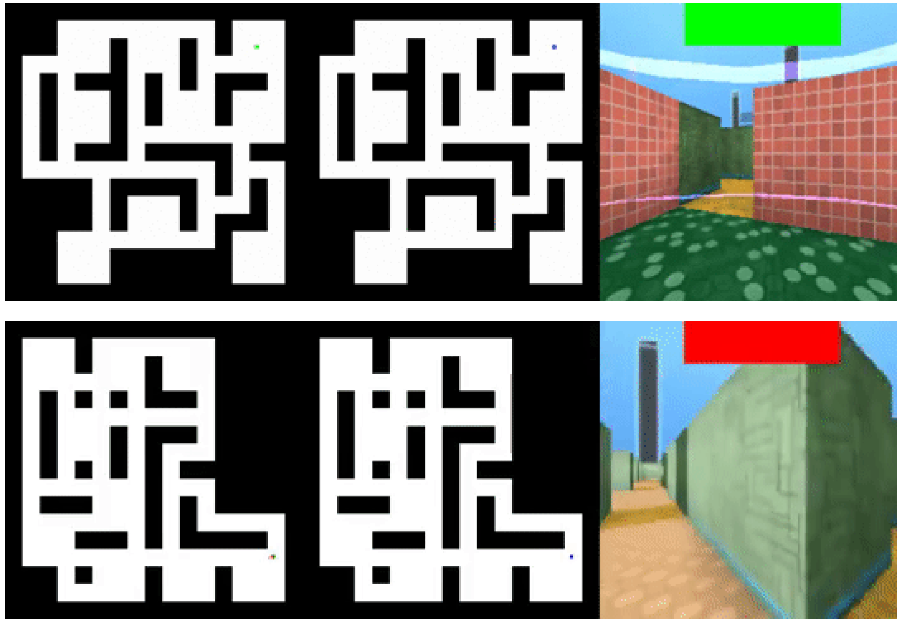


Fig. 2. Reward visualization. The results from Google and ETH's model about "Episodic curiosity through reachability".(See the link: https://ai.googleblog.com/2018/10/curiosity-and-procrastination-in.html)

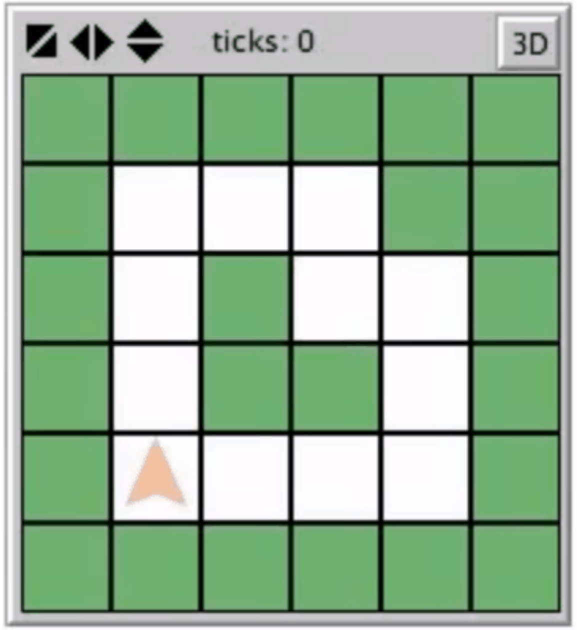


Fig. 3. Demonstration of developmental learning. This result comes from the IDEAL MOOC of Olivier Georgeon, the link is: http://liris.cnrs.fr/ideal/mooc/, the video from: https://www.youtube.com/watch?v=LVZ0cPpmSu8

Imagining the following scenario:

An agent is placed in an **unfamiliar** environment, with only **innate actions** that could let it move around and interact with objects to start the journey of “feeling the world”.

Unlike other interactive scenarios that this interactive process without any **prior** knowledge, nor the final goal for the agent to achieve.

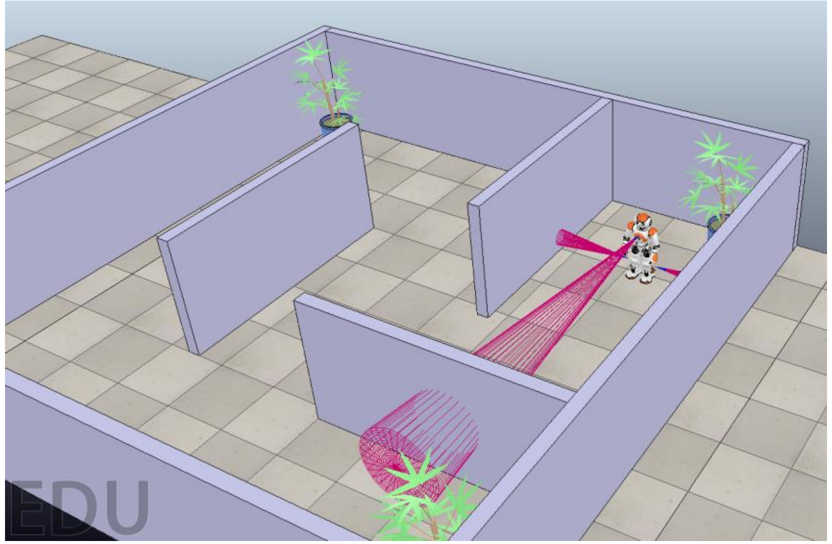


Fig. 4. The initialization environment for the agent to interact

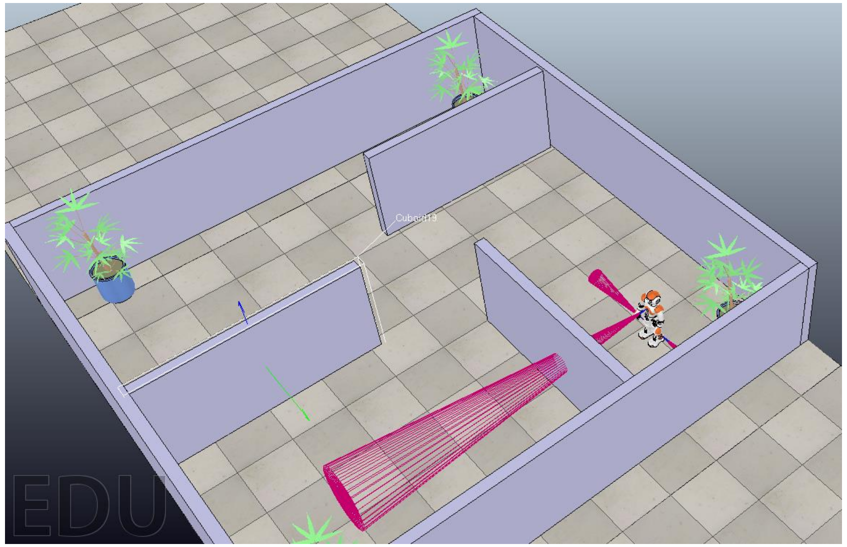


Fig. 5. The agent is placed in the changed environment.

Challenges

How can we design an alternative learning paradigm that satisfies the conditions mentioned above. The agent could **successfully interact with its environment** and learn to **avoid unfavorable interactions** using **regularities** that it has learned.

The agent with structured behaviors it has learned from interactions obtains capabilities of **self-motivated** and **self-adaptive** that can behave in a “intelligent” and **flexible manner under dynamic conditions**.

Our focuses

- **Knowledge construction through interactions** between the agent and the environment. A self-motivated agent could discover and explore regularities in its stream of experiences and to construct knowledge about phenomena, which hypothetical presence in the environment explains these regularities.

- **Higher-level sequence learning with constructivist paradigm**. The agent could increasingly learn elaborated behaviors and gradually organized them in a hierarchy that reflects how the agent exploits the higher-level regularities afforded by the environment.

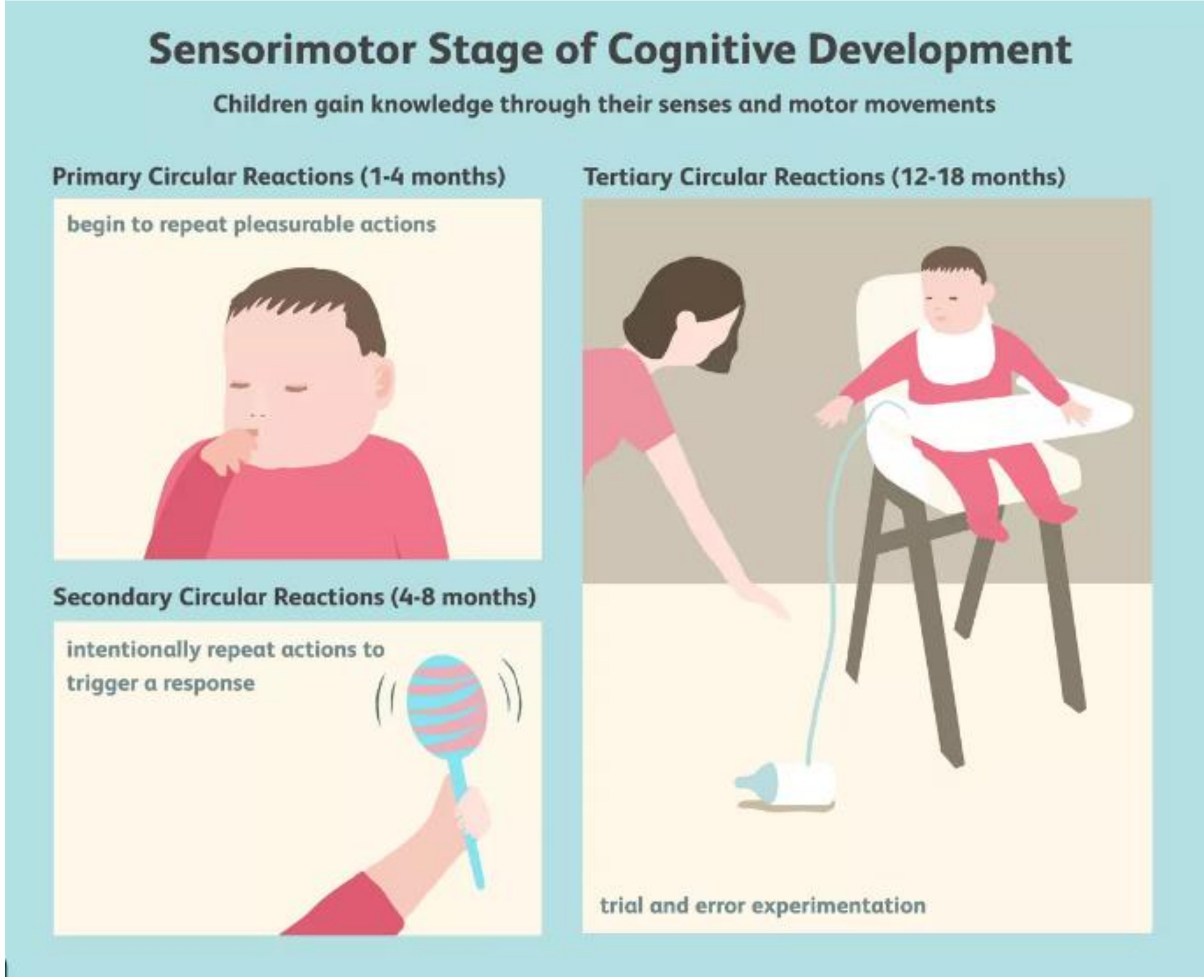
- **Context adaptation and generating proper behaviors**. The agent could understand current interactive situation, learn behavioral patterns for affectively generating proper behaviors or even more optimal structured behaviors for enacting.

Constructivist learning paradigm

Constructivist learning paradigm

- Constructivism as a knowledge acquisition theory which describes the cognitive development of children. It proposes that learning happens as a result of a internal mental representations and external perceptions from interactions.

- During the initial phase of cognitive development, infants exhibit amazing abilities to generate novel behaviors with unfamiliar situations and explore actively to learn the best with lacking extrinsic rewards from the environment.



