Building General-Purpose Robot Autonomy

A Progressive Roadmap

Yuke Zhu

June 16, 2020

Stanford University

NVIDIA

TEXAS
The University of Texas at Austin


Photos from the Internet

Traditional form of automation

Credit: WIRED

General-purpose robot hardware

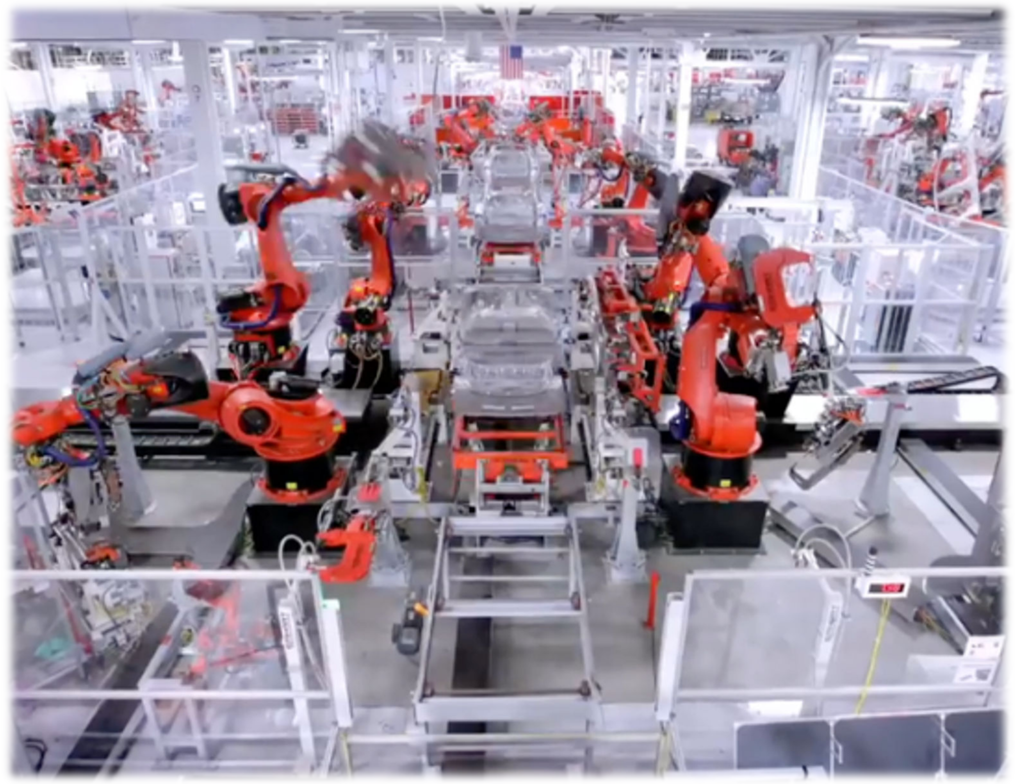
Credit: Kinova Robotics

My Long-Term Research Goal

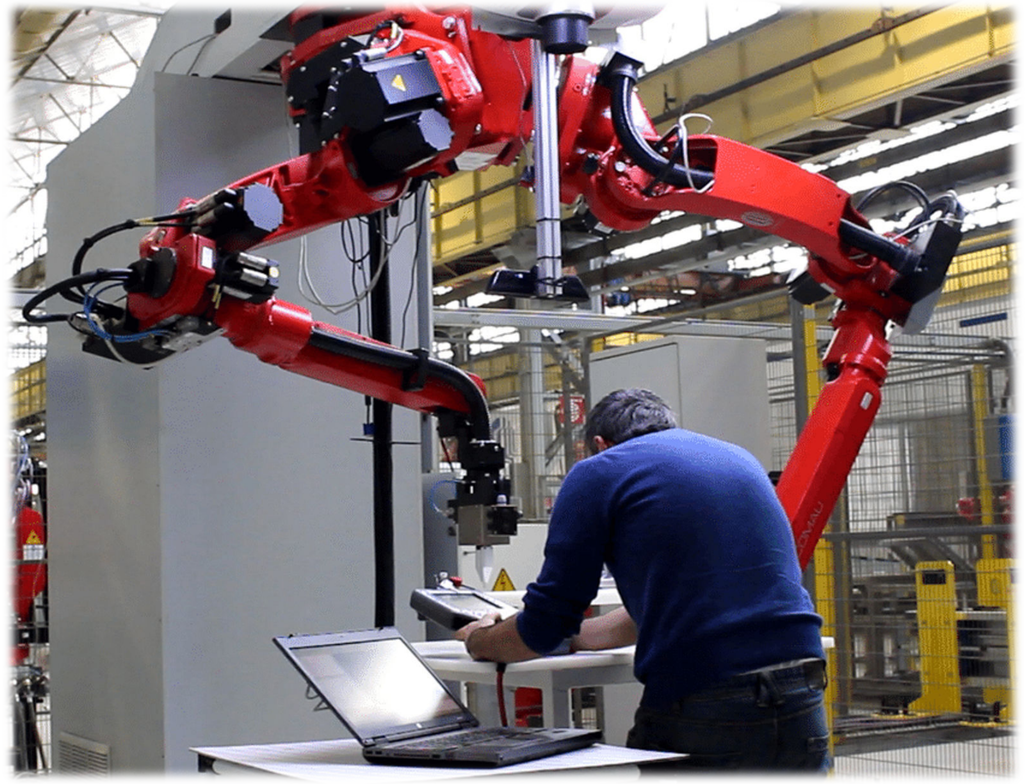
Artificial Intelligence (AI) → Intelligence Augmentation (IA)

building **robot intelligence** to enrich **human intelligence**

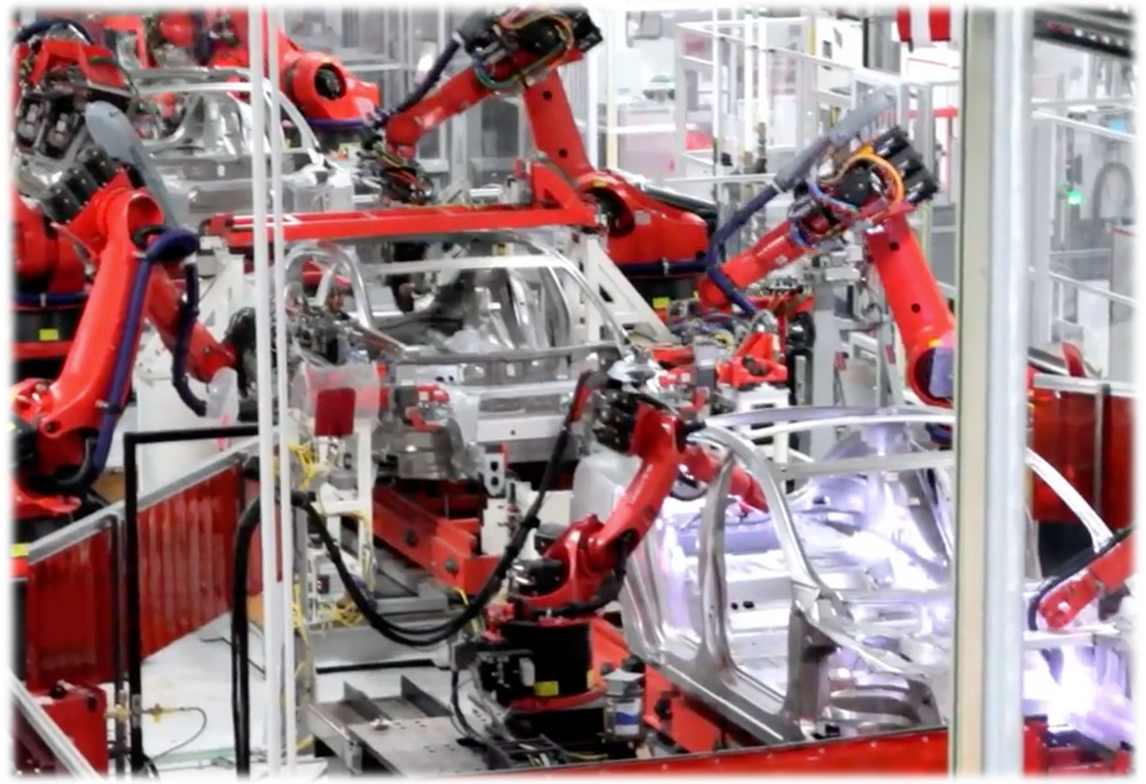
Traditional form of **robot automation**



custom-built robots



human expert programming

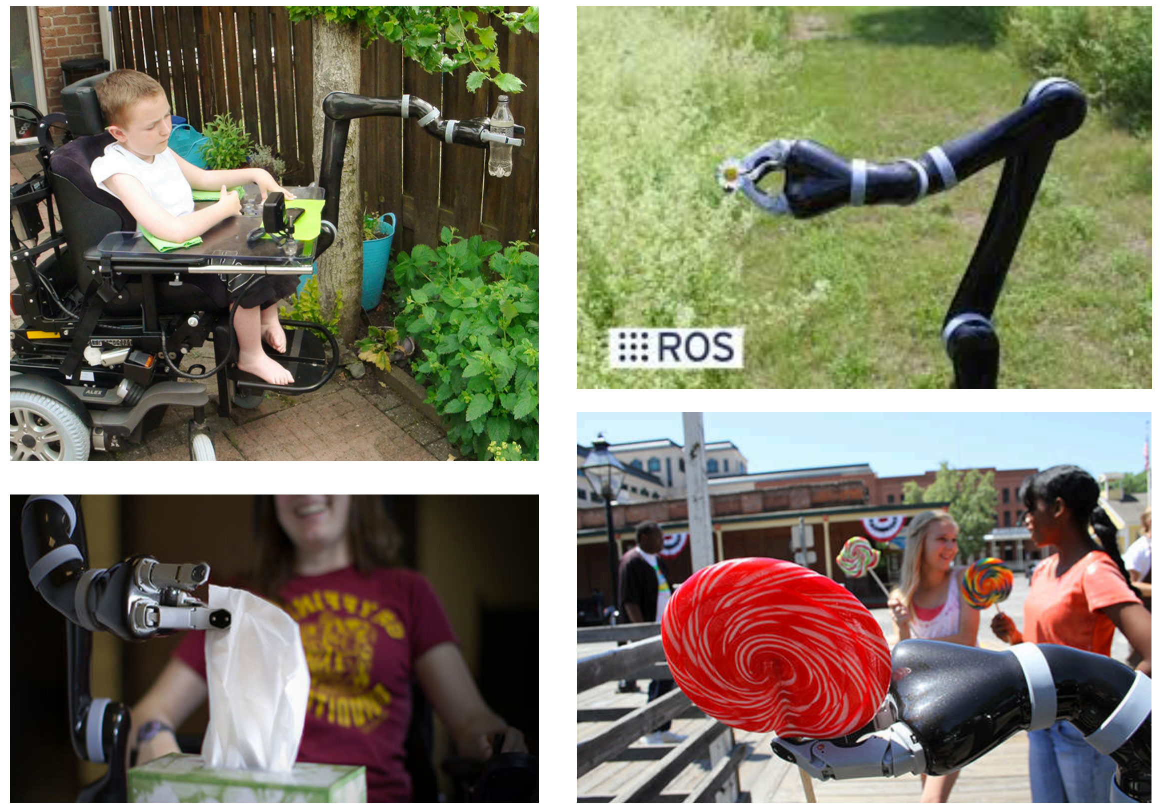


special-purpose behaviors

New form of **robot autonomy**

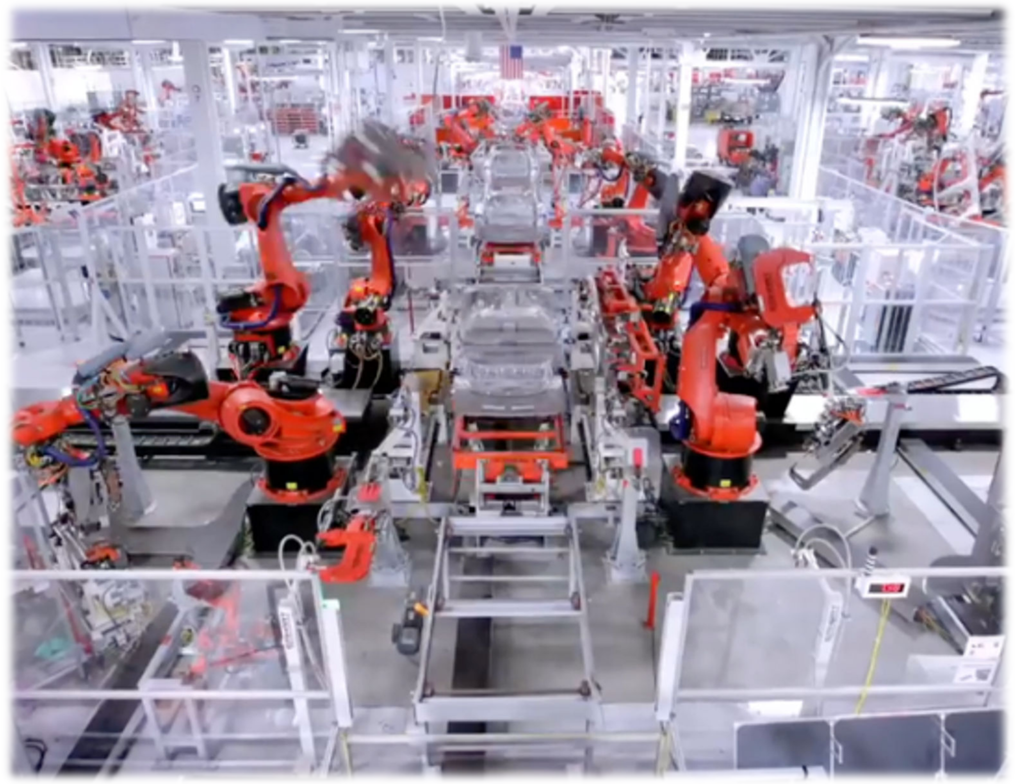
general-purpose robots

?



general-purpose behaviors

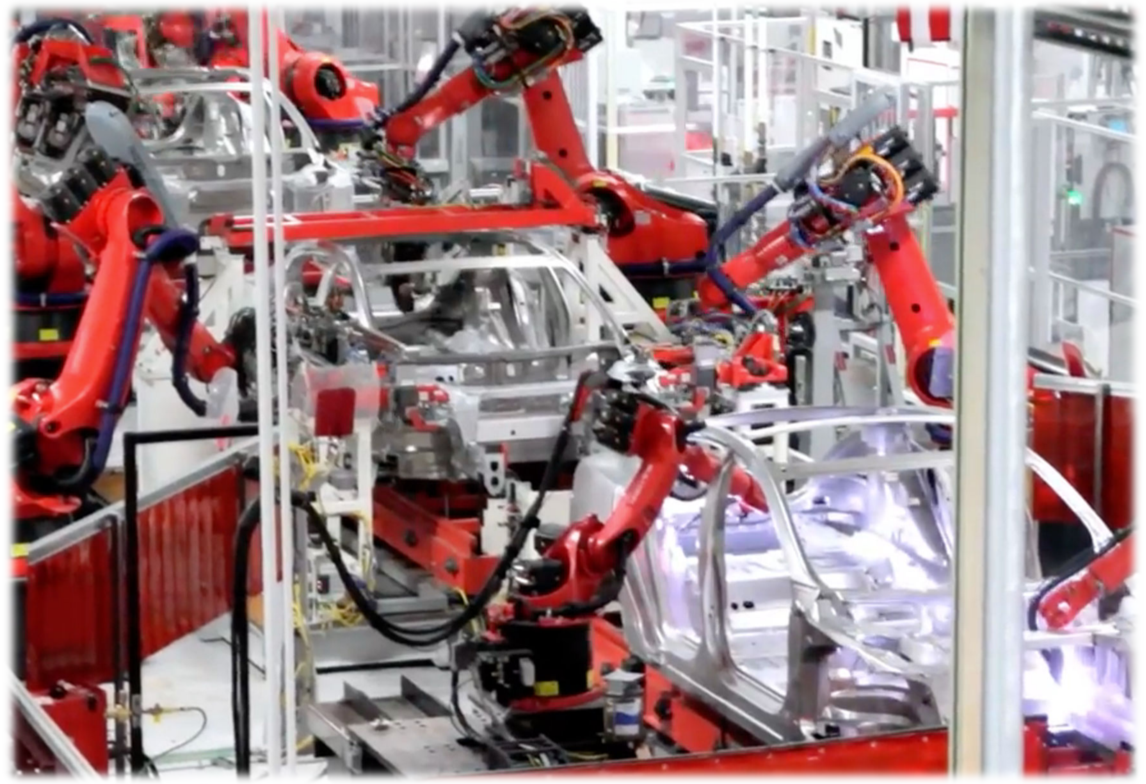
Traditional form of **automation**



custom-built robots

structured environment

human expert programming

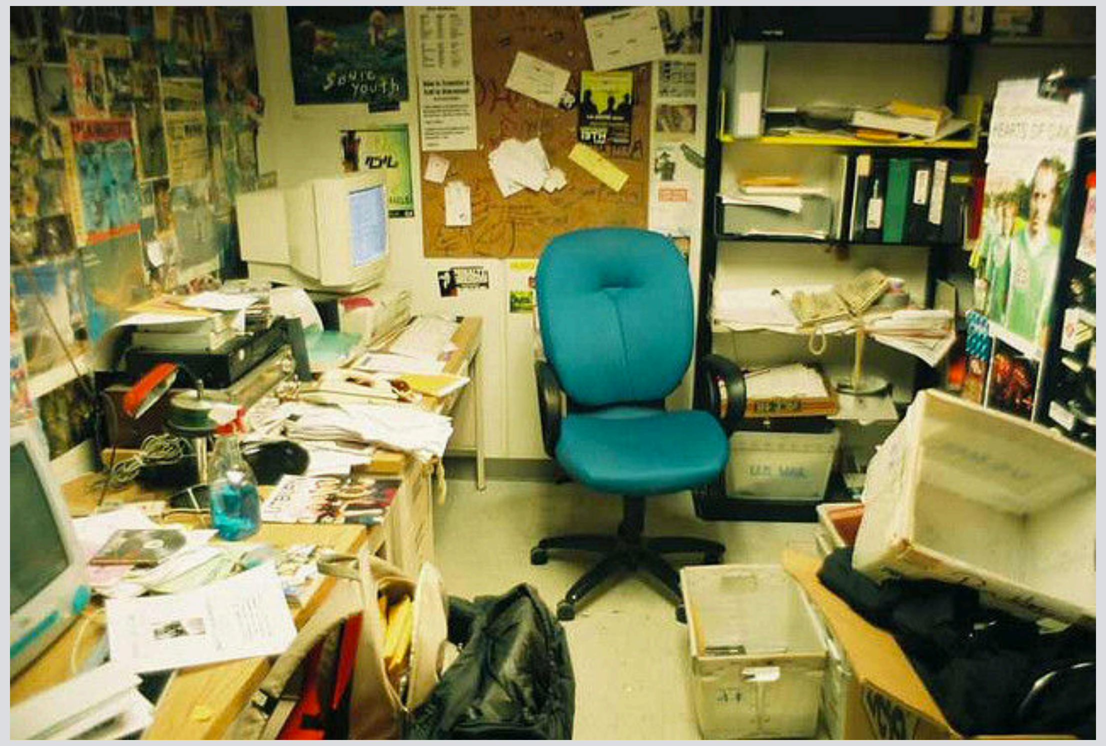


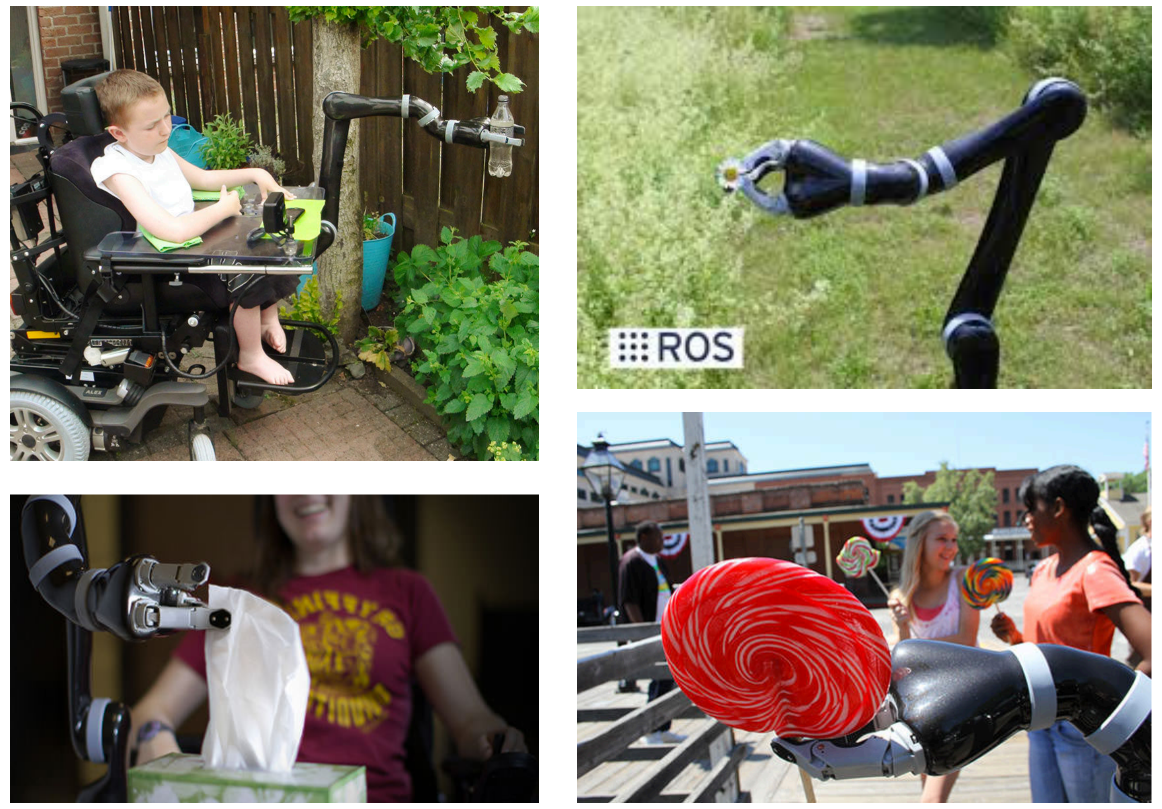
special-purpose behaviors

New form of **automation**

general-purpose robots

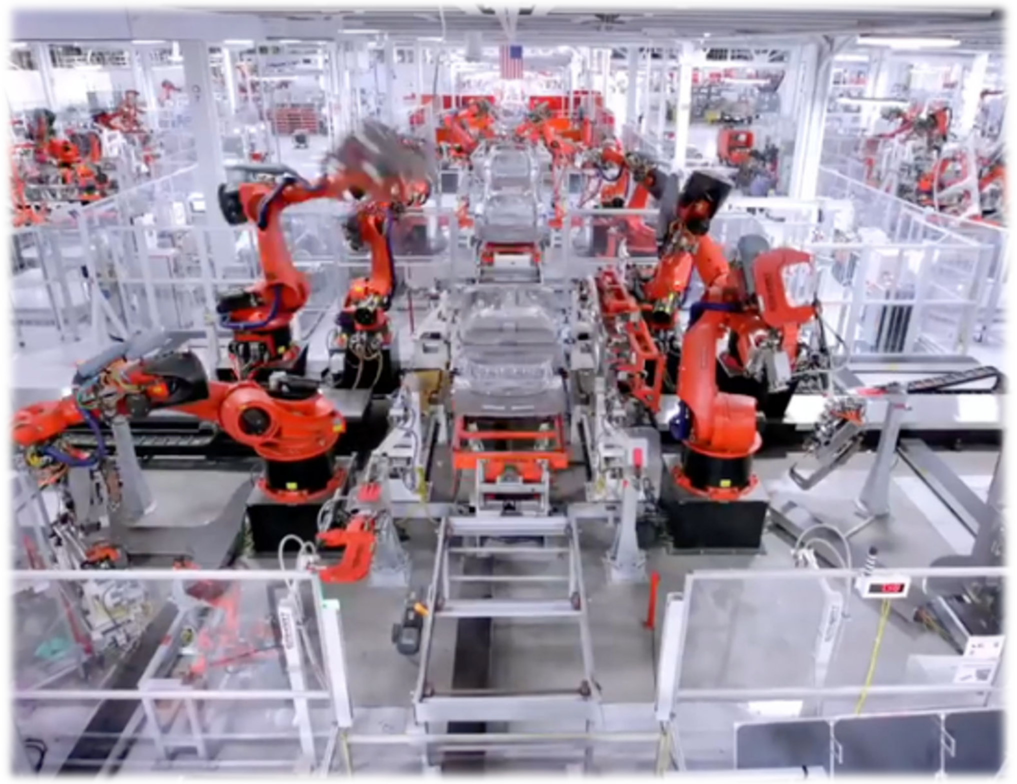
unstructured environment



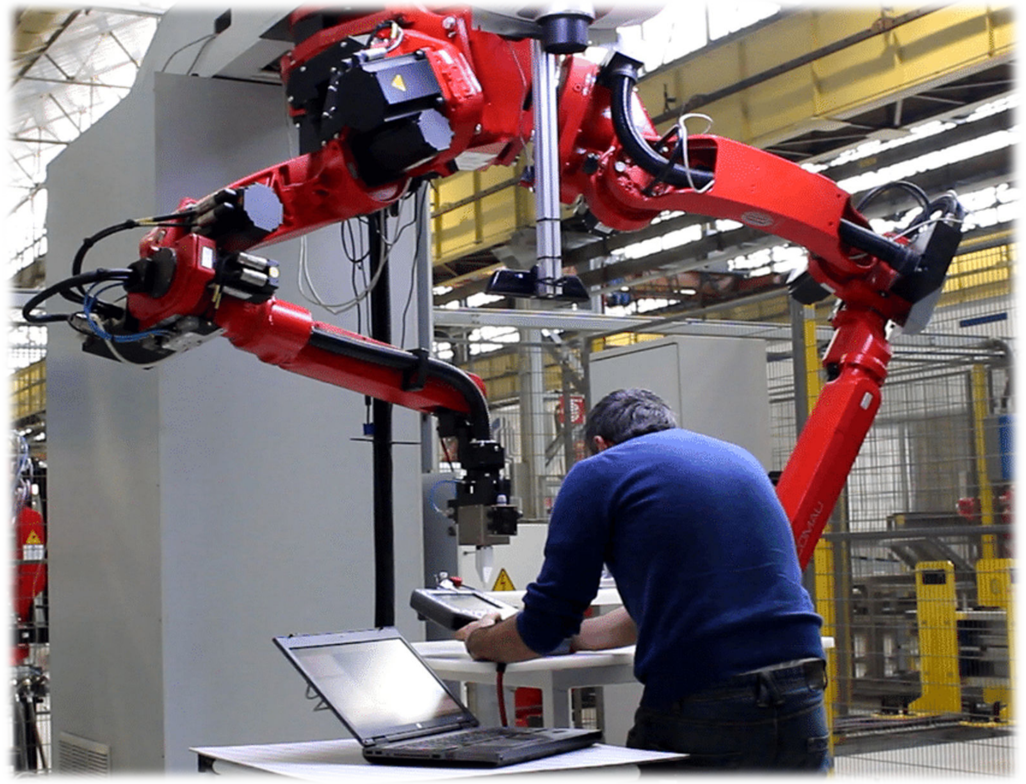


general-purpose behaviors

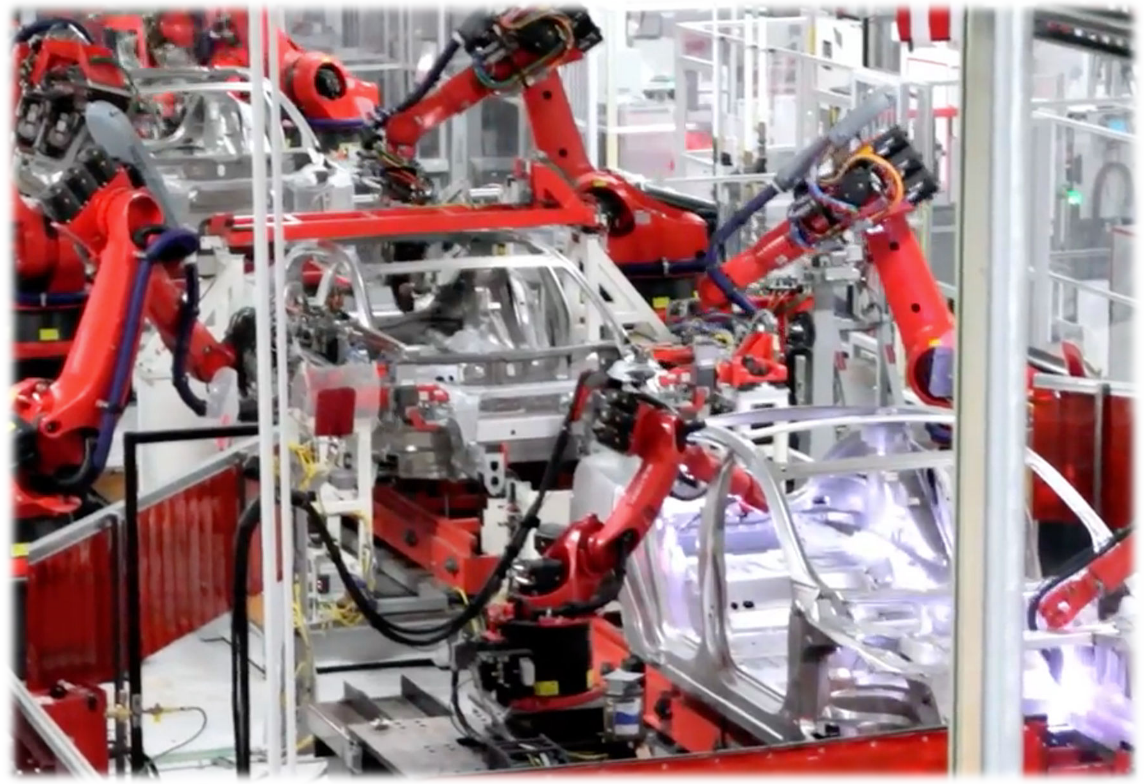
Traditional form of **automation**



custom-built
robots



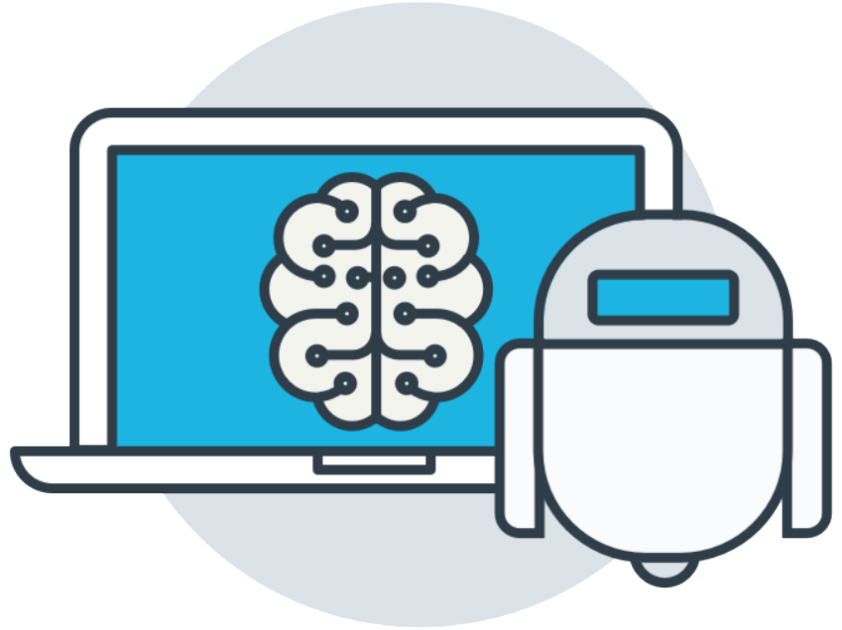
human expert
programming



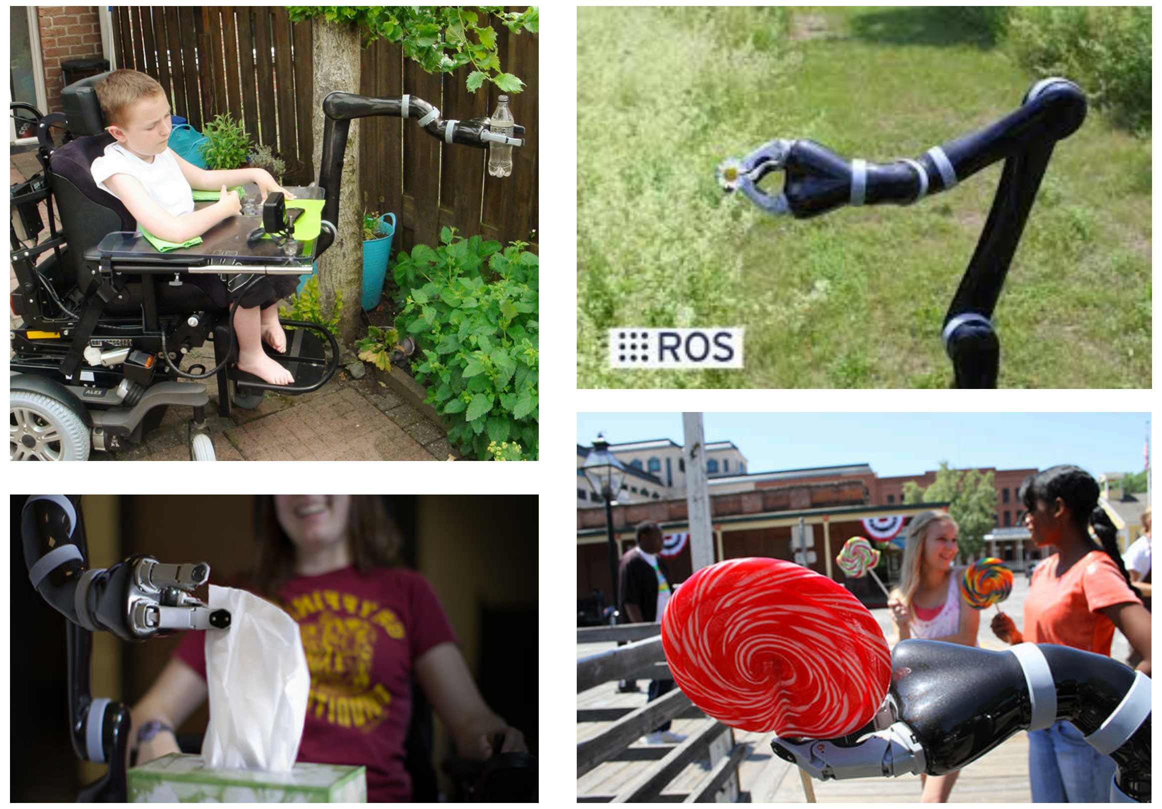
special-purpose
behaviors

New form of **automation**

general-purpose
robots

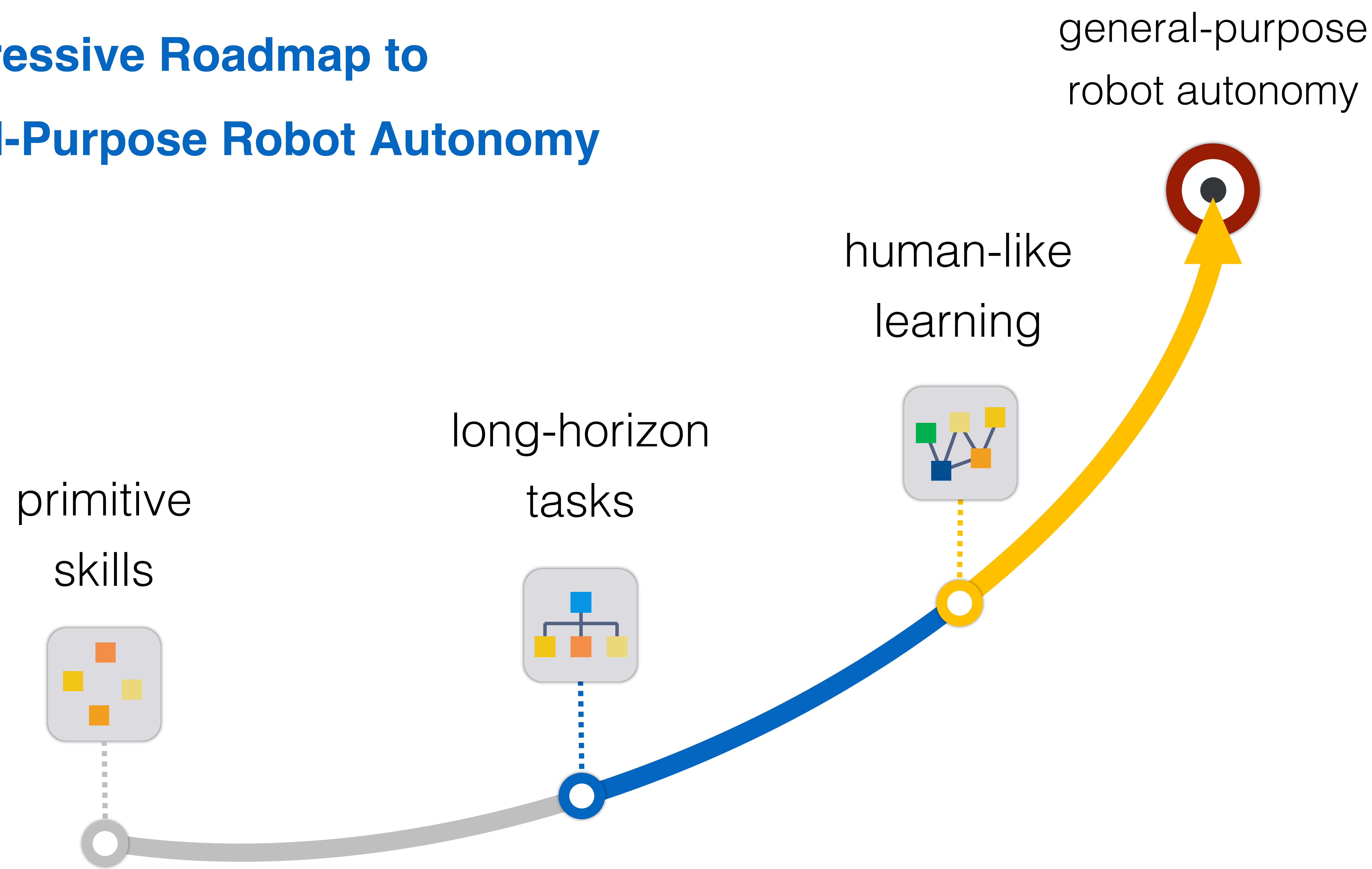


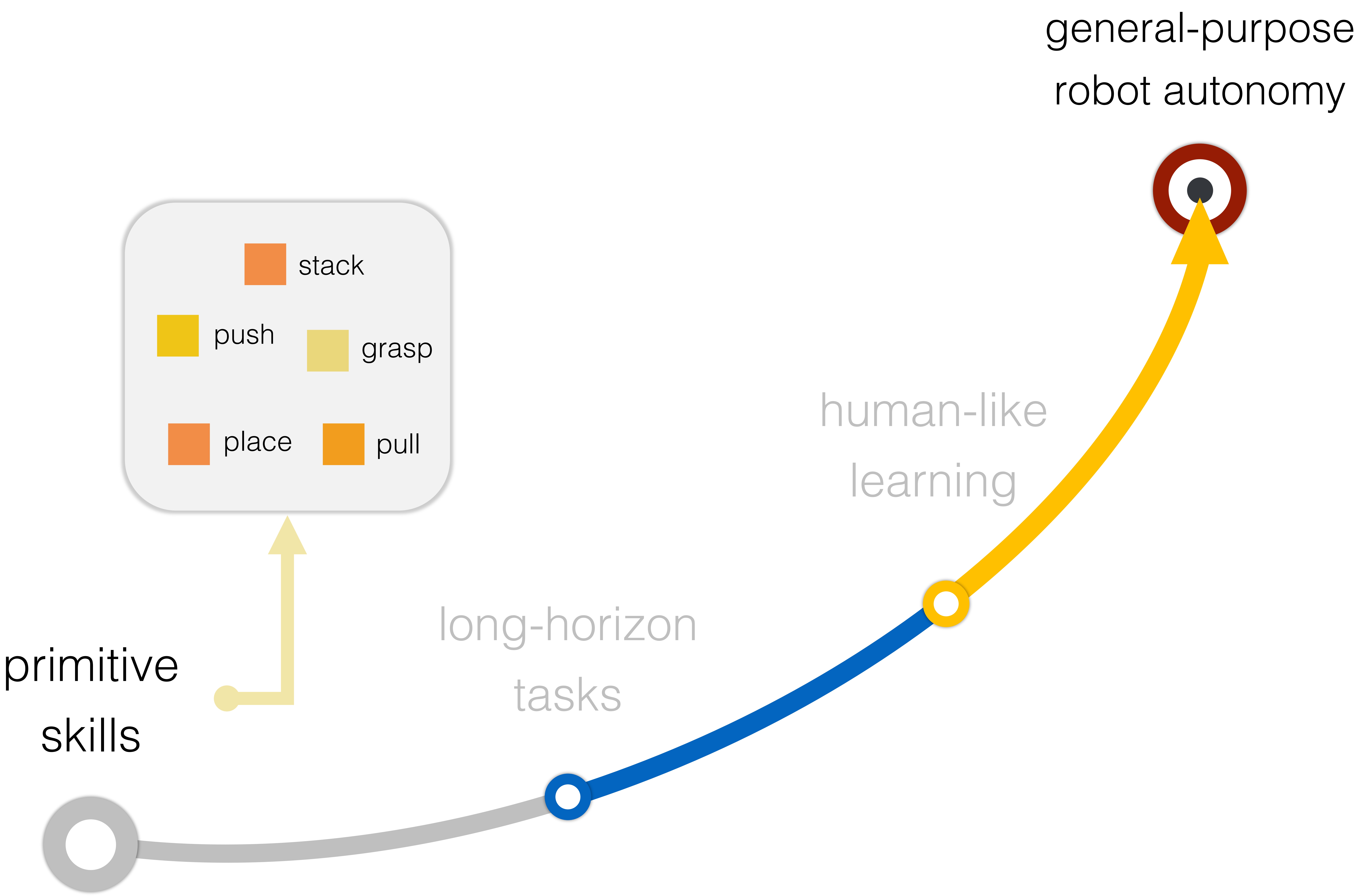
machine learning
& perception

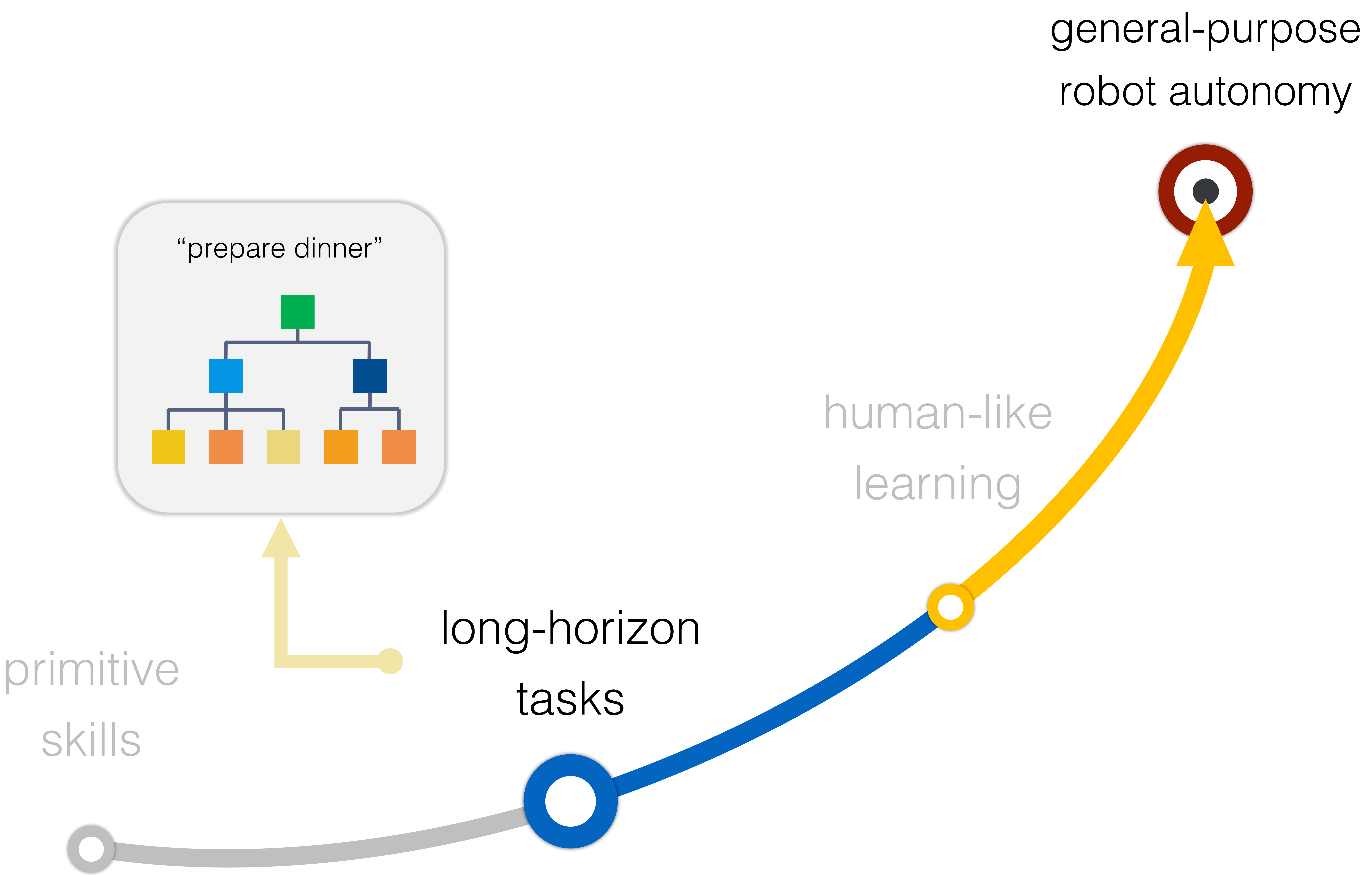


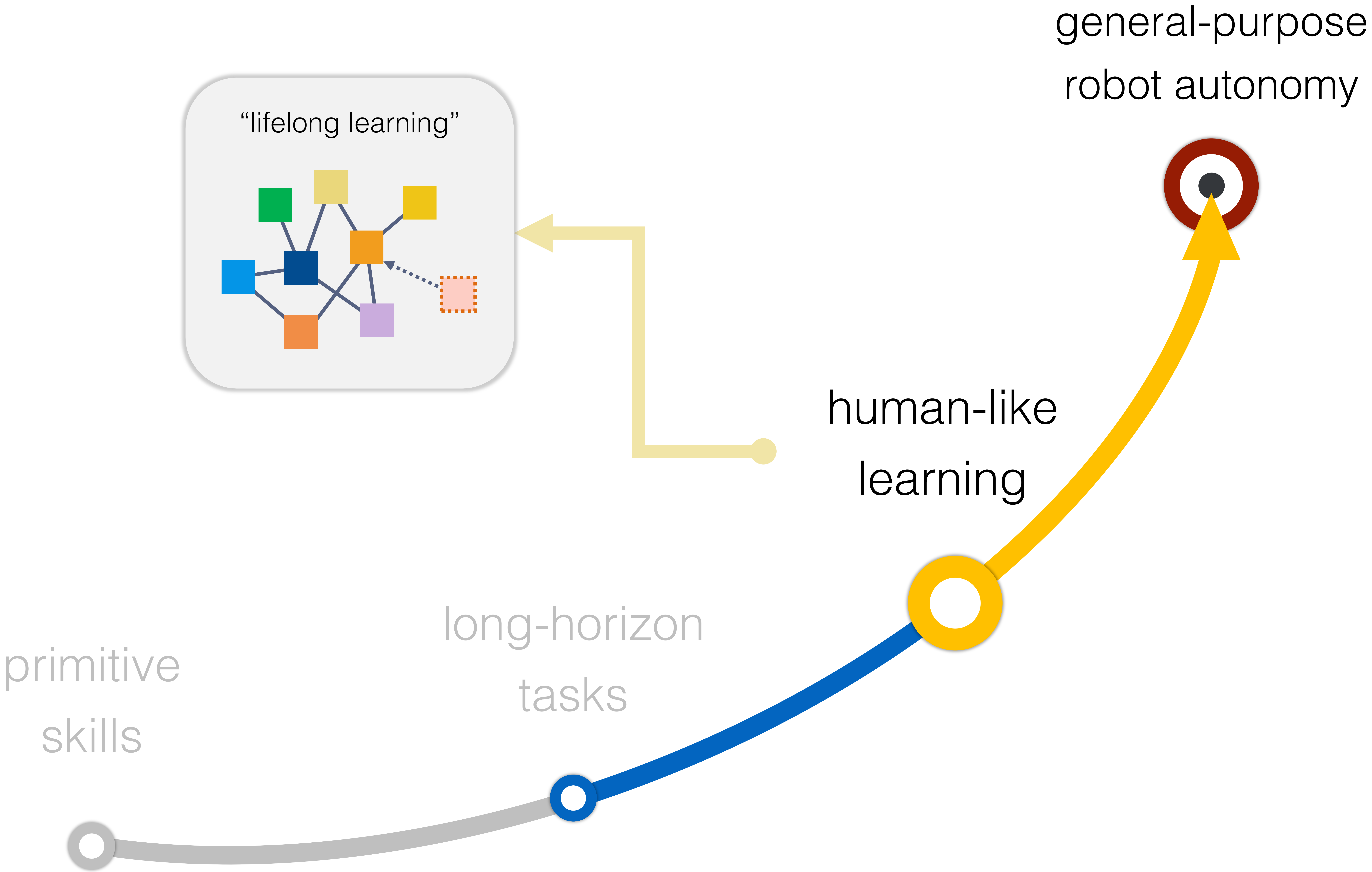
general-purpose
behaviors

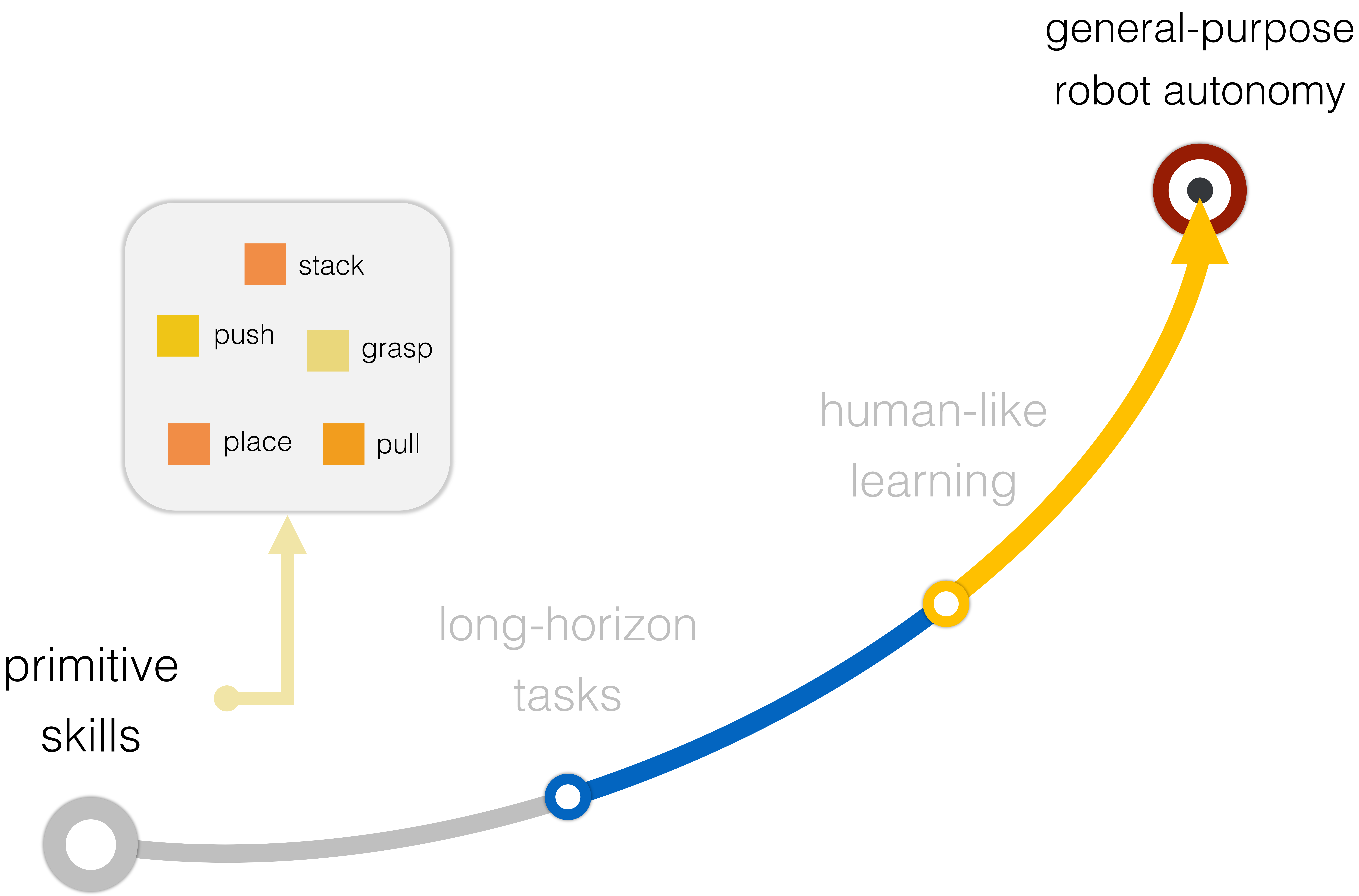
A Progressive Roadmap to General-Purpose Robot Autonomy




general-purpose
robot autonomy
stack
push
grasp
place
pull
human-like
learning
long-horizon
tasks
primitive
skills


general-purpose
robot autonomy
"prepare dinner"
human-like
learning
long-horizon
tasks
primitive
skills


general-purpose
robot autonomy
"lifelong learning"
human-like
learning
long-horizon
tasks
primitive
skills

