Autonomous Robotics

Robot Learning from Demonstration

Josiah Hanna
University of Wisconsin — Madison

Slide credit: Some slides based on Sonia Chernova's course.

Announcements

Work on final projects.

- Extra credit chance: due Tuesday (1 week).
- Let me know if doing anything unusual in submission!

Final reading assignment due Monday.

Reading responses can still be completed.

Check your grades!

Learning Outcomes

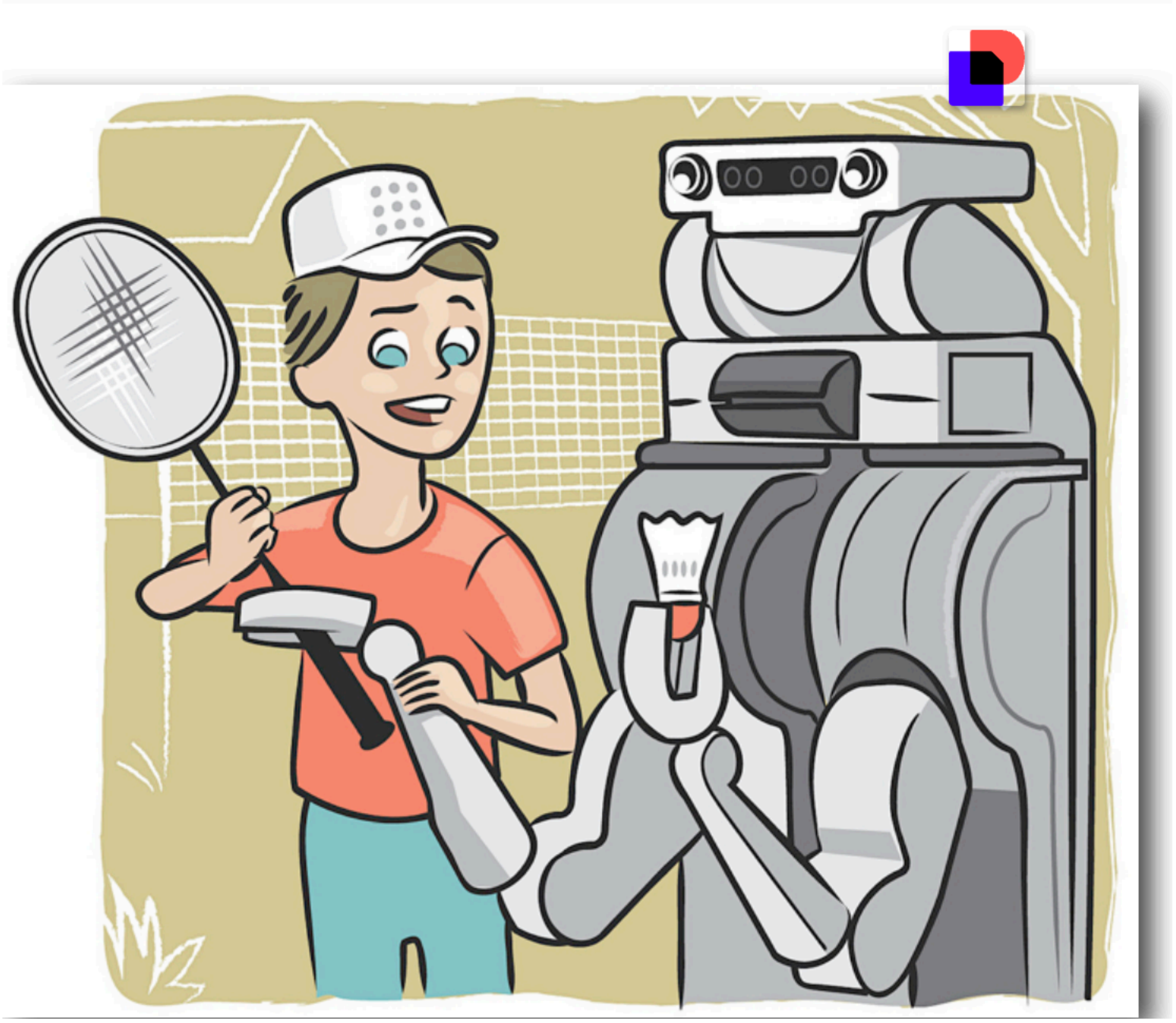
After today's lecture, you will:

- Understand the imitation learning problem setting.

- Be able to describe key challenges that arise when teaching a robot to perform a task.

Motivation

- No need for robot experts.
- Natural way to program robot skills.
- Add new skills on the fly.
- We have lots of data recording people doing things!



Terminology

- Imitation learning (IL)
- Behavior cloning (BC)
- Learning from Demonstration (LfD)
- Programming by Demonstration (PbD)
- Mimicry