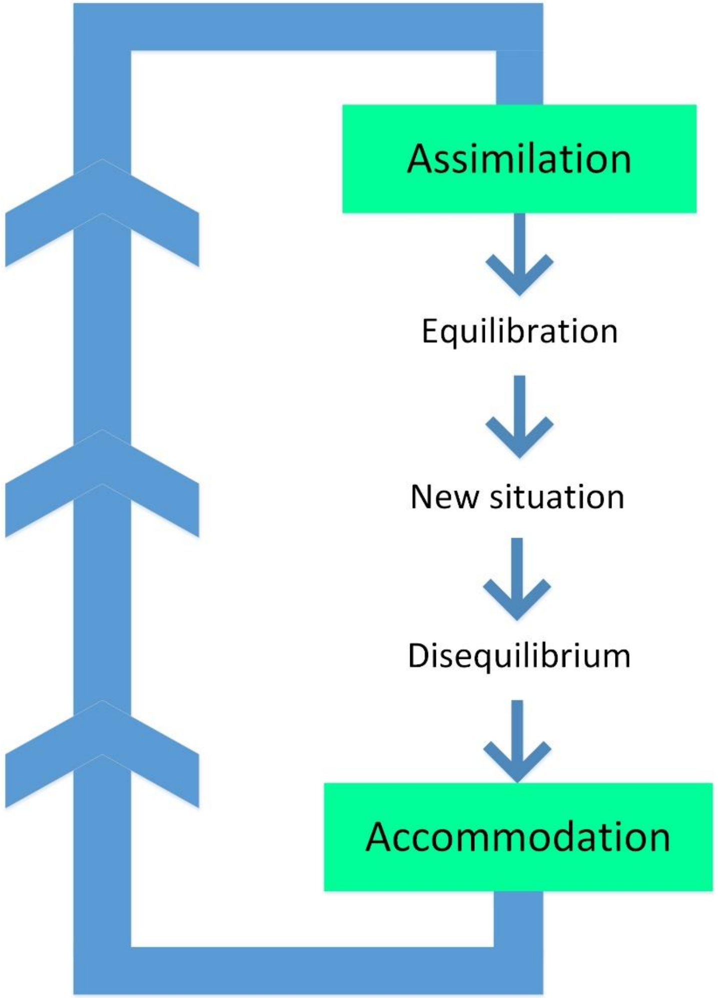


Fig. 6. The initial phase of cognitive development.
Photo source: https://www.verywellmind.com/sensorimotor-stage-of-cognitive-development-2795462

Fig. 7. The three processes in constructivist learning and self-adaption:

The definition of interaction and composite interaction

- Interaction is defined as a tuple of:

$$i_t = \langle e_t, f_t \rangle,$$

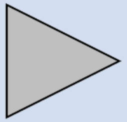
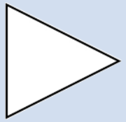
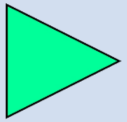
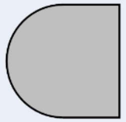
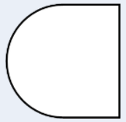
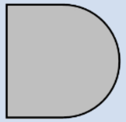
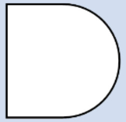
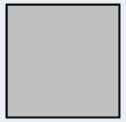
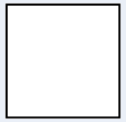
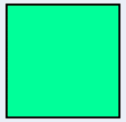
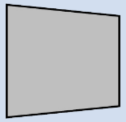
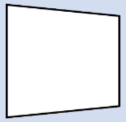
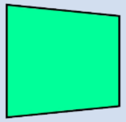
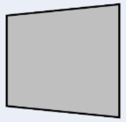
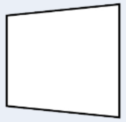
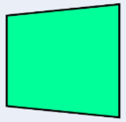
which means the agent performs an experience e_t and receives feedback f_t that composites a given interaction i_t at step t, also we call this the *agent enacts an interaction* i_t.

- The agent intends an interaction i_t^i and receives the enacted interaction i_t^e , then the agent memorizes the two-step **enacted interaction** sequence:

$$c_t = \langle i_{t-1}^e , i_t^e \rangle,$$

as a tuple of < contextInteraction, enactedInteraction >, made by the previous enacted interaction i_{t-1}^e of i_t^e . The interaction i_{t-1}^e is called c_t's pre-interaction, and i_t^e is called c_t's post-interaction.

Primitive interactions

Experiments (or actions)	Icons	Primitive Interactions			
Move forward		Move one step		Bump with wall	
Turn left		Turn left			
Turn right		Turn right			
Touch front		Feel front empty		Feel front wall	
Touch left		Feel left empty		Feel left wall	
Touch right		Feel right empty		Feel right wall	

Constructivist paradigm

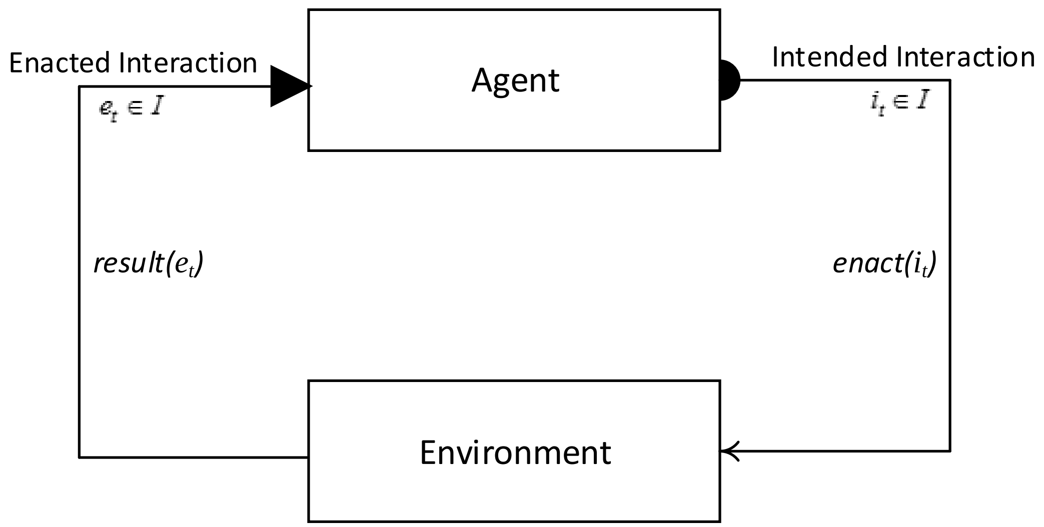


Fig. 8. Constructivist paradigm for the agent to interact with the environment

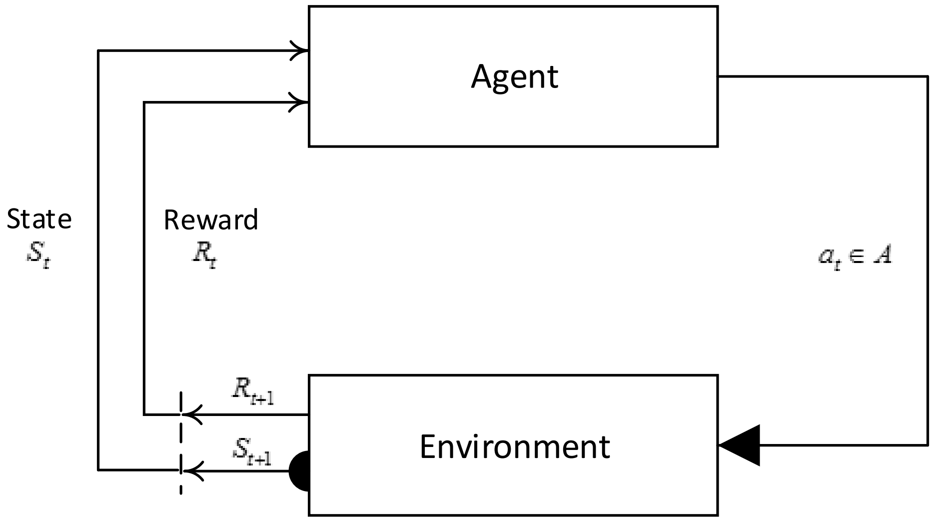


Fig. 9.. The agent-environment interaction in reinforcement learning

Intended interaction and enacted interaction

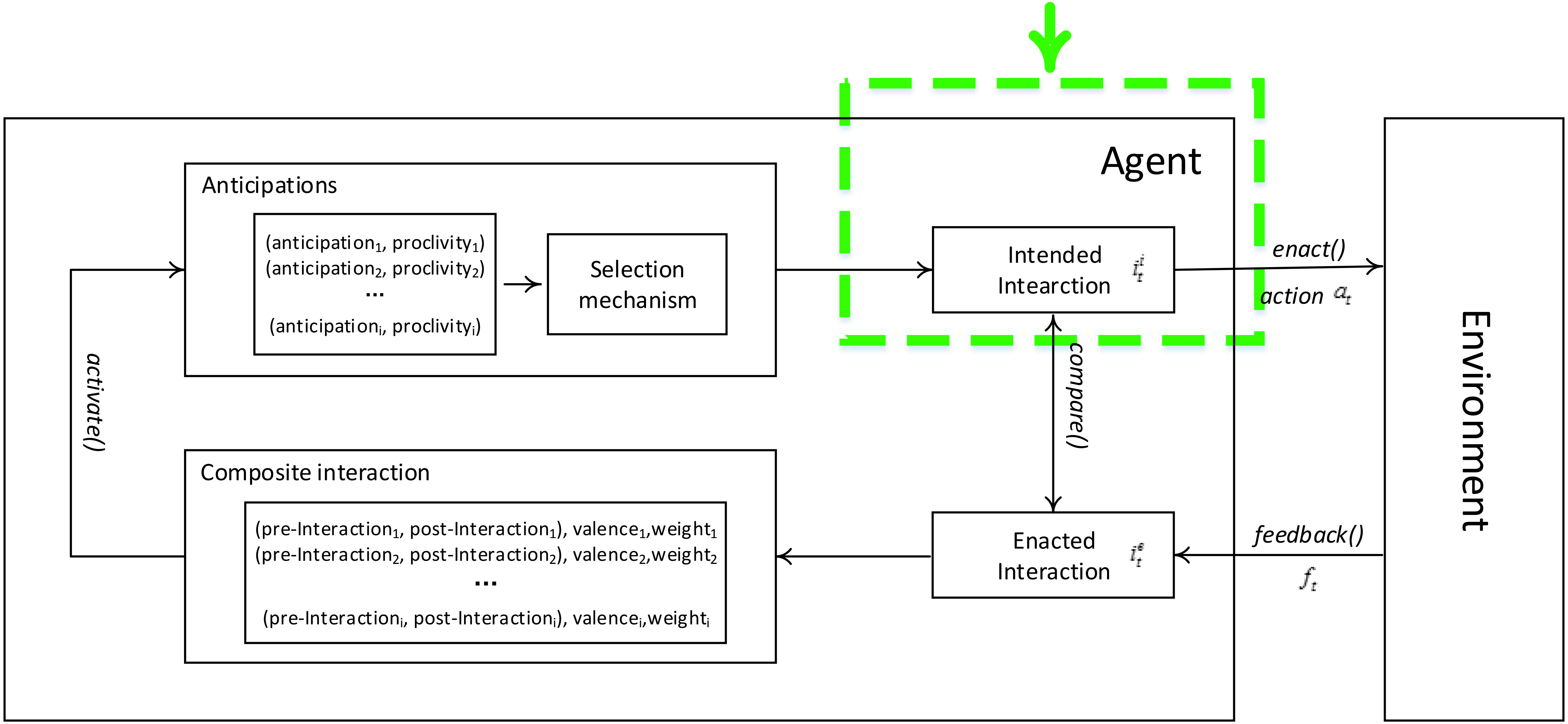
- From the perspective of constructivism:

Intended interaction as it represents the sensorimotor scheme that the agent **intends to enact**, and constitutes the agent's output that is sent to the environment.

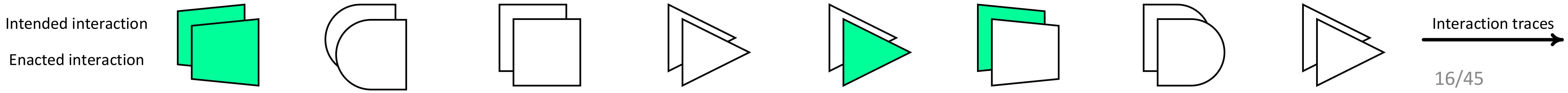
The enacted interaction represents the sensorimotor scheme that the agent records as **actually enacted**, which constitutes the agent's input received from the environment.

If the enacted interaction equals with the intended interaction, then the attempted enaction of intended interaction is considered a *success*, otherwise *failure*.

