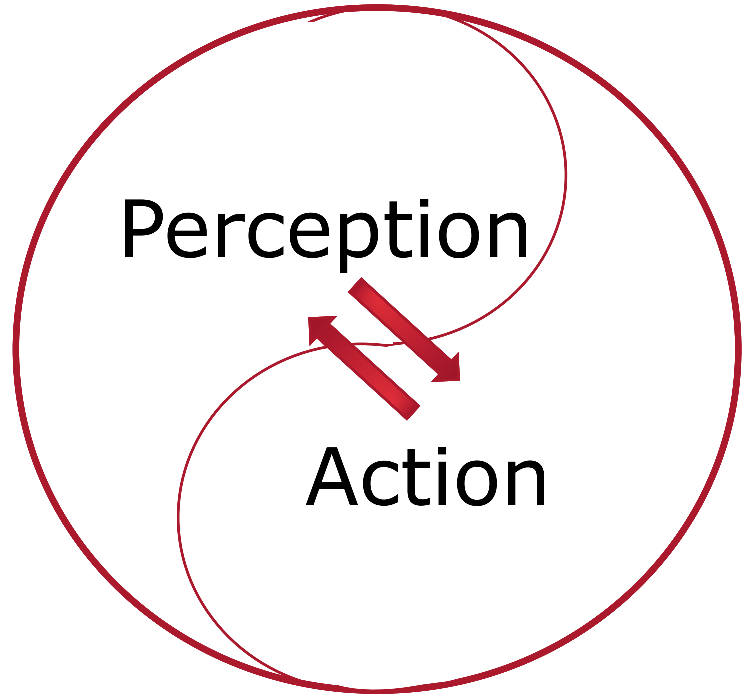
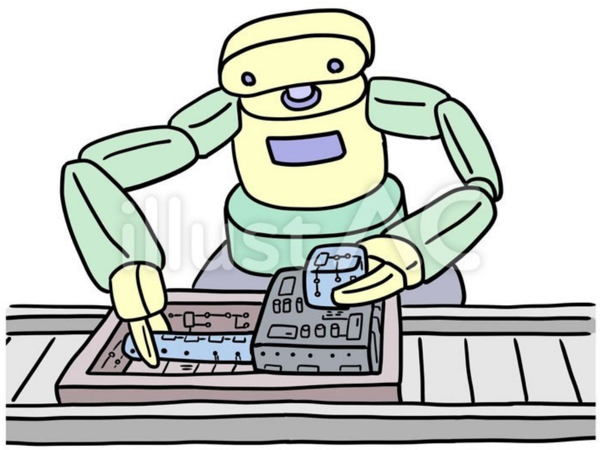
Learning in perception and action loop for efficient manipulation with uncertainty

Jing Xiao
Robotics Engineering Department
Worcester Polytechnic Institute
November 20, 2025

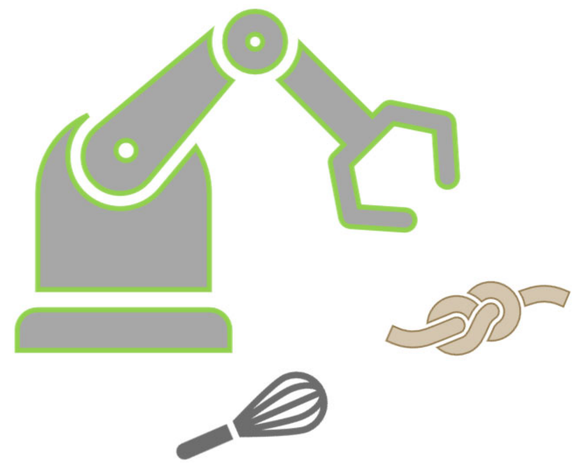


High touch
High precision

Learning in Real-world for Manipulation Tasks



Robotic Assembly

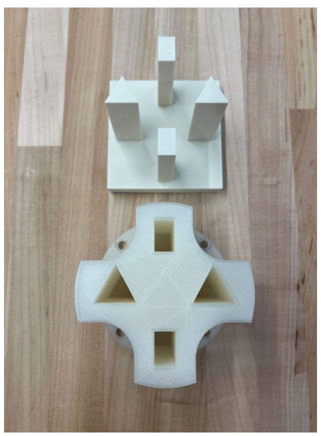


Manipulation

beyond pick-and-place

Robotic Assembly

Challenge: Tight tolerance vs**. Great** uncertainty effect, varied part geometry and complex contact states



Classic approaches:

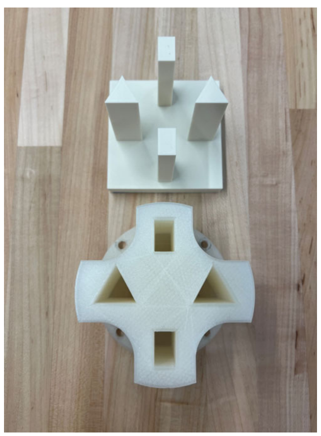
- passive, active compliance, force, impedance control
- simple-shape, single peg-in-hole tasks

Learning-based approaches:

- task-dependent, specific cases

Robotic Assembly

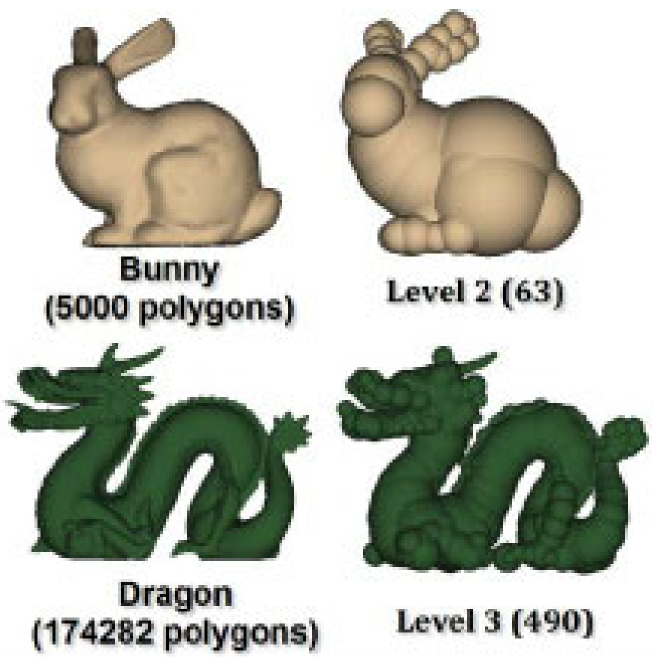
Challenge: Tight tolerance vs**. Great** uncertainty effect, varied part geometry and complex contact states



Can we have a general and efficient online solution?

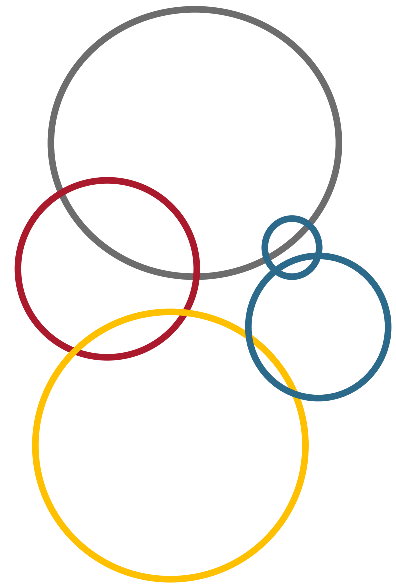
- Efficient representation of contacts
- **Perception** with learning
- Constrained search

Sphere-Tree Representation



[Wang *et al*. IEEE ToH 2011]

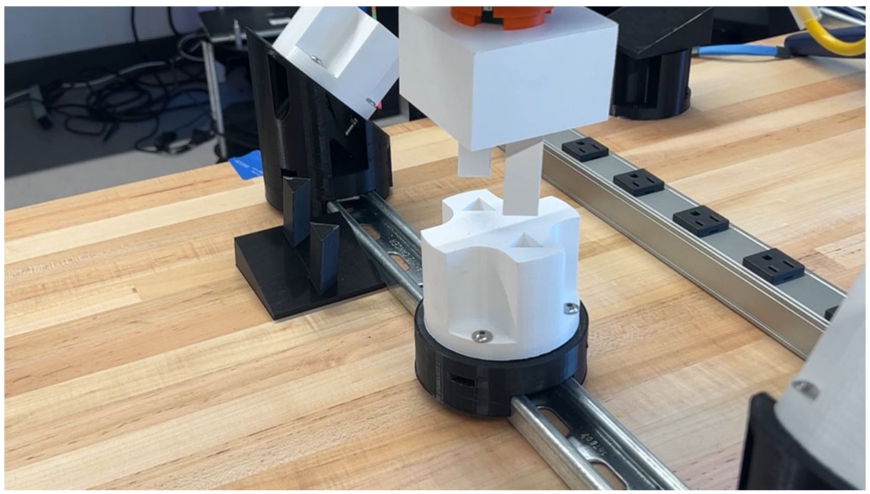
Complex geometry
& contacts



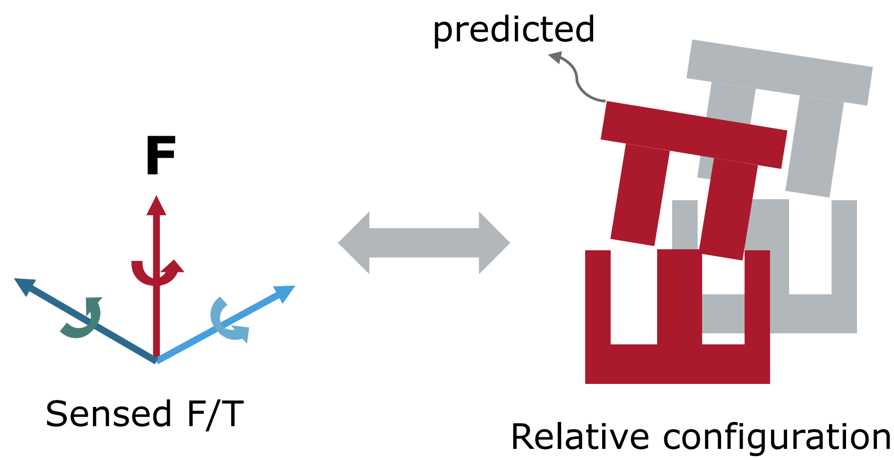
Unified contact
constraints

WPI

“General” Learning to Perceive Contact Config.

