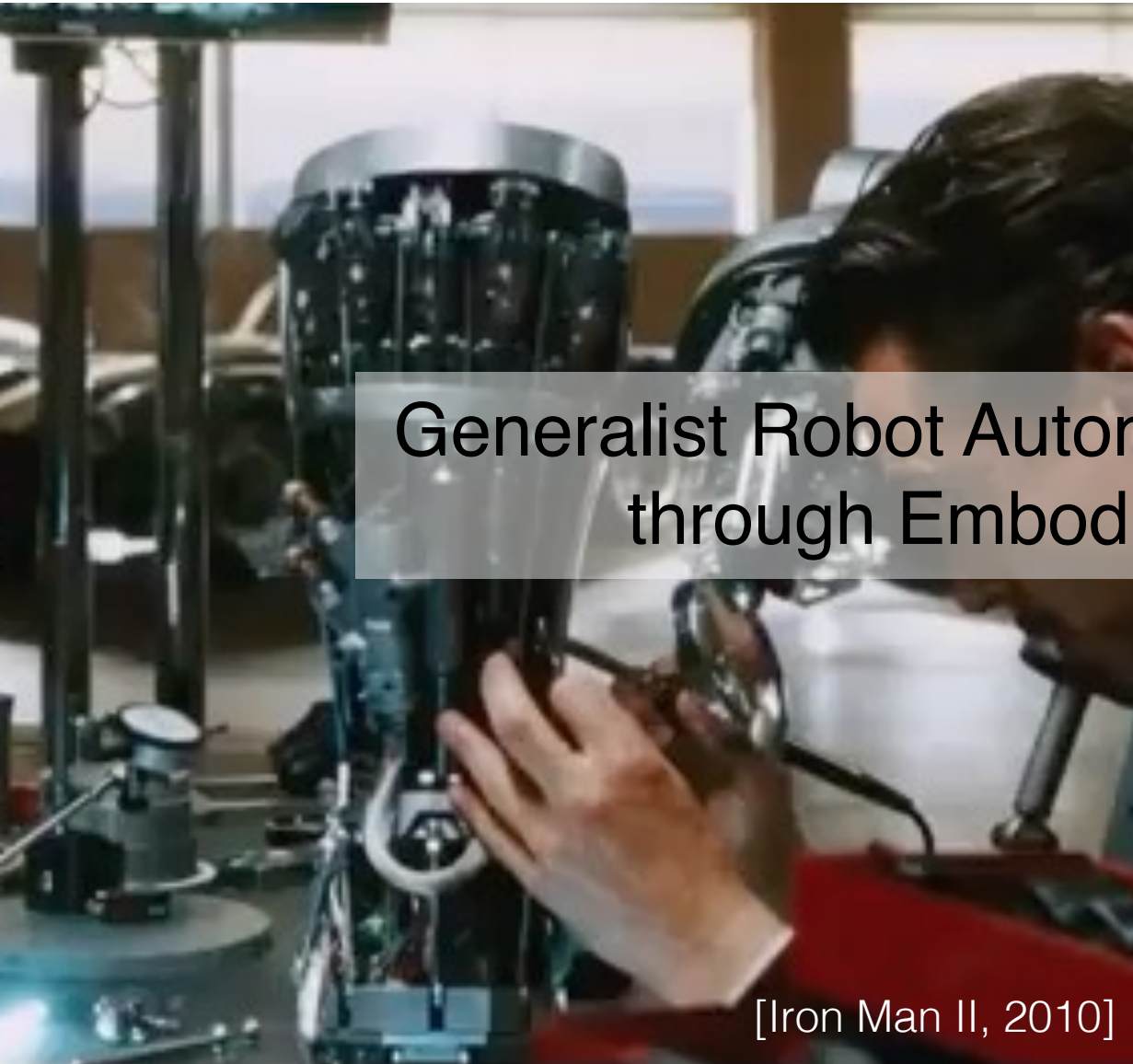


Embodiment-Aware Skill Learning for Diverse Robots

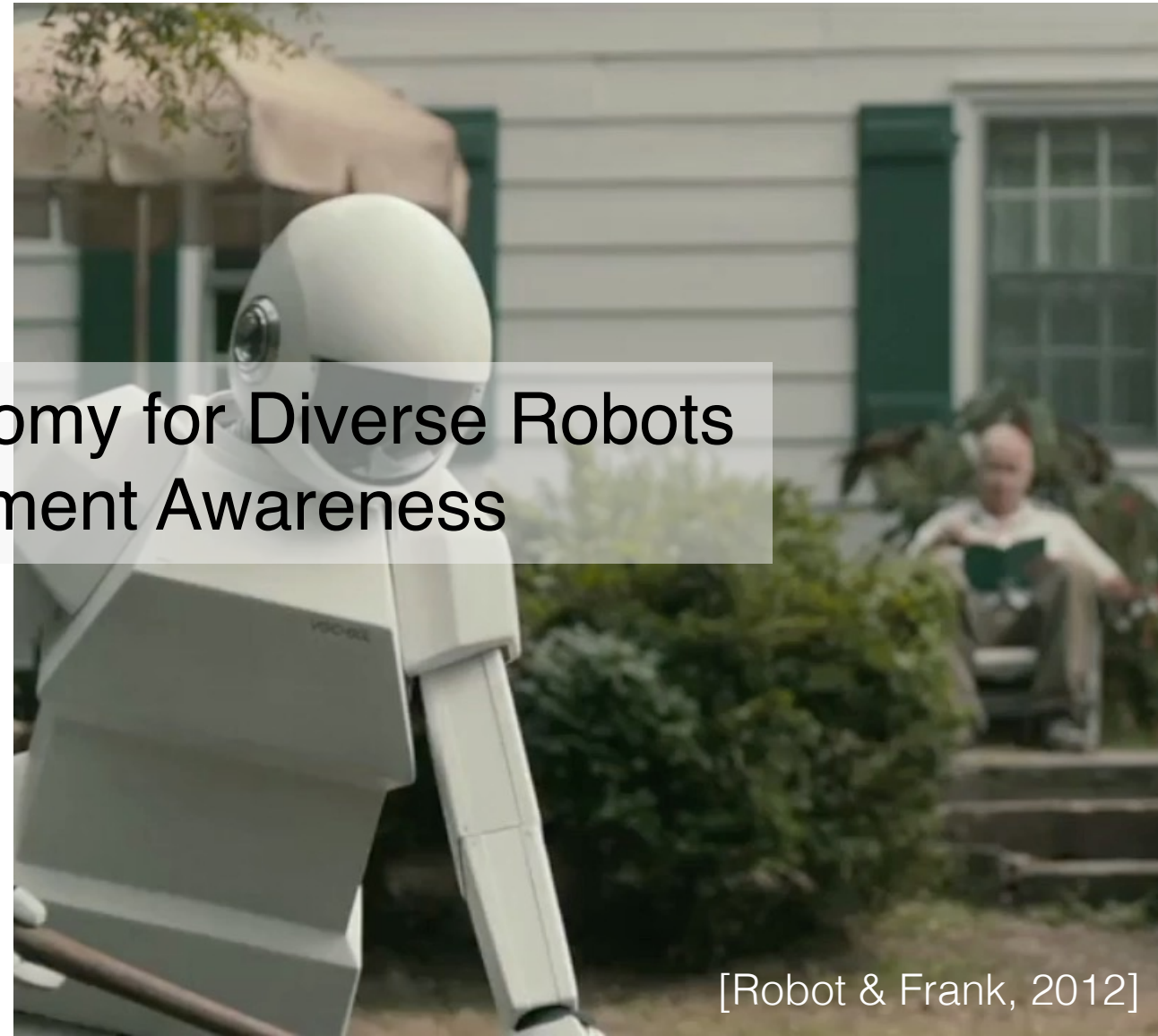
Mingyo Seo
The University of Texas at Austin



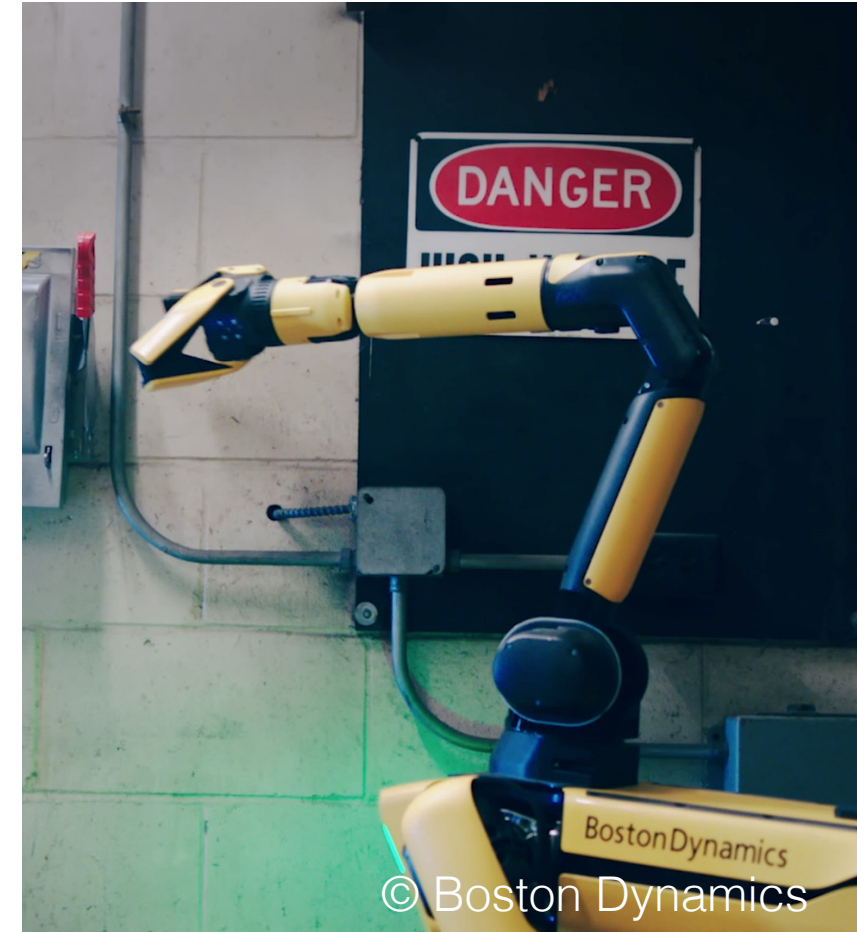
A close-up shot of Tony Stark, played by Robert Downey Jr., working on the chest piece of his Iron Man armor. He is in a workshop, and the armor is a complex, dark, metallic structure with various mechanical components visible. The lighting is focused on the armor and Tony's hands.

Generalist Robot Autonomy for Diverse Robots through Embodiment Awareness

[Iron Man II, 2010]



Diverse Robots Designed to Support Human Activities



Transition to Data-Driven Approaches



Ruled-Based Approaches

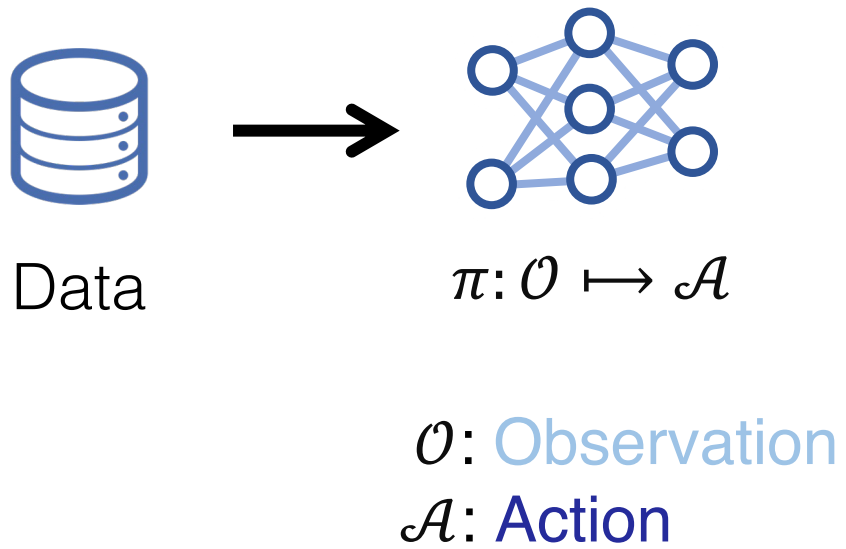
- Task-specific programming
- Limited to constrained settings



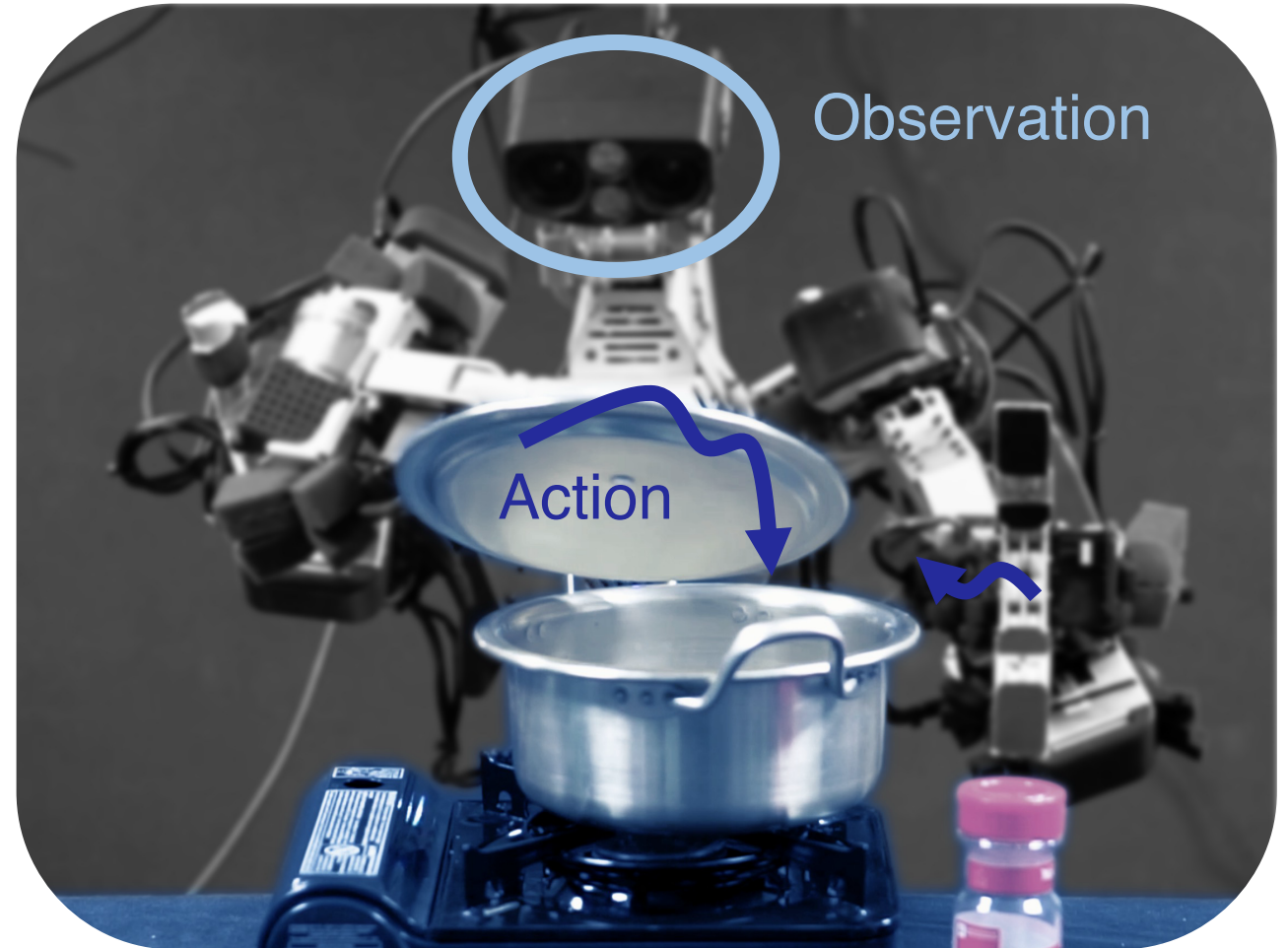
Data-Driven Approaches

- Learning behaviors directly from data
- Adaptable across tasks and settings

Learning Robotic Skills



Highly coupled to the robot platform



Developing Generalizable Skills



What should be **shared**, what must **adapt** to each robot?



Scaling through Extensive Data

Data

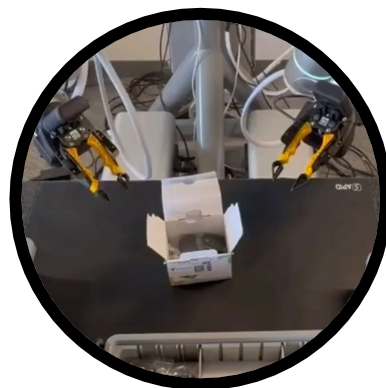
Learning structural differences
from data



OXE
10k hours



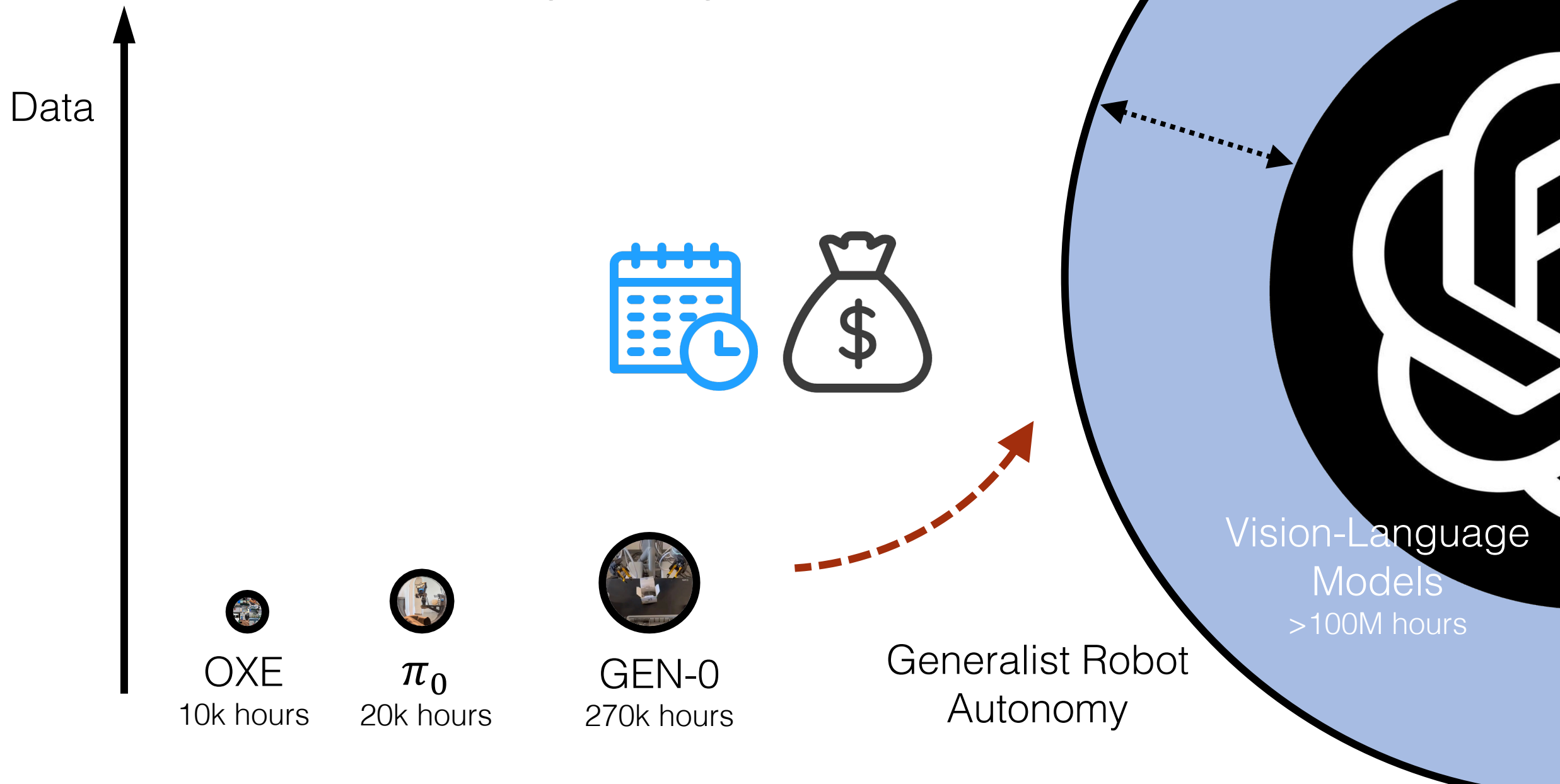
π_0
20k hours



GEN-0
270k hours



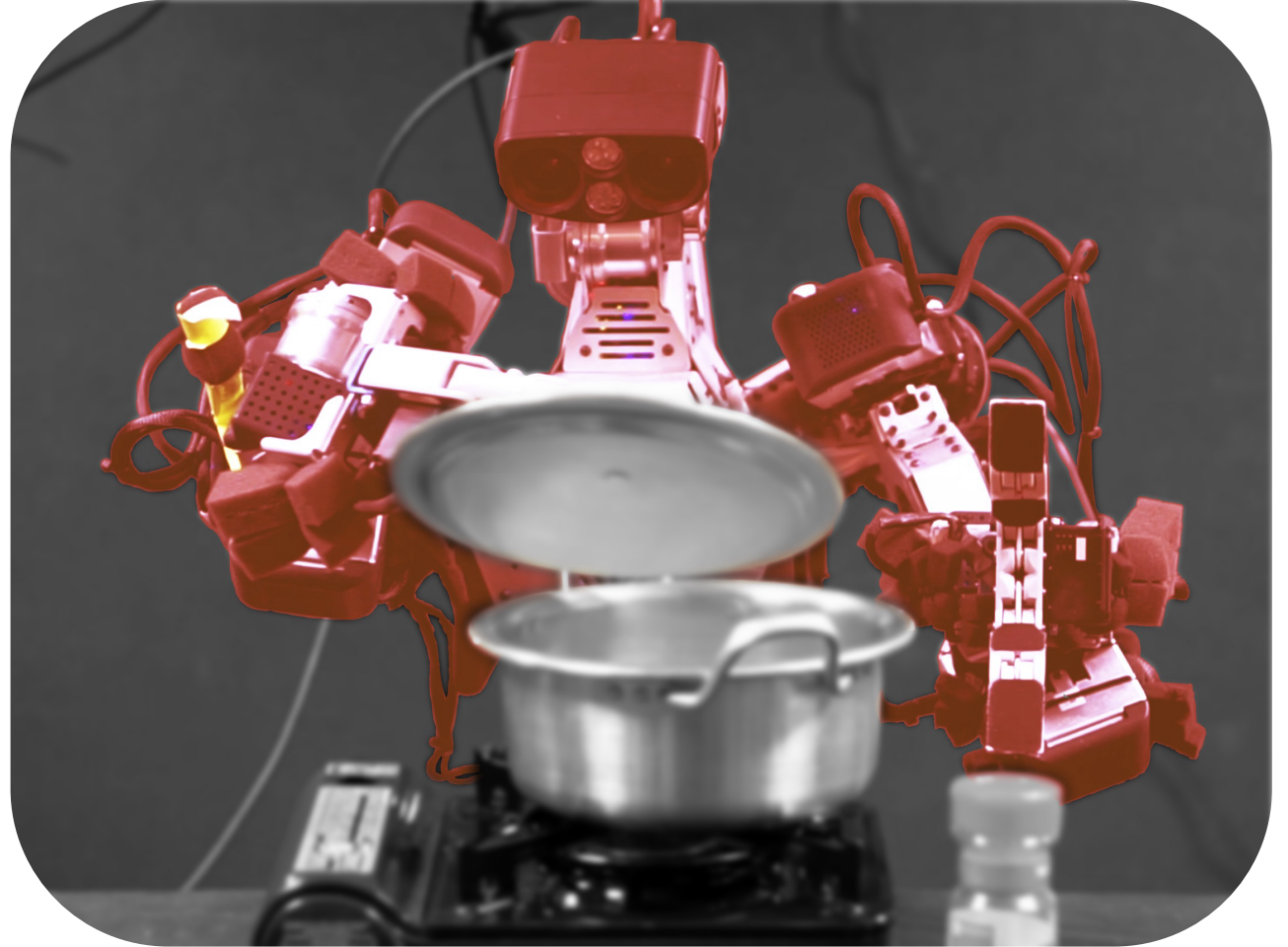
Scaling through Extensive Data



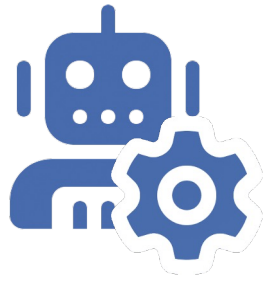
Structural Priors in Robot Systems

-  modifiable **hardware** design
-  structured **control** architecture
-  behaviors structured by **representation**

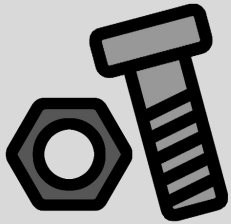
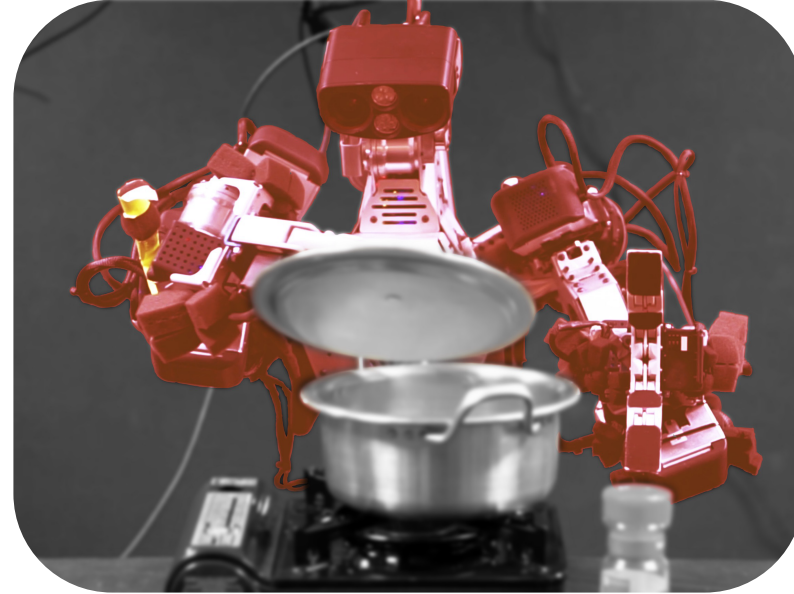
Robot system structure need
not be rediscovered from data