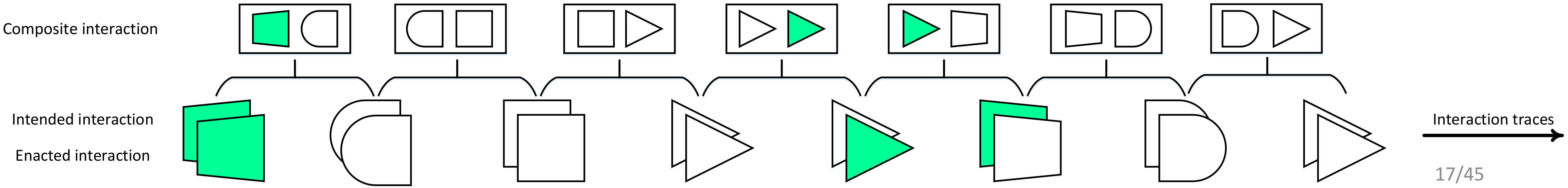
Learning process with constructivist paradigm



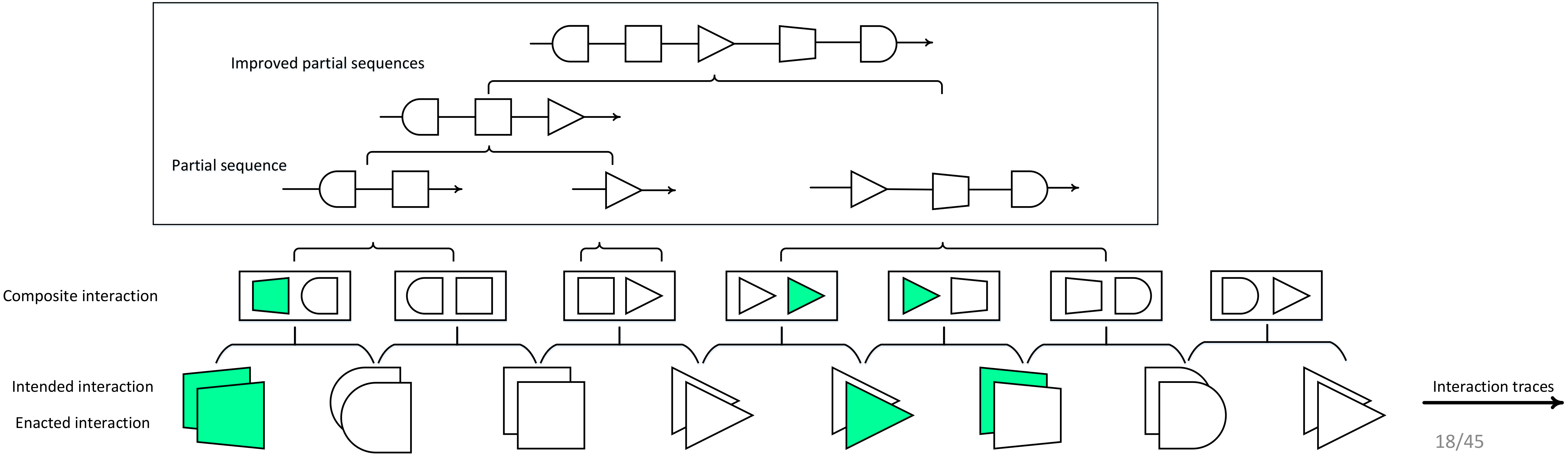
Learning of higher-level structured behaviors with constructivist paradigm



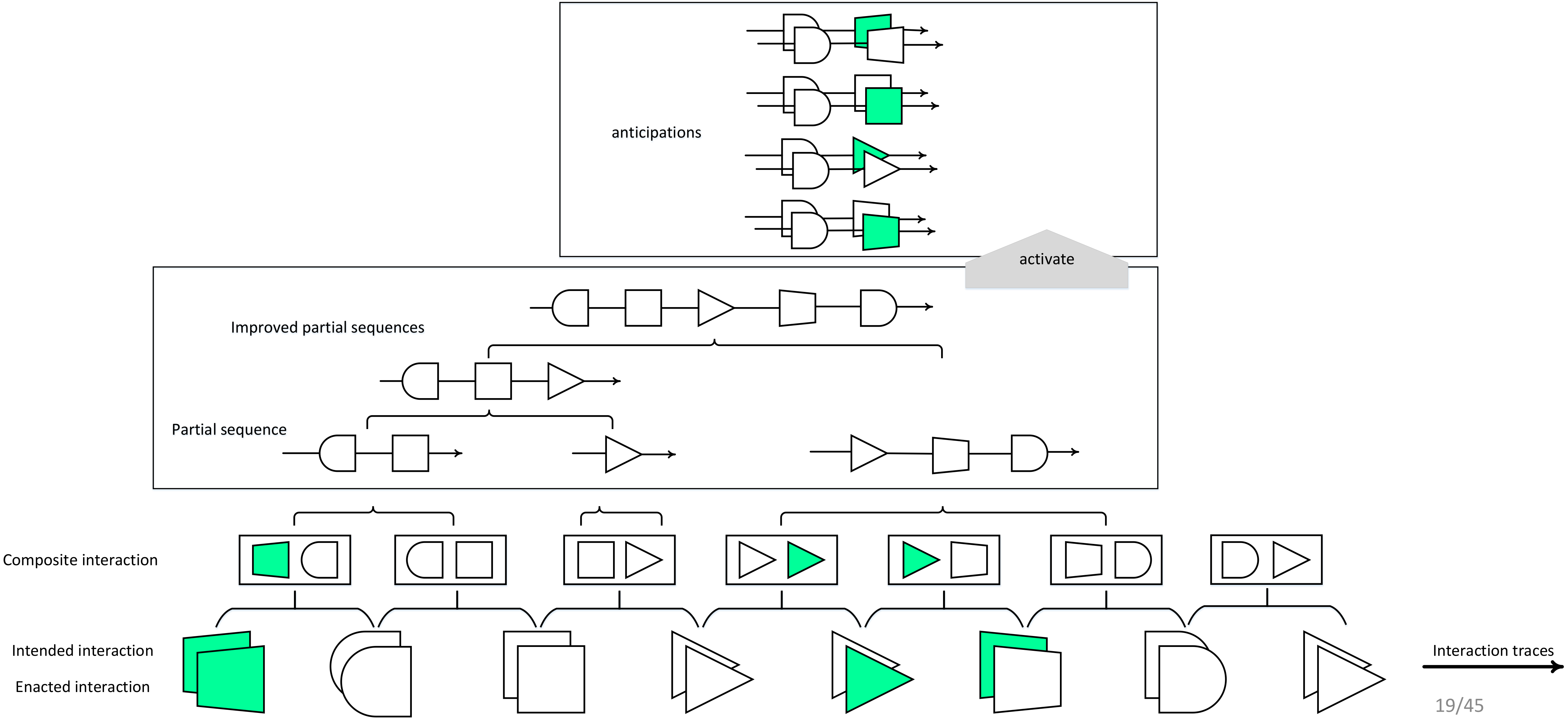
Learning of higher-level structured behaviors with constructivist paradigm



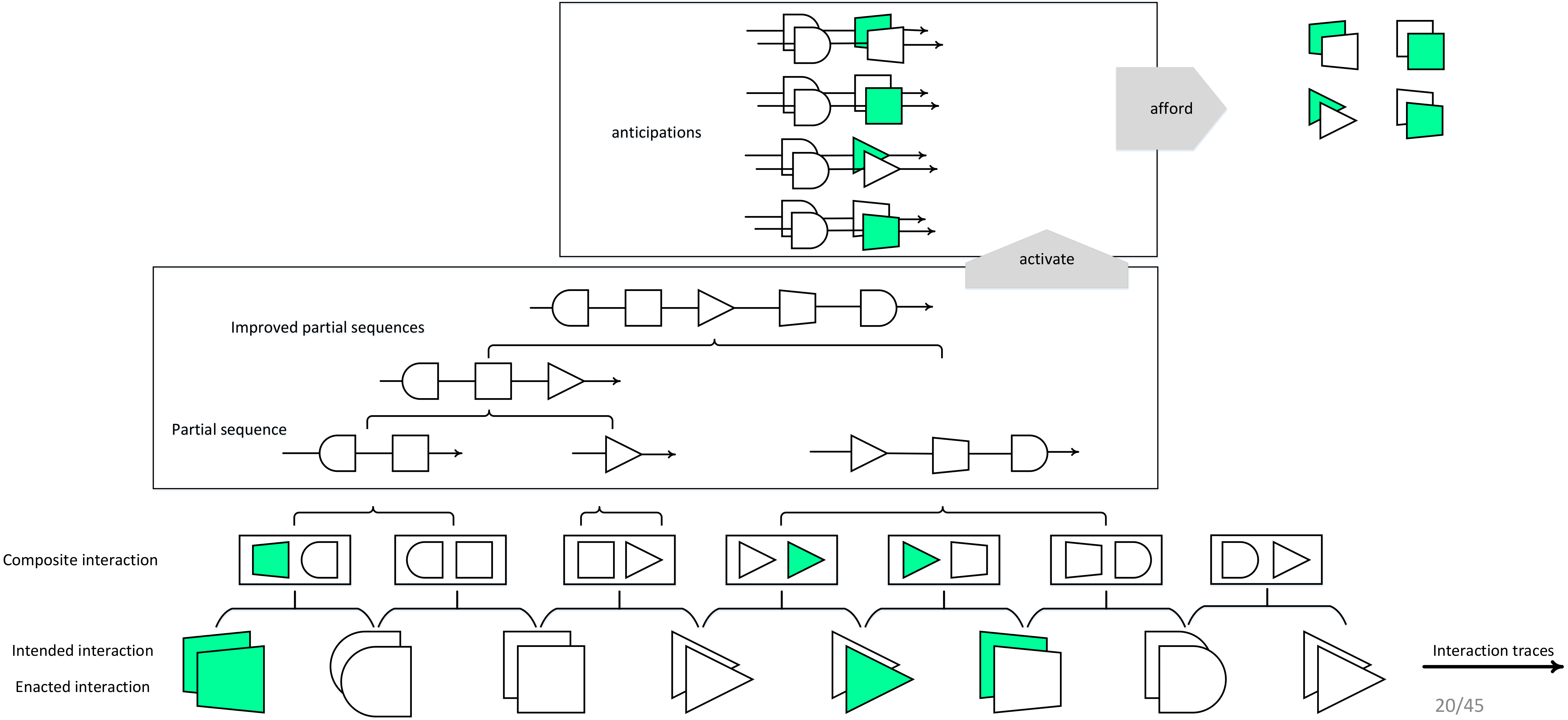
Learning of higher-level structured behaviors with constructivist paradigm



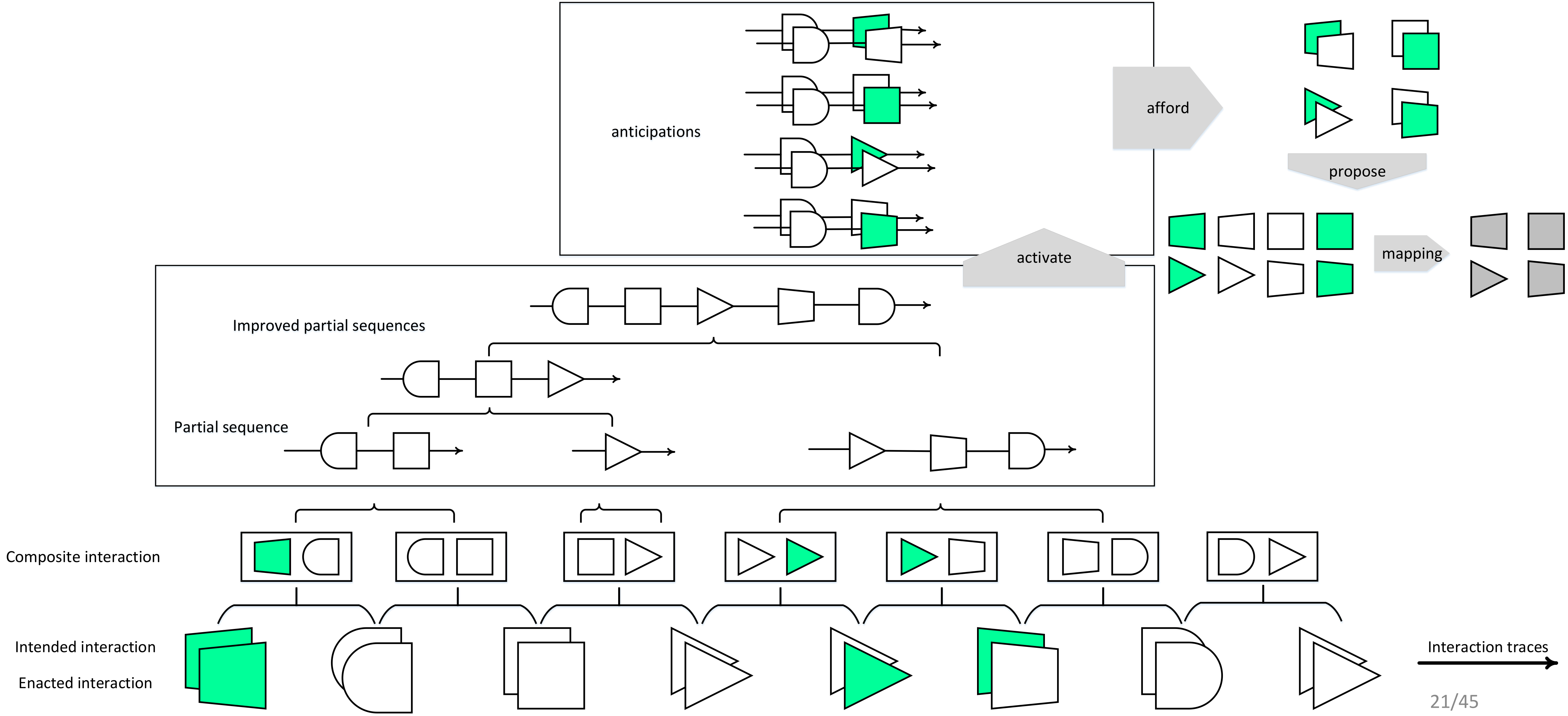
Learning of higher-level structured behaviors with constructivist paradigm