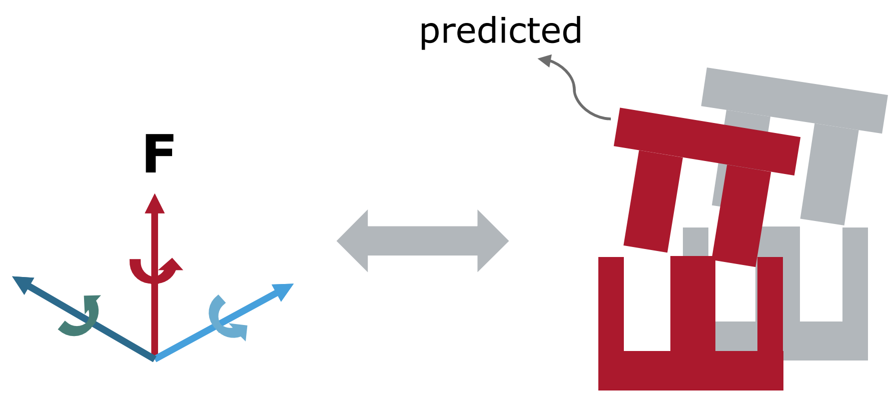


Automatic data collection

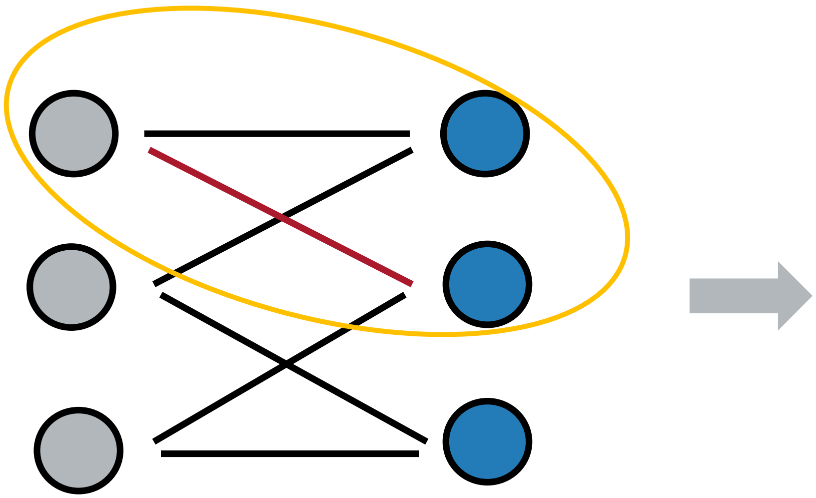


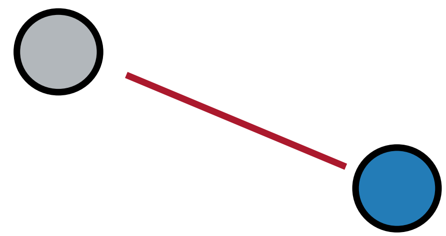
“General” Learning to perceive contact config.

- Object-independent: learned NN applies to different object assembly
- Robot-independent: learned mapping is to relative configurations in Cartesian space



Constrained Search to Resolve Ambiguity





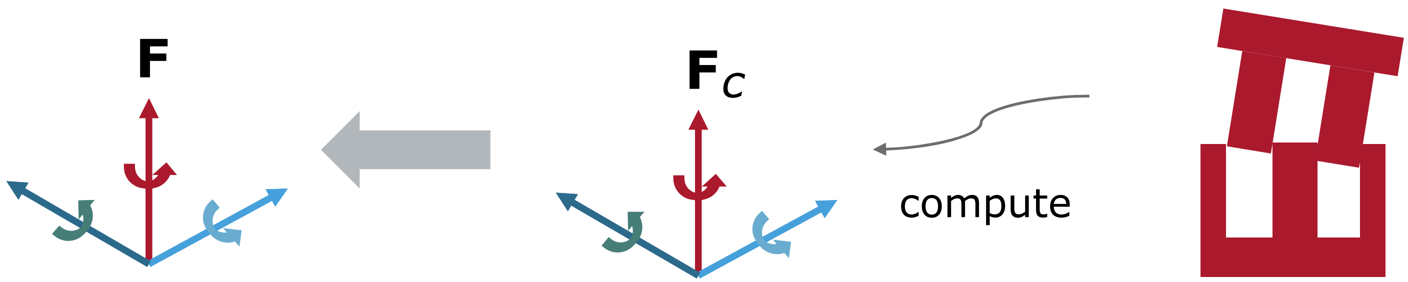
Find the most possible
contact configuration

$$\begin{cases} \text{minimize } E(q, q^*) \\ \text{subject to} \quad S_{obj_1} \bigcap S_{obj_2} = \varnothing \end{cases}$$