Embodiment-Aware Skill Learning



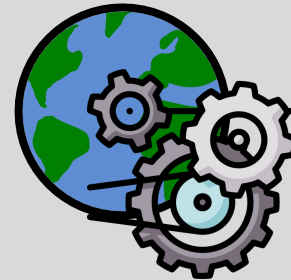
Structural Prior
from Robot Systems



hardware

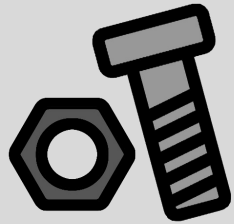
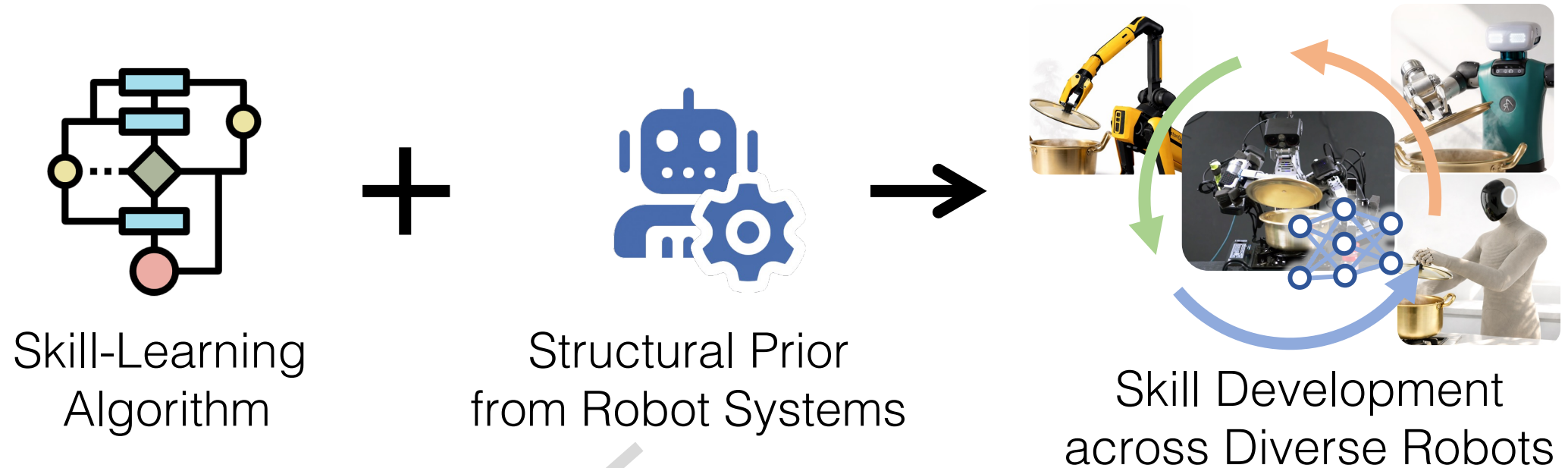


control



representation

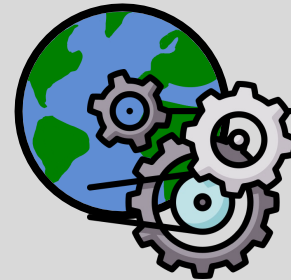
Embodiment-Aware Skill Learning



hardware

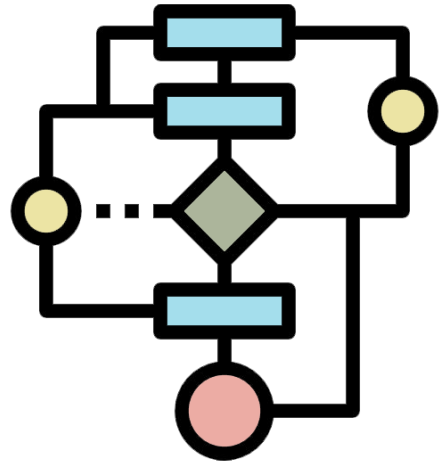


control

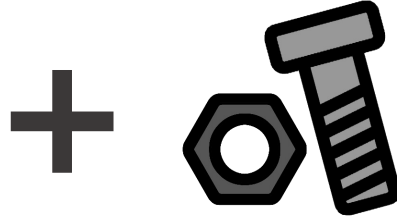


representation

Developing a **holistic, embodiment-aware** approach for learning generalizable skills across diverse robots

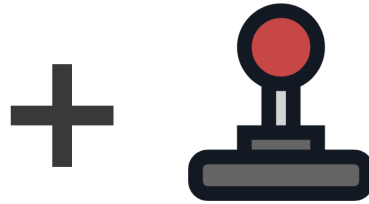


Skill-Learning
Algorithm



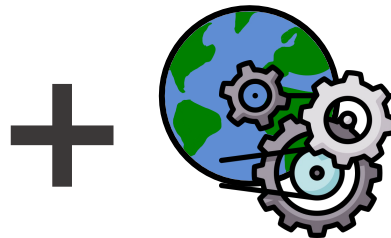
Hardware Design

RA-L 2025
RA-L 2026



Control Architecture

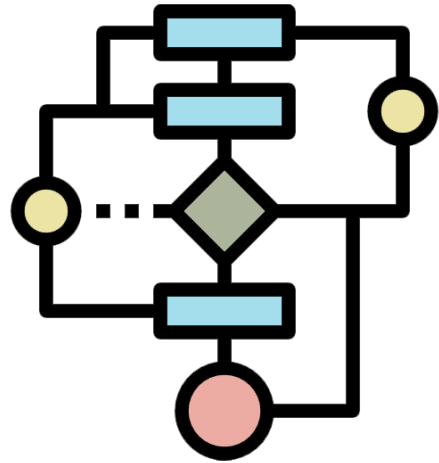
Humanoids 2023: **Best WBC Paper Finalist**
ICRA 2023



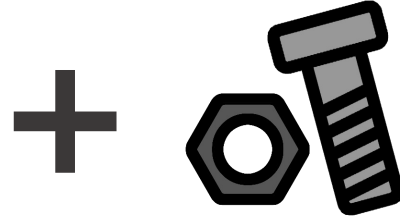
Representation Learning

CoRL 2024
ICRA 2025

Developing a **holistic, embodiment-aware** approach for learning generalizable skills across diverse robots

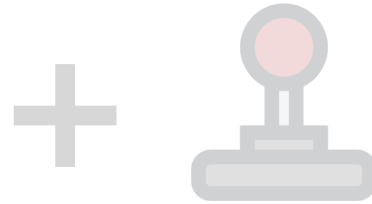


Skill-Learning
Algorithm



Hardware Design

RA-L 2025
RA-L 2026



Control Architecture

Humanoids 2023: **Best WBC Paper Finalist**
ICRA 2023



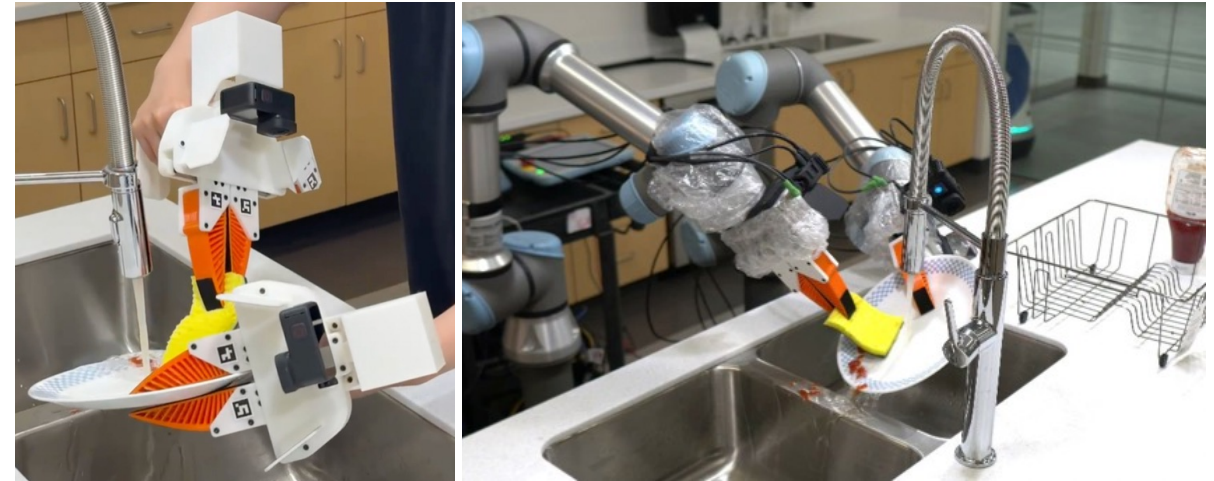
Representation Learning

CoRL 2024
ICRA 2025

Designing Robot Hardware for a Unified Interface



[DexMachina, Zhao et al. 2025]



[UMI, Chi et al. 2024]

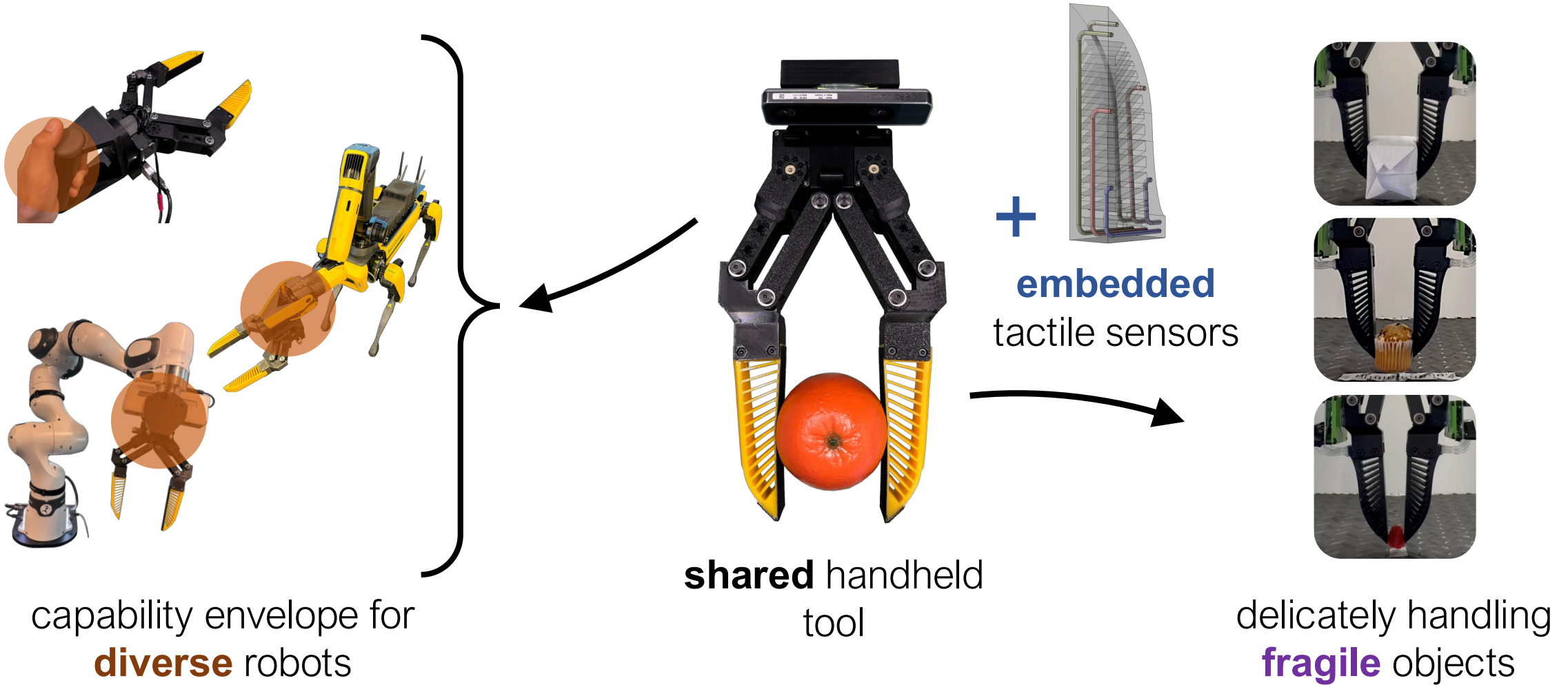
✓ skill transfer across robot hardware

📌 poor generalization to new robots

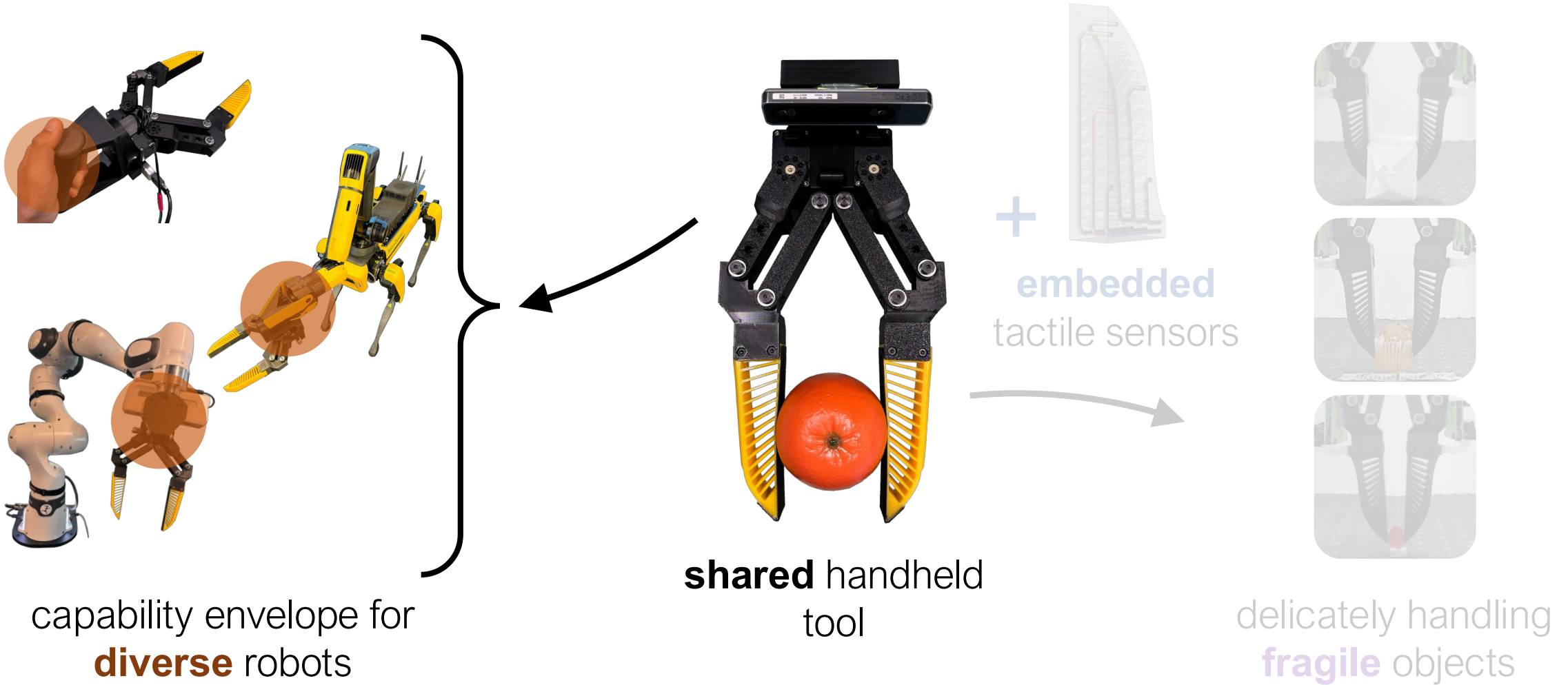
✓ data collection within easier domains

📌 custom hardware per robot

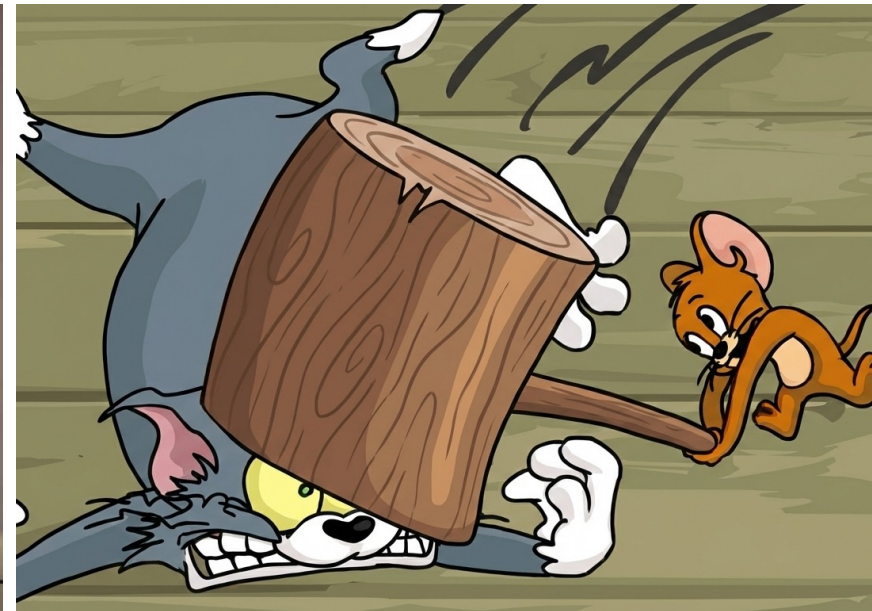
Hardware Design for Consistent Physical Interaction



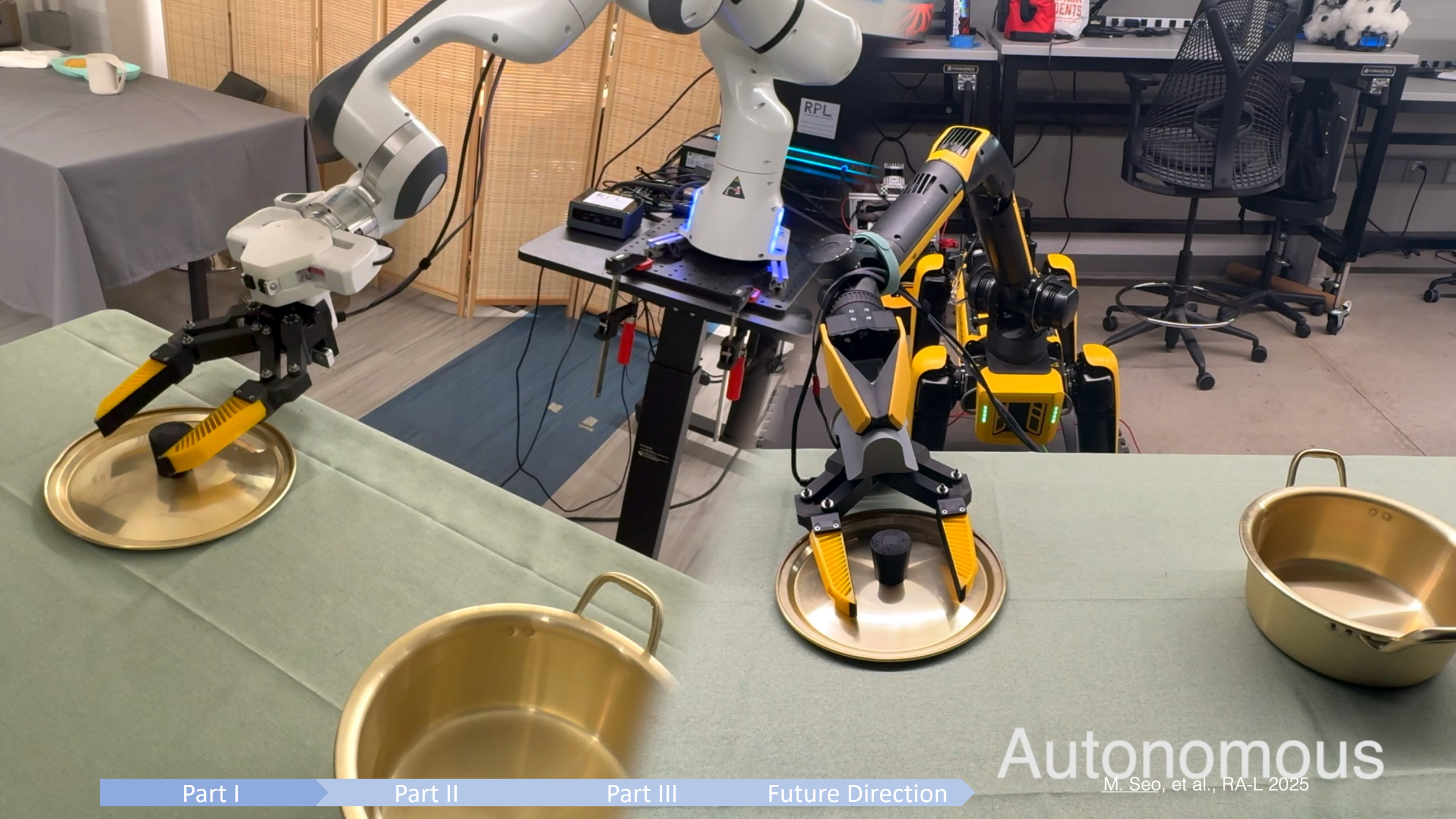
Hardware Design for Consistent Physical Interaction



Integrating a Tool as Part of the Embodiment



serving as a capability envelope



Part I

Part II

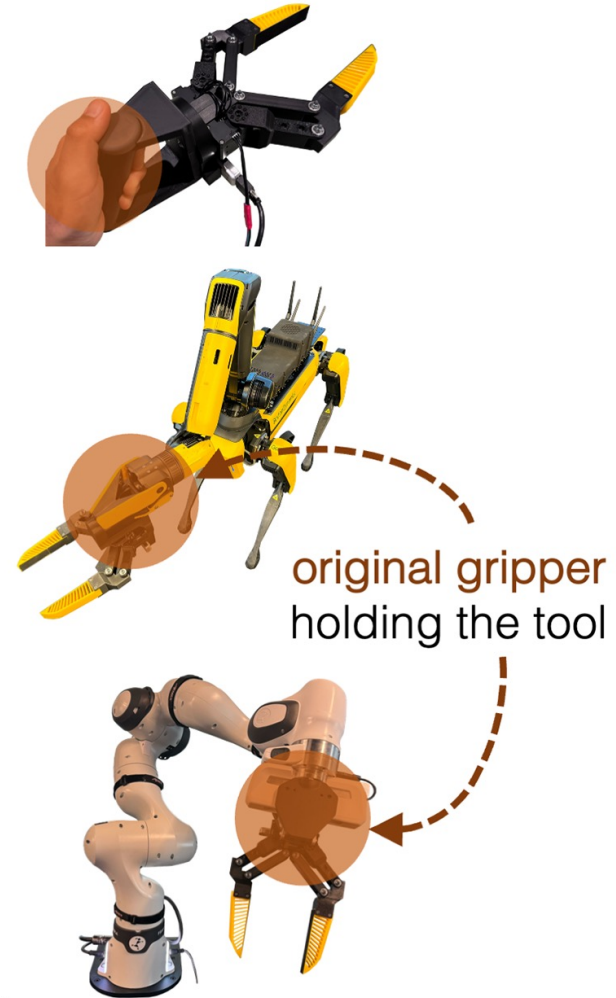
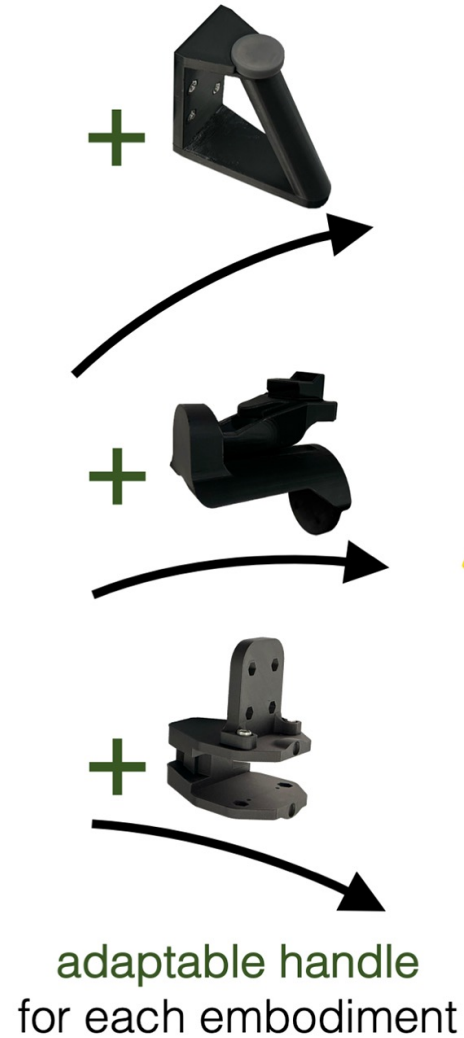
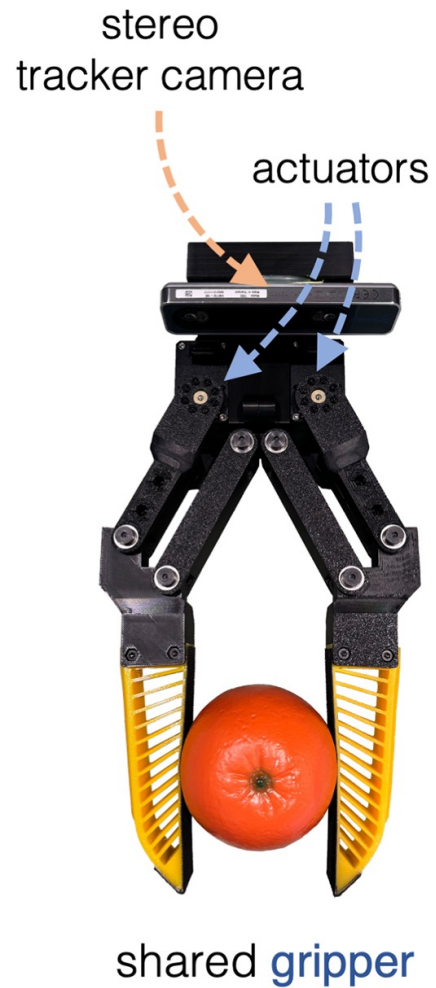
Part III