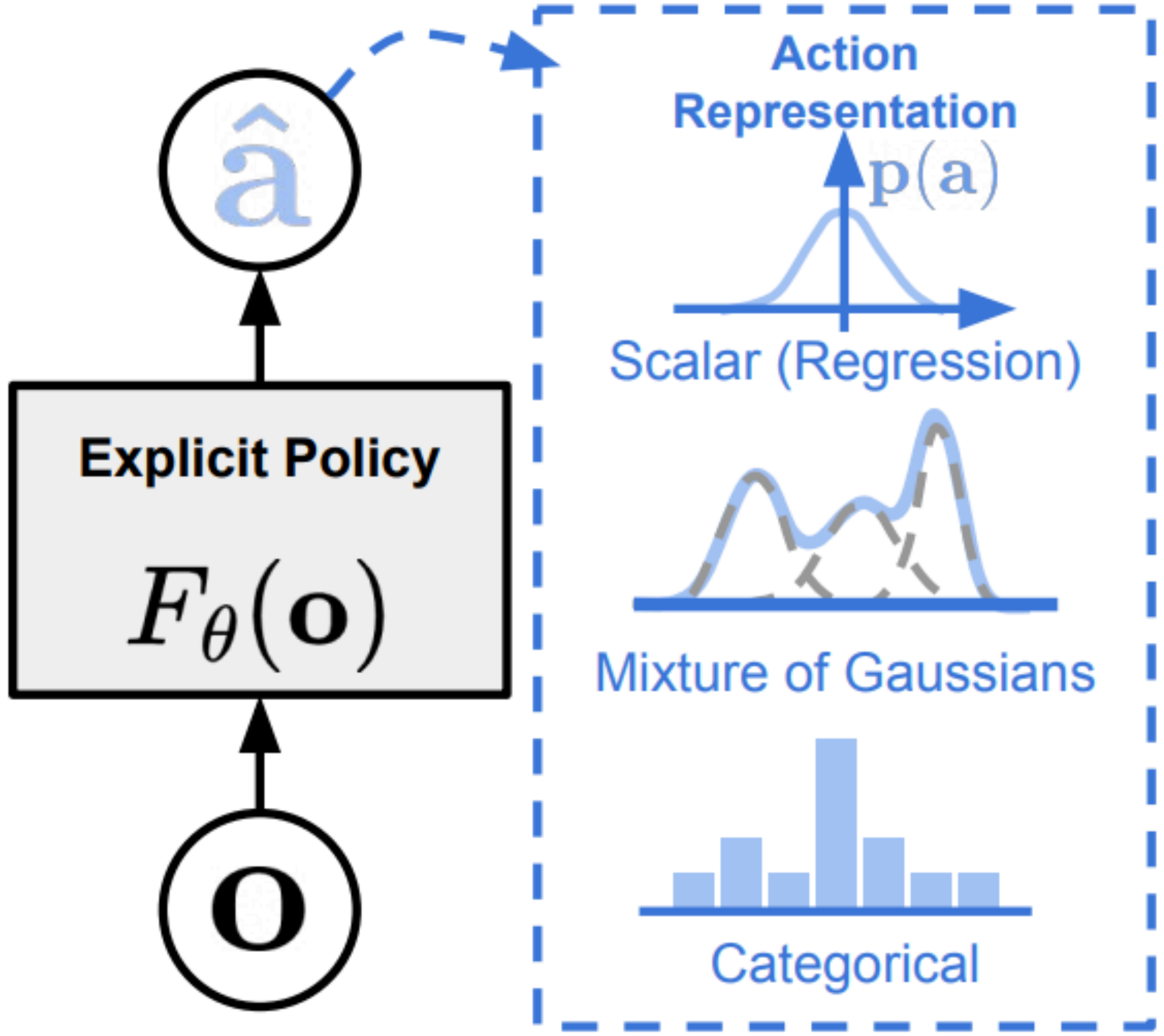
Formalism

- Basic formalism:
  - Given a dataset of the form $\{(s_i, a_i)\}_{i=1}^{m}$.
  - Goal: learn a policy π such that $\pi(s_i) \approx a_i$.
- Intuition: copy the behavior in the dataset.
- There are many variations of this basic setup.
  - Example: BC from Observation: given *state-only* trajectories $\{(s_1, \ldots s_T)\}$, learn a policy that reproduces these trajectories. Why is this hard?

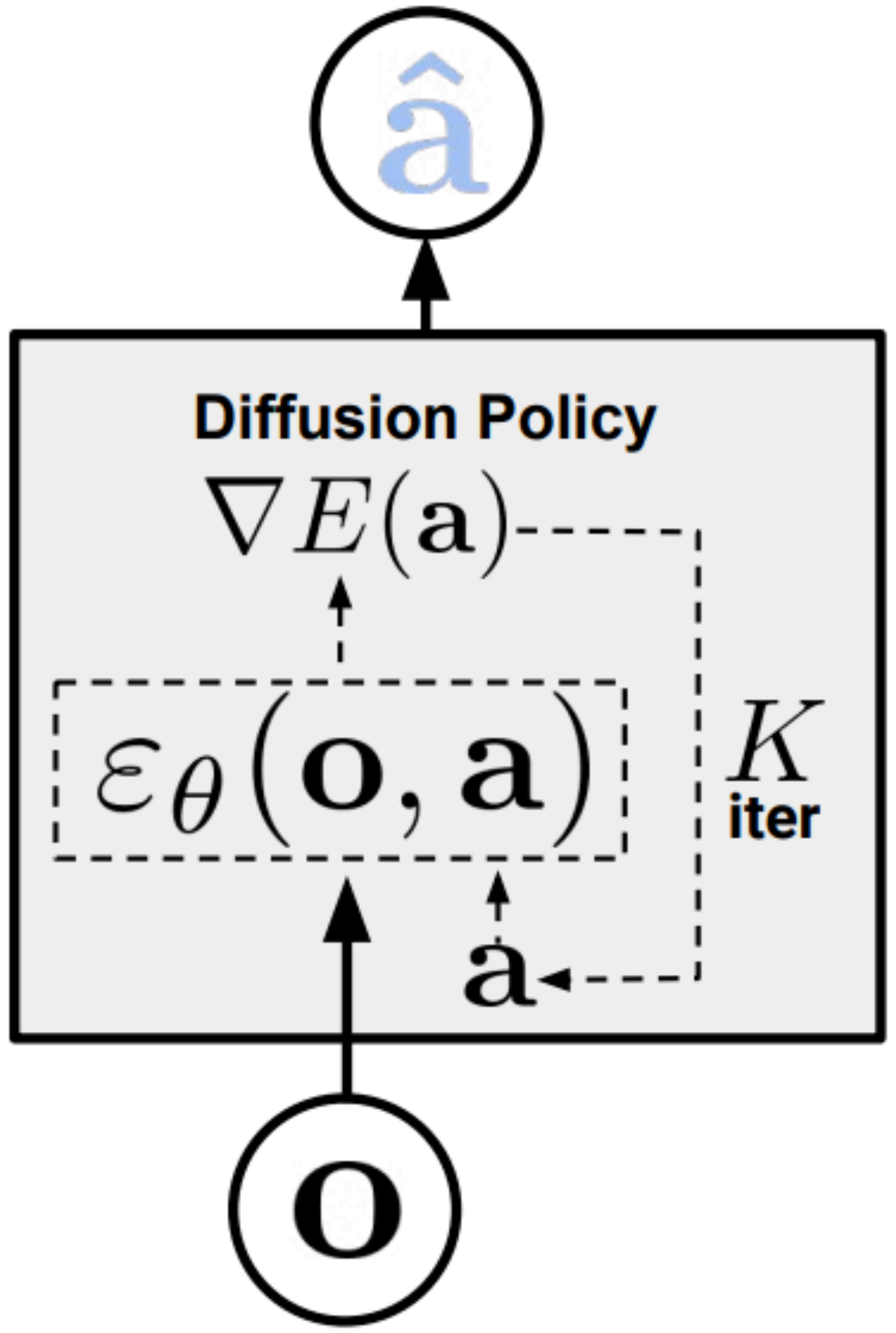
Basic Approach

- Imitation learning is a *supervised* learning problem.
  - States are inputs, demonstrator's actions are labels.
  - Can be classification or regression depending on control space.
- One issue: many supervised learning algorithms assume inputs are independent of each other during training and testing. Problem?
  - Small mistakes by the learner may compound over time.
  - Lack of consistency of actions across time-steps.

