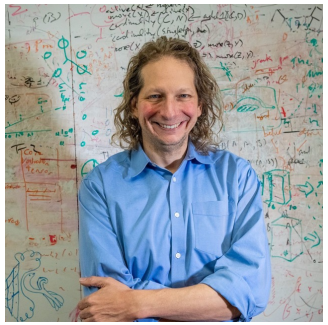
Learning Reusable Manipulation Strategies

Jiayuan Mao

Tomás Lozano-Pérez



Joshua B. Tenenbaum

Leslie Pack Kaelbling

MIT CSAIL
jiayuanm@mit.edu

What can we learn from one demonstration?

A "*strategy*" for pick up the cylinder.

- Push to rotate.
- Exert force on one end so that it tilts.
- Move the bucket.

You might not be able to execute it robustly now, but you have some "*ideas*."

We aim to learn such "strategies" from a single demonstration and apply them compositionally.

Problem Formulation

We have a simulator for object manipulation, an initial state, and a goal.

$$\langle \mathcal{Q}, \quad q_0, \quad \mathcal{G}, \quad \mathcal{U}, \quad \mathcal{T} \rangle$$

State Space | Initial State | Set of Goal States | Actions | Transition Model

Low-Level State:
Object poses, joint angles, velocities, etc.

Low-Level Action:
Joint position/velocity commands

Given a large enough amount of data and search budget, a general search algorithm can conceivably solve it.

Problem Formulation

We have a simulator for object manipulation & *one demonstration*.

What can we learn from the demonstration?

Single Demo

$\langle \mathcal{Q}, \mathcal{U}, \mathcal{T} \rangle$

Simulator

Learning Algo.

"Knowledge" of Manipulation Strategy

Planning Algo.

$q_0, \mathcal{G},$

New Task

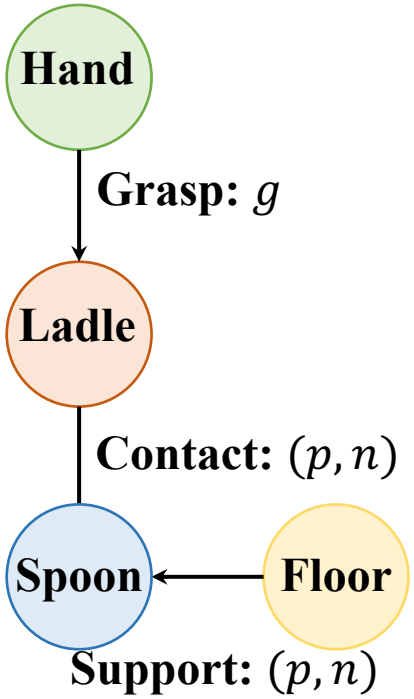
$u_1, u_2, \ldots$

Actions

What can we learn from one demonstration?

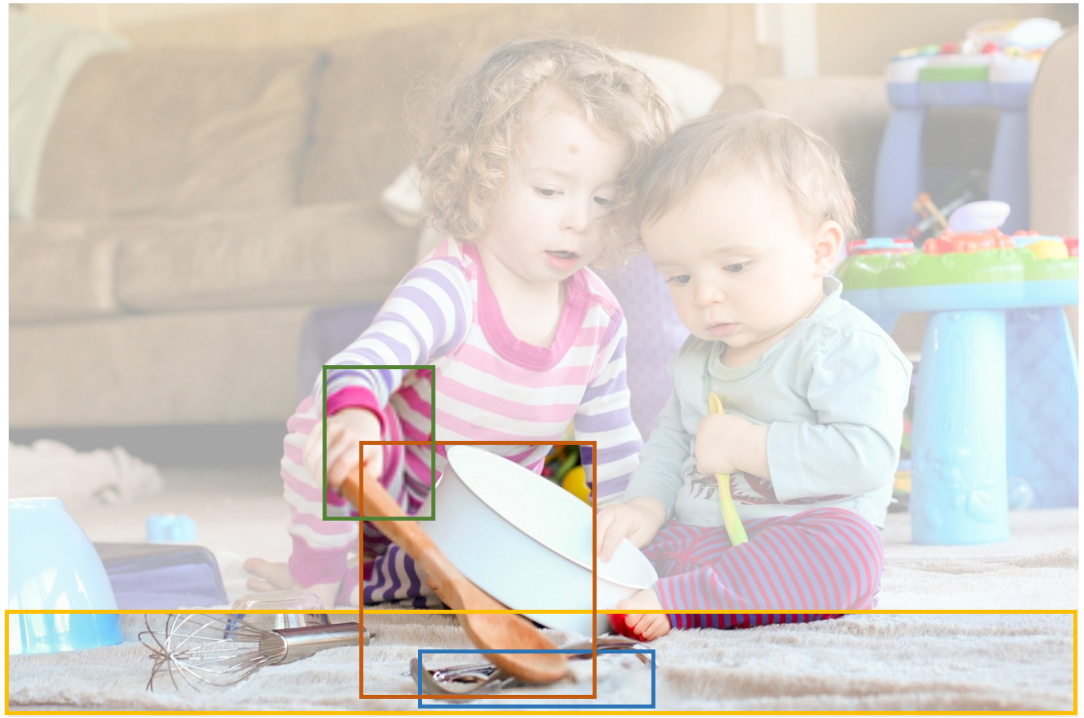
Key idea: some manipulation "strategies" can be modeled by contact relationships among objects.

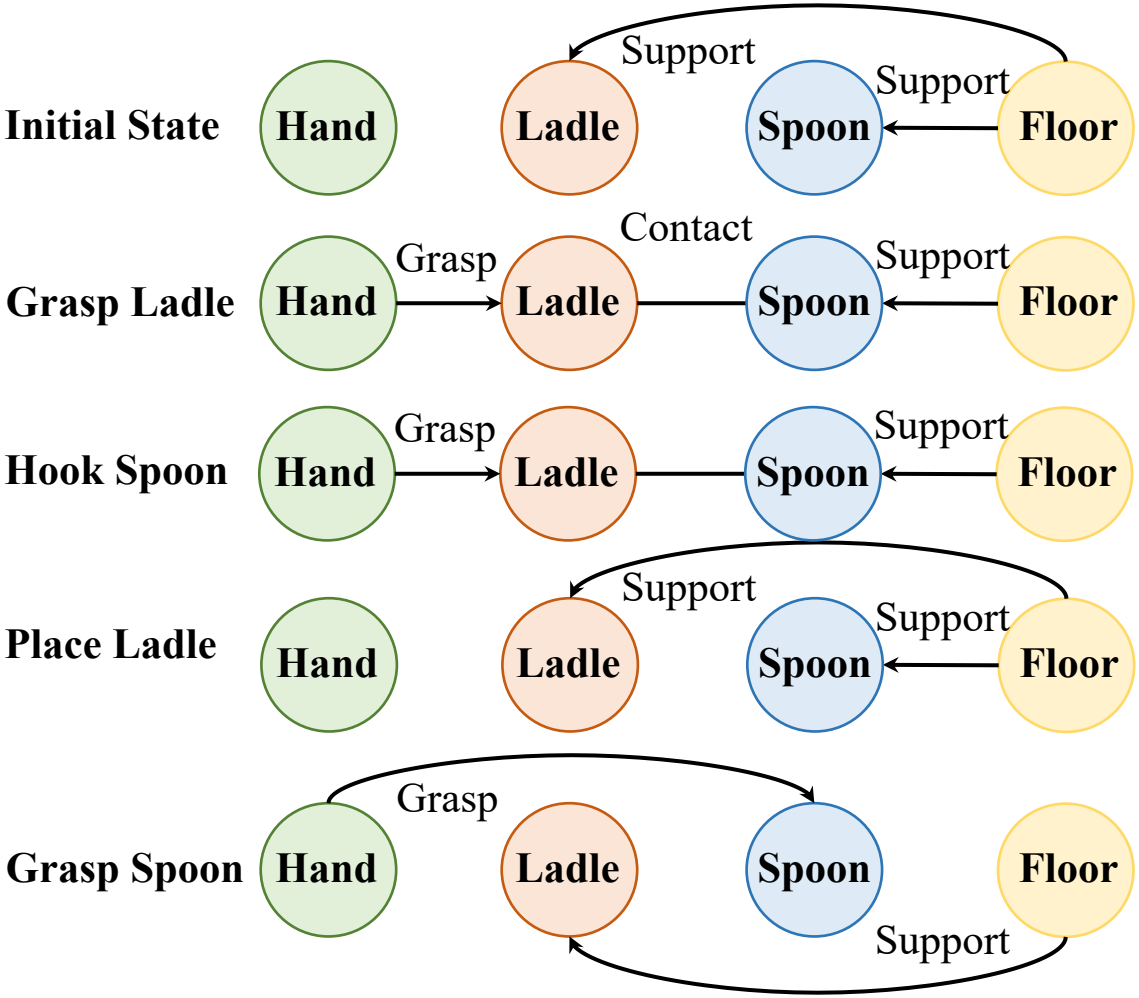
Let's talk about a familiar example: hook-using.



What can we learn from one demonstration?

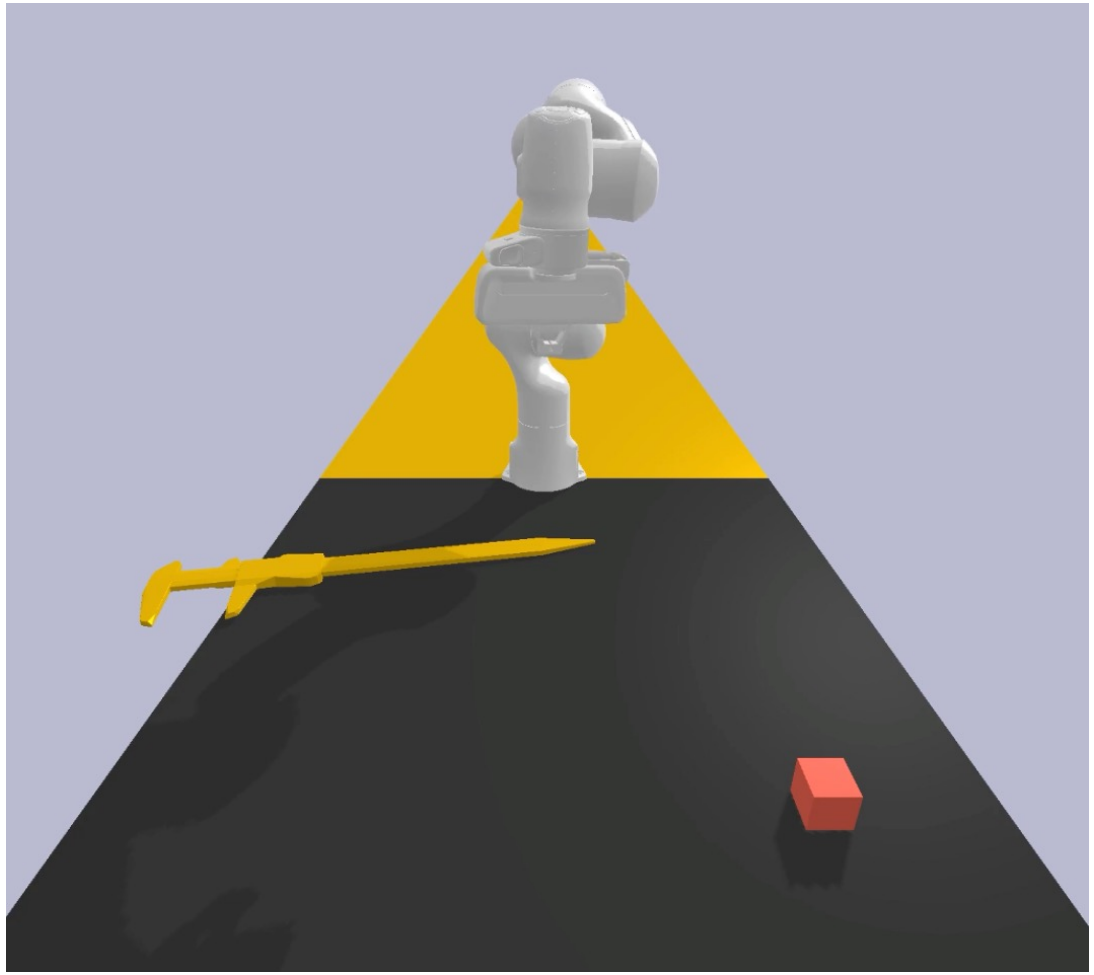
Key idea: some manipulation "strategies" can be modeled by
a sequence of contact relationships among objects.