Future Direction

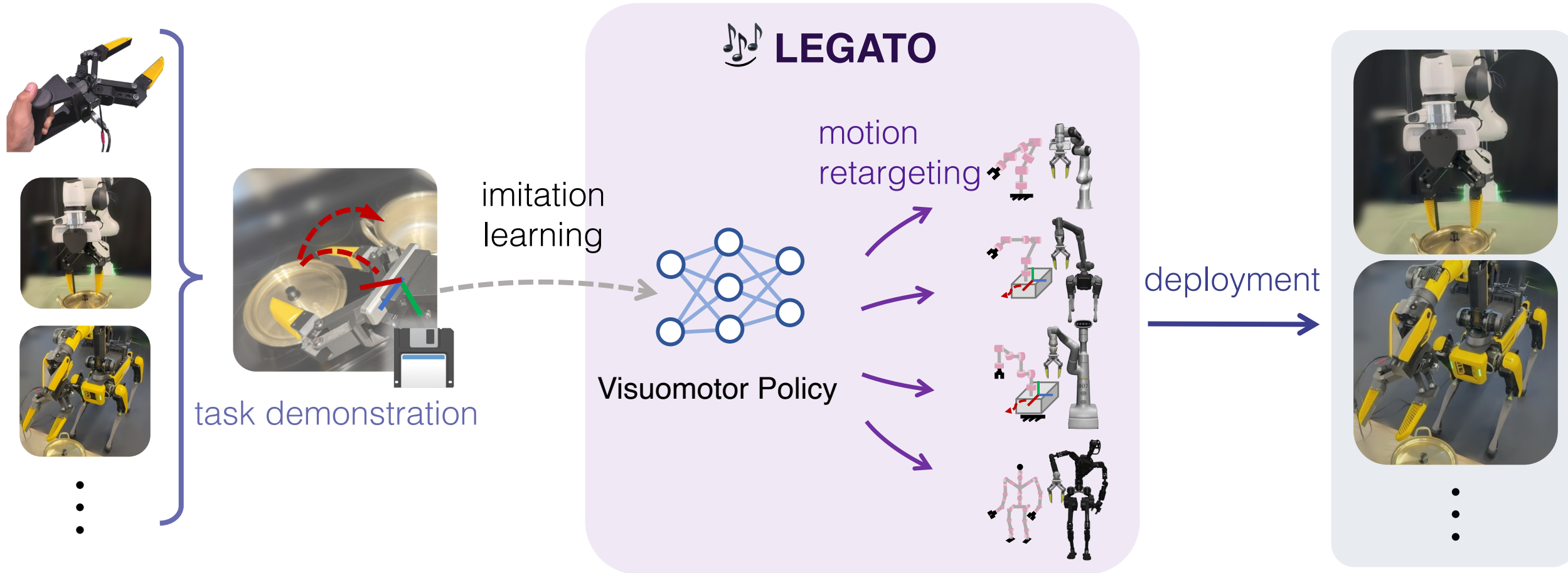
Autonomous

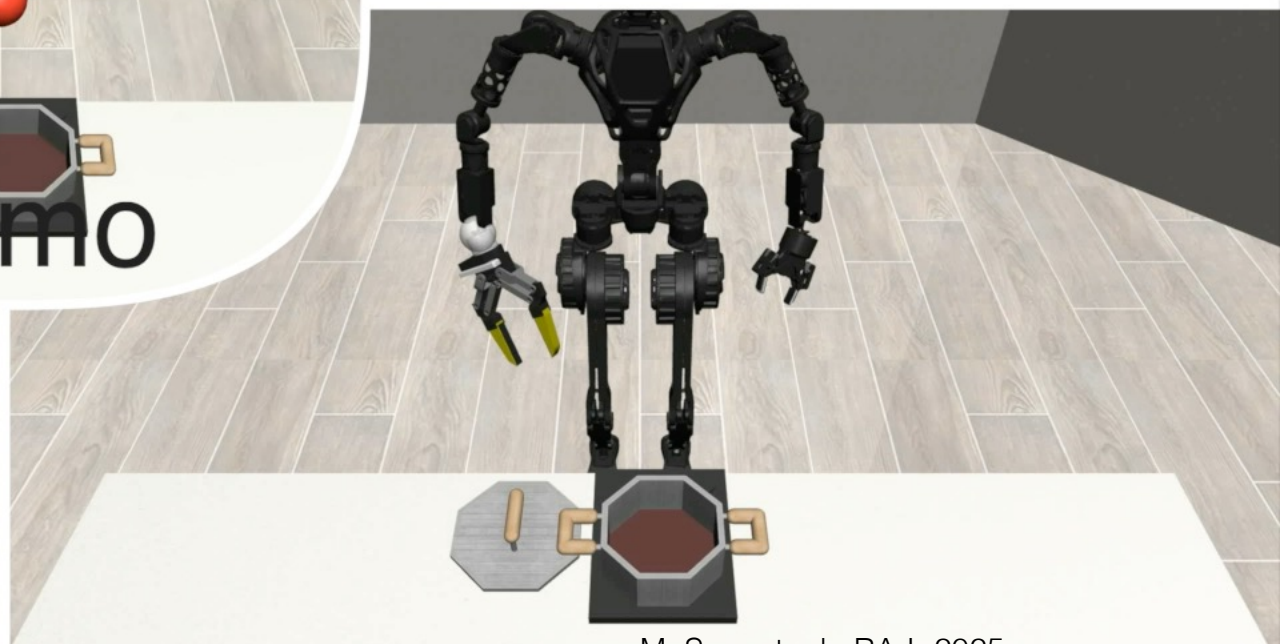
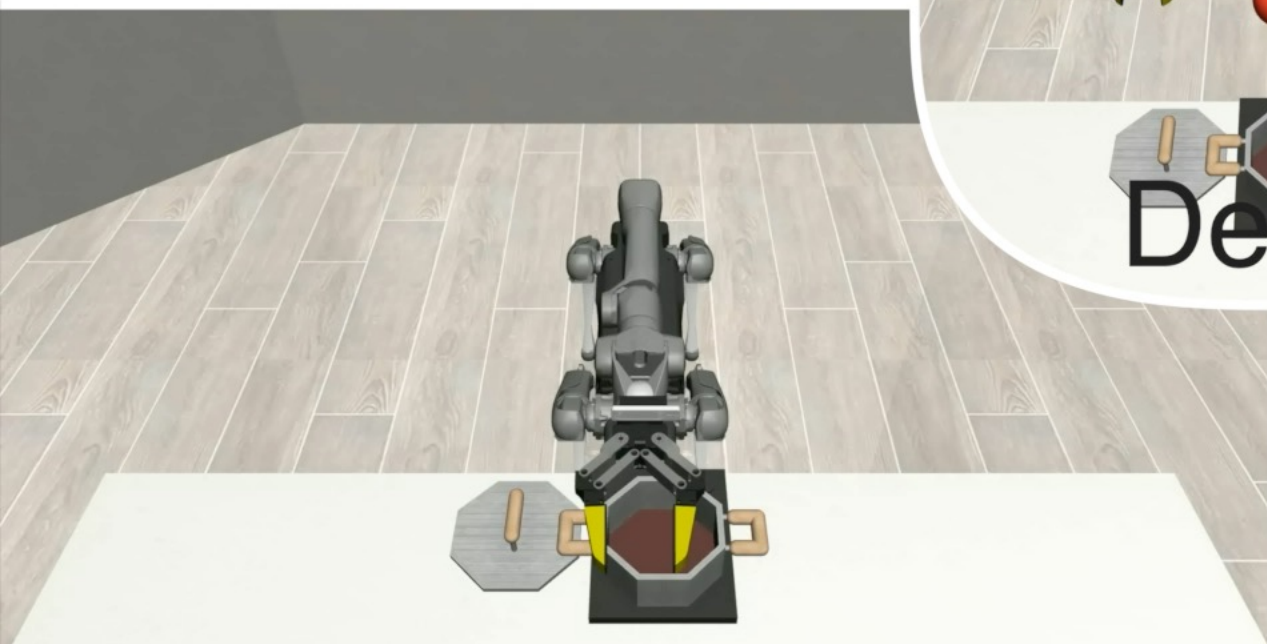
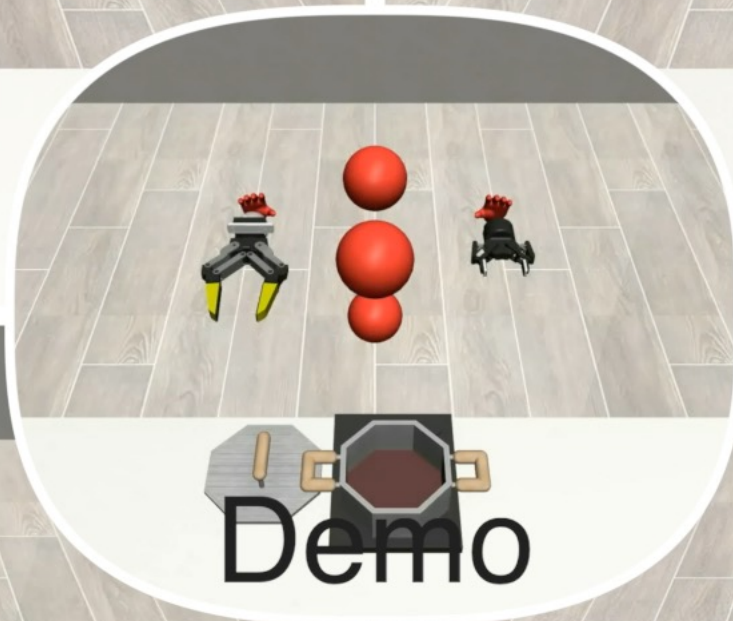
M. Seo, et al., RA-L 2025

LEGATO: Learning with a Handheld Grasping Tool



LEGATO: Learning with a Handheld Grasping Tool





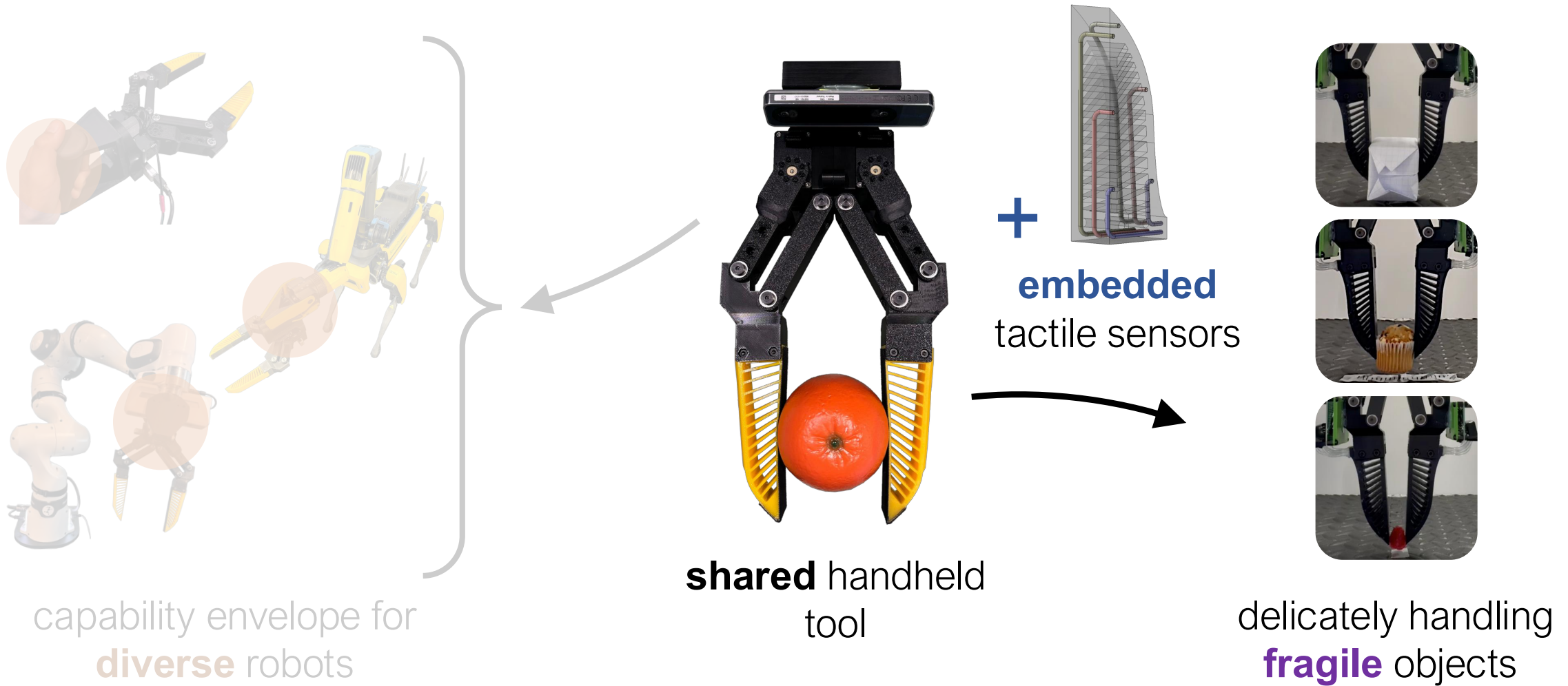
Part I

Part II

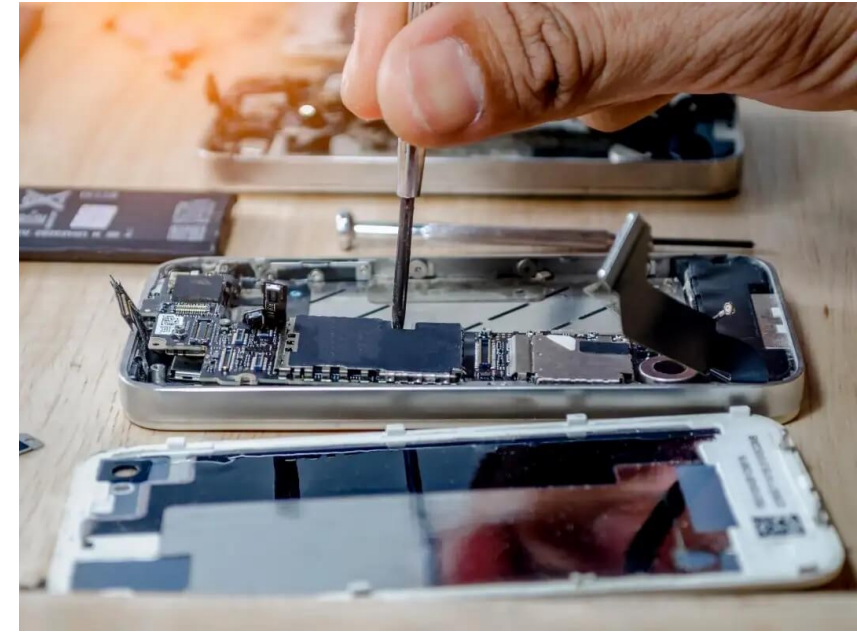
Part III

Future Direction

Hardware Design for Consistent Physical Interaction

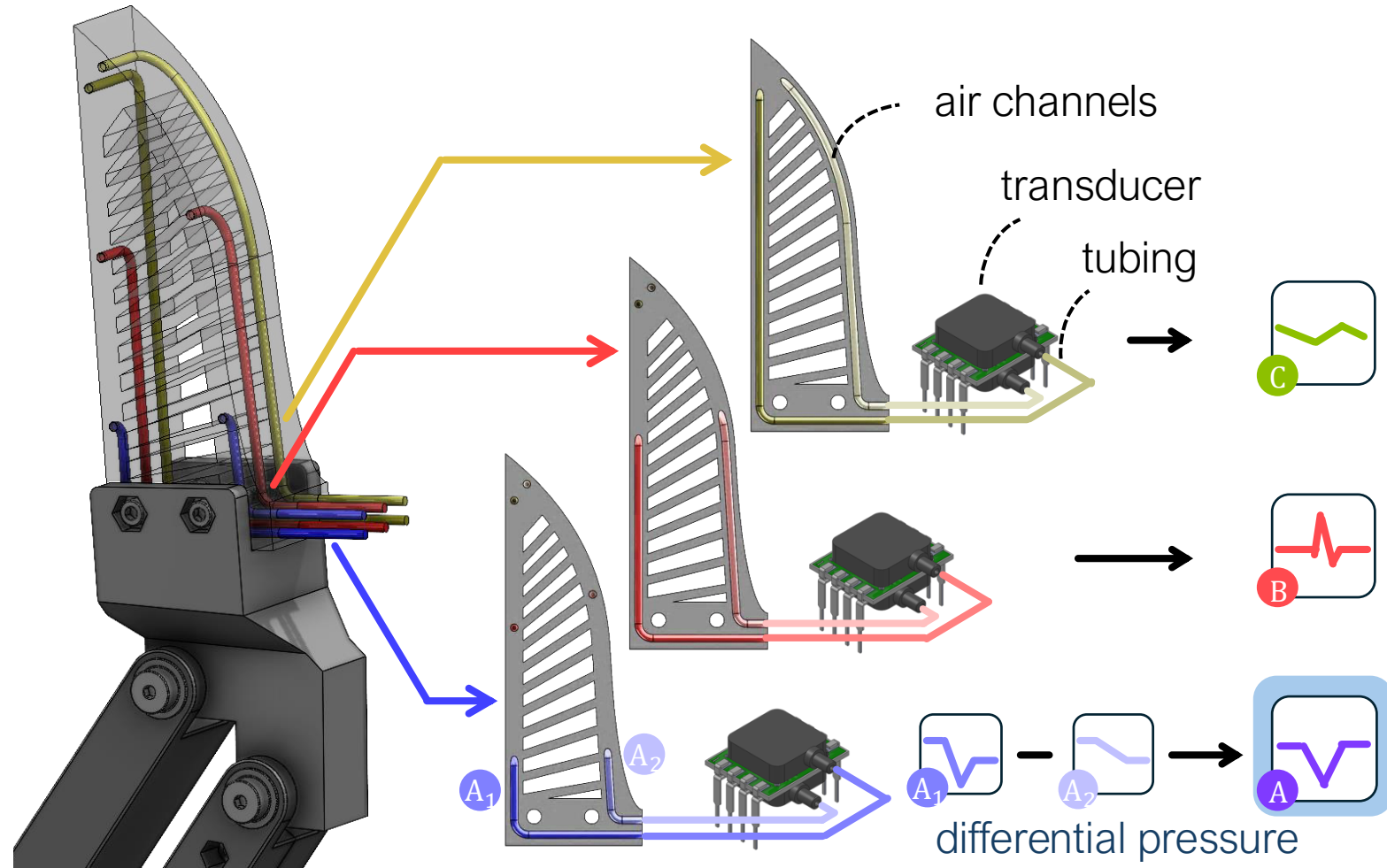


Enhancing Capabilities by Using Tools



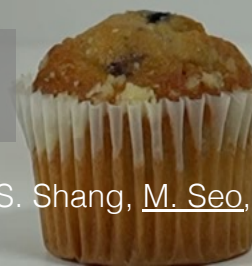
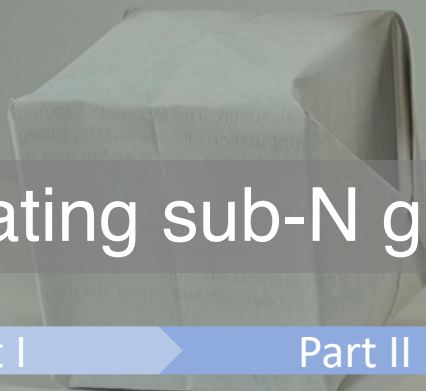
providing new sensing modalities

FORTE: Fragile Object Grasping with Tactile Sensing





FORTE's Force Sensing



Estimating sub-N grasping forces via regression

Part I

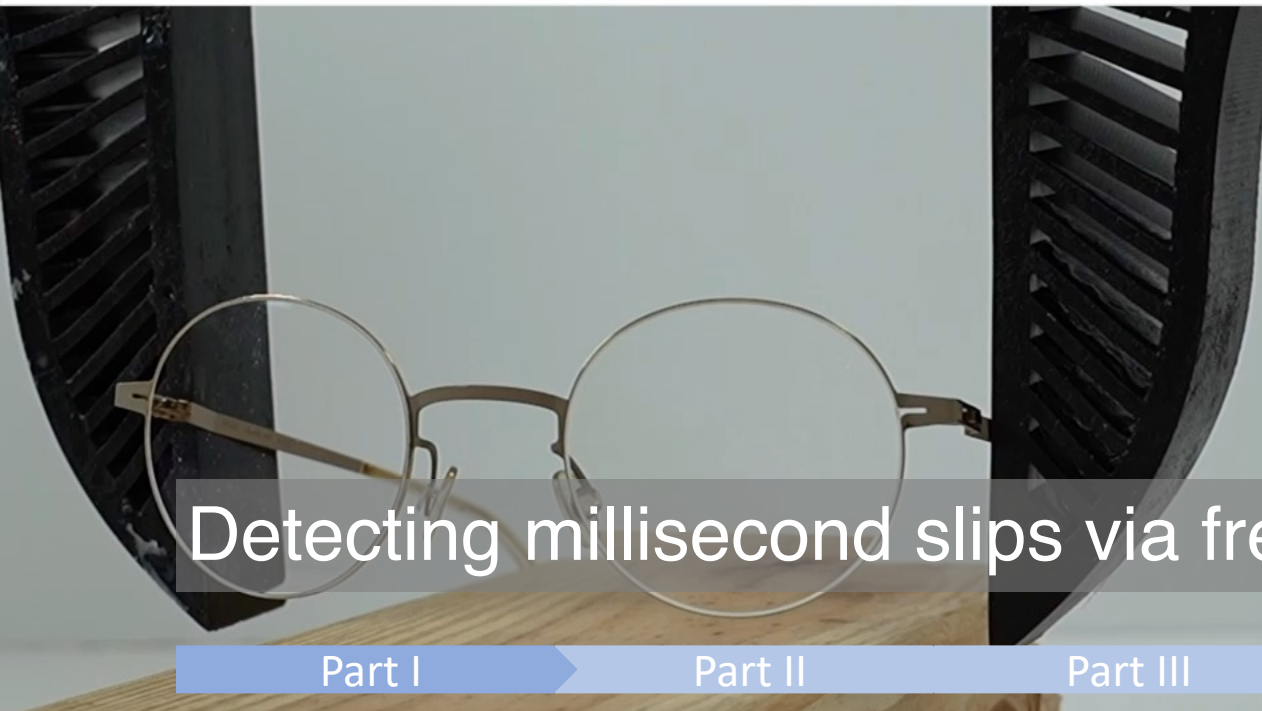
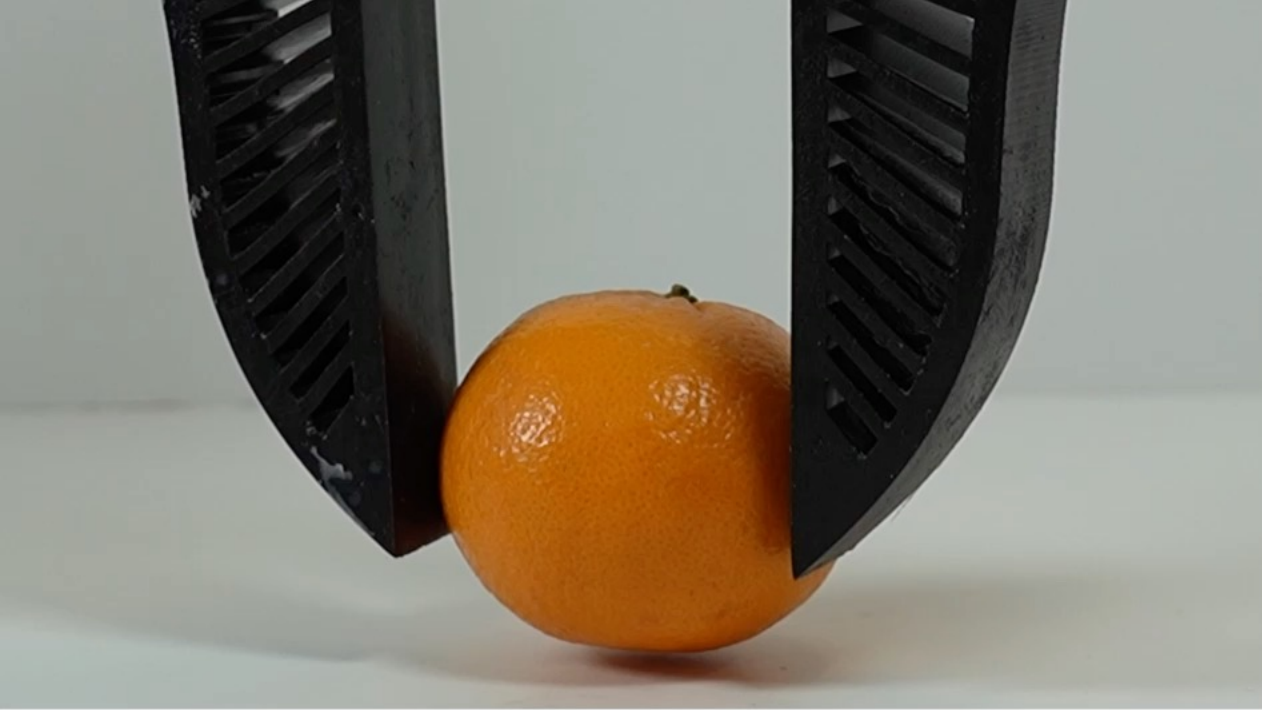
Part II

Part III

Future Direction

S. Shang, M. Seo, et al., RA-L 2026

4x



Detecting millisecond slips via frequency-domain analysis

Part I

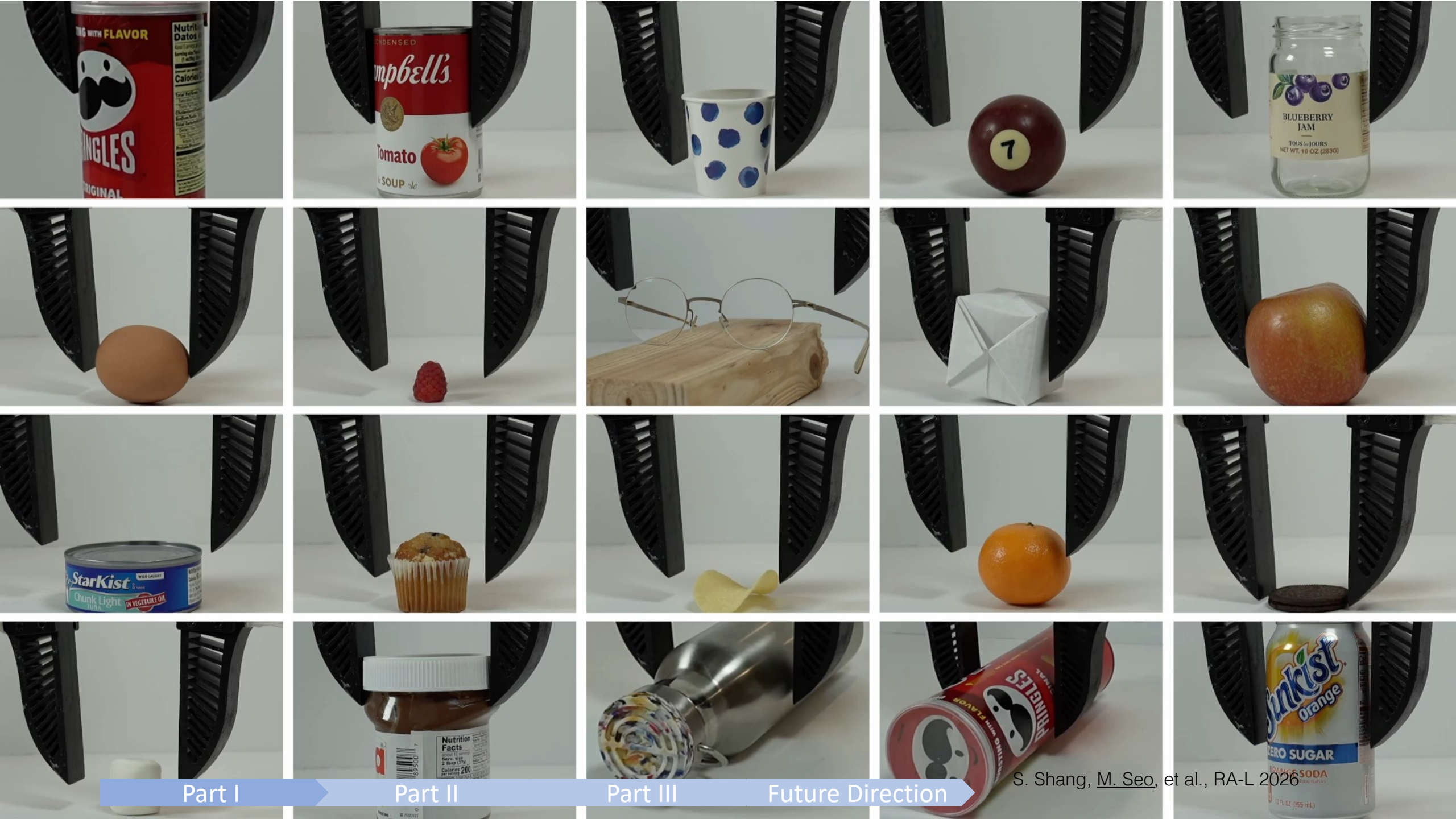
Part II

Part III

Future Direction

S. Shang, M. Seo, et al., RA-L 2023

0.5x



Part I

Part II

Part III

Future Direction

Summary I: Hardware Design for Consistent Physical Interaction



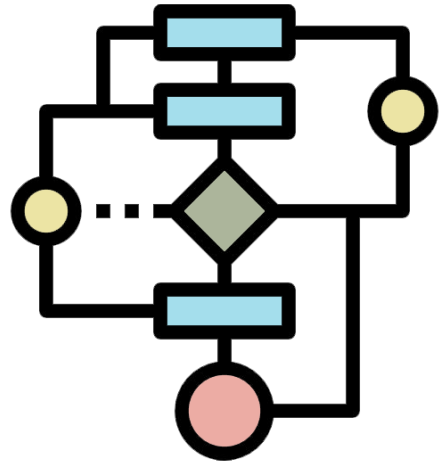
LEGATO

learning with transferrable skills
with handheld grasping tool

FORTE

extending tactile perception
via embedded fluidic sensors

Developing a **holistic, embodiment-aware** approach for learning generalizable skills across diverse robots