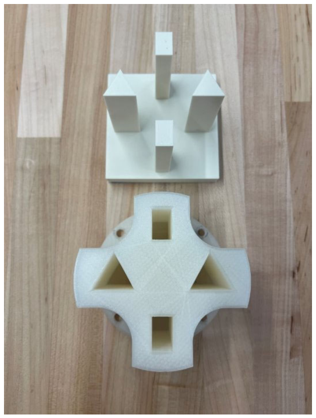
Closing the Loop by Learning

Force Calibration

Best Match

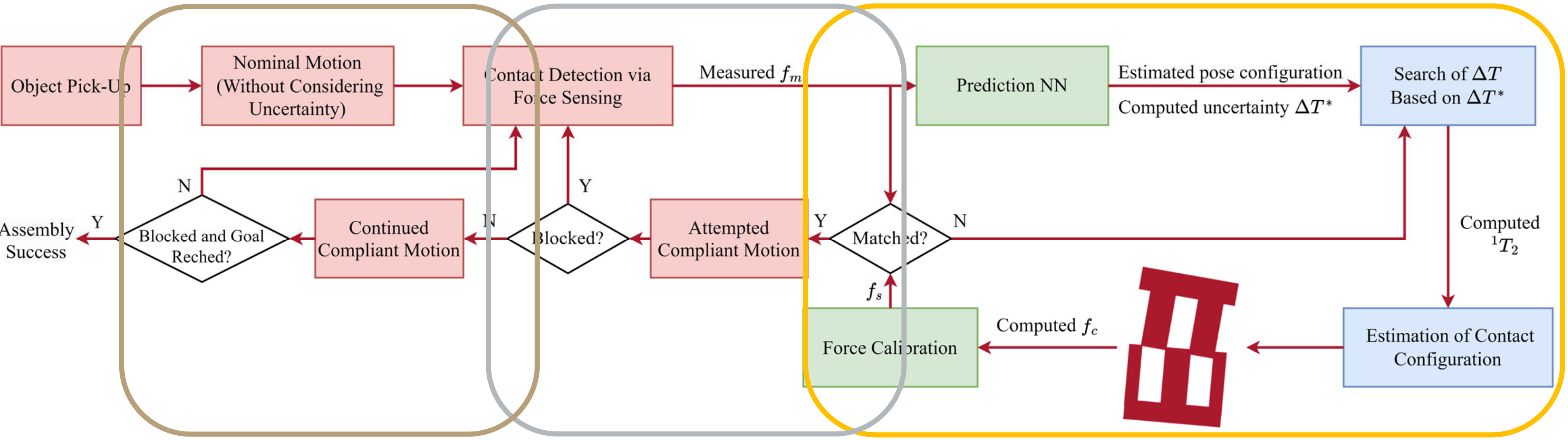


Actual speed



Robust to uncertainty >10 times of task tolerance

Perception-Action Loops



- Red: sensing and robot actions
- **Green**: **learned models**
- **Blue**: **online algorithms**

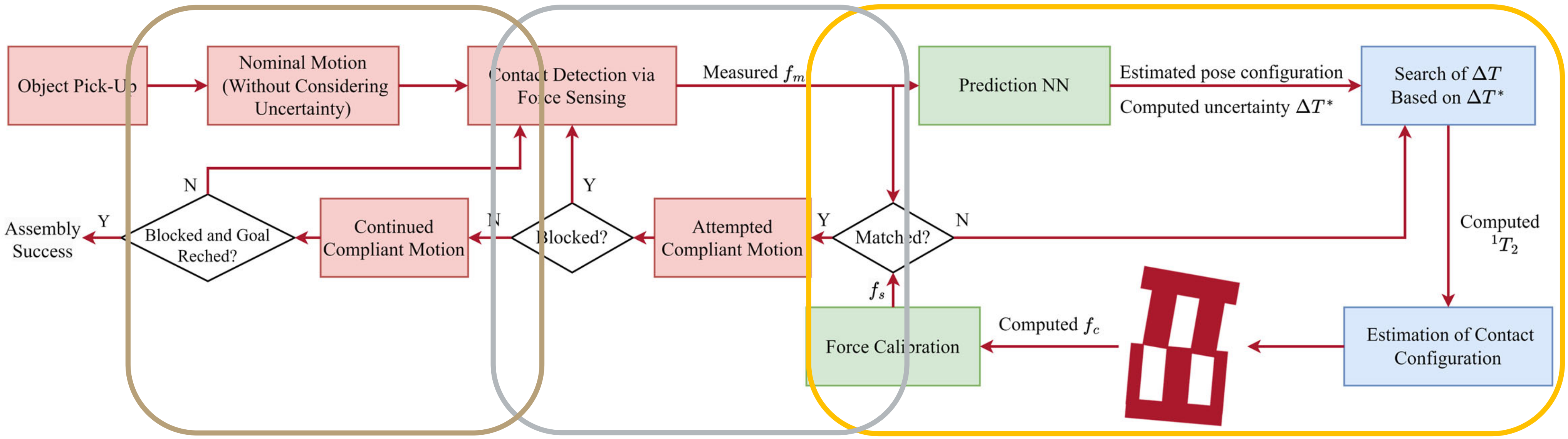
Average runtime for mixed-shape 4-peg-in-hole assembly
(randomized uncertainty offset in each run)