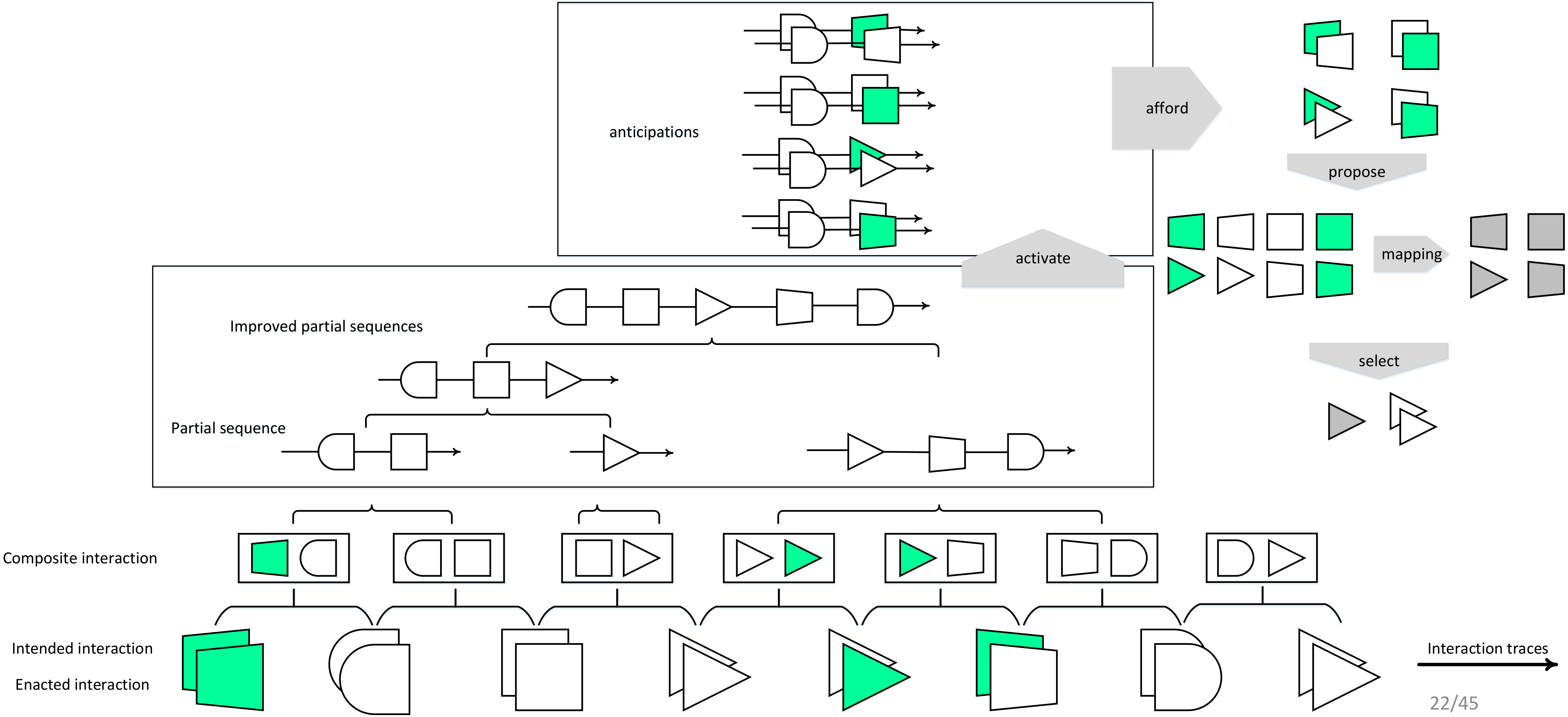
Learning of higher-level structured behaviors with constructivist paradigm



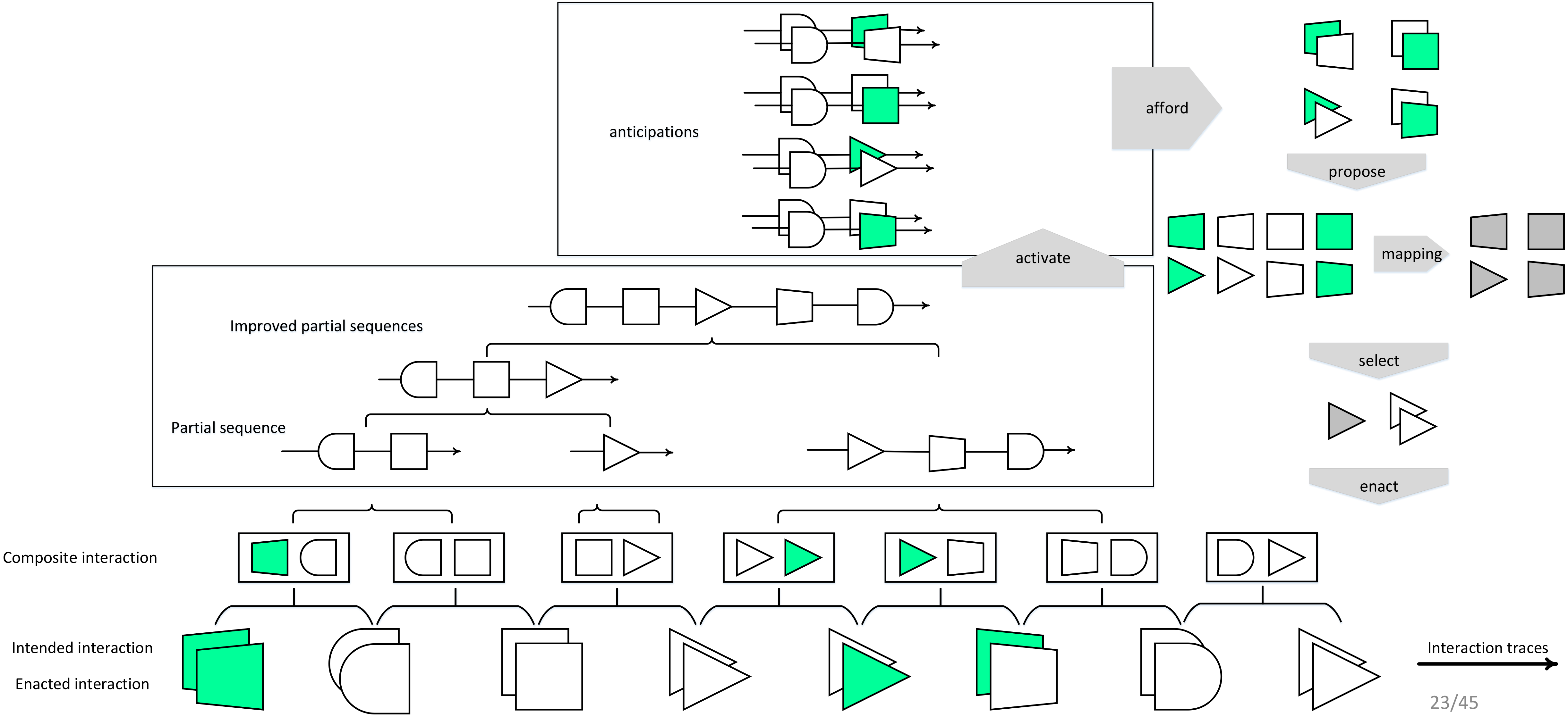
Learning of higher-level structured behaviors with constructivist paradigm



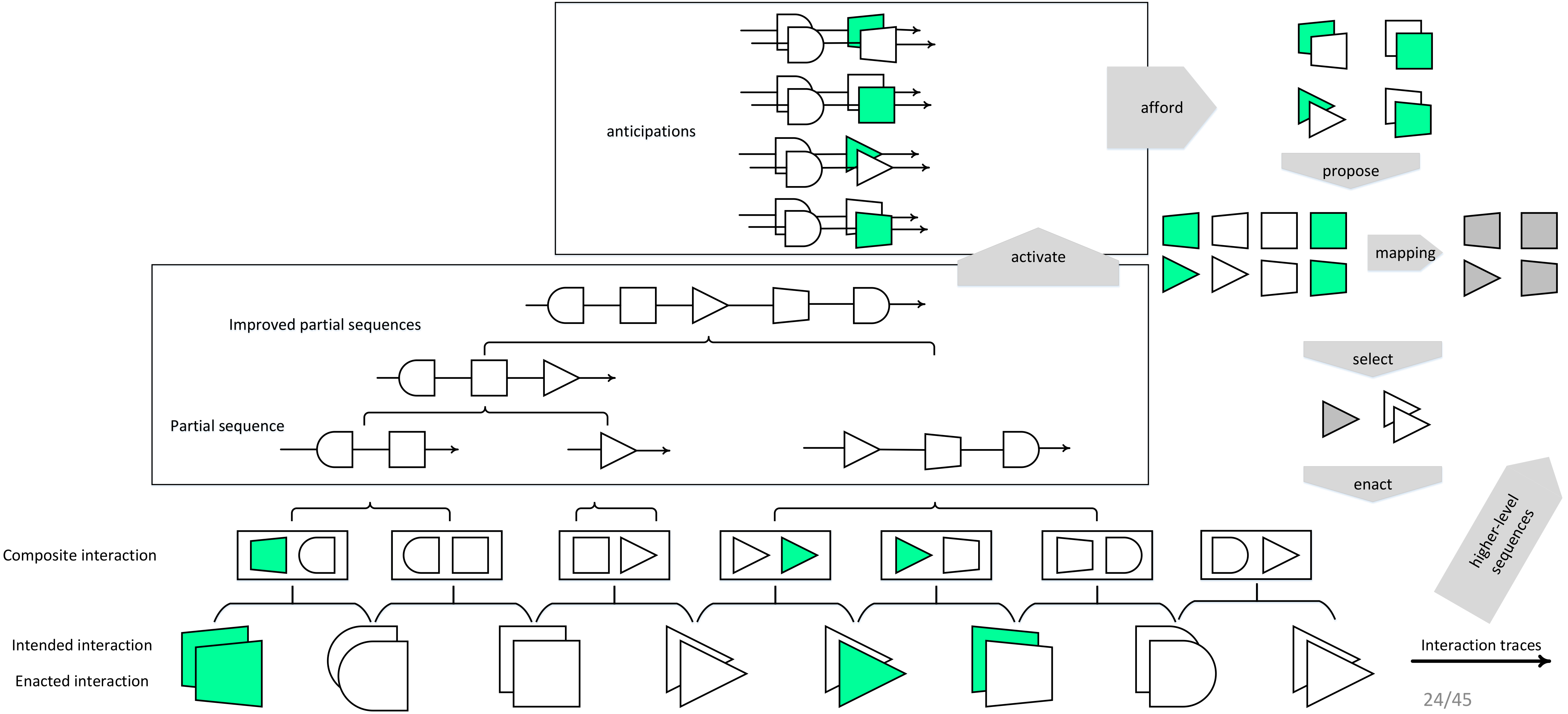
Learning of higher-level structured behaviors with constructivist paradigm