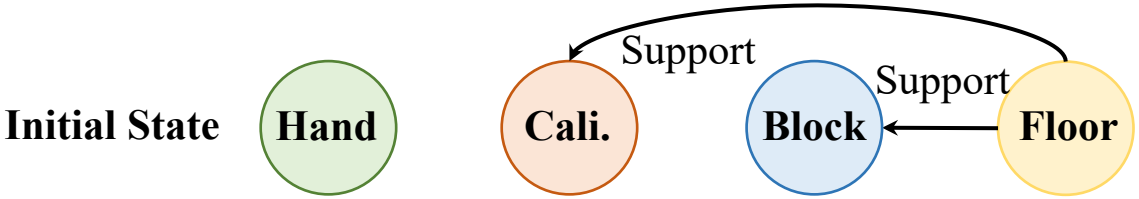




What can we learn from one demonstration?

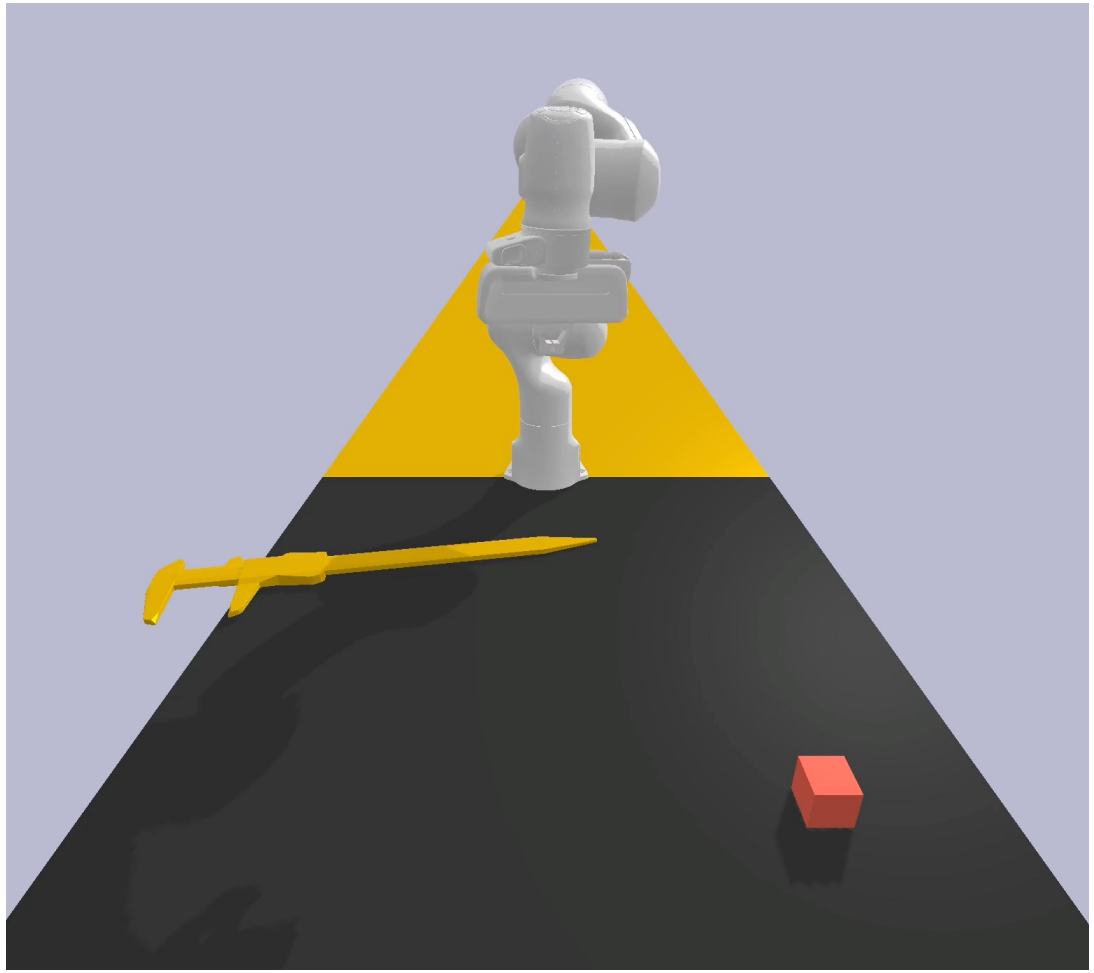
Key idea: some manipulation "strategies" can be modeled by
a sequence of contact relationships among objects.

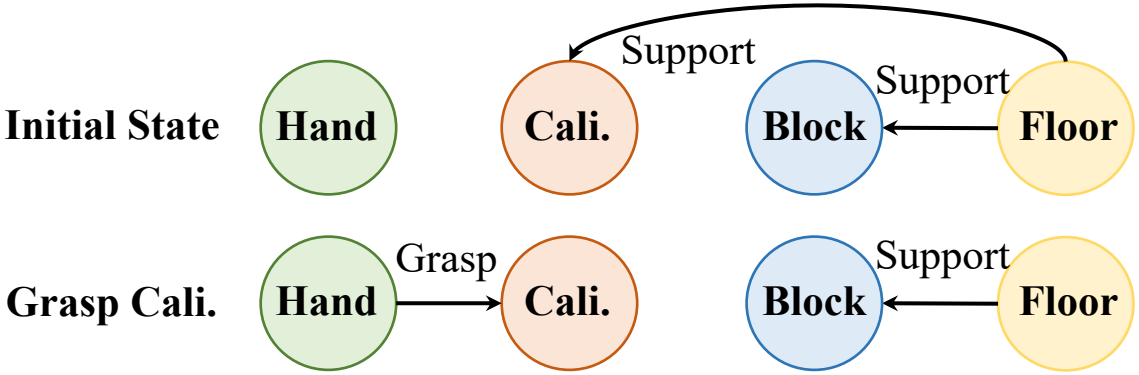




What can we learn from one demonstration?

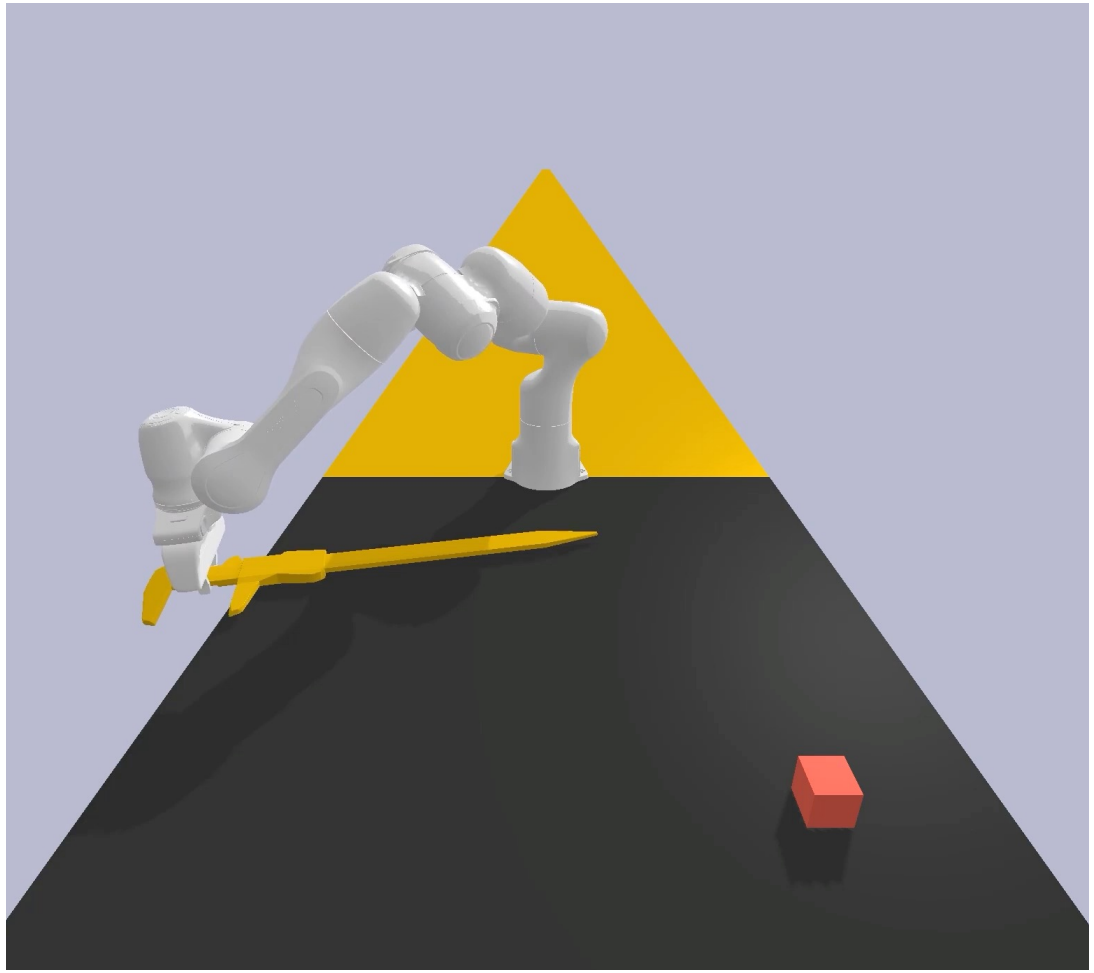
Key idea: some manipulation "strategies" can be modeled by a sequence of contact relationships among objects.

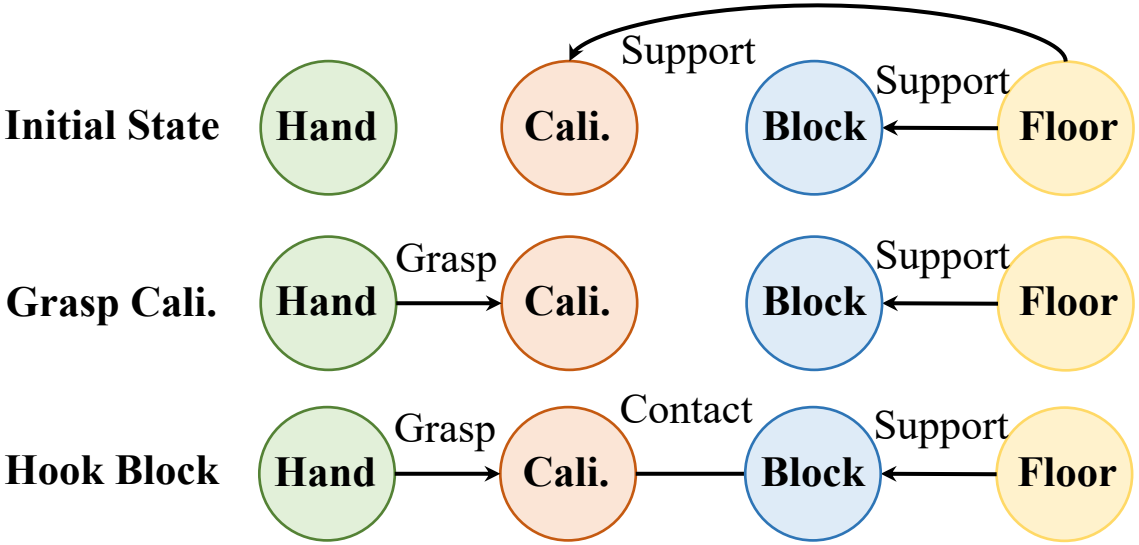




What can we learn from one demonstration?