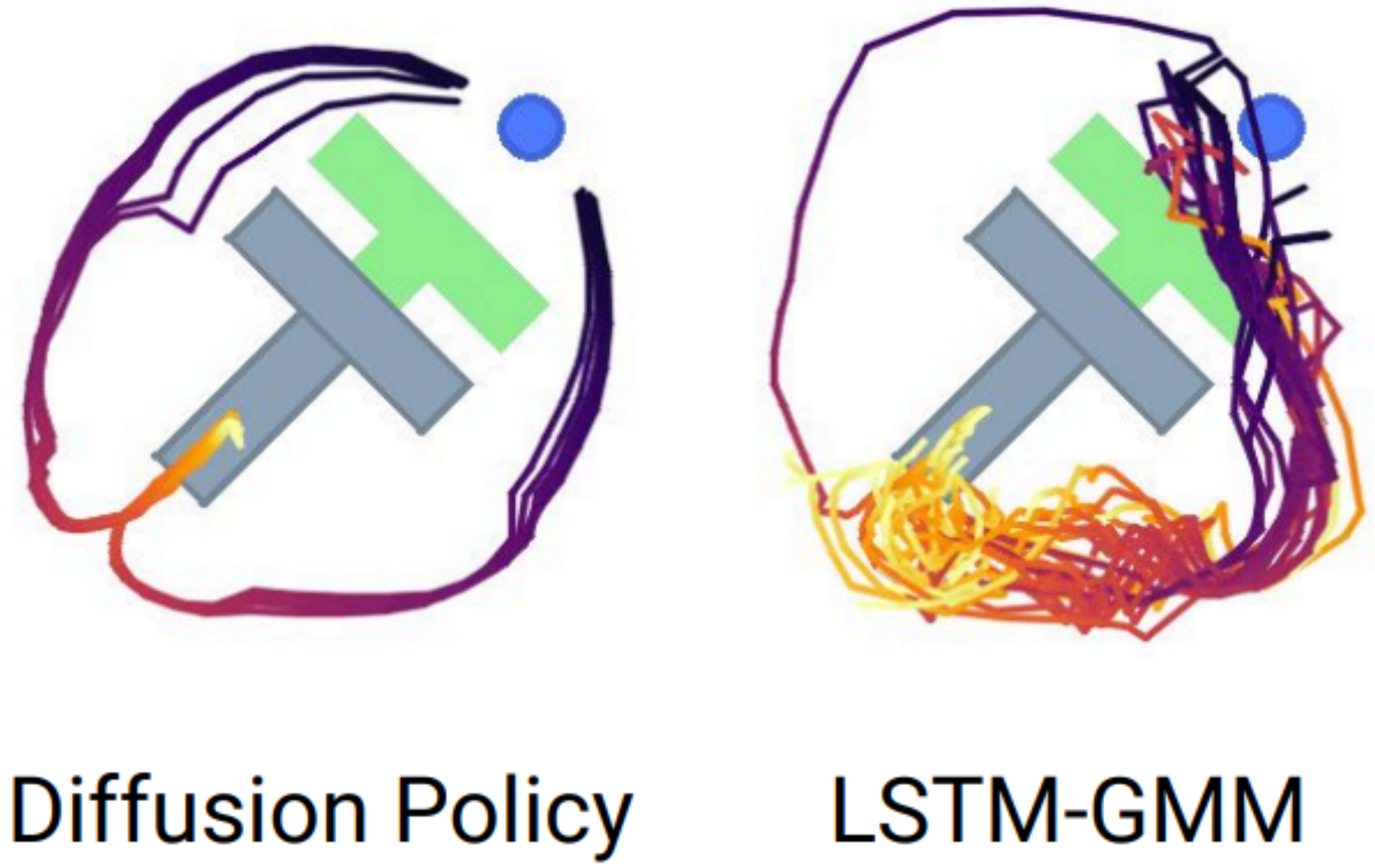
IL as Generative Modelling



Supervised Learning



Diffusion Policy

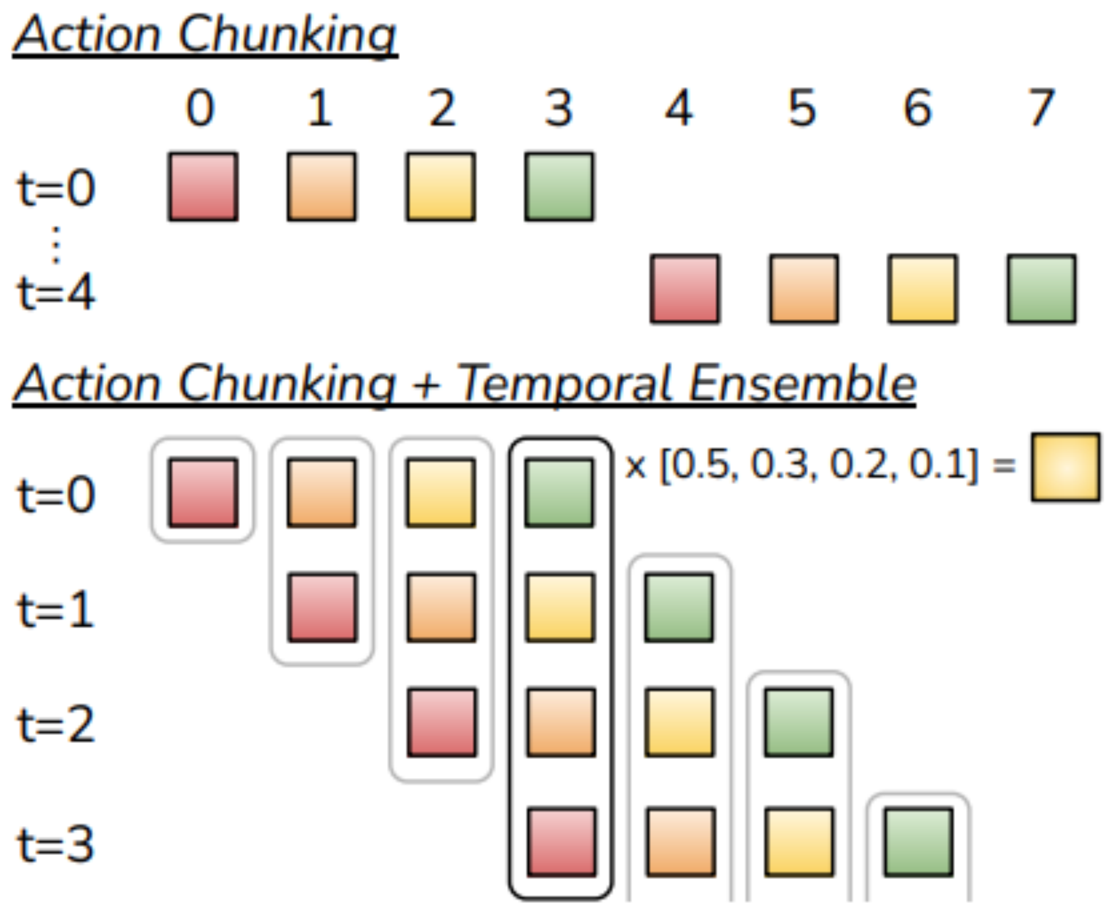


Multi-modal action distributions

Diffusion Policy: Visuomotor Policy Learning via Action Diffusion. Chi et al. 2024.