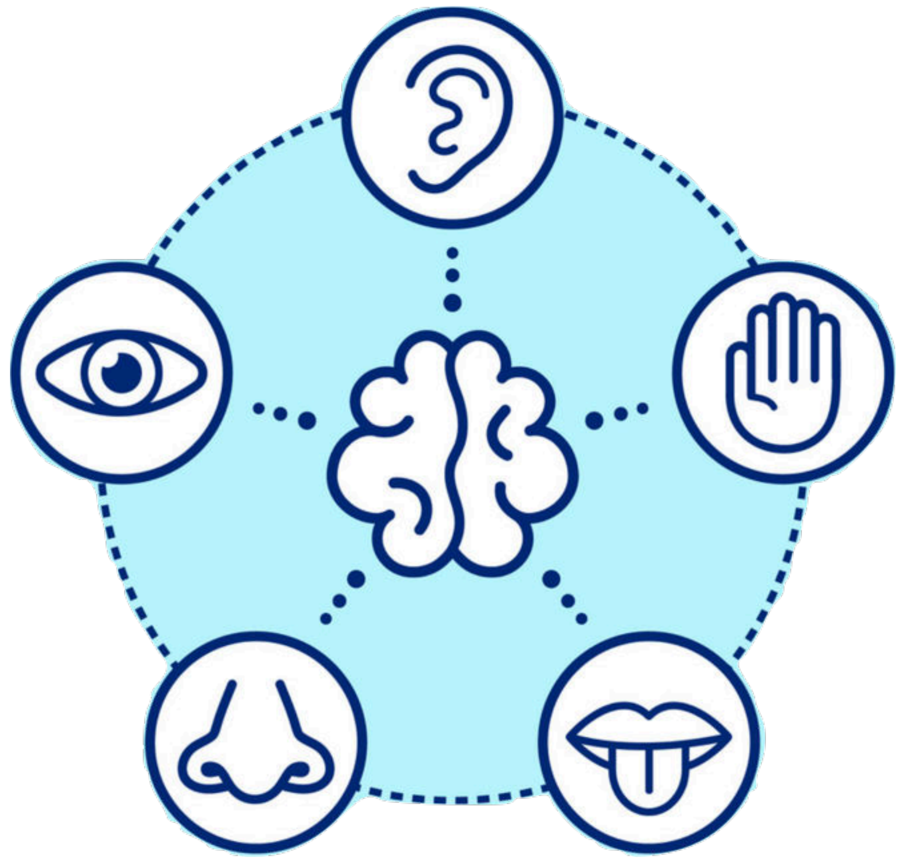

general-purpose
robot autonomy
stack
push
grasp
place
pull
human-like
learning
long-horizon
tasks
primitive
skills

Primitive Skills

perception

action

Credit: BBC Earth Lab

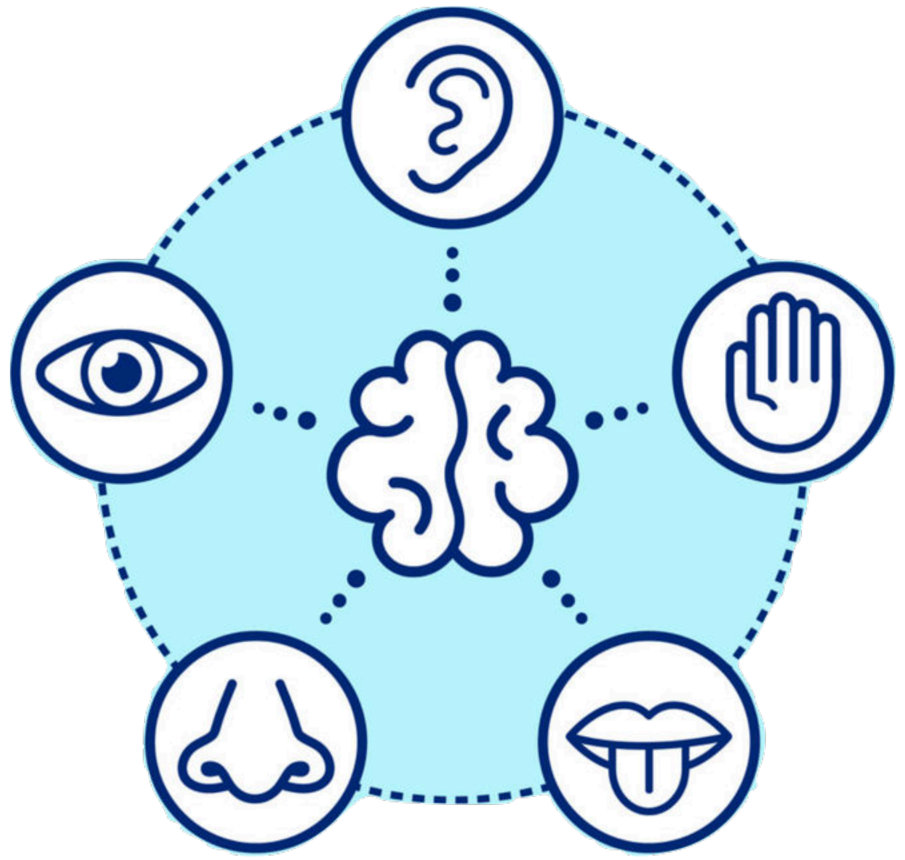


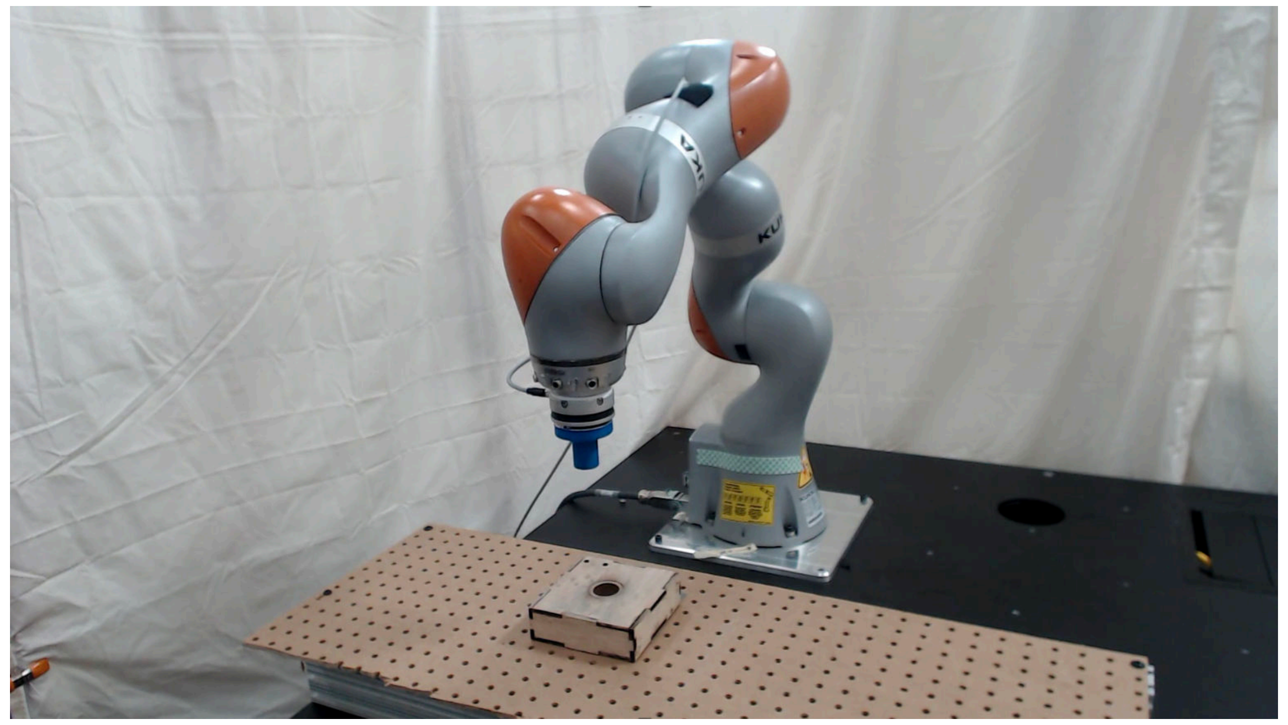
Primitive Skills

perception

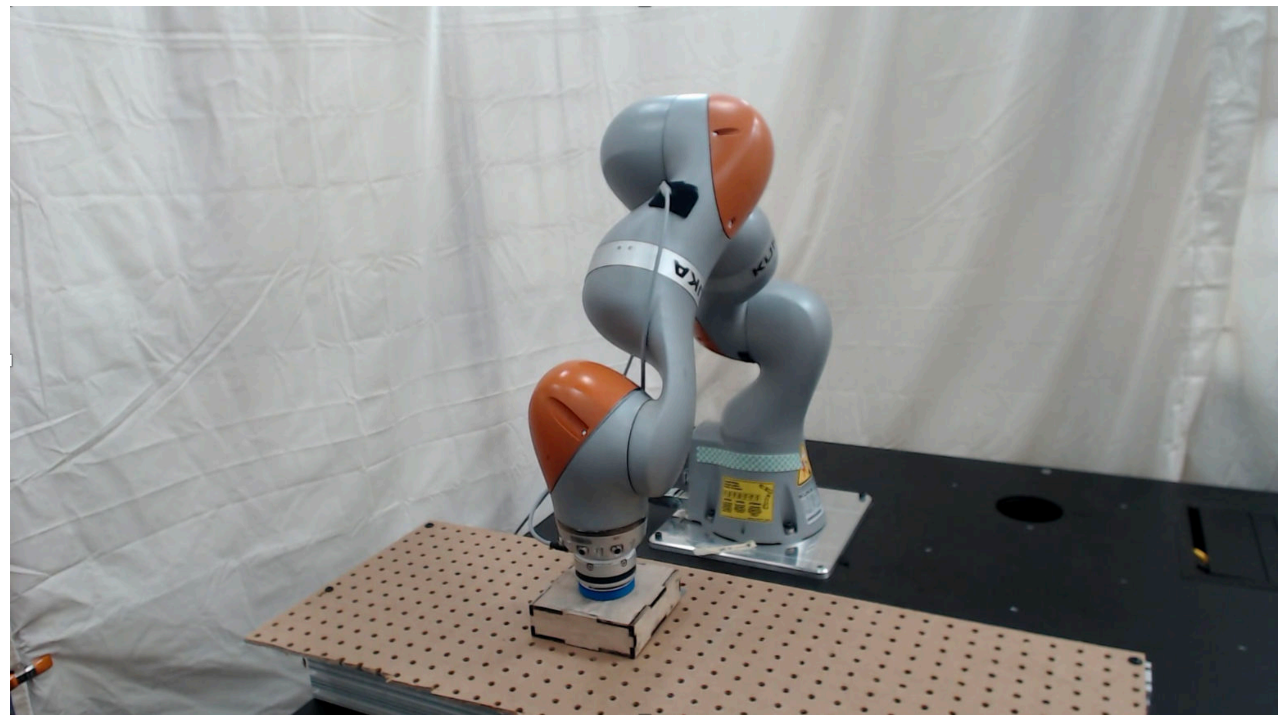
Credit: BBC Earth Lab

action





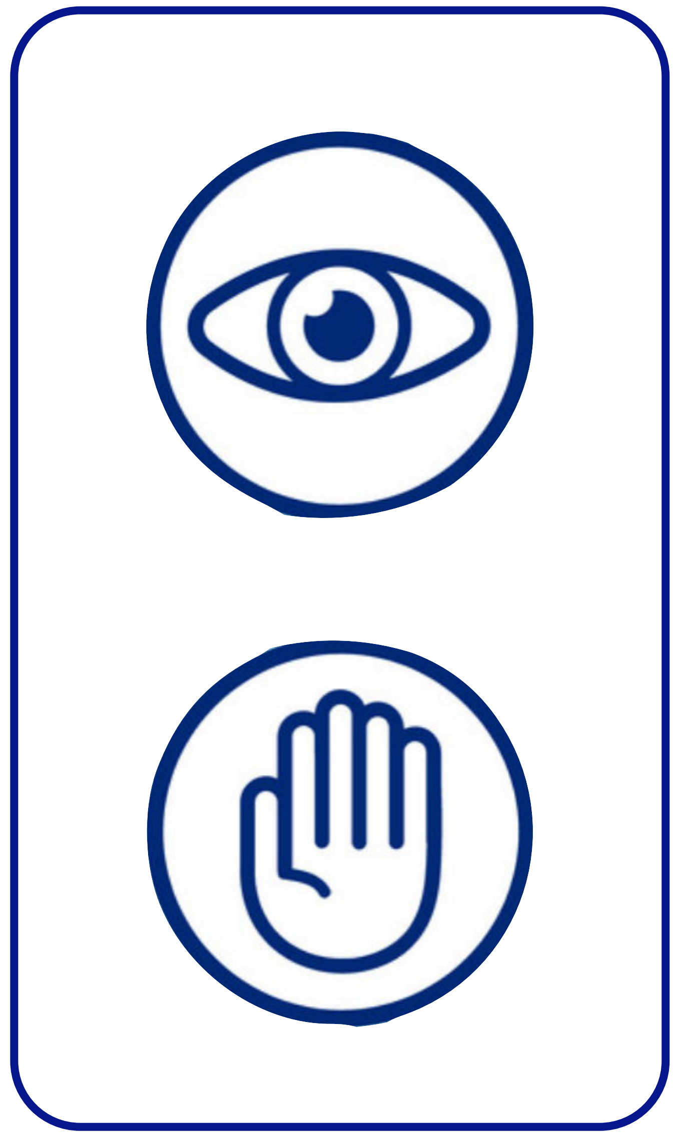
sensory data



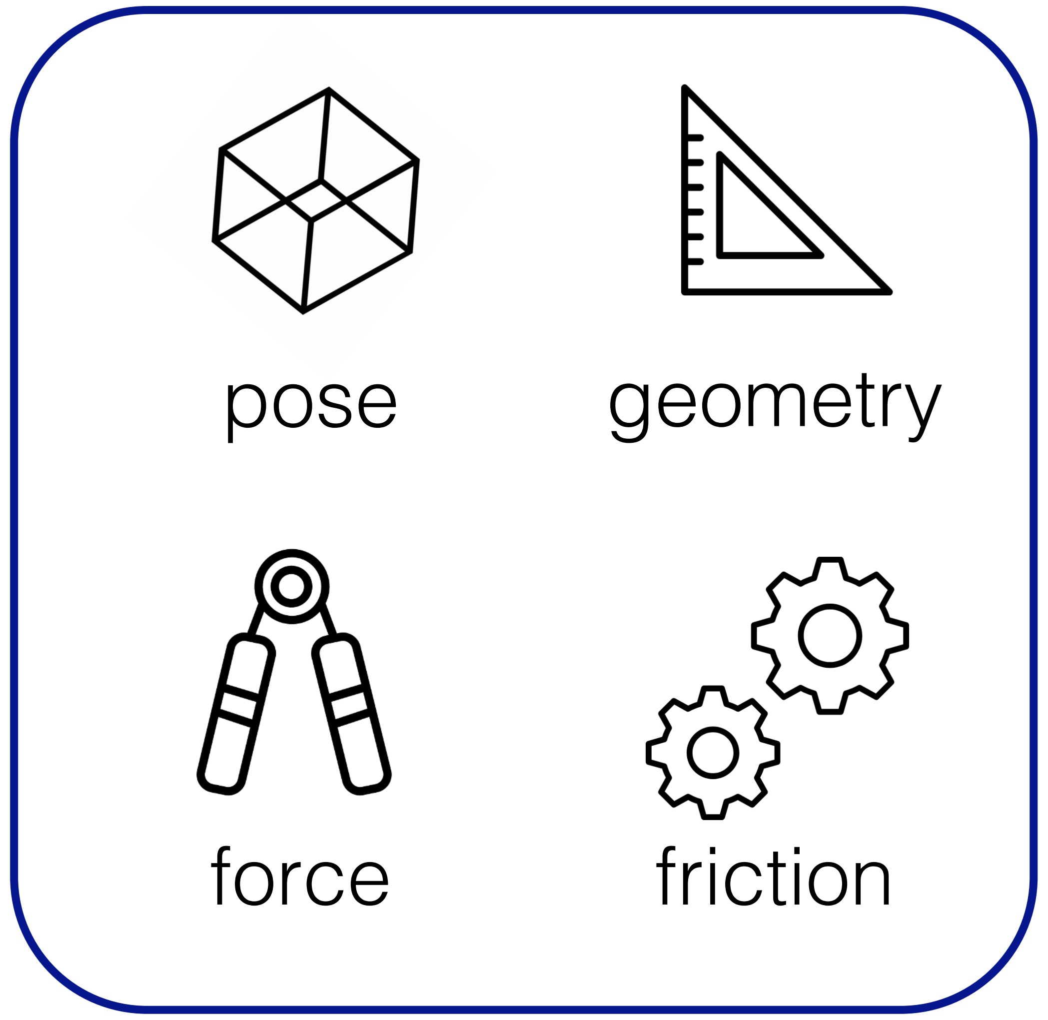
motor command

vision touch

Primitive Skills

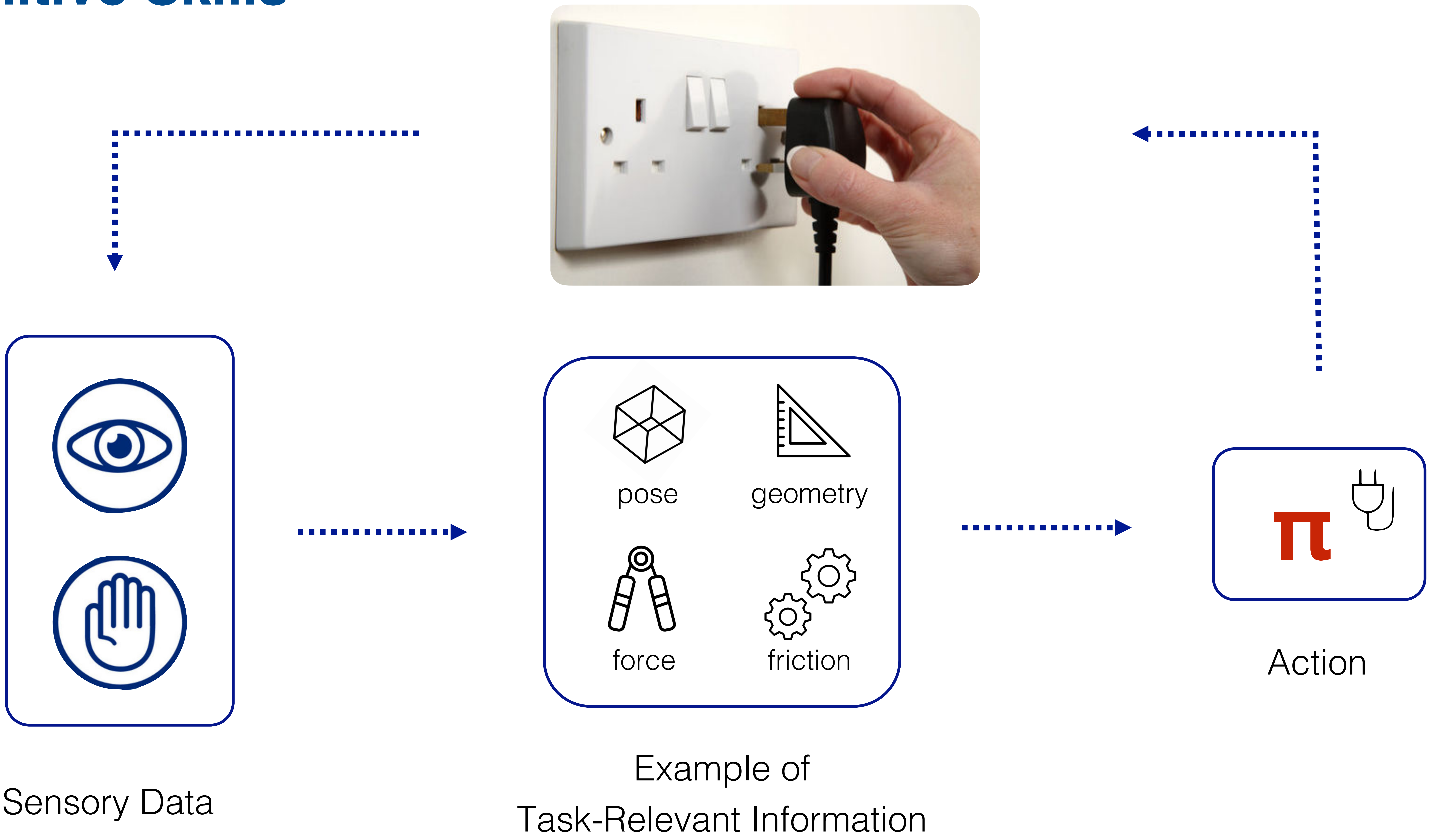


Sensory Data

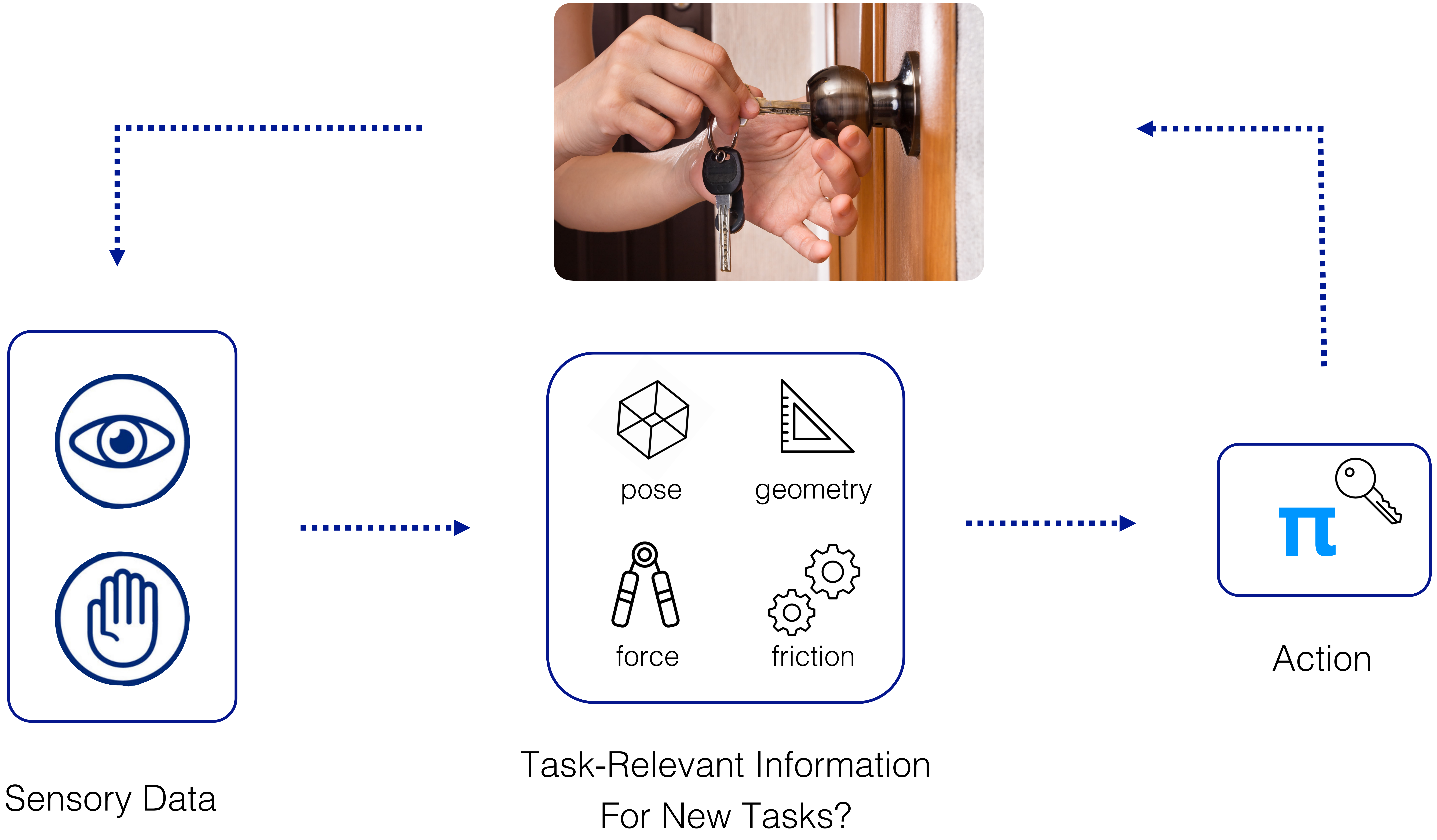


Example of
Task-Relevant Information

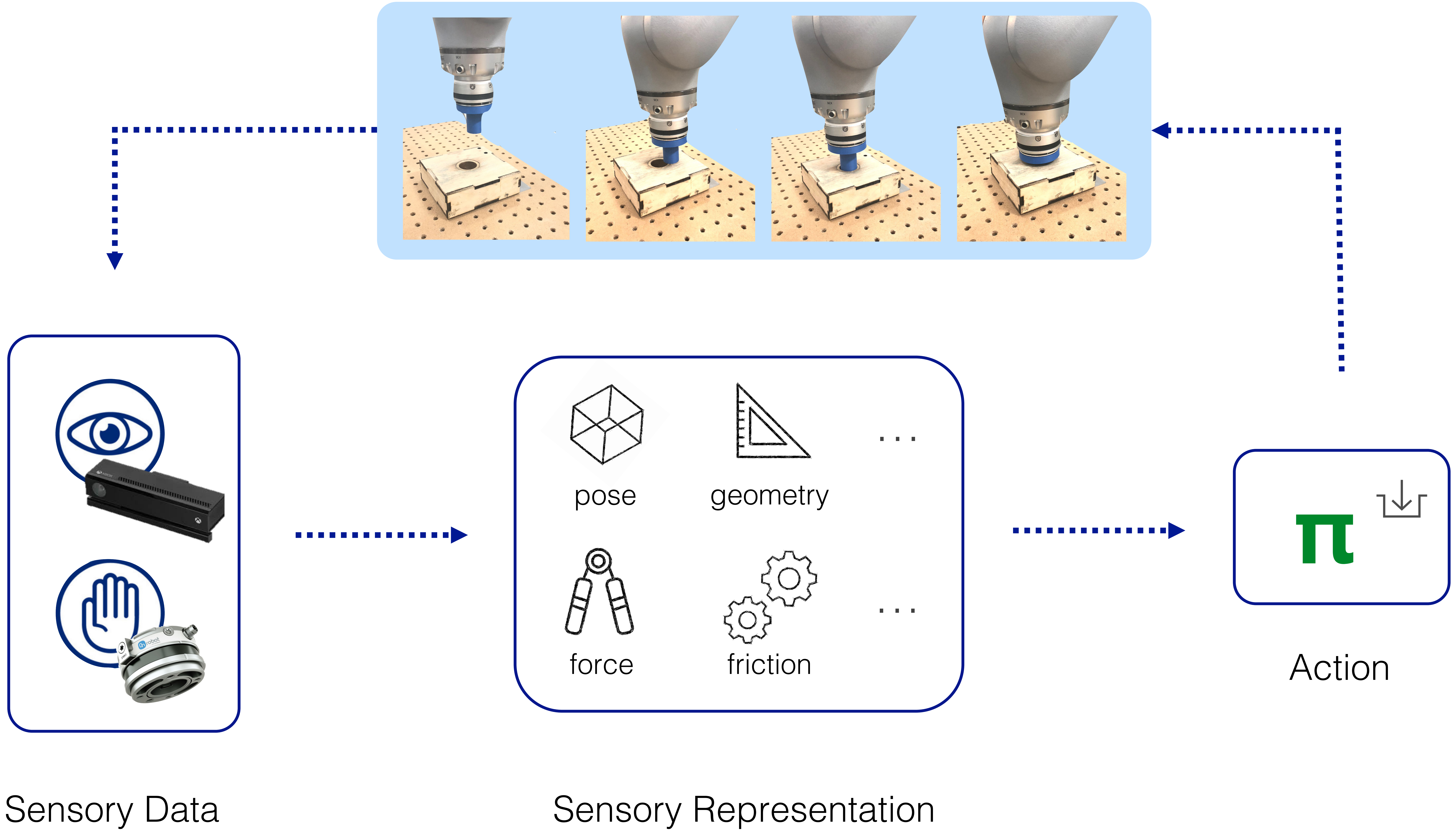
Primitive Skills



Primitive Skills



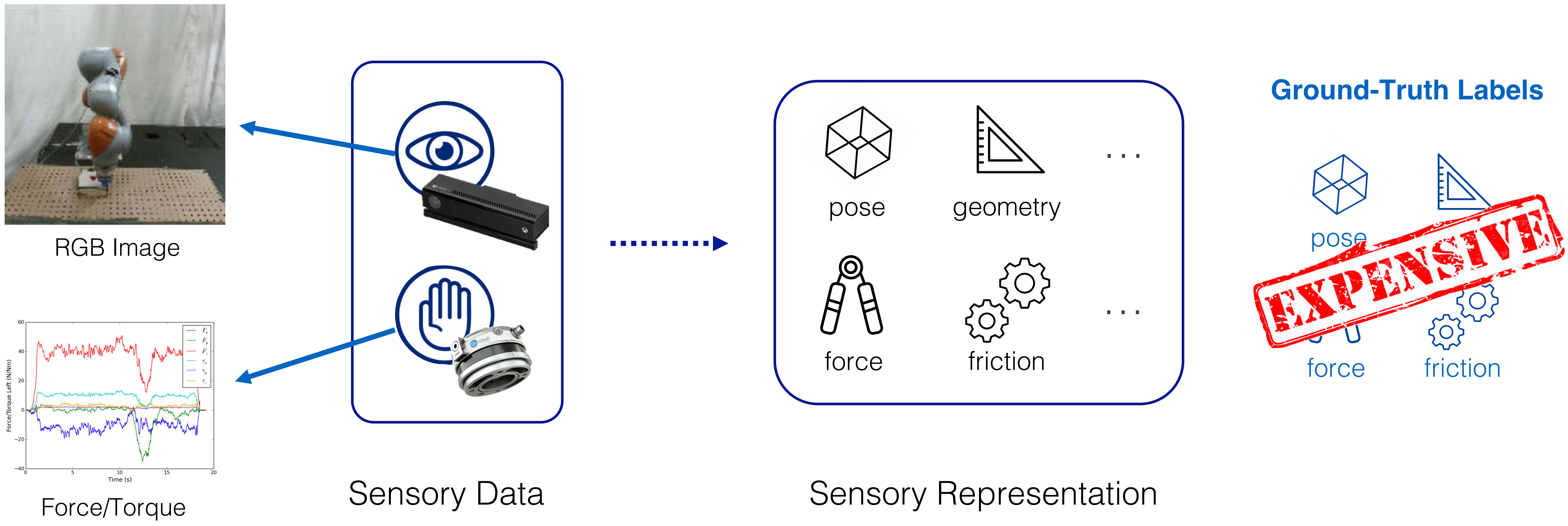
Primitive Skills



Primitive Skills

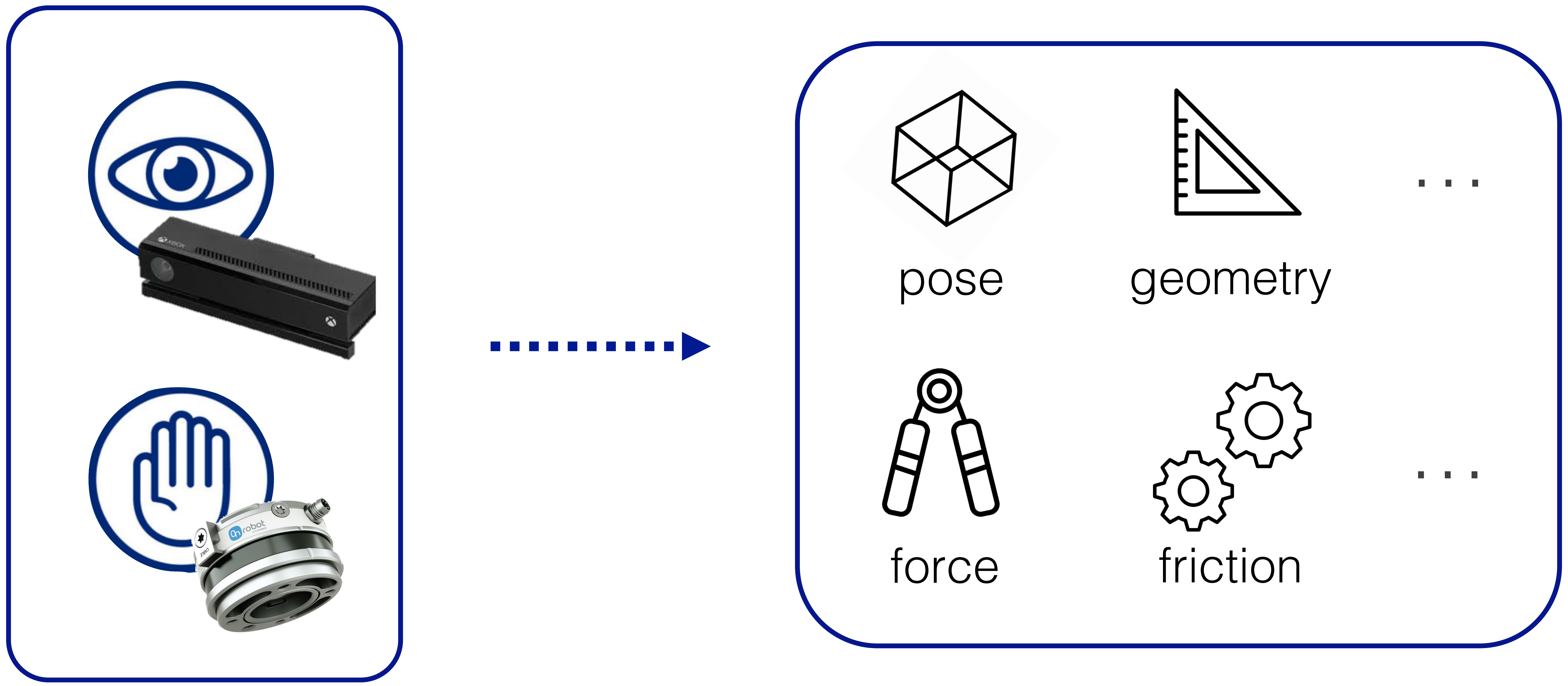
Challenge #1: Raw sensory data are **high-dimensional**, **noisy**, and **multimodal**.

Challenge #2: Manual annotation of supervision is **expensive**.



Primitive Skills

Key idea: **self-supervised representation learning** from raw sensory data

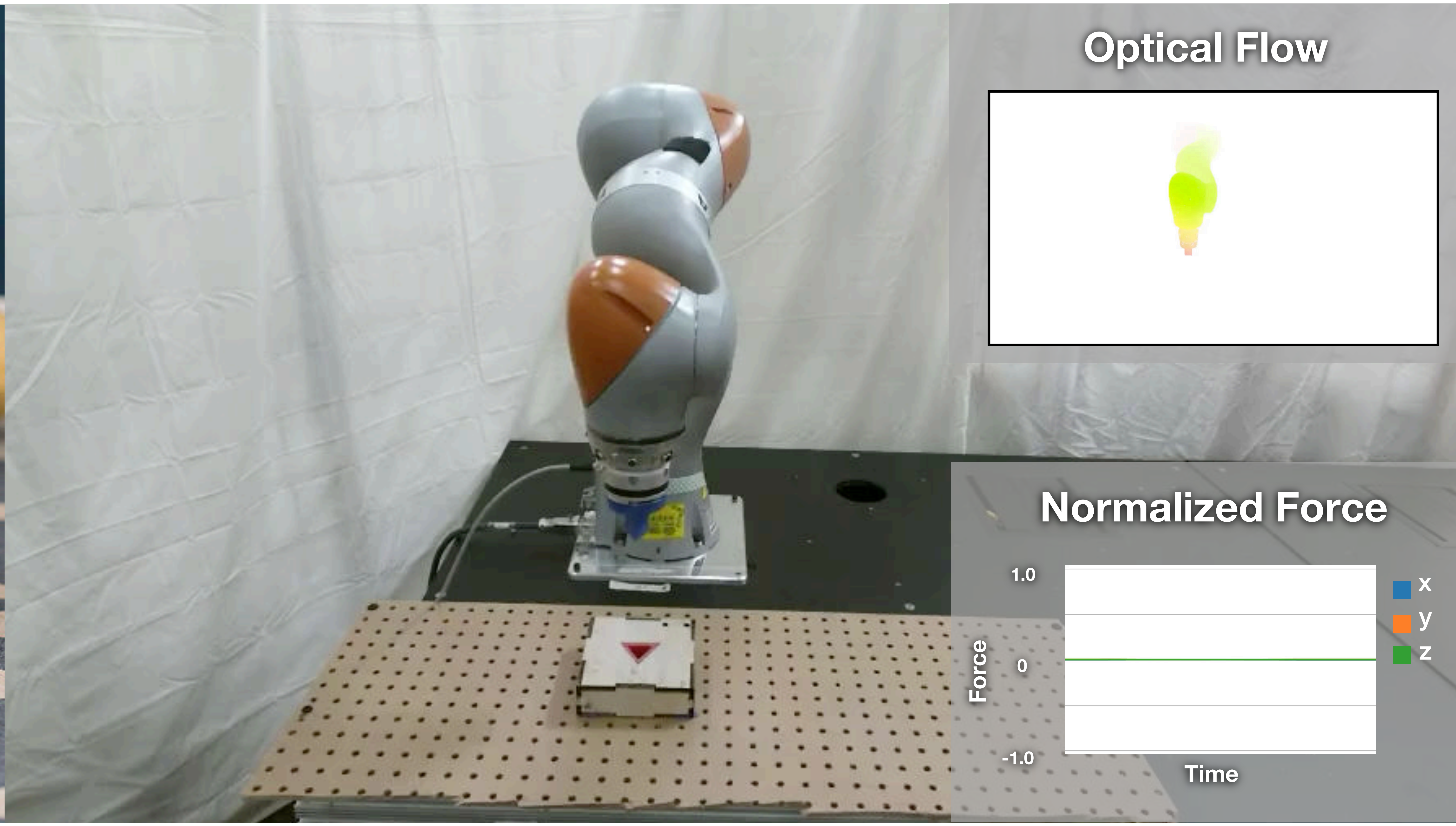


Sensory Data

Sensory Representation

Primitive Skills

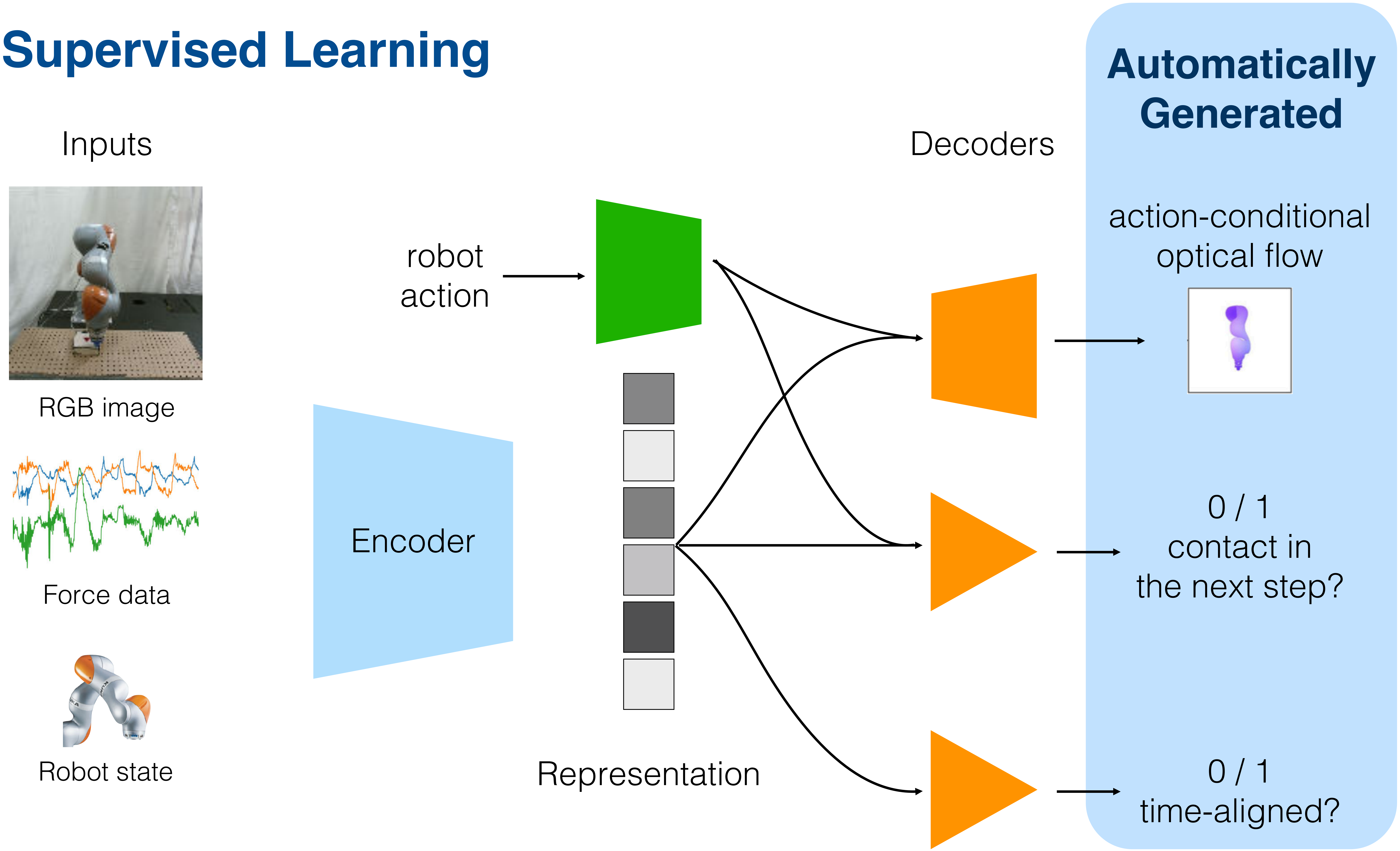
Key idea: self-supervised representation learning from raw sensory data



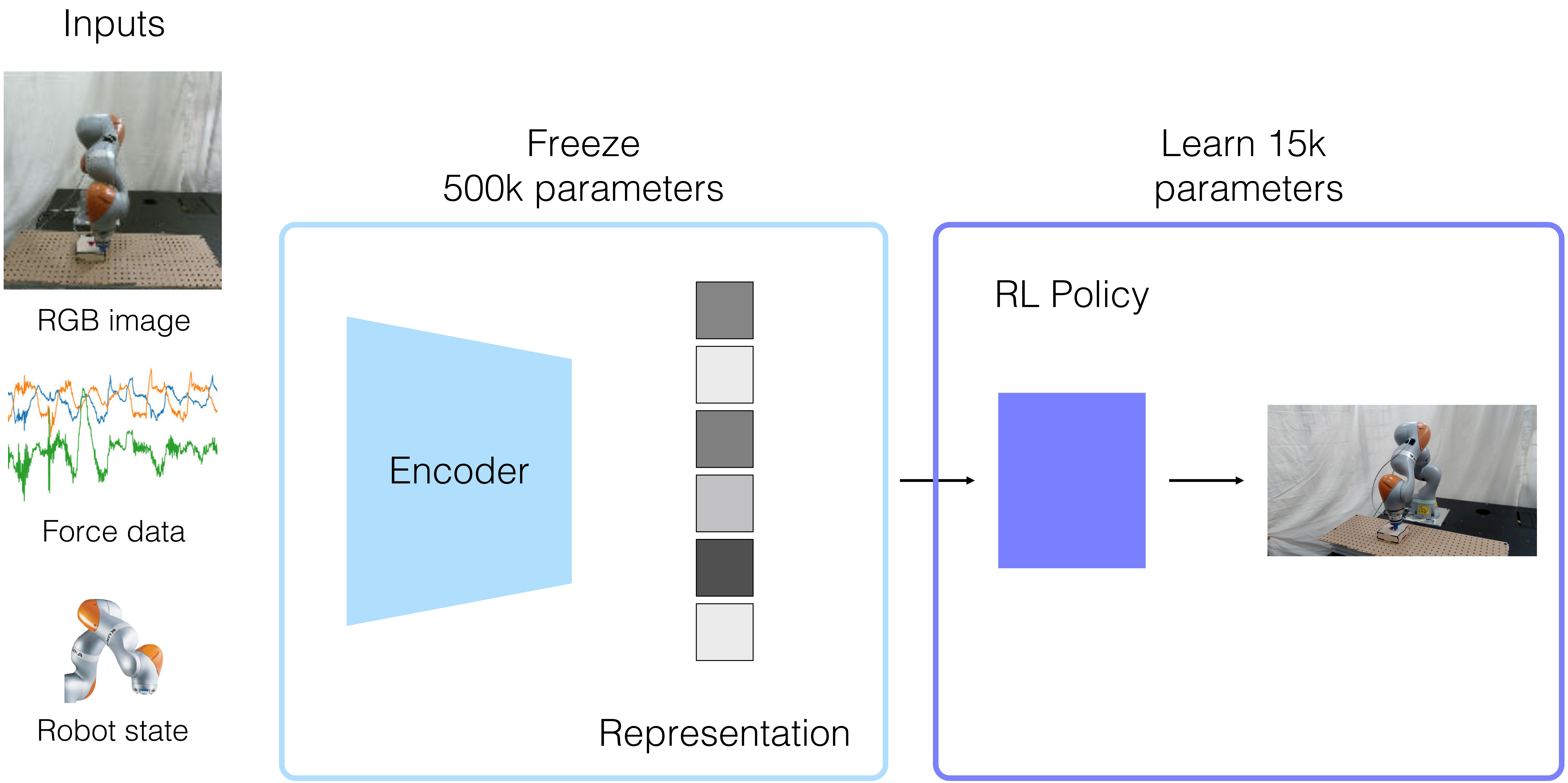
baby learning by playing

robot exploring and collecting data on its own

Self-Supervised Learning

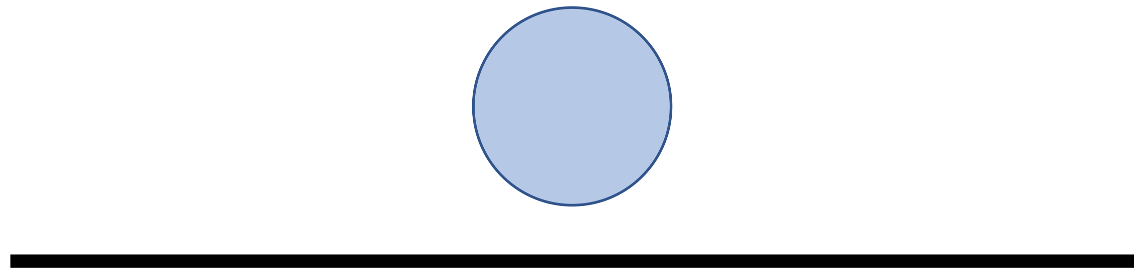


Self-Supervised Learning: Learning sample efficient policies



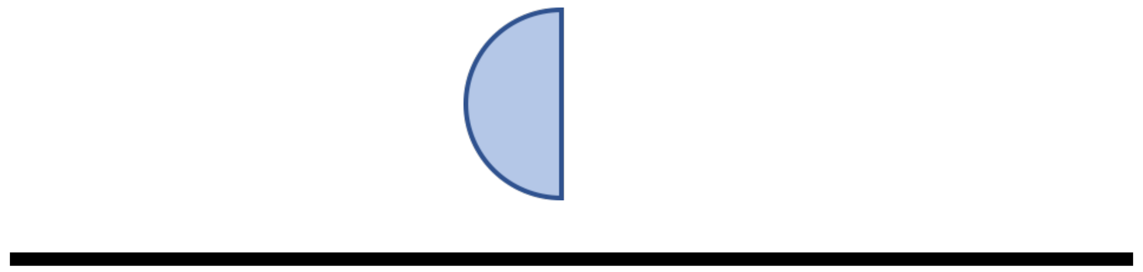
Lee*, Zhu*, et al. “Making Sense of Vision and Touch” ICRA’19

Self-Supervised Learning: We **efficiently** learn policies in 5 hours.