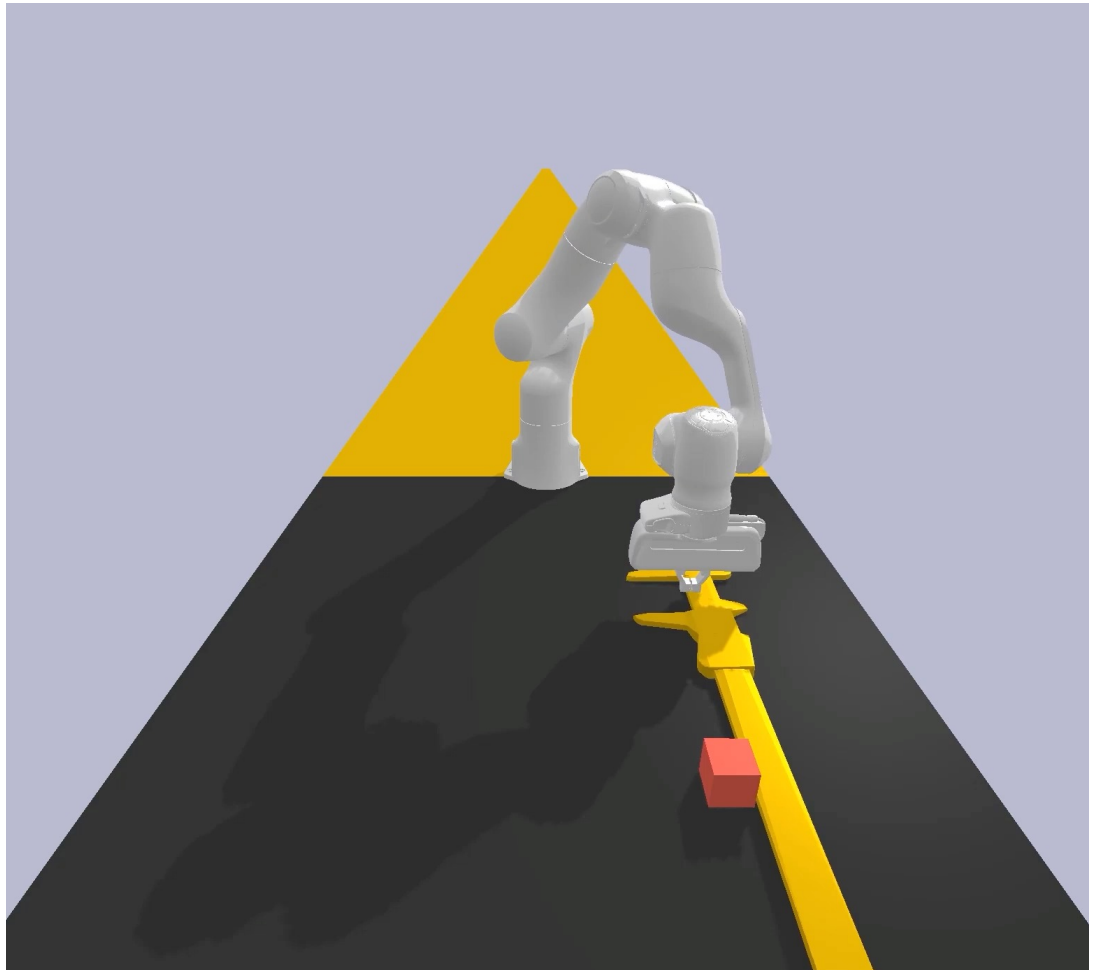
Key idea: some manipulation "strategies" can be modeled by a sequence of contact relationships among objects.

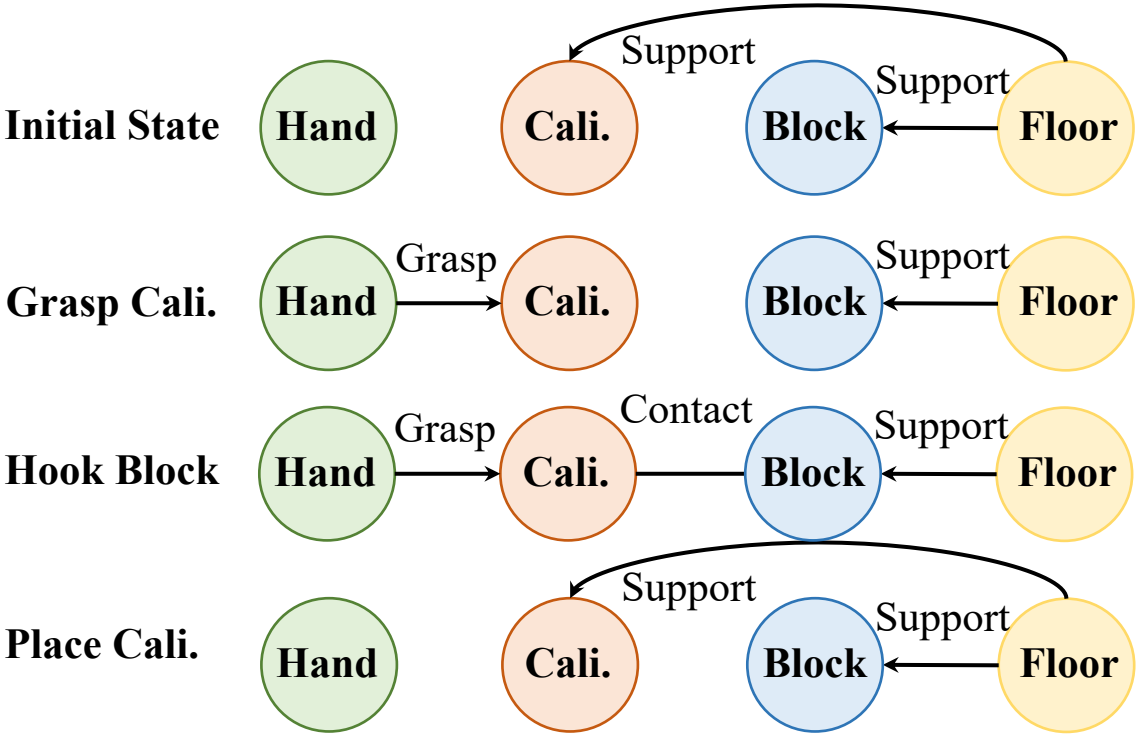




What can we learn from one demonstration?

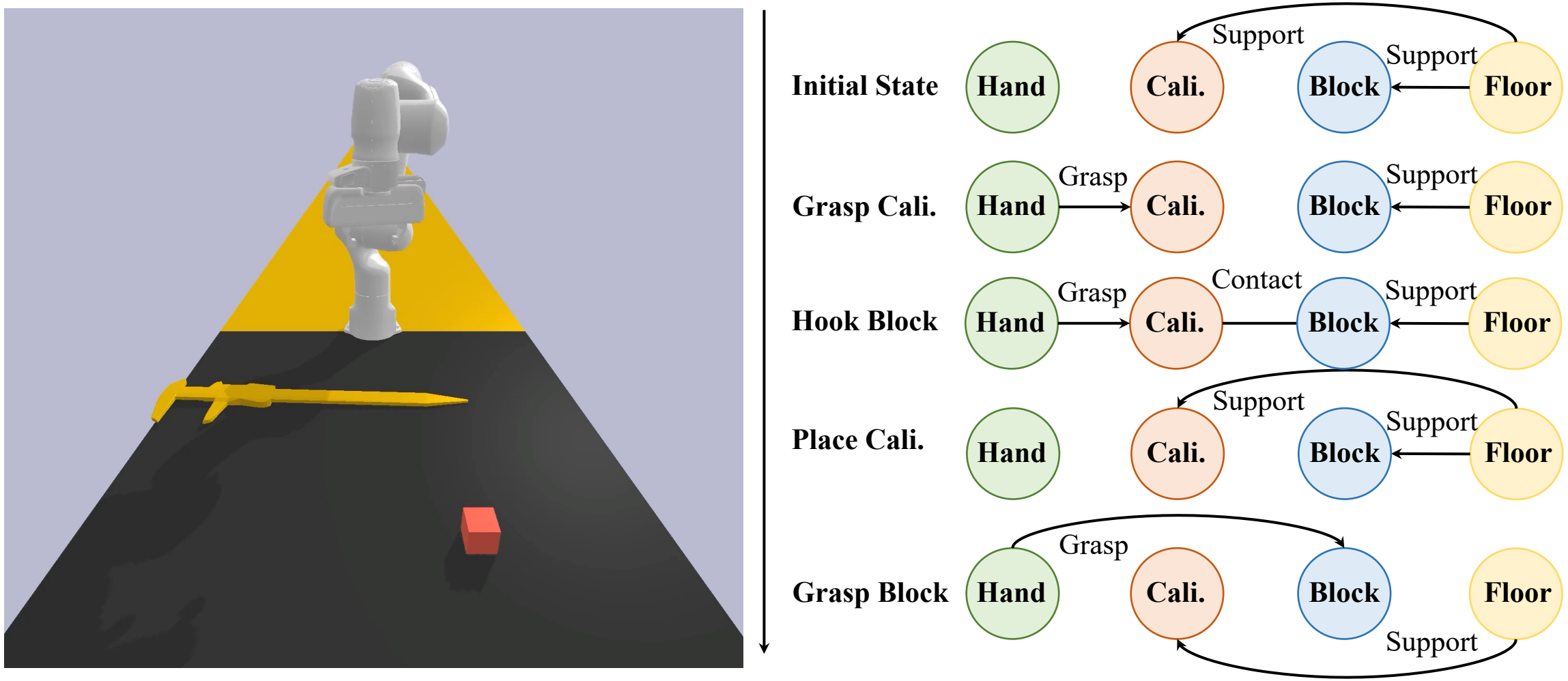
Key idea: some manipulation "strategies" can be modeled by a sequence of contact relationships among objects.





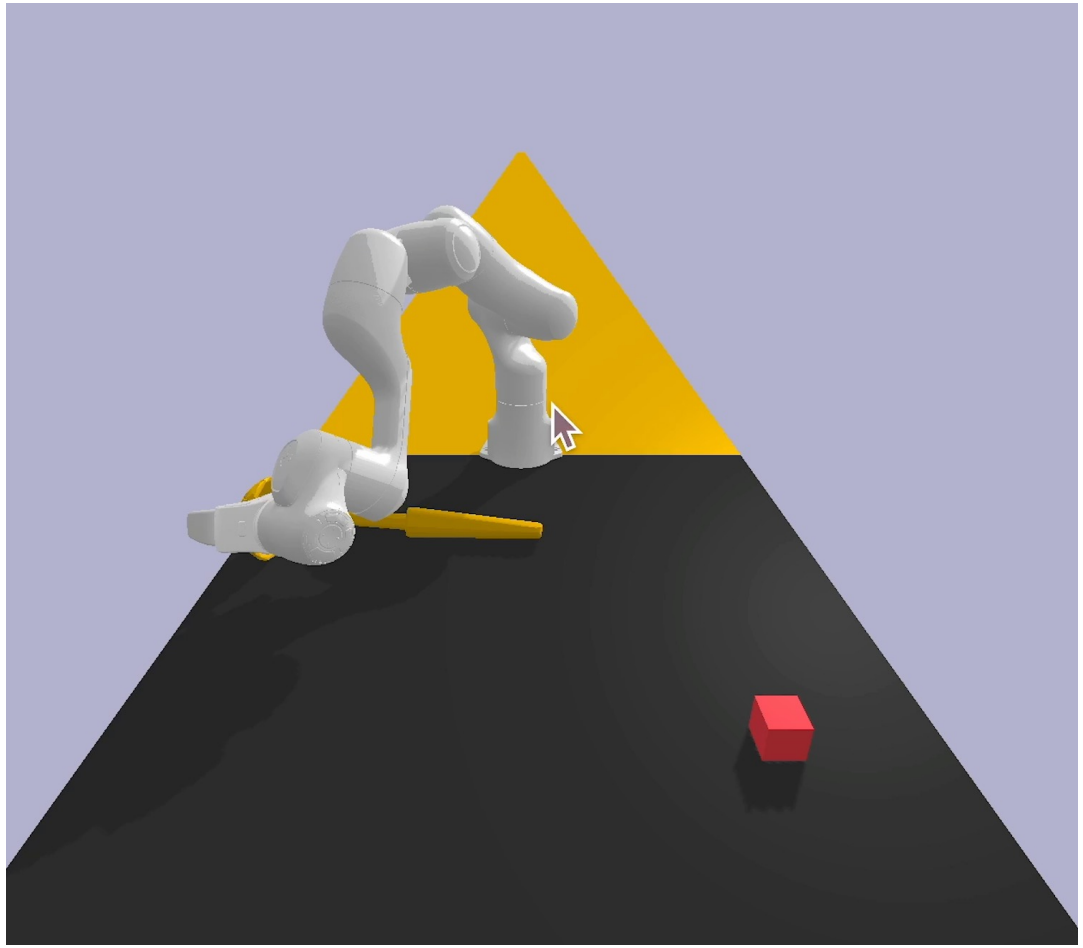
What can we learn from one demonstration?