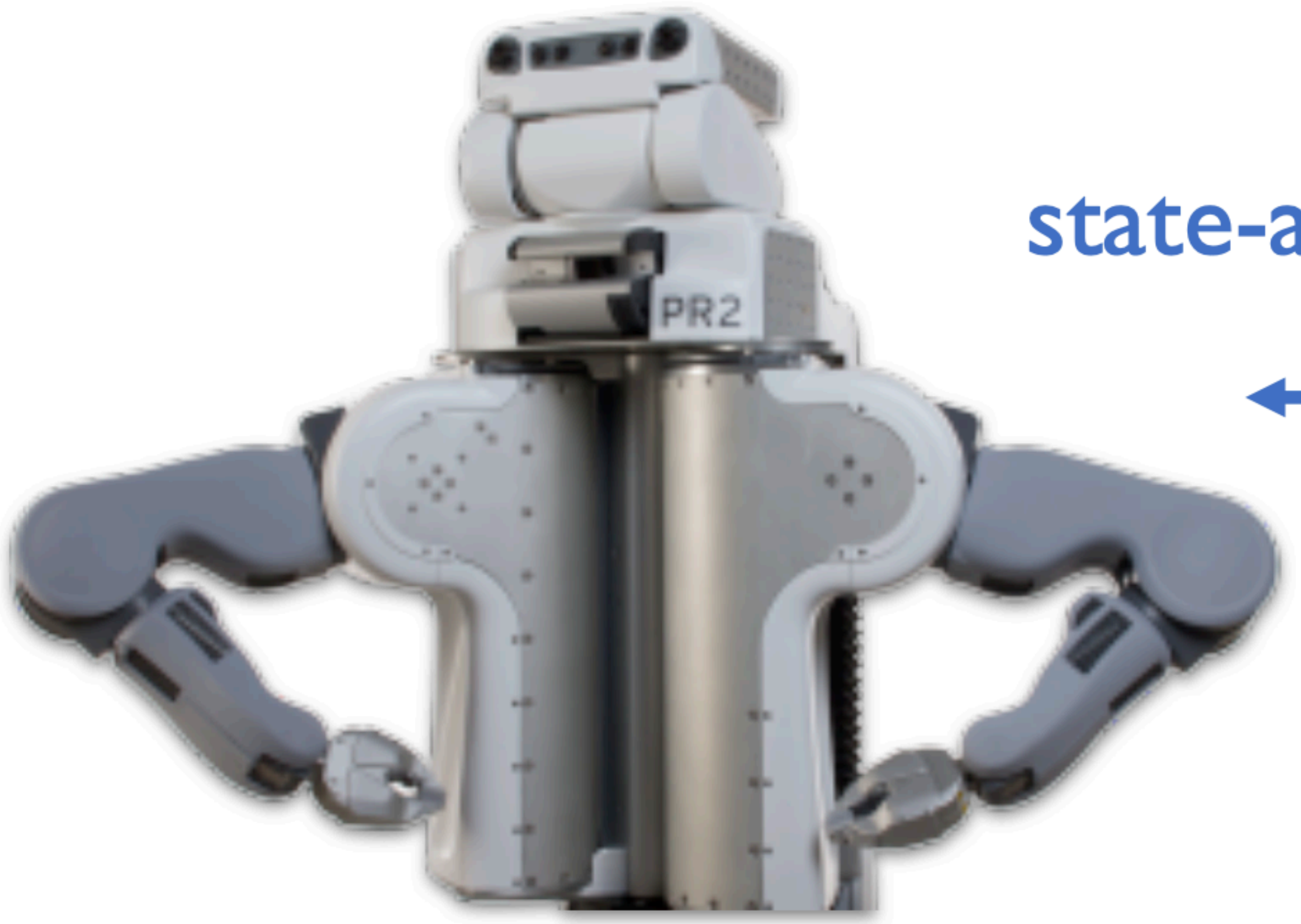
Action Chunking

- Basic set-up only predicts one action at a time.
  - Limitations: actions between time-steps are inconsistent with one another; increased latency.
- Solution: learn to predict a sequence of actions (action chunks).

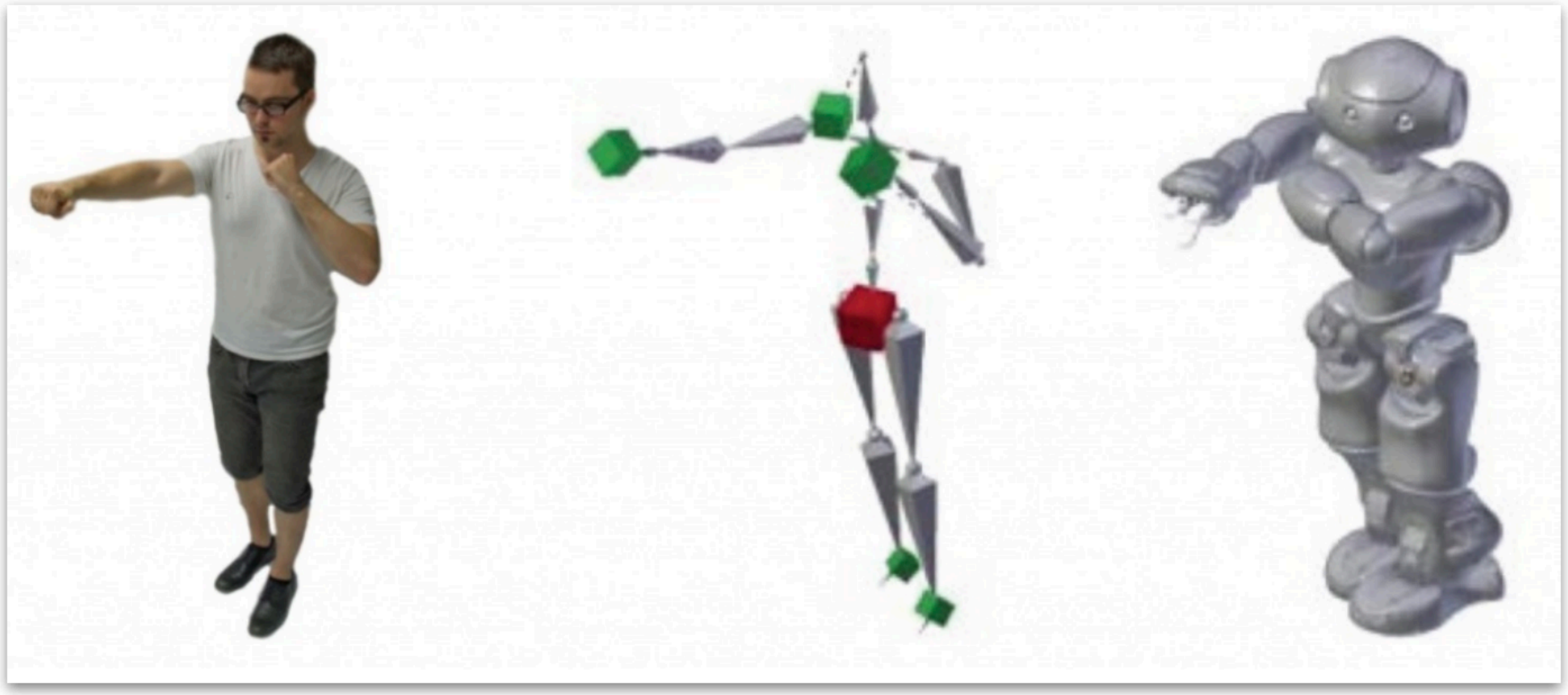


Learning Fine-Grained Bimanual Manipulation with Low-Cost Hardware. Zhao et al. 2023.