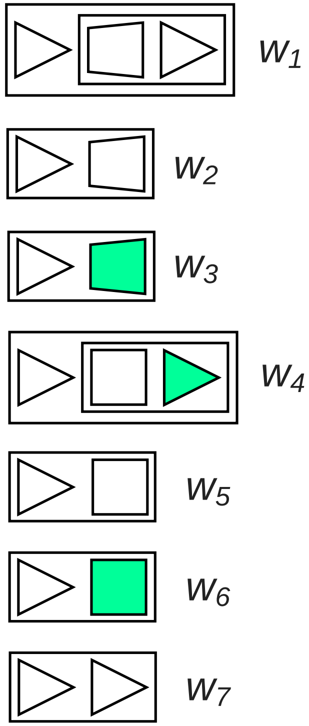
Learning of higher-level structured behaviors with constructivist paradigm



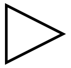
Learning of higher-level structured behaviors with constructivist paradigm



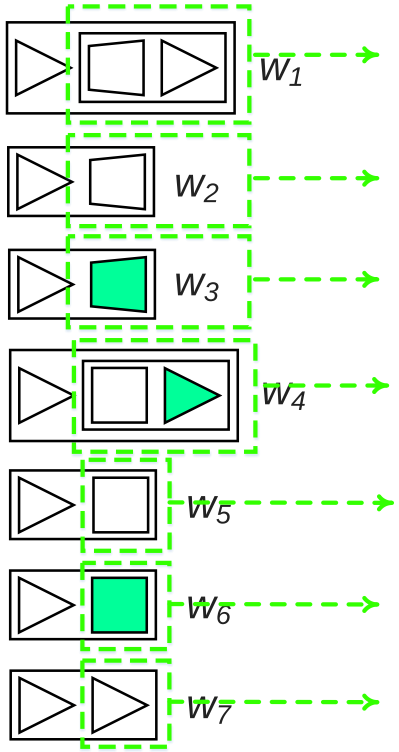
Anticipations creation and selection



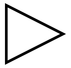
Composite interactions

Current enacted
interaction is: 

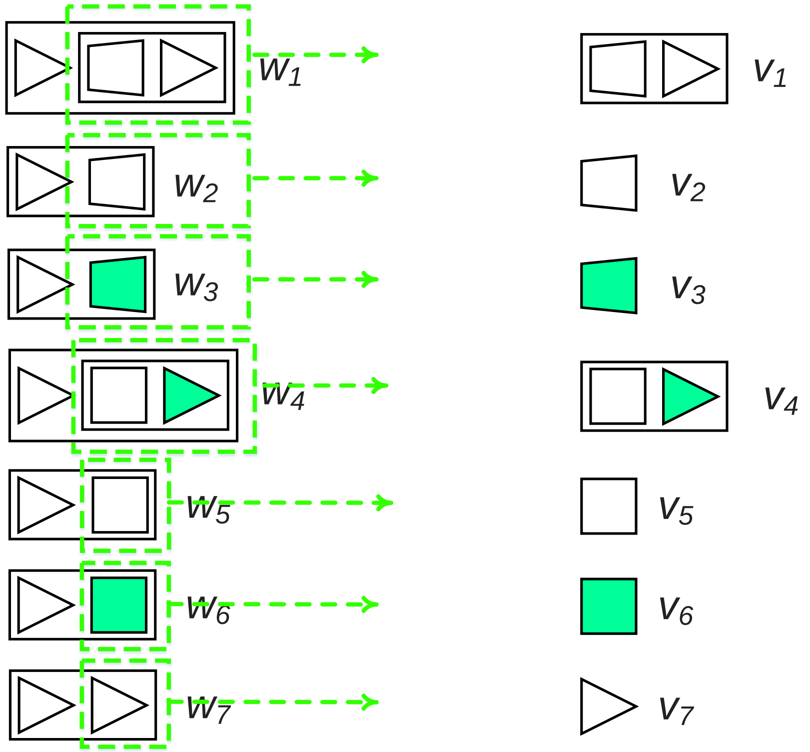
Anticipations creation and selection



Composite interactions

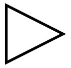
Current enacted interaction is: 

Anticipations creation and selection

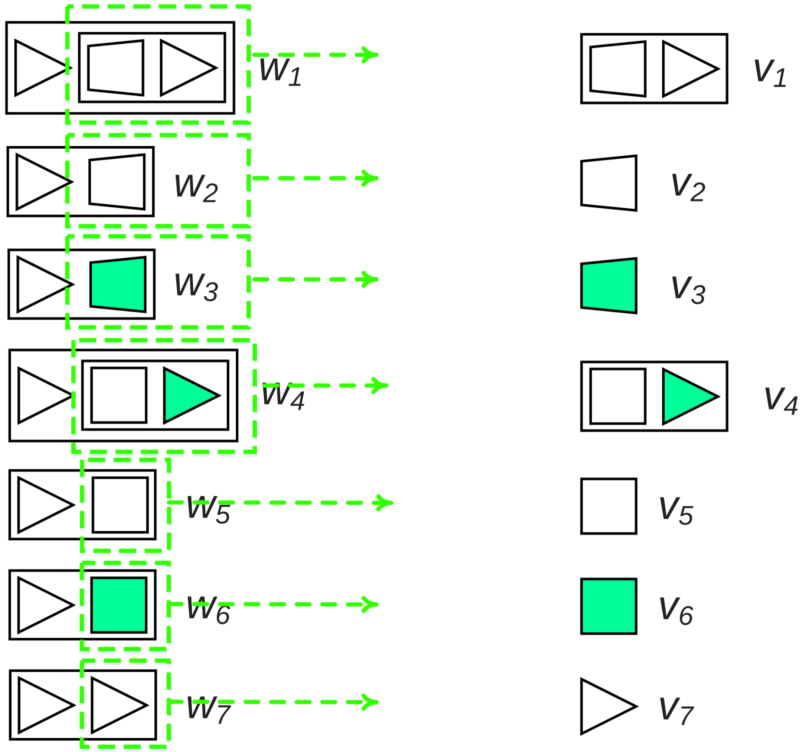


Composite interactions

Anticipations with post-interactions

Current enacted interaction is: 

Anticipations creation and selection



$p_1 = w_1 \text{ X } v_1$

$p_2 = w_2 \text{ X } v_2$

$p_3 = w_3 \text{ X } v_3$

$p_4 = w_4 \text{ X } v_4$

$p_5 = w_5 \text{ X } v_5$

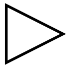
$p_6 = w_6 \text{ X } v_6$

$p_7 = w_7 \text{ X } v_7$

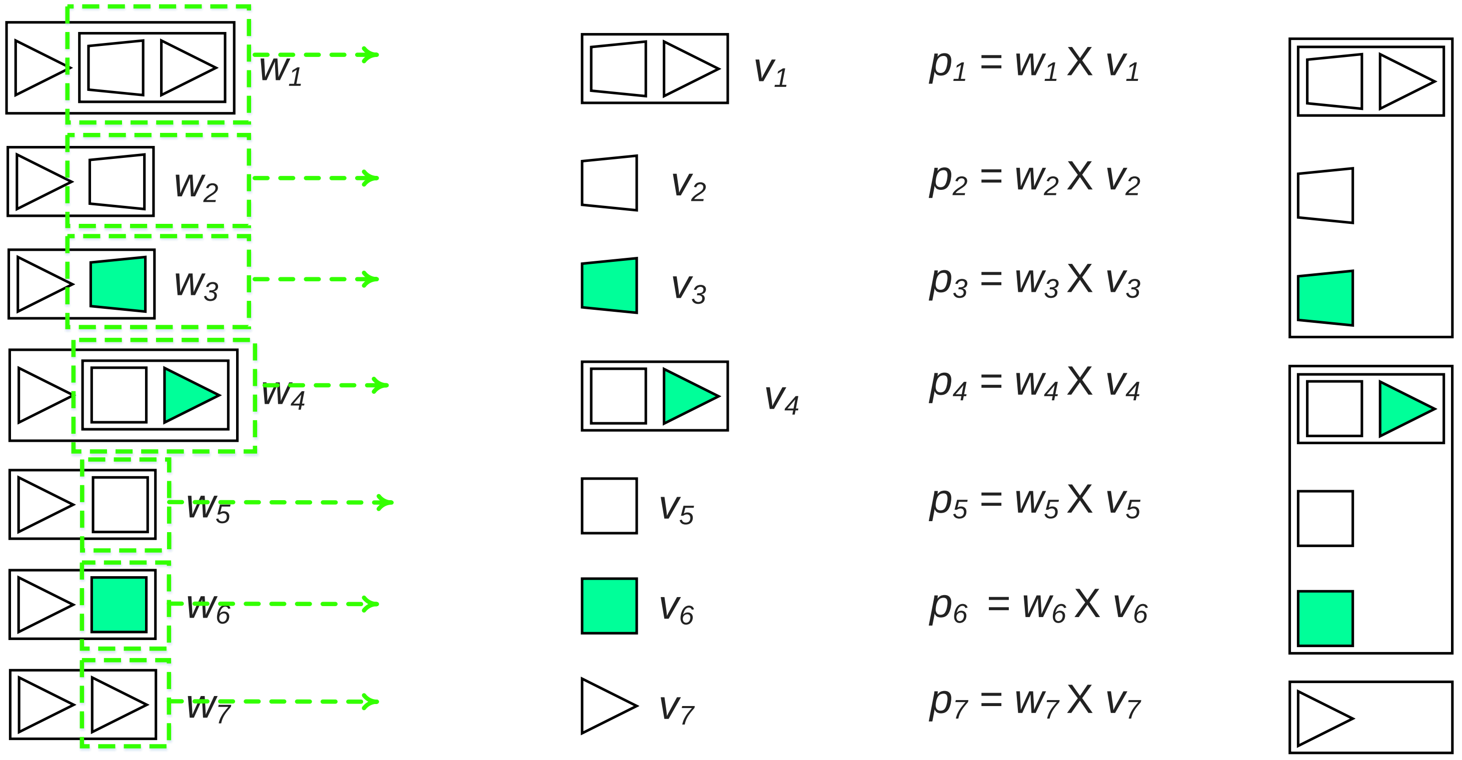
Composite interactions

Anticipations with post-interactions

Proclivities of anticipations calculation

Current enacted interaction is: 