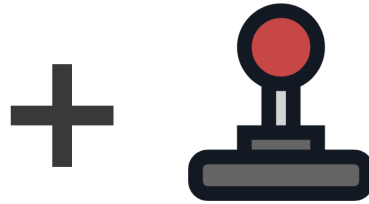


Skill-Learning
Algorithm



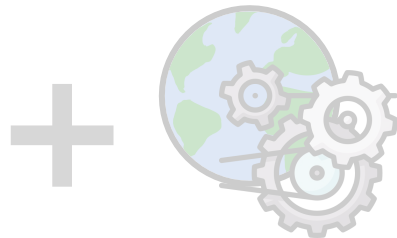
Hardware Design

RA-L 2025
RA-L 2026



Control Architecture

Humanoids 2023: **Best WBC Paper Finalist**
ICRA 2023



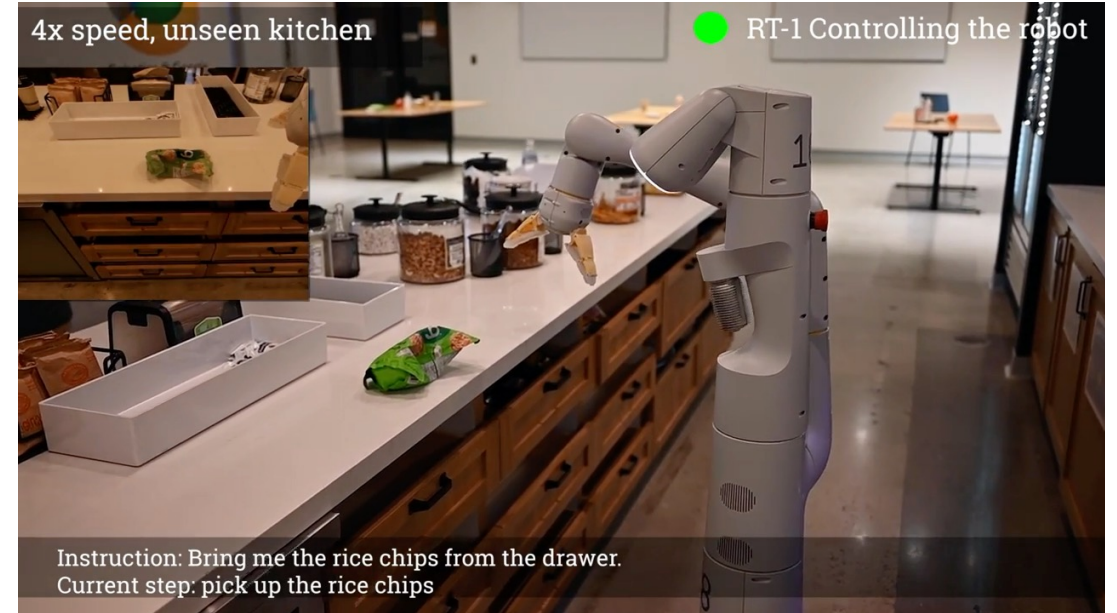
Representation Learning

CoRL 2024
ICRA 2025

Bridging Control and Skill Learning Algorithms



[DRC, Krotkov et al. 2017]

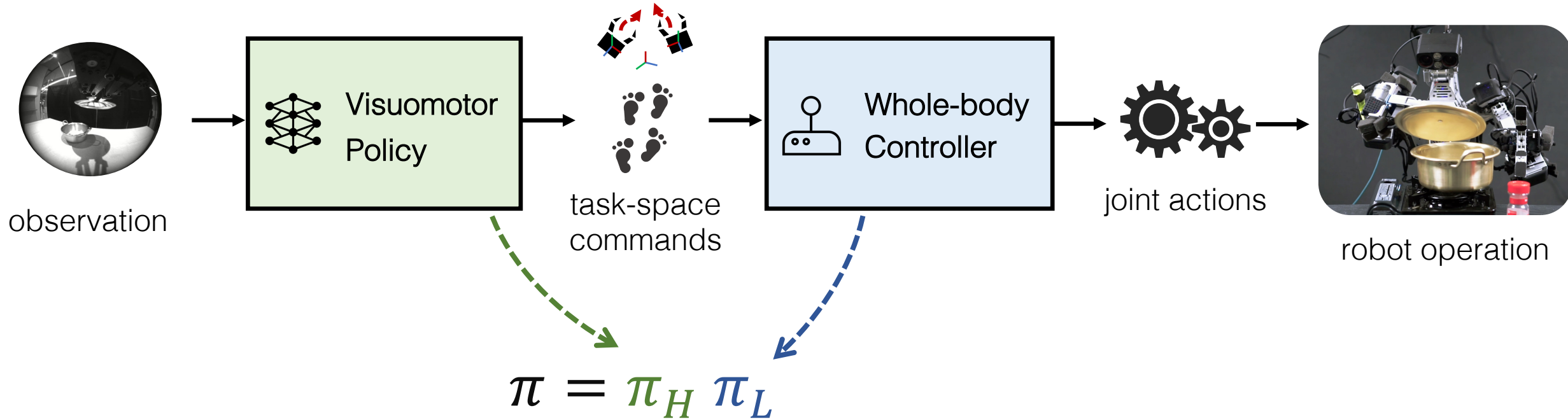


[RT-1, Brohan et al. 2022]

- ✓ advanced robot hardware
- 📌 controlled settings

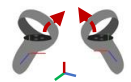
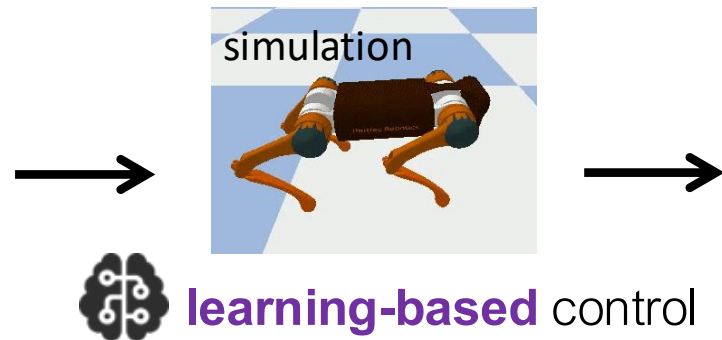
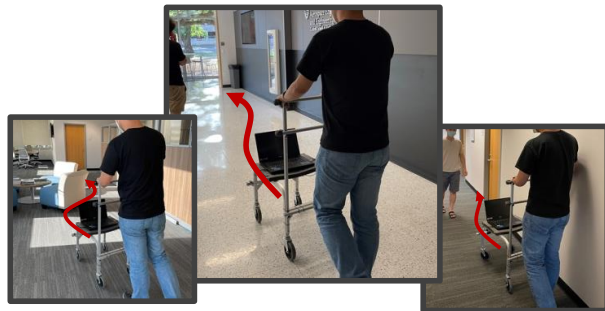
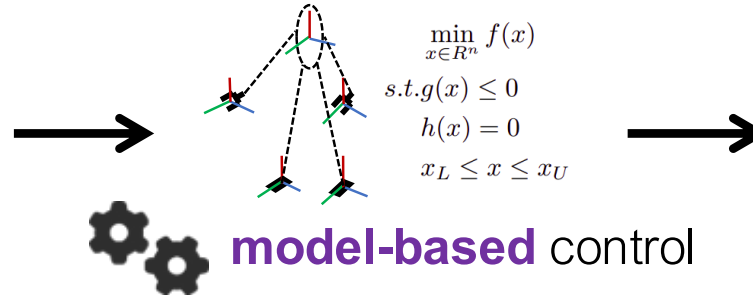
- ✓ complex behaviors
- 📌 simple morphologies

Bridging Control and Skill Learning Algorithms



Control algorithms provide abstractions that simplify visuomotor policy learning.

Bridging Control and Skill Learning Algorithms



human demonstrations



autonomous policy

Part I

Part II

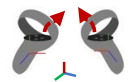
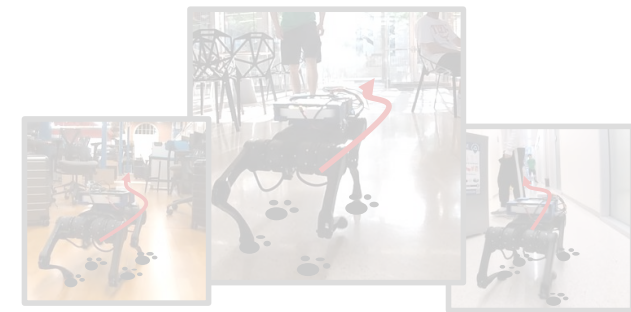
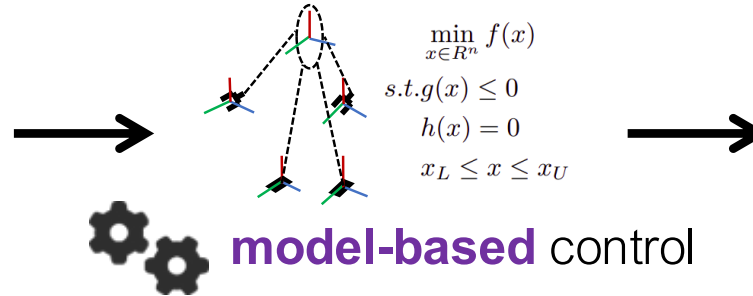
Part III

Future Direction

M. Seo, et al., Humanoids 2023

M. Seo, et al., ICRA 2023

Bridging Control and Skill Learning Algorithms



human demonstrations



autonomous policy

Part I

Part II

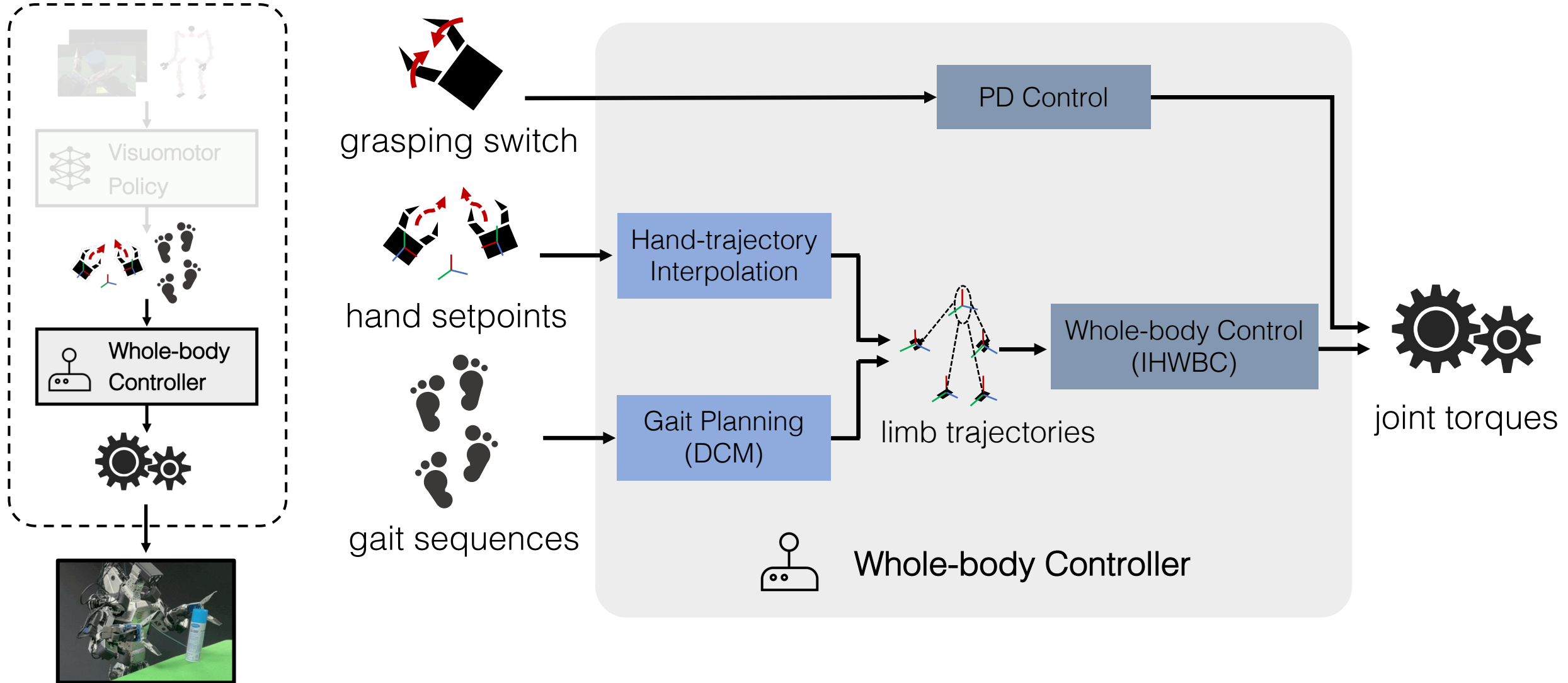
Part III

Future Direction

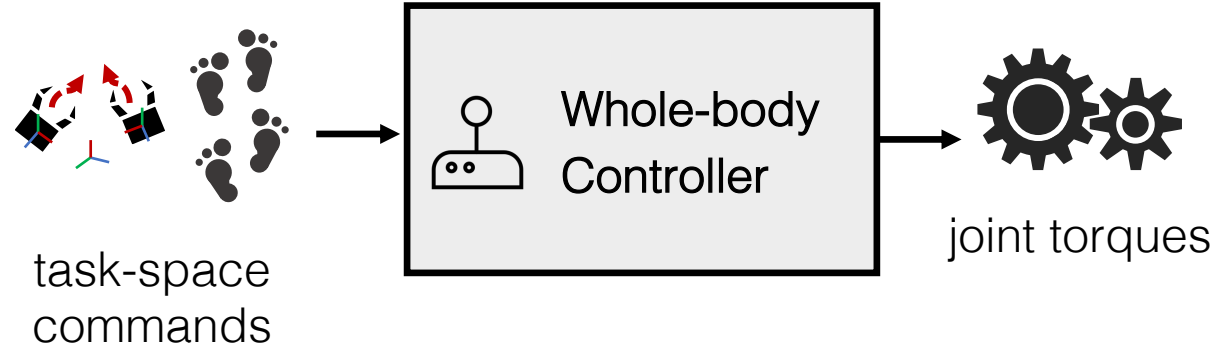
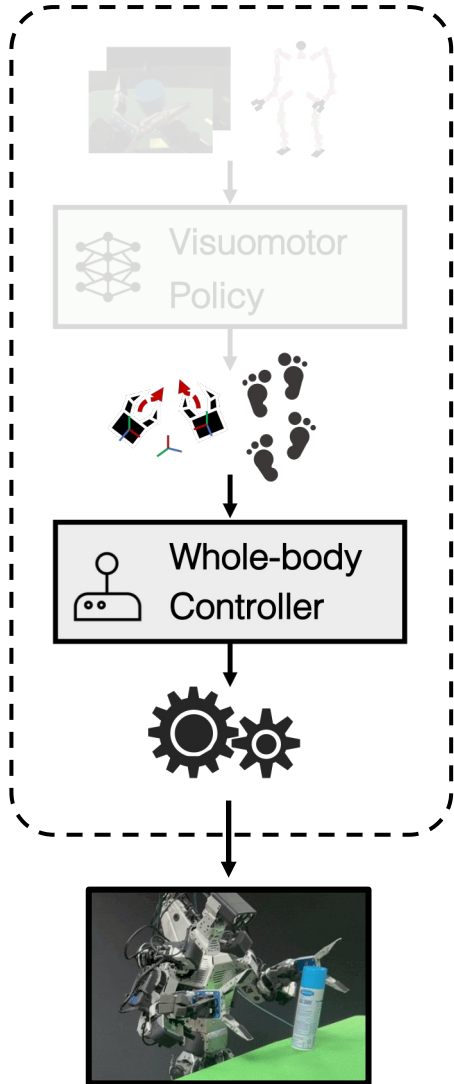
M. Seo, et al., Humanoids 2023

M. Seo, et al., ICRA 2023

TRILL: Teleoperation and Imitation Learning for Loco-manipulation



TRILL: Teleoperation and Imitation Learning for Loco-manipulation



$$\min_{\ddot{q}, f_r} \sum_{i=1}^n w_i \|J_i \ddot{q} + \dot{J}_i \dot{q}_m - \ddot{x}_i^d\|^2 \quad \text{task-space tracking}$$

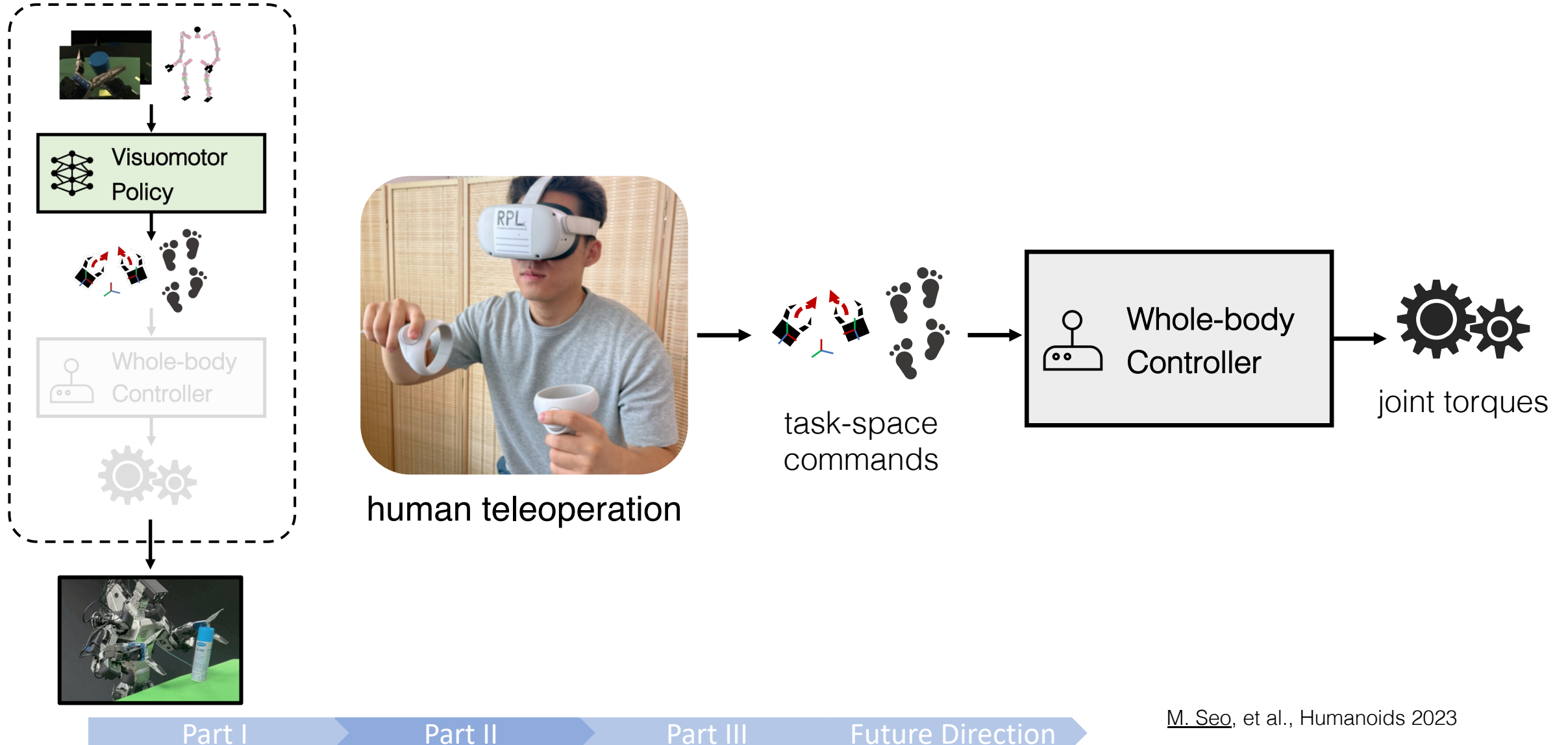
$$+ w_{f_r} \|f_r^d - f_r\|^2 + \lambda_q \|\ddot{q}\|^2 + \lambda_{f_r} \|f_r\|^2, \quad \text{regularization}$$

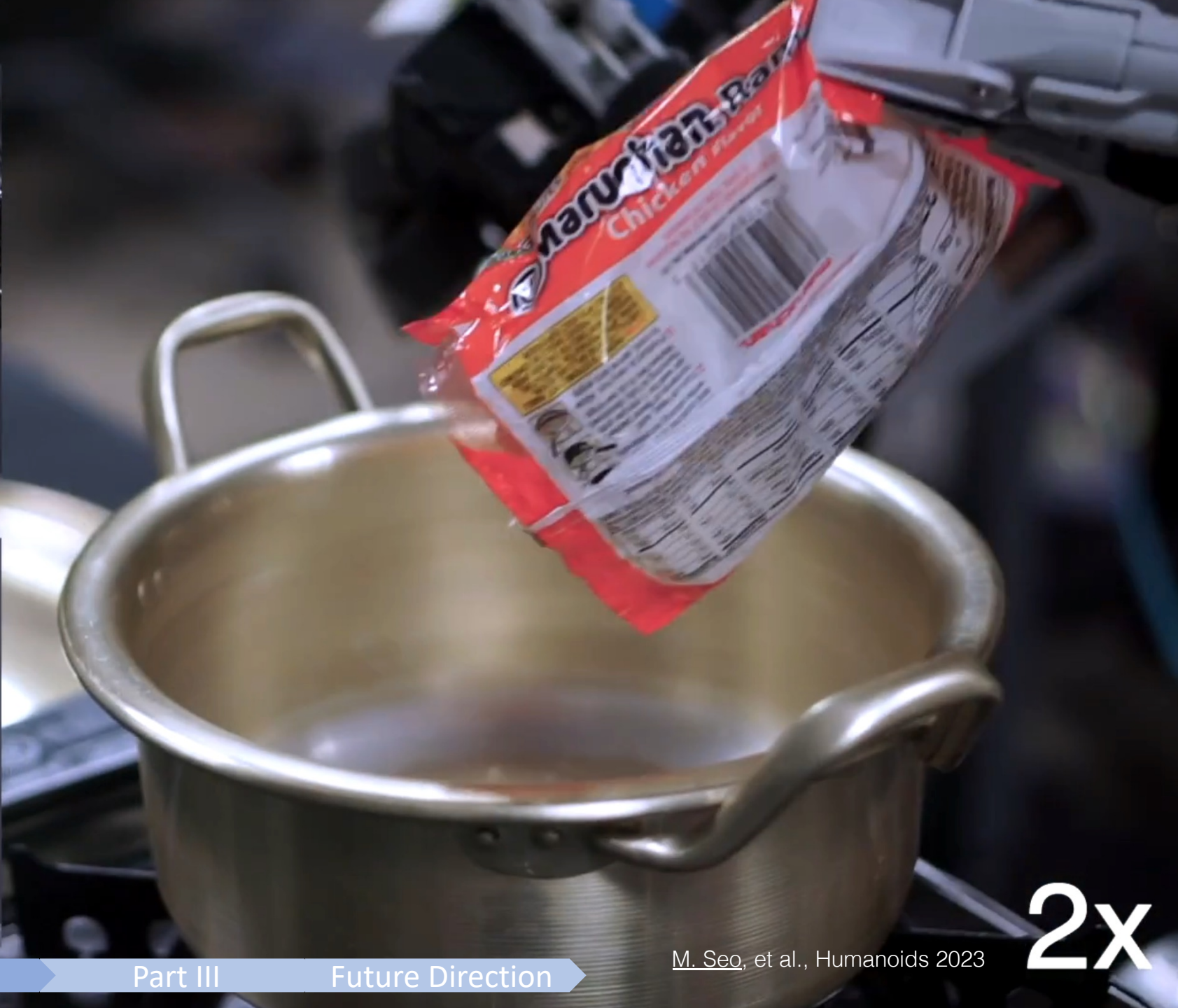
subject to

$$\left. \begin{aligned} S_f(A\ddot{q} + b + g - J^\top f_r) &= 0 \\ U f_r &\geq 0 \\ S_r f_r &\leq f_r^{\max} \end{aligned} \right\} \text{robot dynamics}$$

$$\left. \begin{aligned} \ddot{q}_{\min} &\leq \ddot{q} \leq \ddot{q}_{\max} \\ \tau_{\min} &\leq S_a(A\ddot{q} + b + g - J_c^\top f_r) \leq \tau_{\max} \end{aligned} \right\} \text{joint limits}$$