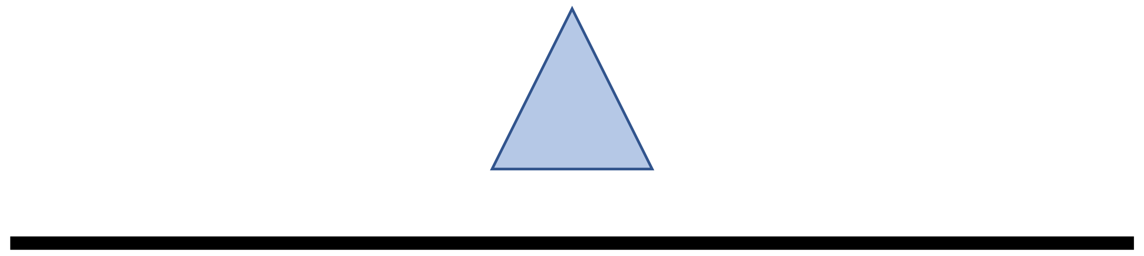
Episode 300

73% success rate



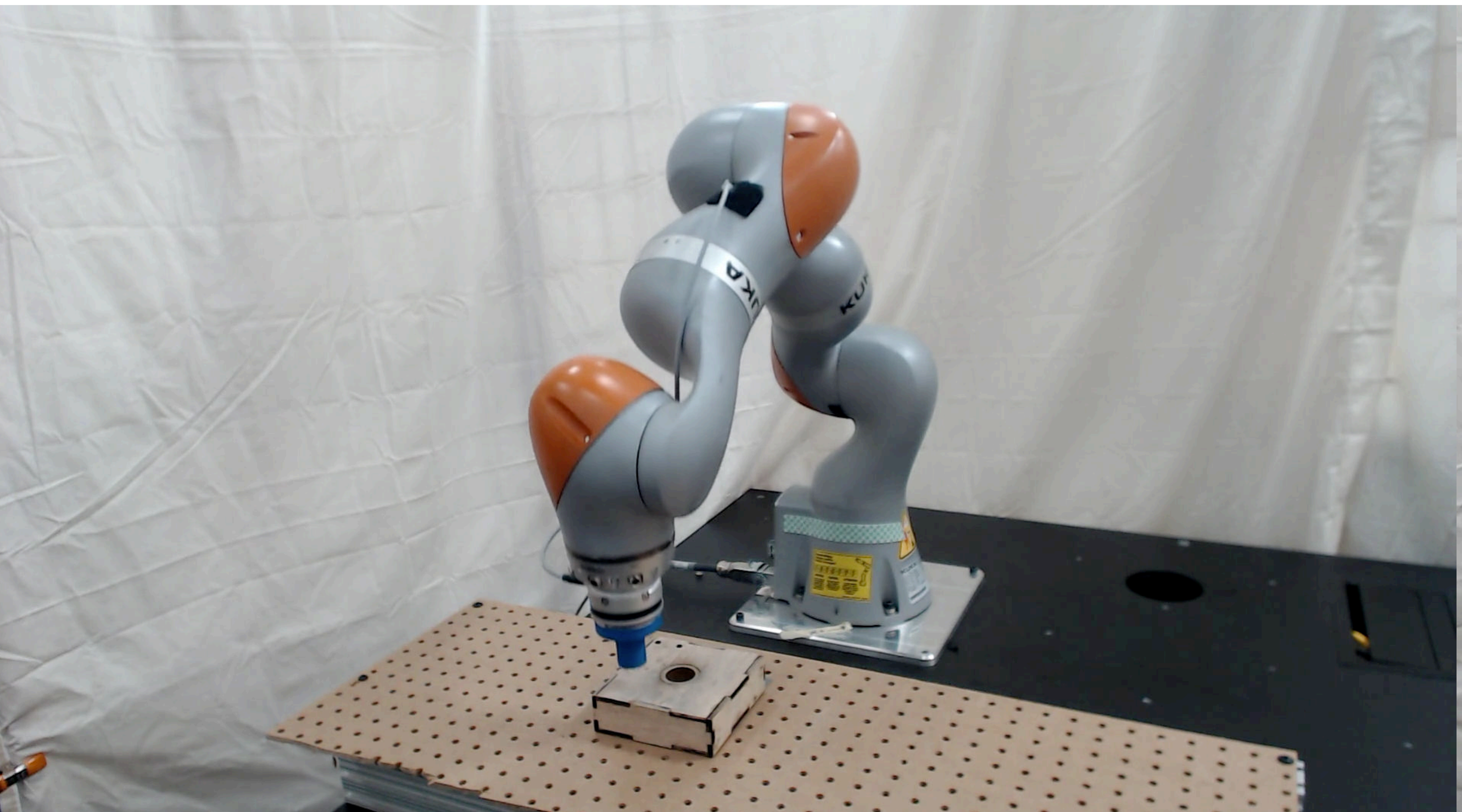
Episode 300

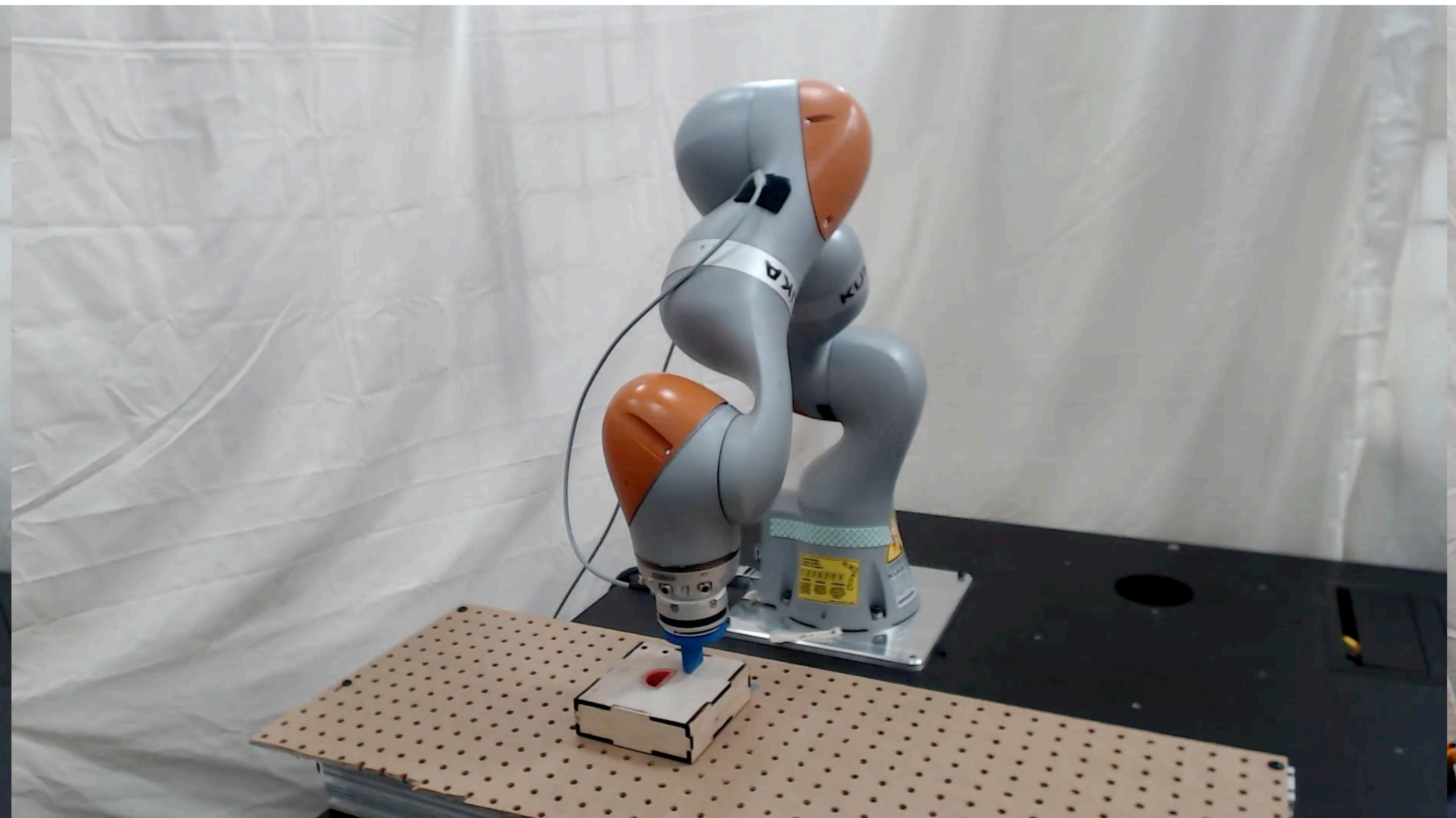
71% success rate

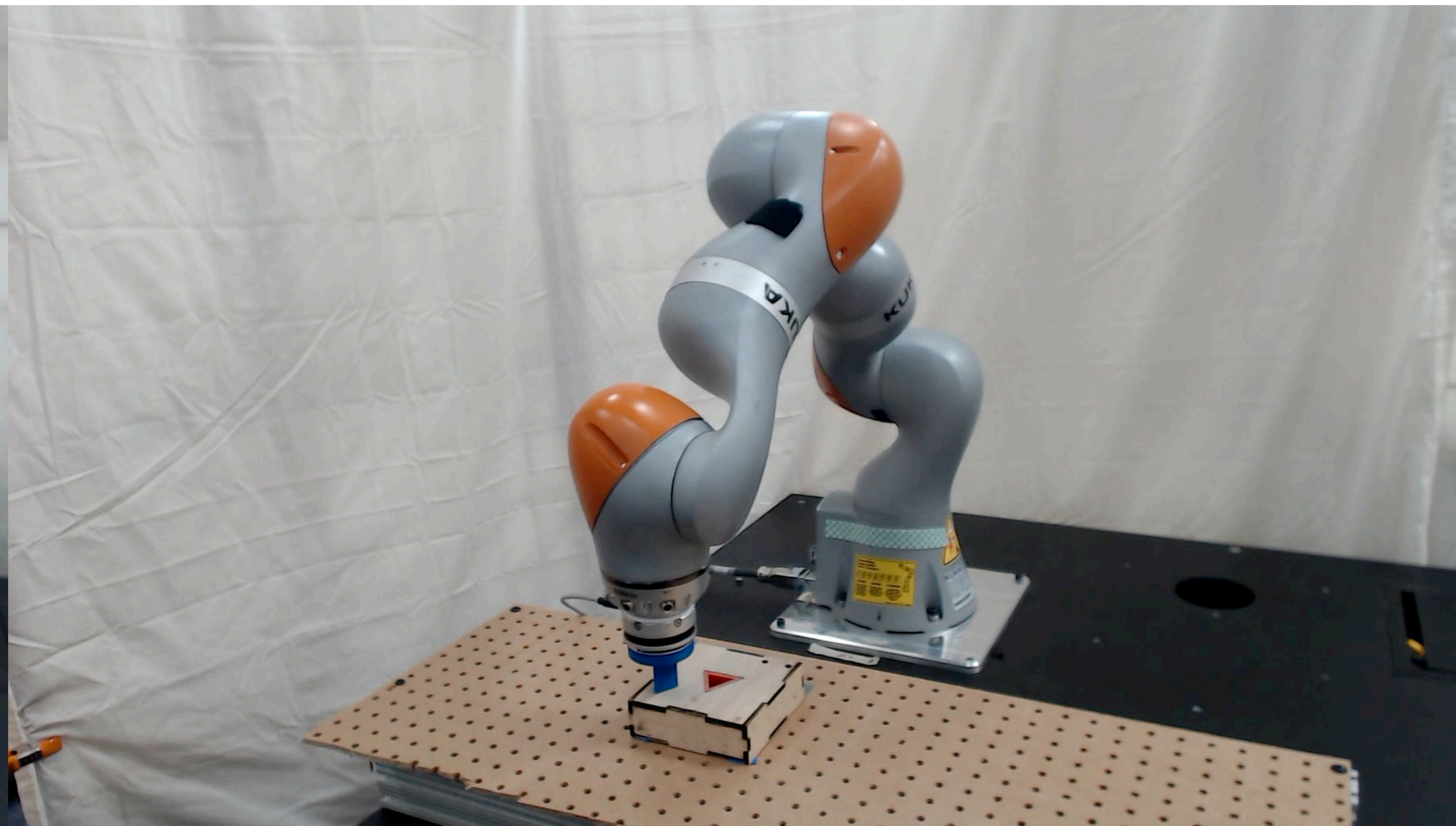


Episode 300

92% success rate



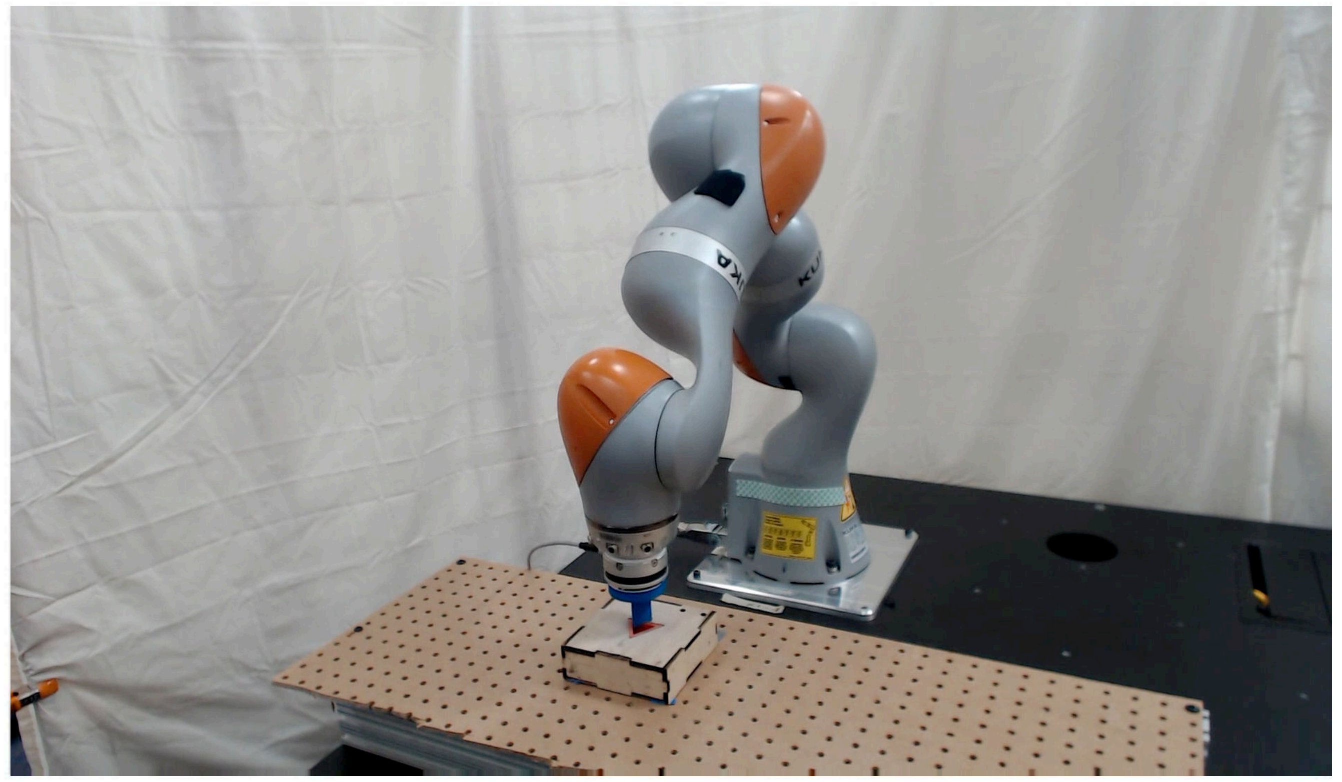





Force Perturbation
Normalized Force
1.0
0.5
0
-0.5
-1.0
Force
Time
x
y
z
x1

Self-Supervised Learning: Does Our Representation Generalize?

92% Success Rate



Tested on

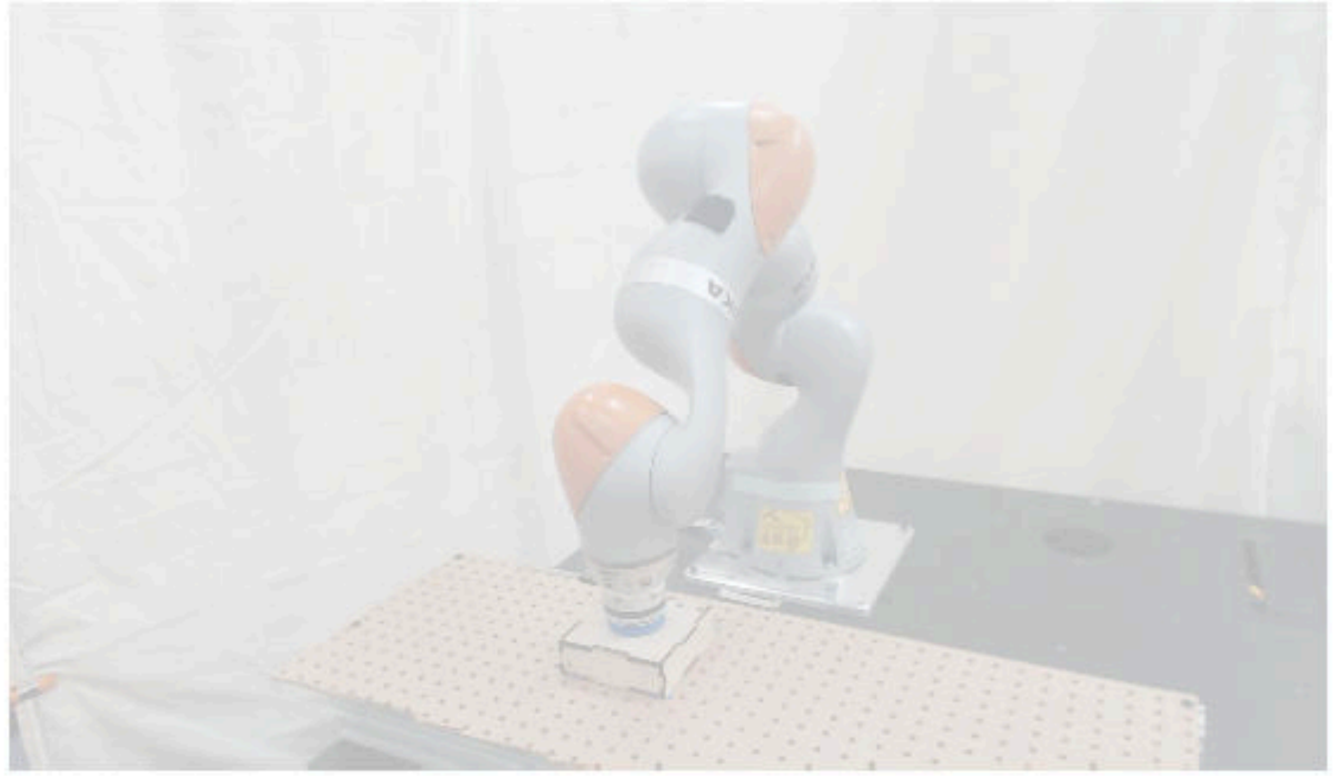
Representation

Policy

Lee*, Zhu*, et al. “Making Sense of Vision and Touch” ICRA’19

Self-Supervised Learning: Policy Transfer

92% Success Rate

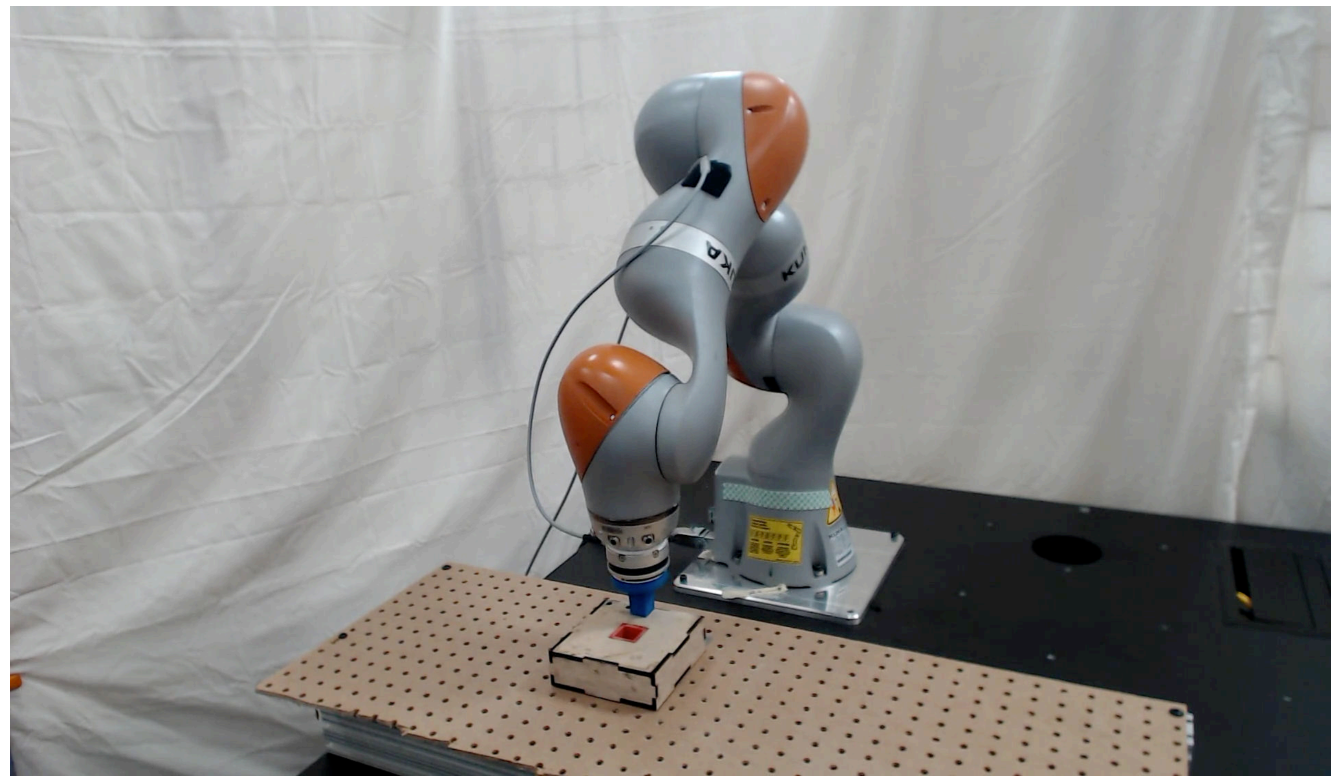


Tested on

Representation

Policy

62% Success Rate



Policy does not transfer

Tested on

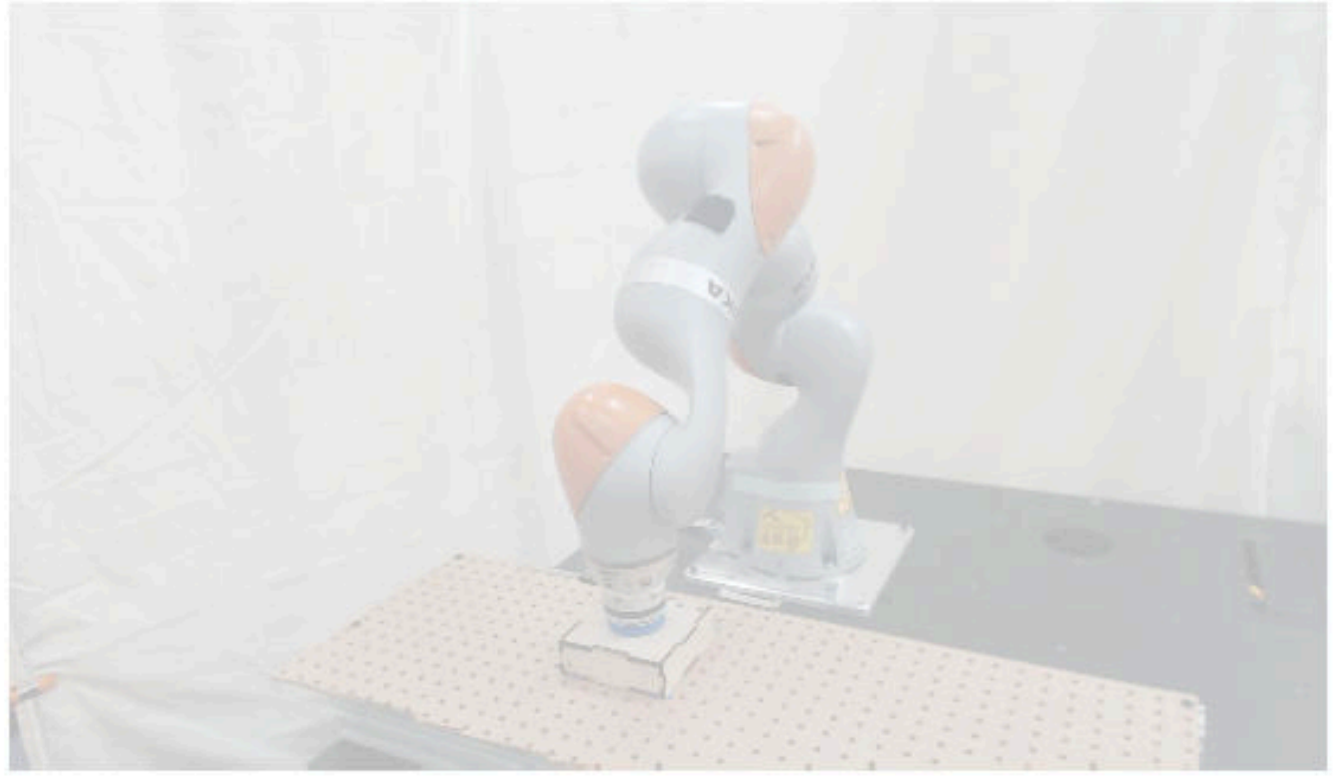
Representation

Policy

Lee*, Zhu*, et al. "Making Sense of Vision and Touch" ICRA'19

Self-Supervised Learning: Representation Transfer

92% Success Rate

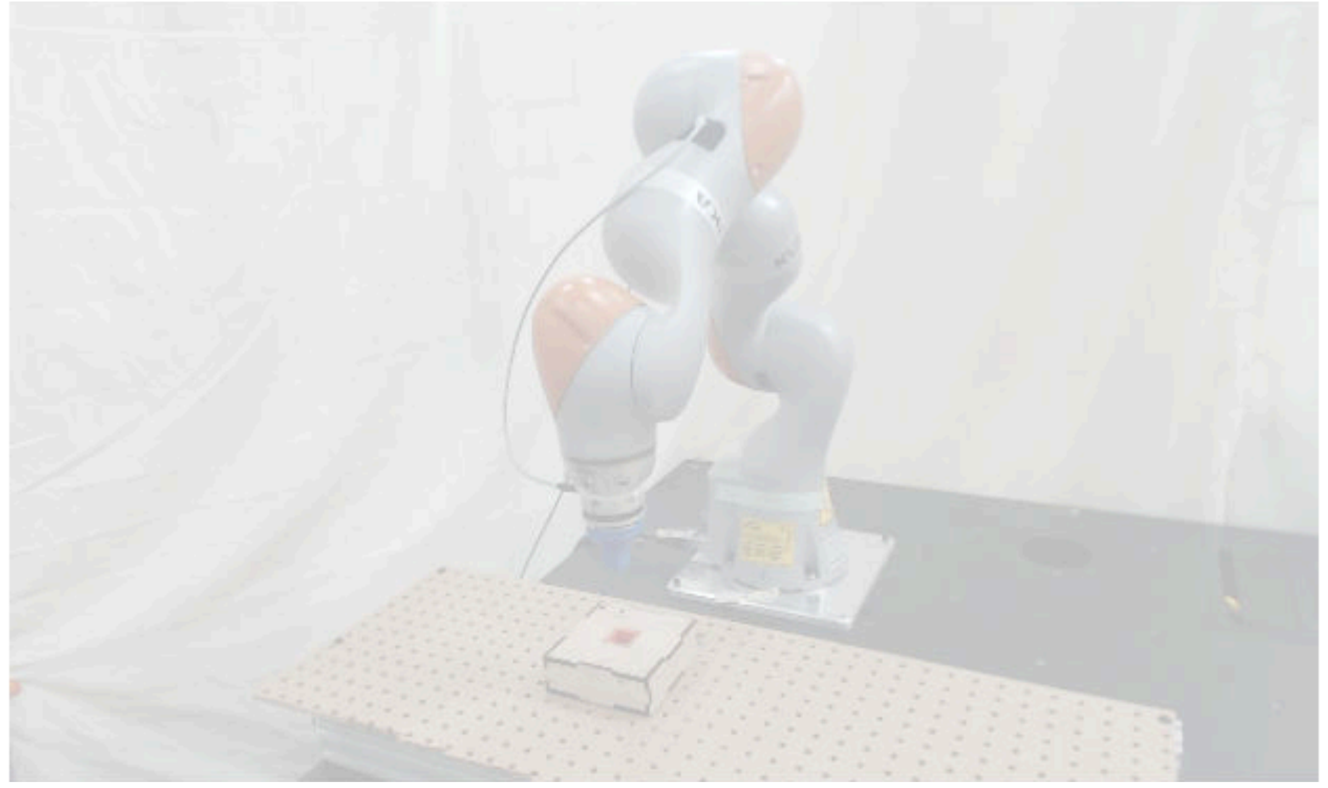


Tested on

Representation

Policy

62% Success Rate



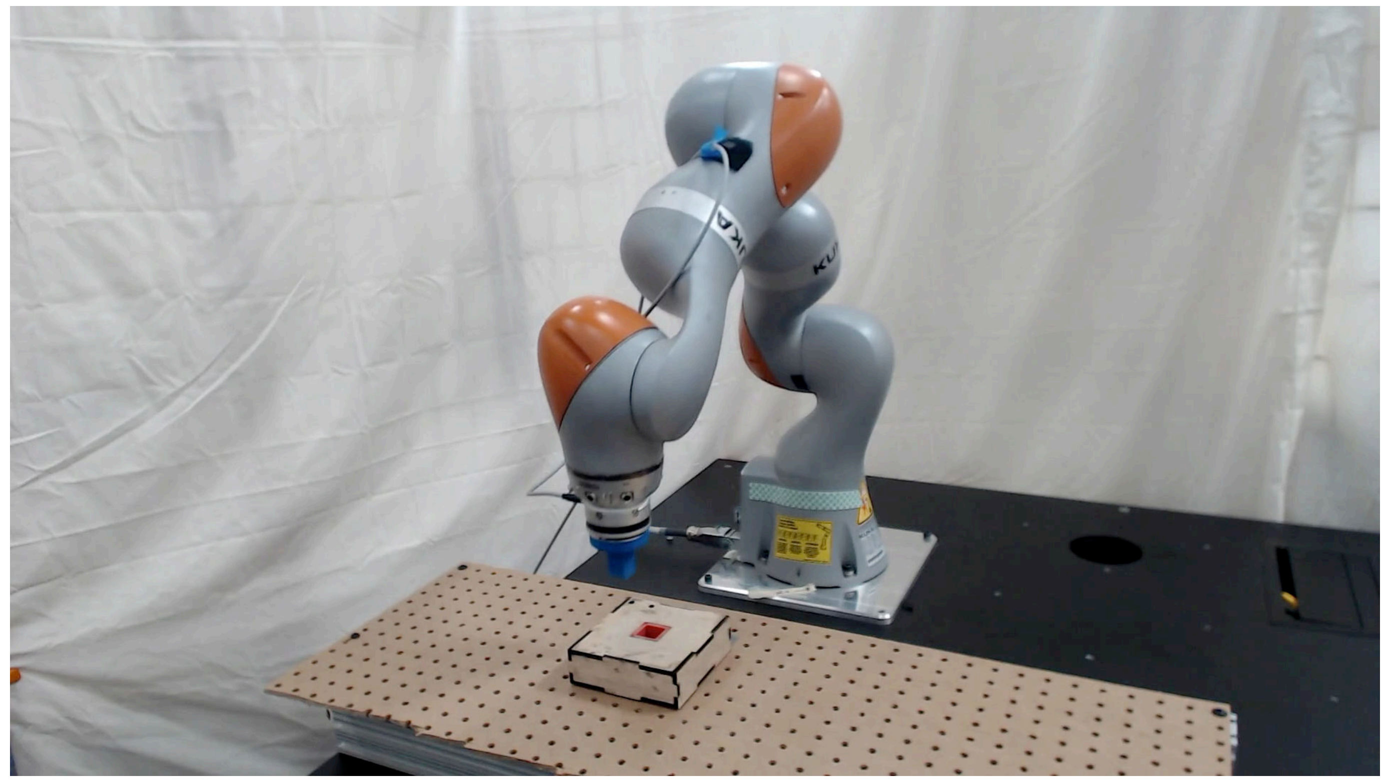
Policy does not transfer

Tested on

Representation

Policy

92% Success Rate



Representation transfers

Tested on

Representation

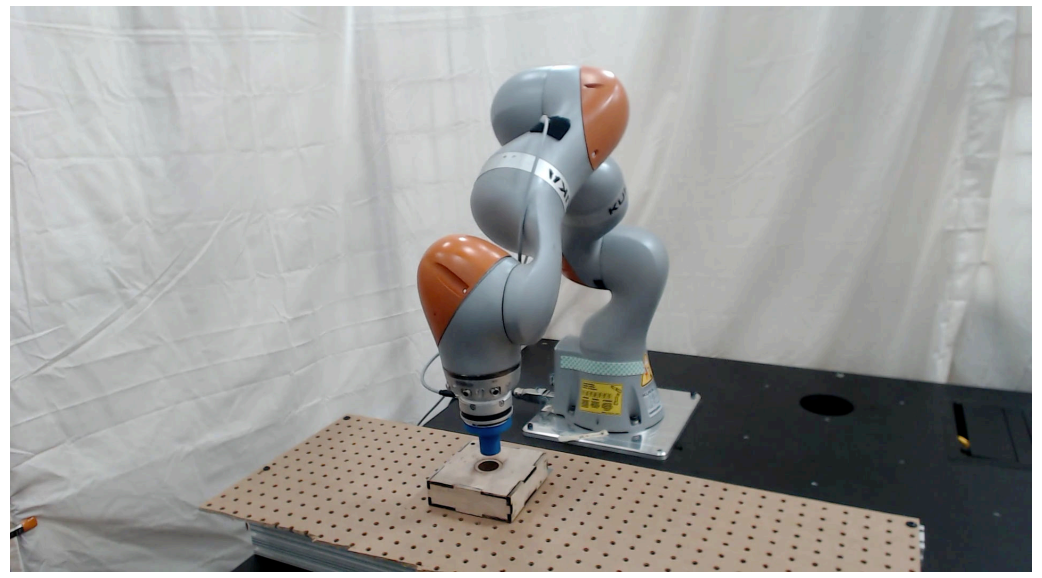
Policy

Lee*, Zhu*, et al. "Making Sense of Vision and Touch" ICRA'19

Primitive Skills: Overview of Our Method

Self-Supervised Data Collection

$$o_{RGB}, o_{force}, o_{robot}$$

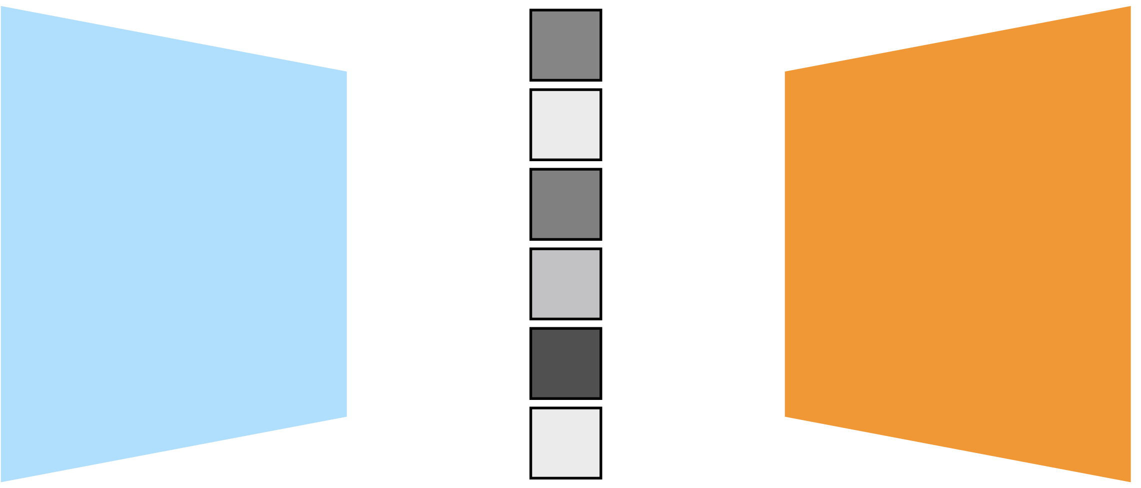


100k data points

90 minutes

Representation Learning

$$f(o_{RGB}, o_{force}, o_{robot})$$

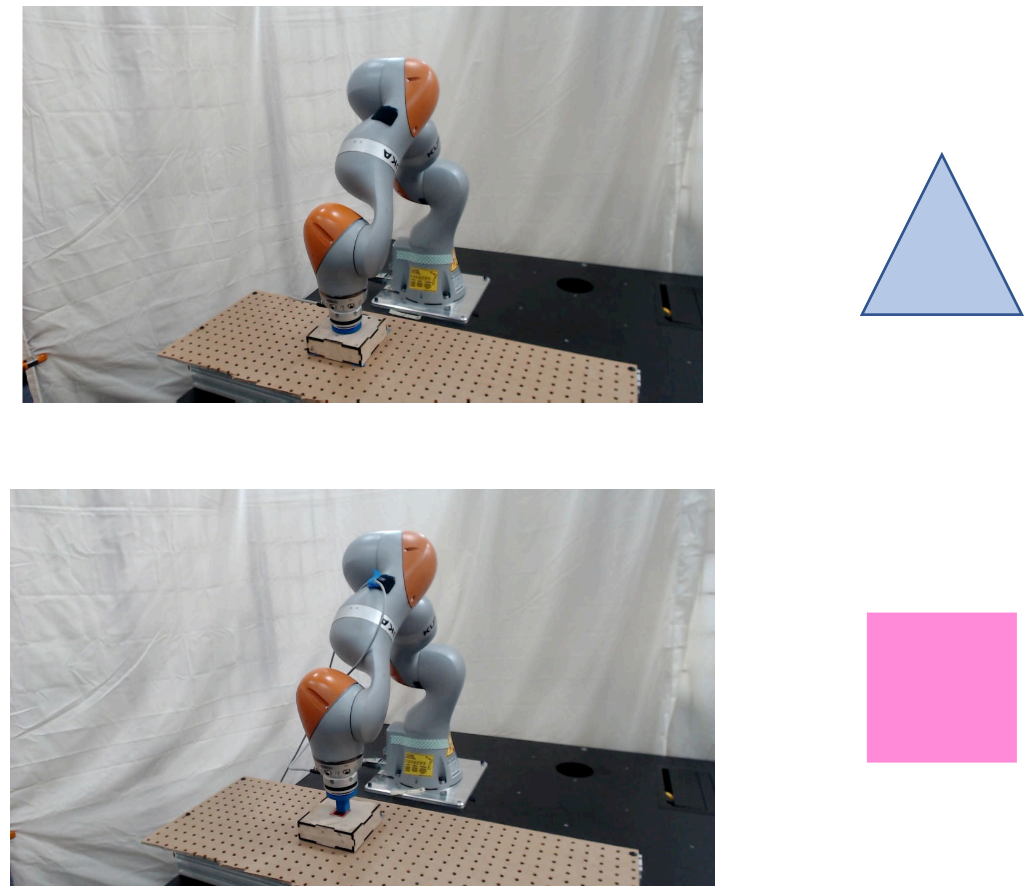


20 epochs on GPU

24 hours

Policy Learning

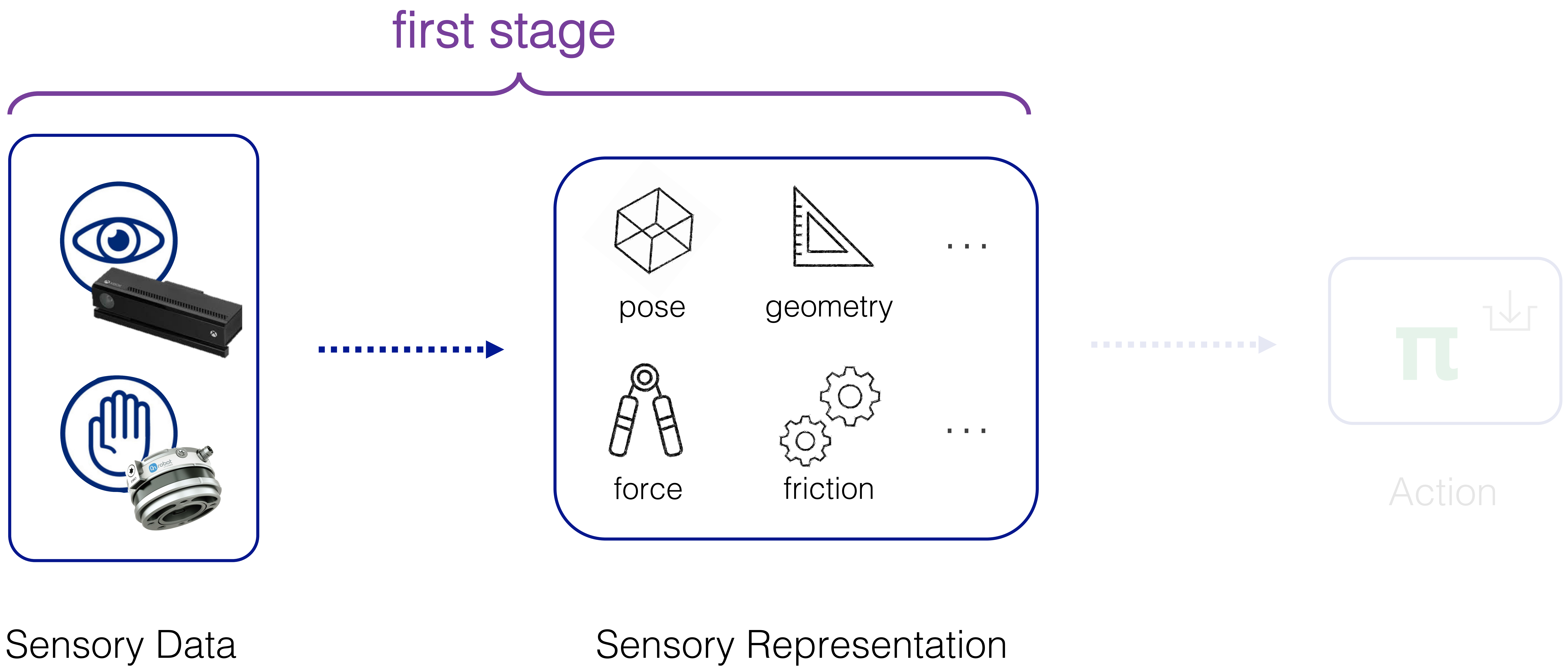
$$\pi(f(\cdot)) = a$$



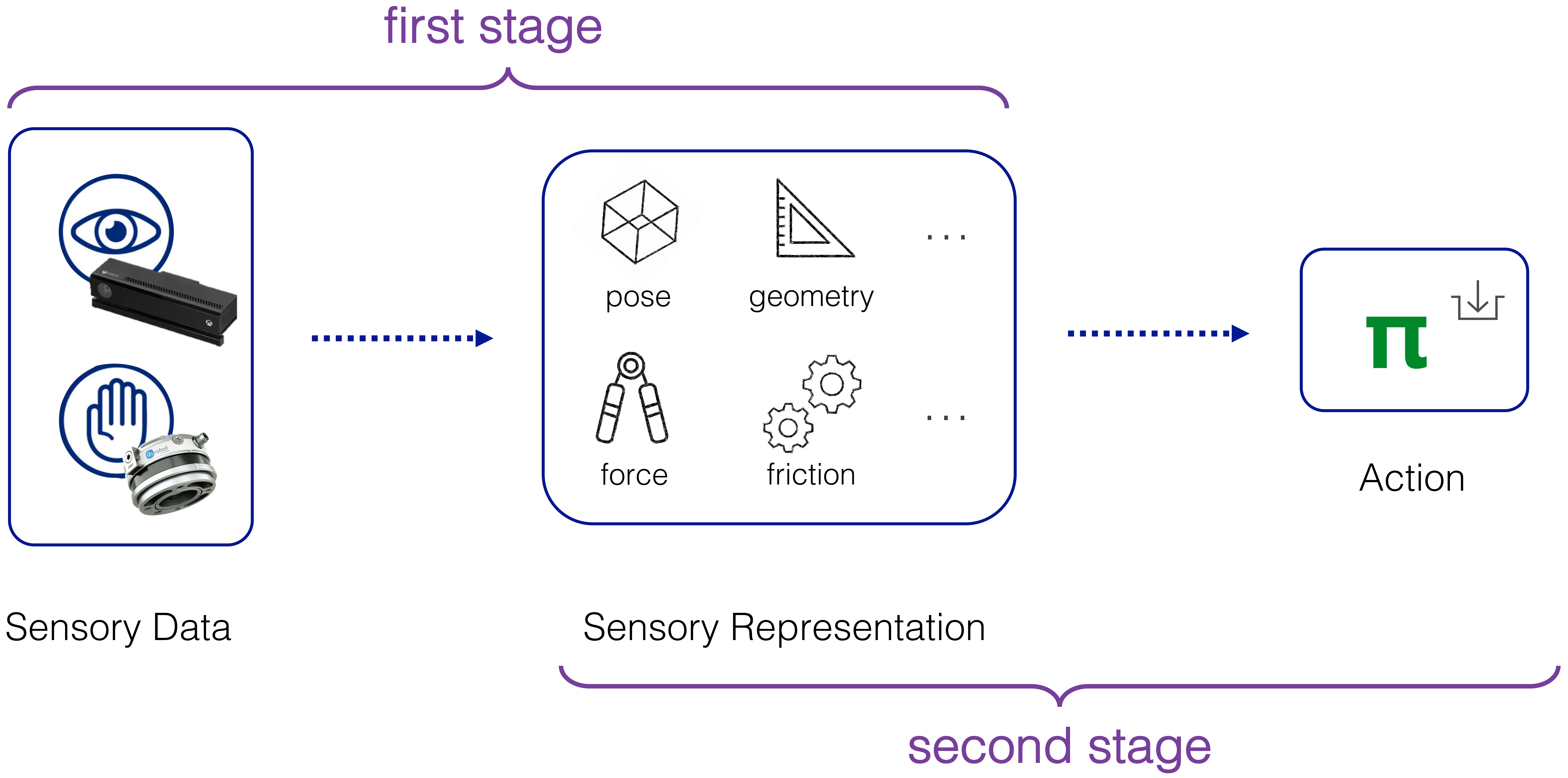
Deep RL

5 hours

Primitive Skills: Self-Supervised Learning

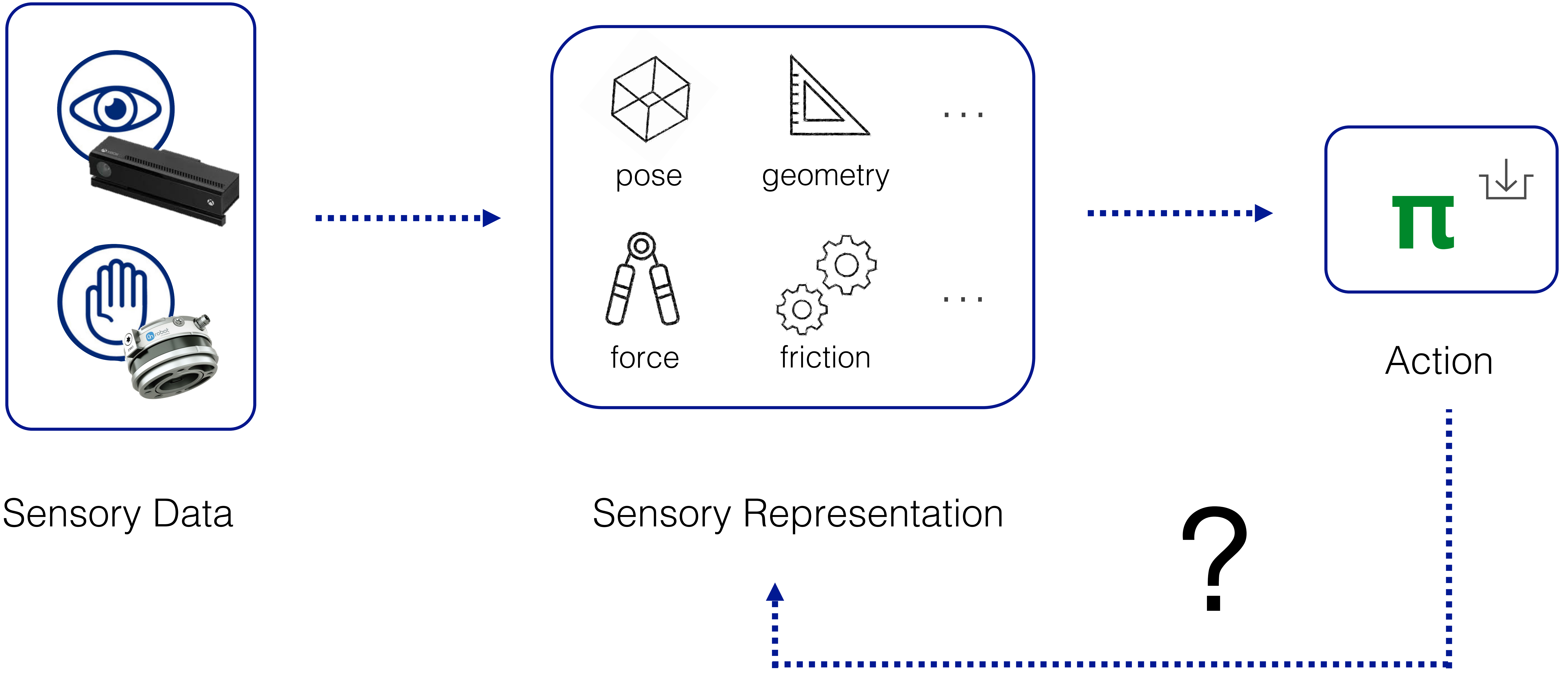


Primitive Skills: Self-Supervised Learning



Primitive Skills: Self-Supervised Learning

Can the downstream task inform the learning of representations?