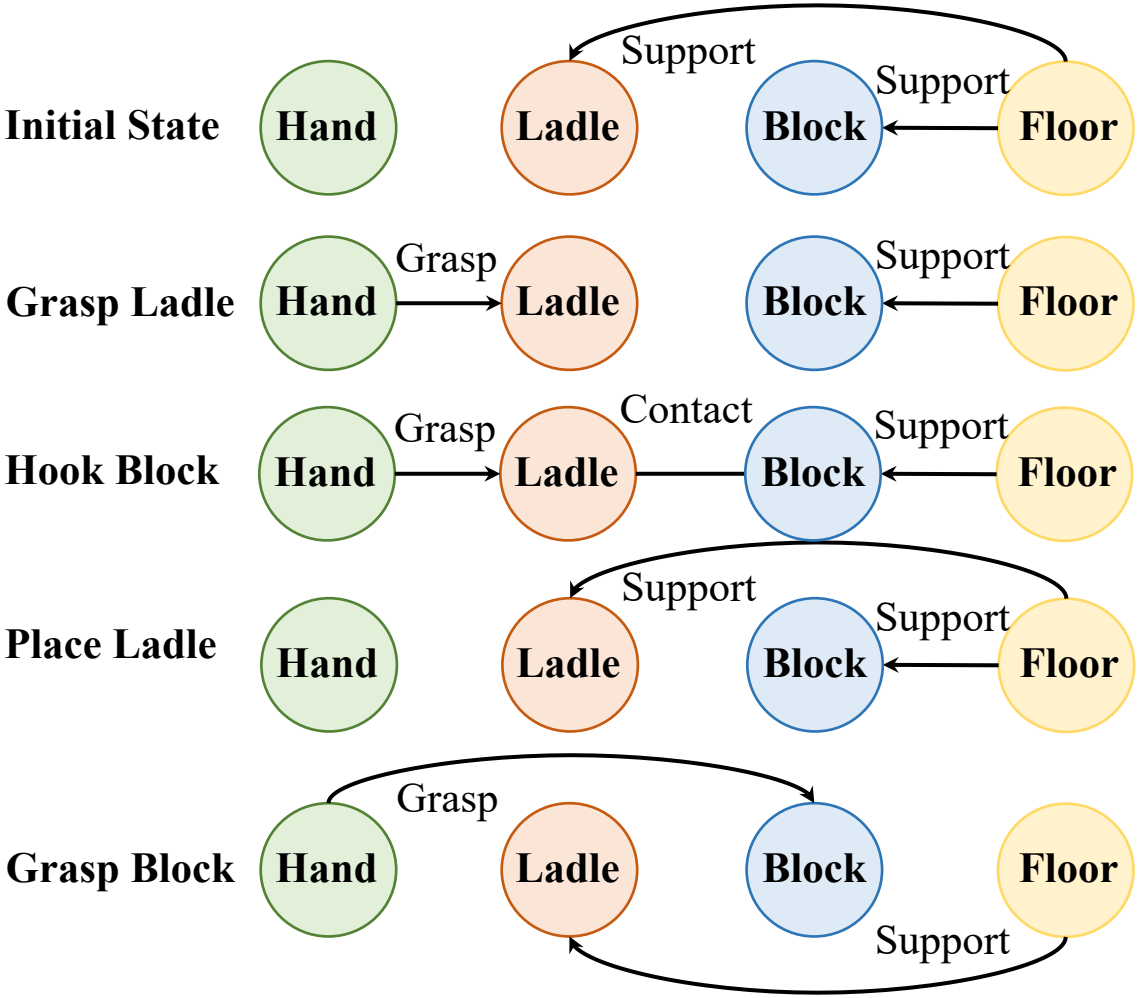
Key idea: some manipulation "strategies" can be modeled by a sequence of contact relationships among objects.



What can we learn from one demonstration?

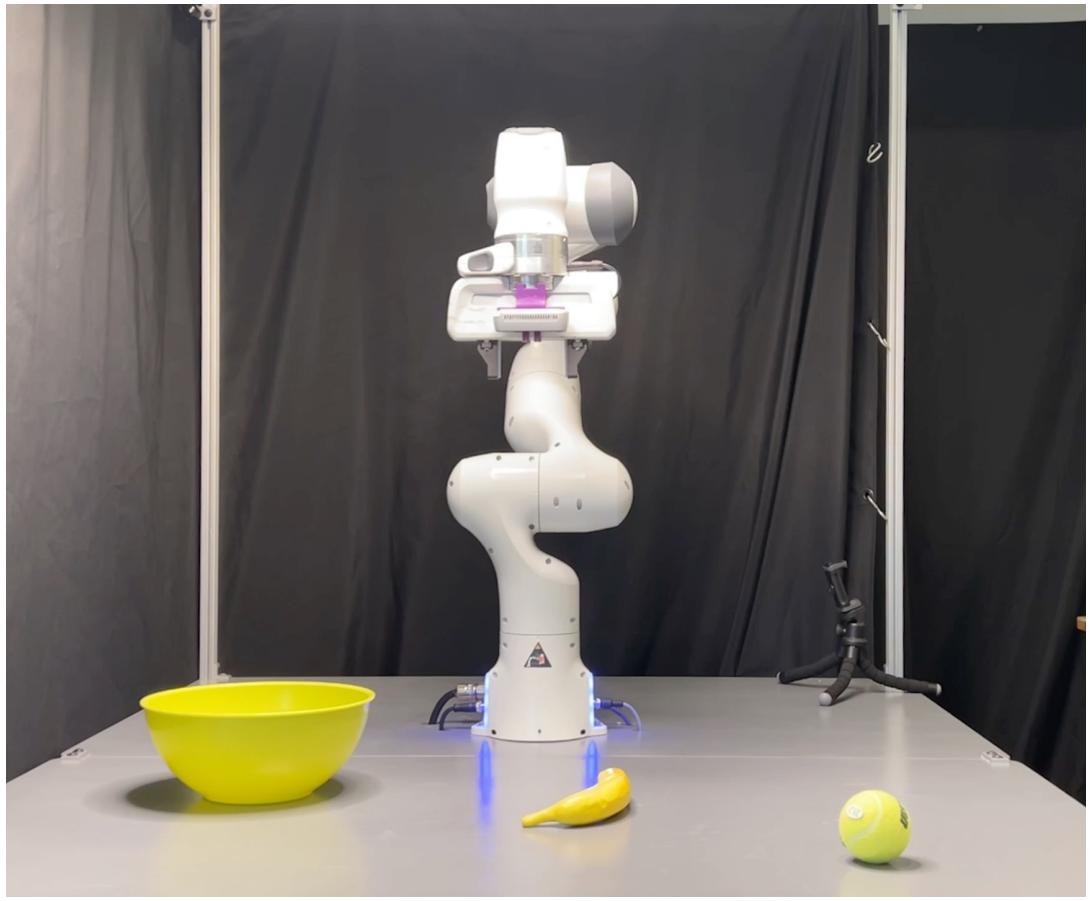
Key idea: some manipulation "strategies" can be modeled by a sequence of contact relationships among objects.

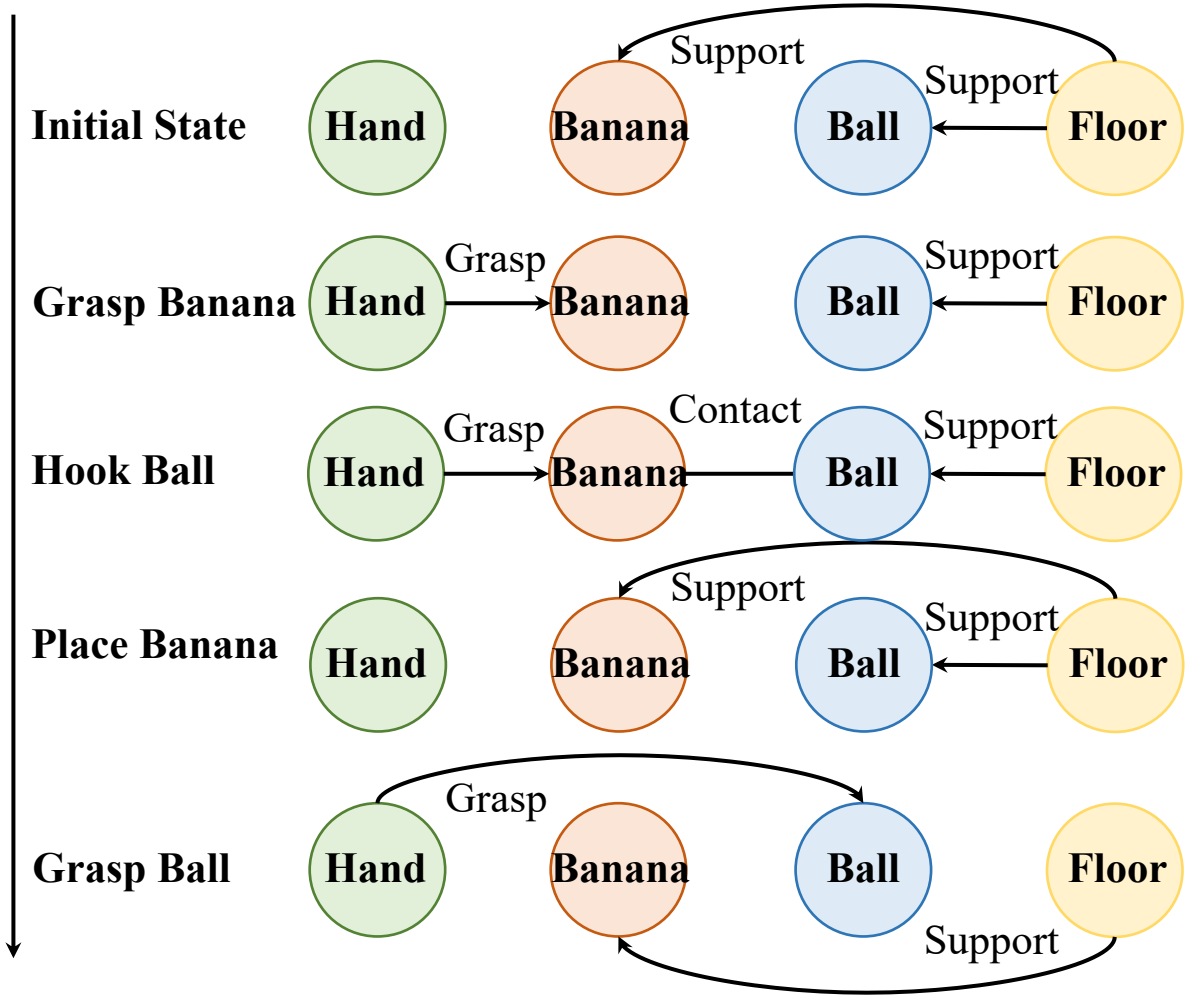




What can we learn from one demonstration?

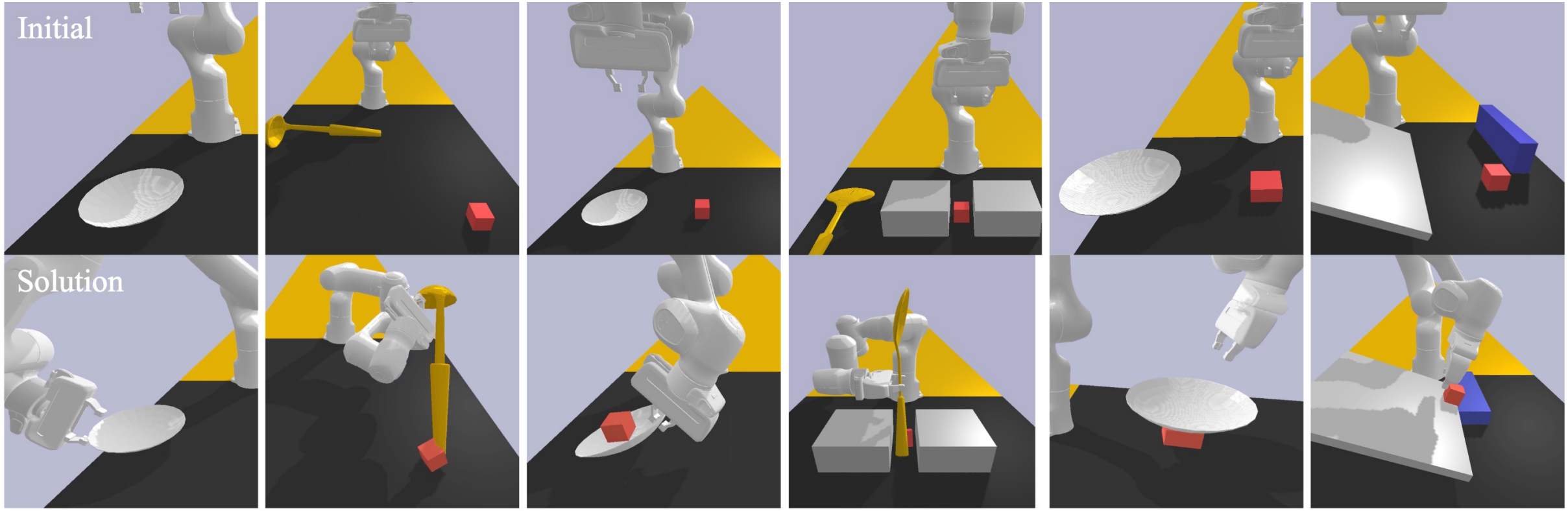
Key idea: some manipulation "strategies" can be modeled by a sequence of contact relationships among objects.





Many strategies can be represented this way.