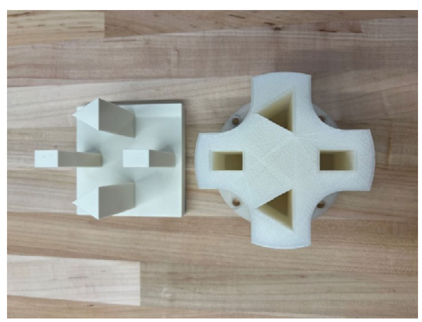
#	Zone	Insertion angle	Average time $\pm$ SD (s) pickup & assembly	Success within 60s
1	$\mathbf{P_1} \rightarrow \mathbf{A_1}$	$0°$	16.0 ± 0.72	20/20
2	$\mathbf{P_1} \rightarrow \mathbf{A_1}$	$20°$	17.4 ± 0.89	19/20
3	$\mathbf{P_1} \rightarrow \mathbf{A_1}$	$25°$	16.2 ± 0.64	19/20
4	$\mathbf{P_1} \rightarrow \mathbf{A_2}$	$0°$	15.8 ± 0.78	20/20
5	$\mathbf{P_2} \rightarrow \mathbf{A_2}$	$25°$	15.8 ± 0.85	20/20
6	$\mathbf{P_2} \rightarrow \mathbf{A_3}$	$0°$	17.6 ± 0.71	20/20
7	$\mathbf{P_2} \rightarrow \mathbf{A_3}$	$90°$	17.2 ± 0.66	18/20
8	$\mathbf{P_2} \rightarrow \mathbf{A_3}$	$25°$	17.8 ± 0.91	19/20

Runtime > 60s
(very large uncertainty)

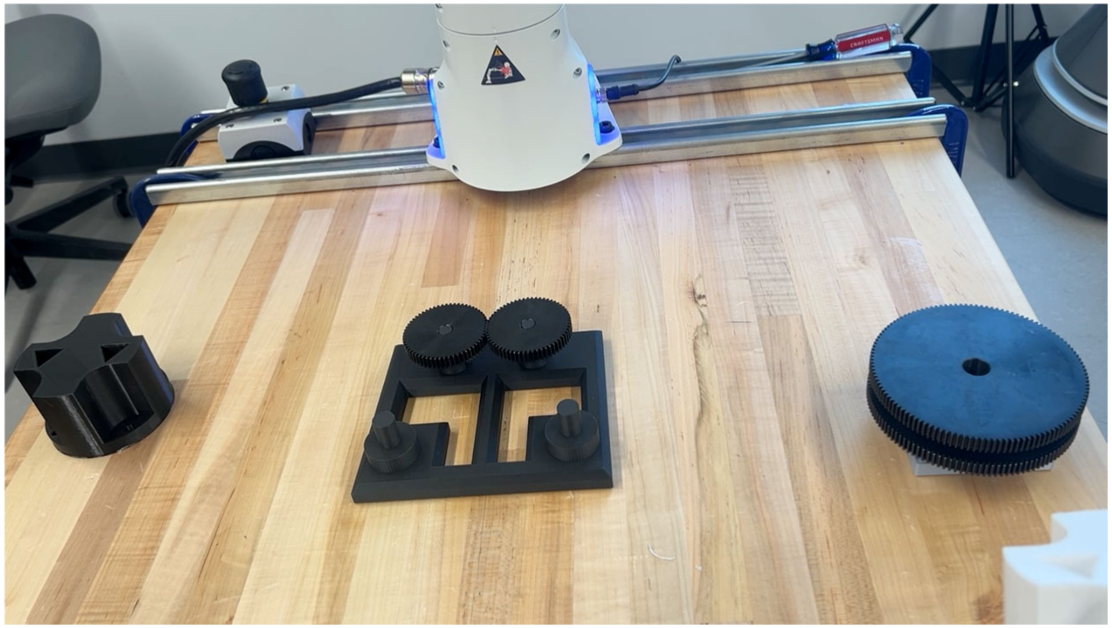
Case	#2	#3	#7	#8
#Runs	1	1	2	1
Max time to succeed (minutes)	2.08	1.3	2.57	3.08