Vision-Based Tool Manipulation

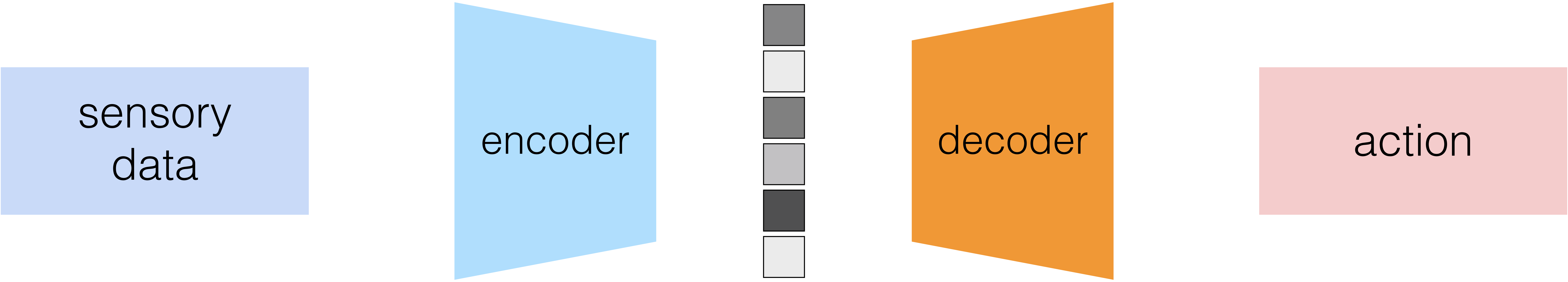
Fang, Zhu, et al., RSS'19

Primitive Skills: Vision-Based Tool Manipulation

sensory data → representation → action

Fang, Zhu, et al., RSS'19; Qin et al. ICRA'20

Primitive Skills: Vision-Based Tool Manipulation

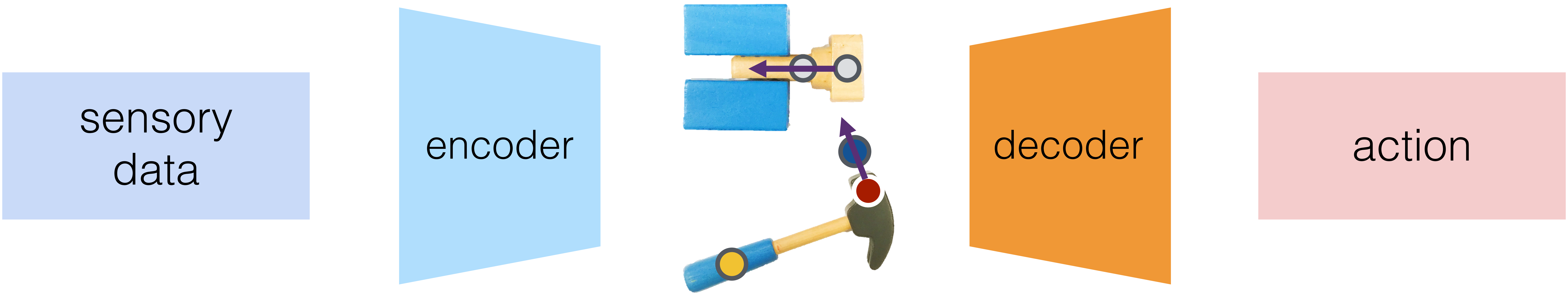


latent representation

Fang, Zhu, et al. “Task-Oriented Grasping” RSS’19

- high-dimensionality
- lack of interpretability

Primitive Skills: Vision-Based Tool Manipulation

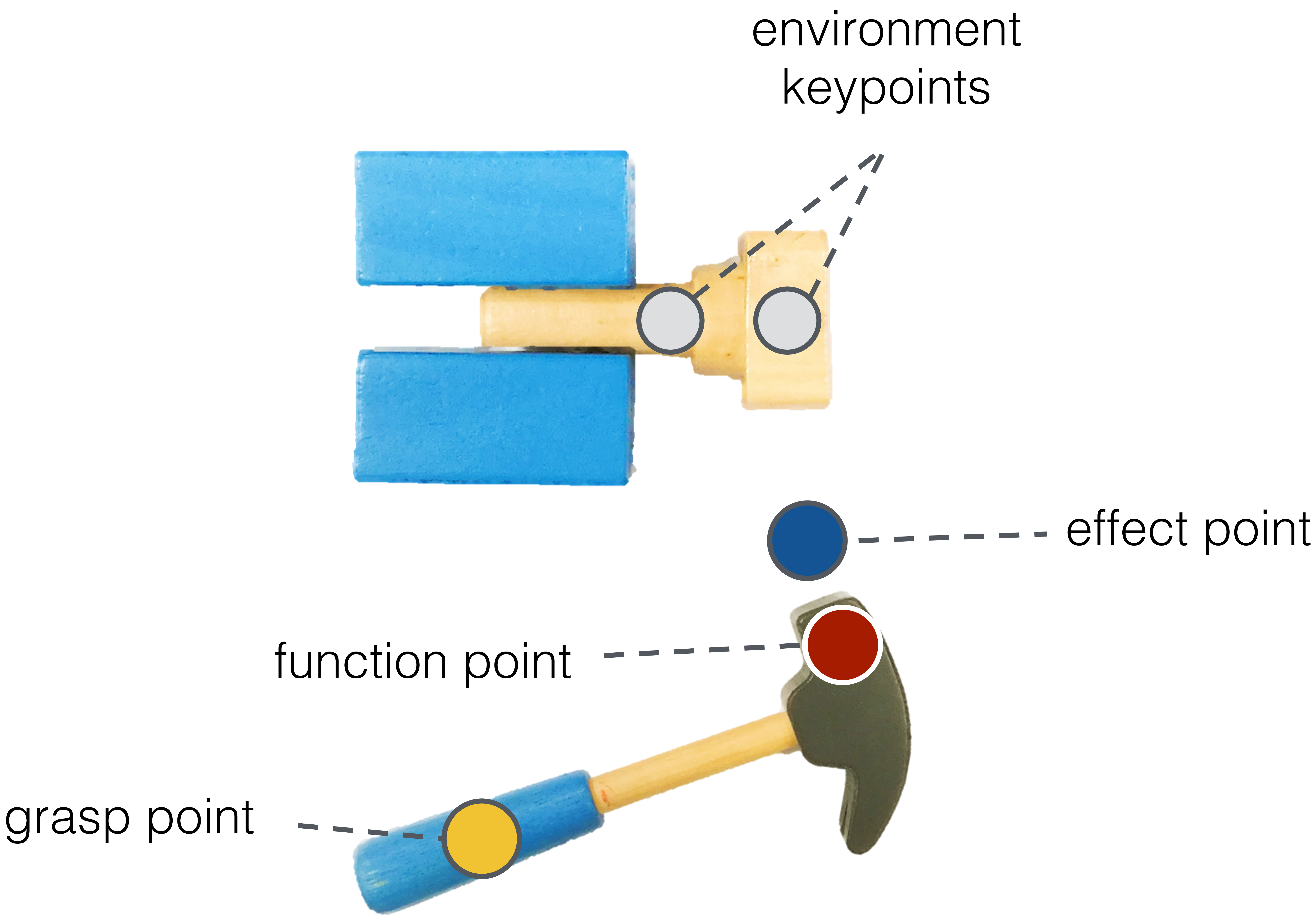


keypoint representation

Qin et al., “KETO” ICRA’20

- compact and informative
- human interpretable

KETO: **Ke**ypoint Representations for **To**ol Manipulation



Qin et al. "KETO" ICRA'20

KETO: **Ke**ypoint Representations for **To**ol Manipulation

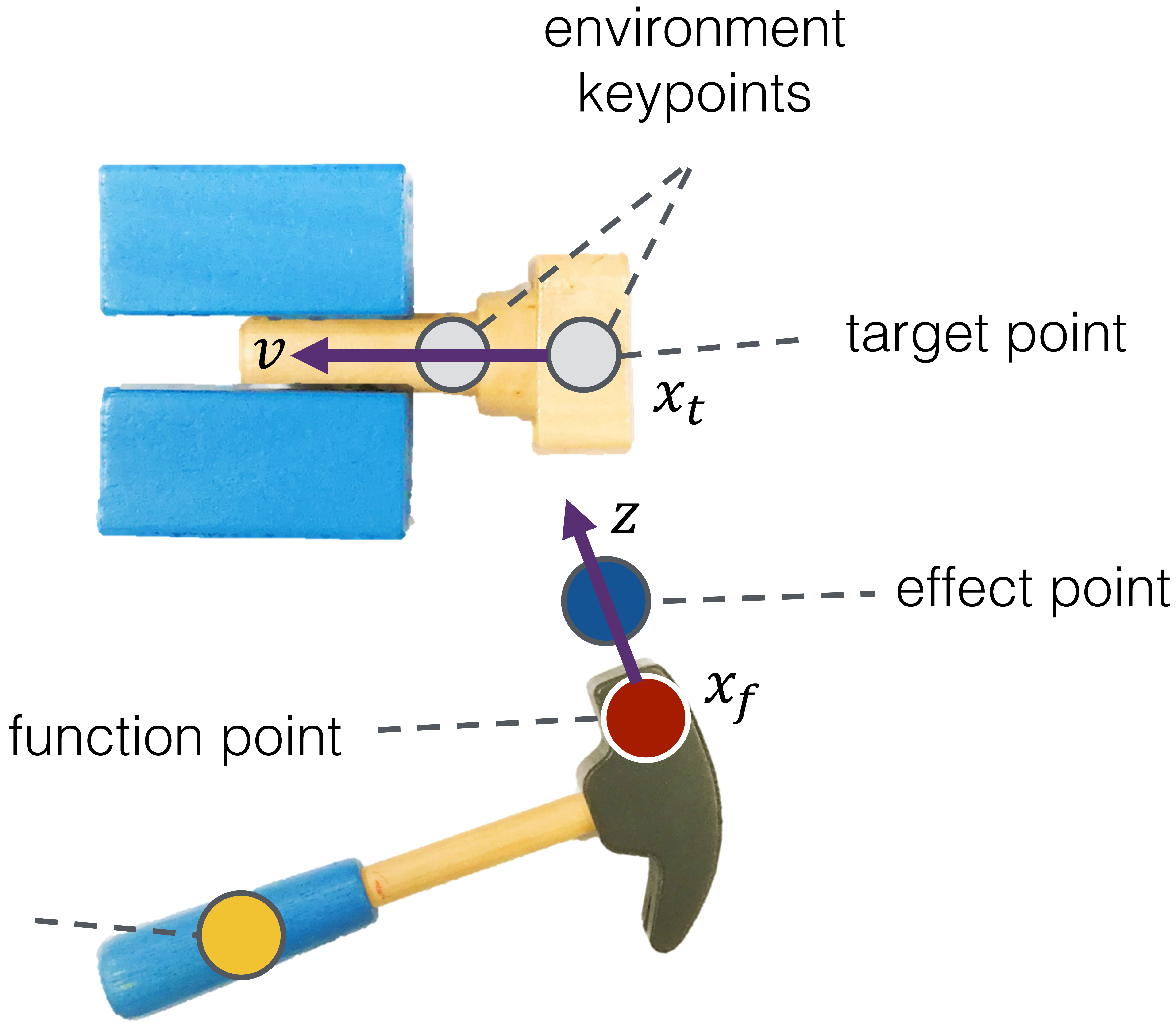
For **hammering**

1. x_t is close to x_f

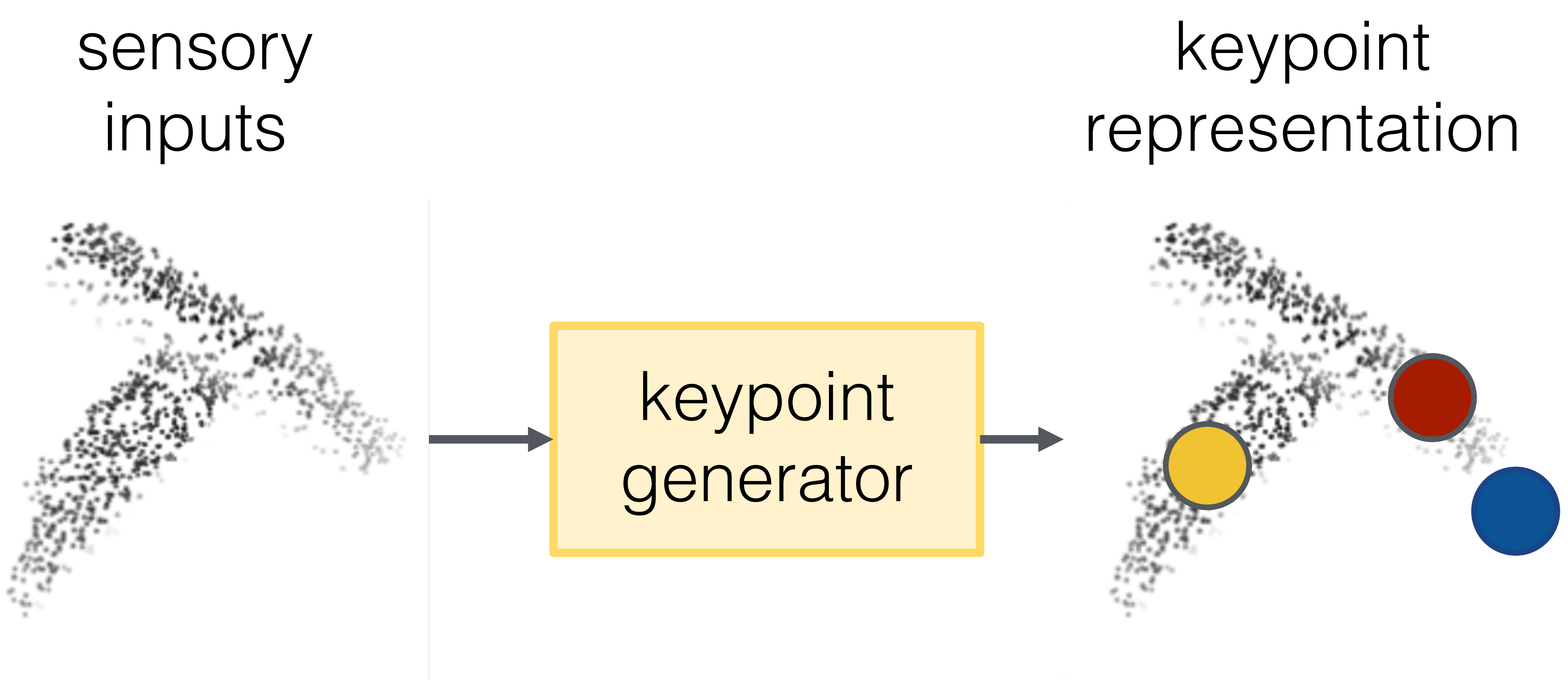
2. Direction of v aligns with z.

$$\max_{p} \; v^T z - \left\|x_f - x_t\right\|^2$$

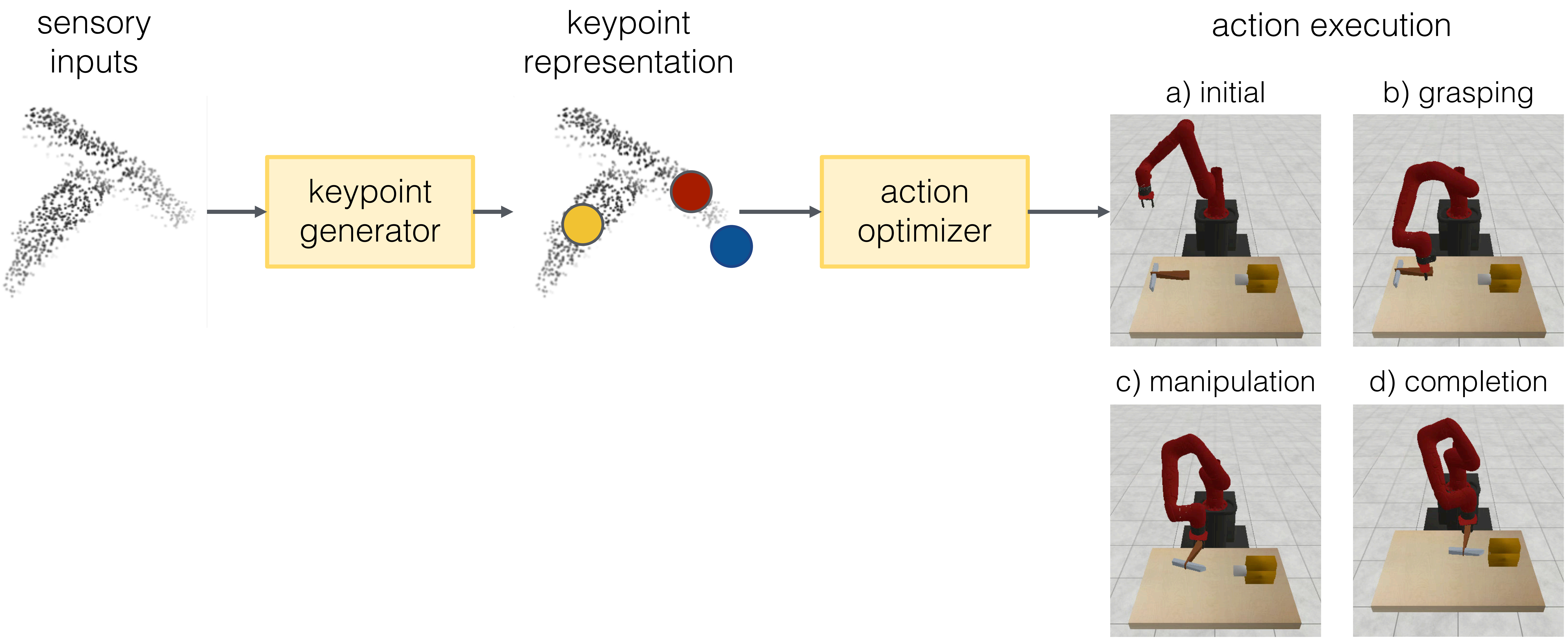
Solving the optimal pose
of object as a QP



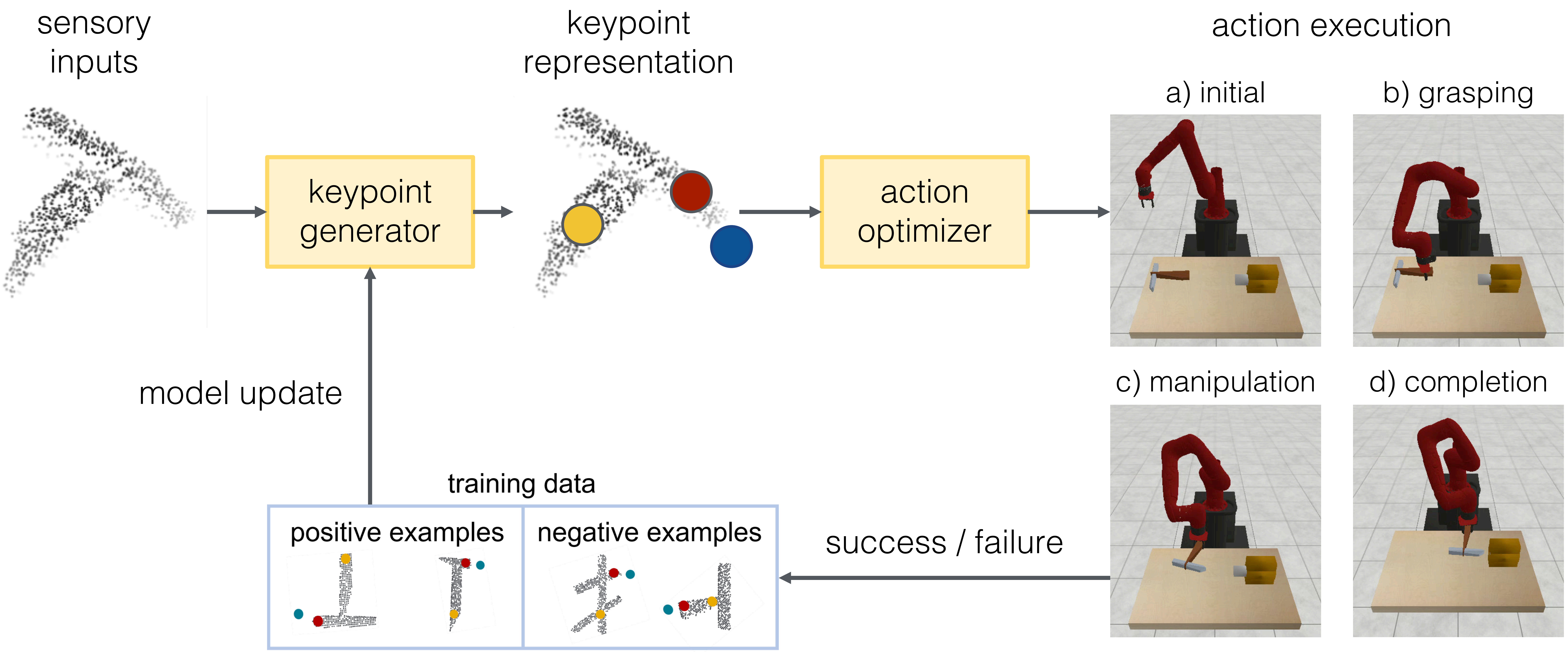
KETO: **Ke**ypoint Representations for **To**ol Manipulation



KETO: **Ke**ypoint Representations for **To**ol Manipulation



KETO: **Ke**ypoint Representations for **To**ol Manipulation

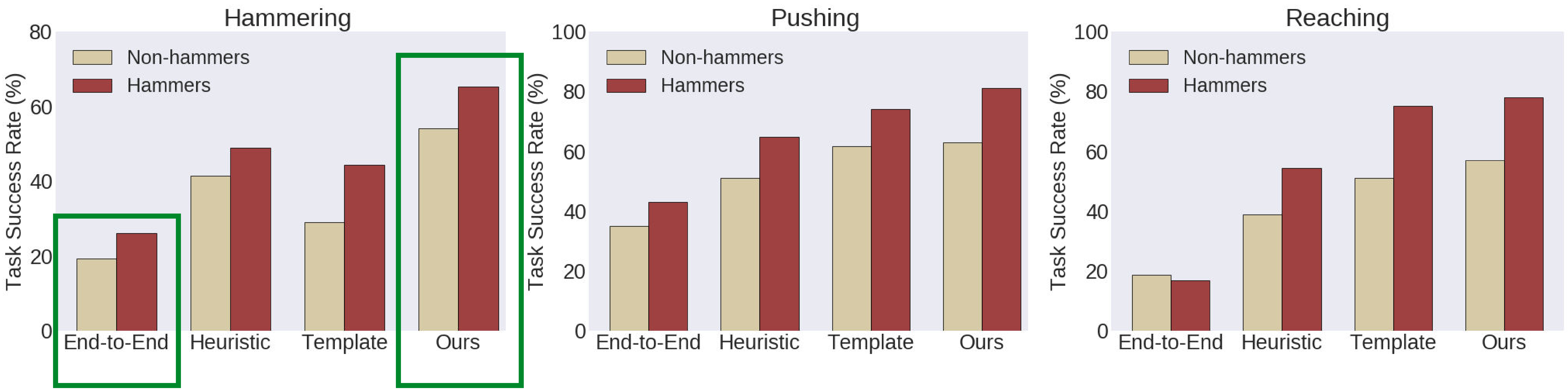


End task performance directly supervises **representation learning**.

Results: Hammering Task

Results: Reaching Task

Results: Quantitative Evaluation

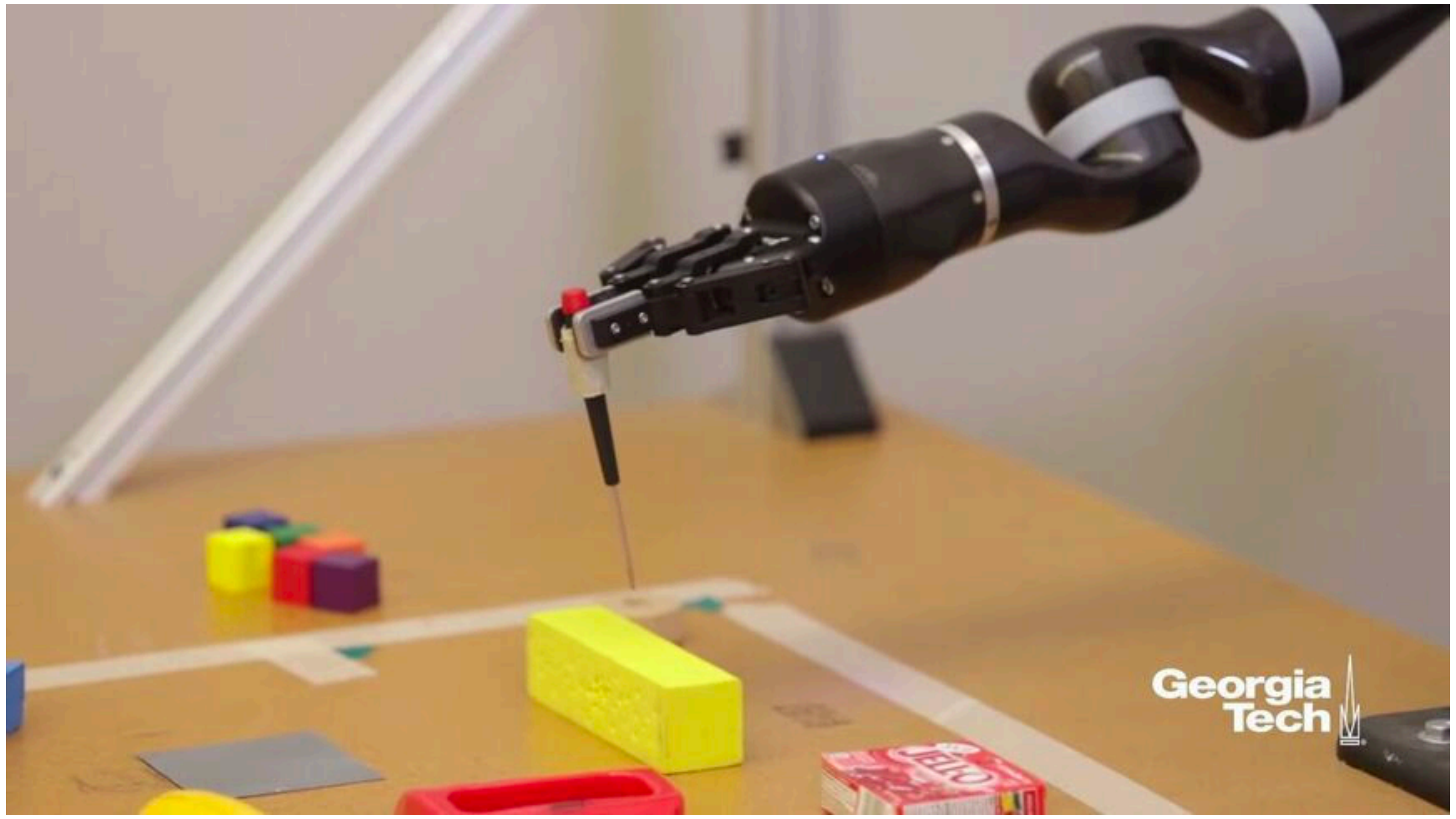


Keypoints as **intermediate representations** of tools are effective.

Qin et al. "KETO" ICRA'20

Tool Creation: Robot MacGyvering

Improvising tools for inventive problem solving



[Nair, Shrivatsav, Erickson, Chernova; RSS'19]

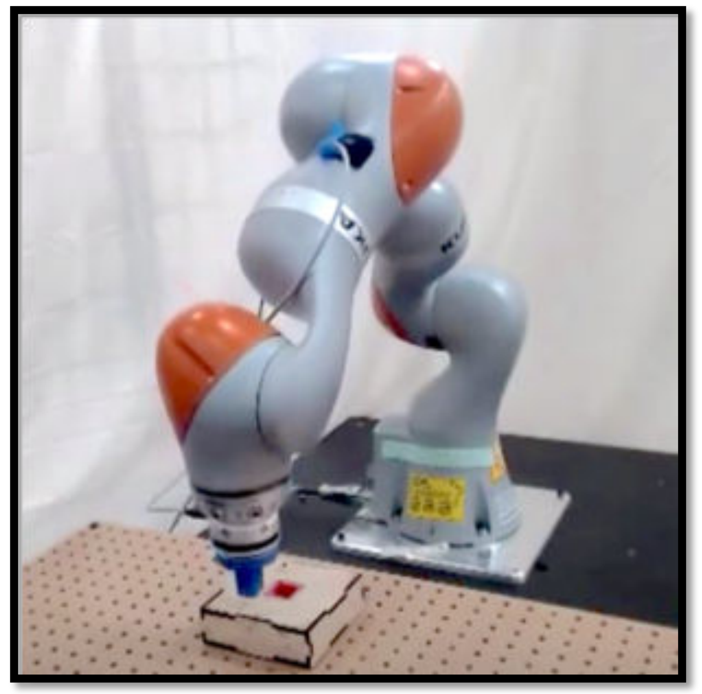
Tool Creation: Robot MacGyvering

Keypoints provides a scaffold for generating tools from object parts.

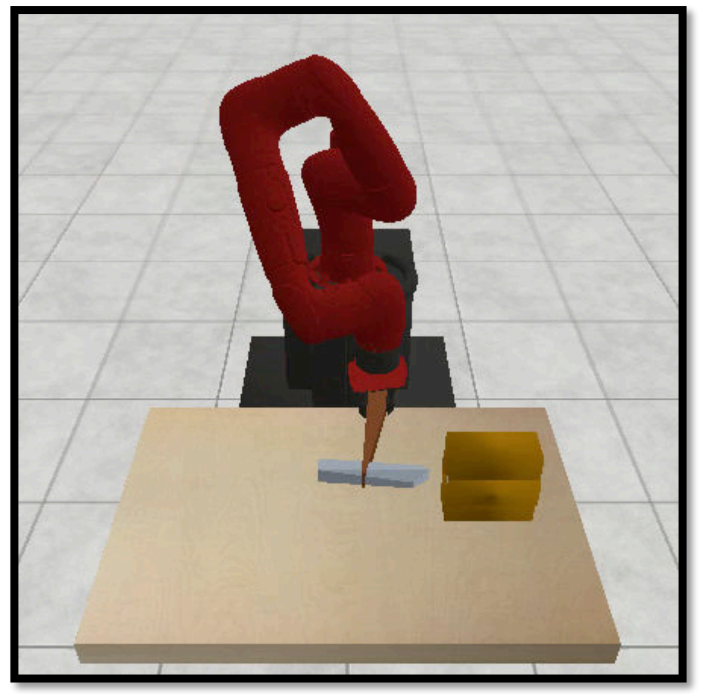
Qin et al. “KETO” ICRA’20

Tool Creation: Robot Creates New Tool for Hammering

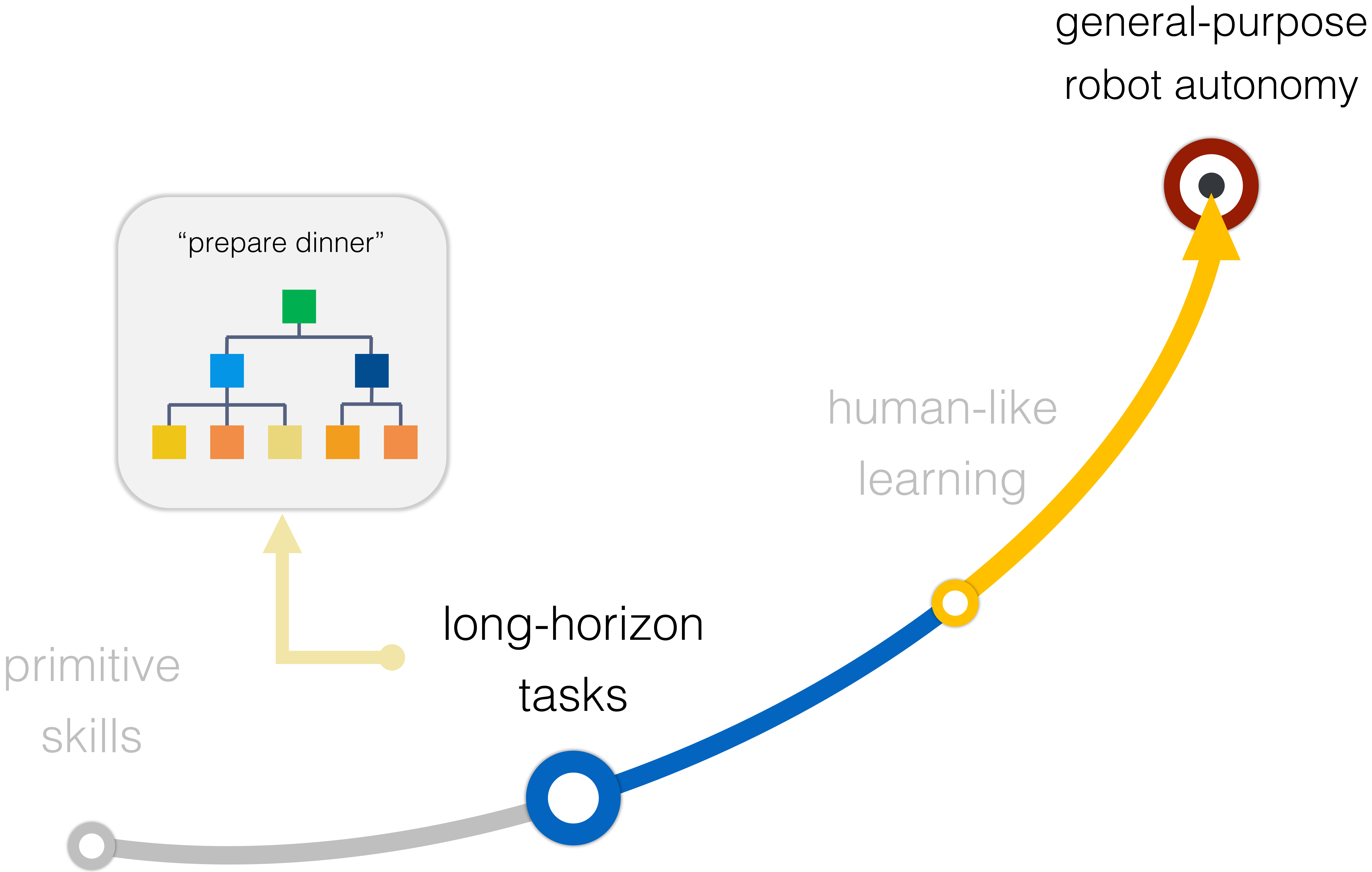
Summary - Part I



Self-supervised learning is a powerful tool to scale up primitive skill learning without human supervision.

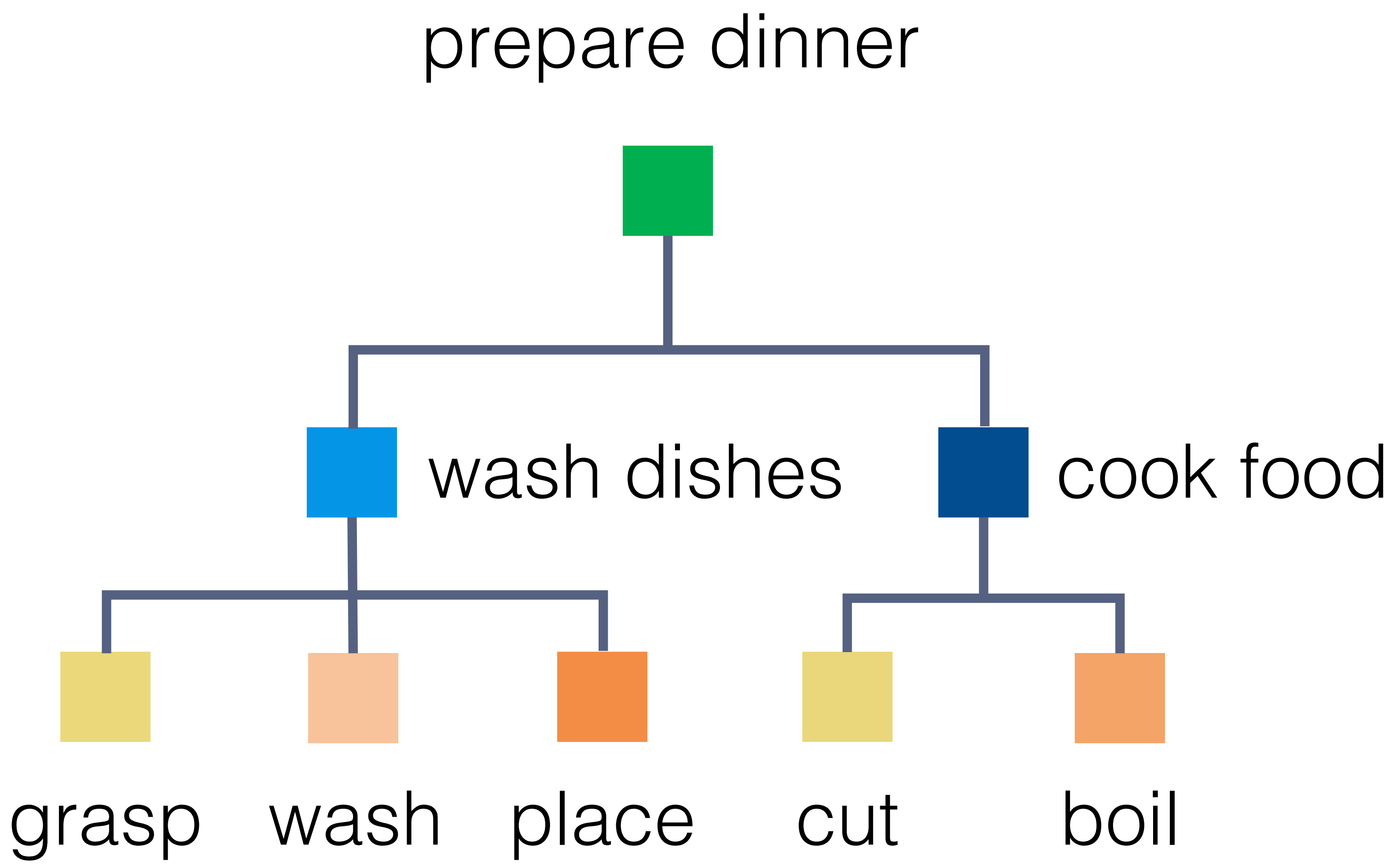


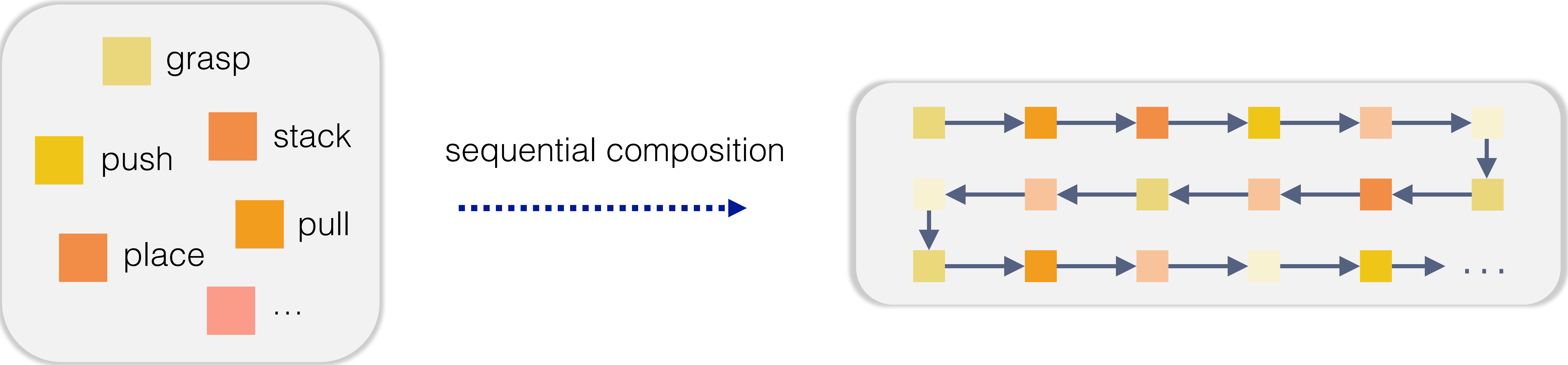
Feedback from downstream tasks and **structural priors** give rise to more compact and informative representations.


general-purpose
robot autonomy
“prepare dinner”
human-like
learning
long-horizon
tasks
primitive
skills

Long-Horizon Tasks

"prepare dinner"



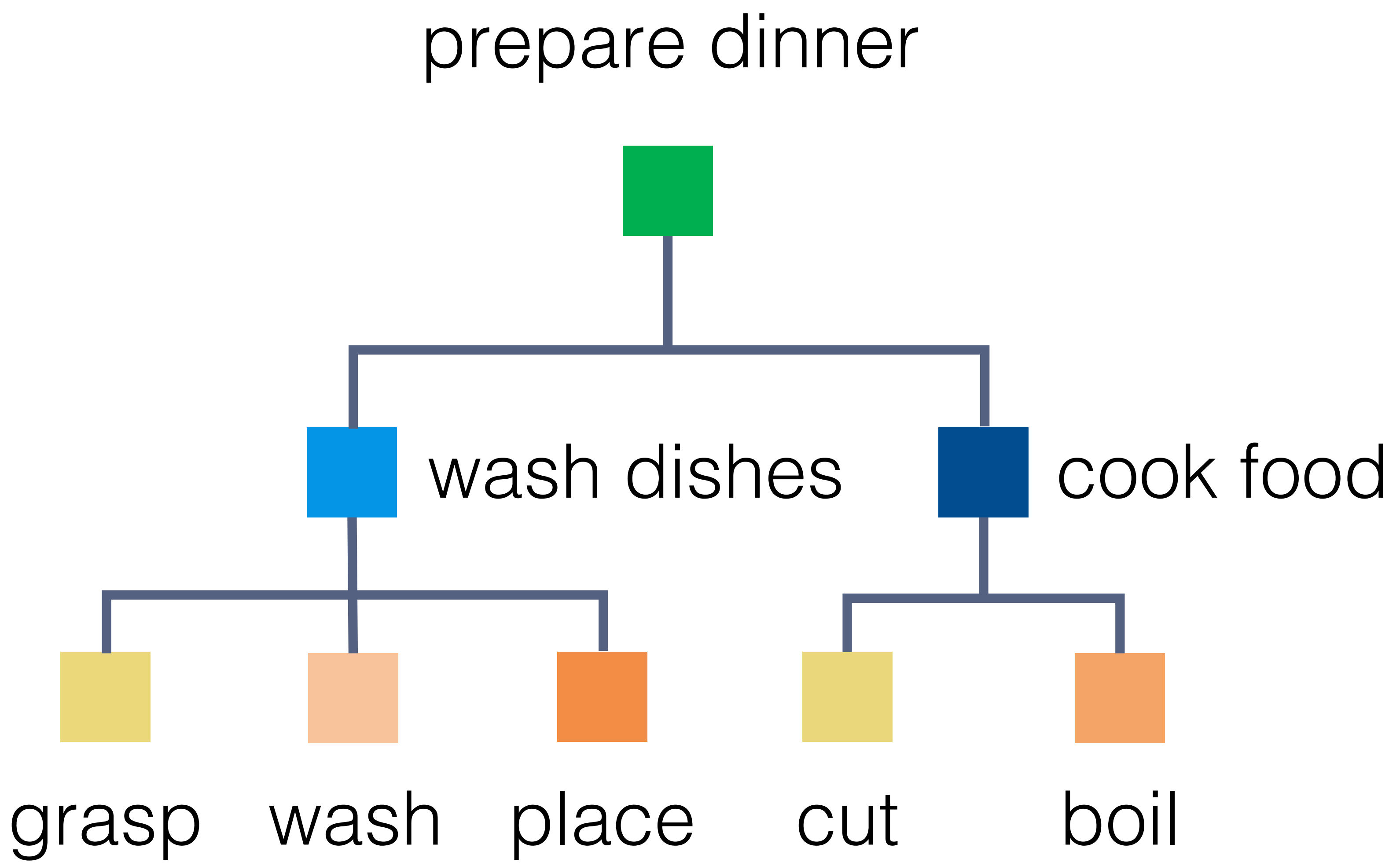


robot primitive skills

Intractable!

Long-Horizon Tasks

"prepare dinner"



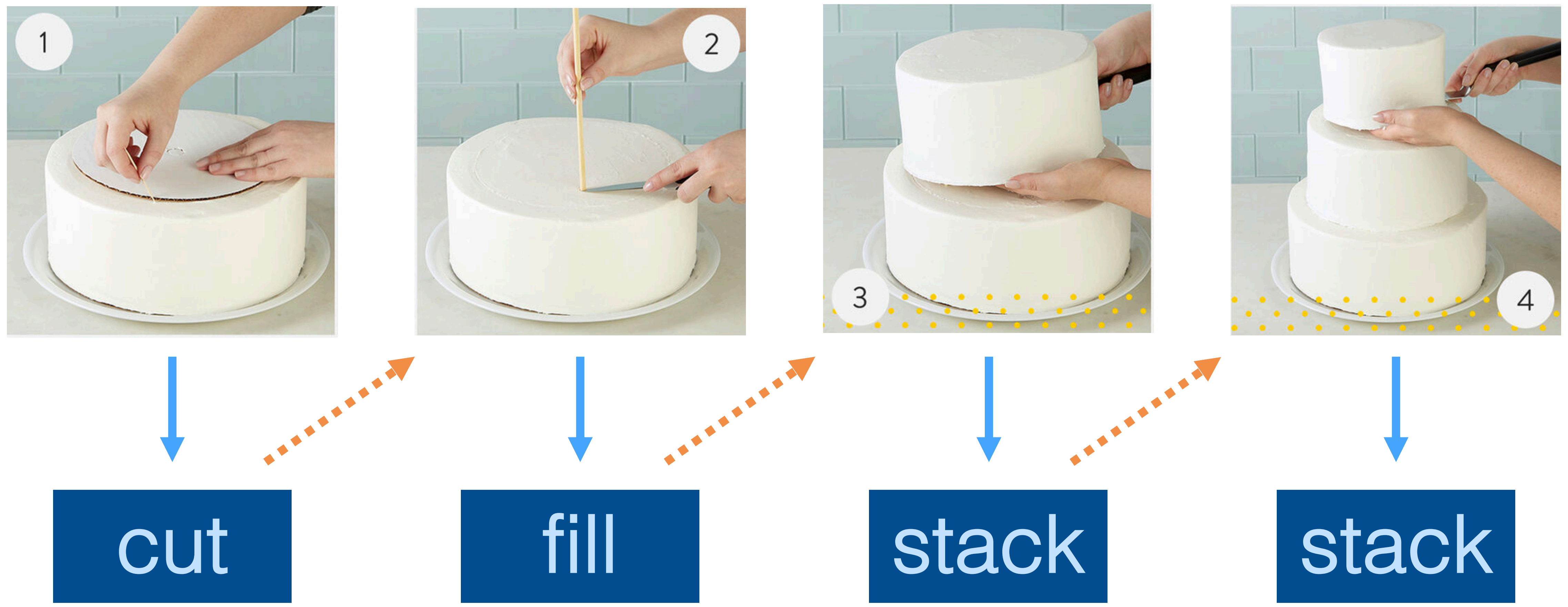
Challenge: **Task complexity** grows exponentially.

Key idea: Leveraging **hierarchy** and **abstraction** of long-horizon tasks

“How to make a cake?”

high-level plan

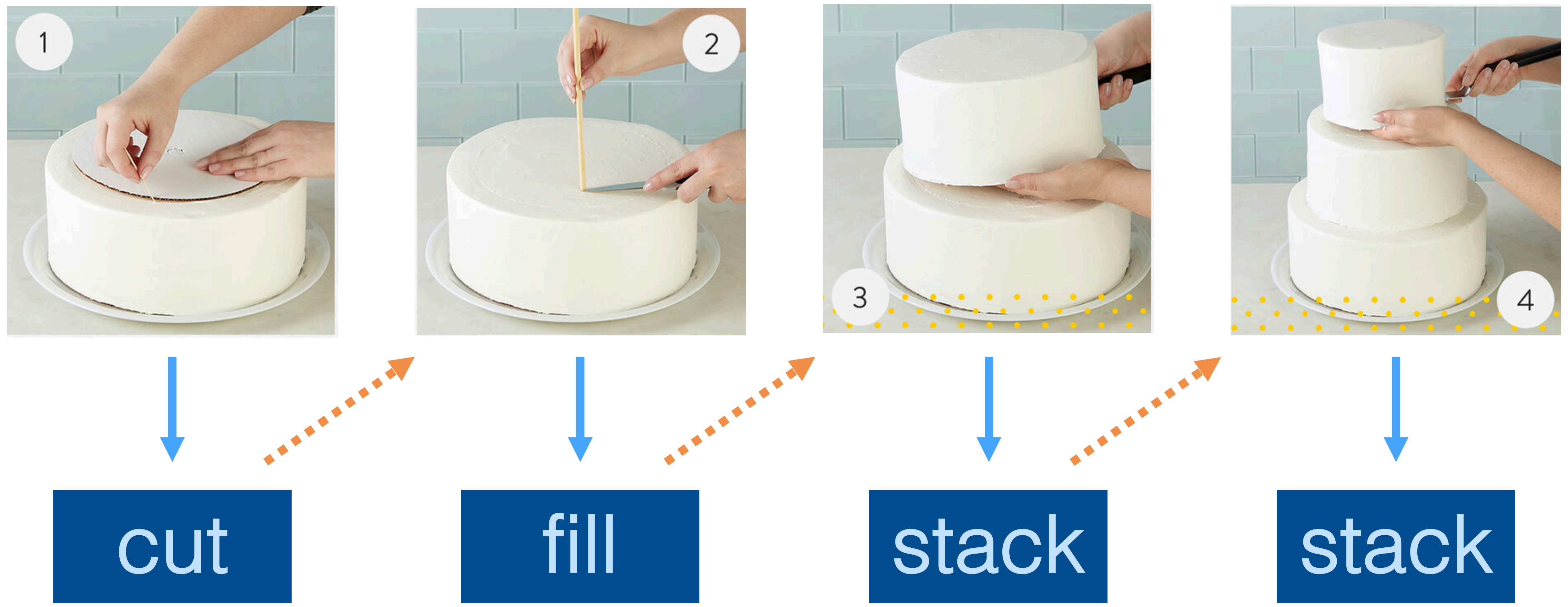
low-level action



"How to make a cake?"