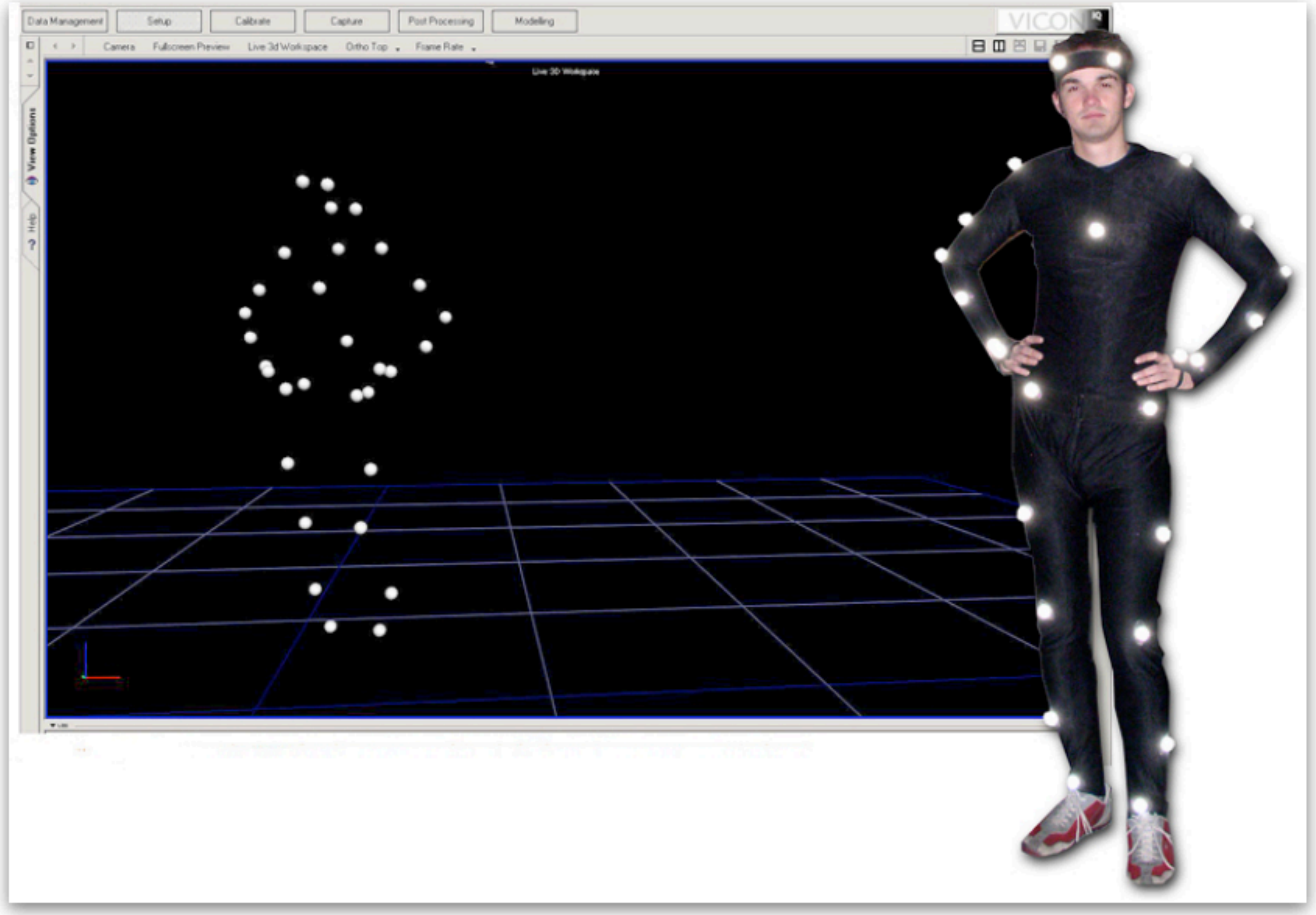
Correspondence Problem



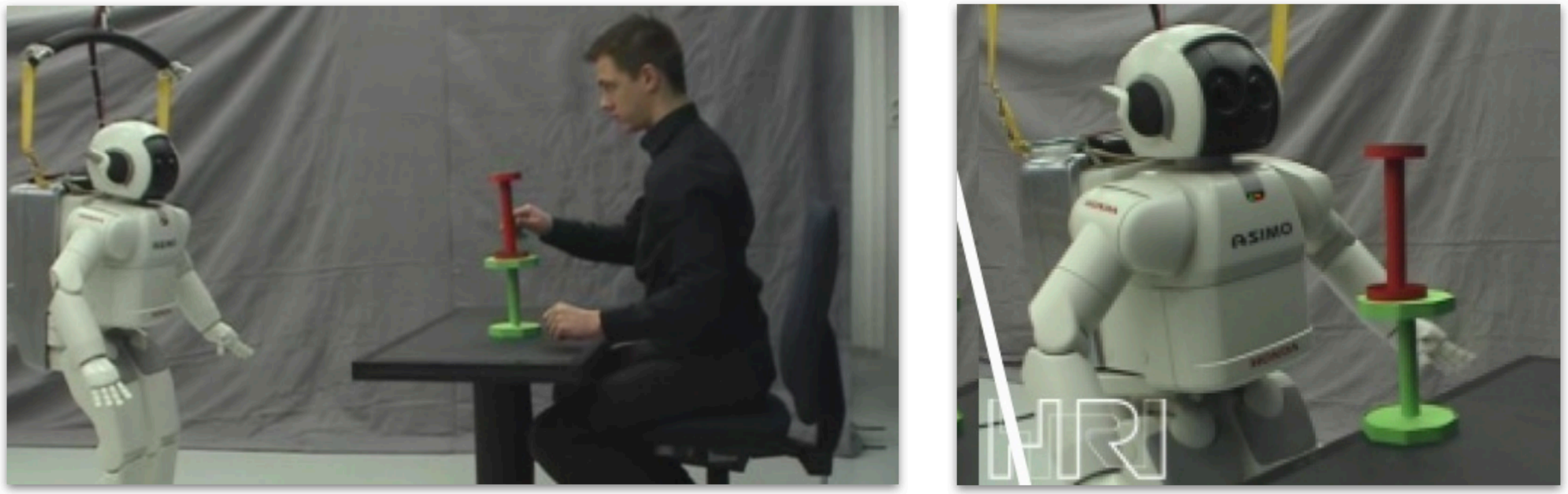
state-action mapping?