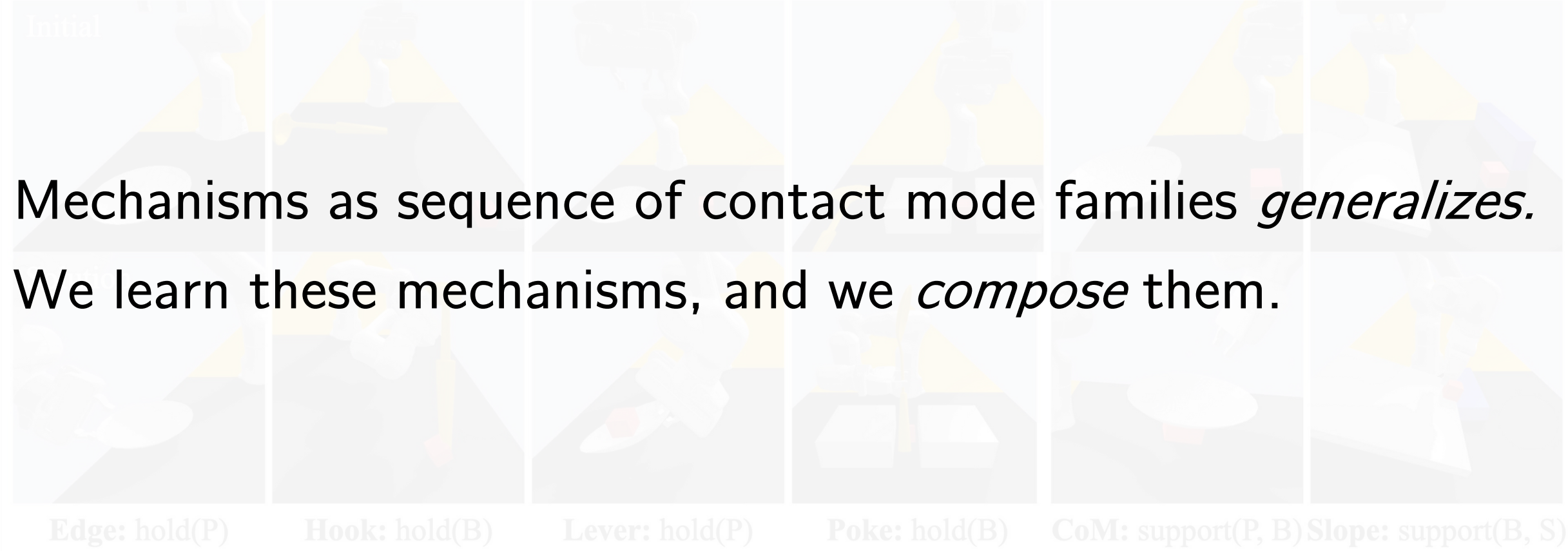
We call these manipulation strategies "*mechanisms.*"



Edge: hold(P) **Hook:** hold(B) **Lever:** hold(P) **Poke:** hold(B) **CoM:** support(P, B) **Slope:** support(B, S)

Many strategies can be represented this way.

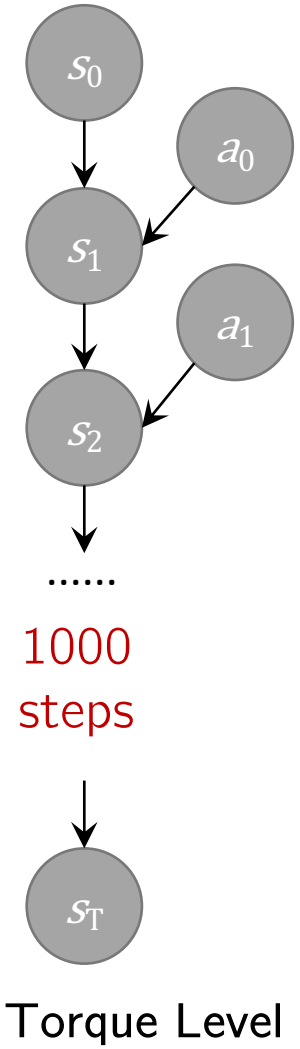
We call these manipulation strategies "*mechanisms.*"



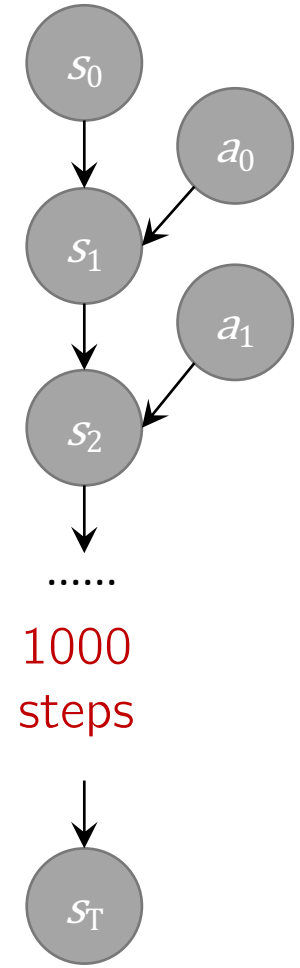
Mechanisms as sequence of contact mode families *generalizes.*

We learn these mechanisms, and we *compose* them.

Motion Planning



Task and Motion Planning



Torque Level

Domain Specification

Predicates describe object poses and properties.

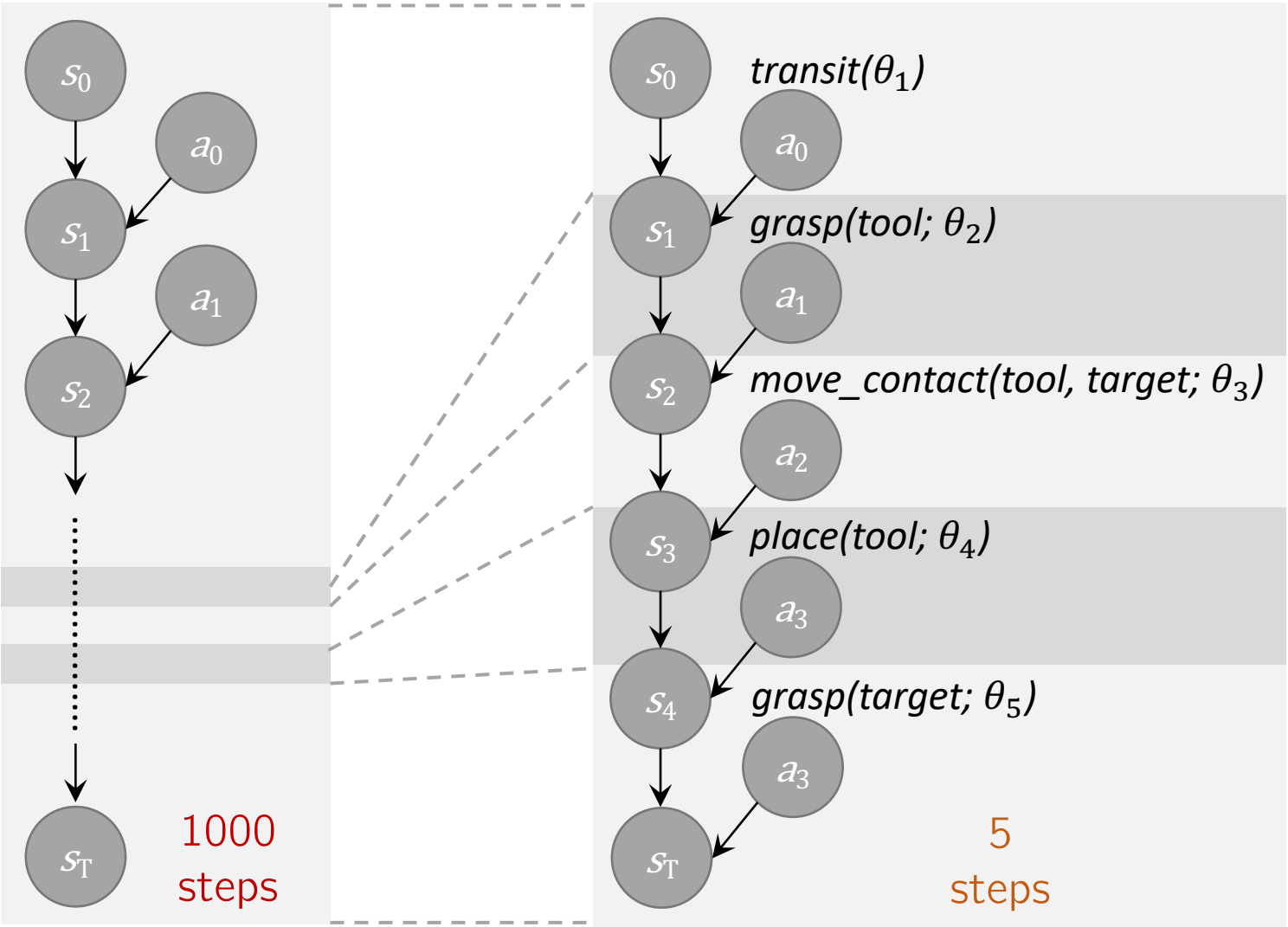
Operators parameterized by continuous parameters model contact mode families: grasp, move, *etc.*
—— they have preconditions, effects, and controllers.

State → Discrete + Continuous Planning Algorithms → Control Cmds.

Goal "Get the spoon."

Our system currently uses 6 operators involving interactions among at most 3 objects. But it can be easily extended to include more complex operators.