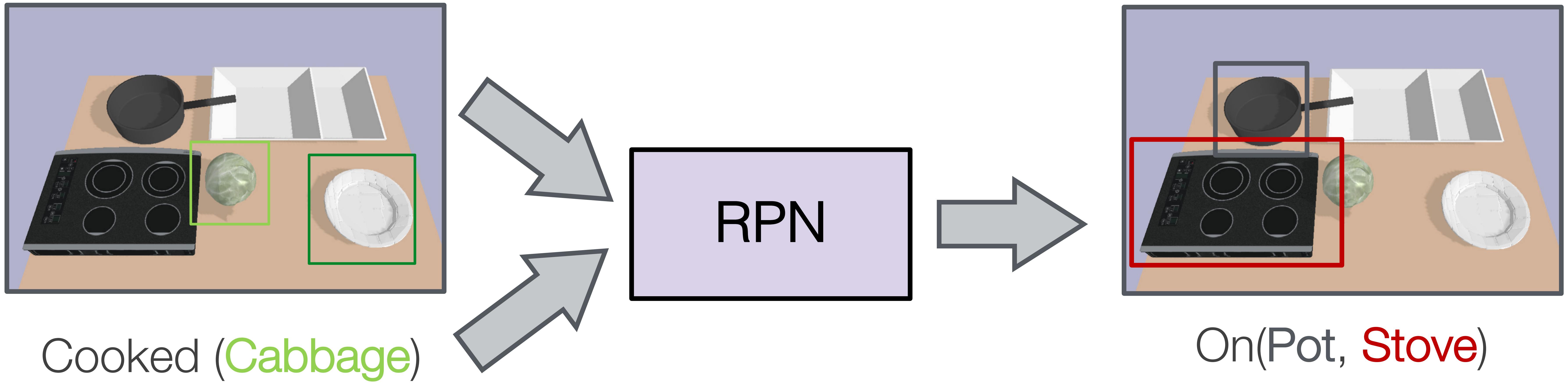
high-level plan

low-level action



Current Observation

Task Goal



Cooked (Cabbage)

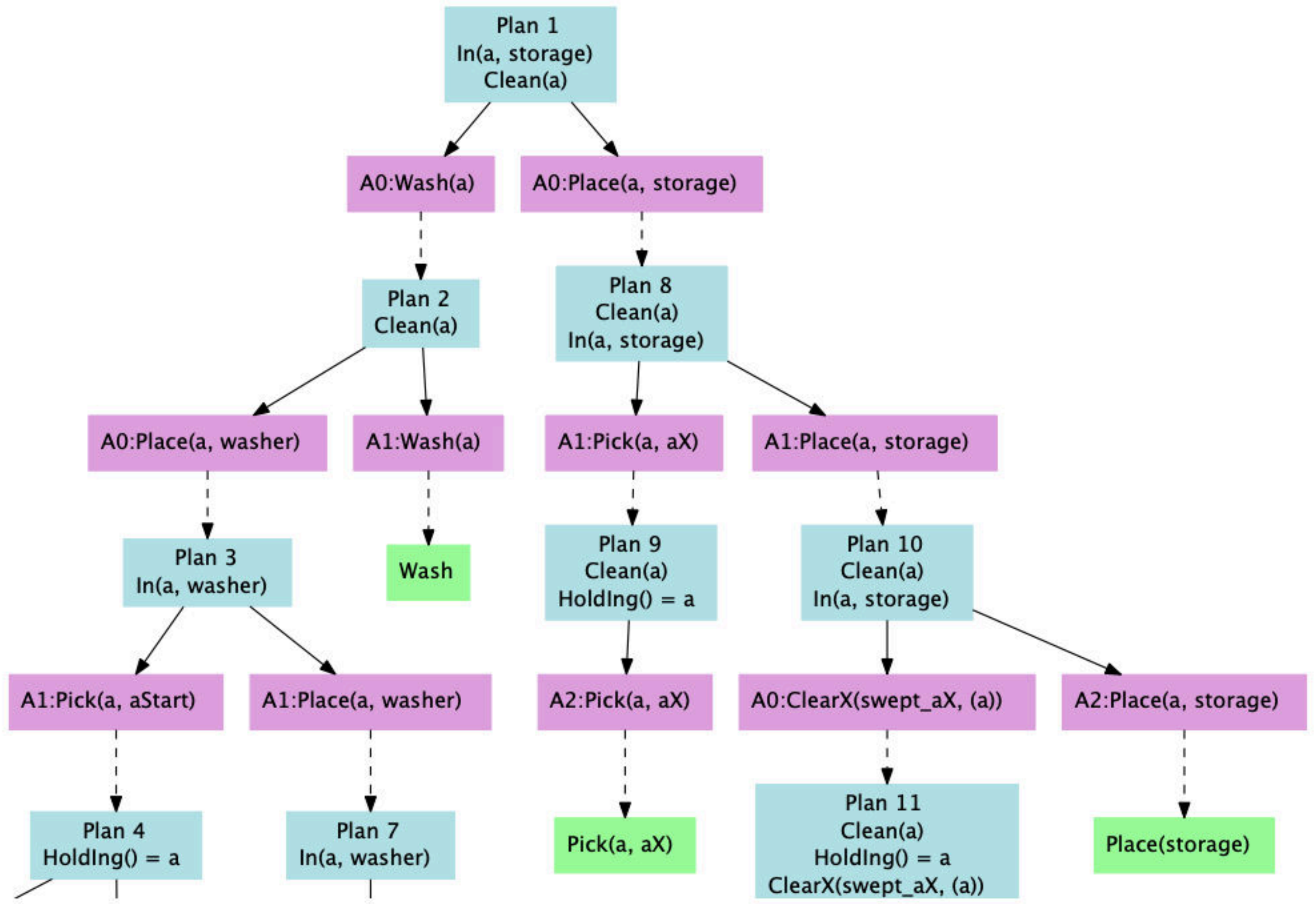
On (Cabbage, Plate)

On(Pot, Stove)

Next Subgoal

Xu et al. "Regression Planning Networks" NeurIPS'19

Regression Planning Network

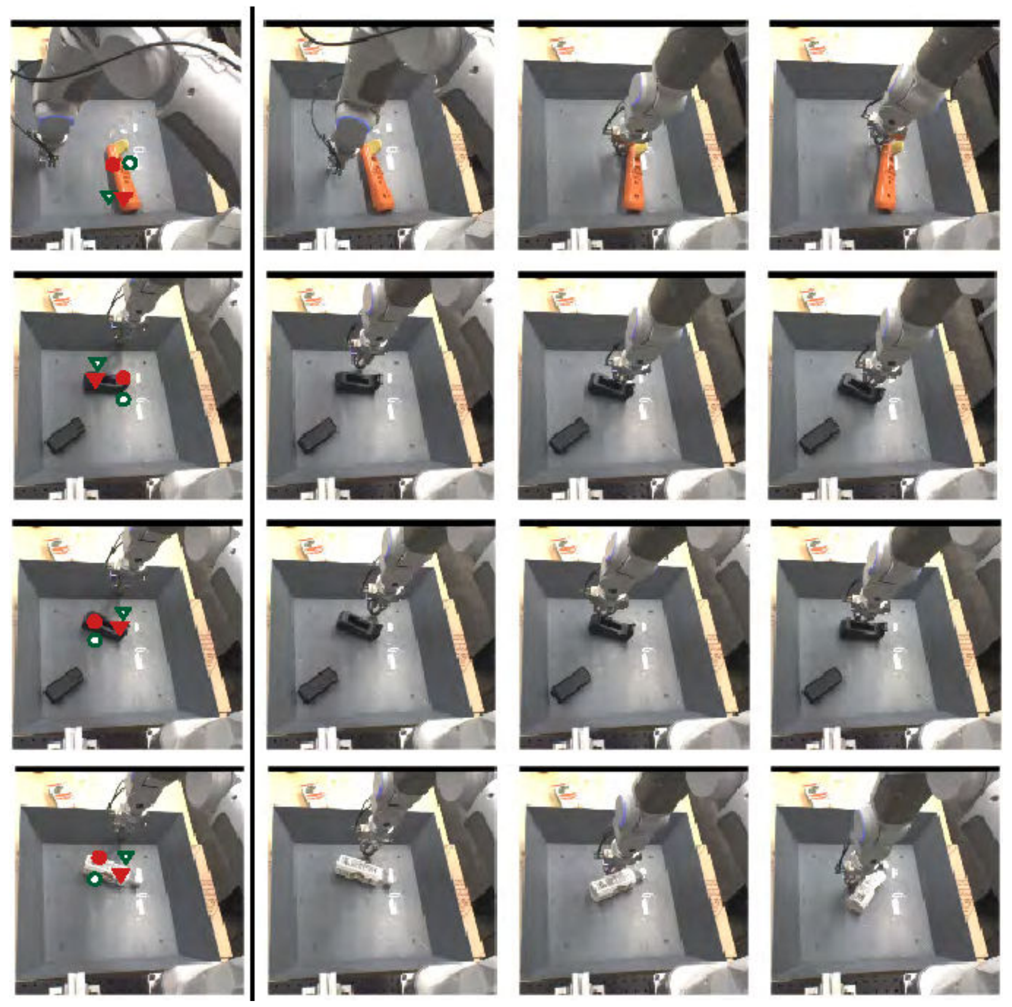


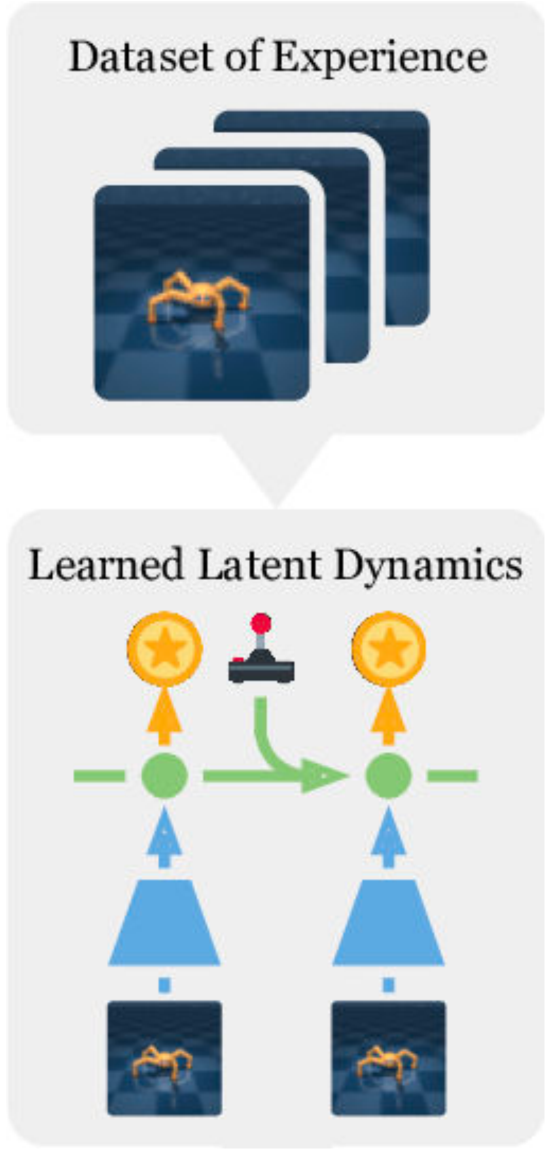
[Waldinger 1975; Korf 1987; Kaelbling ICRA'11]

classical symbolic planning

human-interpretable and long-horizon

symbols and planning domain required





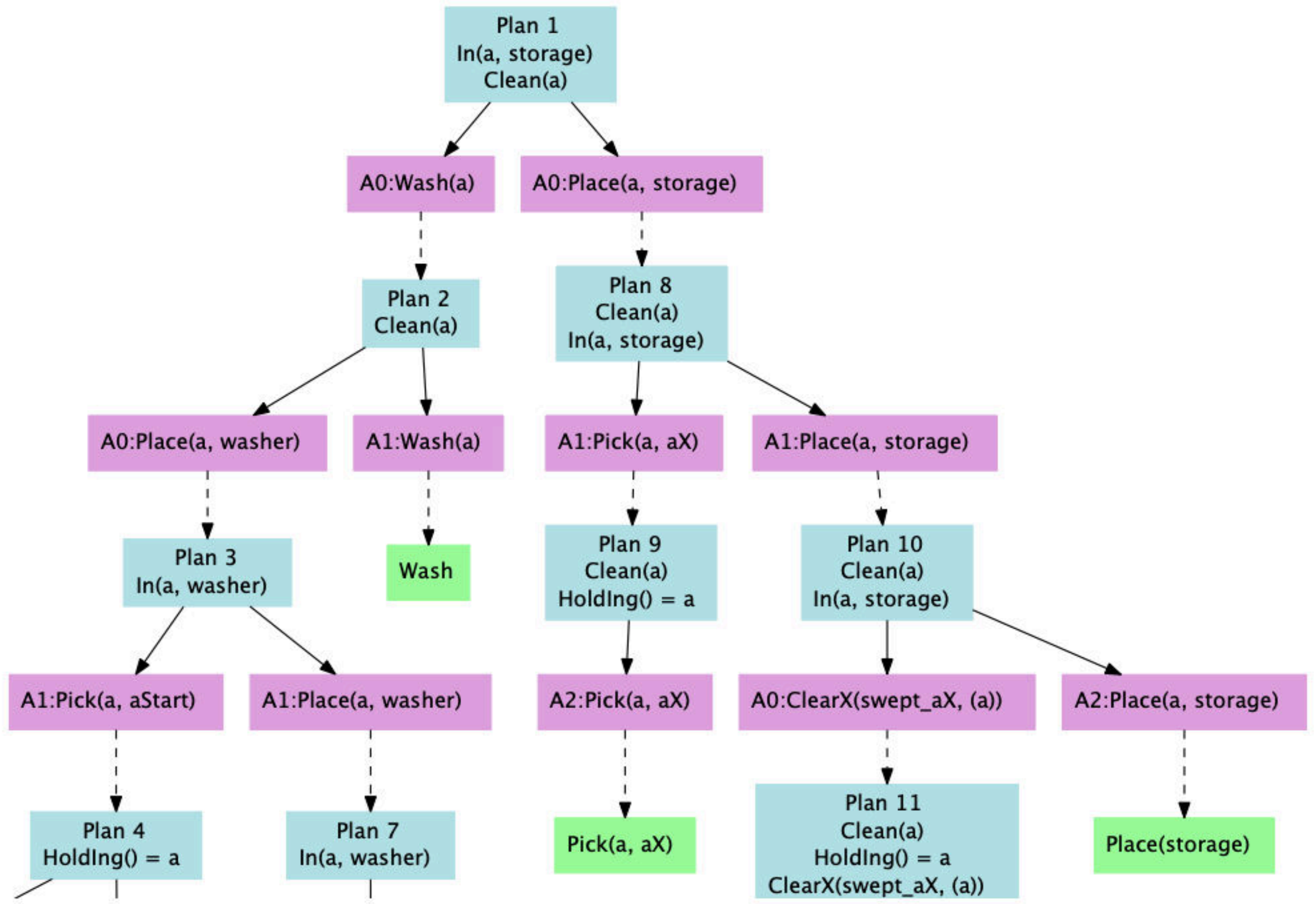
[Finn et al. ICRA'17; Oh et al. NIPS'15; Hafner et al. ICLR'20]

plan from observations

grounded on raw sensory data

myopic sampling, short-horizon tasks

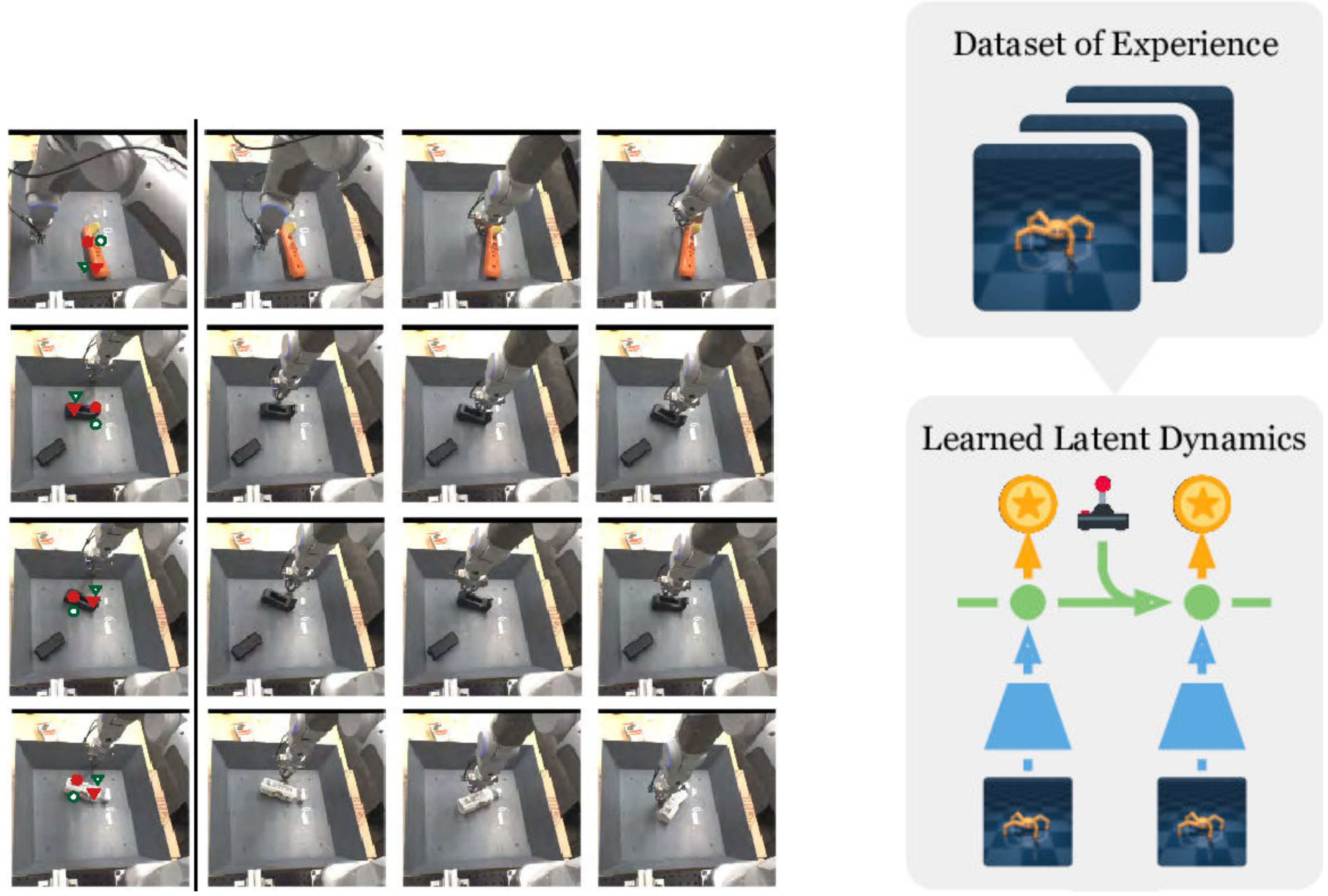
Regression Planning Network



[Waldinger 1975; Korf 1987; Kaelbling ICRA'11]

classical symbolic planning

+

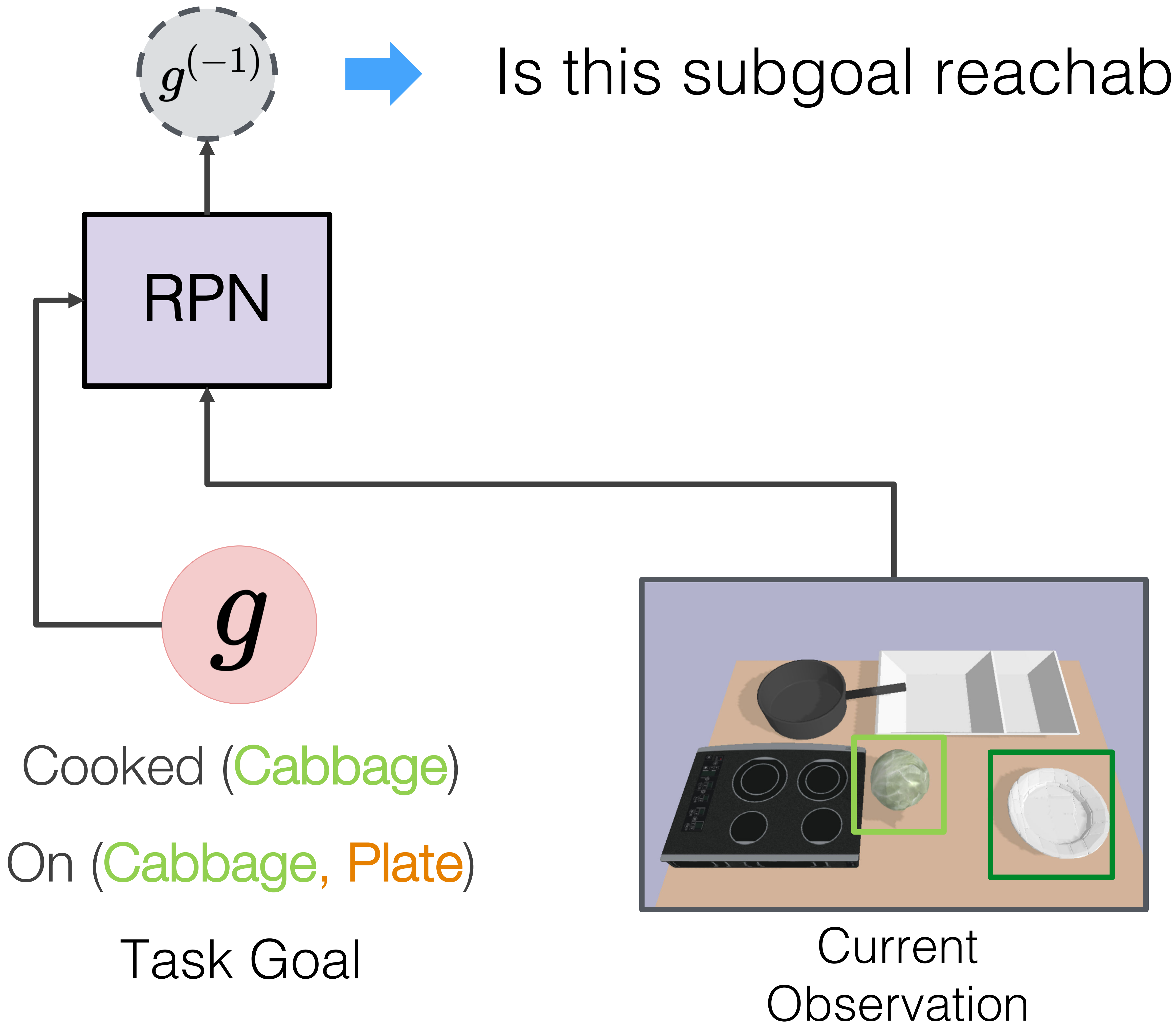


[Finn et al. ICRA'17; Oh et al. NIPS'15; Hafner et al. ICLR'20]

plan from observations

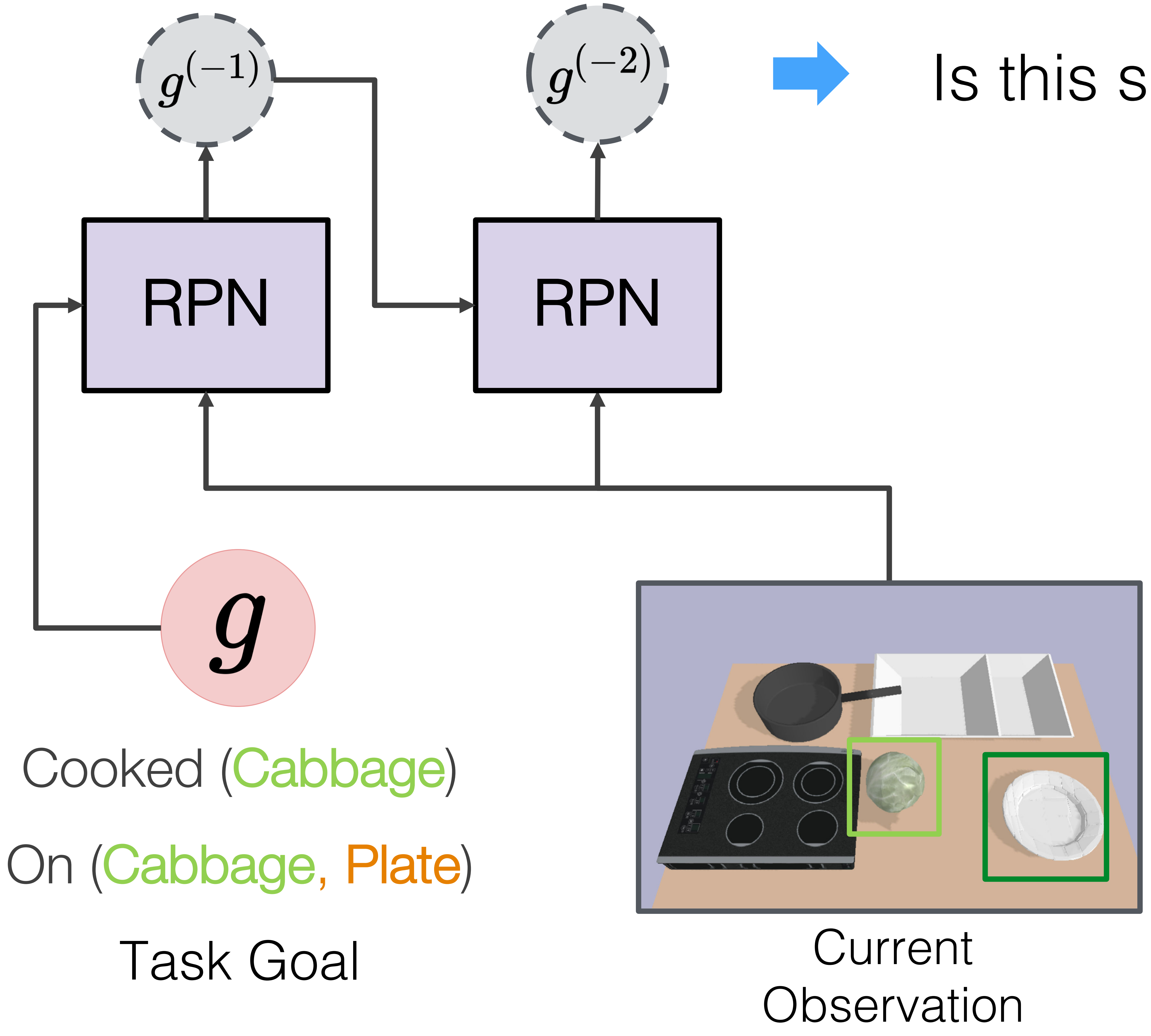
plan backward in a symbolic space conditioning on the visual observation

Regression Planning Network



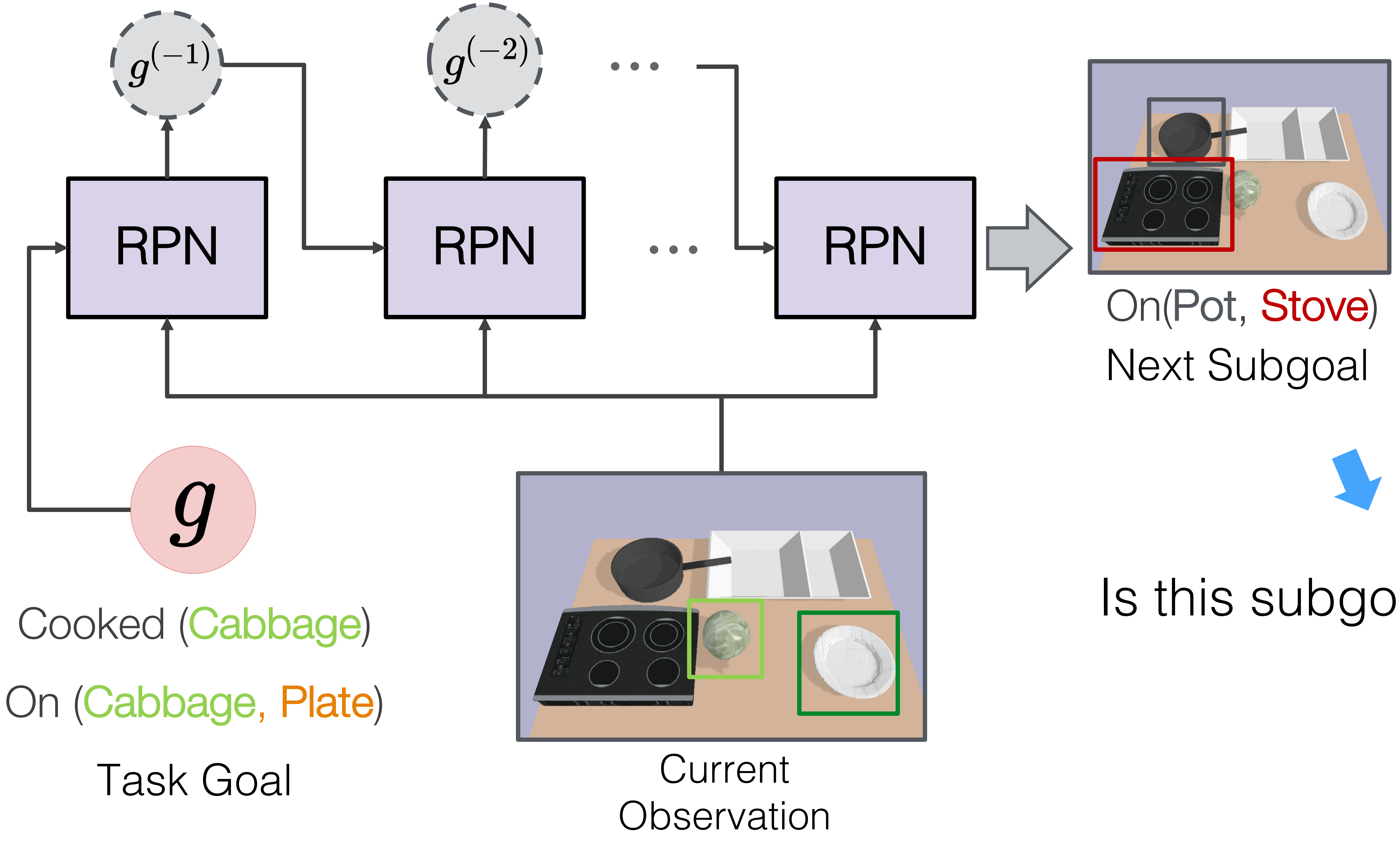
Is this subgoal reachable? **NO!**

Regression Planning Network



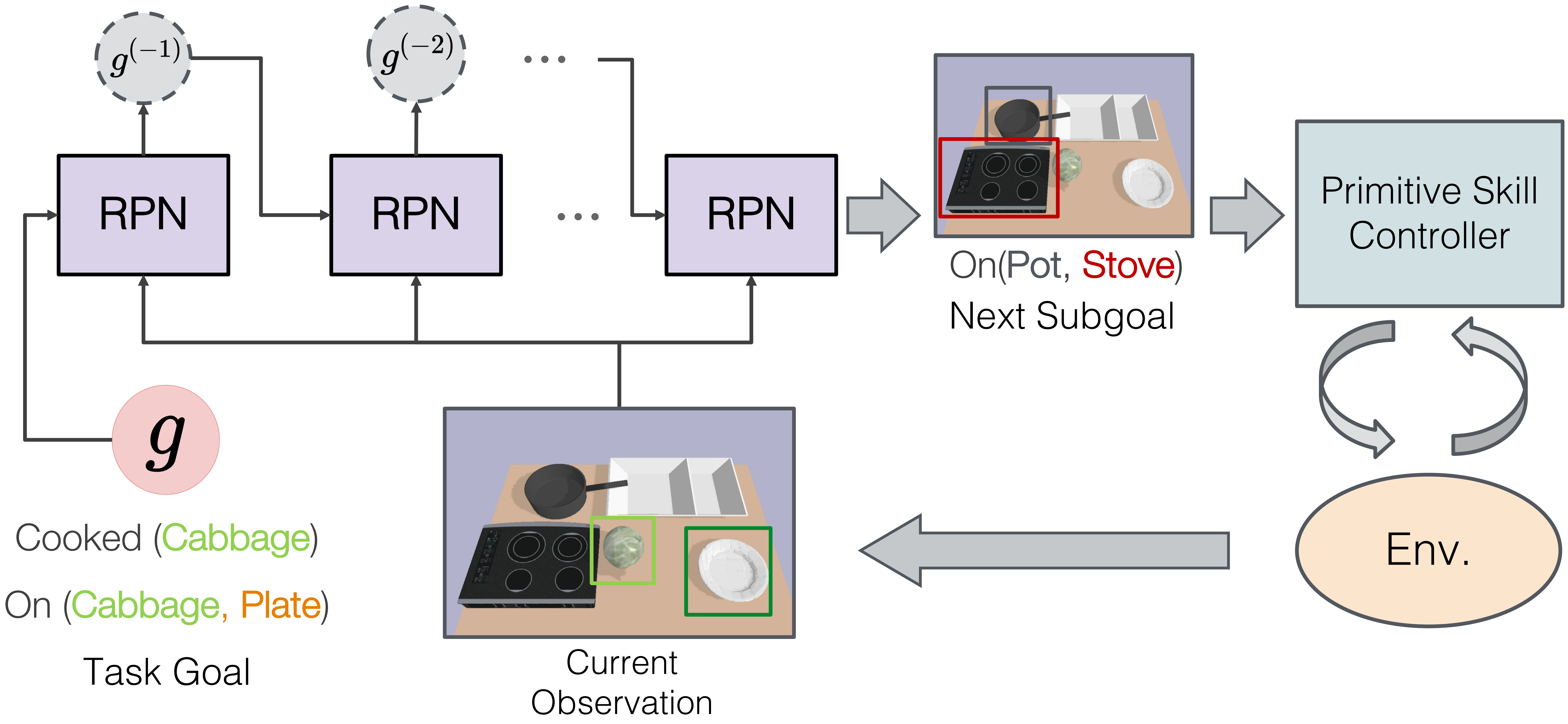
Is this subgoal reachable? **NO!**

Regression Planning Network



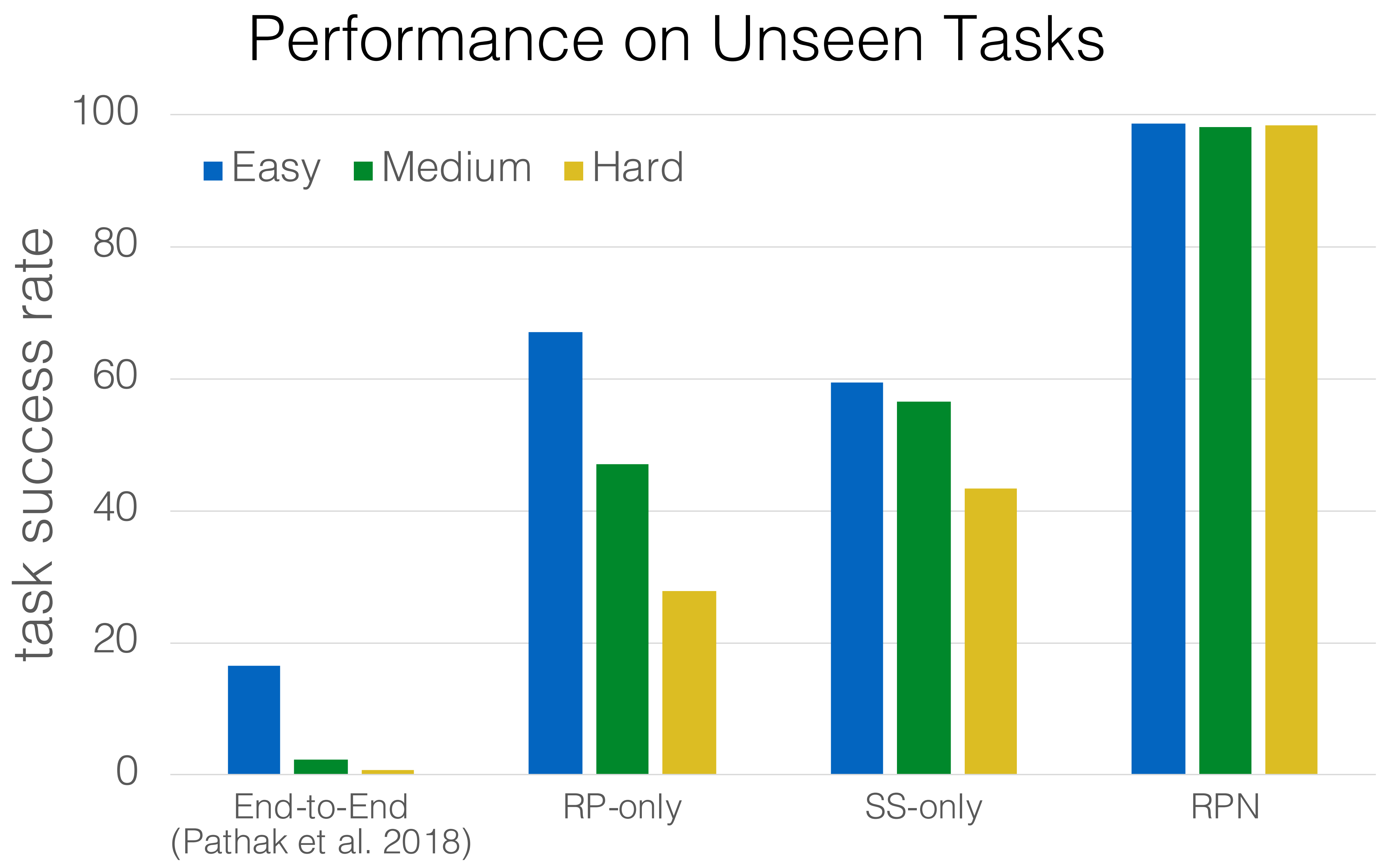
Is this subgoal reachable? **YES!**

Regression Planning Network



Xu et al. “Regression Planning Networks” NeurIPS’19

Regression Planning Network



Qualitative

(cook 3 dishes with 4 ingredients)

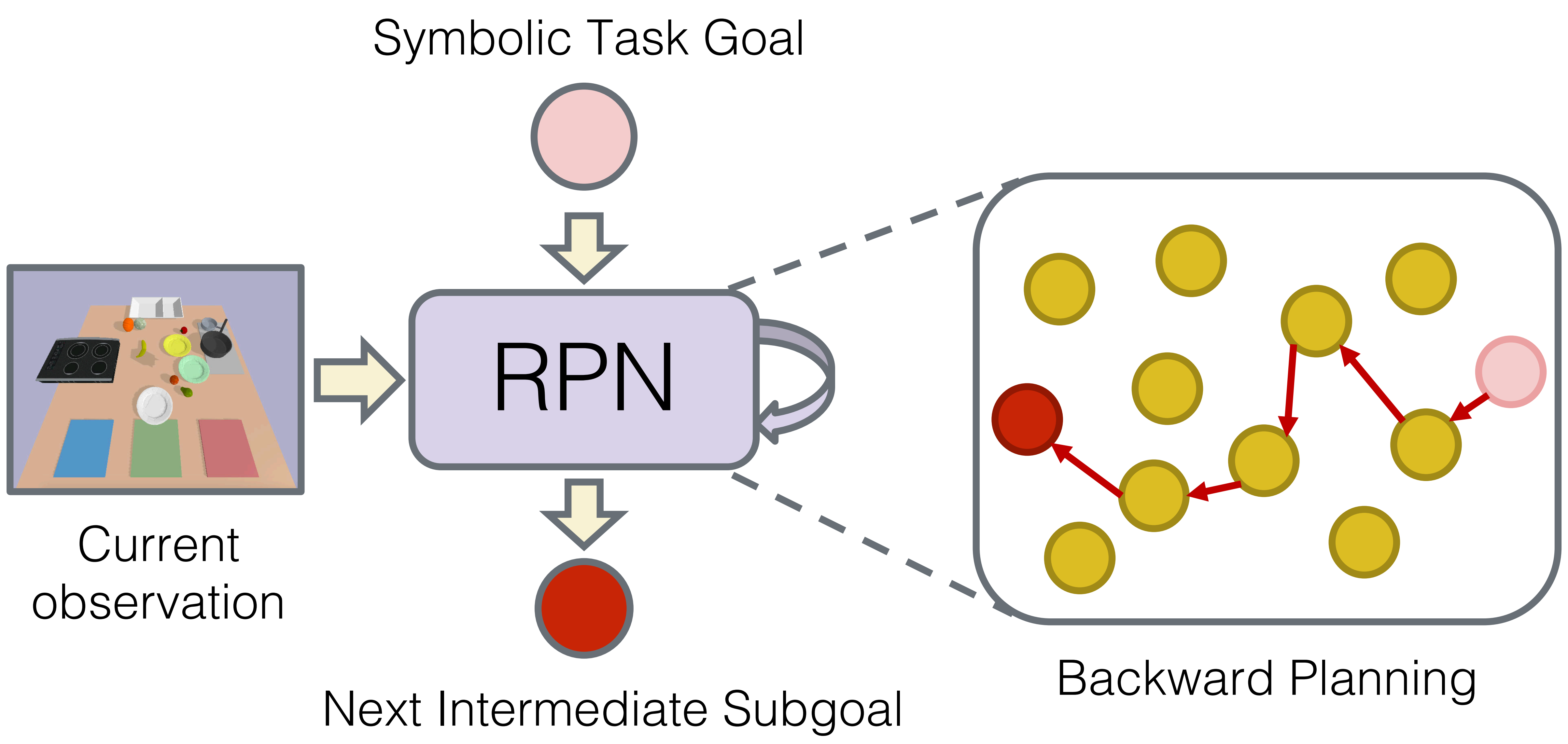
Quantitative

(the higher the better)

Xu et al. “Regression Planning Networks” NeurIPS’19

Regression Planning Network

- Recursively planning backward (regression planning) on **symbolic abstraction**
- Method works on **visual input** without specifying a planning domain
- Learning from video demonstrations and **zero-shot generalization** to new tasks
- Low-level **primitive skills** are modeled as **pre-defined API calls**.



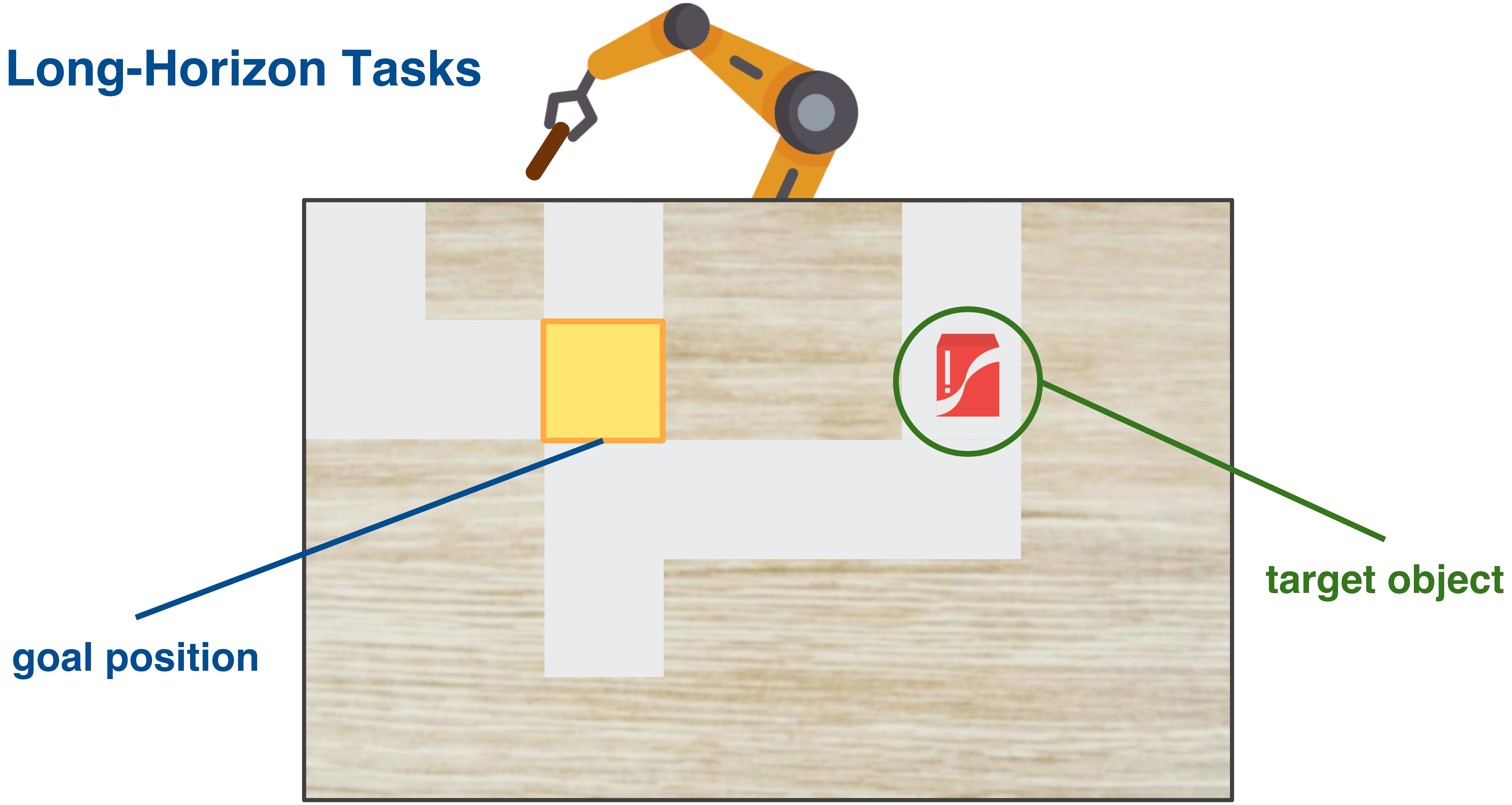
Xu et al. “Regression Planning Networks” NeurIPS’19

Regression Planning Network

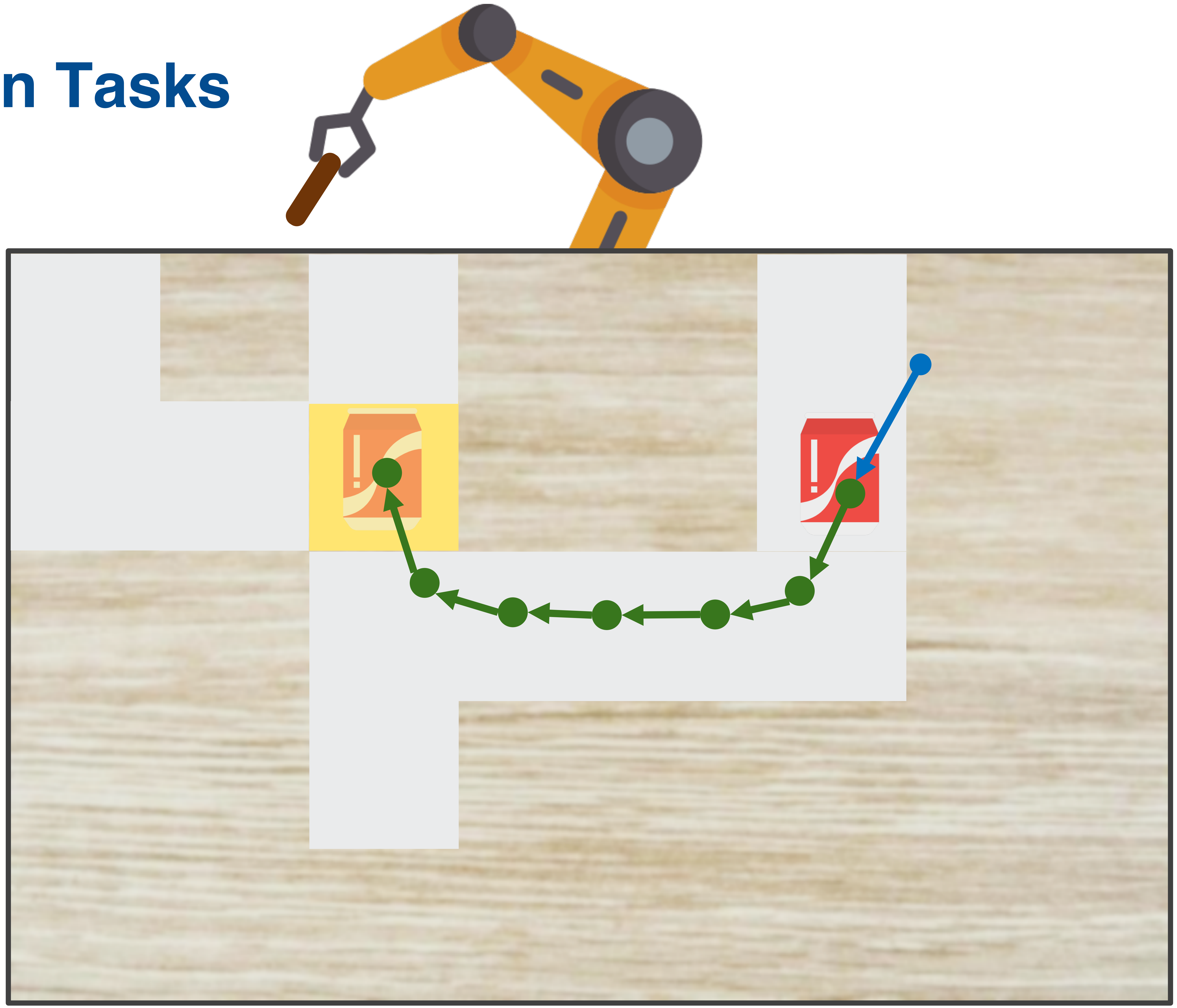
- Recursively planning backward (regression planning) on **symbolic abstraction**
- Method works on **visual input** without specifying a planning domain
- Learning from video demonstrations and **zero-shot generalization** to new tasks
- Low-level **primitive skills** are modeled as **pre-defined API calls**.

Can we learn and plan **primitive skills** and **task plans** jointly?

Xu et al. "Regression Planning Networks" NeurIPS 2019


Long-Horizon Tasks
goal position
target object

Long-Horizon Tasks



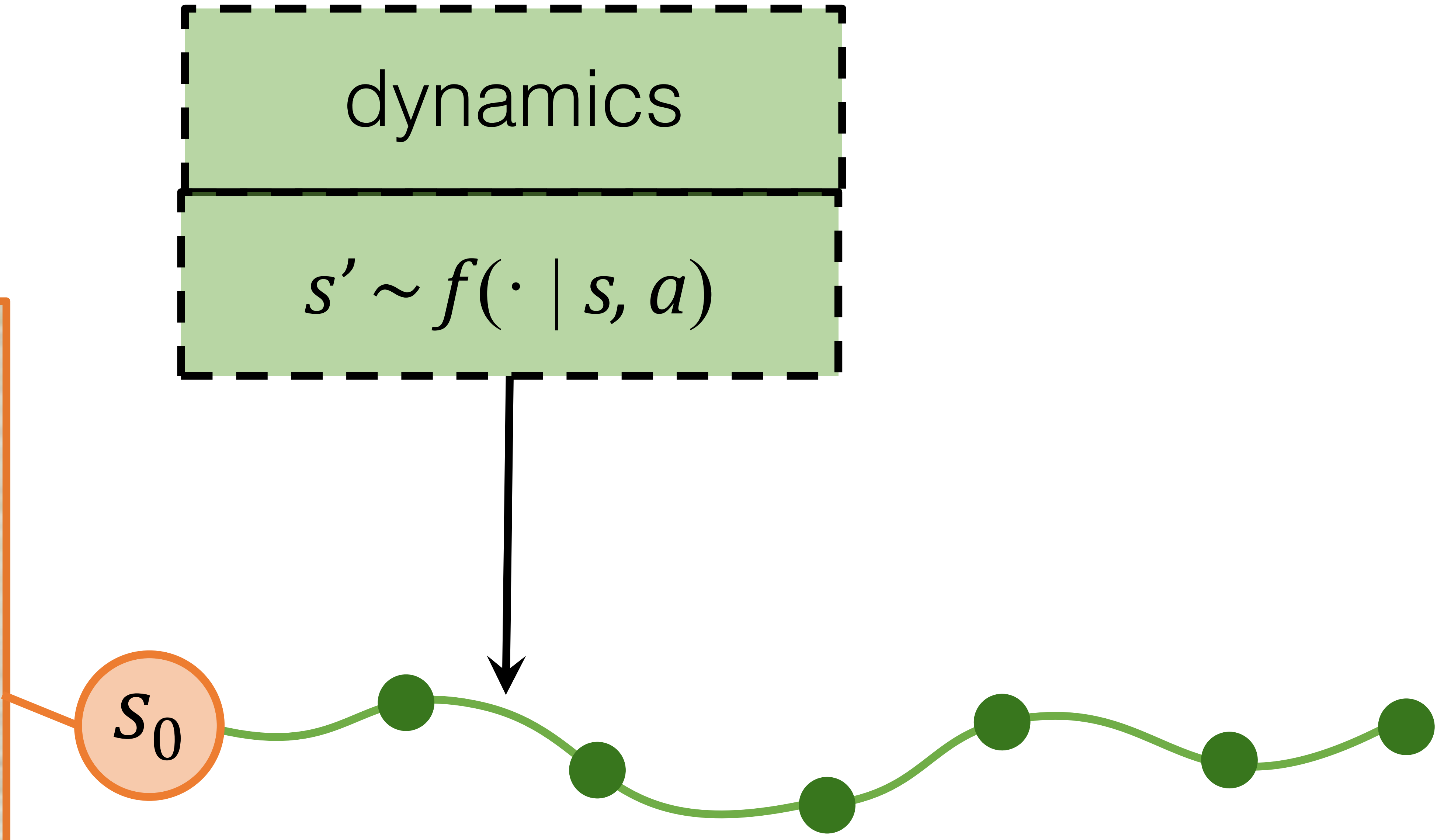
Long-Horizon Tasks

Long-Horizon Tasks

Long-Horizon Tasks: Model-Based Learning

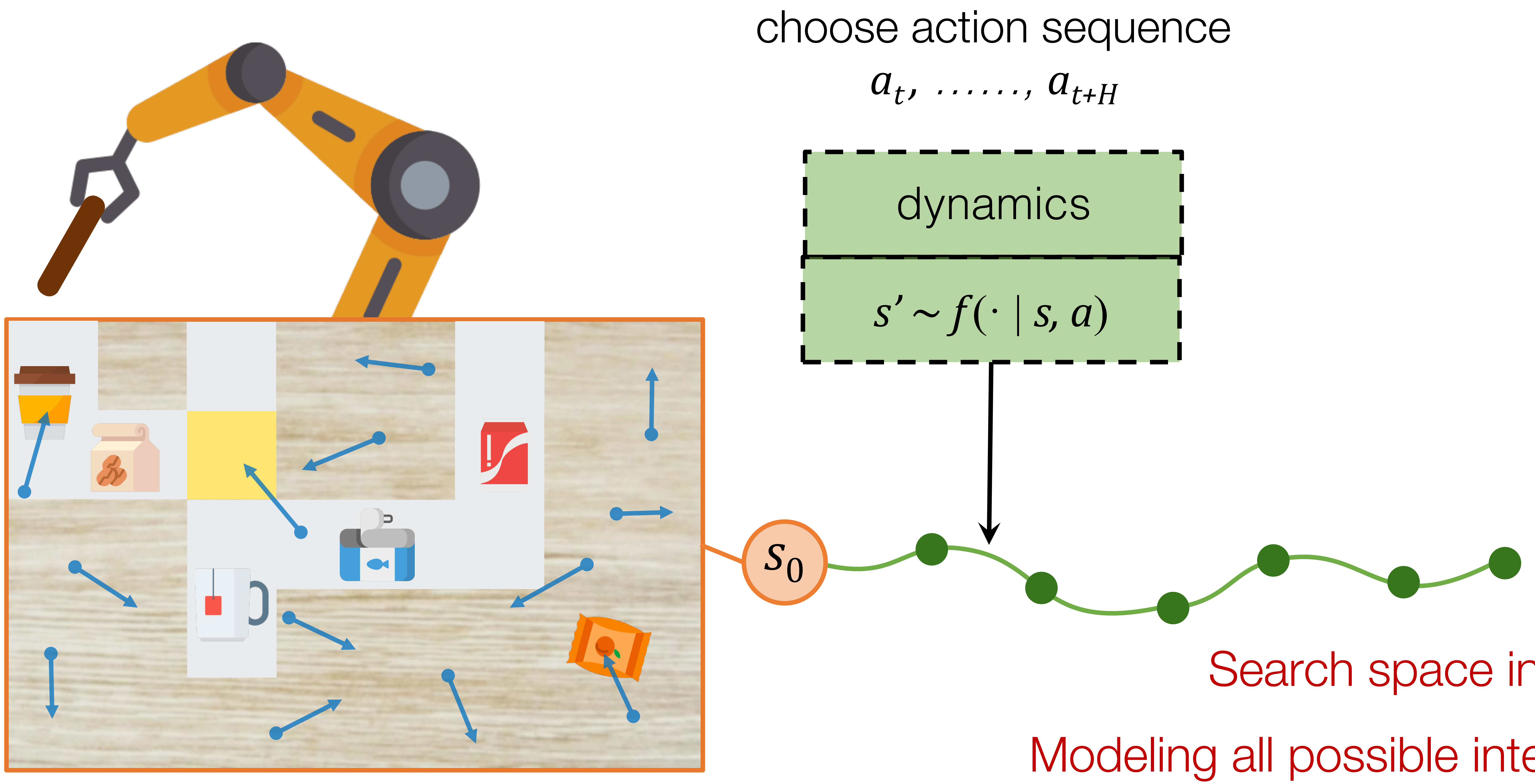
choose action sequence

$a_t, \ldots\ldots, a_{t+H}$



[Deisenroth et al, RSS'07], [Guo et al, NeurIPS'14], [Watter et al, NeurIPS'15], [Finn et al, ICRA'17], ……

Long-Horizon Tasks: Model-Based Learning

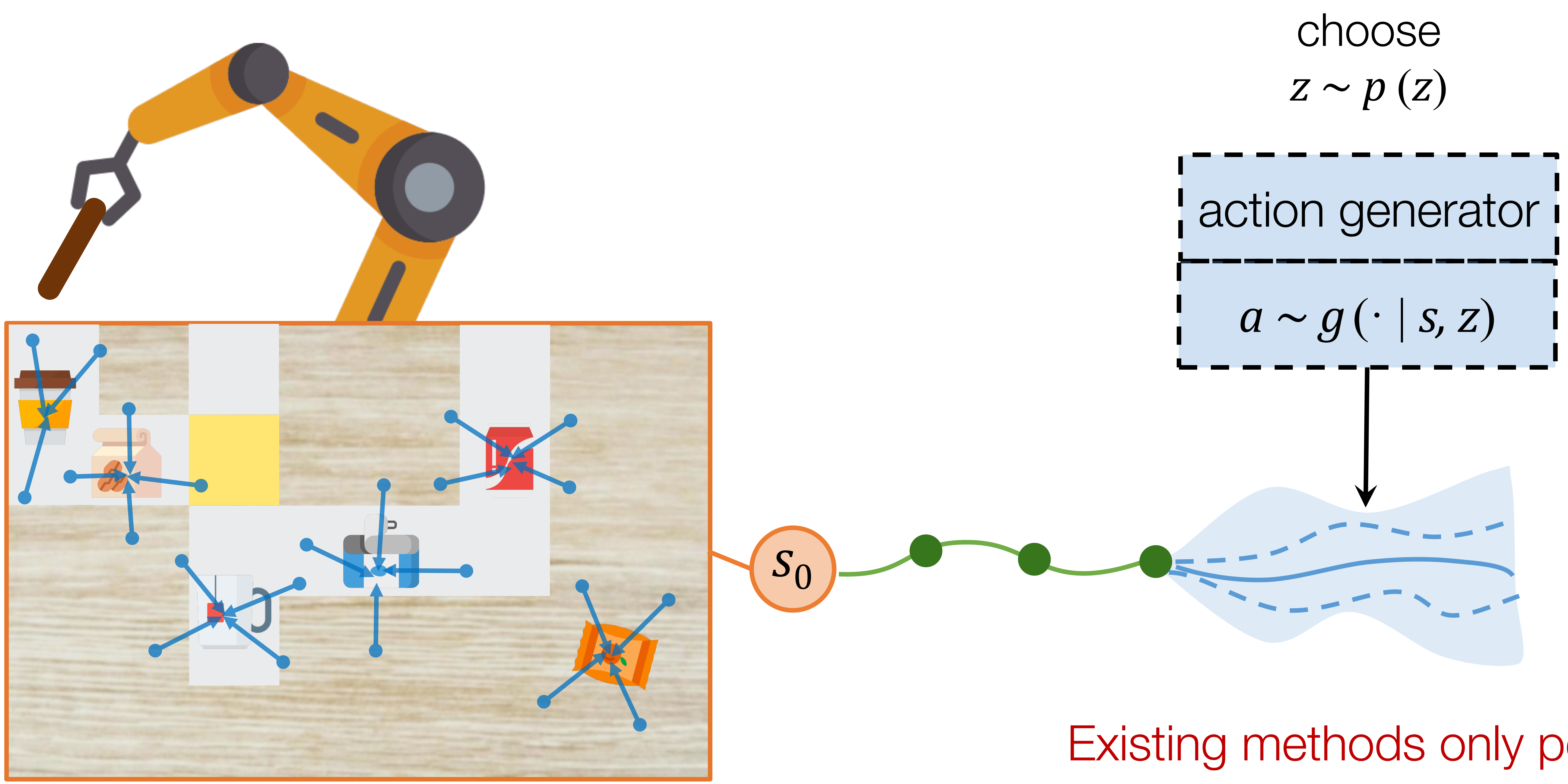


Search space increases exponentially.