TRILL: Teleoperation and Imitation Learning for Loco-manipulation





Part I

Part II

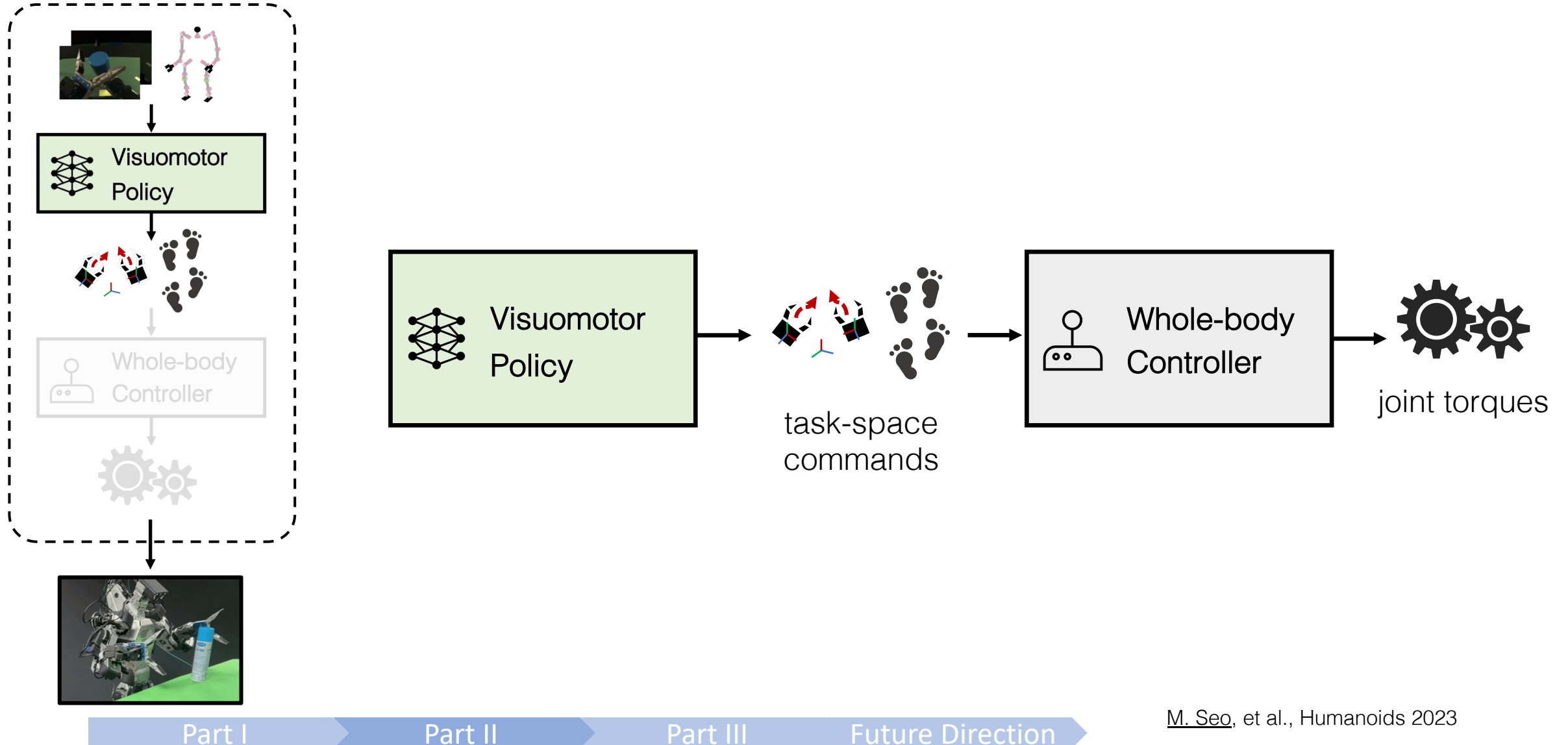
Part III

Future Direction

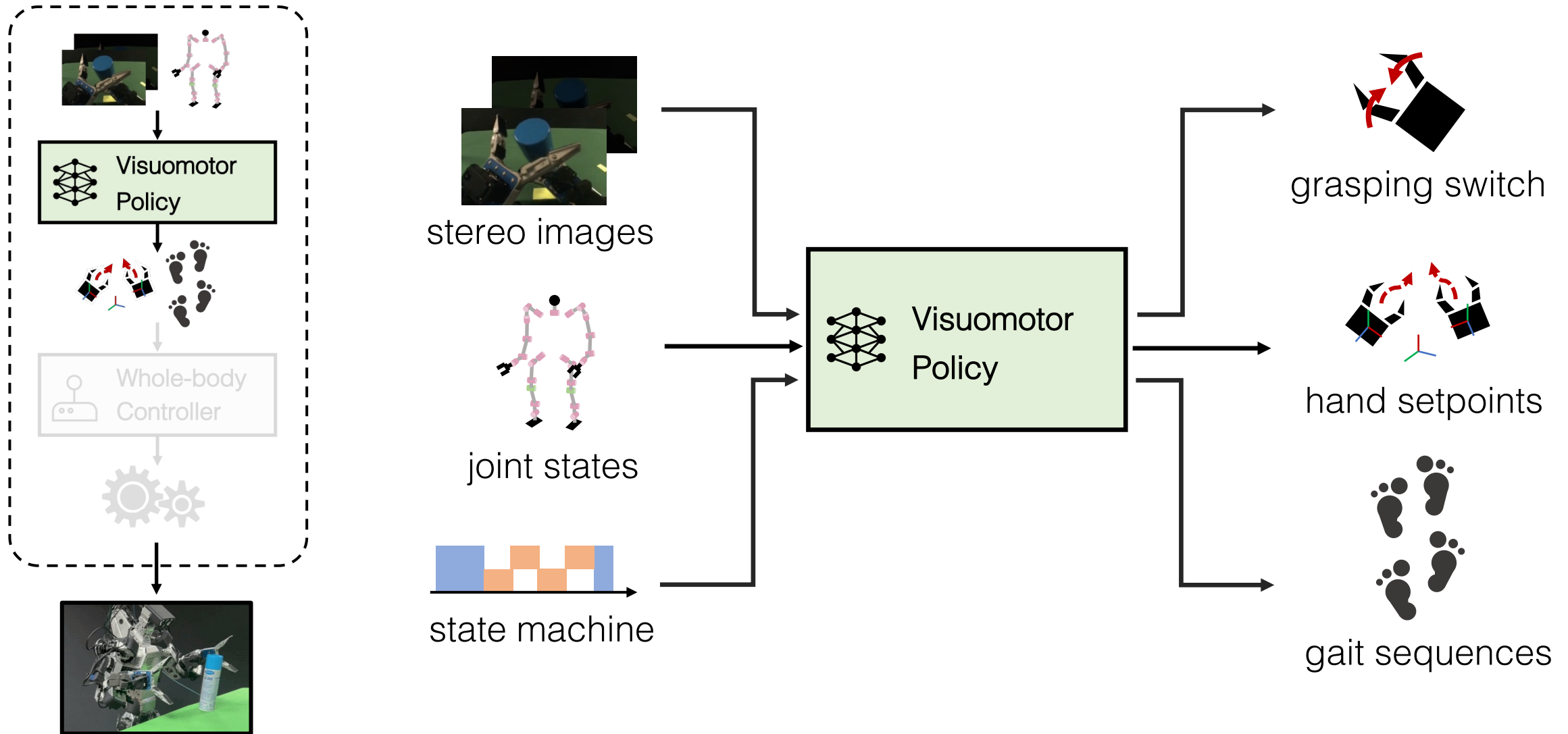
M. Seo, et al., Humanoids 2023

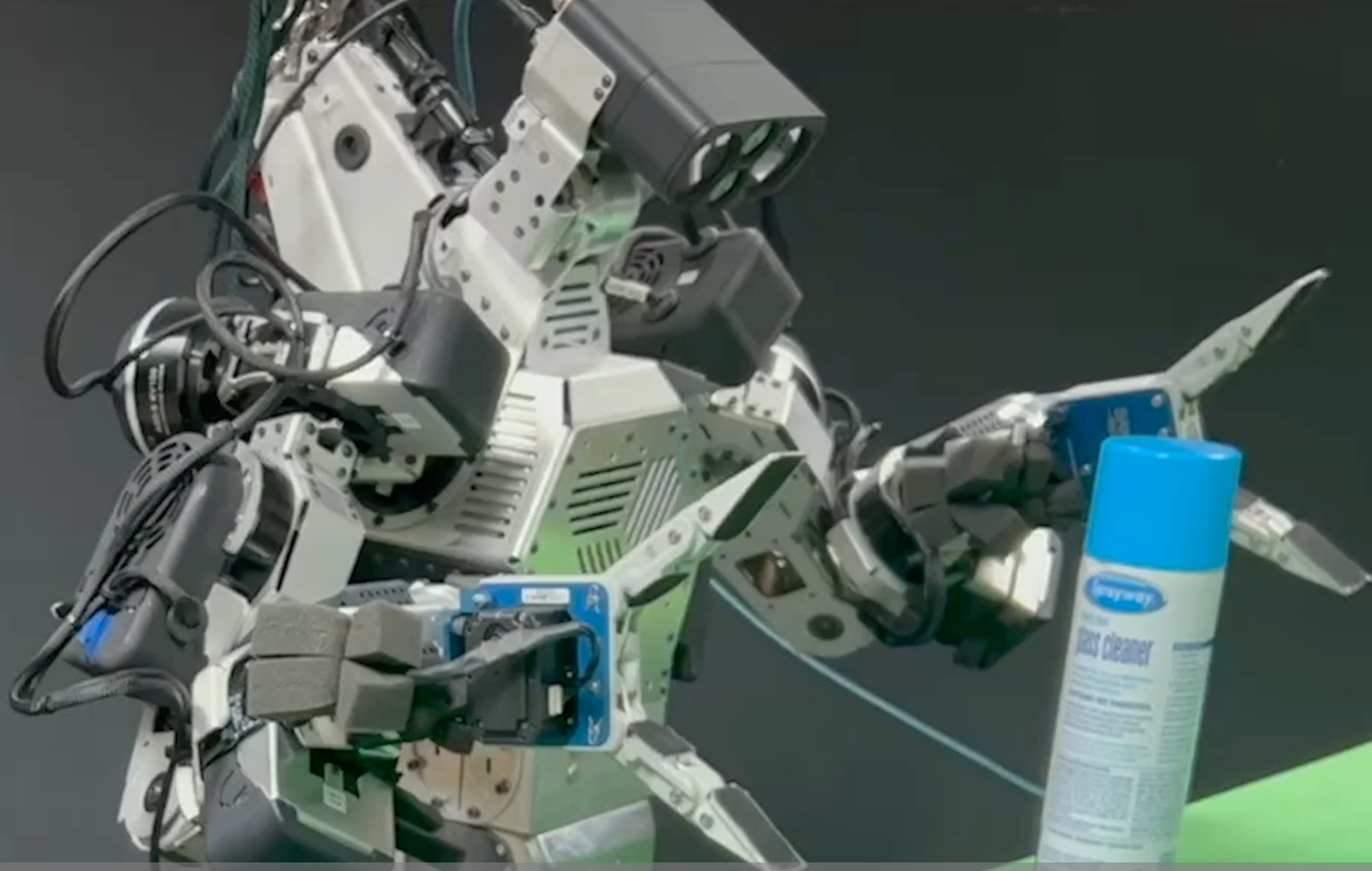
2x

TRILL: Teleoperation and Imitation Learning for Loco-manipulation



TRILL: Teleoperation and Imitation Learning for Loco-manipulation





First successful humanoid visuomotor imitation learning
90% success rate on complex bimanual manipulation



92% success rate on loco-manipulation (sim)

Part I

Part II

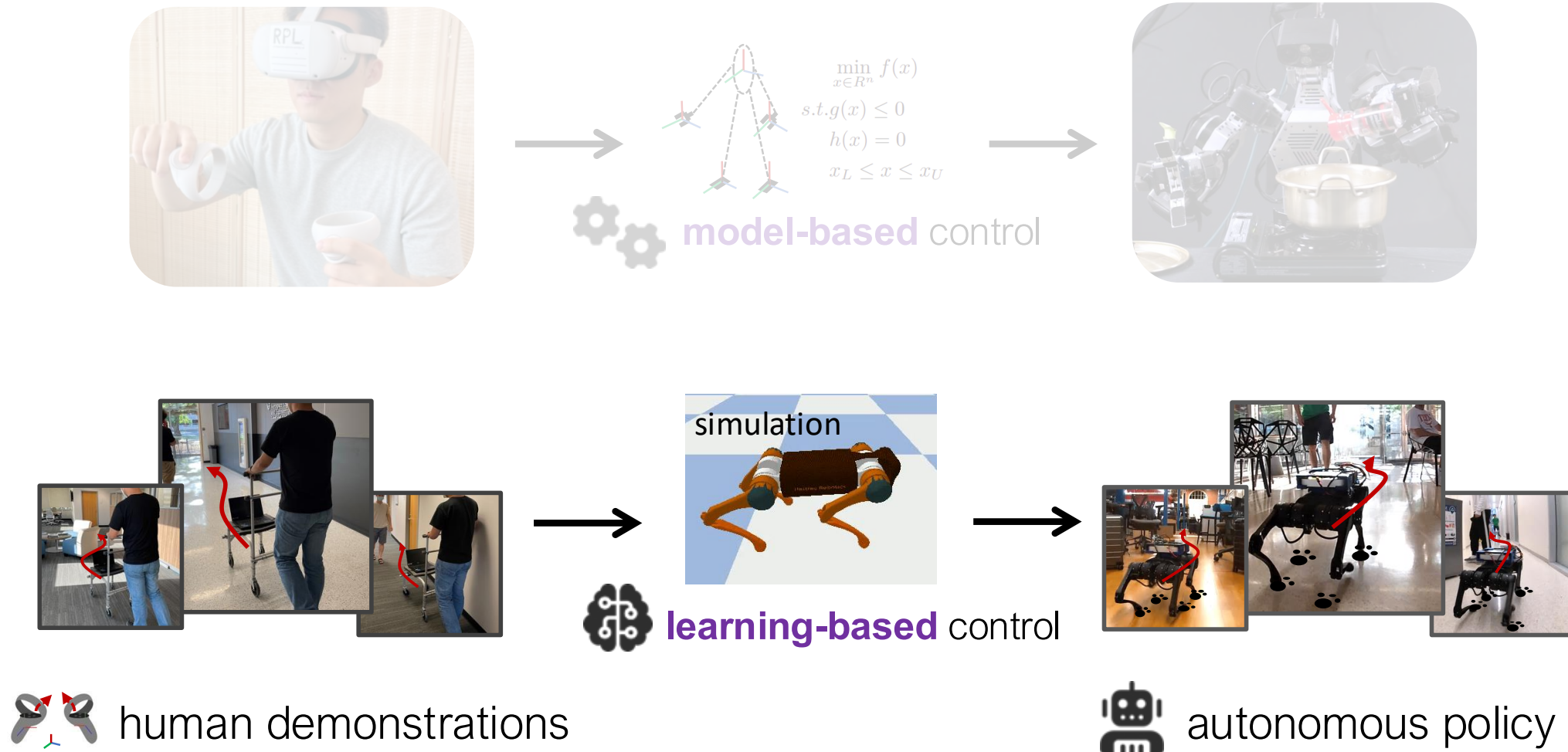
Part III

Future Direction

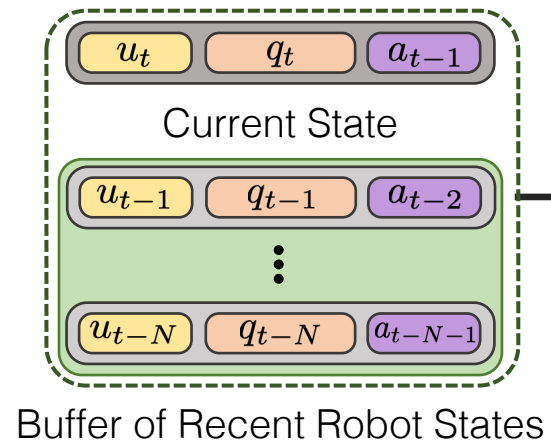
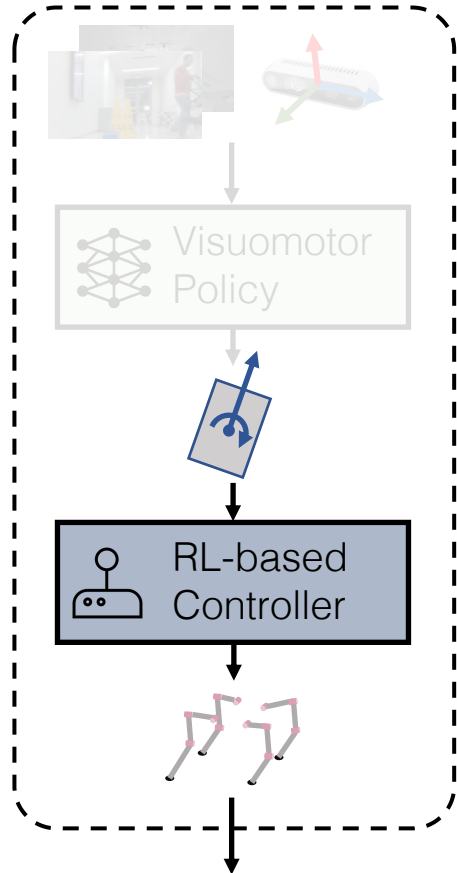
[M. Seo, et al., Humanoids 2023](#)

2x

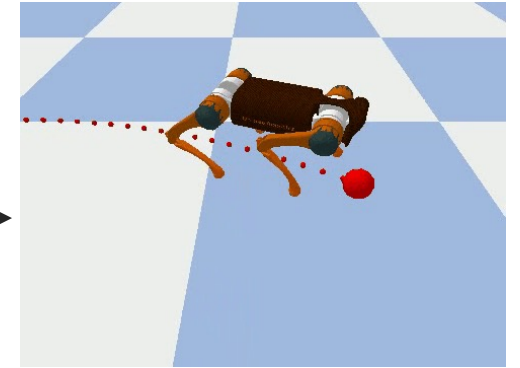
Bridging Control and Skill Learning Algorithms



PRELUDE: Perceptive Locomotion Under Dynamic Environments



joint-space
action

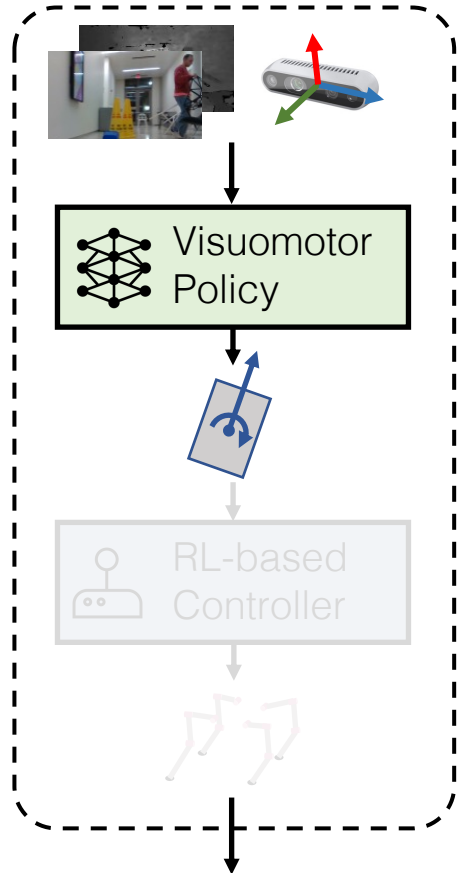


trial and error
in simulation

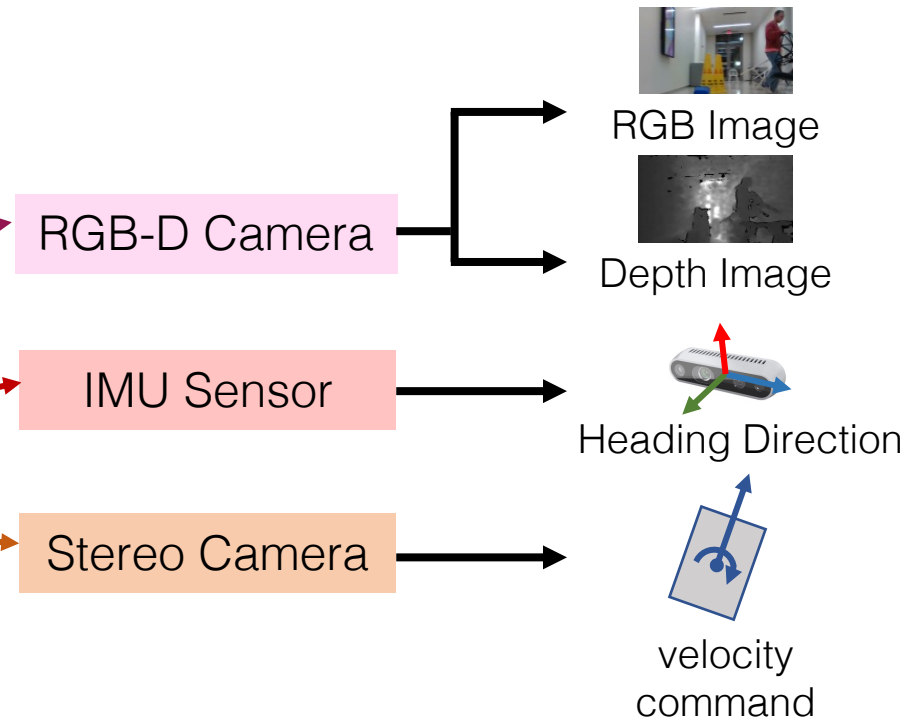
reinforcement learning in **simulation**

- ✓ learning agile locomotion
- ✓ sim-to-real transfer of the control policy

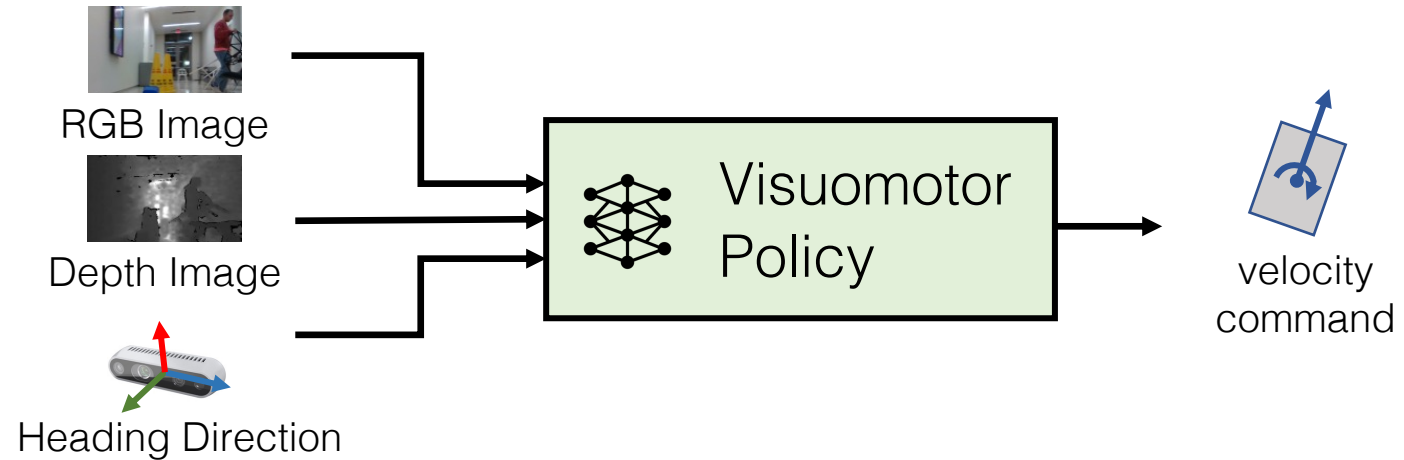
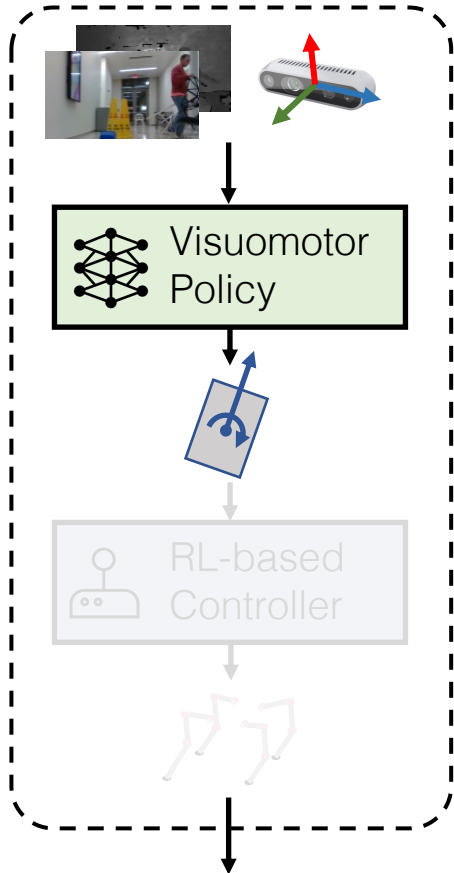
PRELUDE: Perceptive Locomotion Under Dynamic Environments



human demonstrations
of navigation



PRELUDE: Perceptive Locomotion Under Dynamic Environments



imitation learning in **real world**

- ✓ learning complex human navigation behaviors
- ✓ demonstration without robot operation



Part I

Part II

Part III

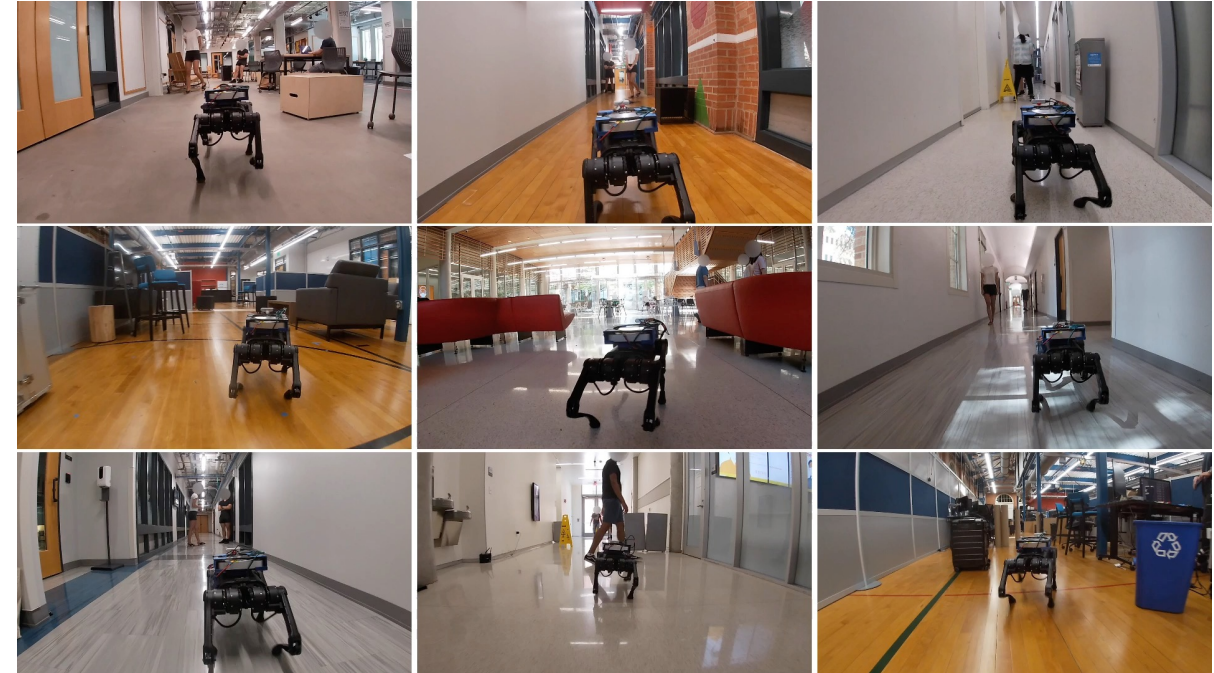
Future Direction

Summary II: Bridging Control and Skill Learning Algorithms



TRILL

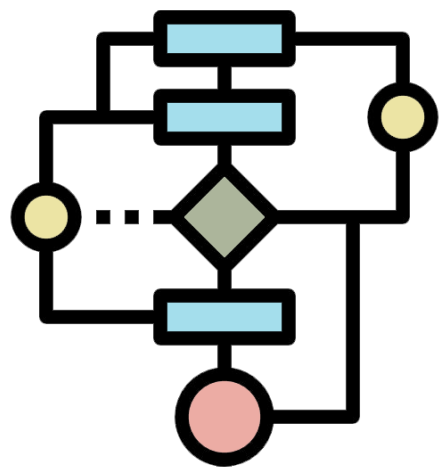
learning dexterous manipulation
with model-based control



PRELUDE

developing agile perceptive locomotion
with learning-based control

Developing a **holistic, embodiment-aware** approach for learning generalizable skills across diverse robots