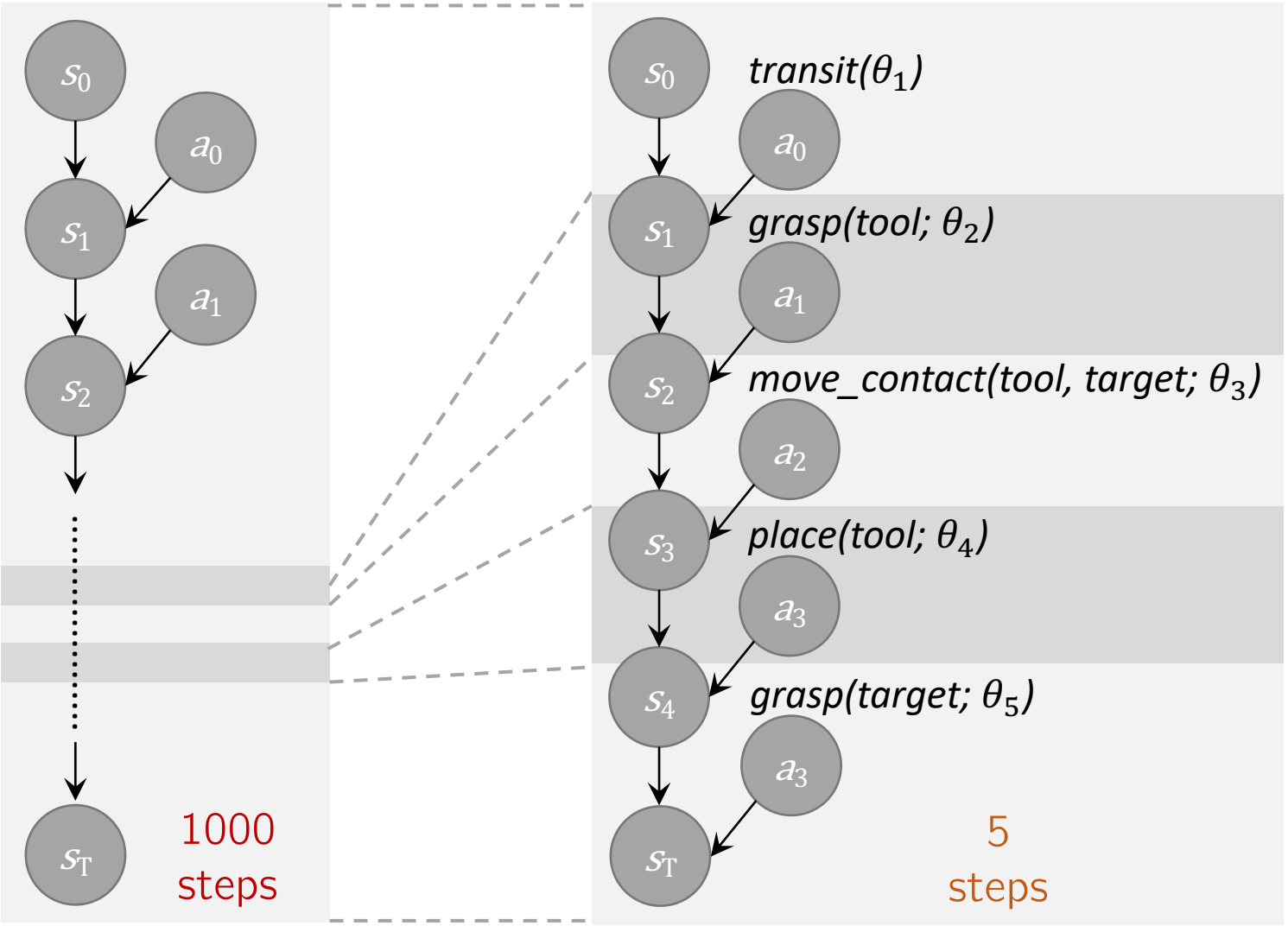
Task and Motion Planning



Torque Level

Contact Mode Level

Task and Motion Planning with Mechanisms

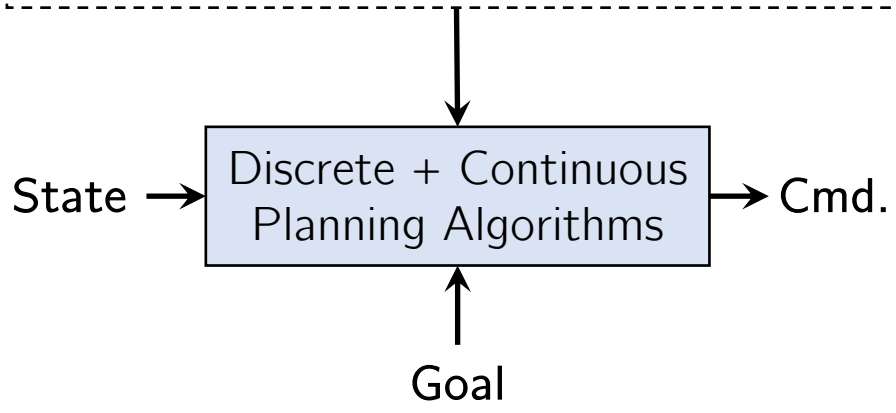


Domain Specification

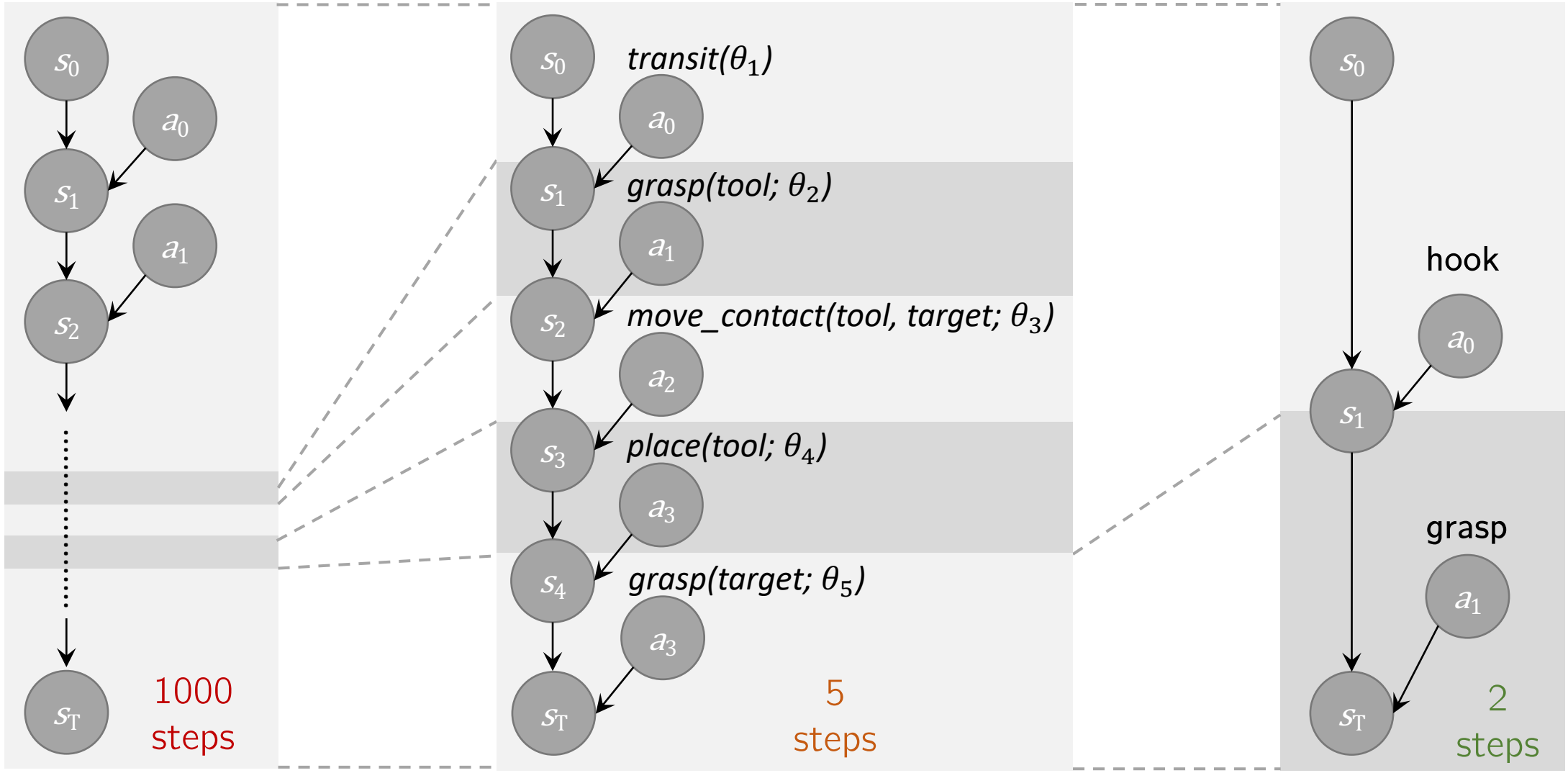
Predicates: poses, properties

Operators: contact mode families

Mechanisms: useful macros of operators, and dedicated samplers for continuous parameters



Task and Motion Planning



Torque Level

Contact Mode Level

Mechanism Level

Overview of the Framework

There are two **learning problems**:

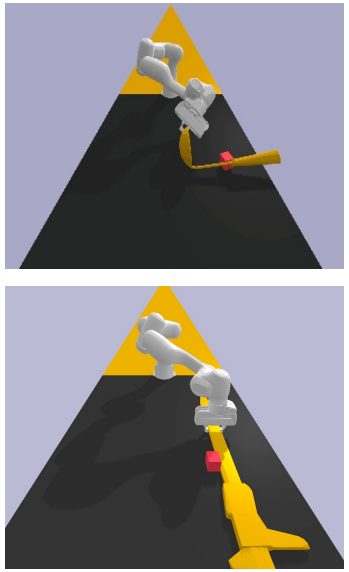
1. Learning of the contact mode sequence.

2. Learning samplers for parameters of the contact modes: where to grasp, how to move, *etc.*

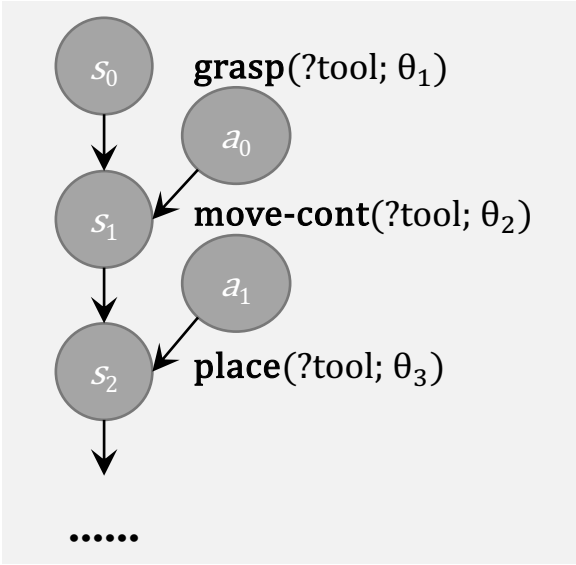
Single Demo

```
(:macro hook
 :parameters (
  ?tool - item
  ?target - item
  ?support - item
 )
 :certified (
  (holding ?target)
 )
 ...)
```

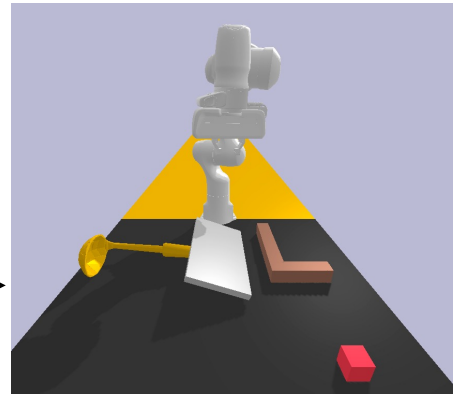
Contact Modes and Goals



Self-Play



Learned Contact Distributions



Goal:
Block on the Slope

Compositional Planning

Step 1: Recover Mode Sequences

There are two **learning problems**:

1. Learning of the contact mode sequence.

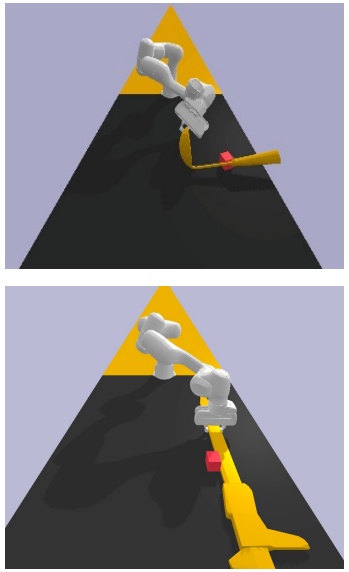
   We will recover it from the single demonstration.

2. Learning samplers for parameters of the contact modes: where to grasp, how to move, *etc.*

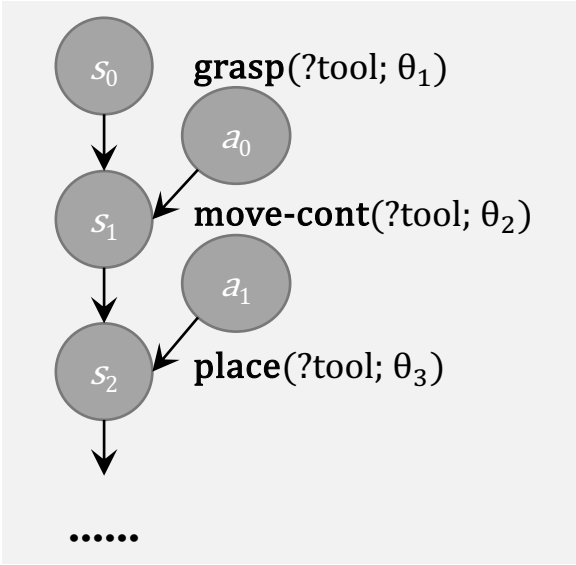
Single Demo

```
(:macro hook
 :parameters (
  ?tool - item
  ?target - item
  ?support - item
 )
 :certified (
  (holding ?target)
 )
 ...)
```