Modeling all possible interactions is intractable.

[Deisenroth et al, RSS'07], [Guo et al, NeurIPS'14], [Watter et al, NeurIPS'15], [Finn et al, ICRA'17], ……

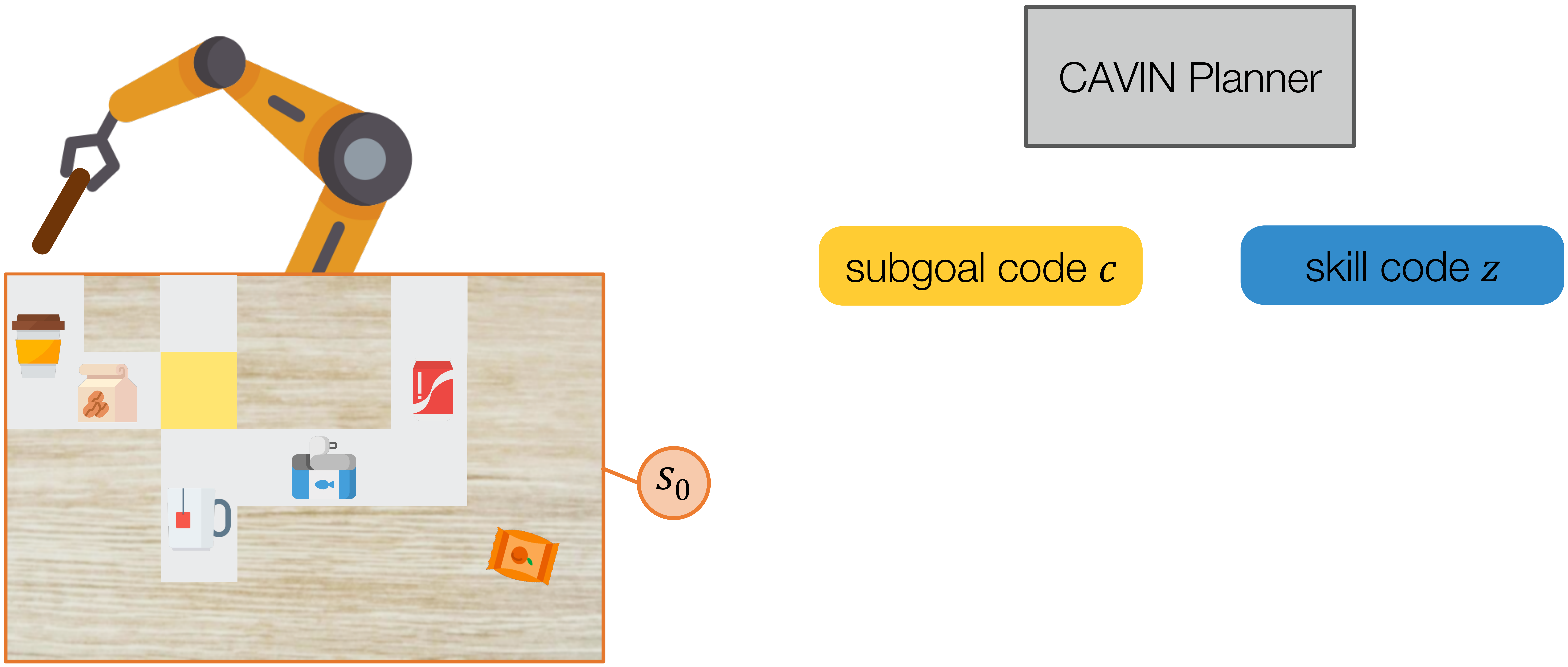
Long-Horizon Tasks: Planning in Learned Latent Spaces



Existing methods only perform flat planning.

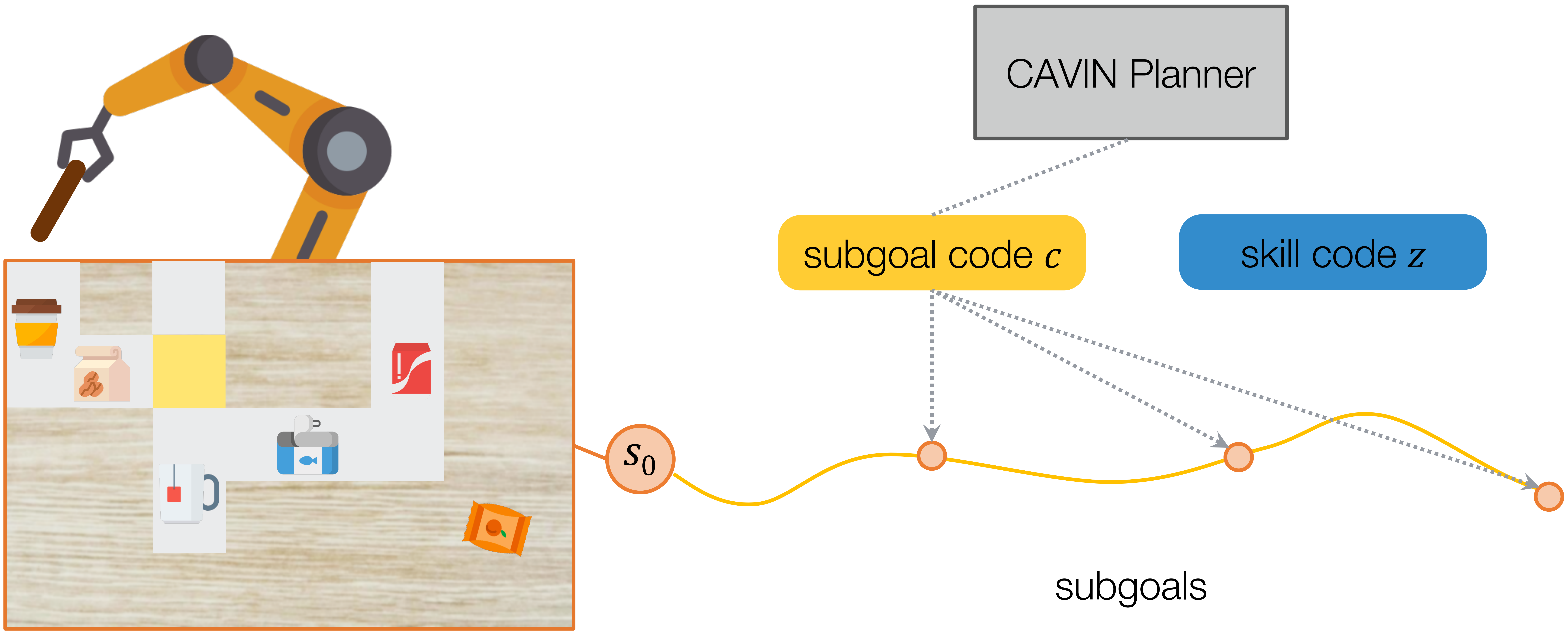
[Ichter et al, ICRA'19], [Kurutach et al, NeurIPS'18], [Co-Reyes et al, ICML'18]

Long-Horizon Tasks: Hierarchical Planning in Latent Spaces



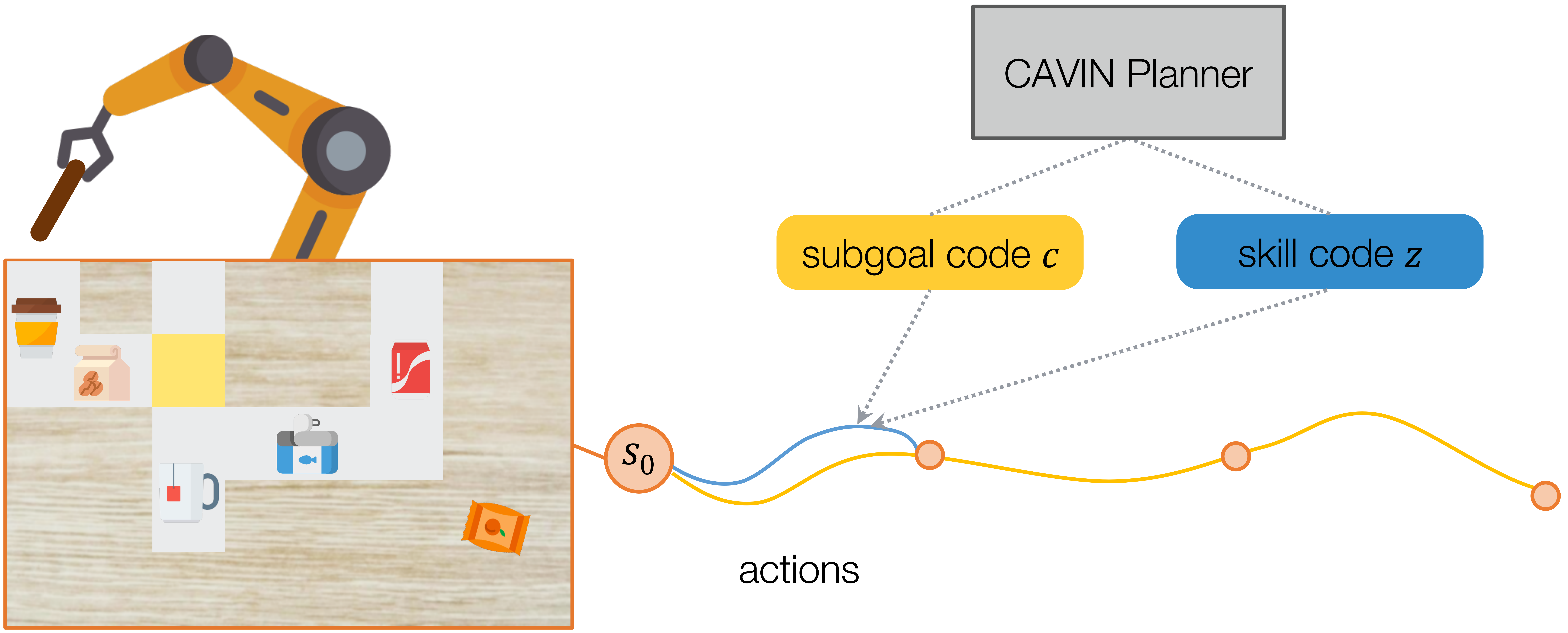
Fang et al. “CAVIN” CoRL’19

Long-Horizon Tasks: Hierarchical Planning in Latent Spaces



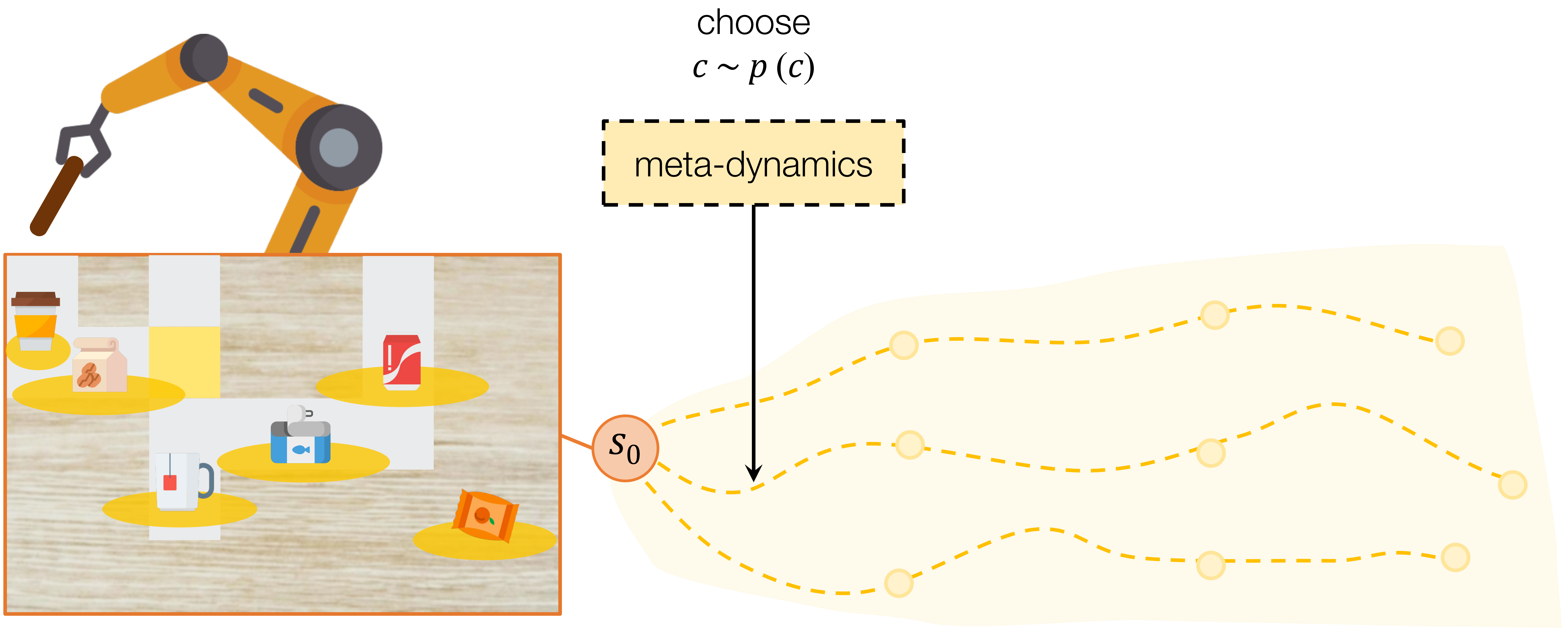
Fang et al. "CAVIN" CoRL'19

Long-Horizon Tasks: Hierarchical Planning in Latent Spaces



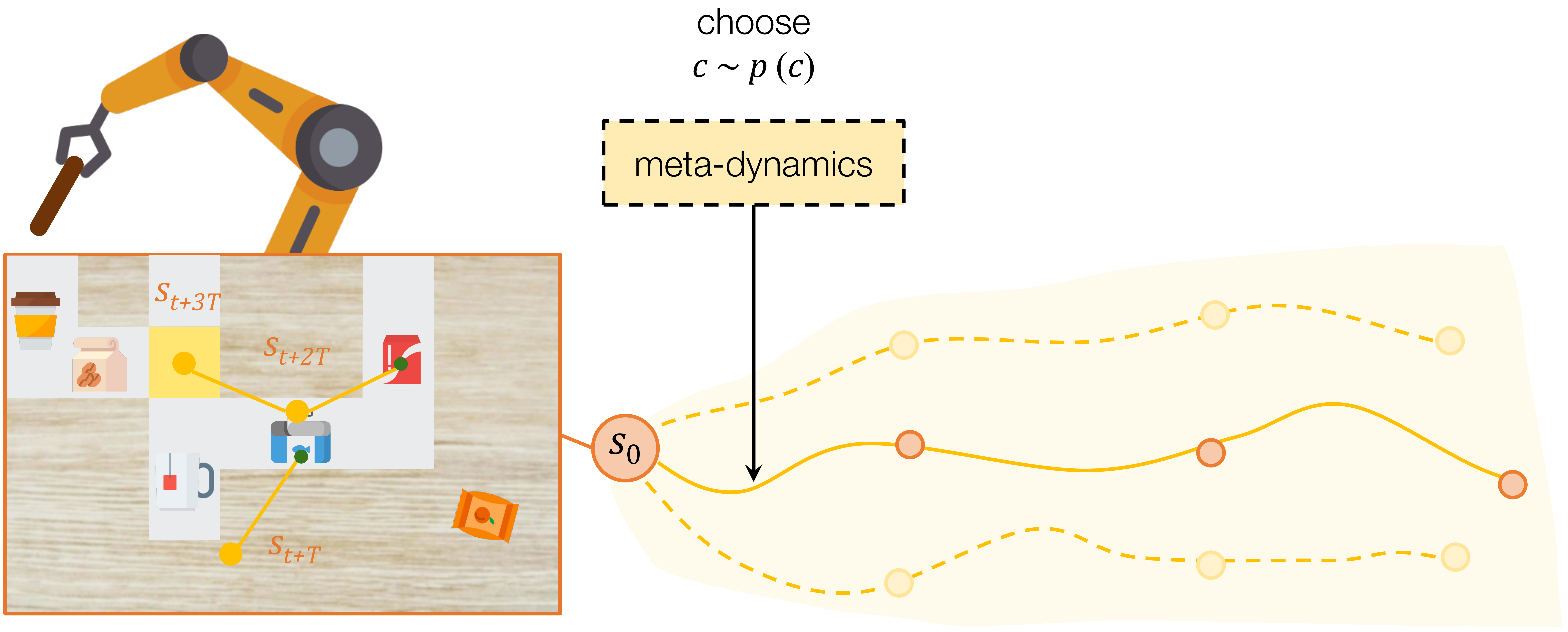
Fang et al. “CAVIN” CoRL’19

Long-Horizon Tasks: Hierarchical Planning in Latent Spaces



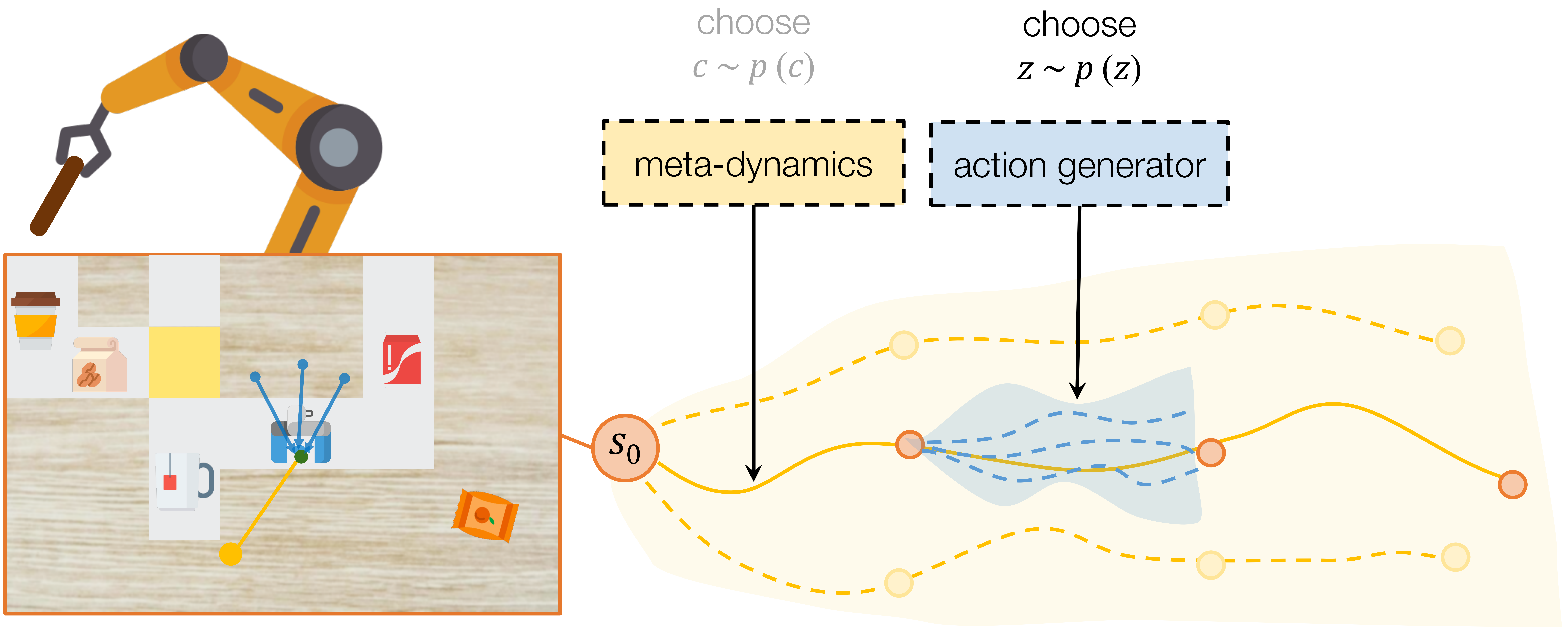
Fang et al. “CAVIN” CoRL’19

Long-Horizon Tasks: Hierarchical Planning in Latent Spaces



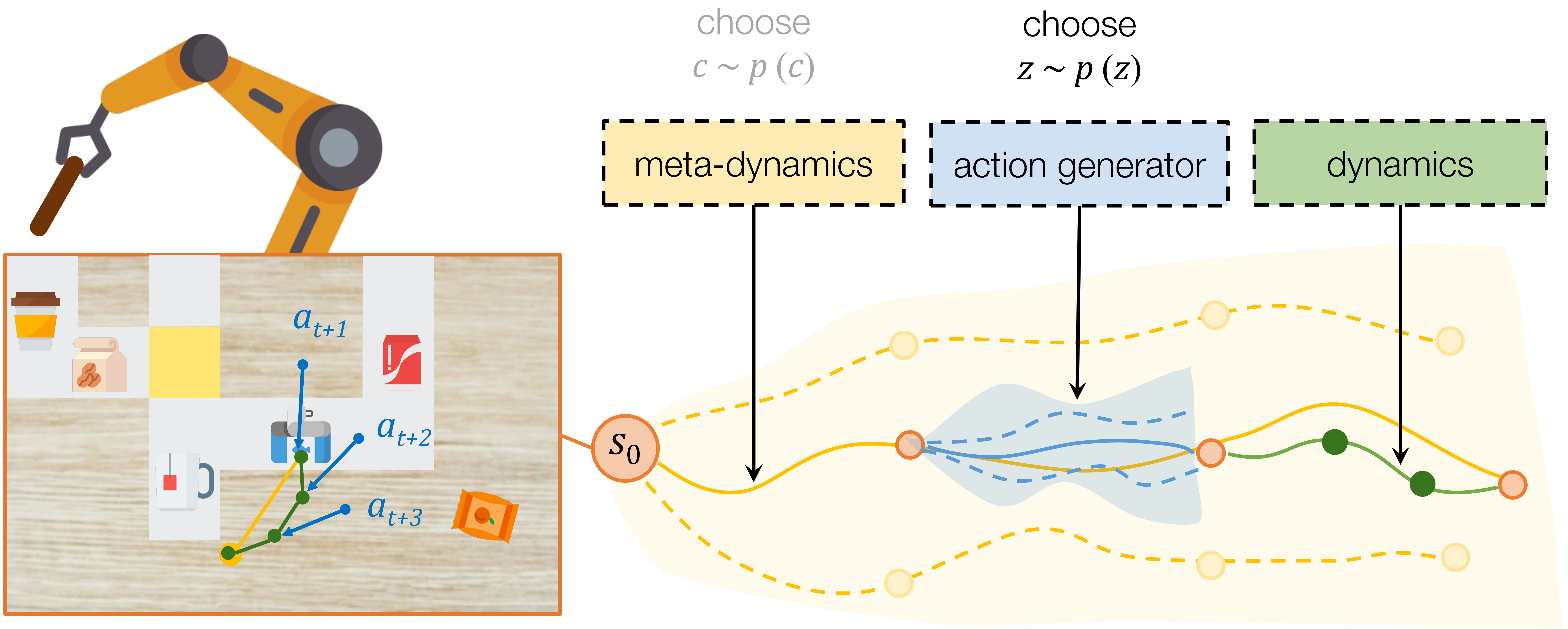
Fang et al. "CAVIN" CoRL'19

Long-Horizon Tasks: Hierarchical Planning in Latent Spaces



Fang et al. "CAVIN" CoRL'19

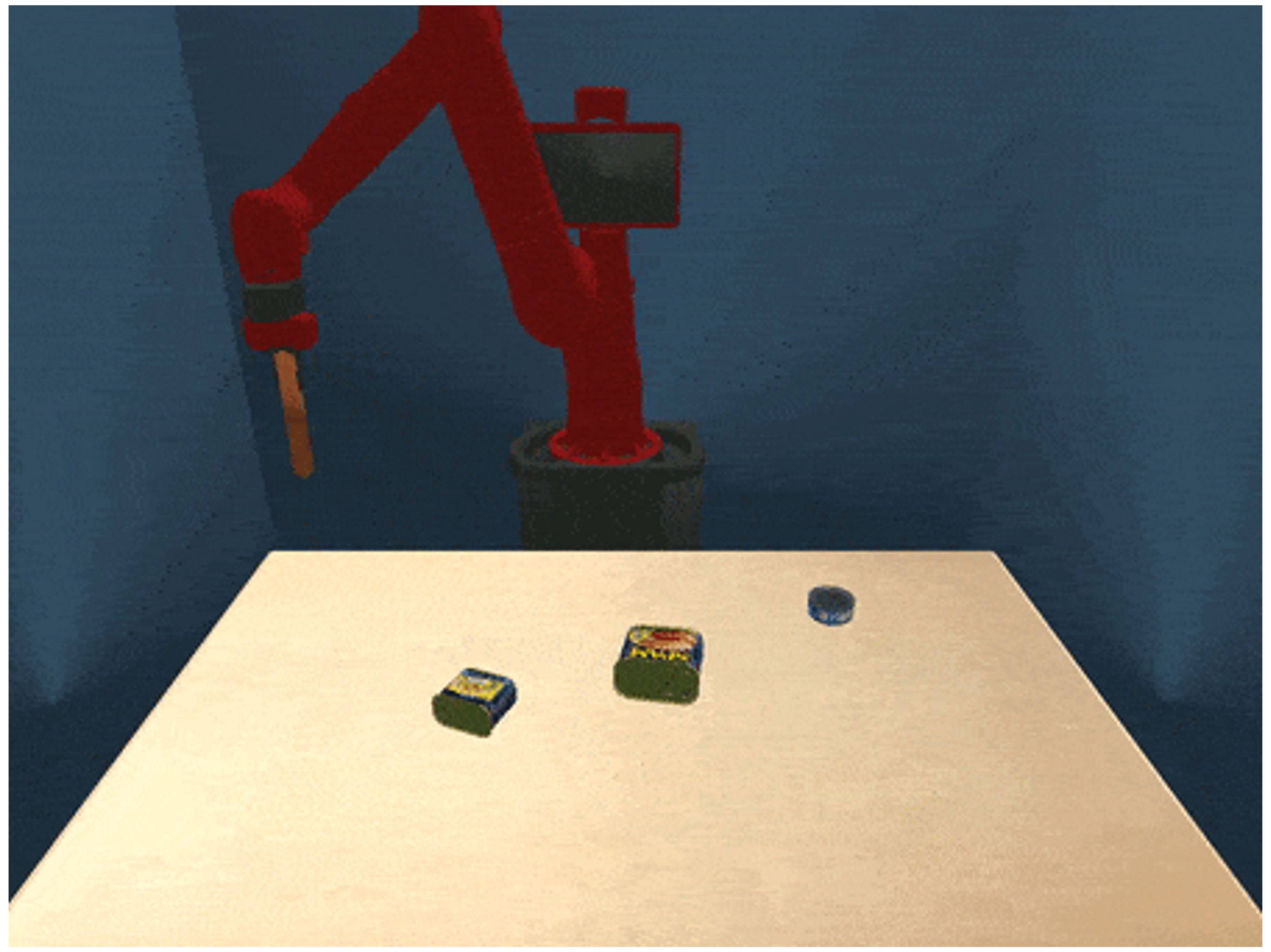
Long-Horizon Tasks: Hierarchical Planning in Latent Spaces

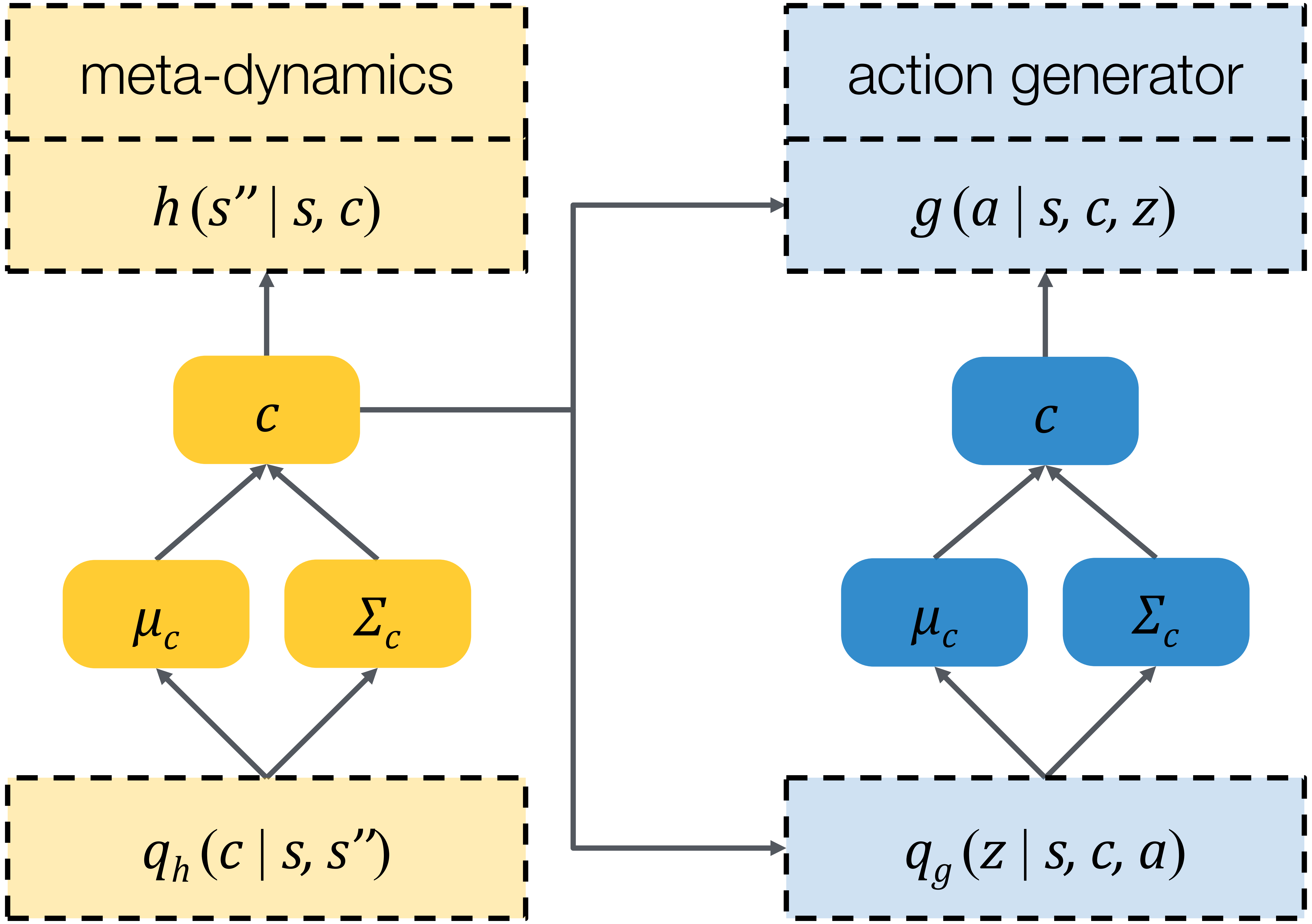


Fang et al. “CAVIN” CoRL’19

Long-Horizon Tasks: Hierarchical Planning in Latent Spaces

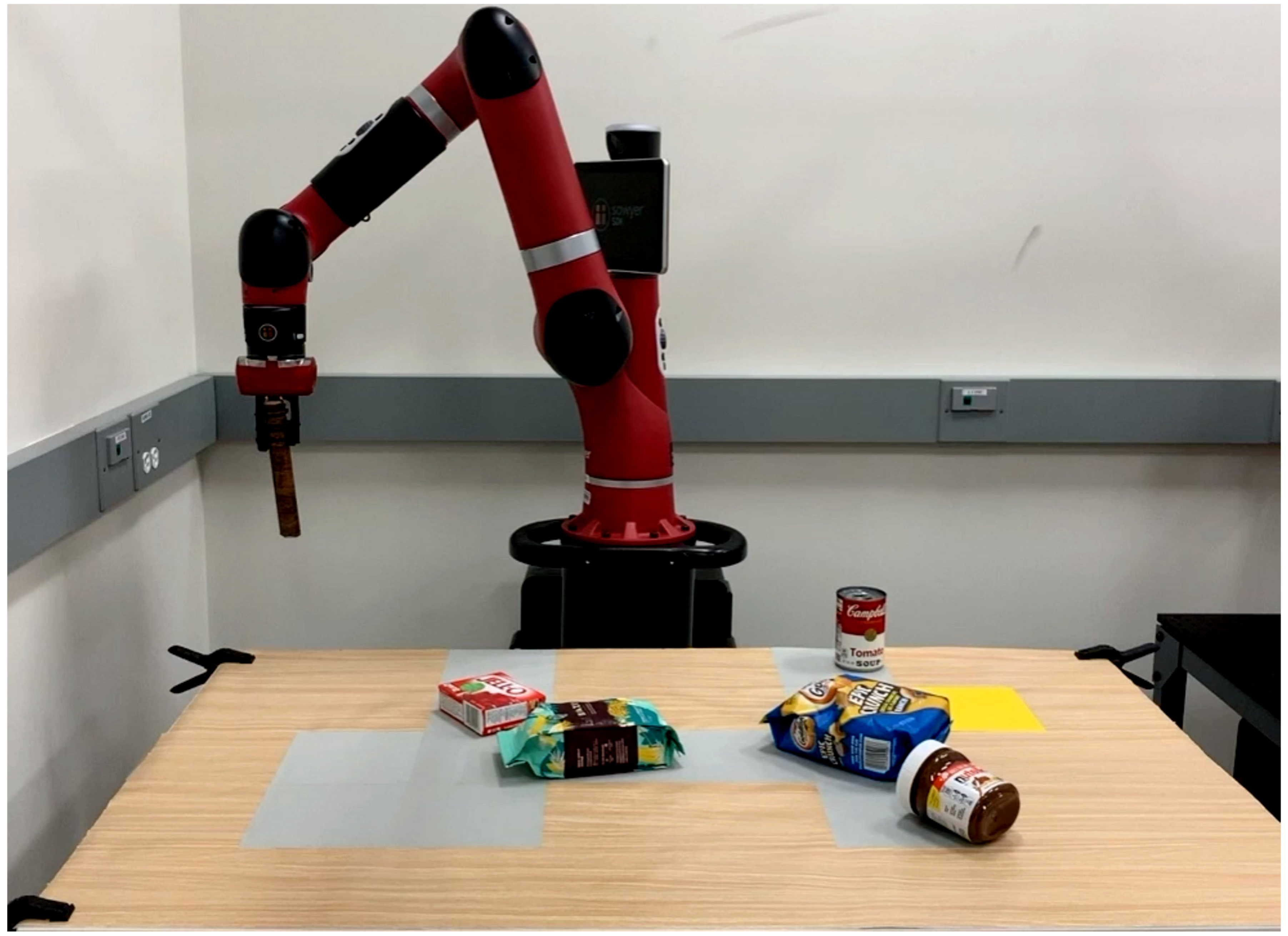
task-agnostic interaction



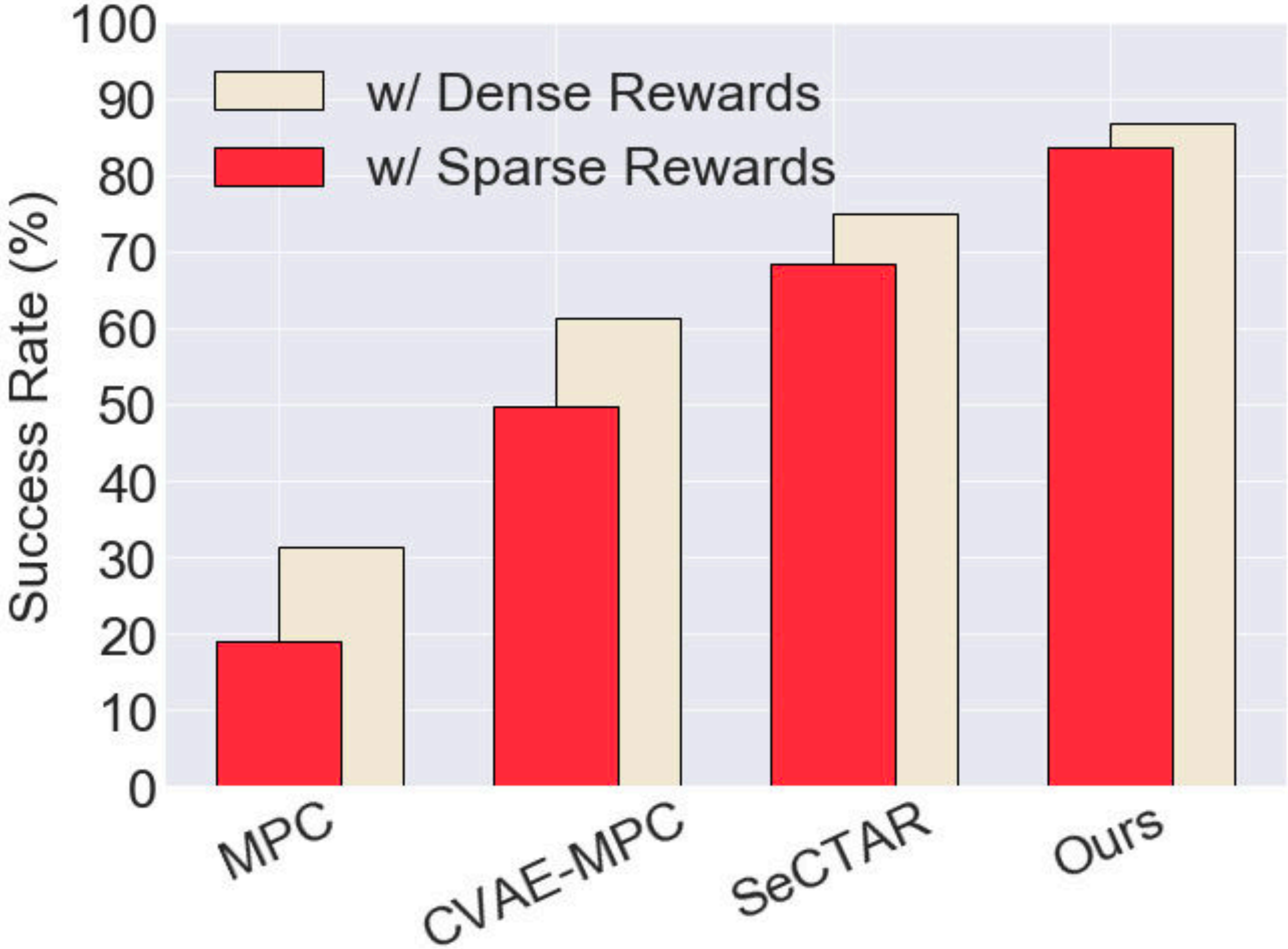


Fang et al. “CAVIN” CoRL’19

Long-Horizon Tasks: Cascaded Variational Inference



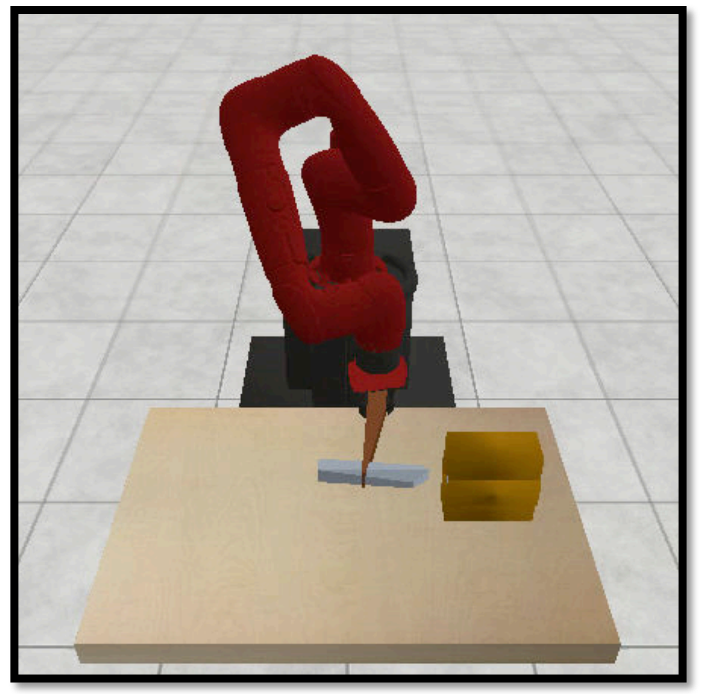
“move away obstacles”



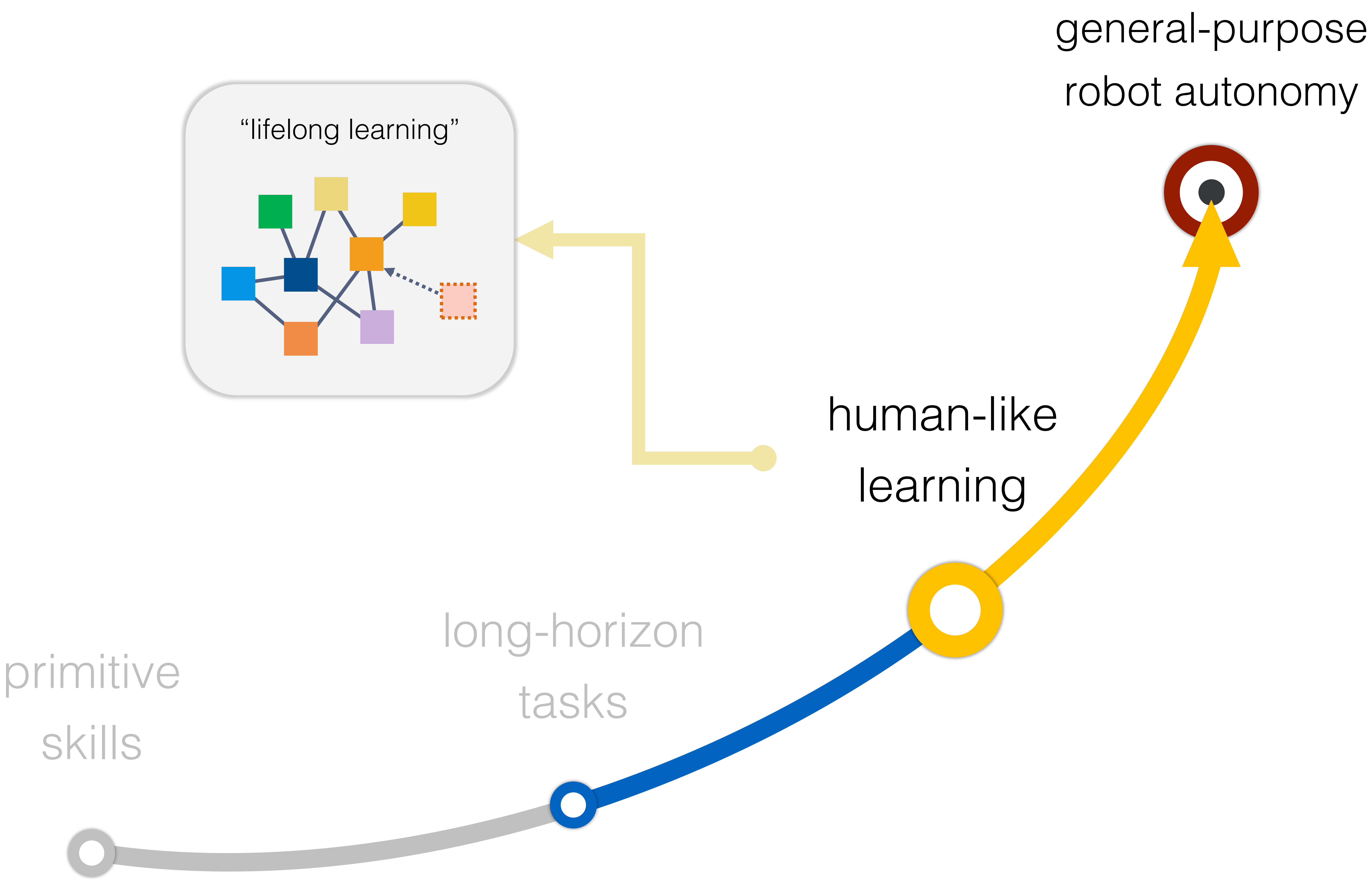
Fang et al. “CAVIN” CoRL’19

Summary - Part II

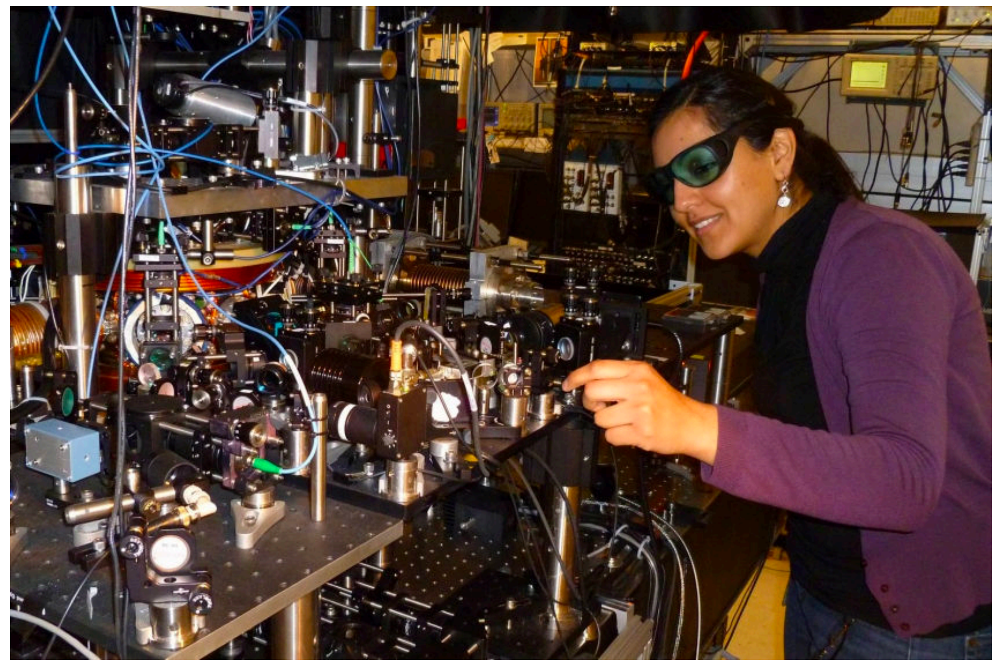
Hierarchical planning and **symbolic abstraction** scale up to long-horizon manipulation tasks.



High-level plans and **low-level skills** can be learned jointly from task-agnostic interactions.