Anticipations creation and selection



Composite interactions

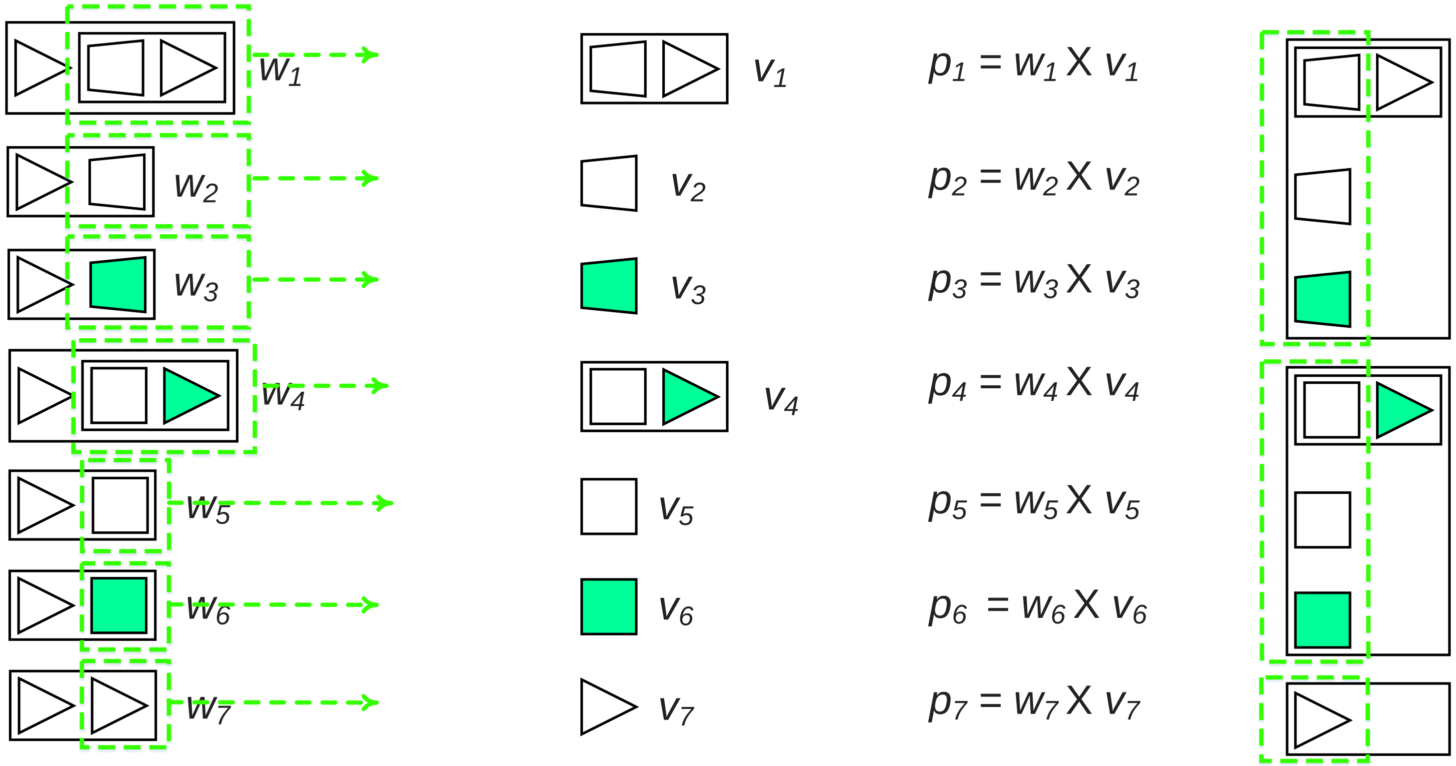
Current enacted interaction is: ▷

Anticipations with post-interactions

Proclivities of anticipations calculation

Partial similar anticipations

Anticipations creation and selection



Composite interactions

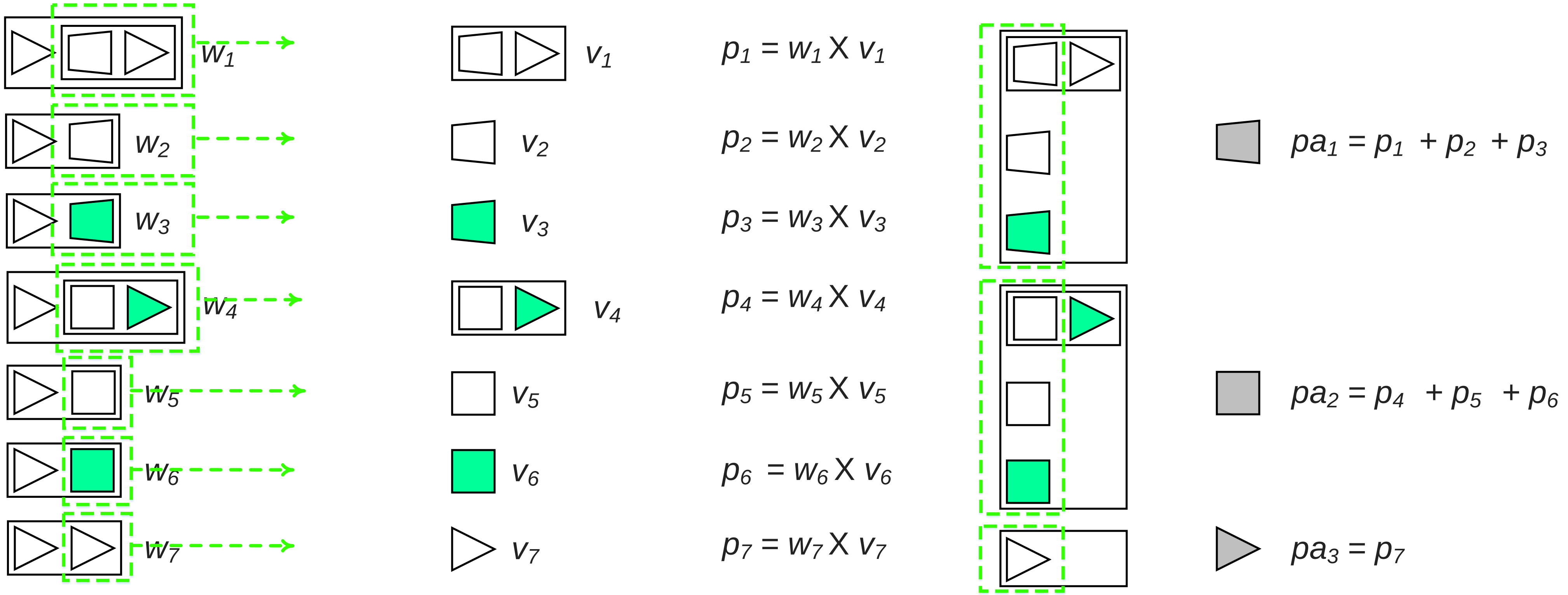
Anticipations with post-interactions

Proclivities of anticipations calculation

Partial similar anticipations

Current enacted interaction is:

Anticipations creation and selection



Composite interactions

Anticipations with post-interactions

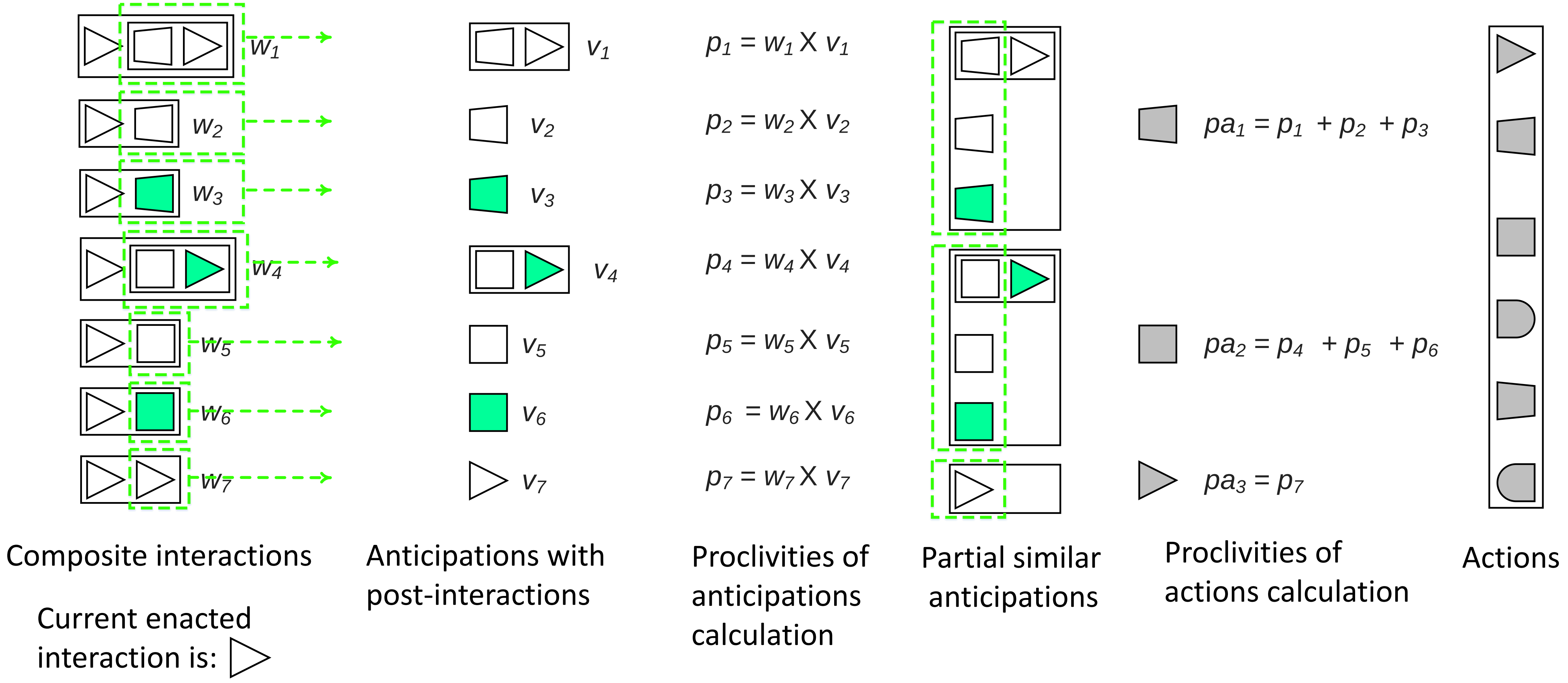
Proclivities of anticipations calculation