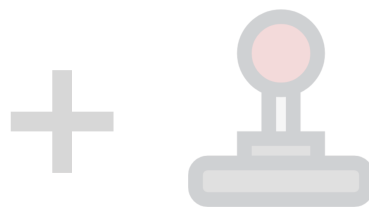


Skill-Learning
Algorithm



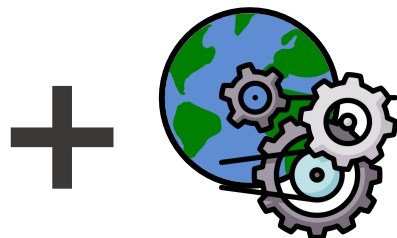
Hardware Design

RA-L 2025
RA-L 2026



Control Architecture

Humanoids 2023: **Best WBC Paper Finalist**
ICRA 2023



Representation Learning

CoRL 2024
ICRA 2025

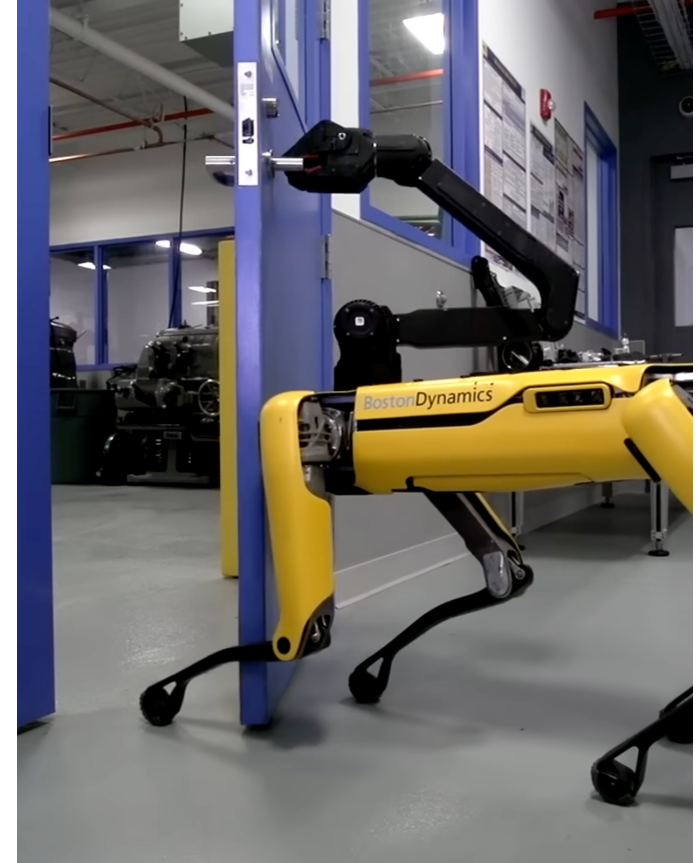
Part I

Part II

Part III

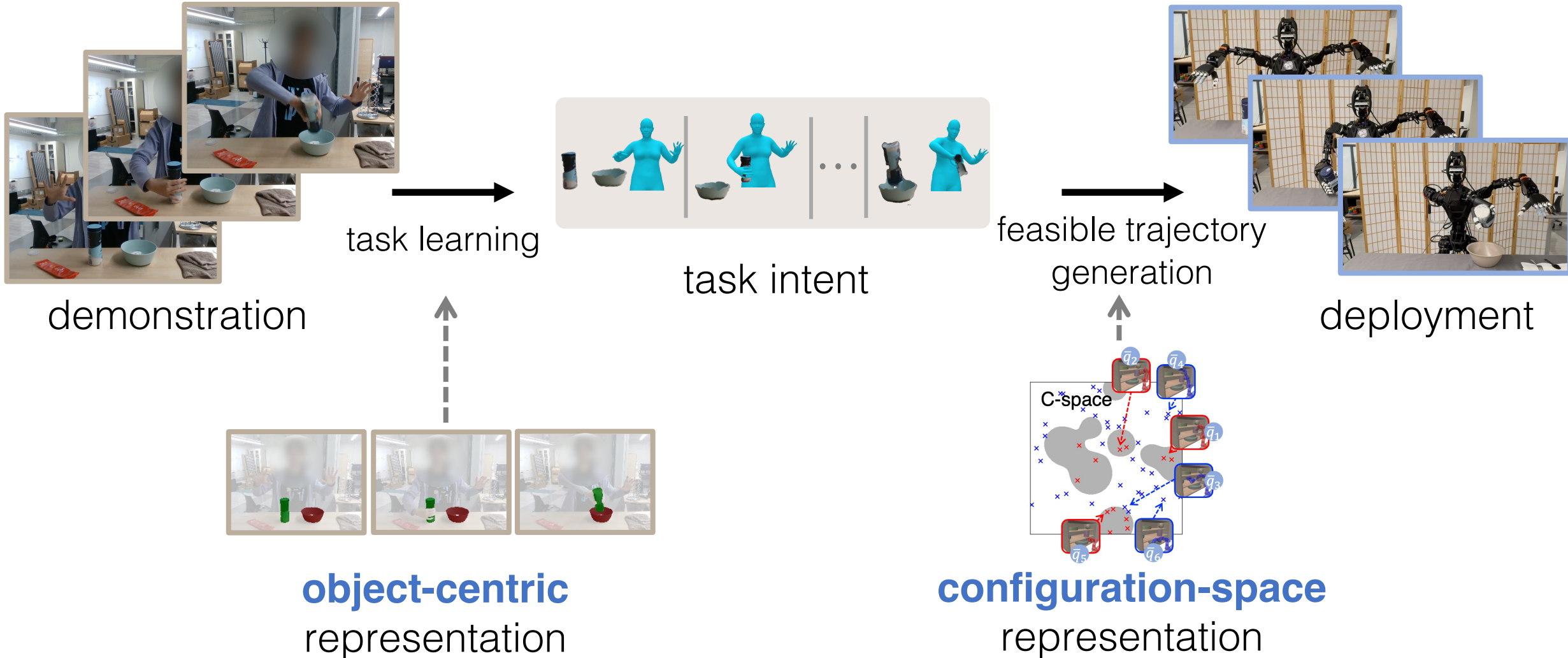
Future Direction

Learning Embodiment-Aware Representation

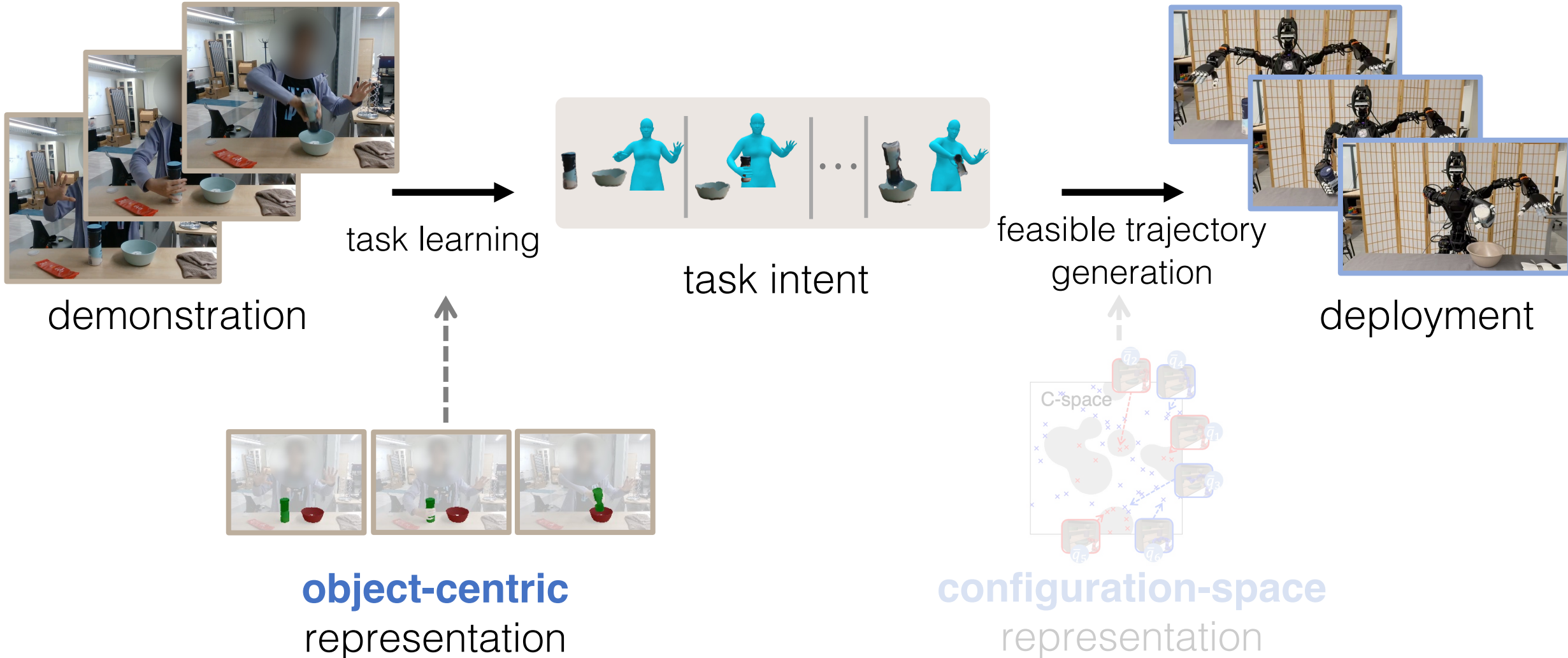


same task, different behavior across embodiments

Learning Embodiment-Aware Representation



Learning Embodiment-Aware Representation





Part I

Part II

Part III

Future Direction

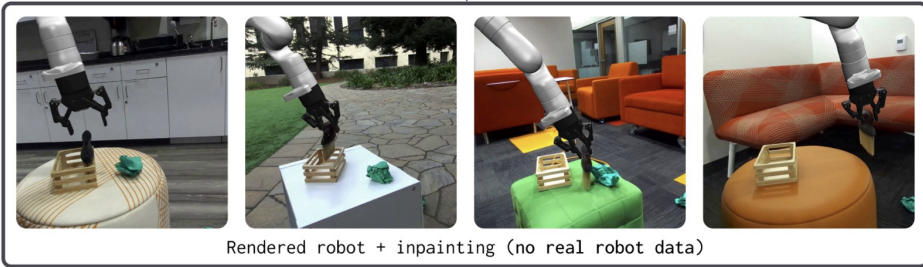
J. Li, ..., M. Seo, et al., CoRL 2024

Capturing Task Intent from Demonstration

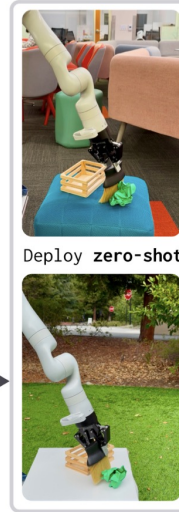
Collect human videos in many scenes



Data-edit



Train π

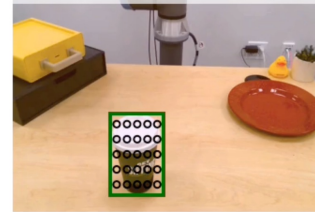


[Phantom, Lepert et al. 2025]

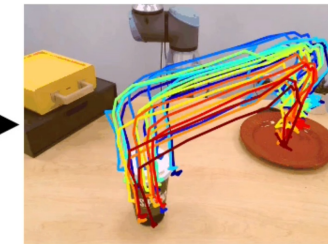
"Put the green cup into the red plate"



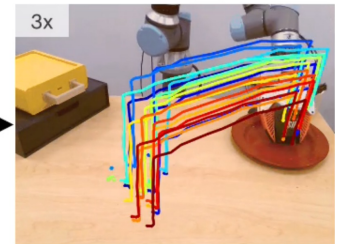
Initial Keypoints Sampling



Generated Flow



Robot Execution



Actions from flow-conditioned diffusion policy

[Im2Flow2Act, Xu et al. 2024]

✓ enforced observation consistency across embodiments

✓ object-centric trajectory representations

📌 manual post-processing of demonstration data

📌 lack of embodiment awareness

Part I

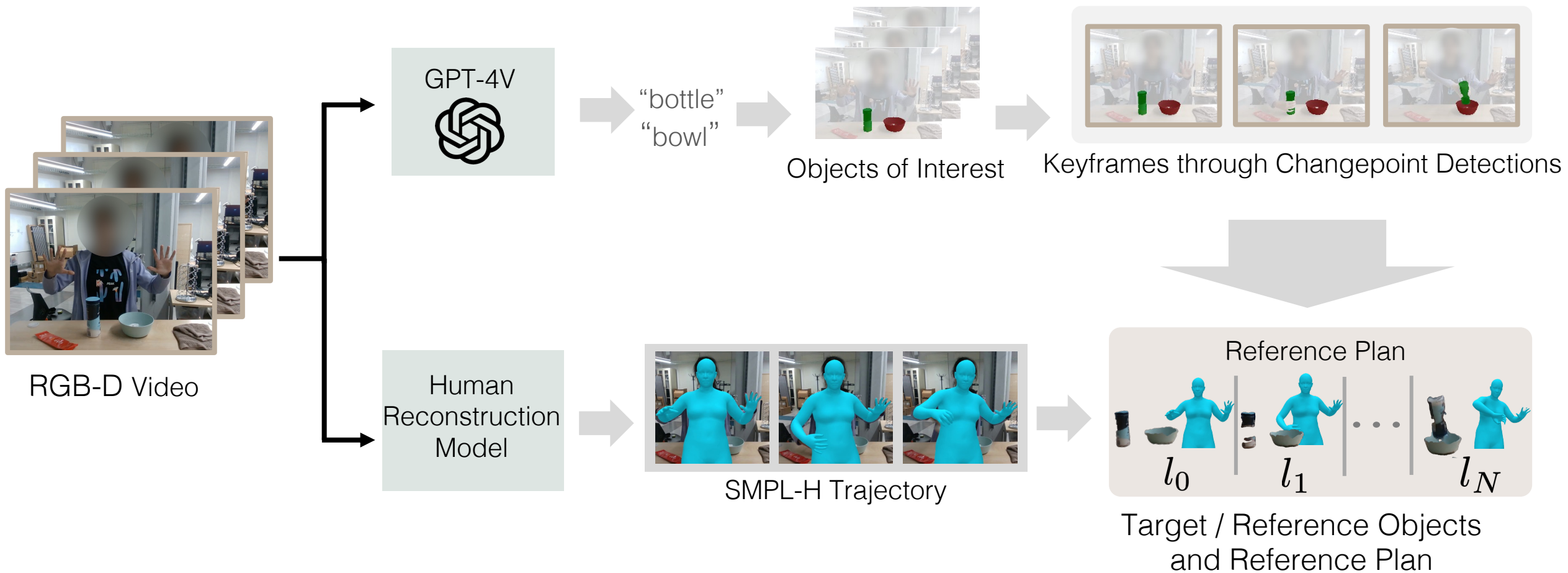
Part II

Part III

Future Direction

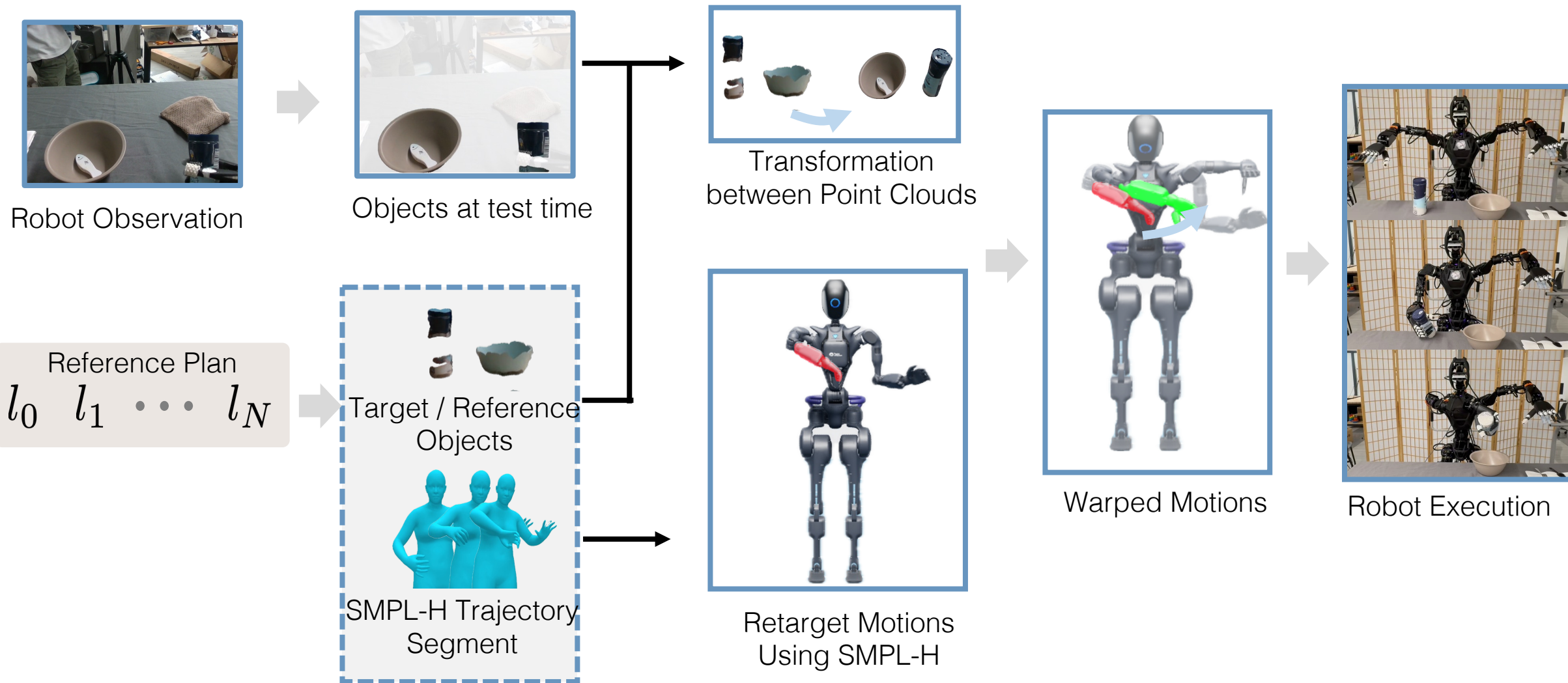
OKAMI:

Object-aware Kinematic Retargeting for humanoid Manipulation Imitation

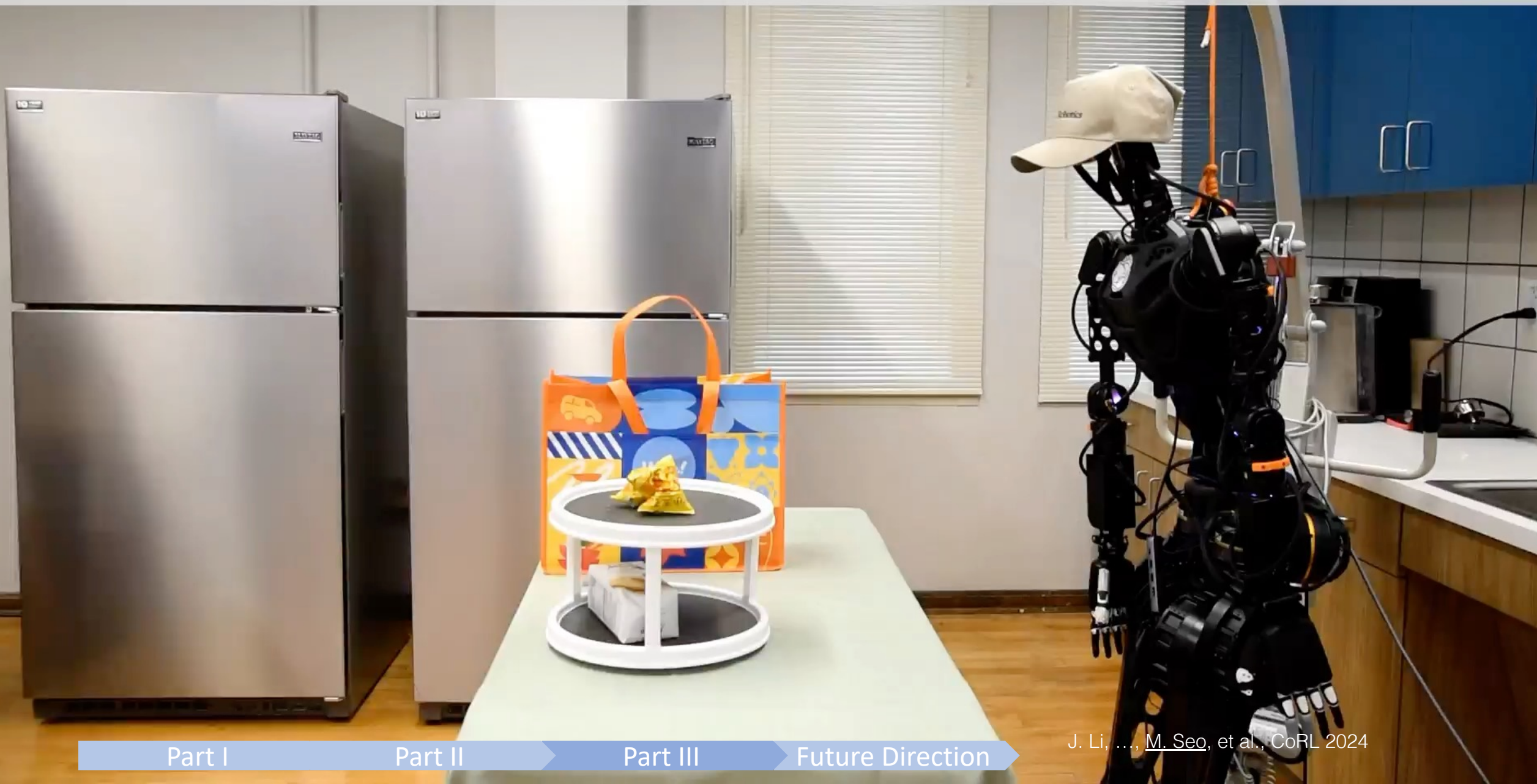


OKAMI:

Object-aware Kinematic Retargeting for humanoid Manipulation Imitation



The humanoid watches a human performing a task.



Part I

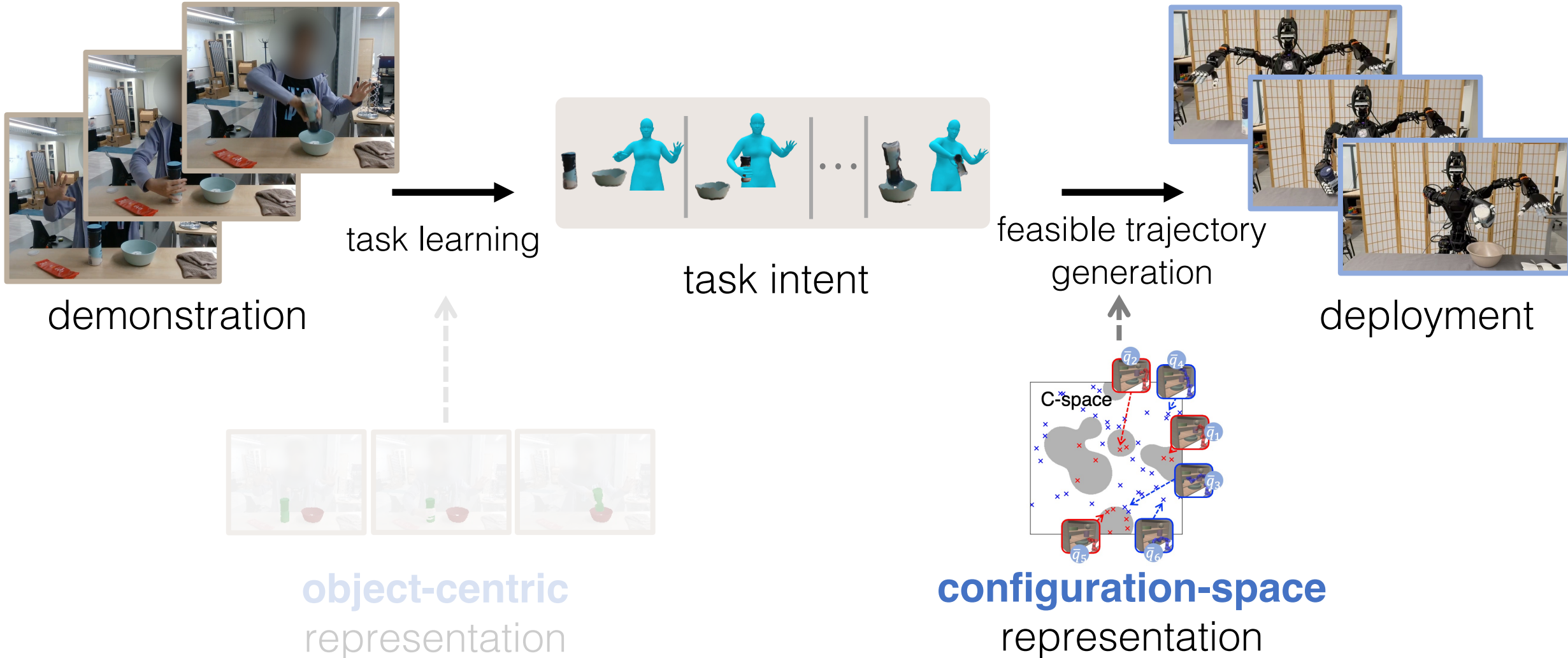
Part II

Part III

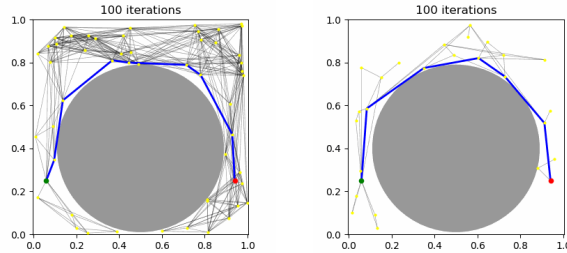
Future Direction

J. Li, ..., M. Seo, et al., CoRL 2024

Learning Embodiment-Aware Representation

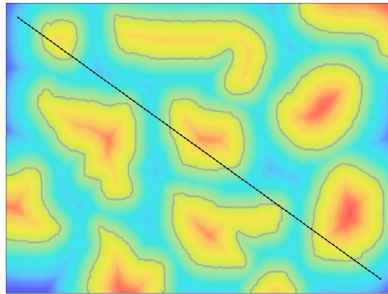


Challenges in Motion Planning



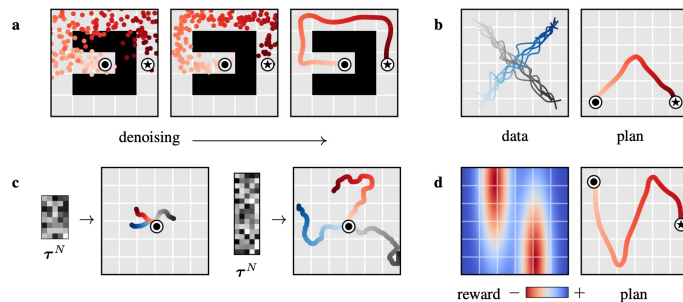
sample-based planners [1,2]

🔍 narrow-passage, high dimensionality



optimization-based planners [3]

🔍 dependence on seed trajectories, local minima



generative models [4]

✓ data-driven trajectory sampling priors

🔍 poor constraint handling, weak generalization

[1] N. Kavraki, et al., Probabilistic roadmaps for path planning in high-dimensional configuration spaces, 1996