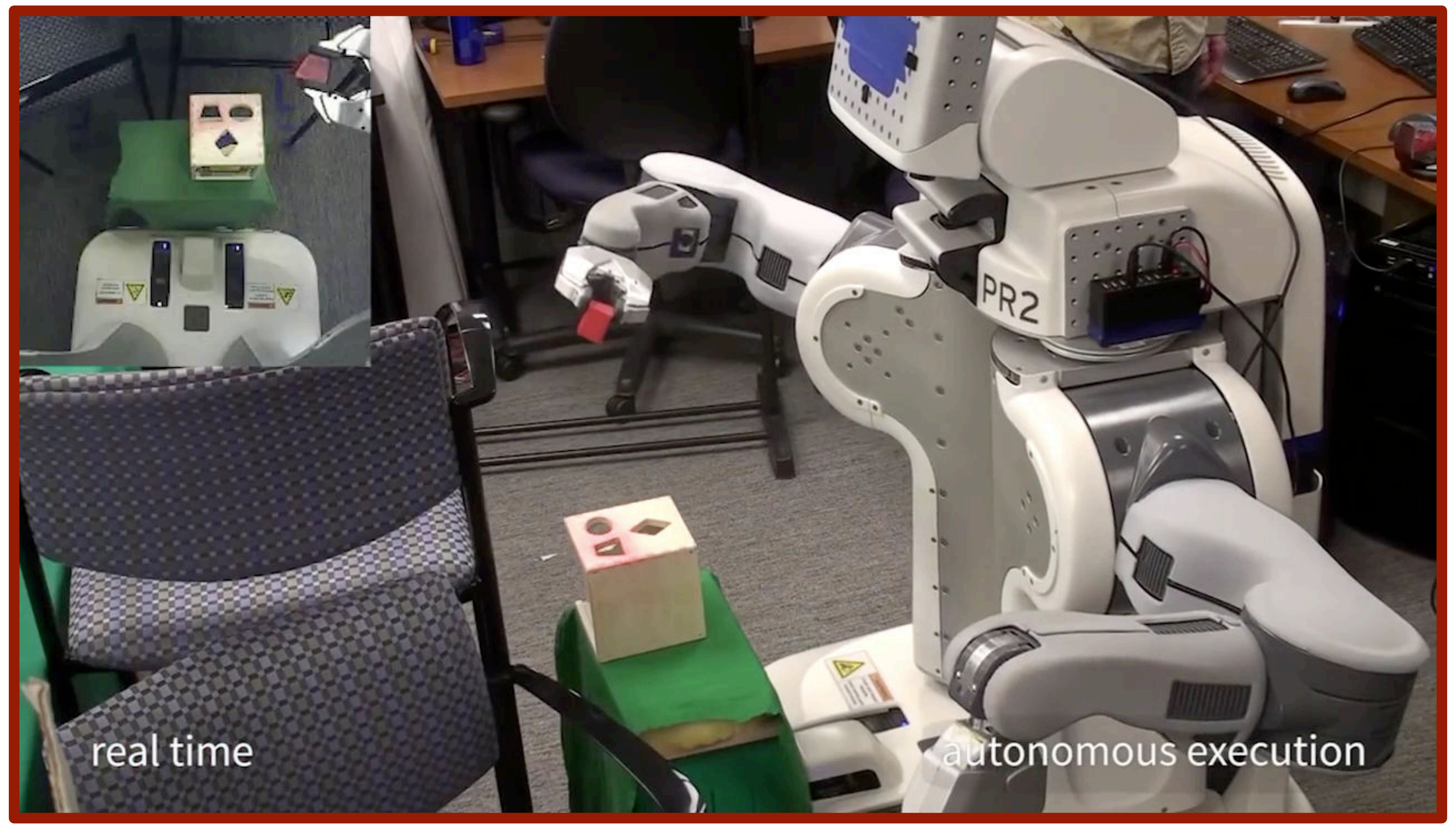

general-purpose
robot autonomy
“lifelong learning”
human-like
learning
long-horizon
tasks
primitive
skills

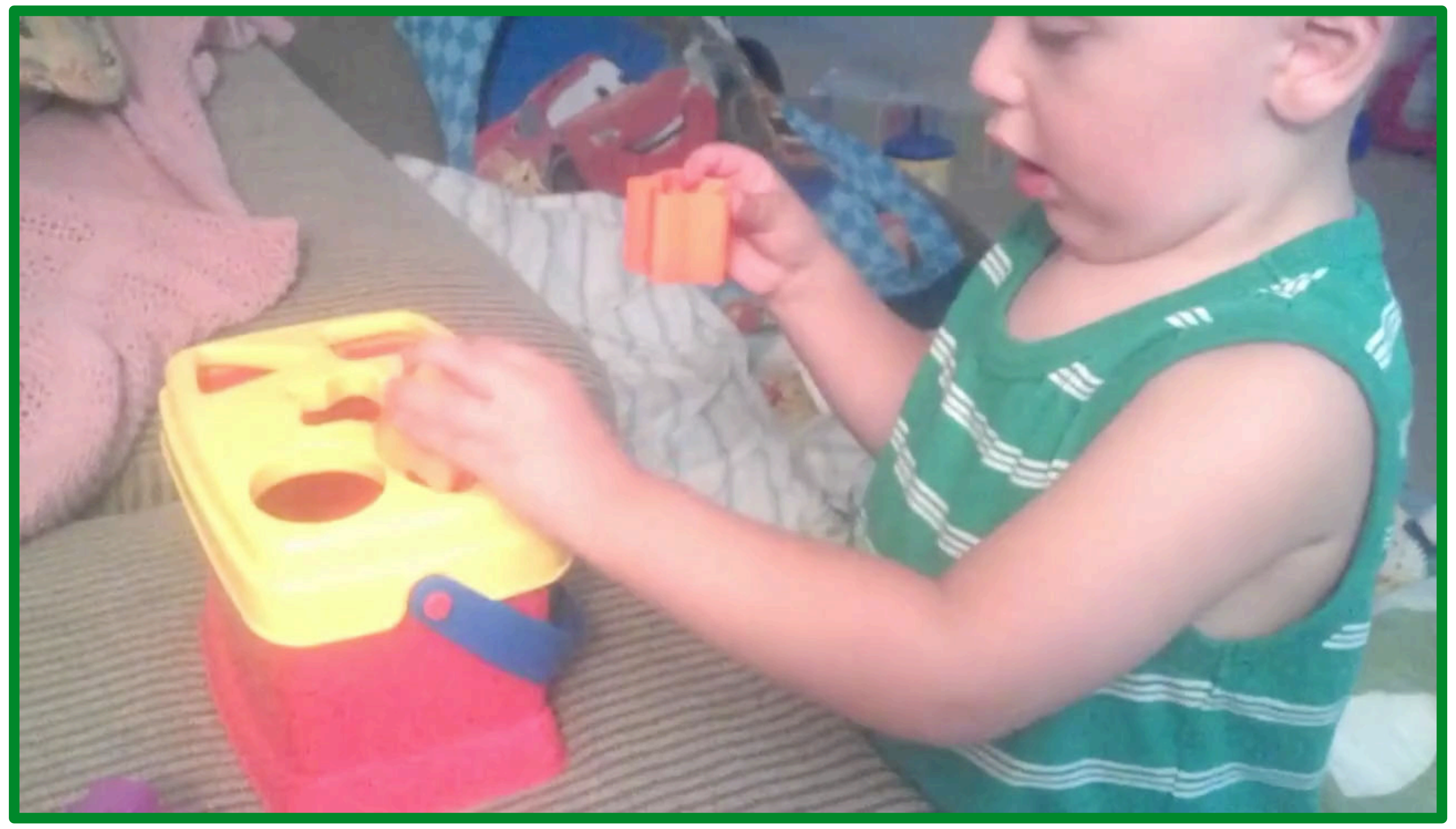
Human-Like Learning: A Lifelong Process



Learning as a **lifelong process** of **active exploration** and **model building**

Human-Like Learning: Harvesting Human Ingenuity



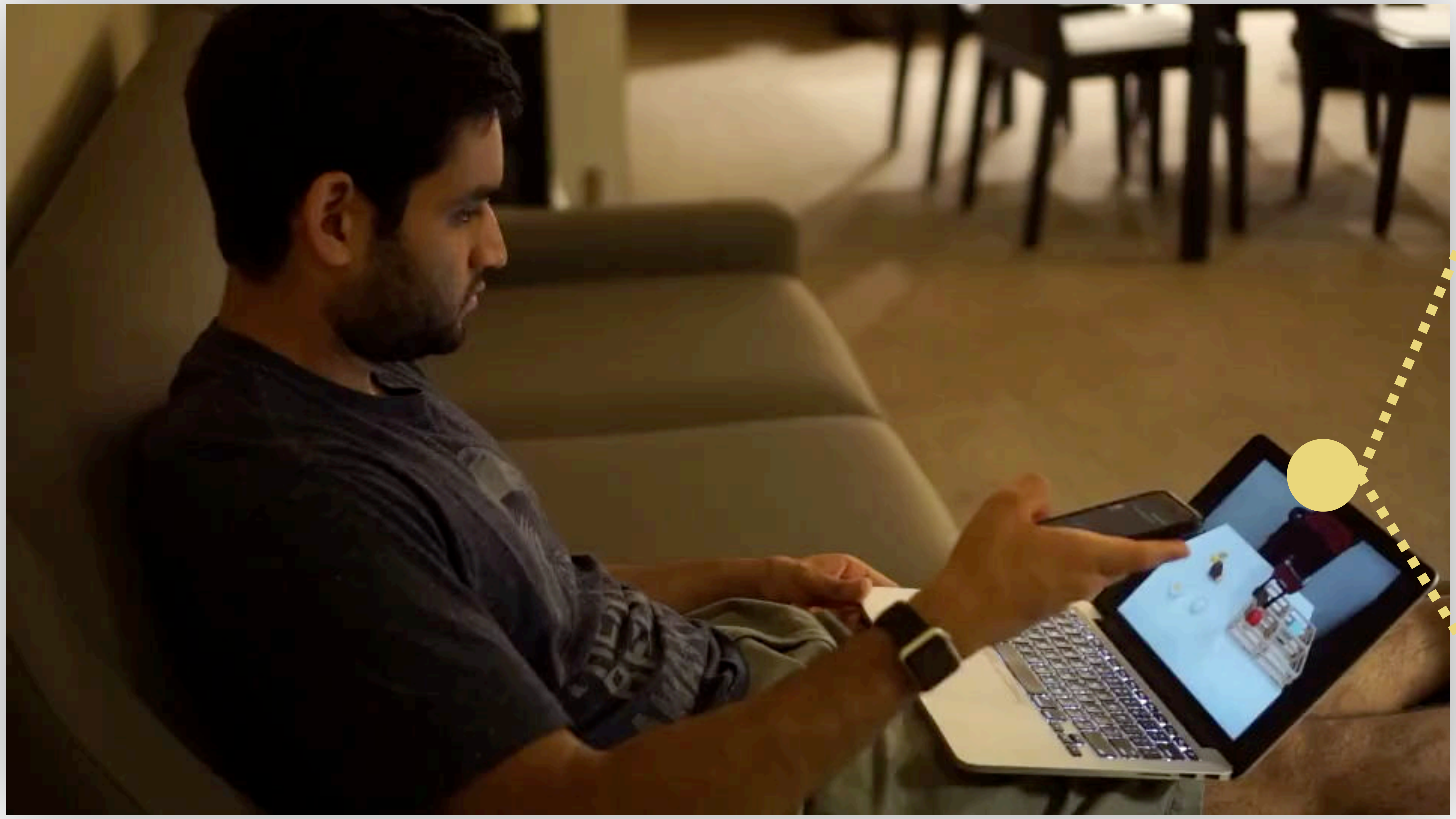


- ✗ Narrow-minded
- ✗ Limited object manipulation

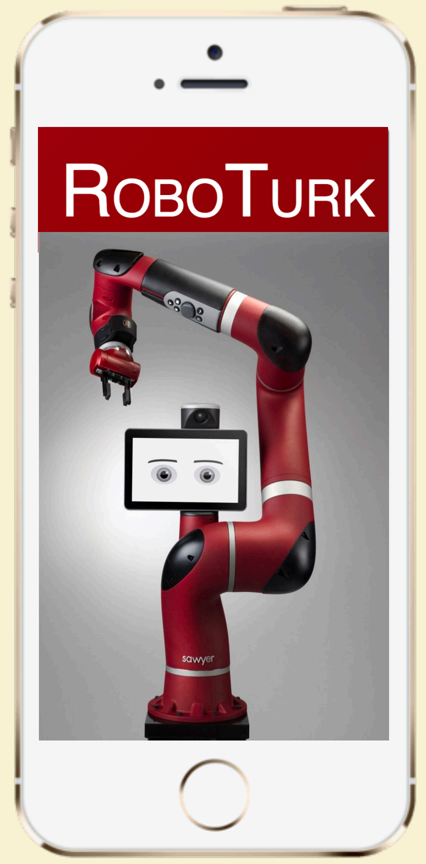
- ✓ Creative solutions
- ✓ Rich object manipulation

Human-Like Learning: Harvesting Human Ingenuity

RoboTurk: Crowdsourcing Platform for Large-Scale Teleoperation

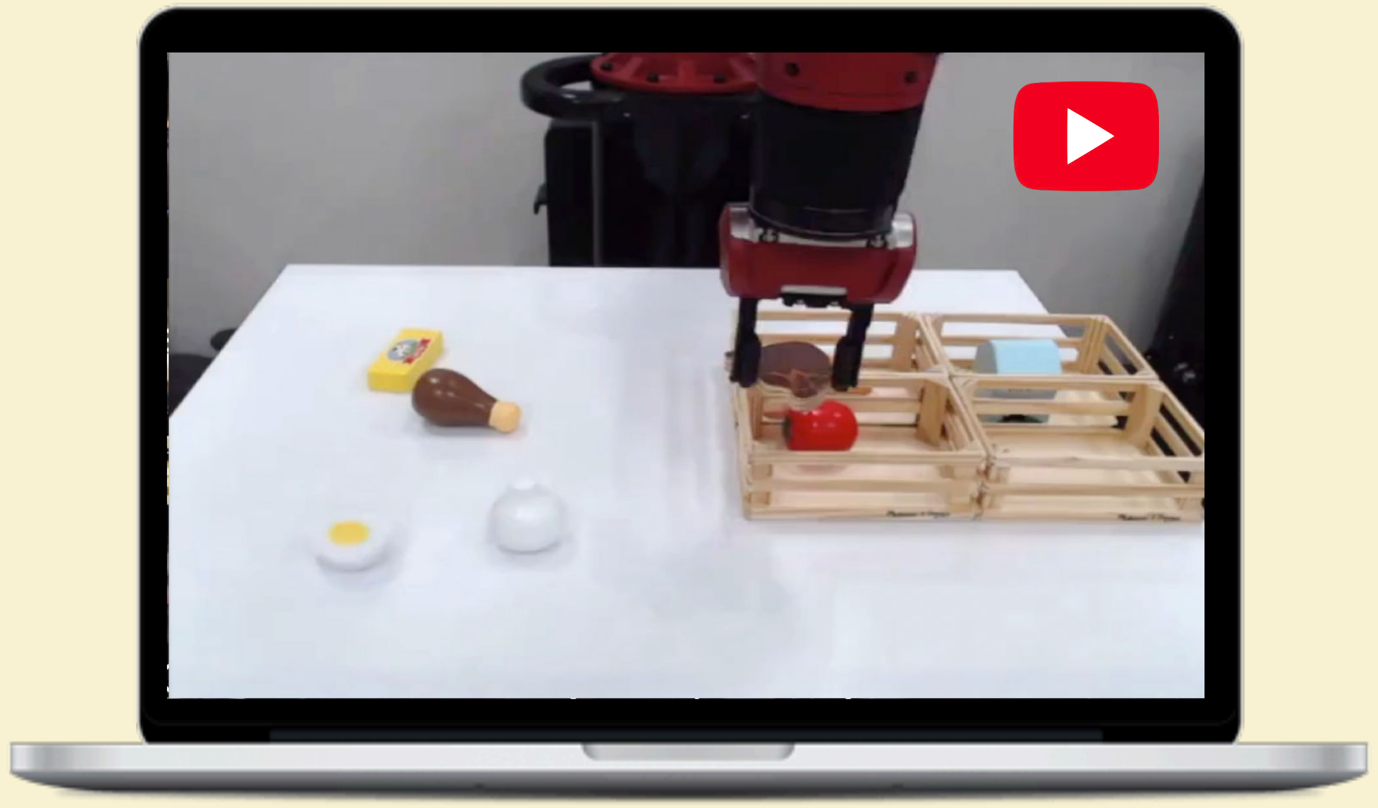


RoboTurk in action



6-DoF controller

+



real-time streaming from remote robot

roboturk.stanford.edu

Mandlekar et al. CoRL 2018
Mandlekar et al. IROS 2019


Kinect
Robot
Webcam
RoboTurk Systems

Human Ingenuity in Solution Strategies

Human Ingenuity in Solution Strategies

Emergent Strategies: Alternate Cups and Bowls

Emergent Strategies: Flip Bowl for Base

Emergent Strategies: 3 Cups for a Stable Platform

RoboTurk
Real Robot Dataset

111 hours of robot demonstrations

1 week of data collection

3 dexterous manipulation tasks

54 non-expert users

2144 demonstrations

10x larger than prior work

roboturk.stanford.edu/realrobotdataset

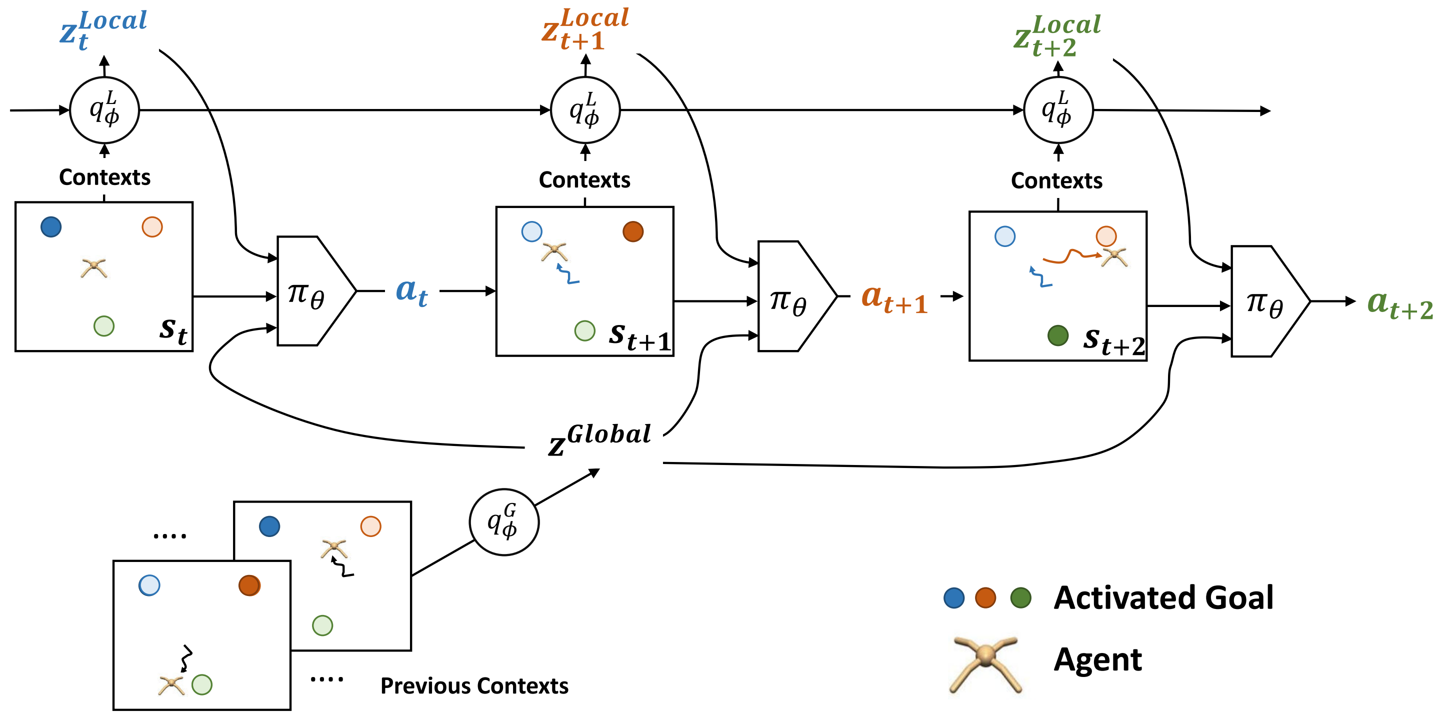
Mandlekar et al. IROS 2019

Human-Like Learning: Three Key Ingredients

A **human-like learning agent** will

Learning to Learn

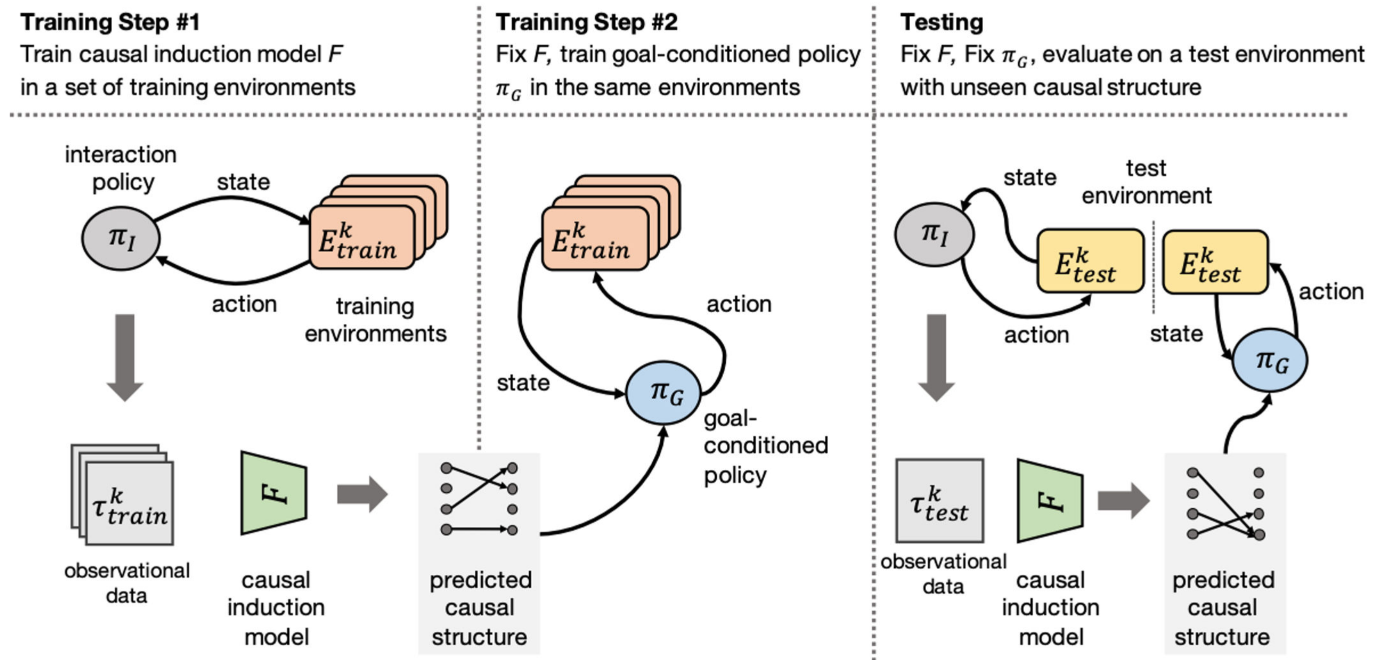
Re-use prior knowledge to learn and adapt fast



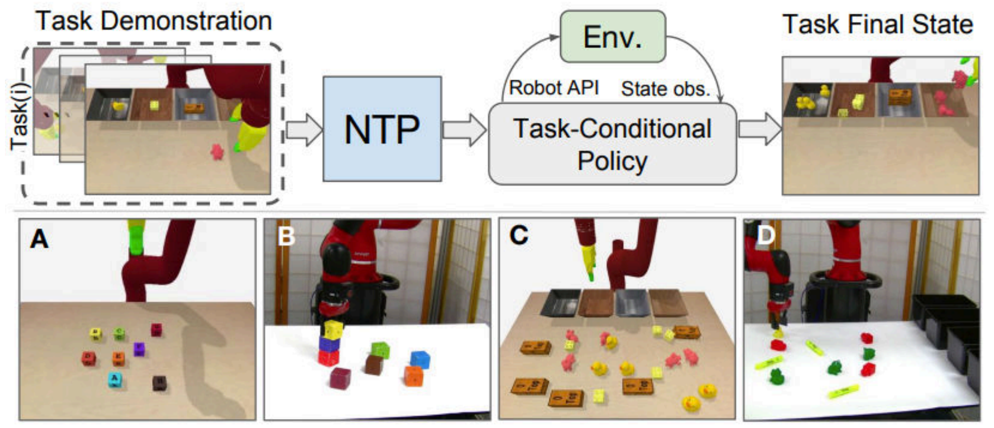
[Ren et al. UAI'20 (to appear)]

Causal Understanding

Reason about causal and effect from interaction for interaction



[Nair et al. NeurIPS'19 CausalML]



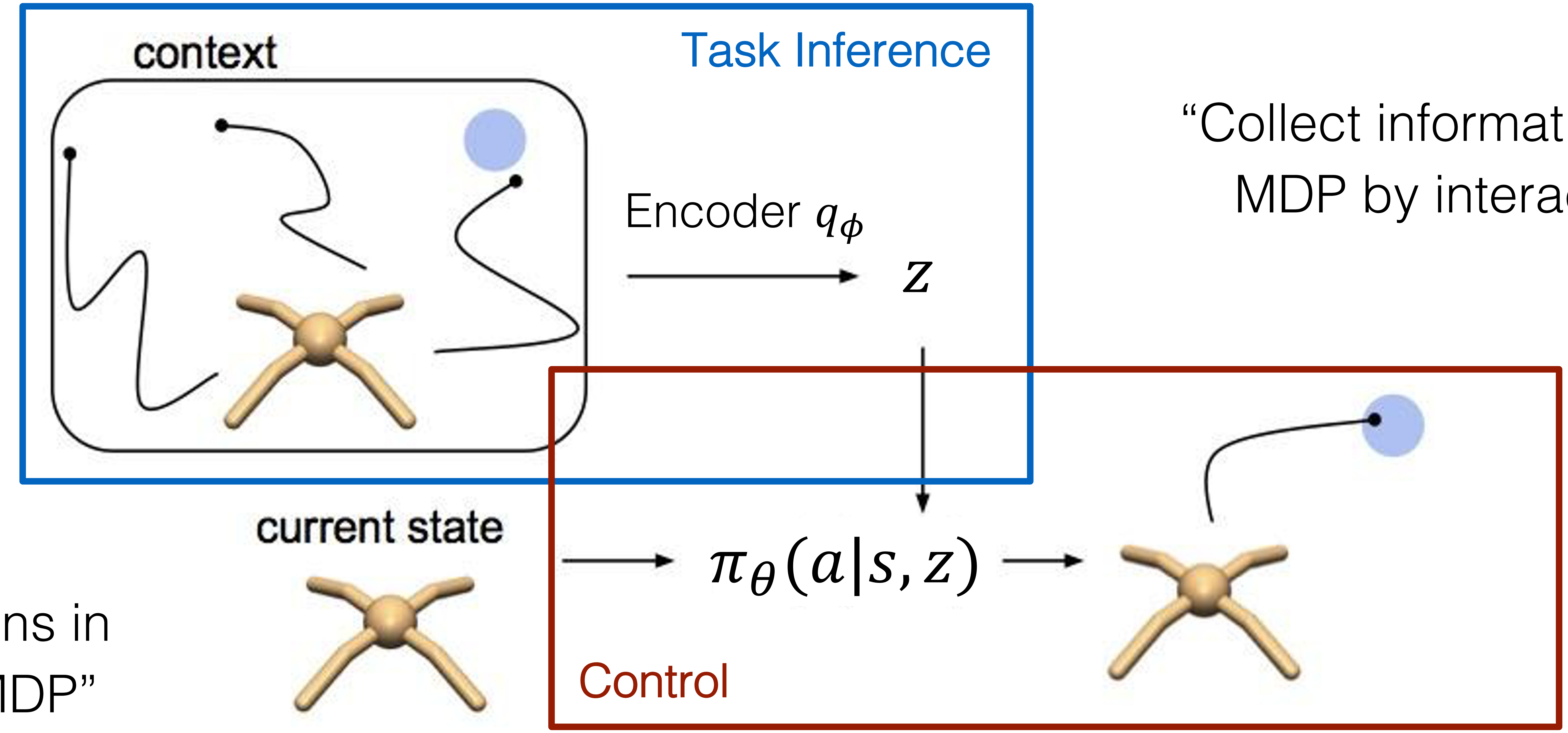
[Xu*, Nair*, et al. ICRA'18; Huang*, Nair*, Xu*, et al. CVPR'19]

Compositionality

Capture the compositional structure of semantics and tasks

Human-Like Learning: Learning to Learn

Meta-reinforcement learning through online task inference



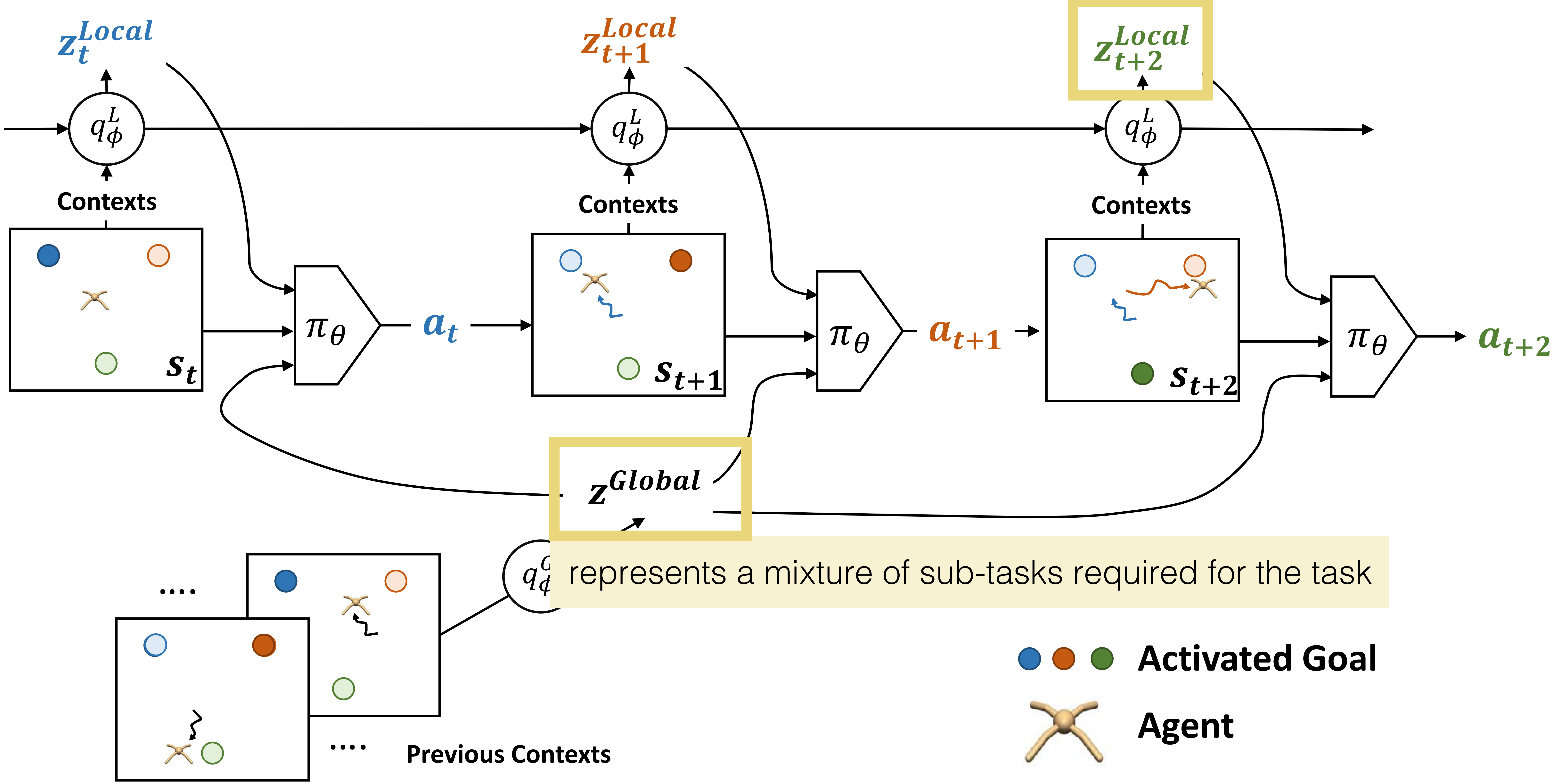
"Collect information from MDP by interaction"

"Make decisions in the inferred MDP"

Ren et al. "OCEAN" UAI'20

Human-Like Learning: Learning to Learn

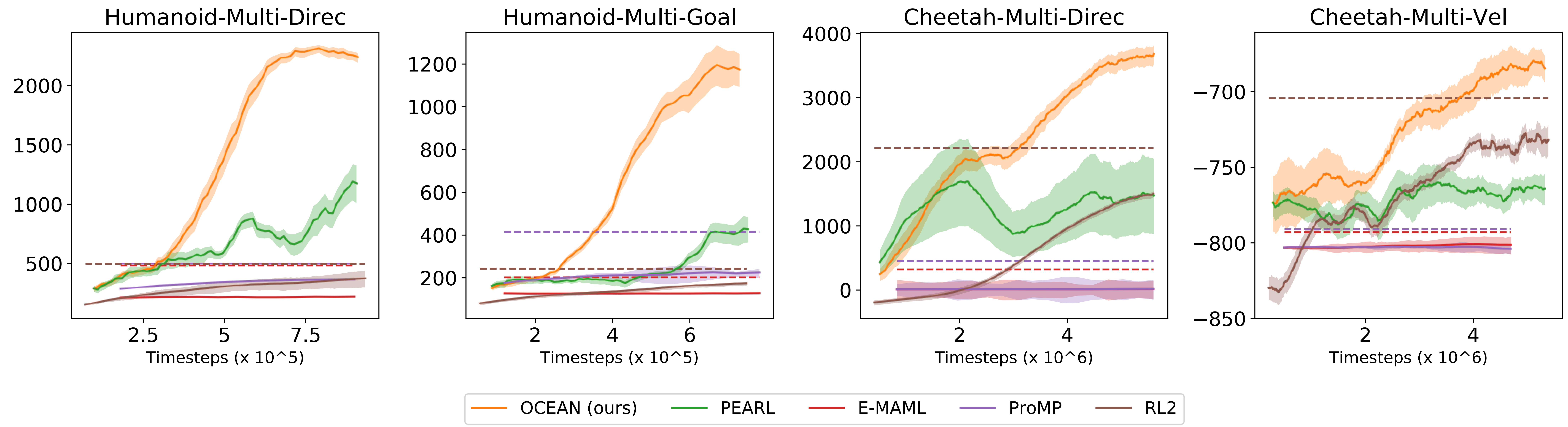
captures transitions between the subtasks



represents a mixture of sub-tasks required for the task

Ren et al. "OCEAN" UAI'20

Human-Like Learning: Learning to Learn

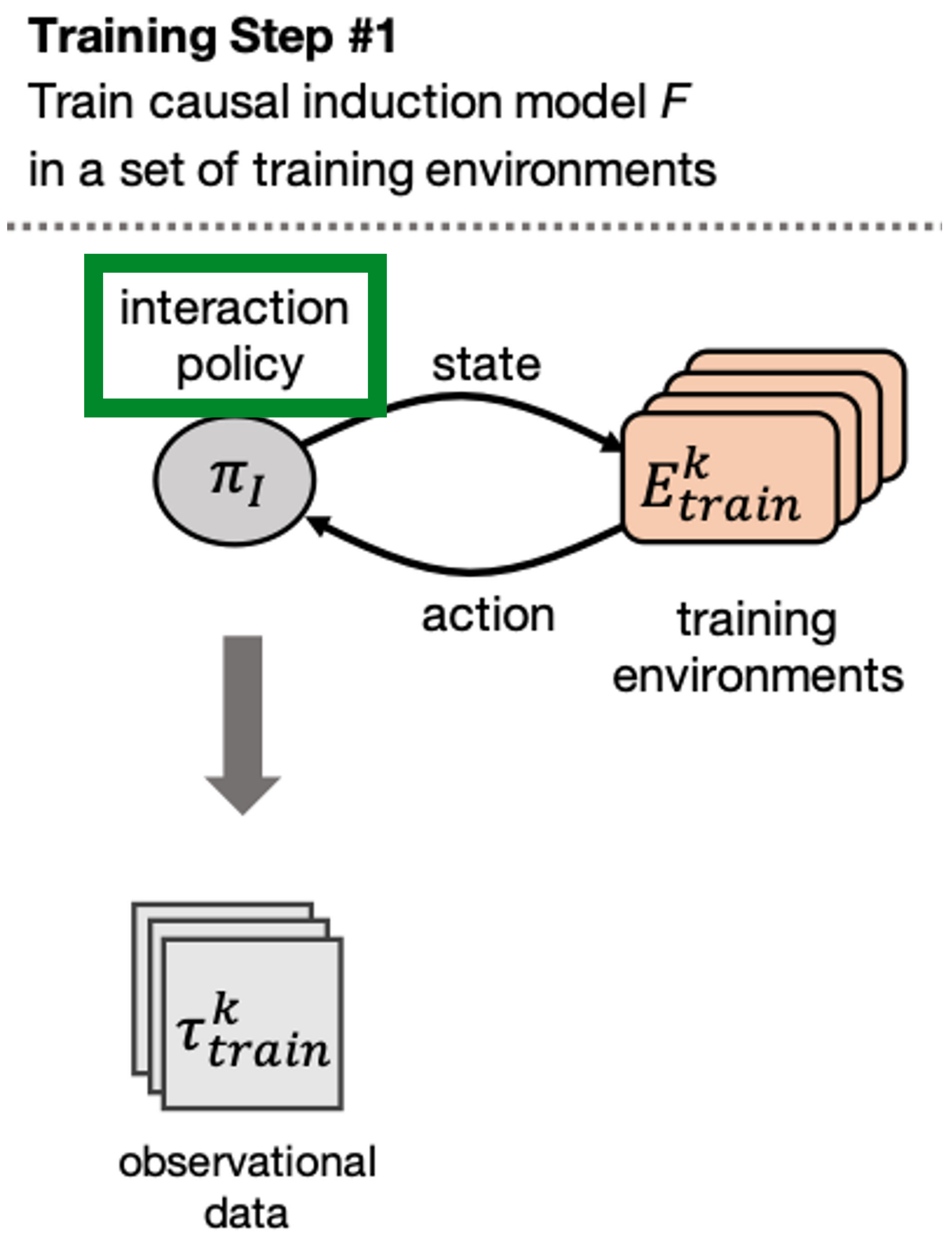


OCEAN is especially effective in **long-horizon tasks** that involve a sequence of primitive skills.

Ren et al. "OCEAN" UAI'20

Human-Like Learning: Causal Understanding

Learning causal models from interaction for goal-directed tasks in visual environments

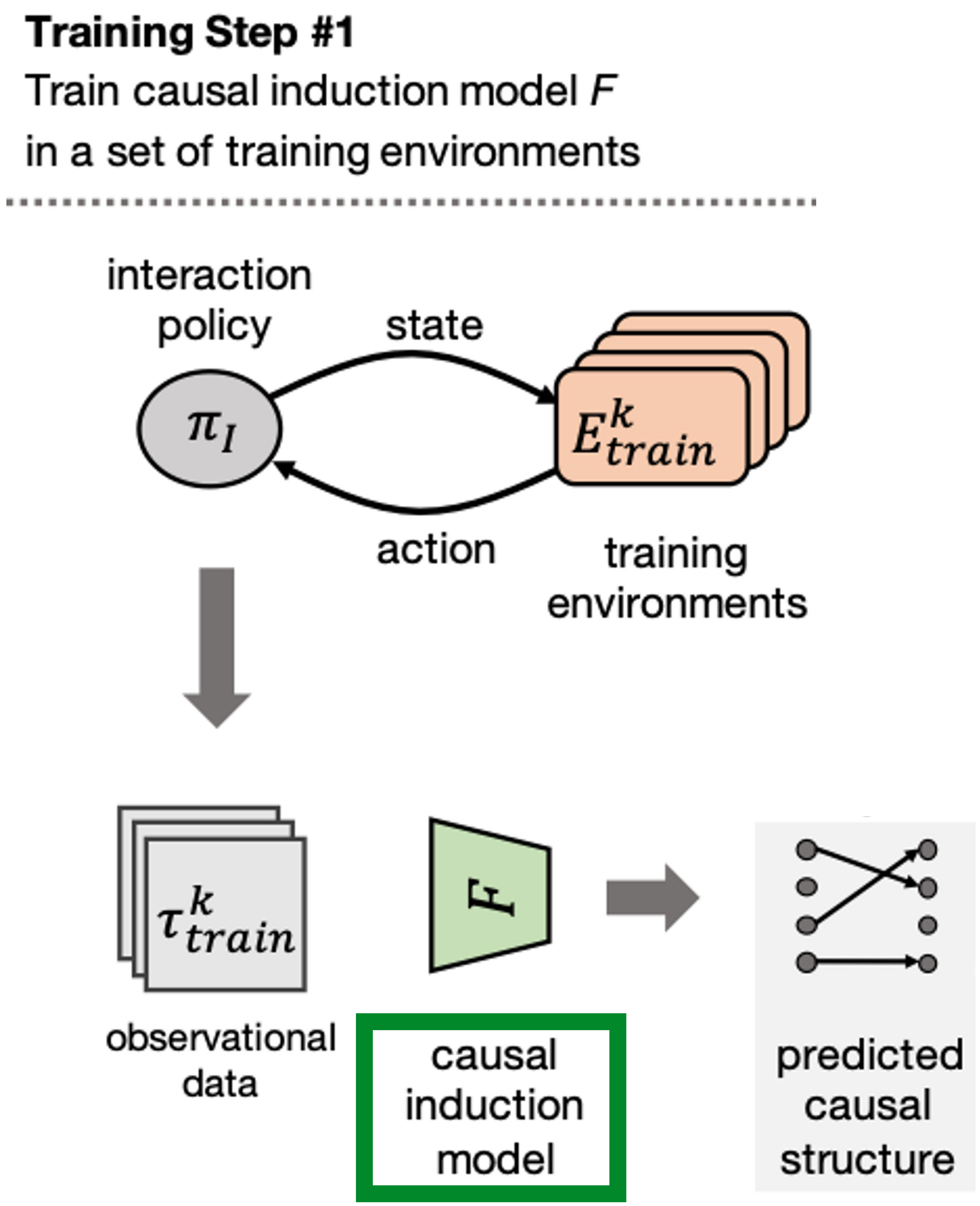


Interaction policy π_I collects observational data in environment

Nair, Zhu, et al., NeurIPS 2019 Causal ML Workshop

Human-Like Learning: Causal Understanding

Learning causal models from interaction for goal-directed tasks in visual environments

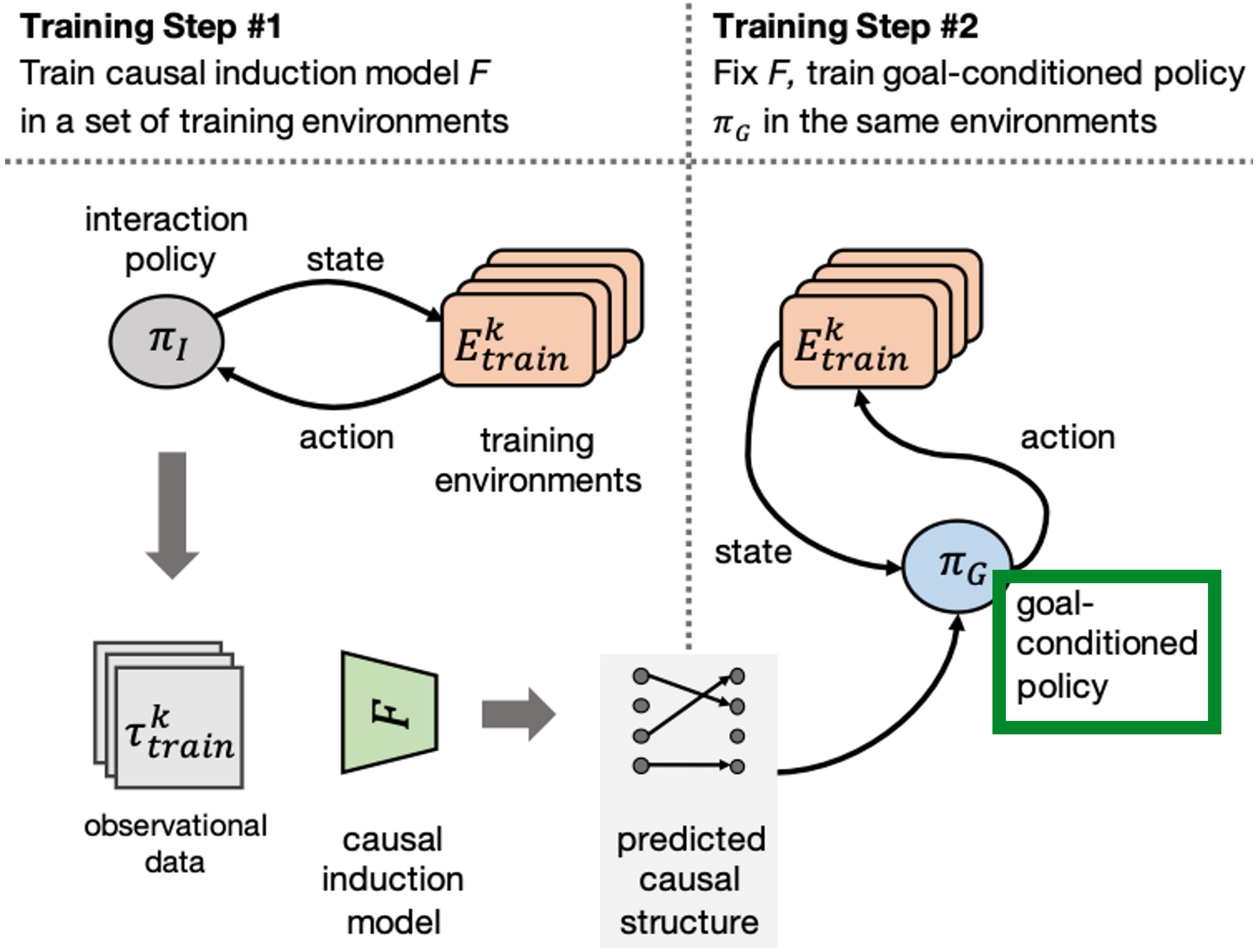


Causal induction model F predicts causal graph from observational data.

Nair, Zhu, et al., NeurIPS 2019 Causal ML Workshop

Human-Like Learning: Causal Understanding

Learning causal models from interaction for goal-directed tasks in visual environments

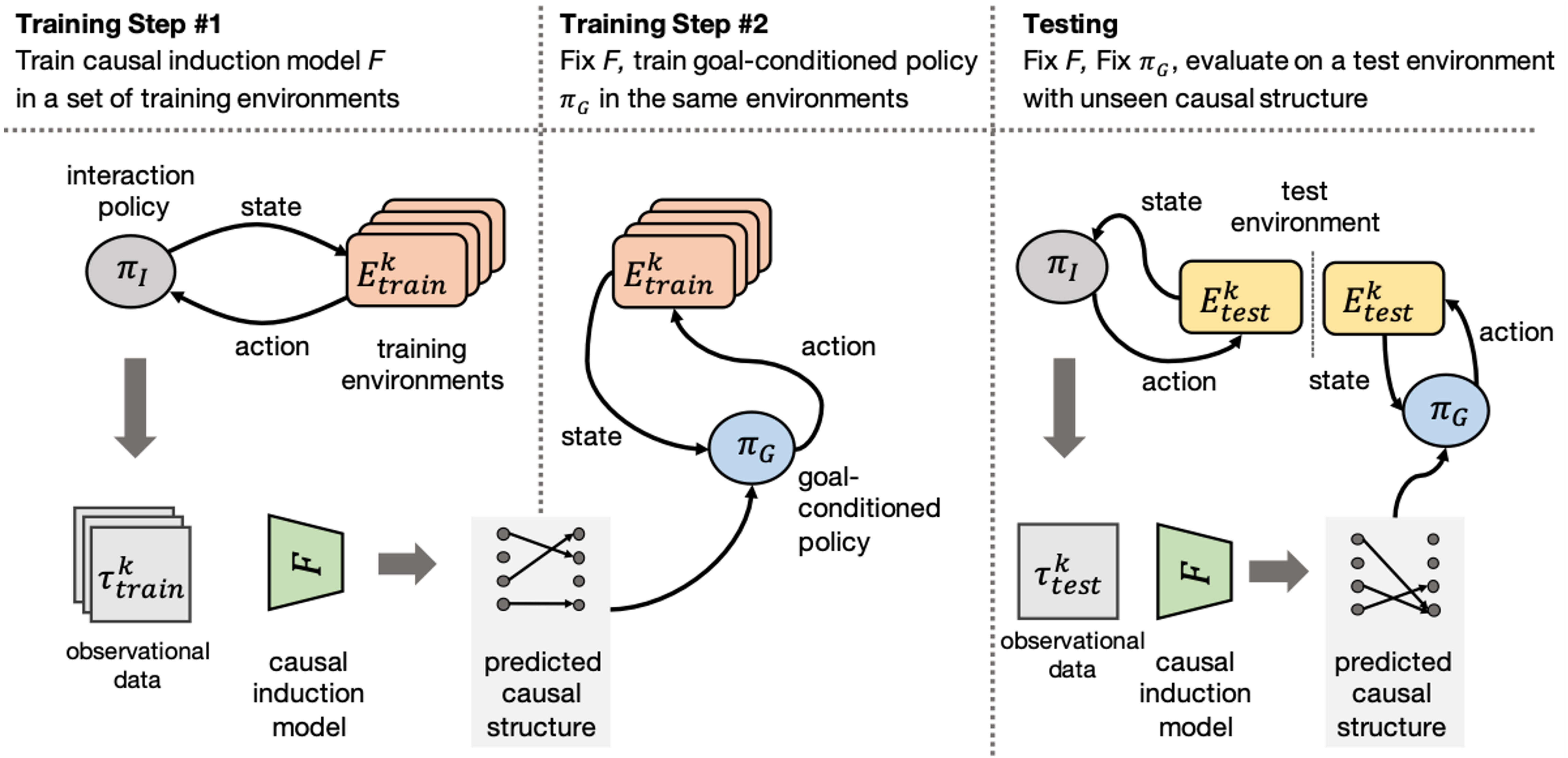


Conditioned on causal graph, **goal conditioned policy π_G** tries to complete the tasks in environment.