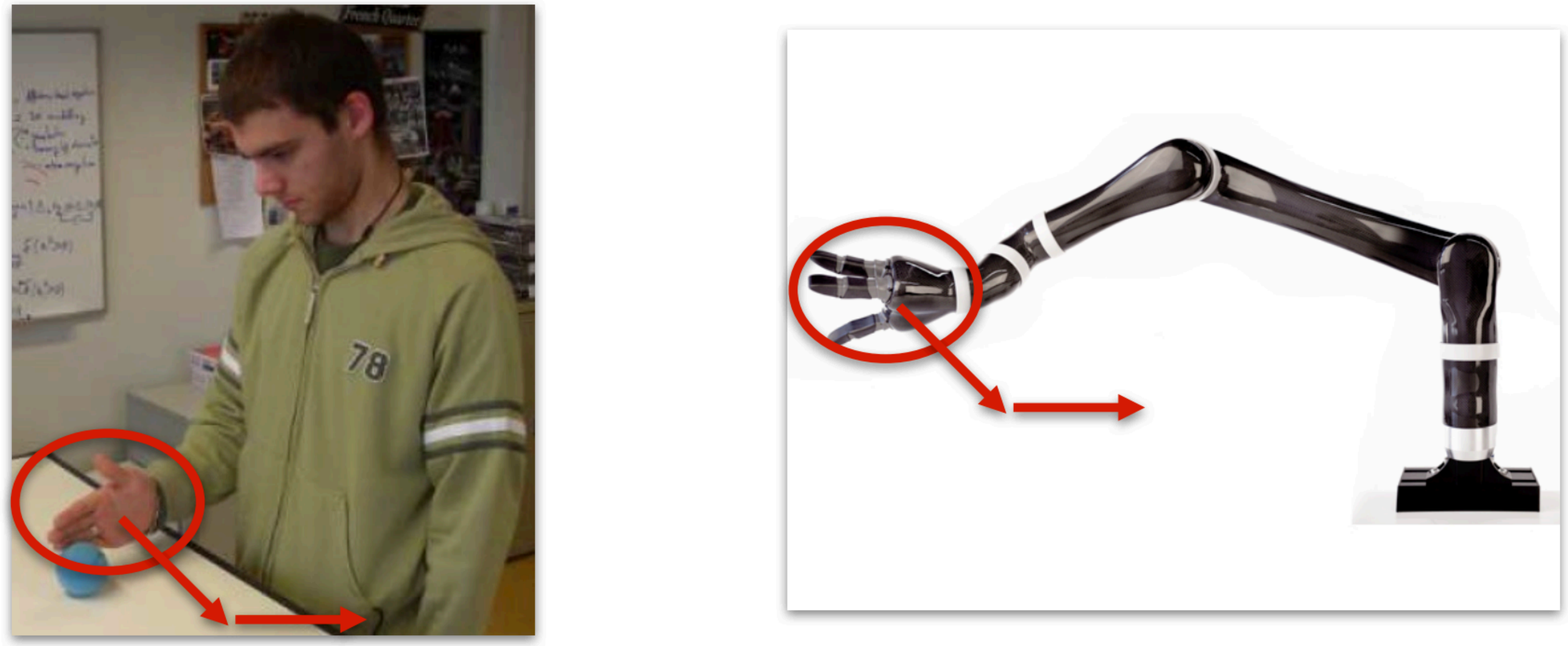
Motion Capture



Mapping between Humans and Robots



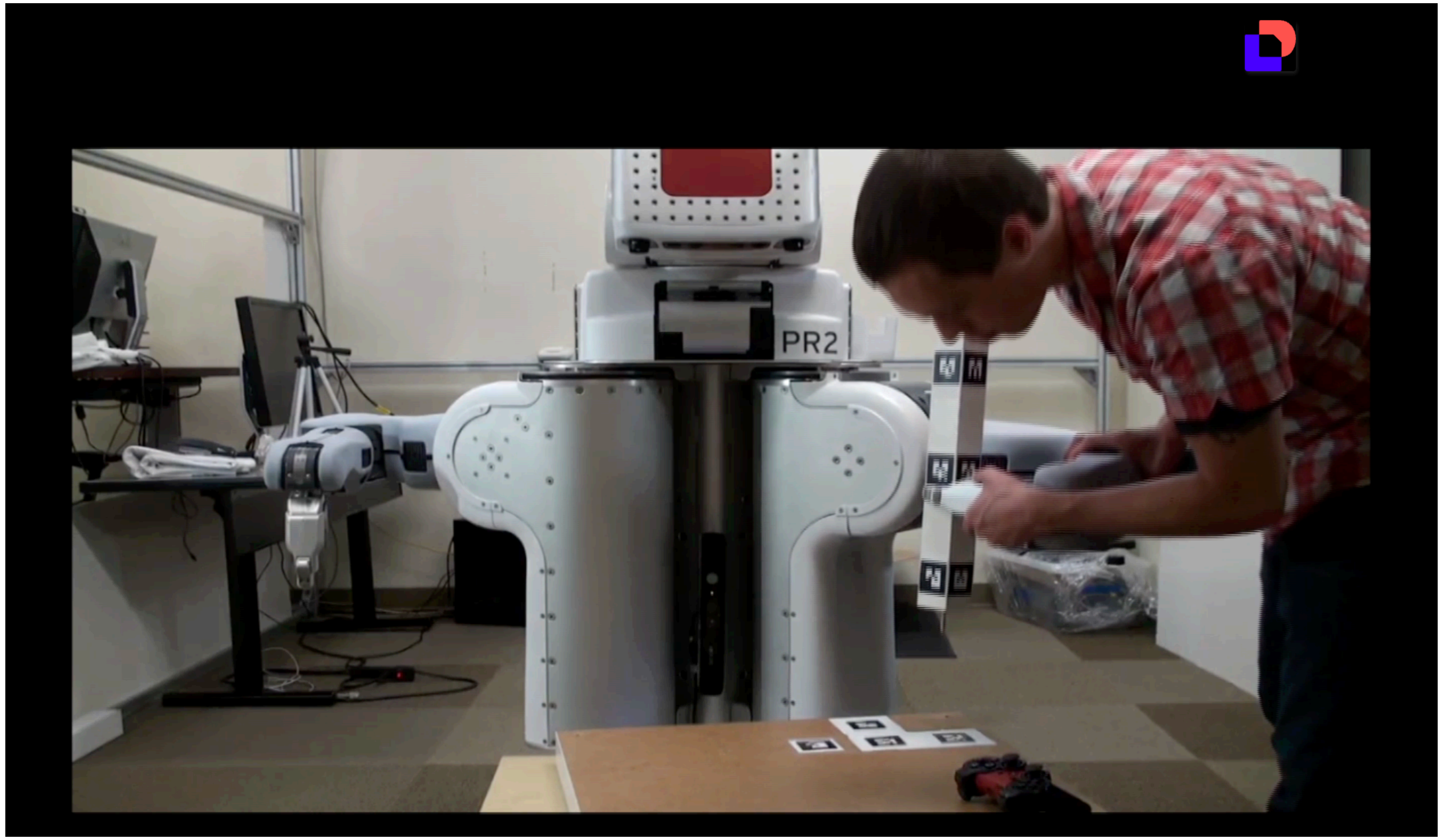
Object-based



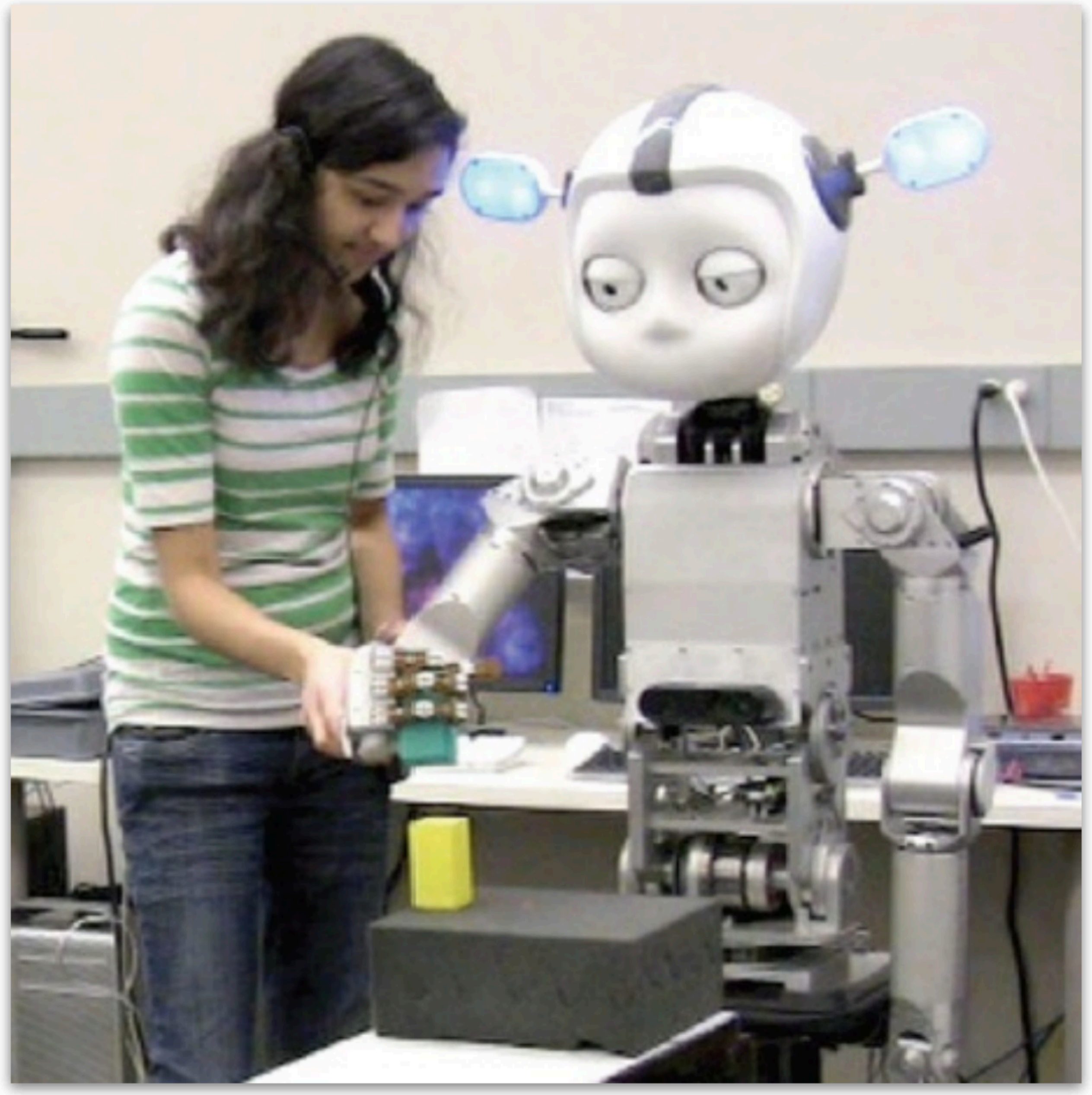
End effector-based

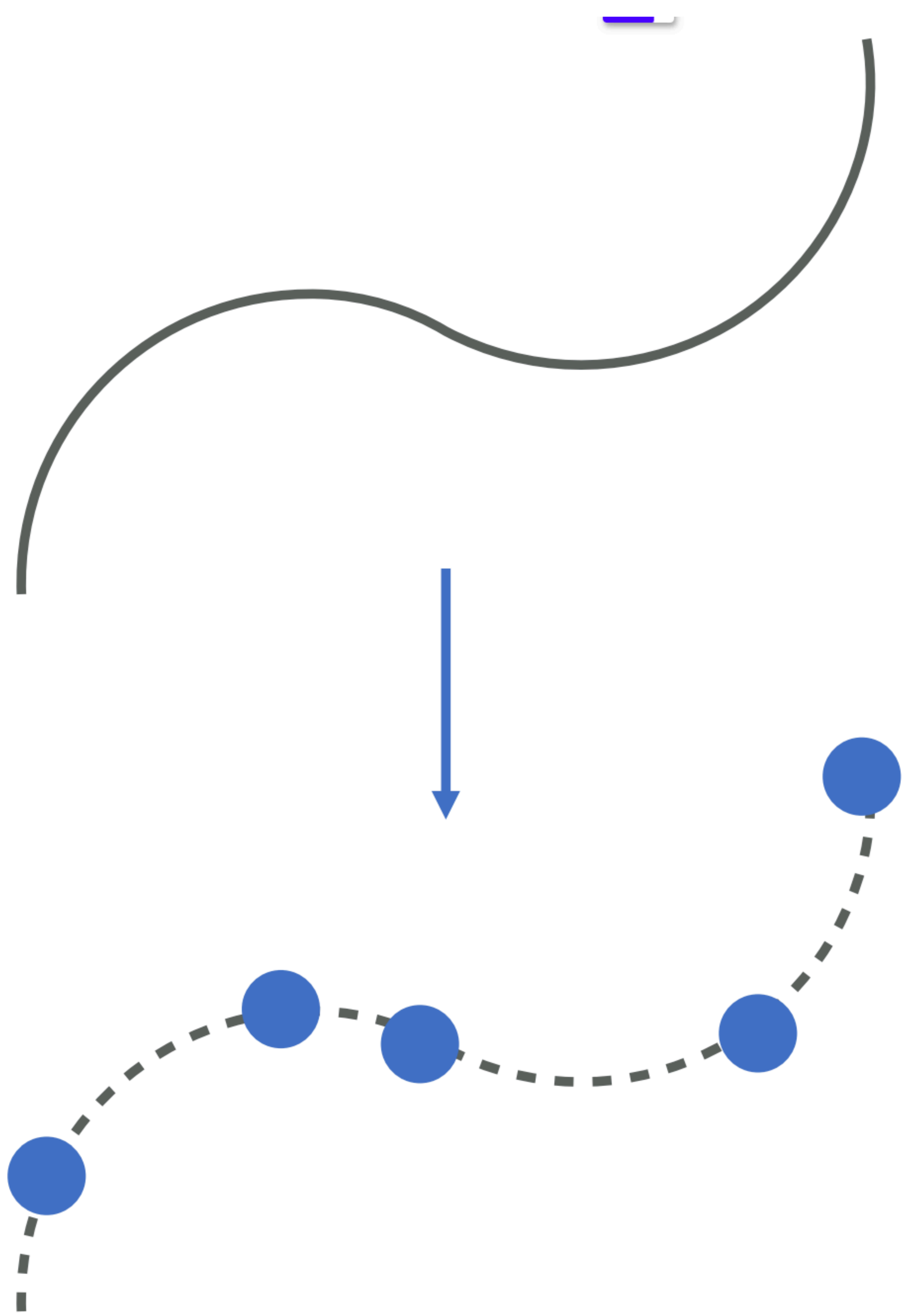