Perception-Action Loops



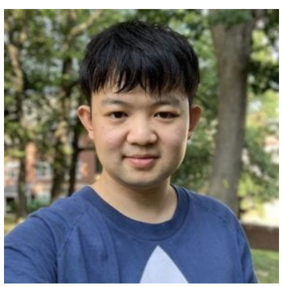


Actual speed



Zoom-out view

Zoom-in view



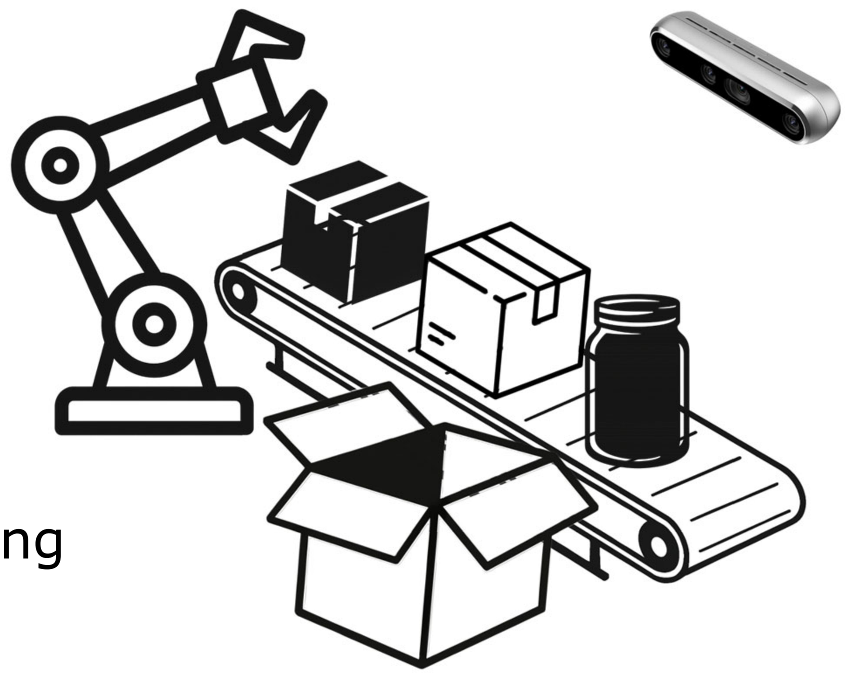
Clearance:

- Position: 0.25mm
- Orientation: 0.016rad

Robust to uncertainty >
10 times of task tolerance

Shichen Cao and Jing Xiao, "On efficient and flexible autonomous robotic insertion assembly in the presence of uncertainty," *IEEE Robotics and Automation Letters*, 2024.

Dynamic Dense Packing of Novel Objects



RGB-D perception
& model building

Dynamic packing
planning

Robotic insertion assembly

Real-time Dense Packing of Novel Objects

RGB-D Perception

- Object model
- Object Pose

Rough, imprecise model

→

Dense Packing Planning

- Maximizing space
- Optimizing motion path

Use rough models

→

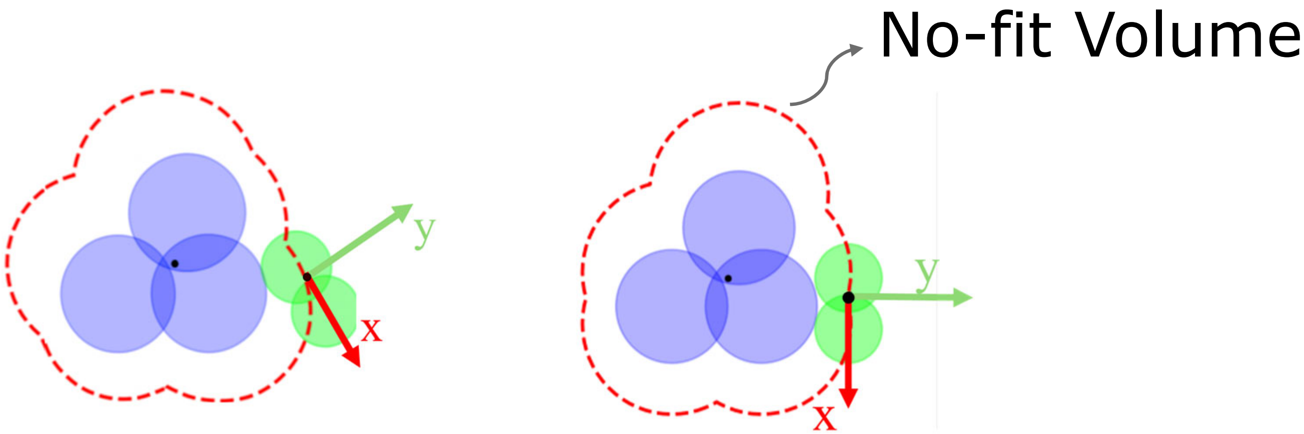
Insertion with F/T Sensing

- Contact state prediction
- Compliant motion

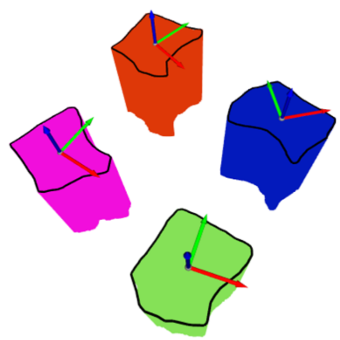
Handle effects of inaccuracies and uncertainties