Nair, Zhu, et al., NeurIPS 2019 Causal ML Workshop

Human-Like Learning: Causal Understanding

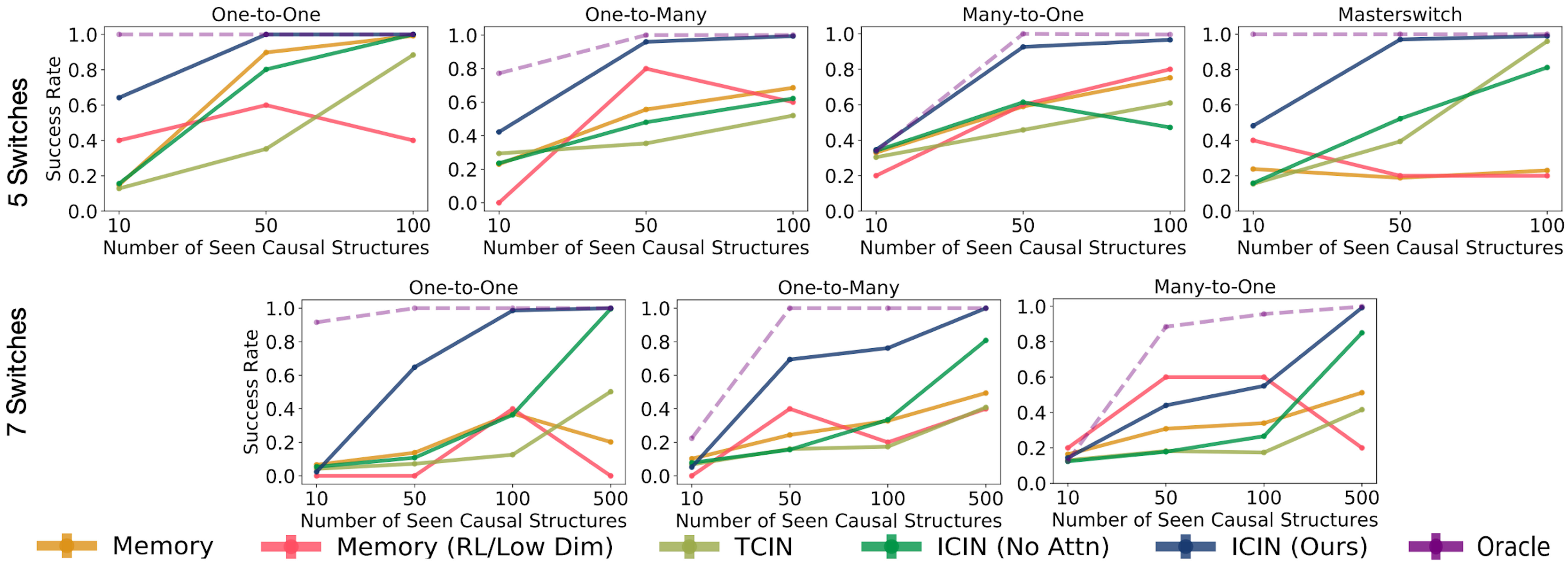
Learning causal models from interaction for goal-directed tasks in visual environments



Nair, Zhu, et al., NeurIPS 2019 Causal ML Workshop

Human-Like Learning: Causal Understanding

Learning causal models from interaction for goal-directed tasks in visual environments

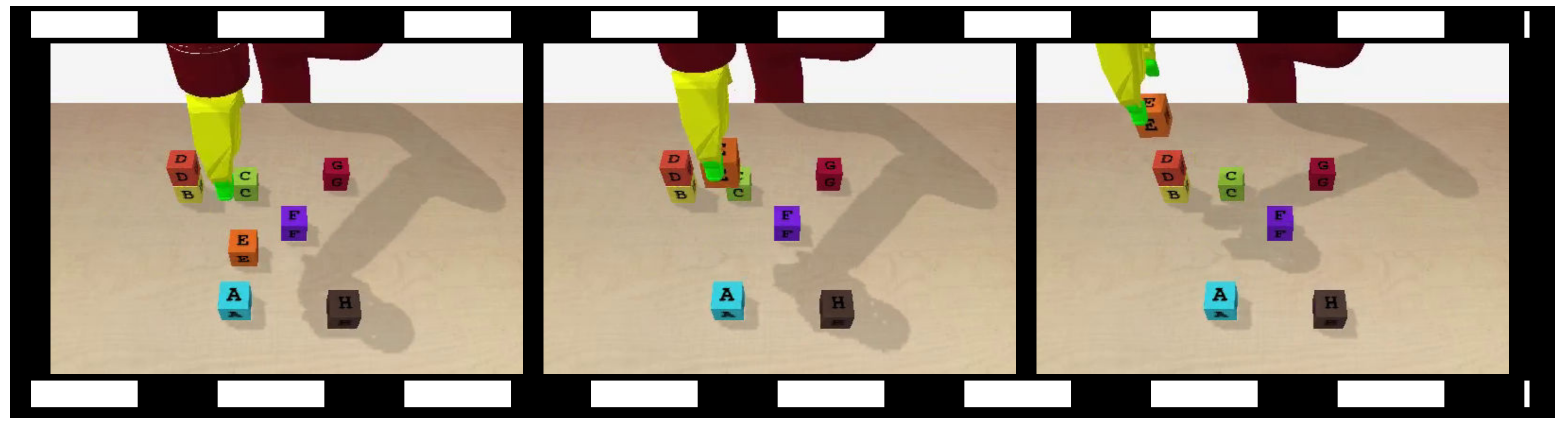


policy success rate in (unseen) light-switch environments

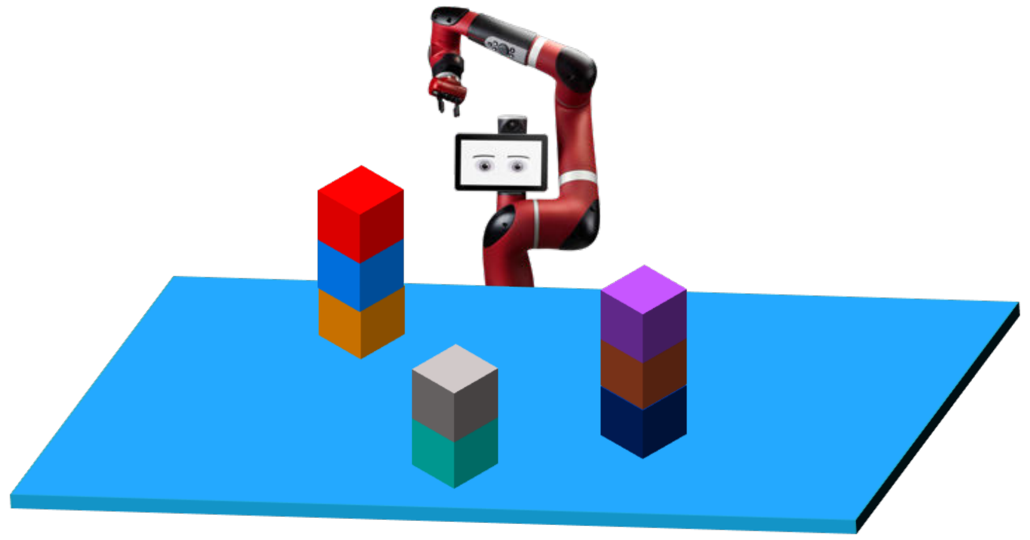
Nair, Zhu, et al., NeurIPS 2019 Causal ML Workshop

Human-Like Learning: Compositionality

Modeling complex tasks as compositional program structures



meta-learning model



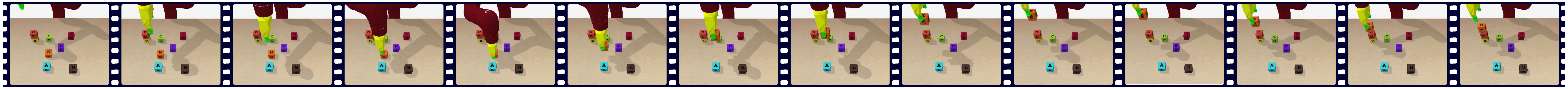
single video demonstration

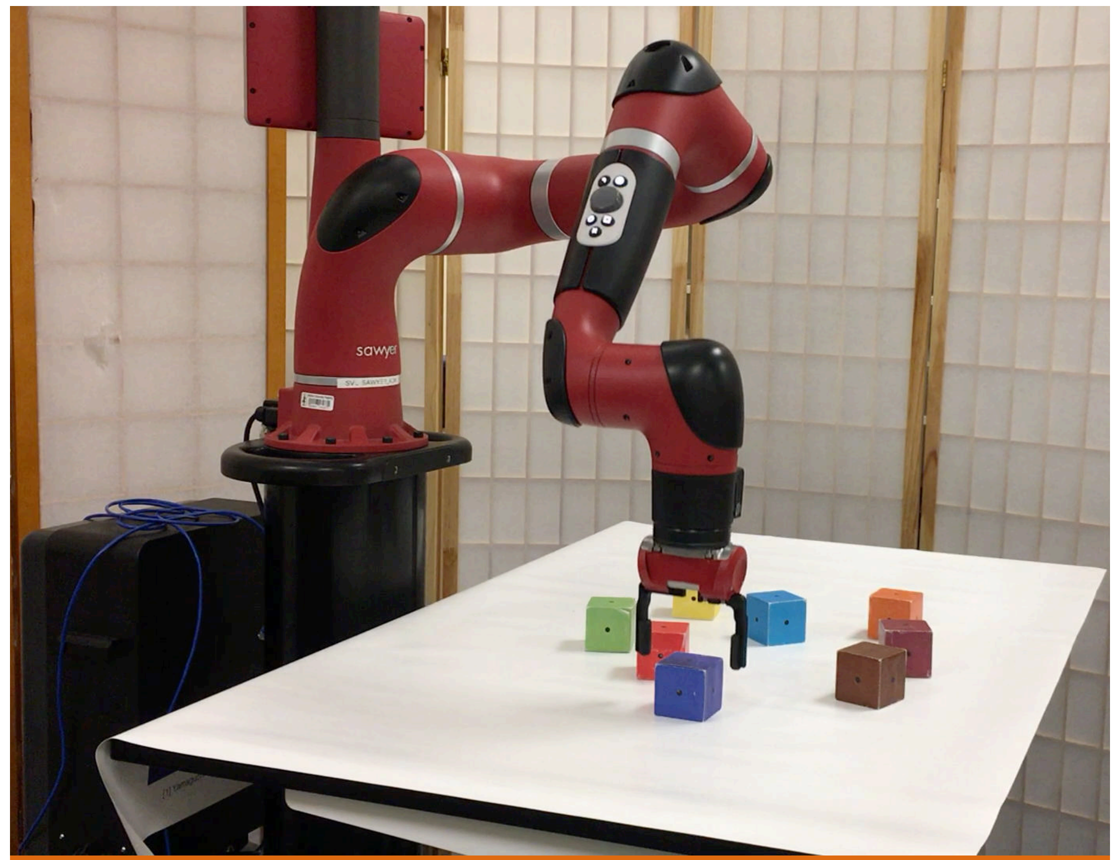
policy for the demonstrated task

Xu*, Nair*, et al. ICRA'18; Huang*, Nair*, Xu*, et al. CVPR'19; Huang et al. IROS'19

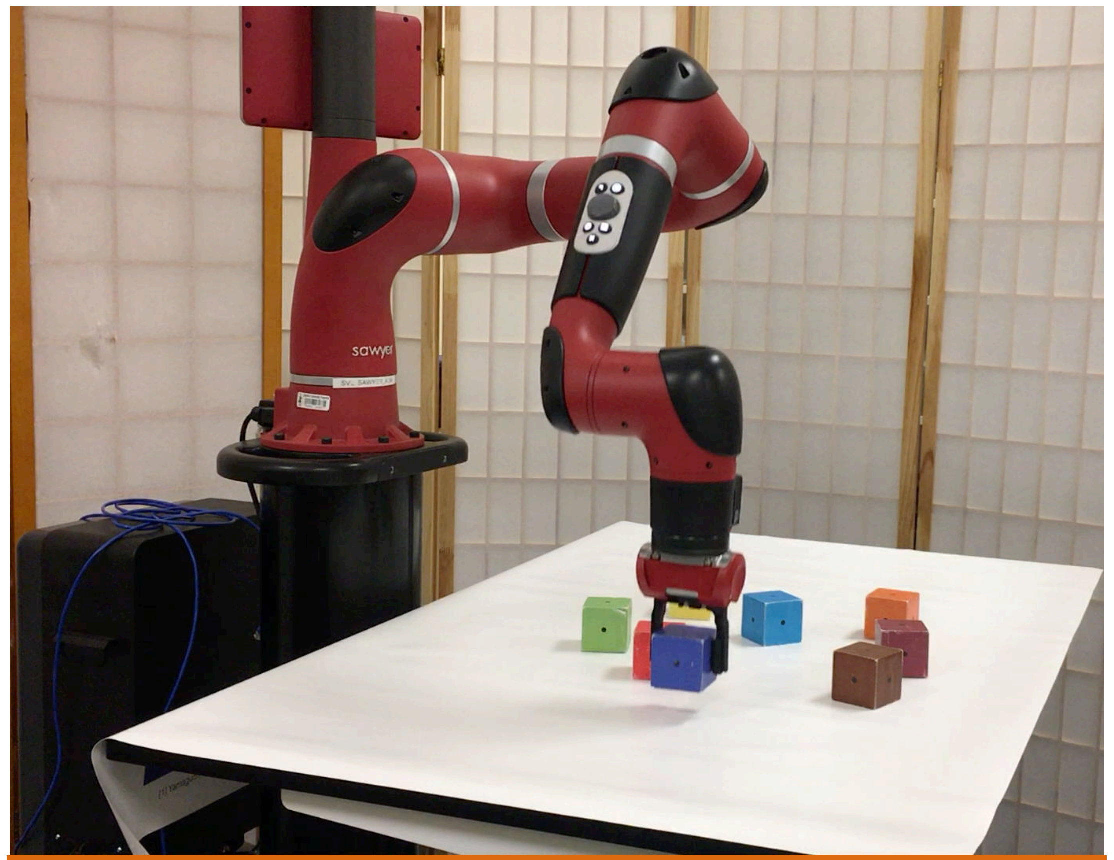
Xu*, Nair*, et al. ICRA 2018

Neural Task Programming (NTP): Hierarchical Policy Learning as Neural Program Induction

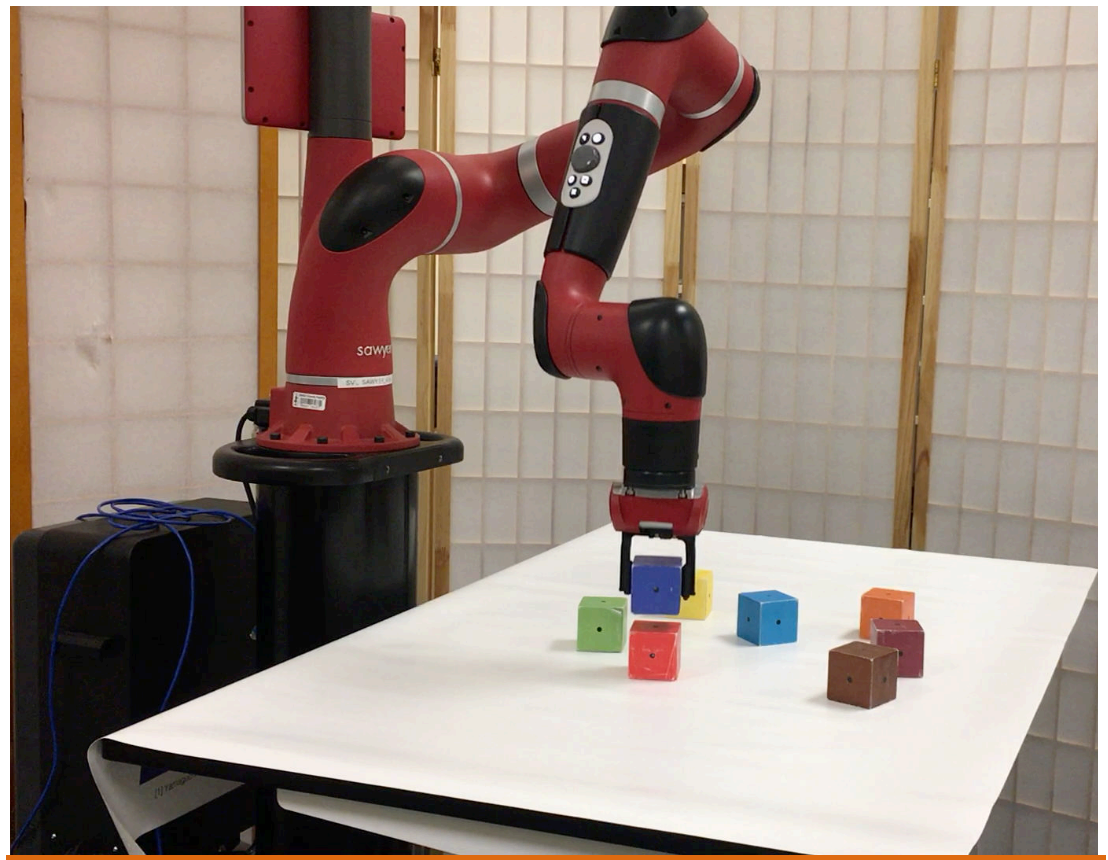




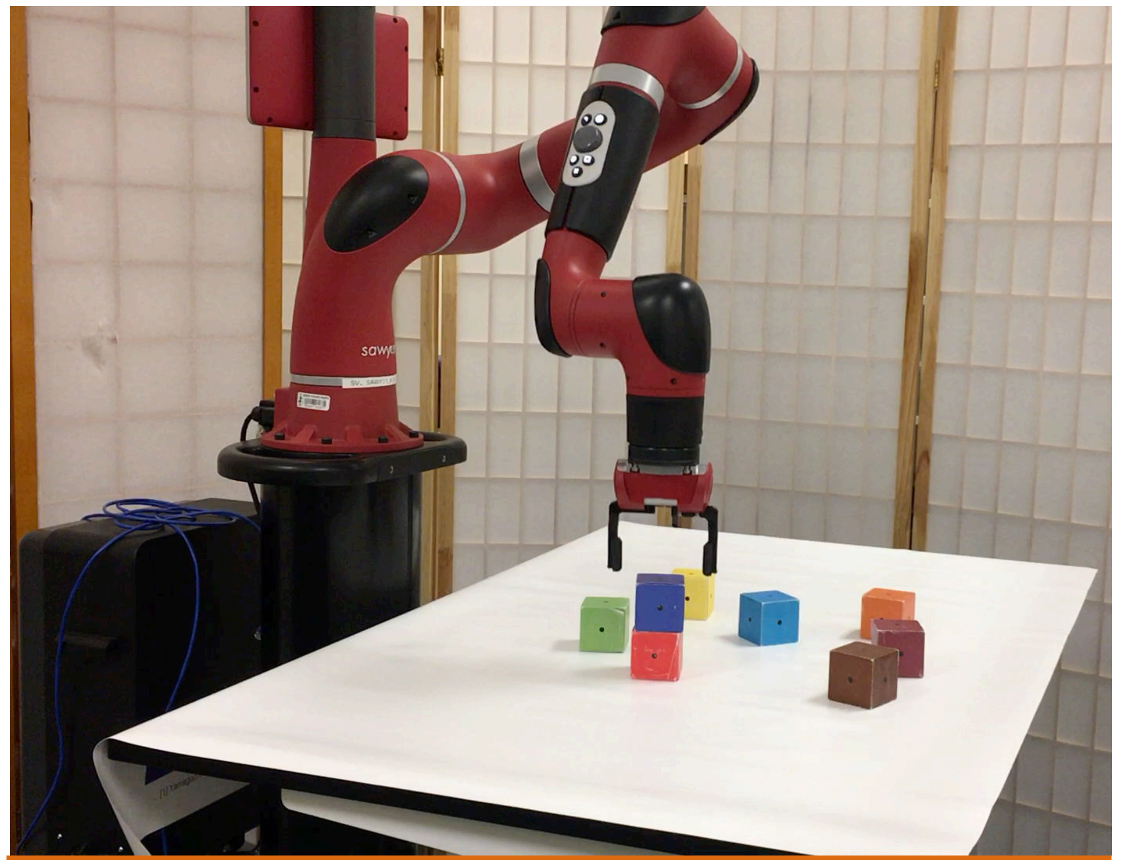
Move_to (Blue)



Grip (Blue)



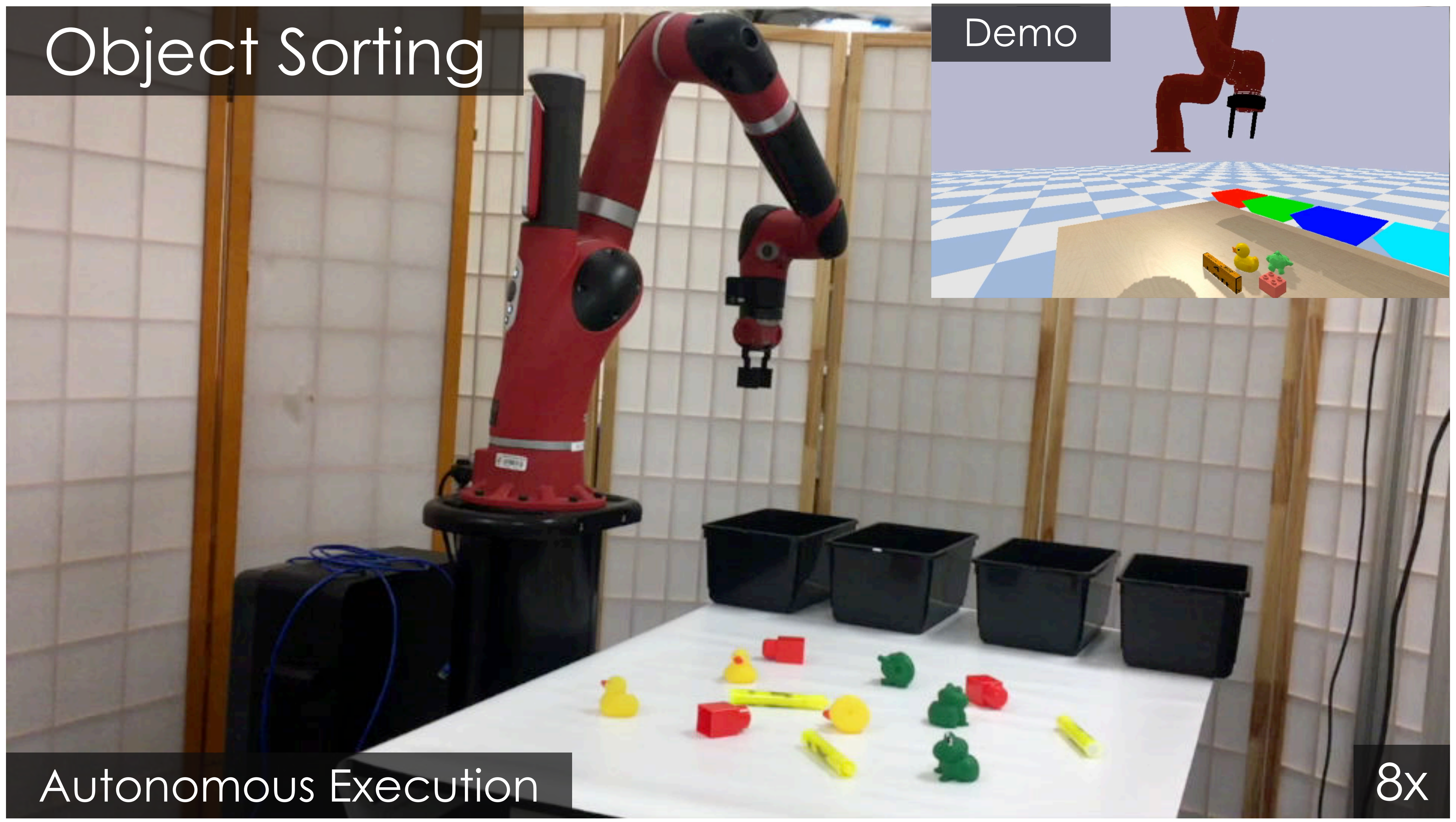
Move_to (Red)



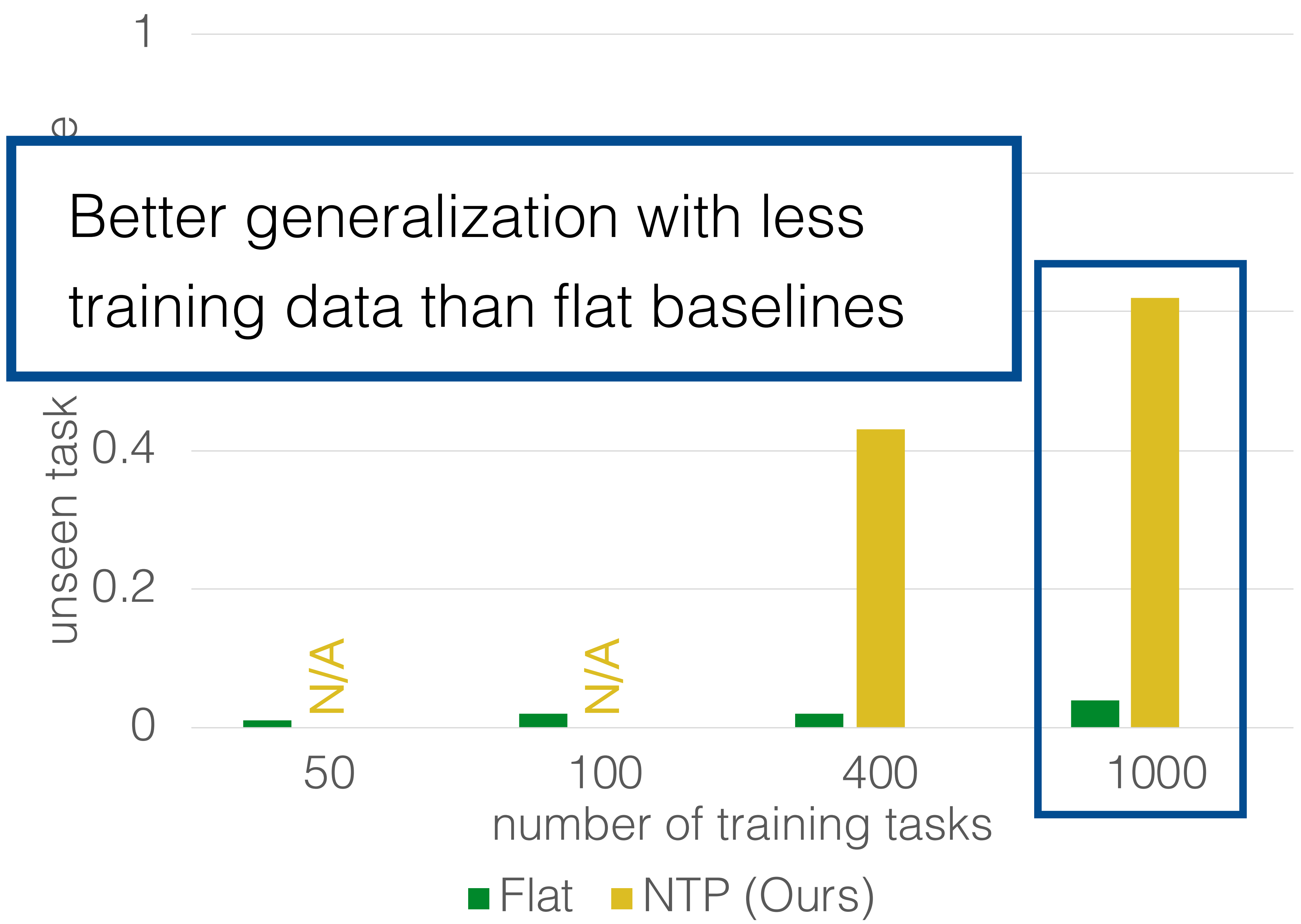
Release( )

Human-Like Learning: Compositionality

Modeling complex tasks as compositional program structures



Qualitative

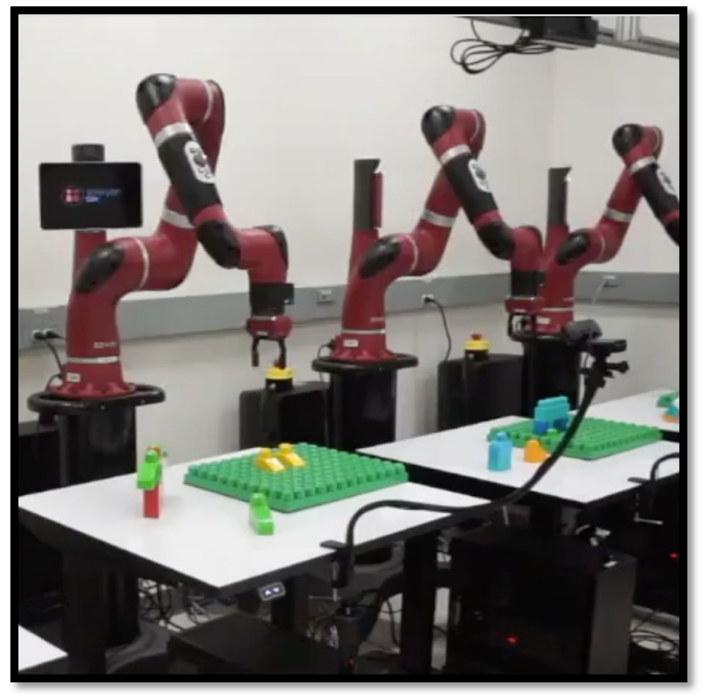


Quantitative

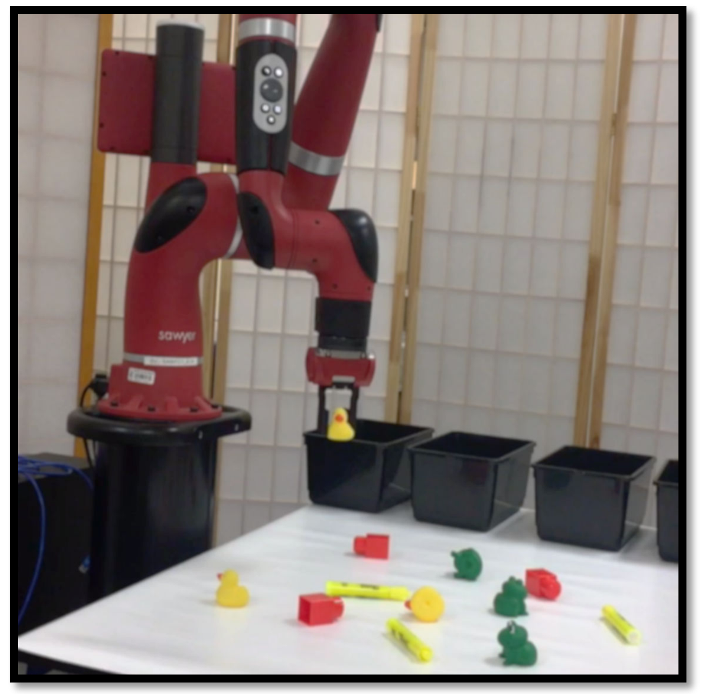
(the higher the better)

Xu*, Nair*, et al. ICRA 2018

Summary - Part III

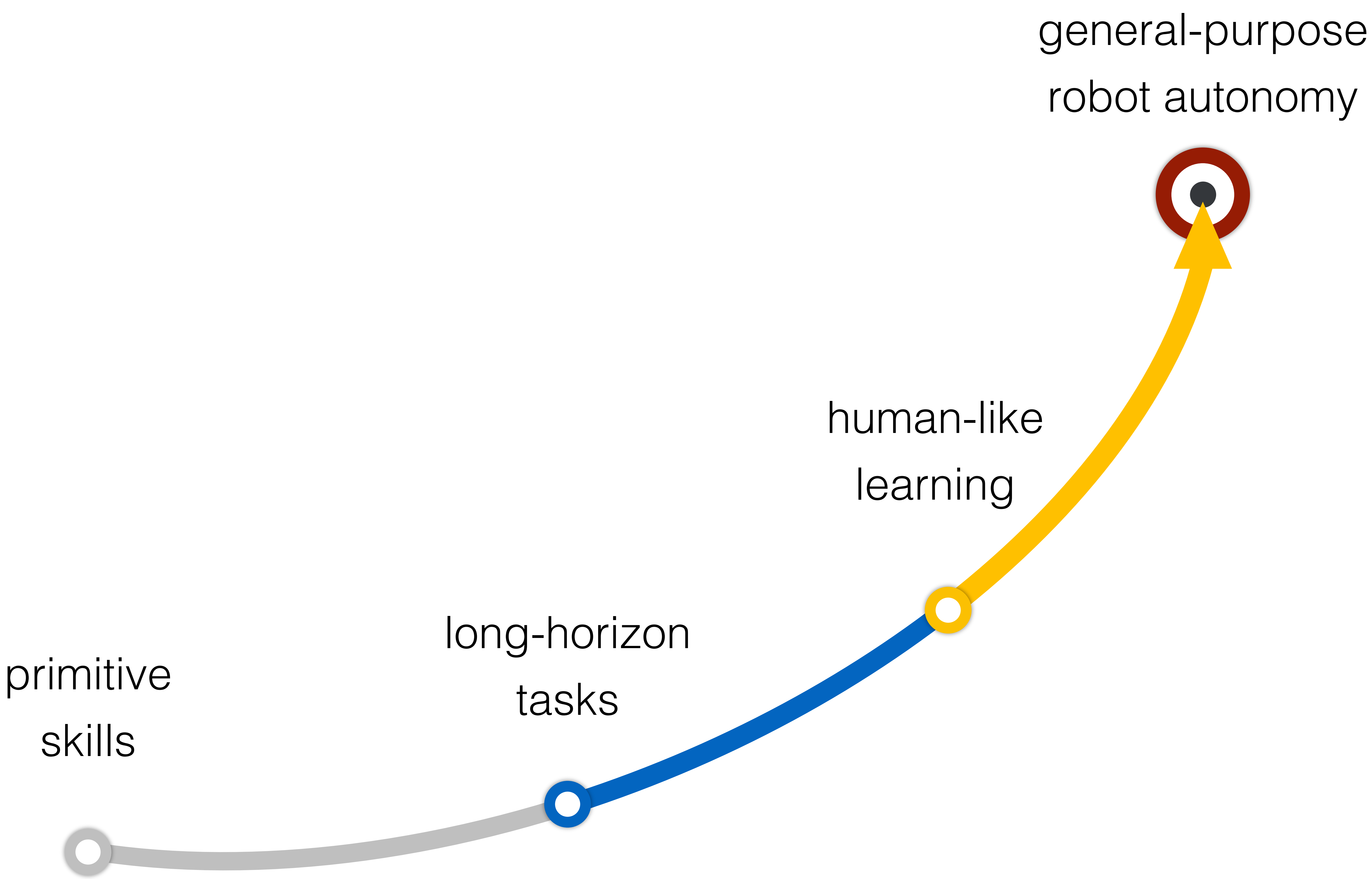


Learning from humans: Harvesting **human ingenuity** through teleoperated crowdsourcing with **RoboTurk**



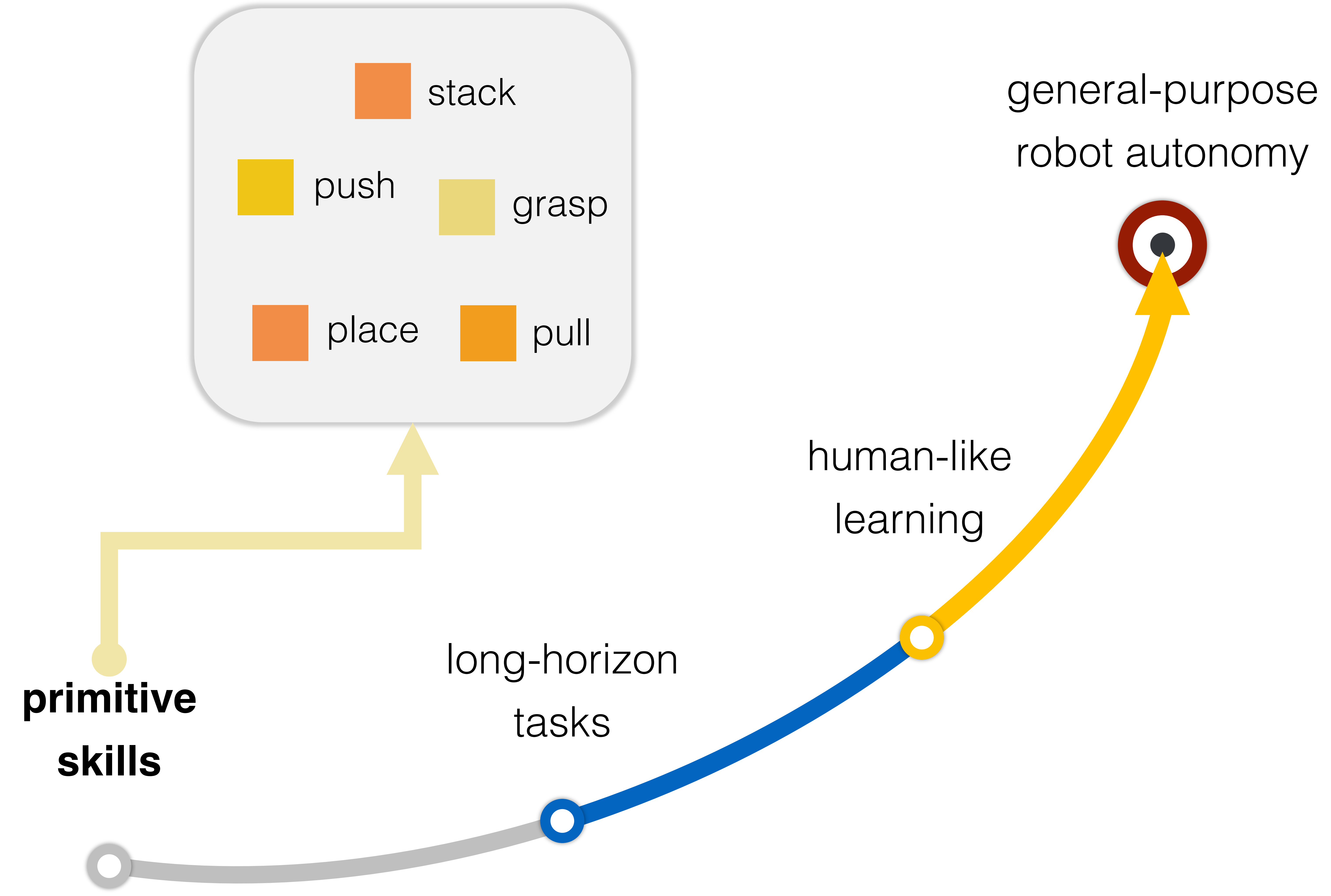
Learning like a human: Building agents that **learn to learn**, reason about **causal & effect**, and exploit **compositionality**

A Progressive Roadmap to General-Purpose Robot Autonomy



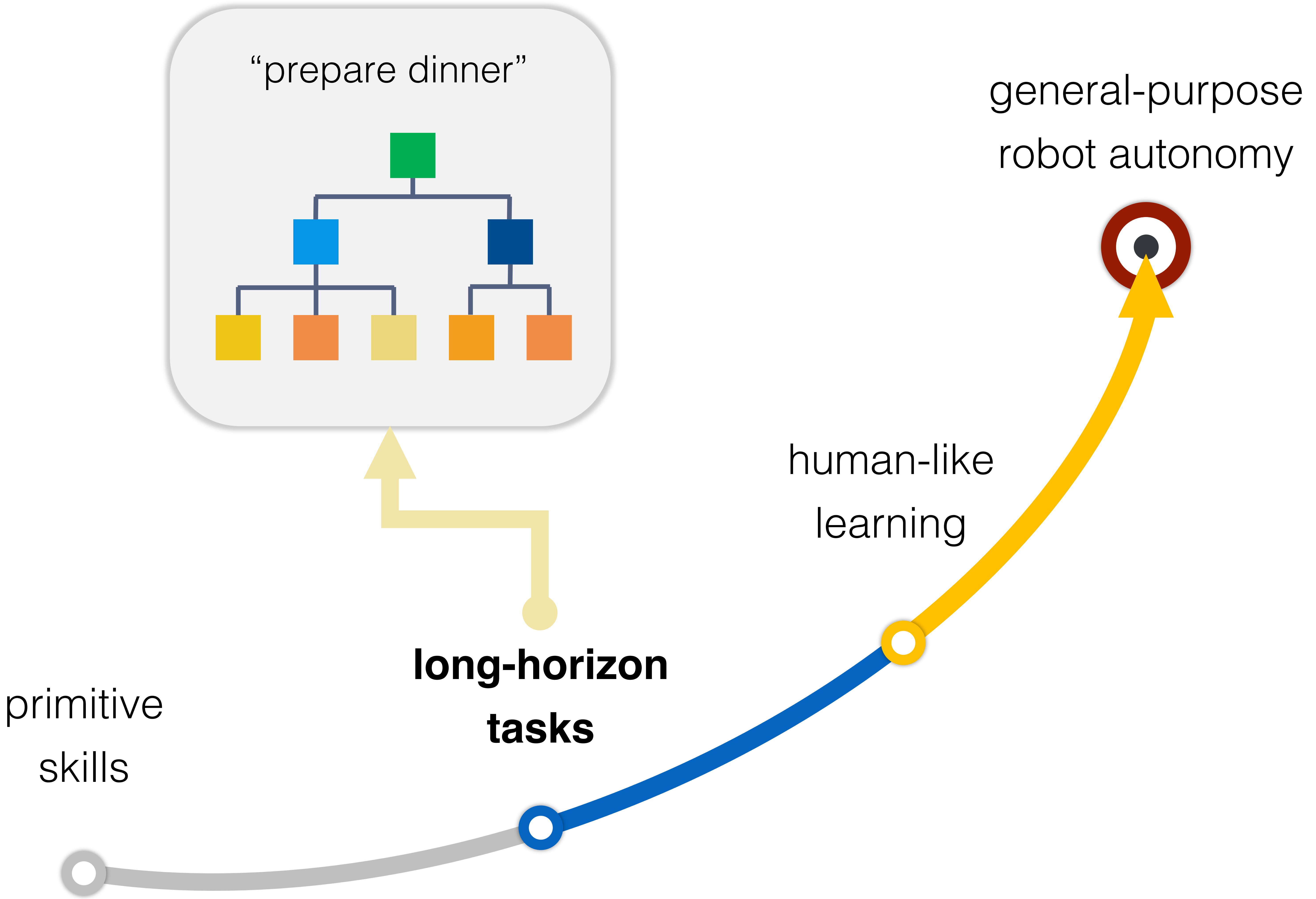
A Progressive Roadmap to General-Purpose Robot Autonomy

- **self-supervised learning** of primitive skills from raw sensory input



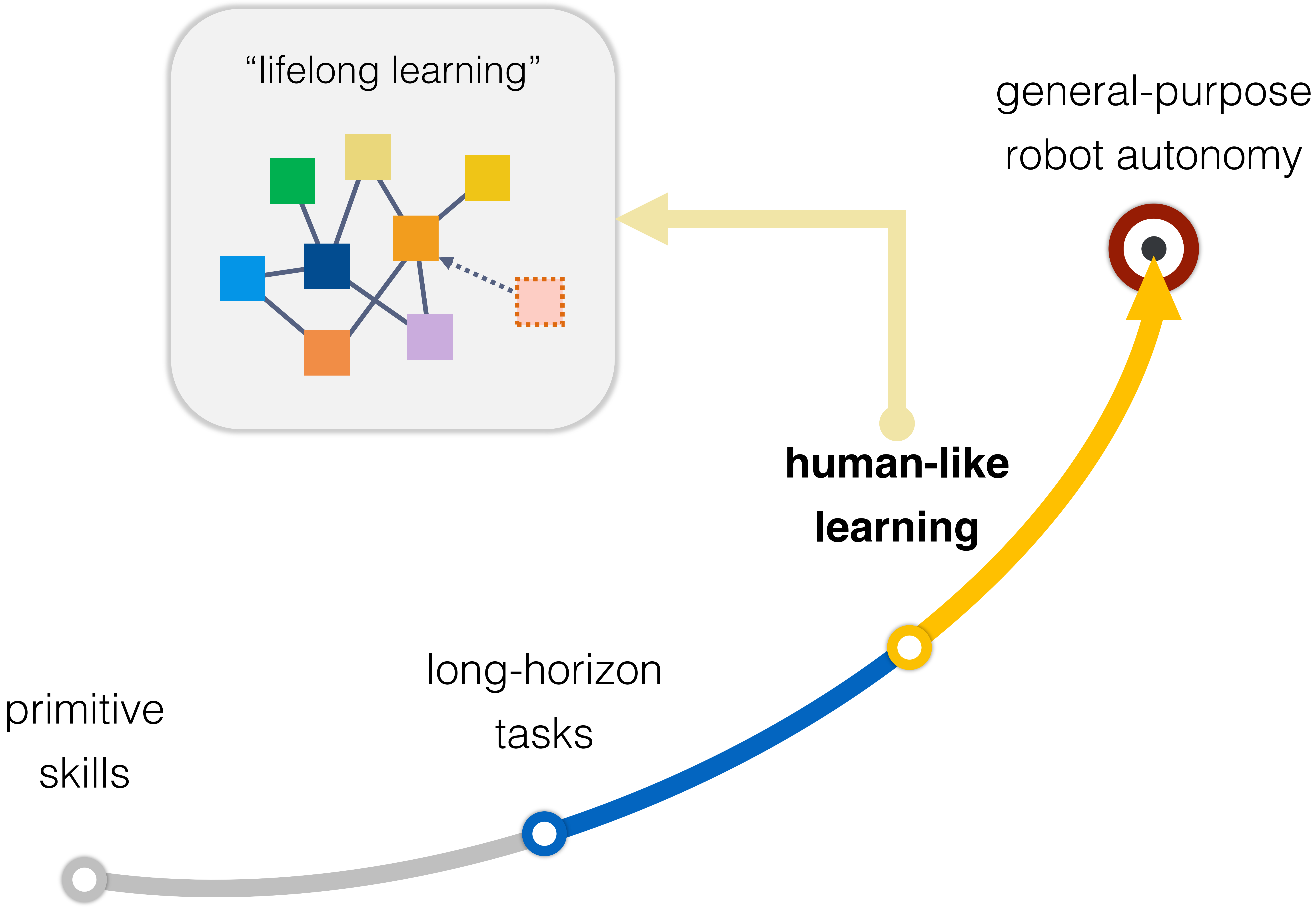
A Progressive Roadmap to General-Purpose Robot Autonomy

- **self-supervised learning** of primitive skills from raw sensory input
- scaling to long-horizon tasks through **hierarchy** and **abstraction**



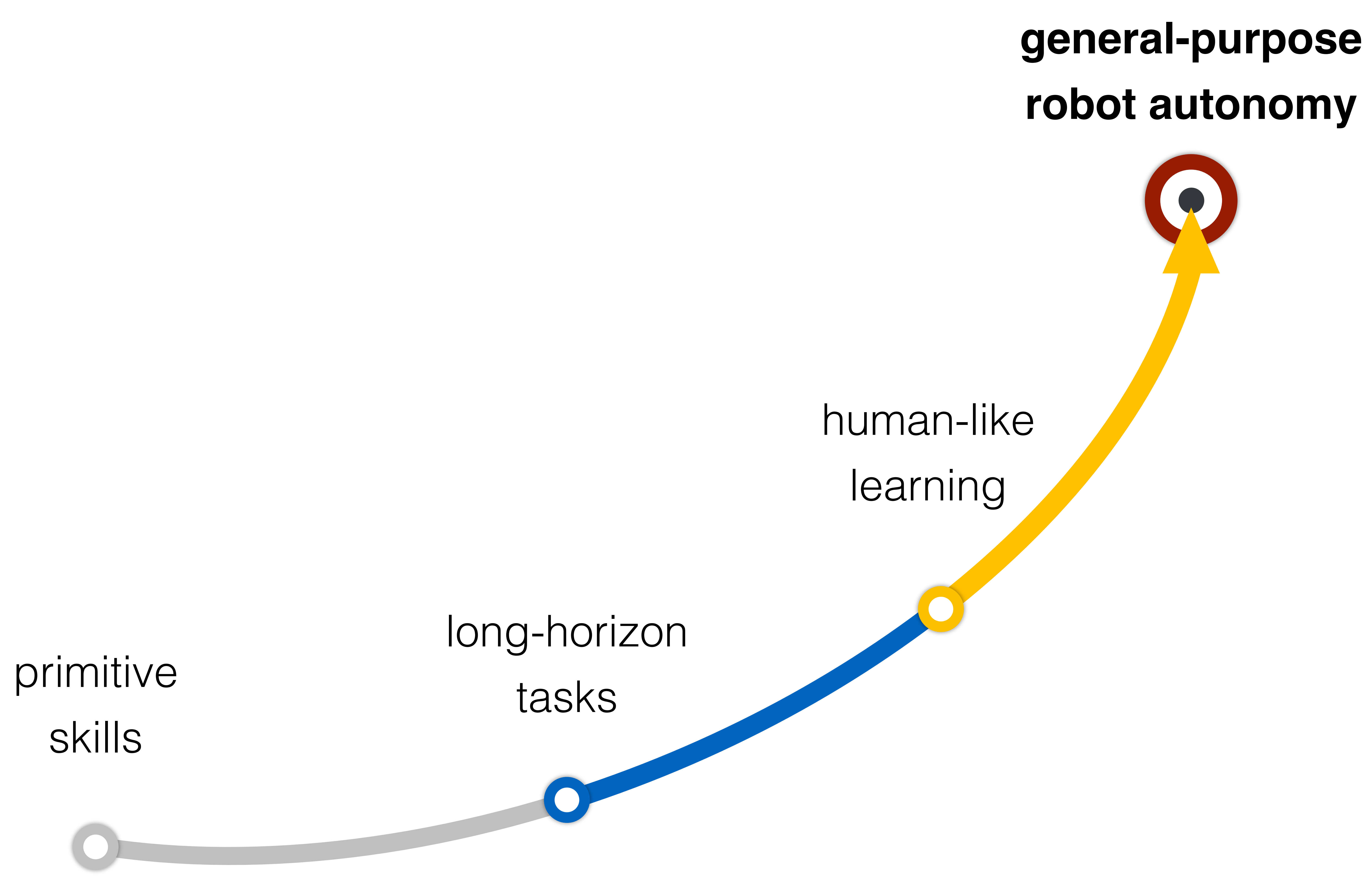
A Progressive Roadmap to General-Purpose Robot Autonomy

- **self-supervised learning** of primitive skills from raw sensory input
- scaling to long-horizon tasks through **hierarchy** and **abstraction**
- human-like learning via **active exploration** and **model building**



A Progressive Roadmap to General-Purpose Robot Autonomy

- **self-supervised learning** of primitive skills from raw sensory input
- scaling to long-horizon tasks through **hierarchy** and **abstraction**
- human-like learning via **active exploration** and **model building**



A Progressive Roadmap to General-Purpose Robot Autonomy

- **self-supervised learning** of primitive skills from raw sensory input
- scaling to long-horizon tasks through **hierarchy** and **abstraction**
- human-like learning via **active exploration** and **model building**

Acknowledgements

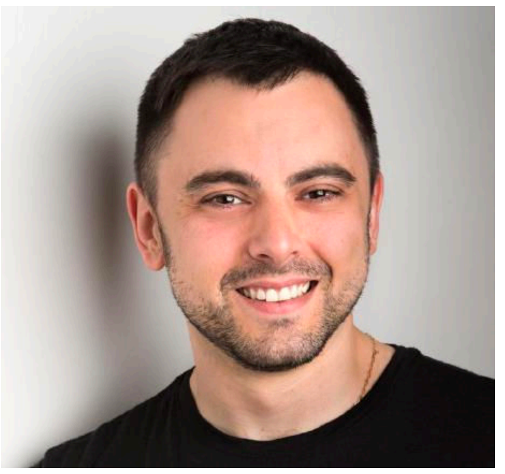

Fei-Fei Li


Silvio Savarese


Jeannette Bohg

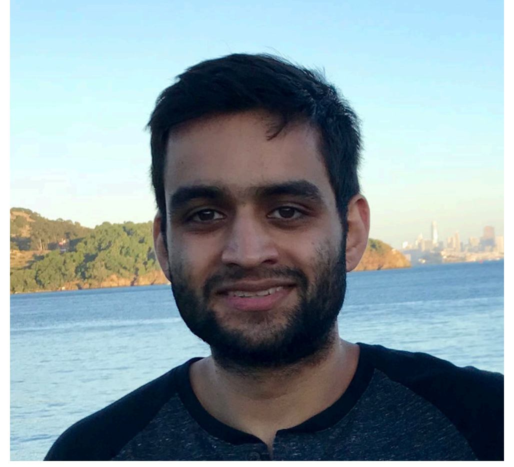

Animesh Garg


Anima Anandkumar

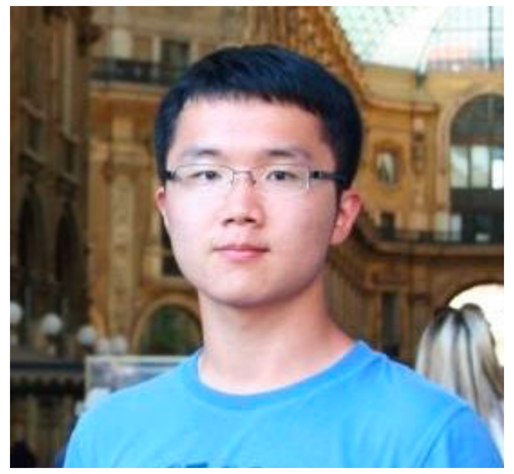

Roberto Martín-Martín


Danfei Xu


Michelle Lee


Ajay Mandlekar


De-An Huang


Kuan Fang*


Zengyi Qin


Hongyu Ren


Suraj Nair

TOYOTA RESEARCH INSTITUTE
Google
DEPARTMENT OF THE NAVY ONR Science & Technology
JD.COM

A Progressive Roadmap to General-Purpose Robot Autonomy

Yuke Zhu

@yukez

yukez@cs.utexas.edu

- **self-supervised learning** of primitive skills from raw sensory input
- scaling to long-horizon tasks through **hierarchy** and **abstraction**
- human-like learning via **active exploration** and **model building**