Teleoperation

Kinesthetic Teaching



Keyframe Demonstrations

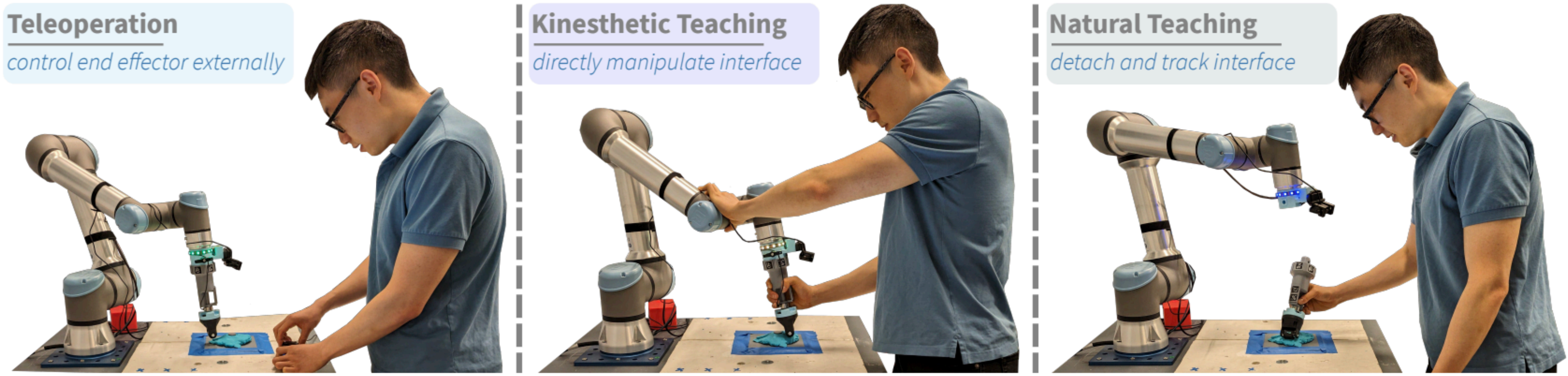




Interface Design

Mobile ALOHA: Learning Bimanual Mobile Manipulation with Low-Cost Whole-Body Teleoperation

Fu et al. 2024.