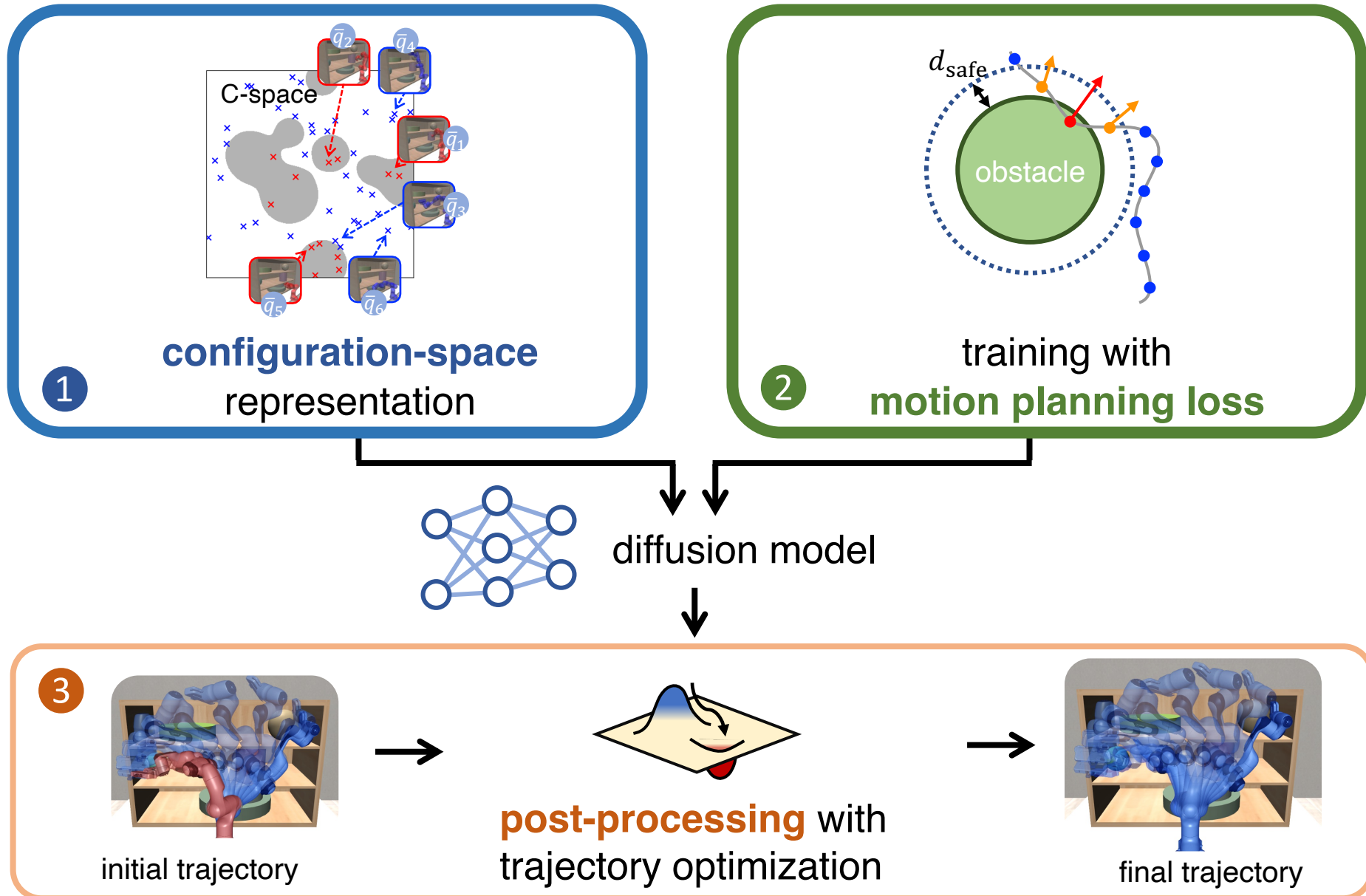
[2] S. LaValle, et al., Rapidly-exploring random trees: A new tool for path planning, 1998

[3] J. Schulman, et al., Motion Planning with Sequential Convex Optimization and Convex Collision Checking, 2013

[4] M. Janner, et al., Planning with Diffusion for Flexible Behavior Synthesis, 2022

PRESTO:

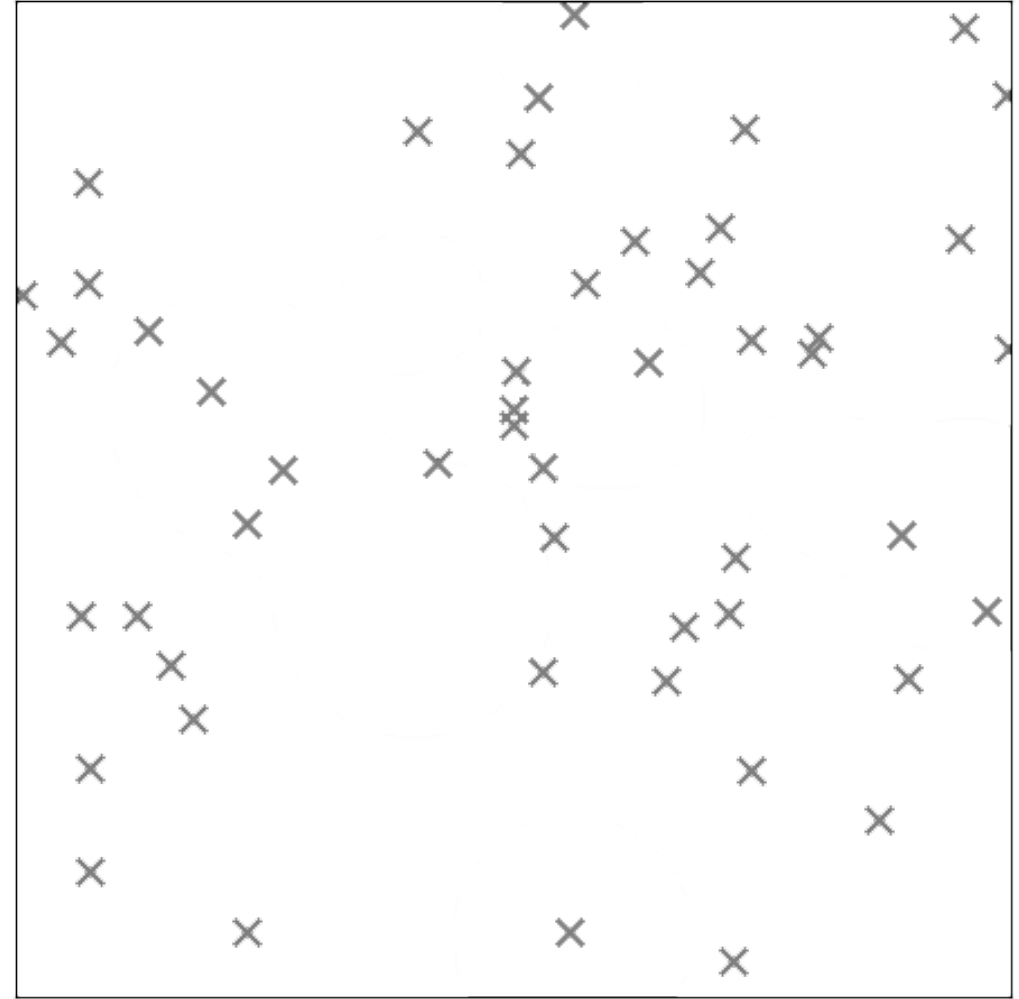
Planning with Environment Representation, Sampling, and Trajectory Optimization



PRESTO:

Planning with Environment Representation, Sampling, and Trajectory Optimization

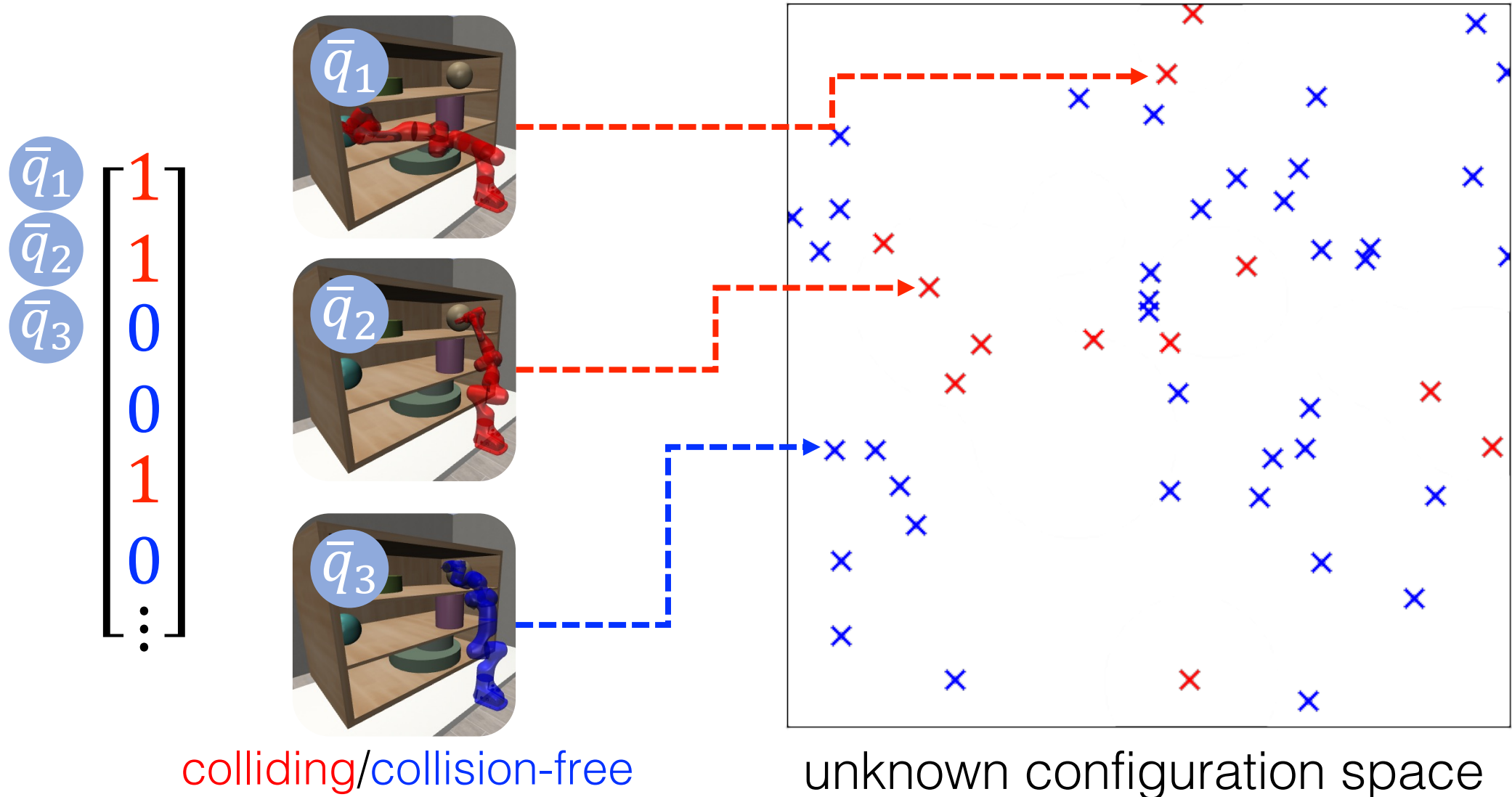
task-related
key configurations



unknown configuration space

PRESTO:

Planning with Environment Representation, Sampling, and Trajectory Optimization

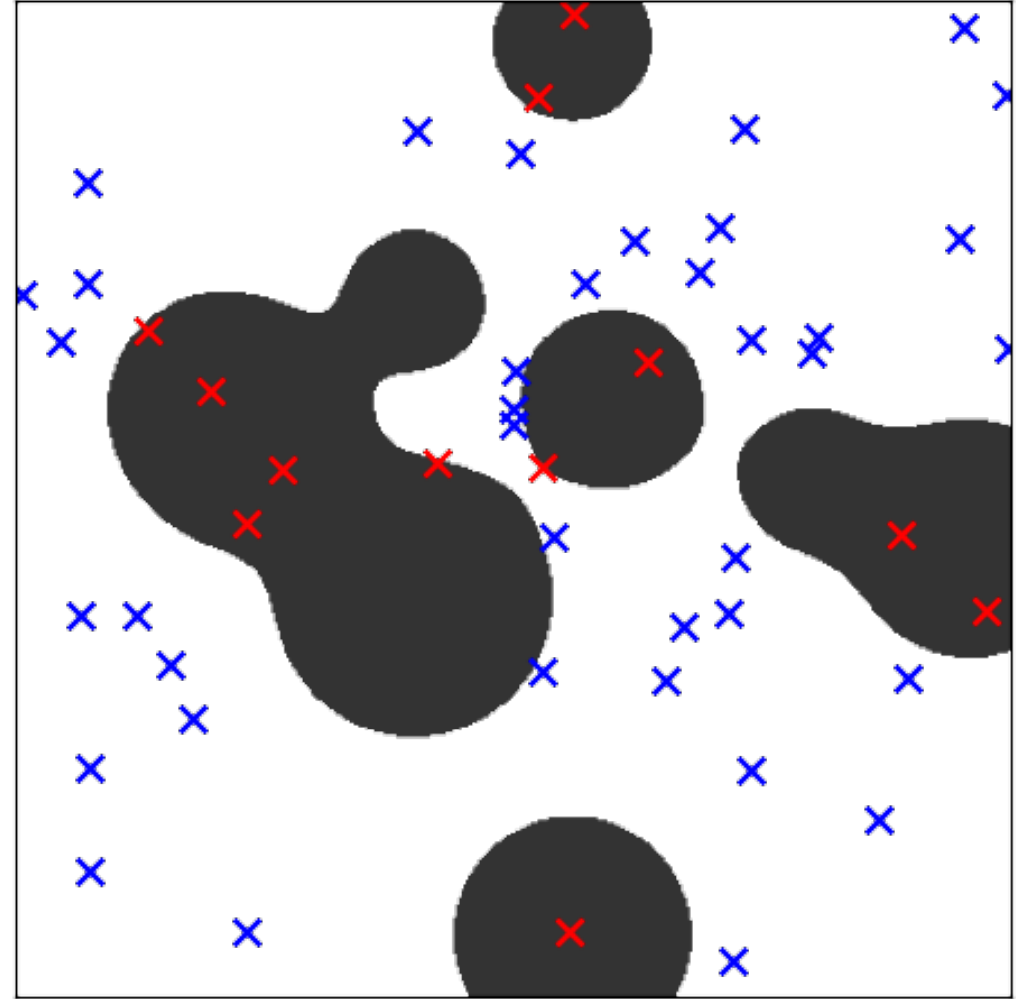


PRESTO:

Planning with Environment Representation, Sampling, and Trajectory Optimization

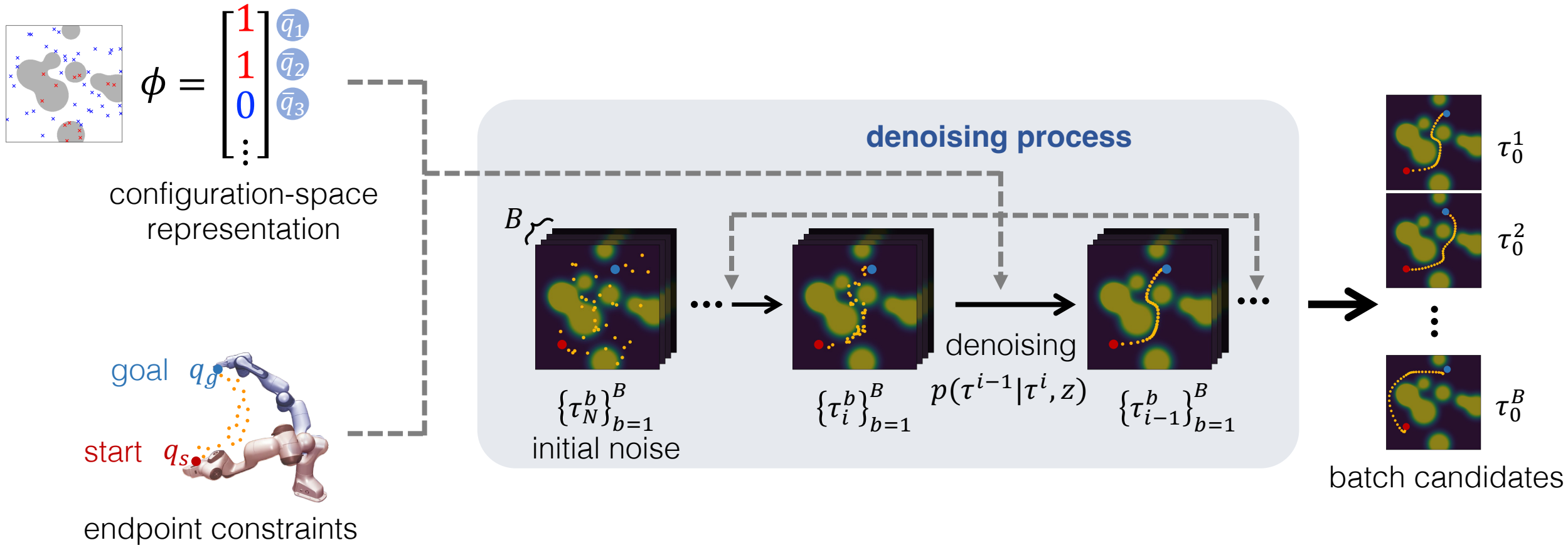
$$\phi = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ \vdots \end{bmatrix}$$

configuration-space
representation



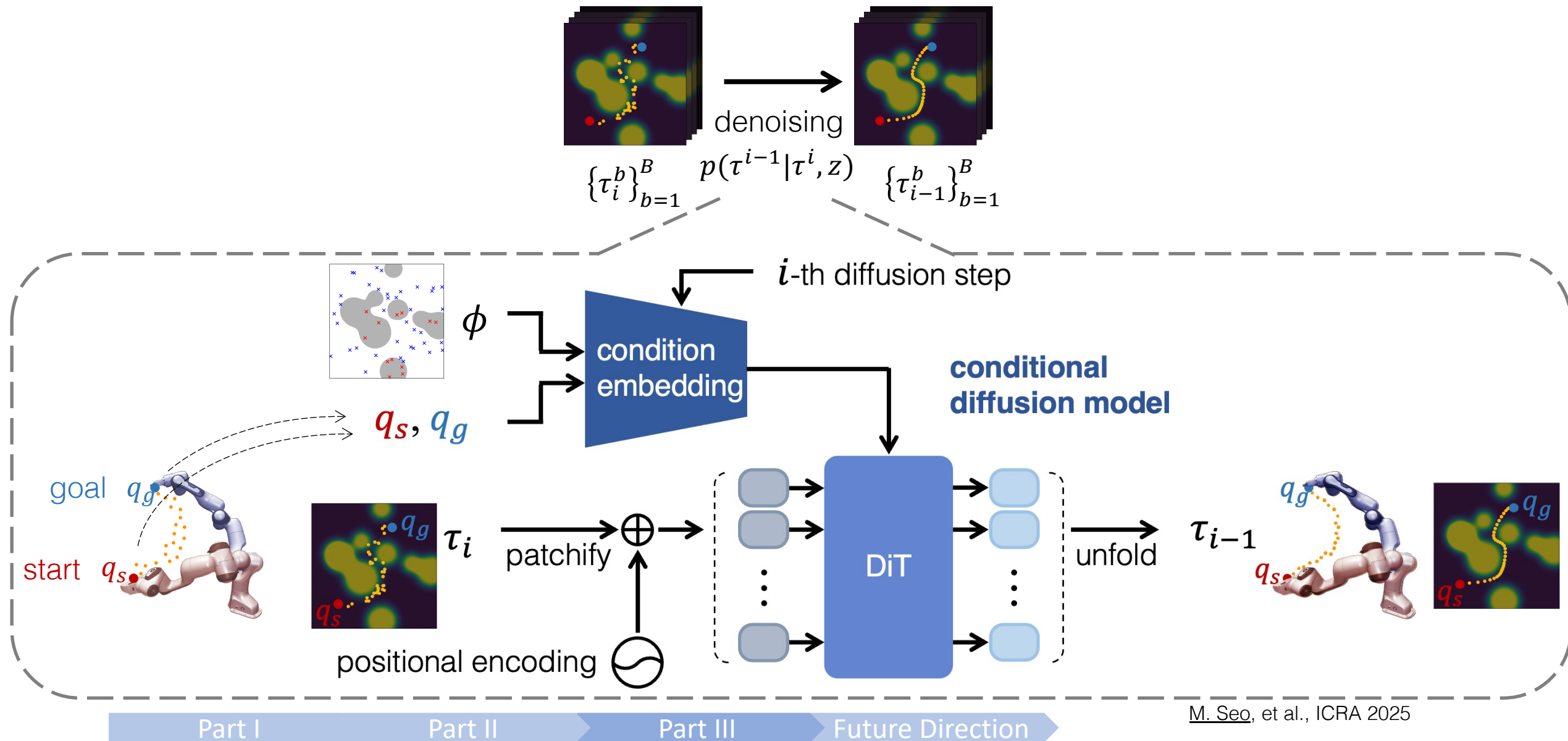
PRESTO:

Planning with Environment Representation, Sampling, and Trajectory Optimization



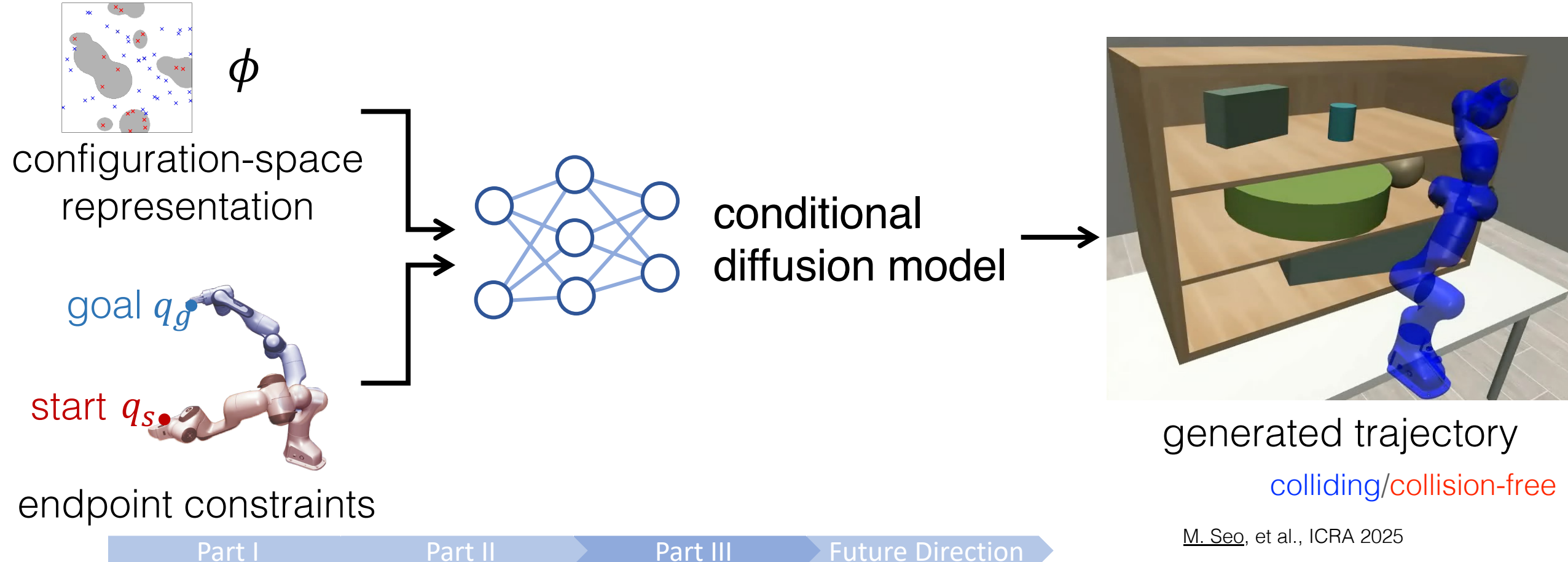
PRESTO:

Planning with Environment Representation, Sampling, and Trajectory Optimization



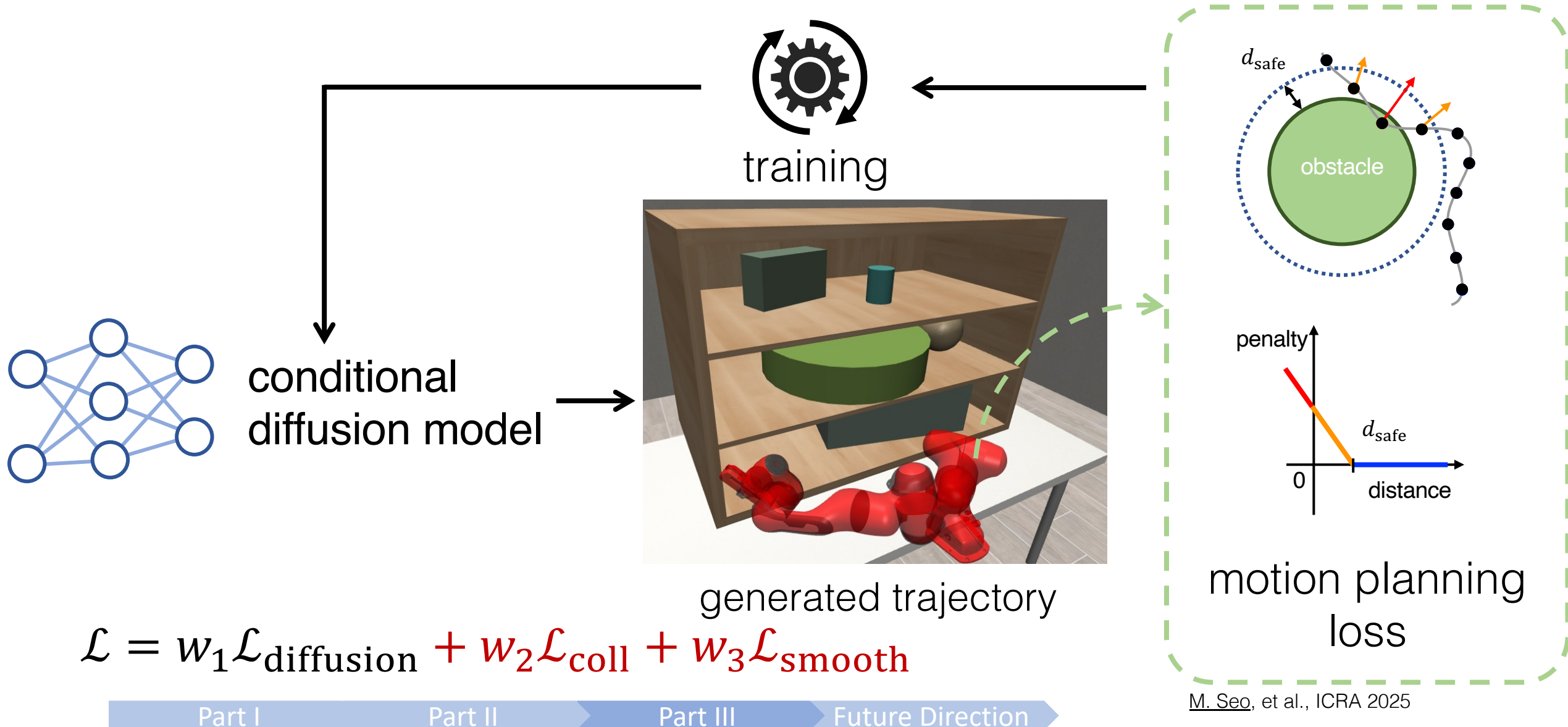
PRESTO:

Planning with Environment Representation, Sampling, and Trajectory Optimization



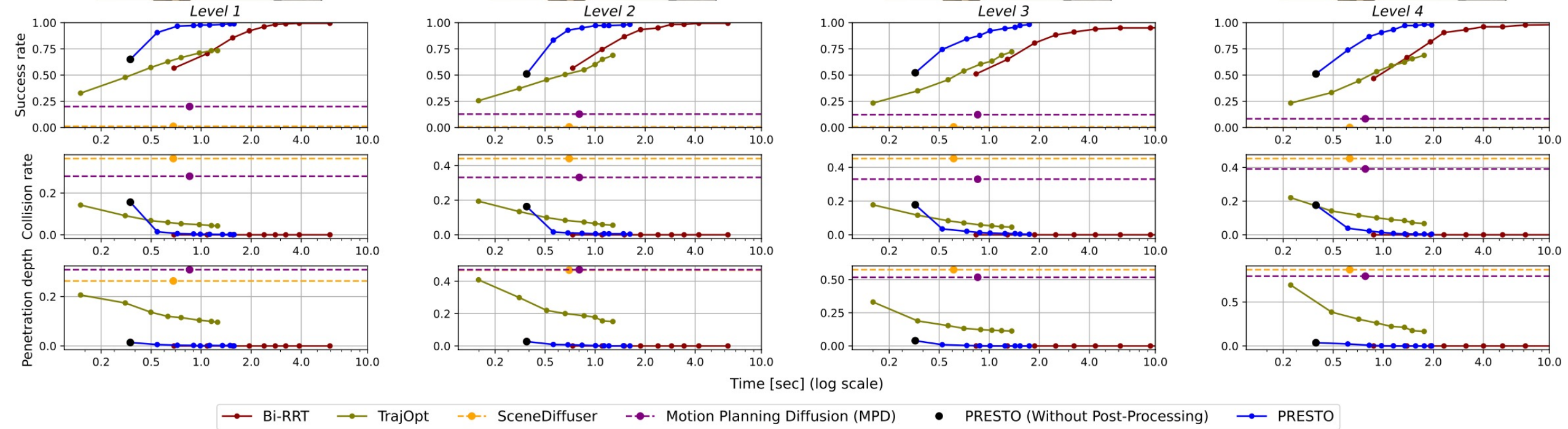
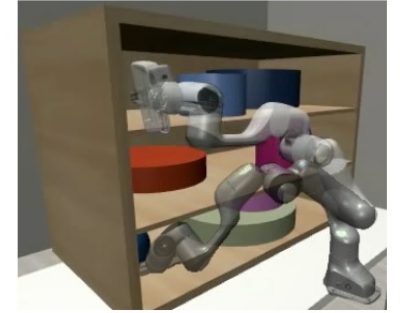
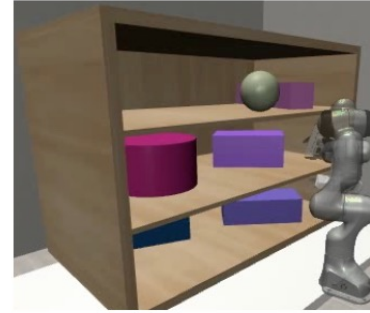
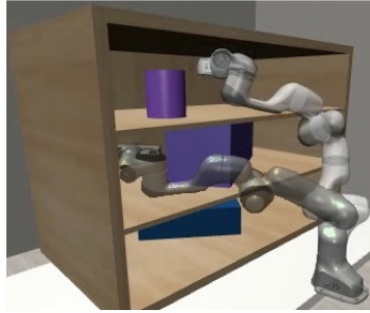
PRESTO:

Planning with Environment Representation, Sampling, and Trajectory Optimization



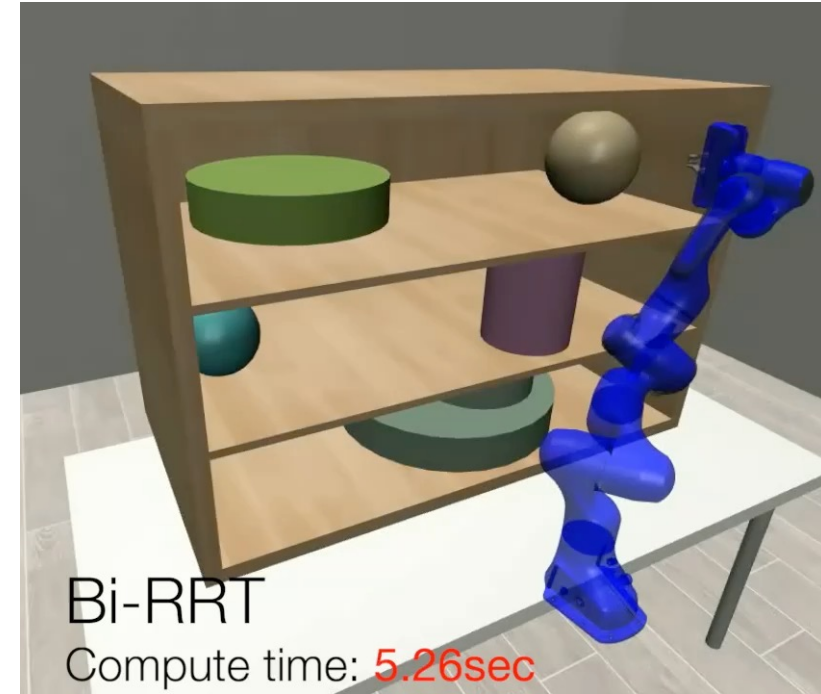
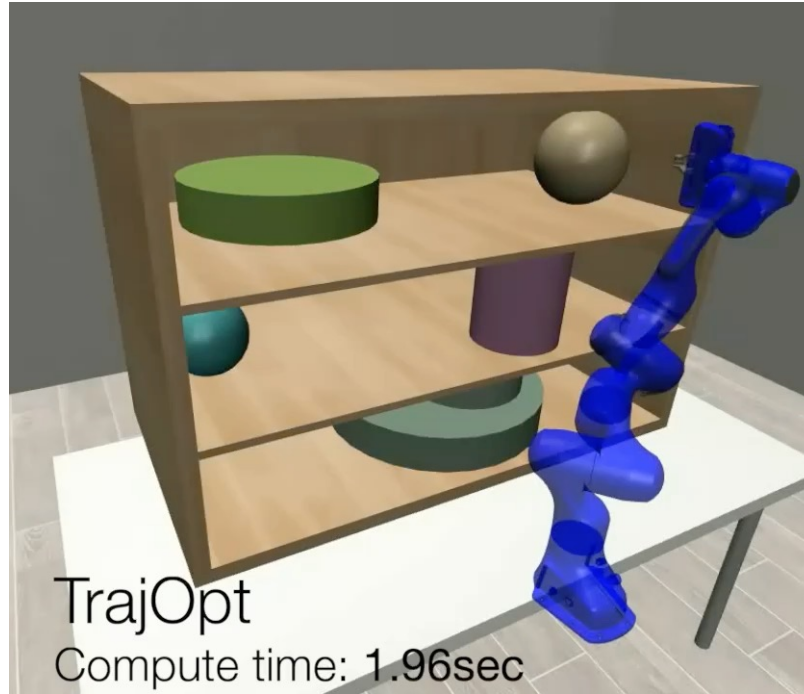
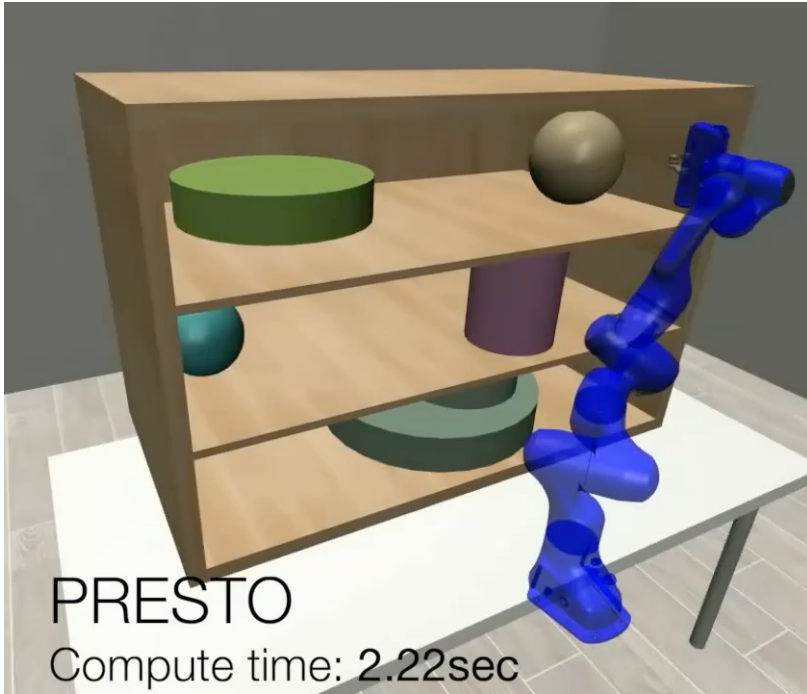
PRESTO:

Planning with Environment Representation, Sampling, and Trajectory Optimization



PRESTO:

Planning with Environment Representation, Sampling, and Trajectory Optimization



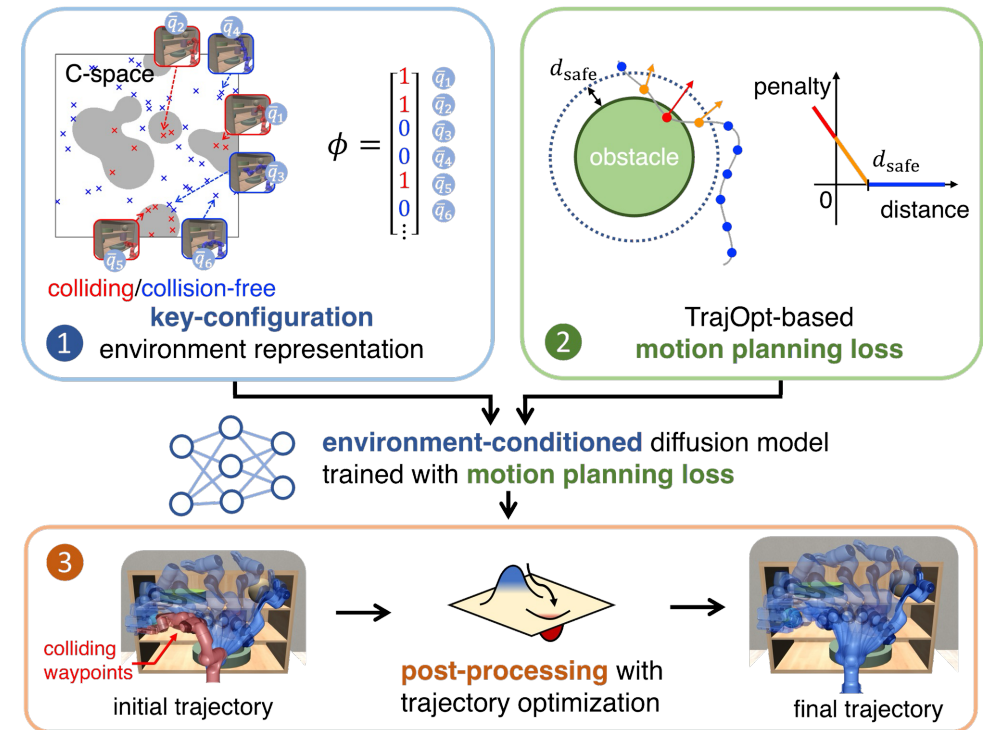
colliding/collision-free

Summary III: Learning Embodiment-Aware Representation



OKAMI

object-centric representation for flexible
cross-embodiment skill transfer



PRESTO

configuration-space representation
for trajectory generation

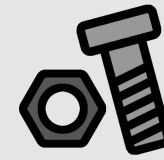
Ongoing and Future Research

generalist
robot autonomy

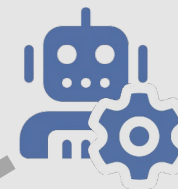


skill learning
across diverse robots

embodiment-aware skill learning



hardware



structural
prior



control



representation

Part I

Part II

Part III

Future Direction

Future Research Agenda

generalist
robot autonomy

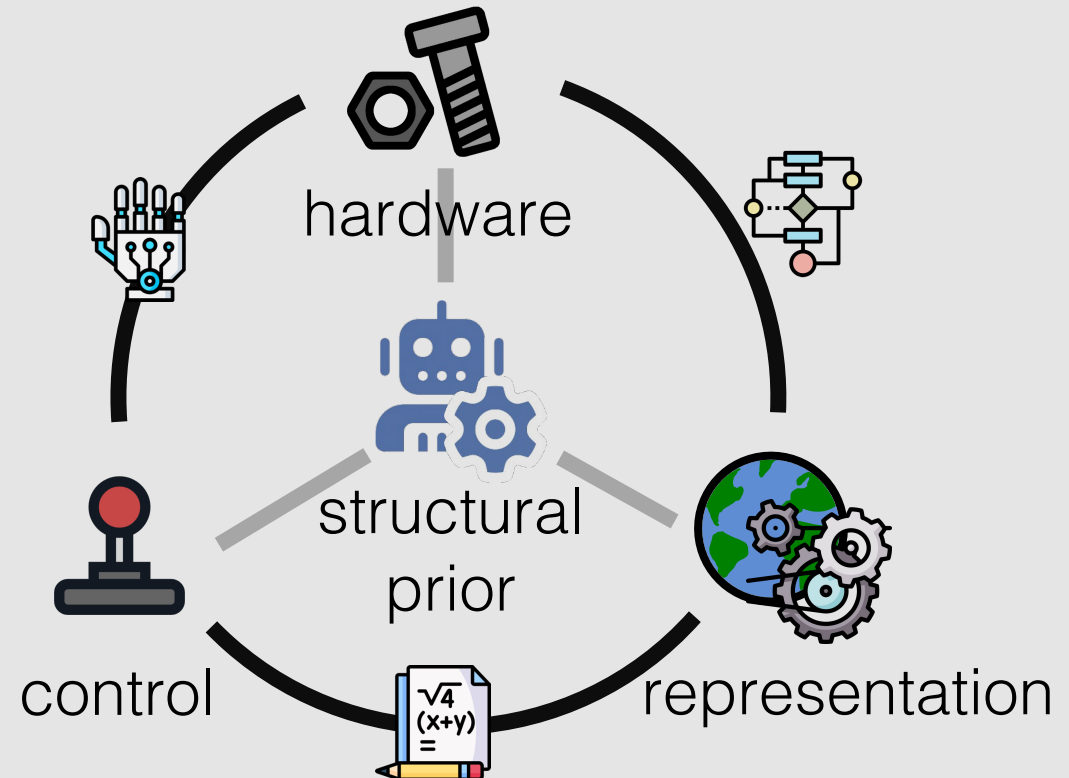


skill composition
across diverse robots



skill learning
across diverse robots

embodiment-aware skill learning



Part I

Part II

Part III

Future Direction

Future Research Agenda

generalist
robot autonomy

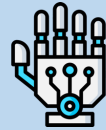


skill composition
across diverse robots

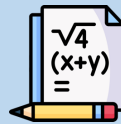


skill learning
across diverse robots

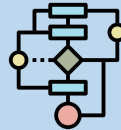
embodiment-aware skill composition



hardware for task diversity



principle of **control** for skill learning



collective **knowledge** sharing

Part I

Part II

Part III

Future Direction

Future Research Agenda

generalist
robot autonomy

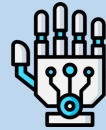


skill composition
across diverse robots



skill learning
across diverse robots

embodiment-aware skill composition



hardware for task diversity



principle of control for skill learning



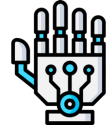
collective knowledge sharing

Part I

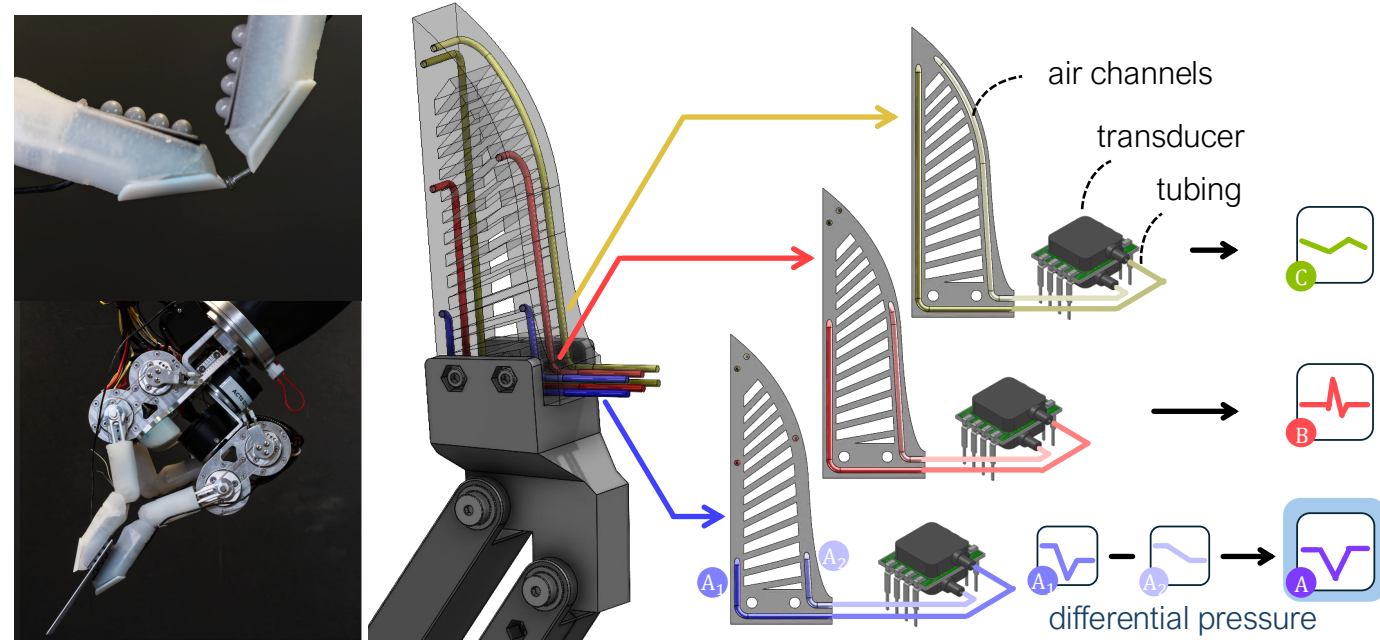
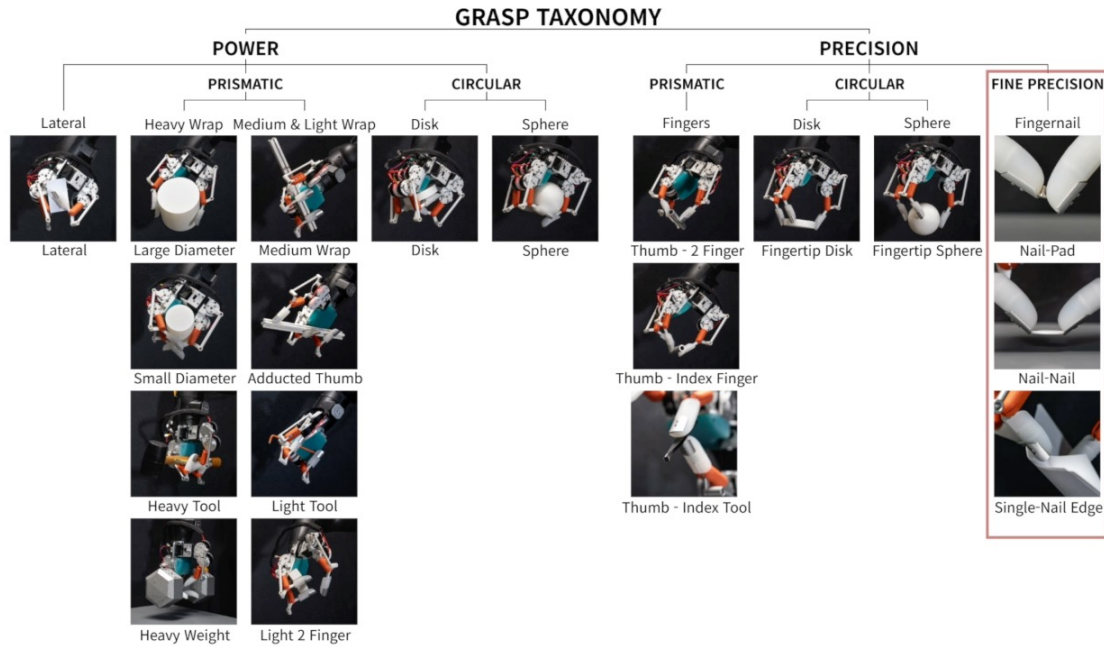
Part II

Part III

Future Direction



Designing Hardware for Task Diversity



- 💡 grasp taxonomies and sensor placement for learning dexterous skills
- 💡 universal multi-finger hands with integrated tactile sensing

D. Kang, A. Kim, M. Seo, et al., 2026
A. Kim, D. Kang, M. Seo, et al., 2026
S. Shang, M. Seo, et al., RA-L 2025

Part I

Part II

Part III

Future Direction

Future Research Agenda

generalist
robot autonomy



skill composition
across diverse robots

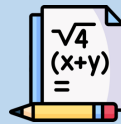


skill learning
across diverse robots

embodiment-aware skill composition



hardware for task diversity



principle of control for skill learning



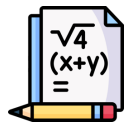
collective knowledge sharing

Part I

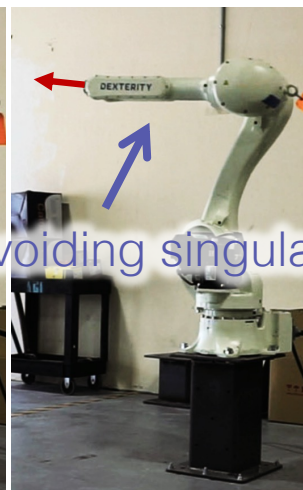
Part II

Part III

Future Direction



Devising Principles of **Control** for Skill Learning



- 💡 theoretical and empirical principles underlying control methods in skill learning
- 💡 co-design of control architecture and skill learning algorithms

Future Research Agenda

generalist
robot autonomy



skill composition
across diverse robots



skill learning
across diverse robots

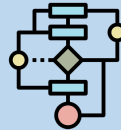
embodiment-aware skill composition



hardware for task diversity



principle of control for skill learning



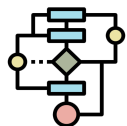
collective knowledge sharing

Part I

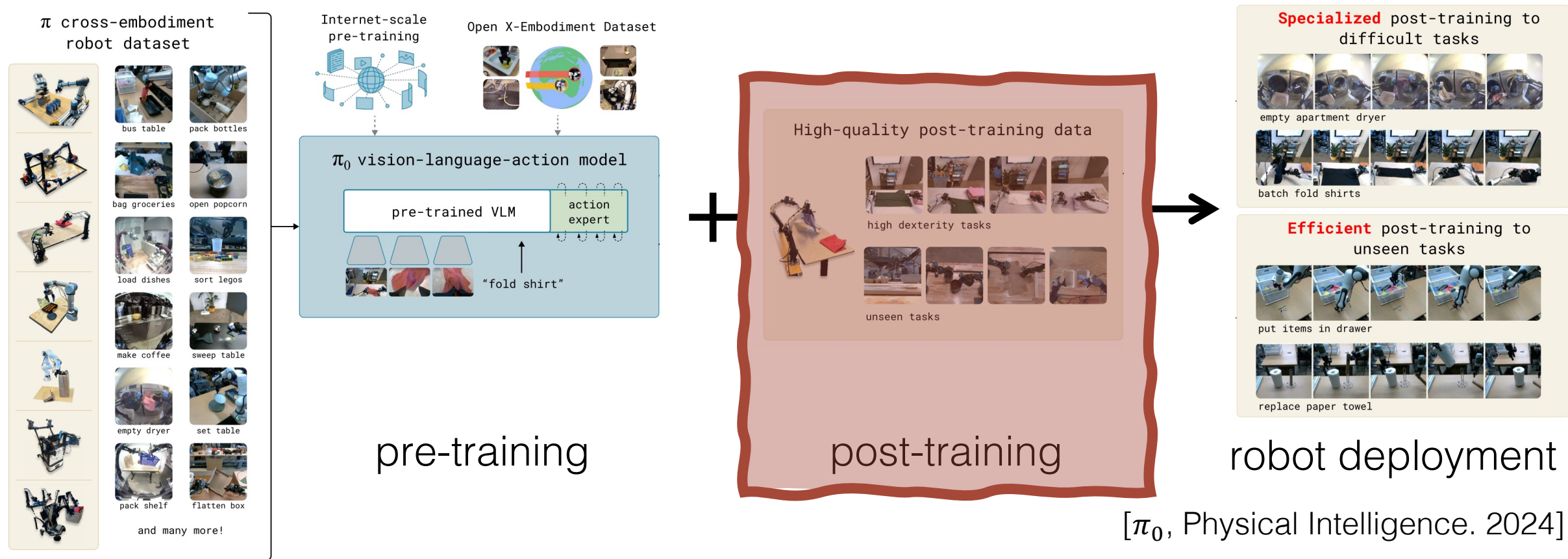
Part II

Part III

Future Direction



Scaling Knowledge Sharing with Embodiment Awareness



- 💡 extending embodiment-aware learning to diverse long-horizon tasks
- 💡 cost-efficient data-collection for robotic foundation models

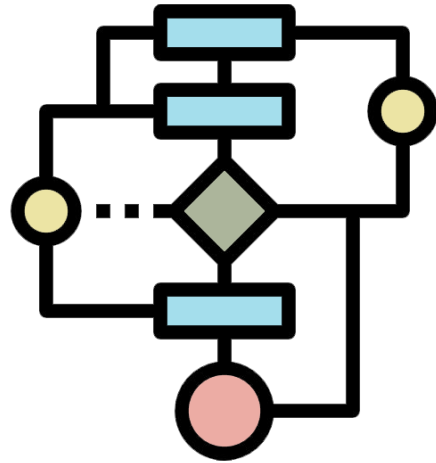
Part I

Part II

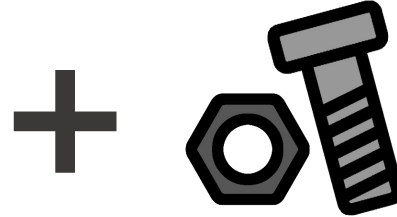
Part III

Future Direction

Developing a **holistic, embodiment-aware** approach for learning generalizable skills across diverse robots

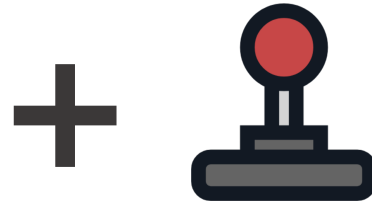


Skill-Learning
Algorithm



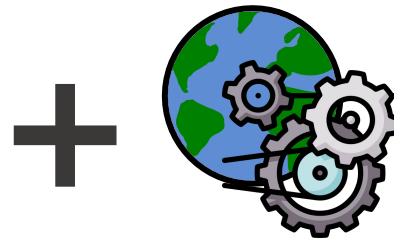
Hardware Design

RA-L 2025
RA-L 2026



Control Architecture

Humanoids 2023: **Best WBC Paper Finalist**
ICRA 2023



Representation Learning

CoRL 2024
ICRA 2025

Acknowledgement