Partial similar anticipations

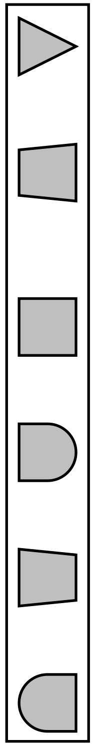
Proclivities of actions calculation

Current enacted interaction is: ▷

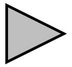
Anticipations creation and selection



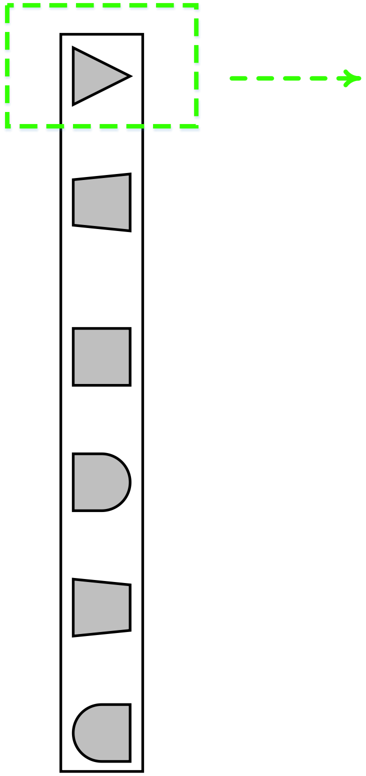
Enacting intended Interaction

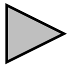


Actions

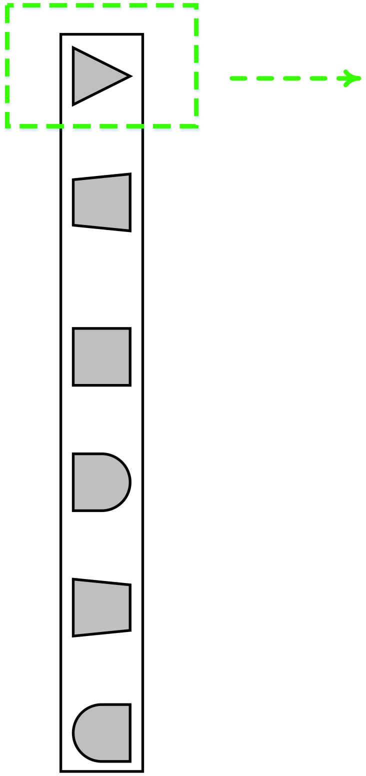
action: 
has the biggest
proclivity

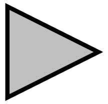
Enacting intended Interaction

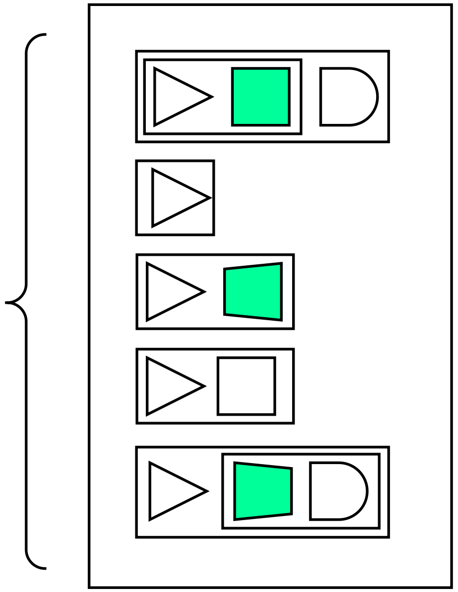


action: 
has the biggest
proclivity

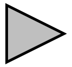
Enacting intended Interaction





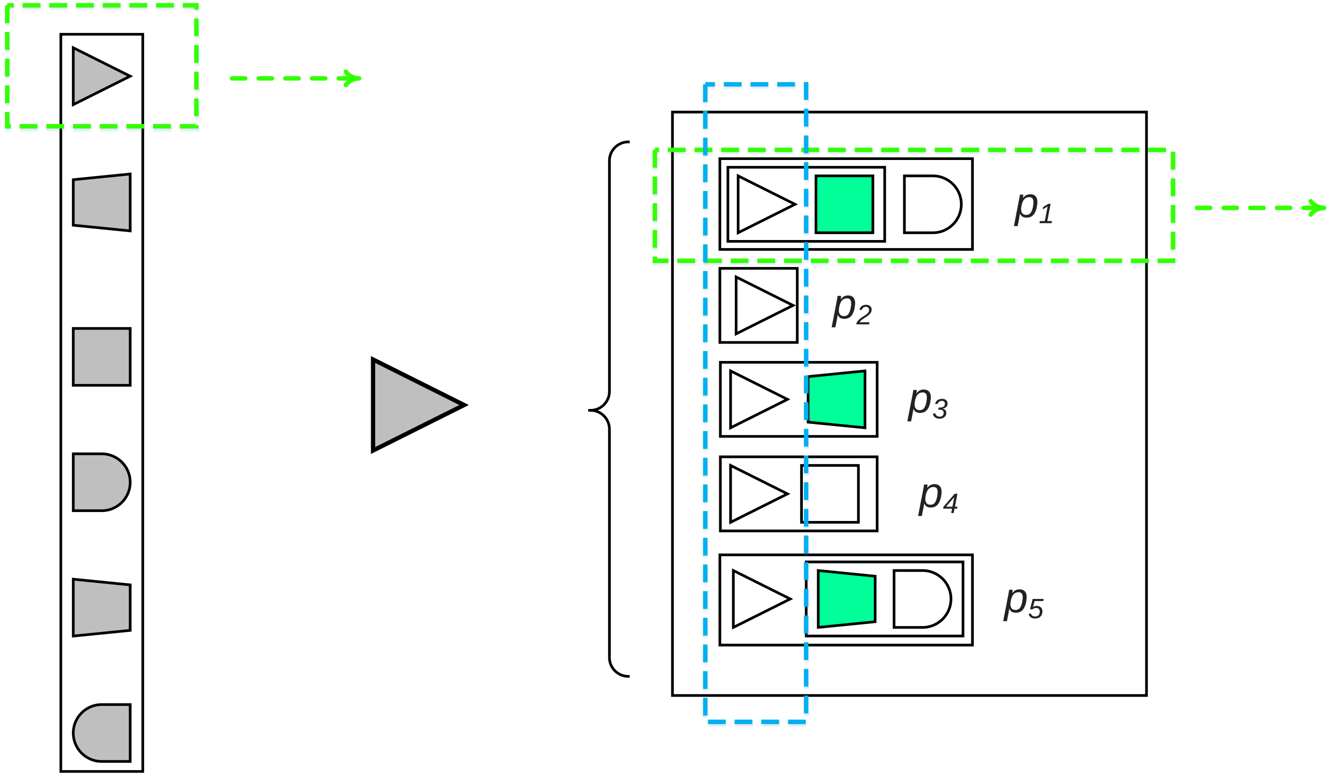


Actions

action: 
has the biggest
proclivity

Participation interactions
for intending

Enacting intended Interaction

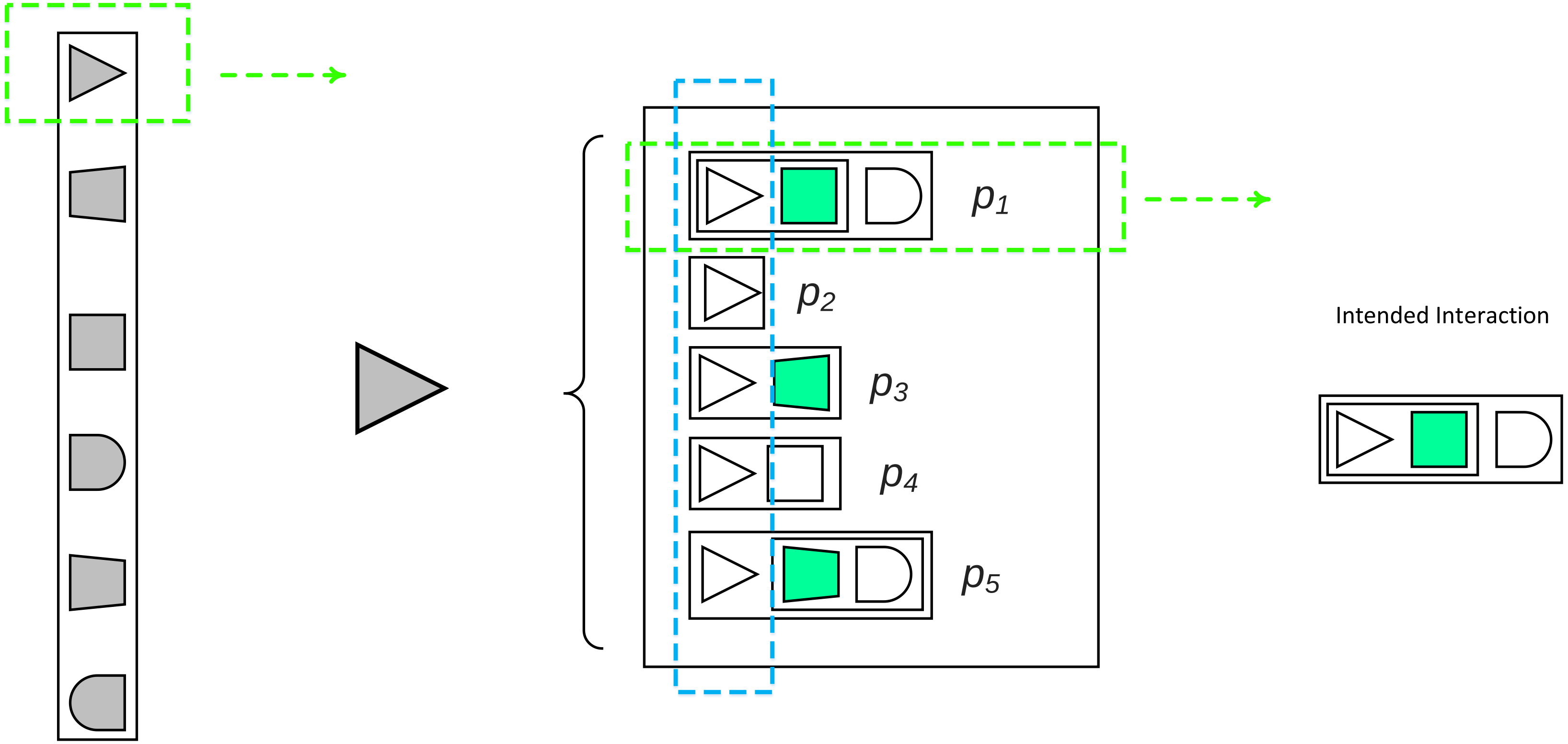


Actions

action:
has the biggest
proclivity

Participation interactions
for intending

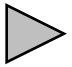
Enacting intended Interaction



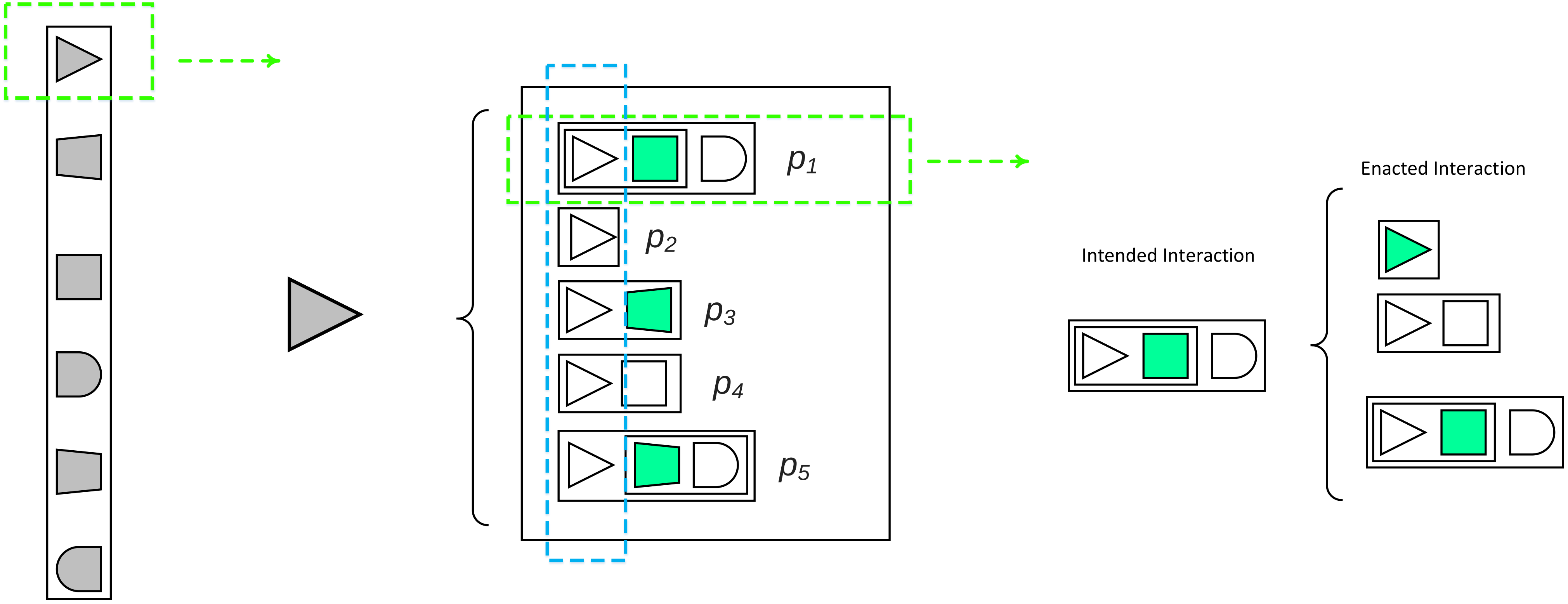
Actions

Participation interactions for intending

Intended interaction

action: 
has the biggest proclivity

Enacting intended Interaction

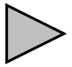


Actions

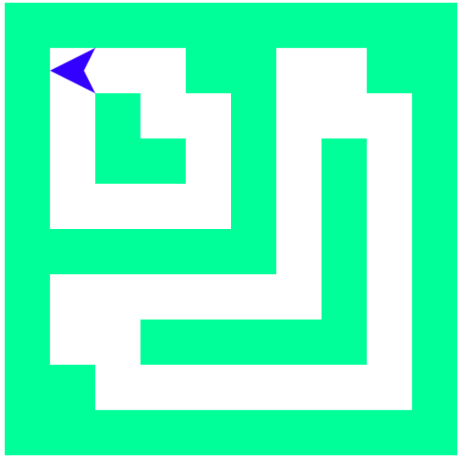
Participation interactions for intending