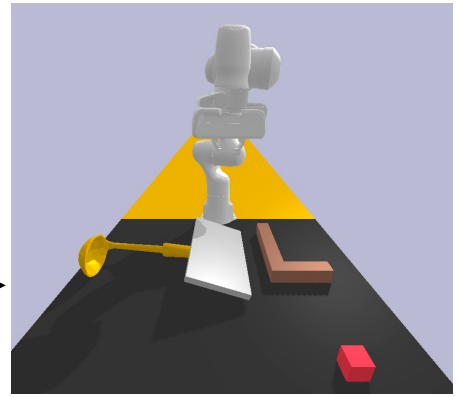
Contact Modes and Goals



Self-Play



Learned Contact Distributions



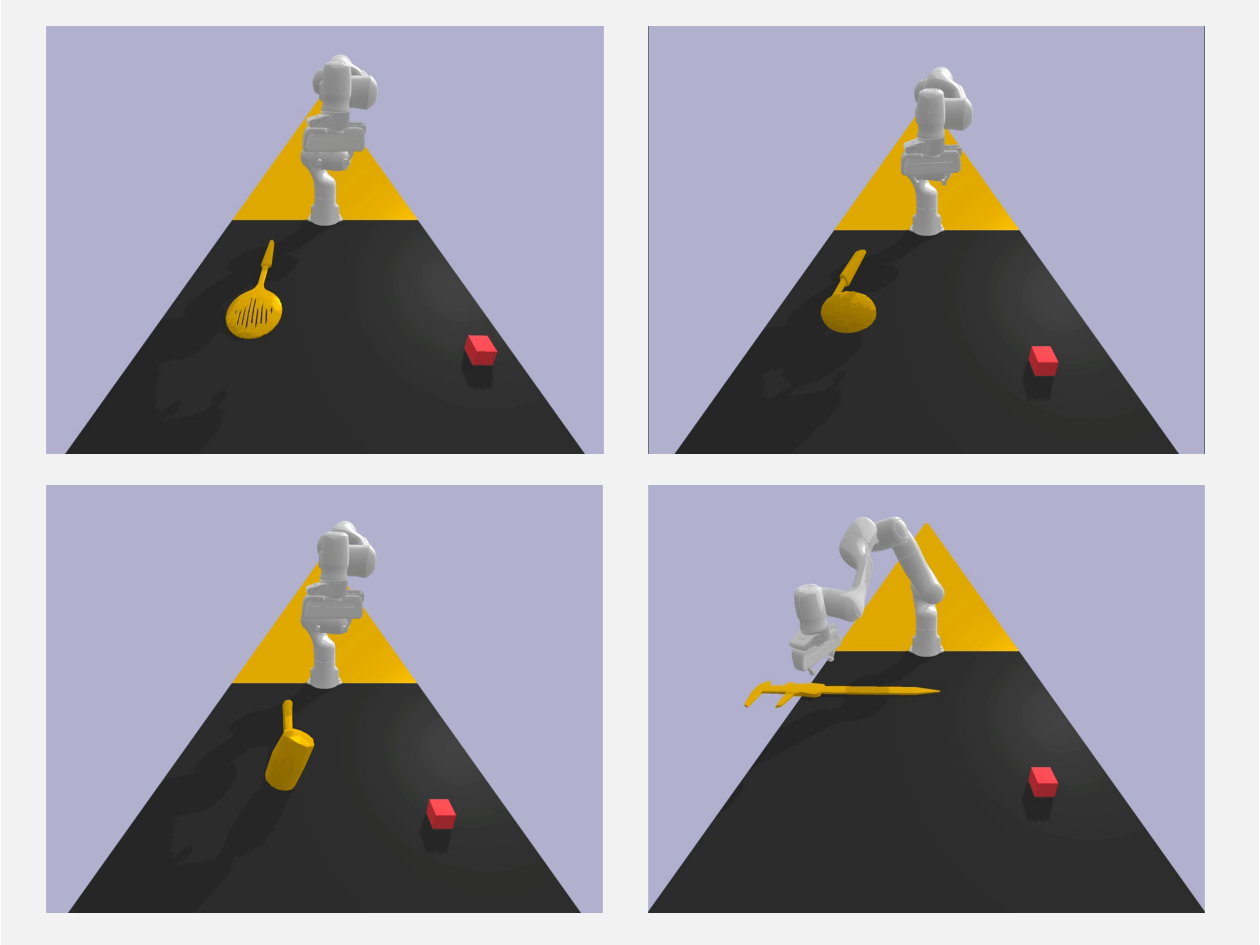
Goal:
Block on the Slope

Compositional Planning

Step 2: Learn Mechanism-Specific Samplers

We will learn those samplers (parameter generators) from self-plays.

Contact Modes and Goals →

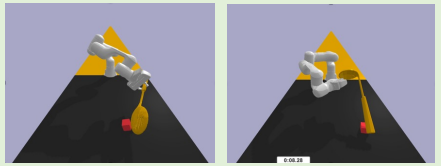


Self-Play with Randomly Sampled Objects and Poses

→

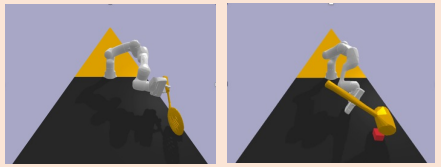
Dataset

Successful Trials



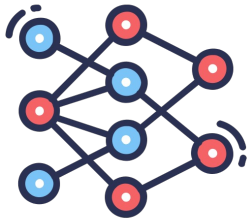
.....

Failed Trials



.....

↓

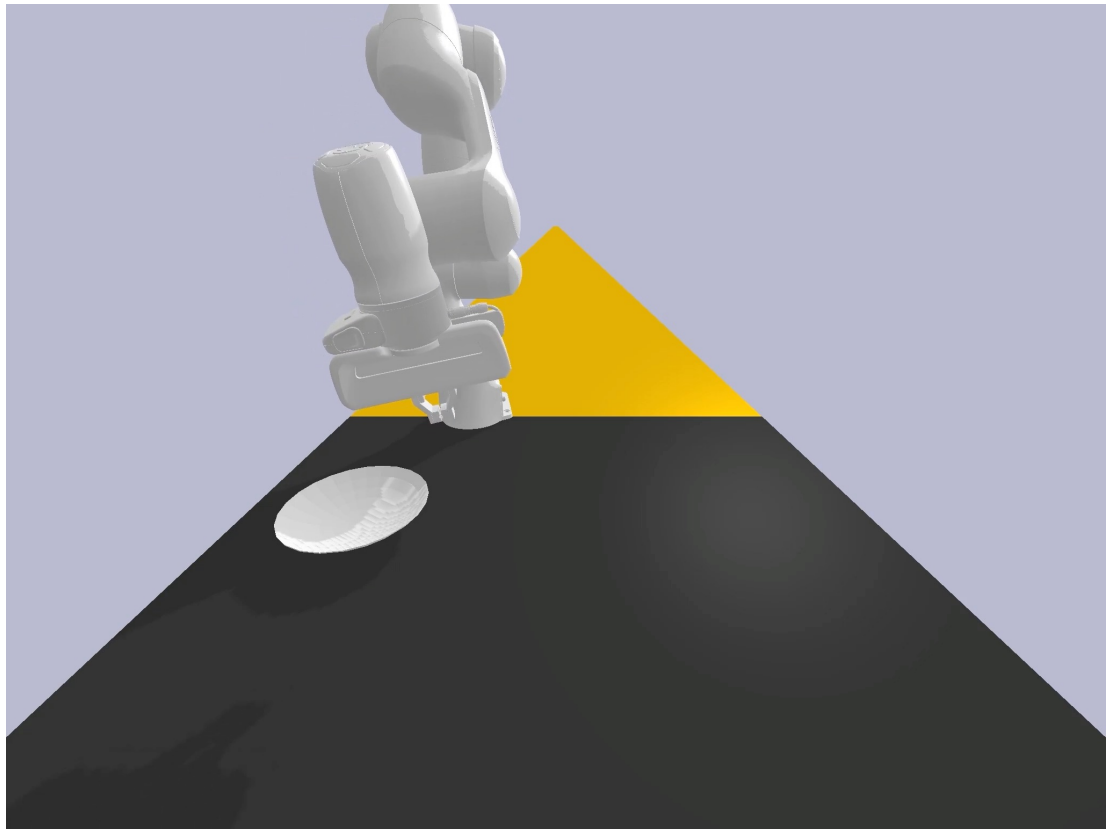


NN-Based Sampler

Learning Mechanisms Improves Efficiency

Method	Edge	Hook	Lever	Poking	CoM	Slope&Blocker
Basis Ops Only	89.45$\pm$5.53	>600	523.18$\pm$9.22	>600	19.30$\pm$2.82	>600
Ours (Macro+Sampler)	**0.57**$\pm$0.05	**3.84**$\pm$1.56	1.55$\pm$0.29	**97.76**$\pm$10.67	**0.97**$\pm$0.09	**4.11**$\pm$0.94

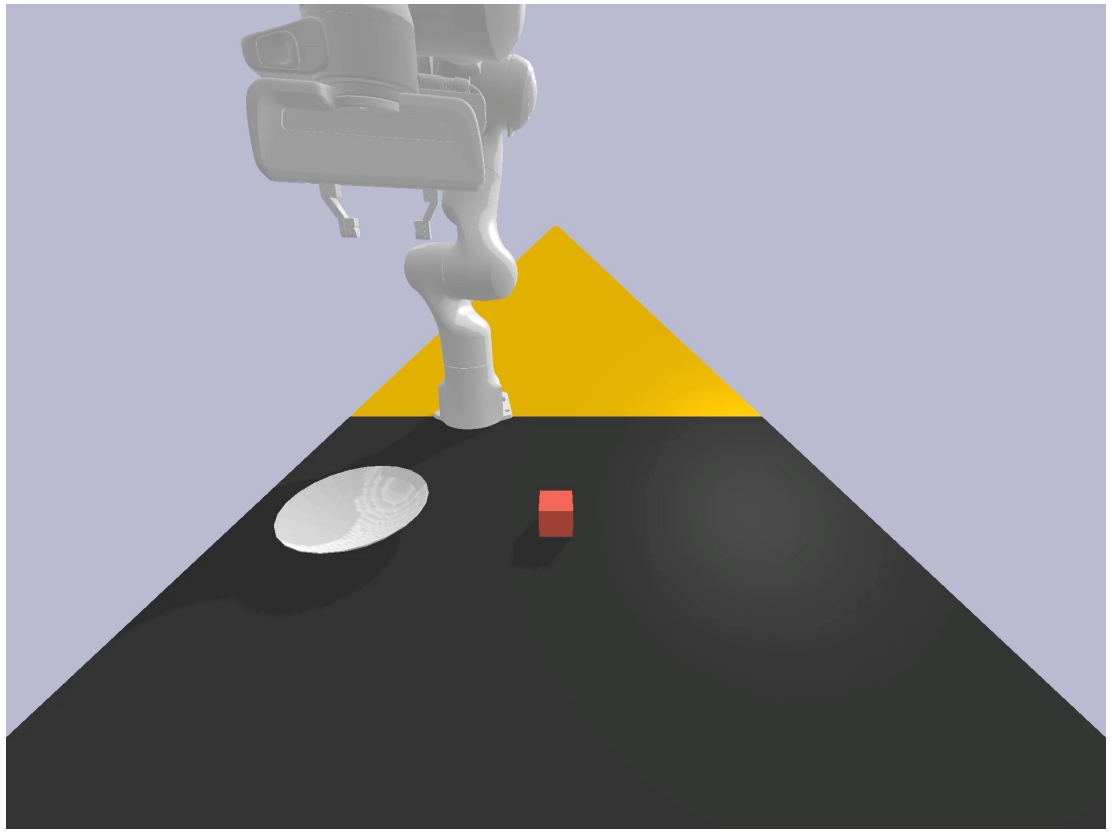
Learning Mechanisms Improves Efficiency



Goal:
`holding(plate)`

Method	Edge	Hook	Lever	Poking	CoM	Slope&Blocker
Basis Ops Only	89.45±5.53	>600	523.18±9.22	>600	19.30±2.82	>600
Ours (Macro+Sampler)	**0.57**±0.05	**3.84**±1.56	1.55±0.29	**97.76**±10.67	**0.97**±0.09	**4.11**±0.94