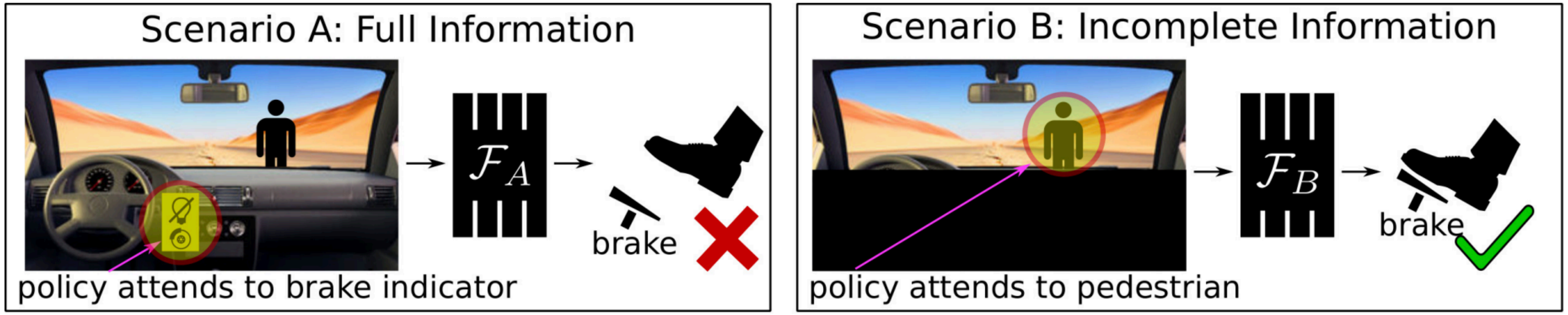


Versatile Demonstration Interface: Toward More Flexible Robot Demonstration Collection.
Hagenow et al. 2025.

Causal Confusion

- The demonstrator is running some policy that depends on some state variables that the observer can observe in the world.

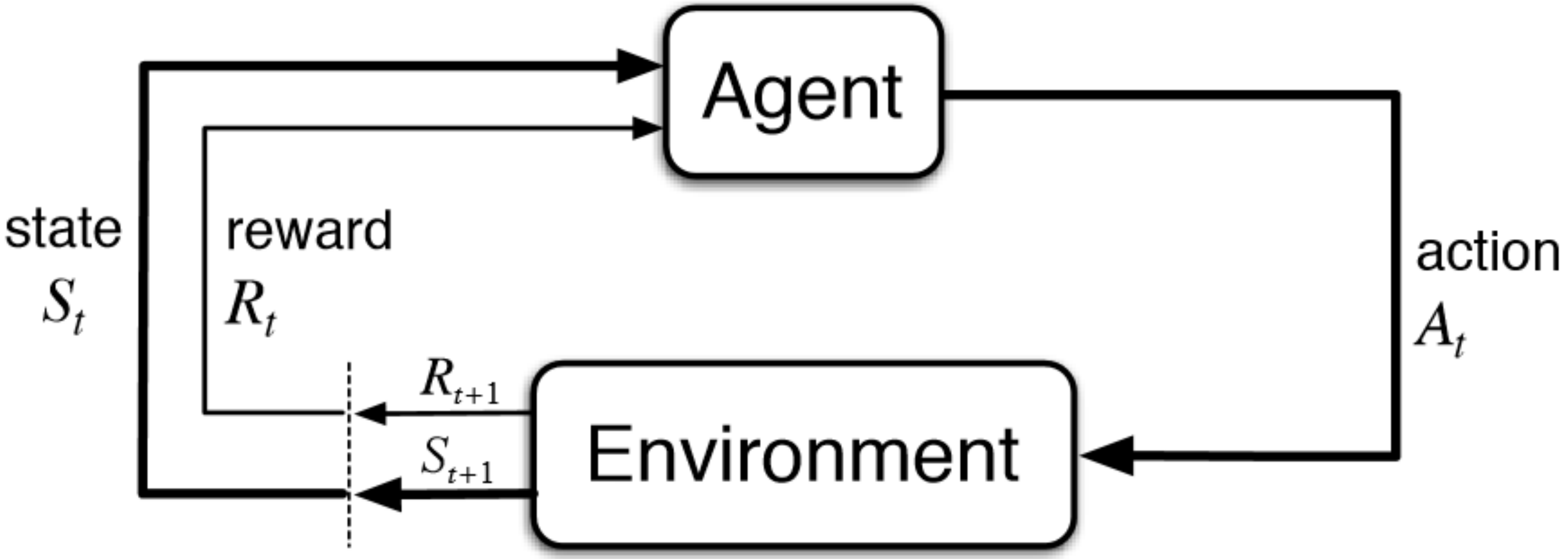
- Depending on its sensors, the robot can observe / estimate some set of variables that may be different than the demonstrators.



Inverse Reinforcement Learning

- Basic formalism:
  - Given a dataset of the form $\{(s_i, a_i)\}_{i=1}^m$.
  - Goal: learn a reward function such that $\pi^\star(s) \approx a$, where $\pi^\star$ is the optimal policy for the learned reward function.
- Intuition: learn demonstrator's intention rather than simply mimicking their actions. Why useful?
  - Potentially generalizes better to new scenarios.
  - Can attempt to surpass the demonstrator.

Inverse Reinforcement Learning



$$\ldots S_t, A_t, S_{t+1}, A_{t+1}, \ldots$$

$$S_{t+1} \sim p(\cdot \mid S_t, A_t)$$

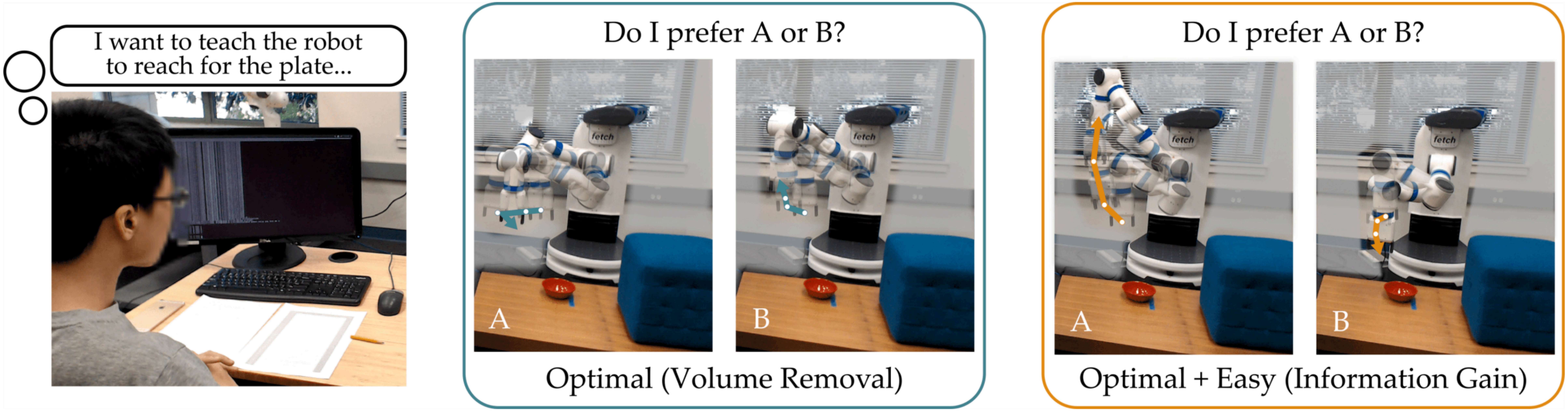
$$A_{t+1} \leftarrow \pi(S_{t+1})$$

Inverse Reinforcement Learning

Helicopter tricks

[Abbeel et al. 2007]

Inverse Reinforcement Learning



Asking Easy Questions: A User-Friendly Approach to Active Reward Learning. Biyik et al. 2020

Summary

Today we covered:

1. The imitation learning problem, issues, and approaches.

2. Introduced and discussed inverse reinforcement learning.

Action Items

Final project.

Begin final reading!