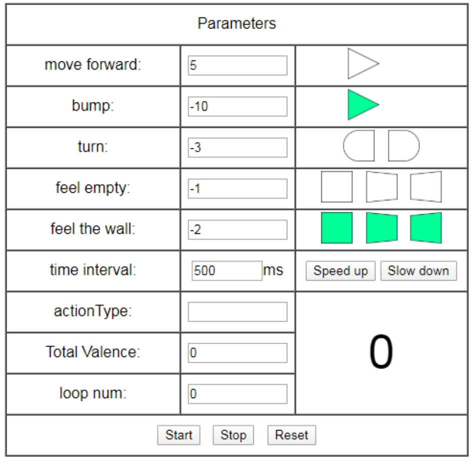
Intended interaction

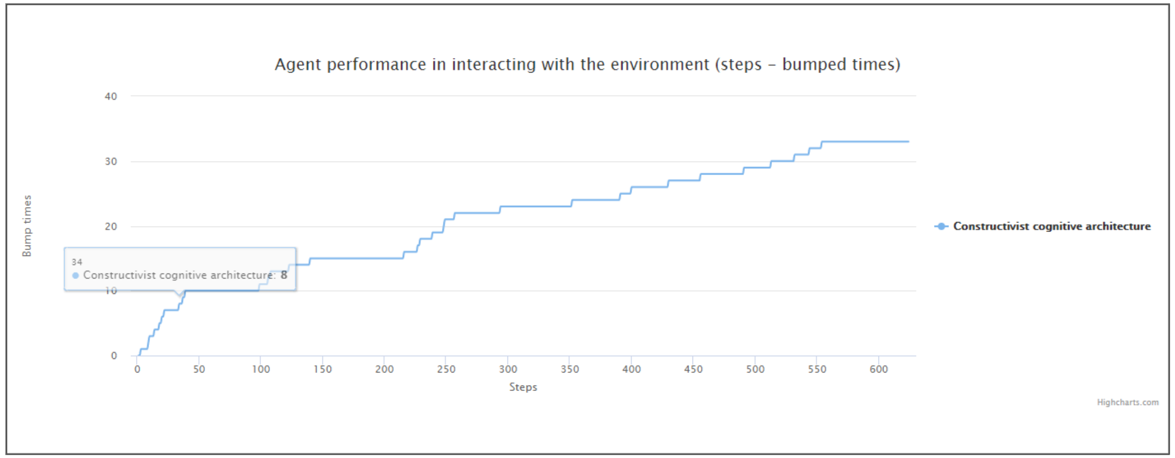
Possible enacted interactions

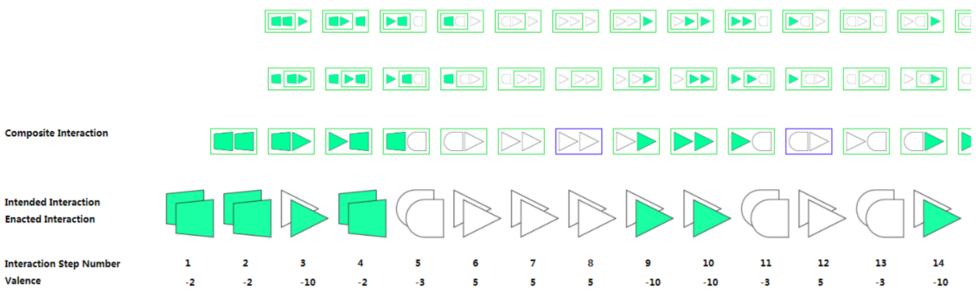
action: 
has the biggest proclivity

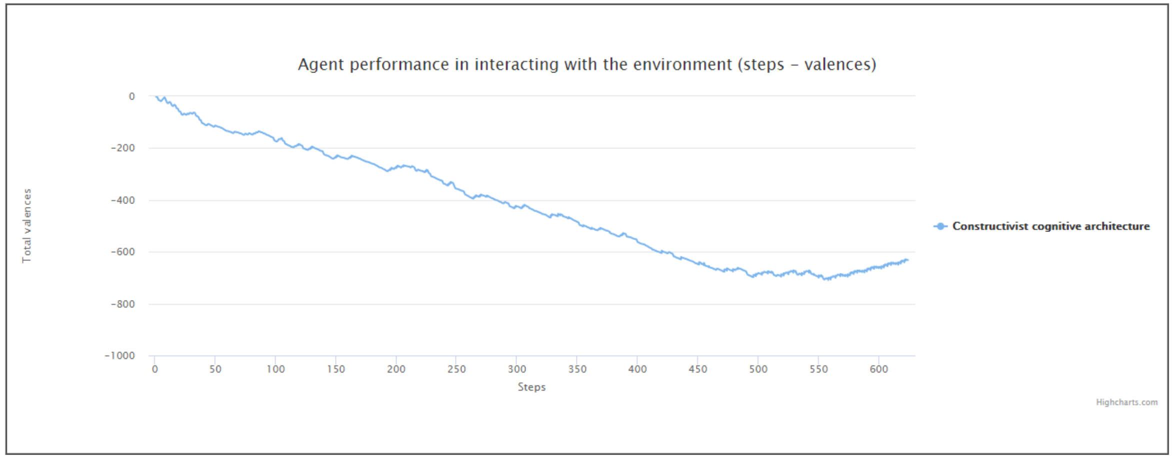
Toolkit of Generating and Analyzing Interaction Traces (GAIT)



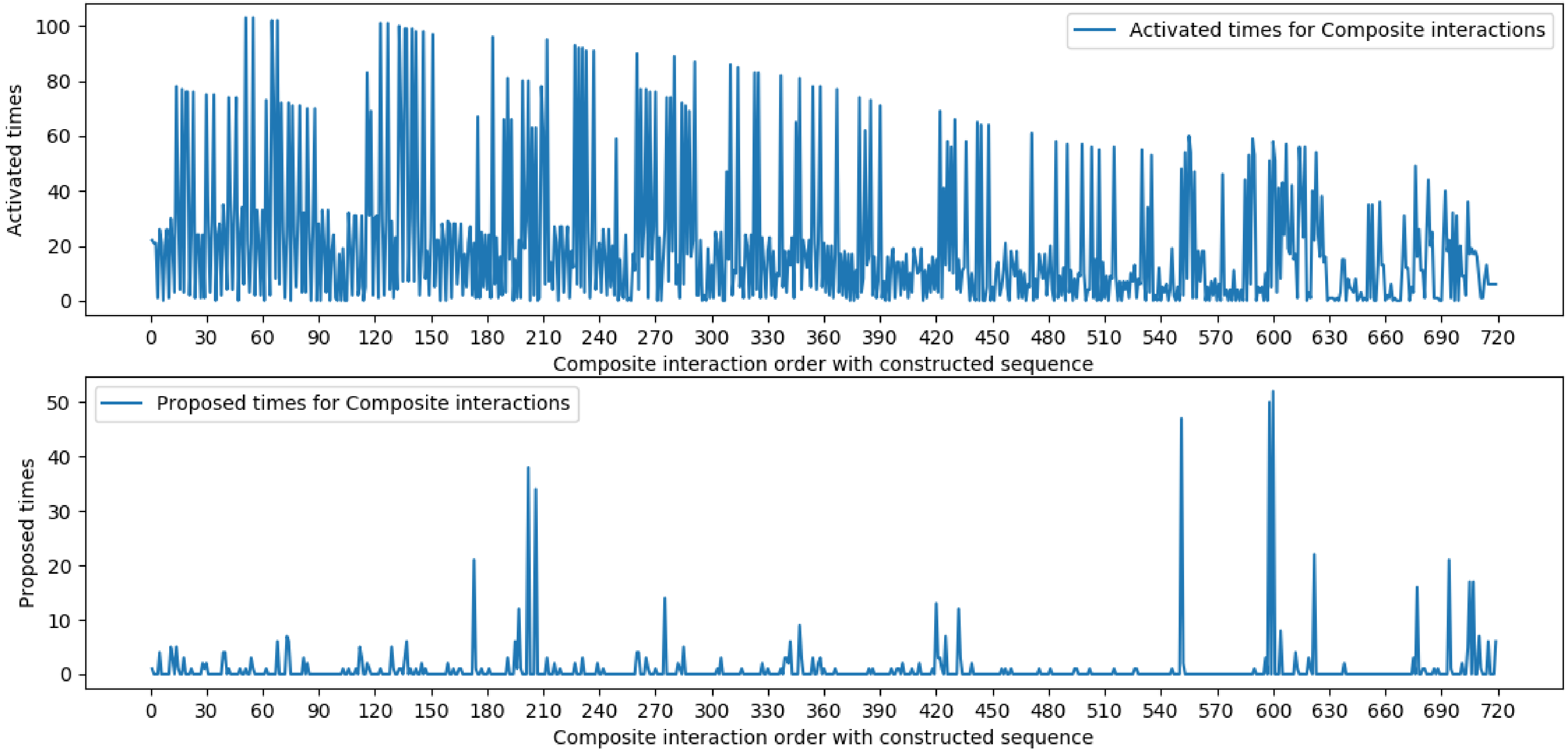




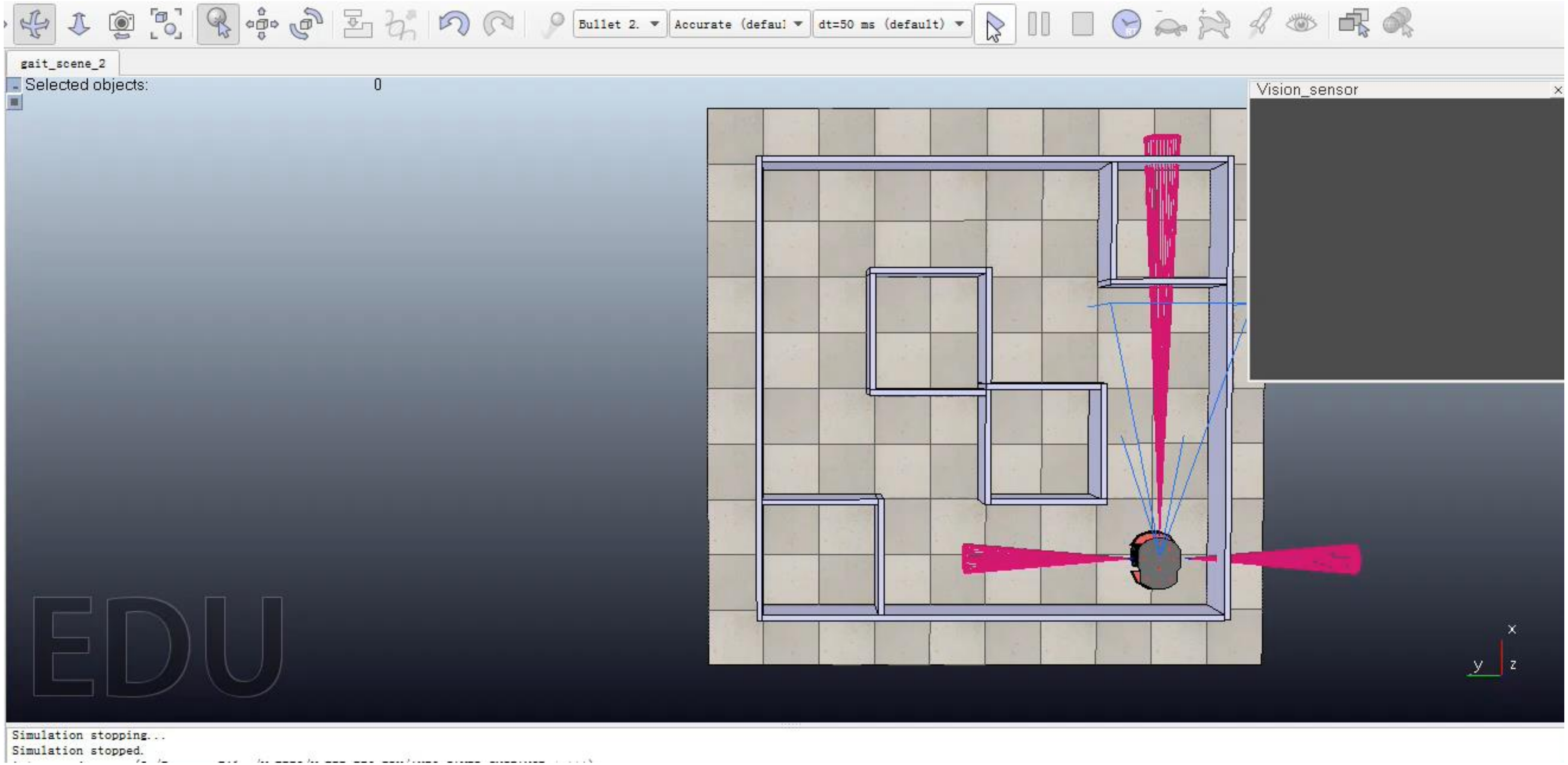




Results: usage rate of the composite interactions



A simulation GAIT with robot in VREP



Contributions

- Propose a bottom-up hierarchical sequence learning algorithm with constructivist paradigm, as a solution for autonomous and continuous learning of environment representations and agent's self-adaptation.

- Design and develop an implementation of toolkit for agent autonomously generating and analyzing interaction (GAIT) at run-time, which facilitates to observe the detailed learning process for agent interacting with the environment and each structured behaviors it has learned within each decision-making.

Perspective

- Agent has to retrospect all previous learned composite interactions to retrieve the ones whose pre-interactions are matched with the current enacted interaction in each decision-making.
- The valence assignment for different experiments is an important issue in constructivist learning.
- With memorizing patterns that could improve the learning efficiency and eliminating composite interactions that probably will not use to simplify the activation and proposition processes in the future.
- Application of GAIT: Prepare the interfaces for the Integration of algorithm and framework GAIT with Robots in ROS and VREP.

Sources

- Sources of framework with installation guide in GitHub:

  https://github.com/xuejianyong/Interaction_Traces_Analysis_Toolkit

- The simulation of GAIT in VREP with Python

  https://www.youtube.com/watch?v=w93ThiBKU2k&t=8s

Merci