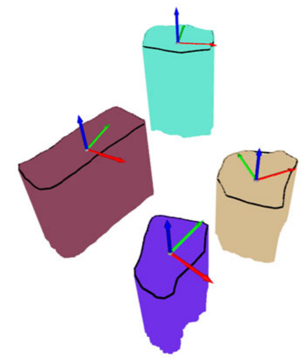
Packing Planning

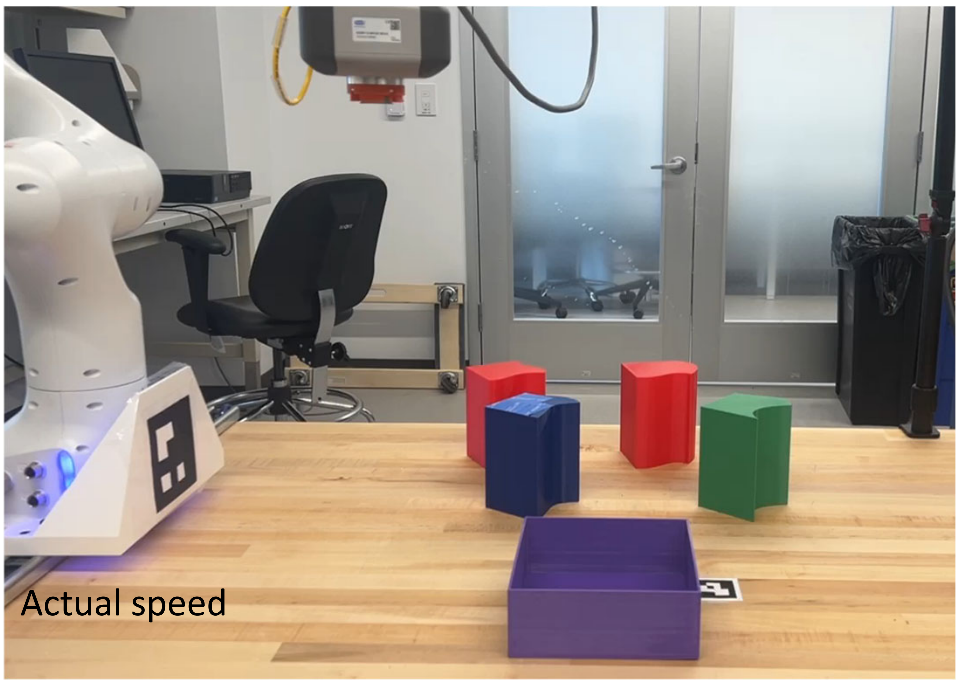
Sequencing: prioritize objects short, large, difficult to fit