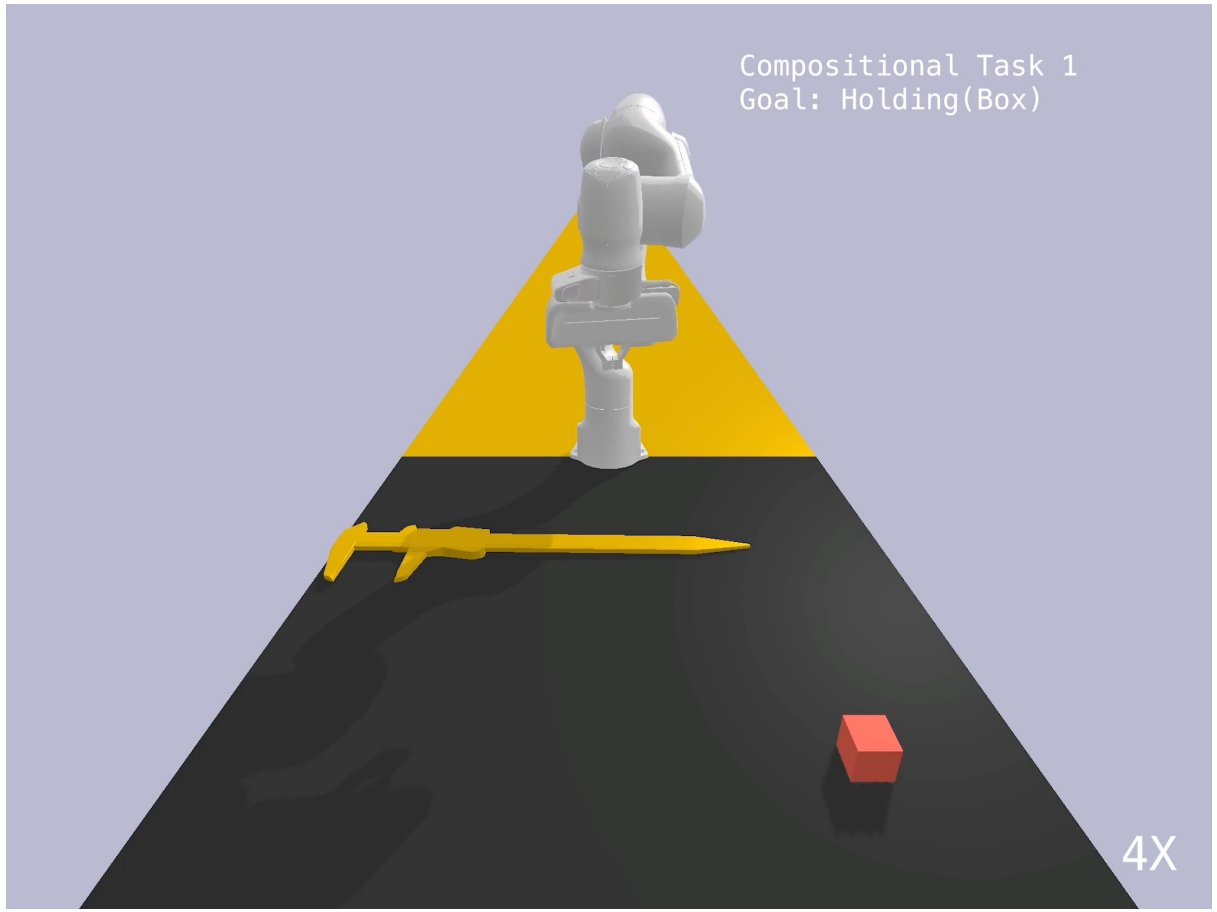
Learning Mechanisms Improves Efficiency



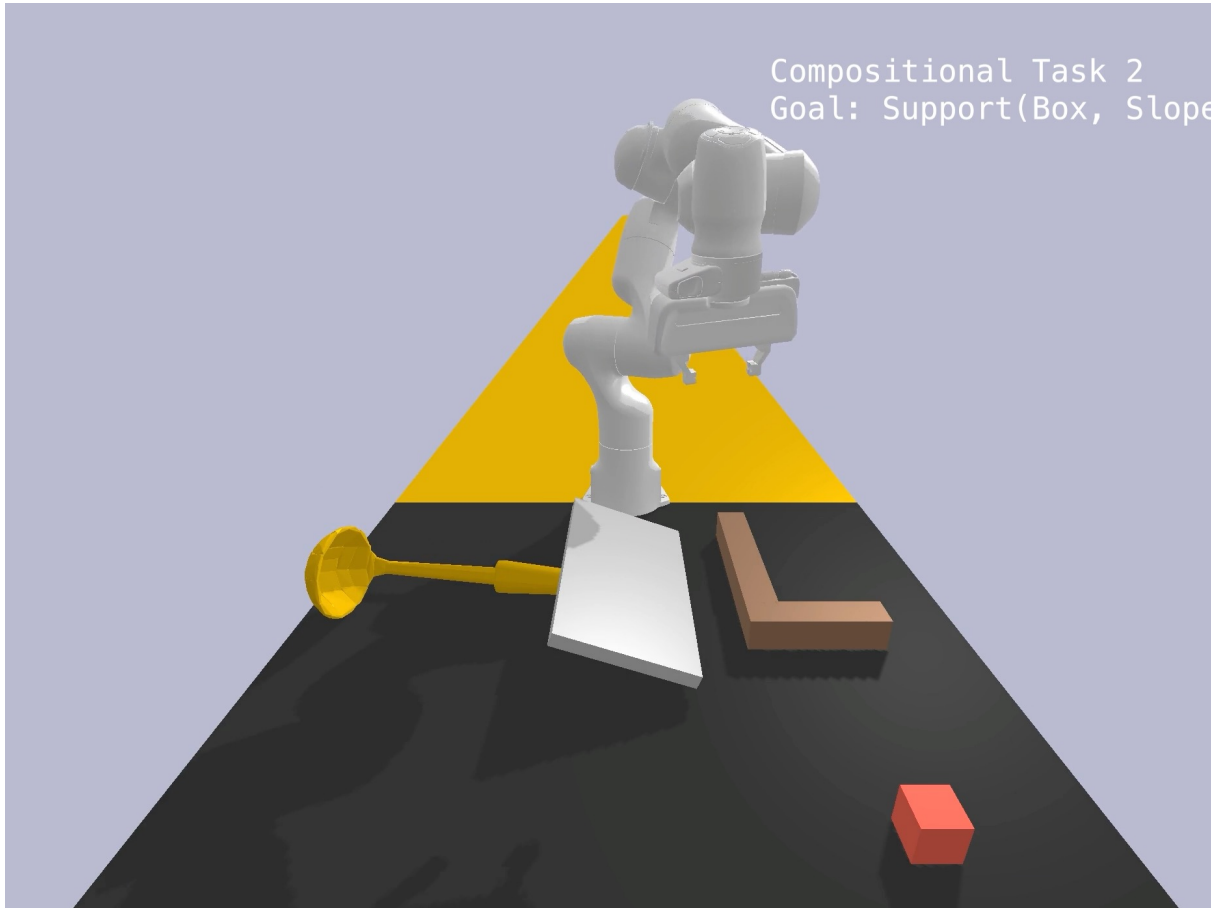
Goal:
`holding(plate)`

Method	Edge	Hook	Lever	Poking	CoM	Slope&Blocker
Basis Ops Only	89.45±5.53	>600	523.18±9.22	>600	19.30±2.82	>600
Ours (Macro+Sampler)	**0.57**±0.05	**3.84**±1.56	1.55±0.29	**97.76**±10.67	**0.97**±0.09	**4.11**±0.94

The learned mechanisms can be composed.

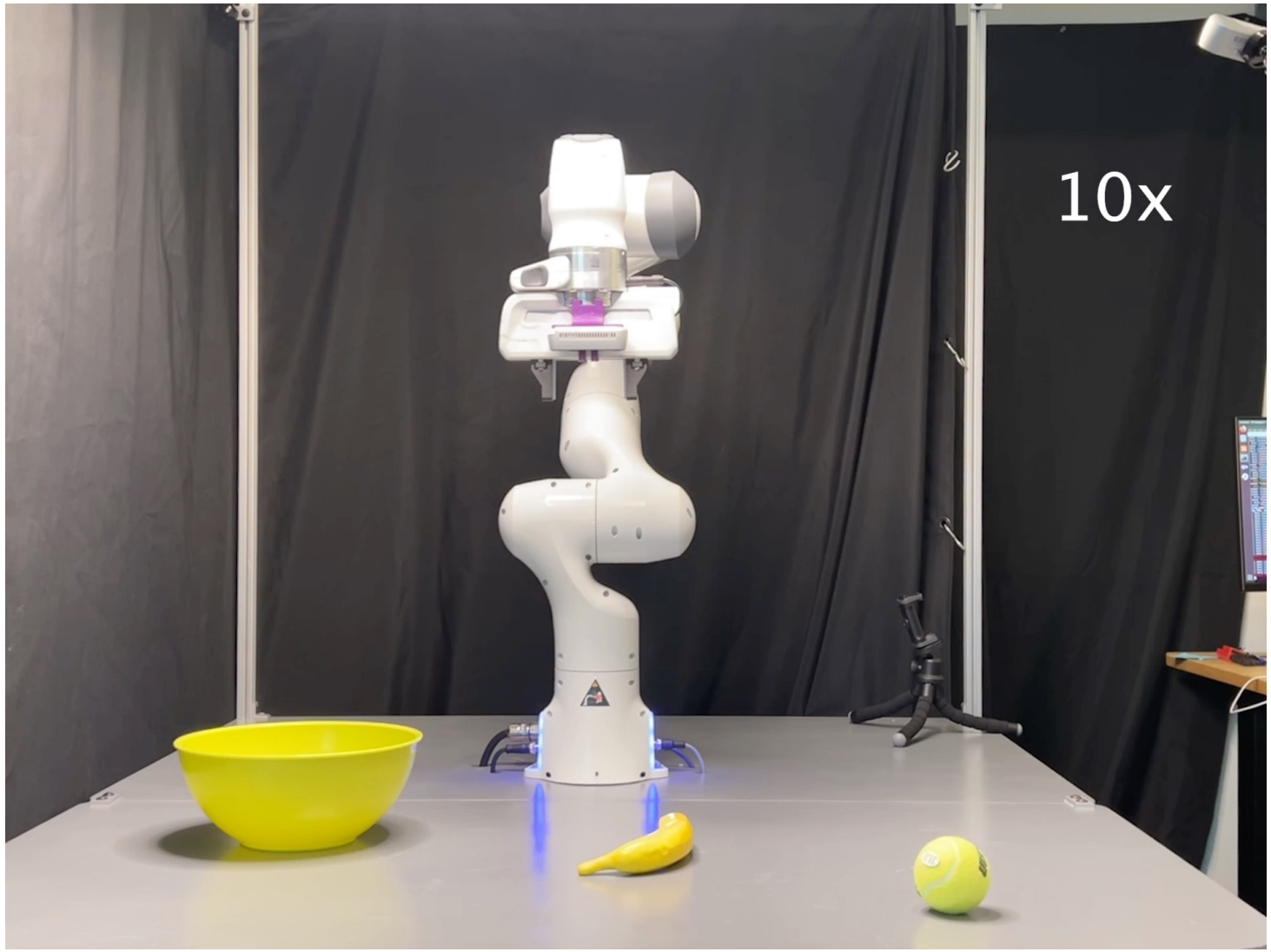


Goal: `holding(box)`
The caliper is too flat to be grasped.



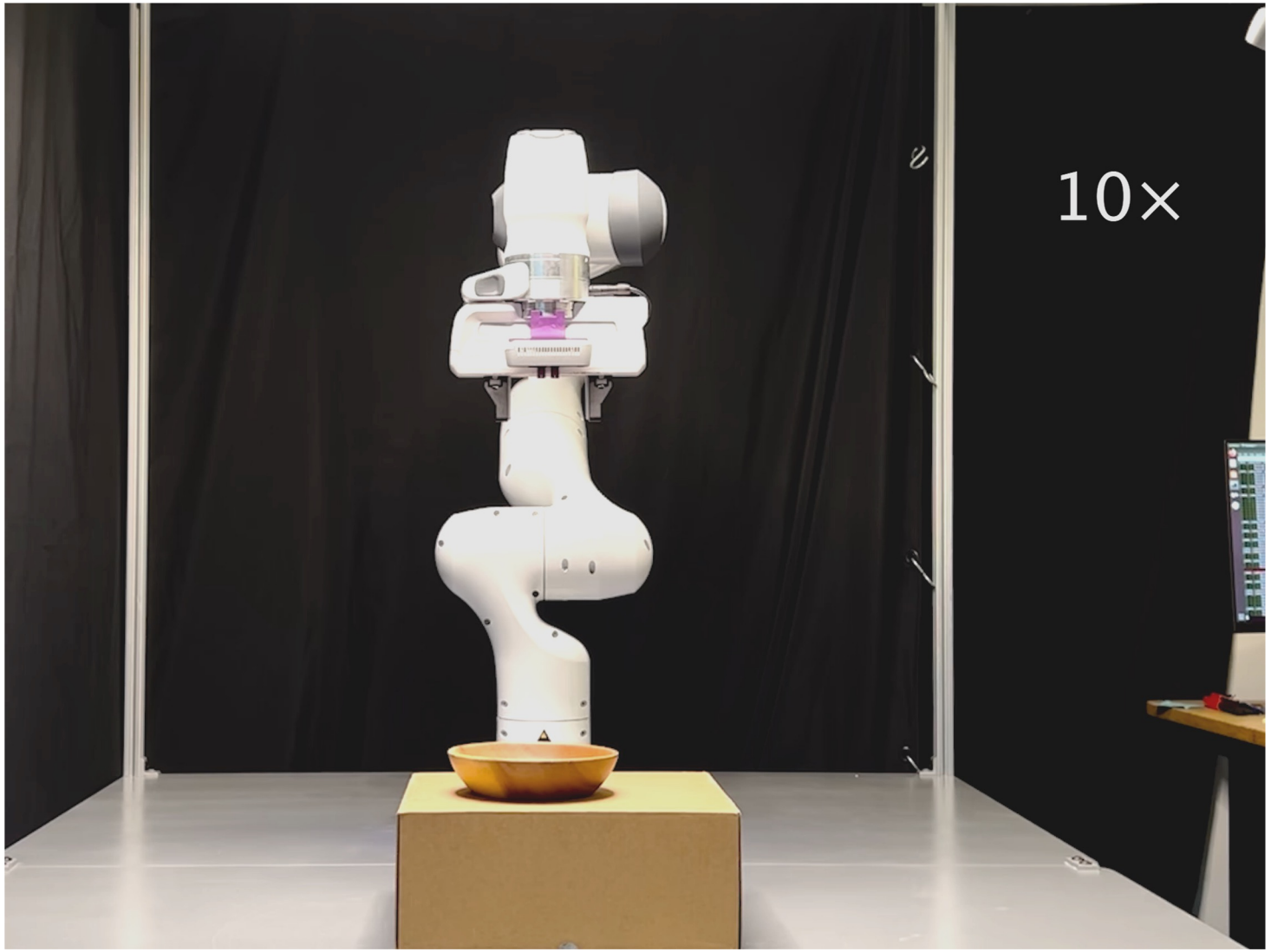
Goal: `on(box, ramp)`
Box may slide down the ramp.

...and they can be deployed on physical robots.



Goal: `in(ball, bowl)`
The robot uses the banana to hook.

...and they can be deployed on physical robots.



Goal: `holding(plate)`
The robot pushes it to the edge.

Learning Reusable Manipulation Strategies

Key idea: manipulation *mechanisms* can be modeled by
a sequence of contact relationships among objects.

- Principles:
  - contact mode sequences are generalizable.
  - learning specific samplers to accelerate planning.

- Advantages:
  - strong generalization.
  - strong compositionality.
  - strong extendability by including different basis operators.