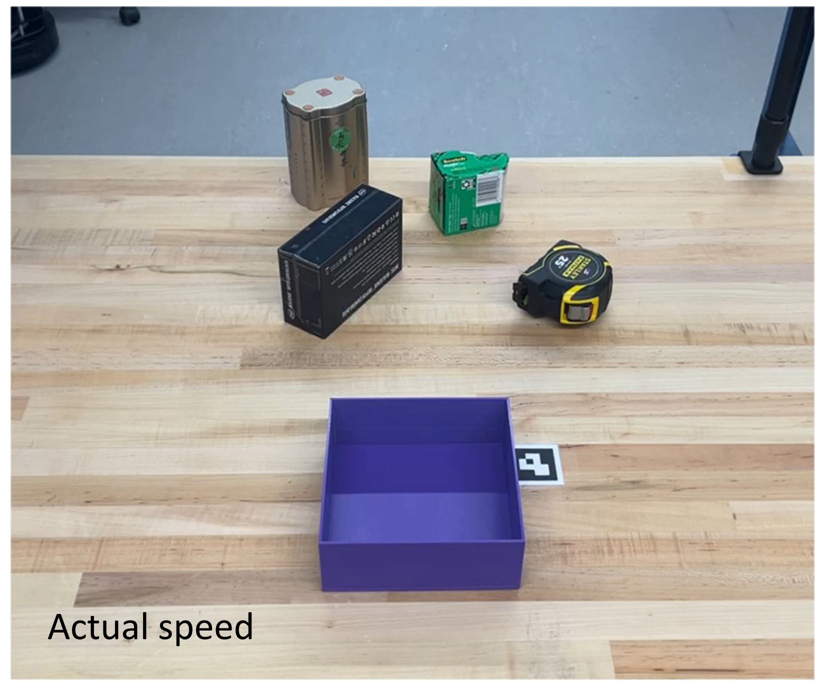


Pose optimization: maximize available space

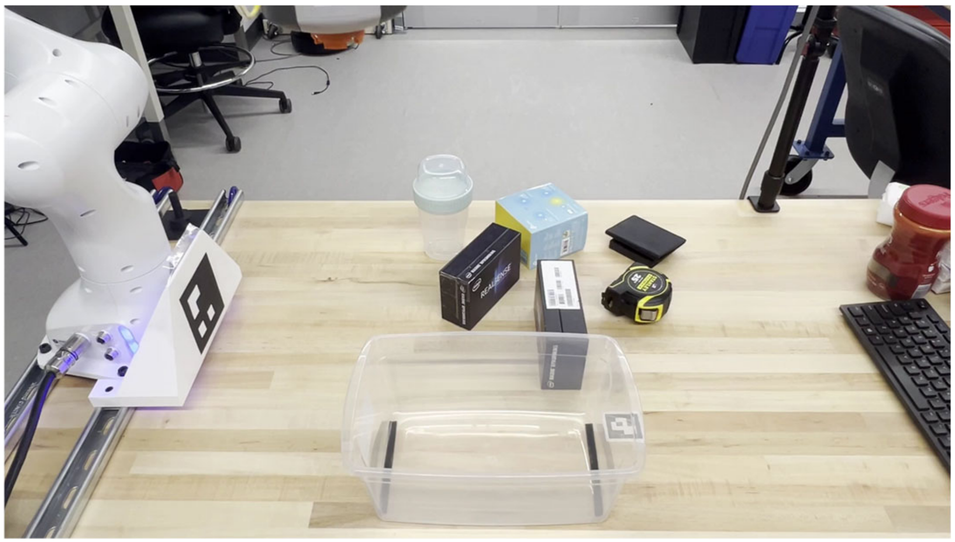


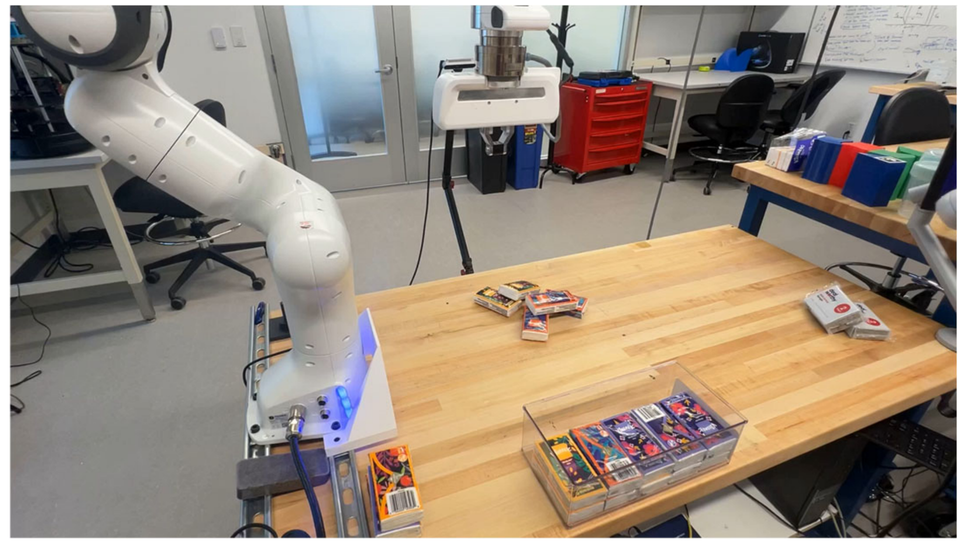



Actual speed

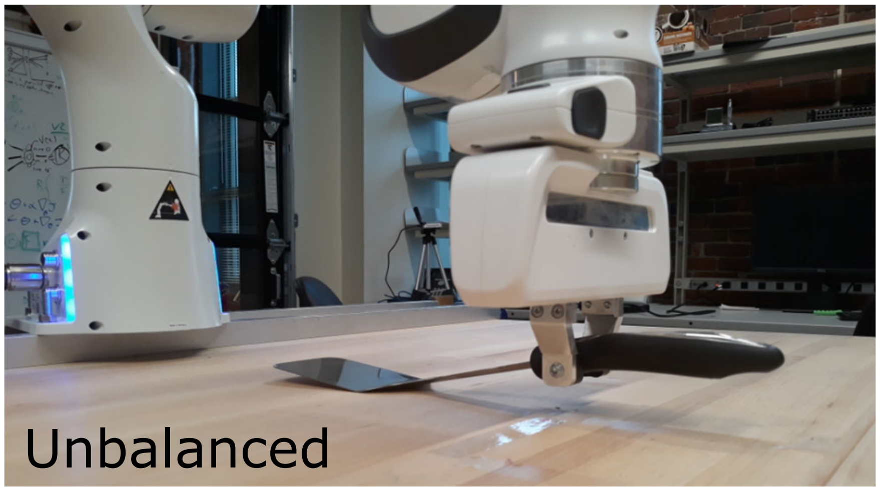

Actual speed

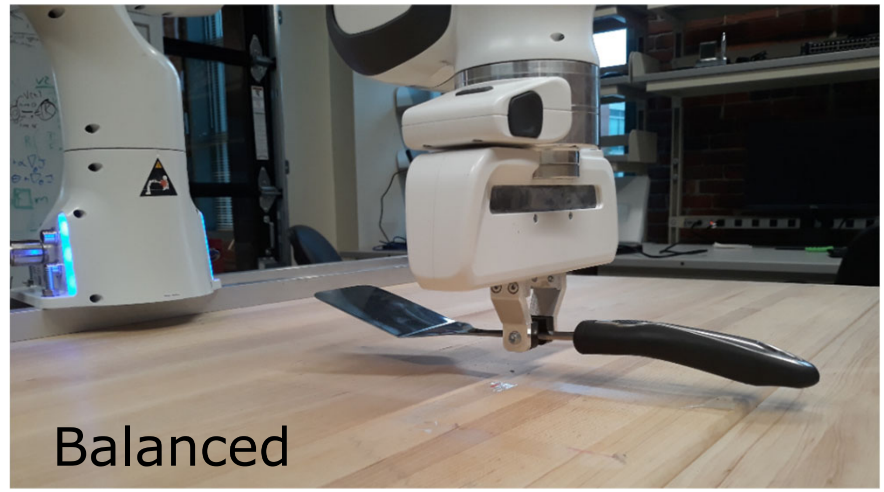
Multi-layer and Deformable Packing





Learning through Automatic Manipulation

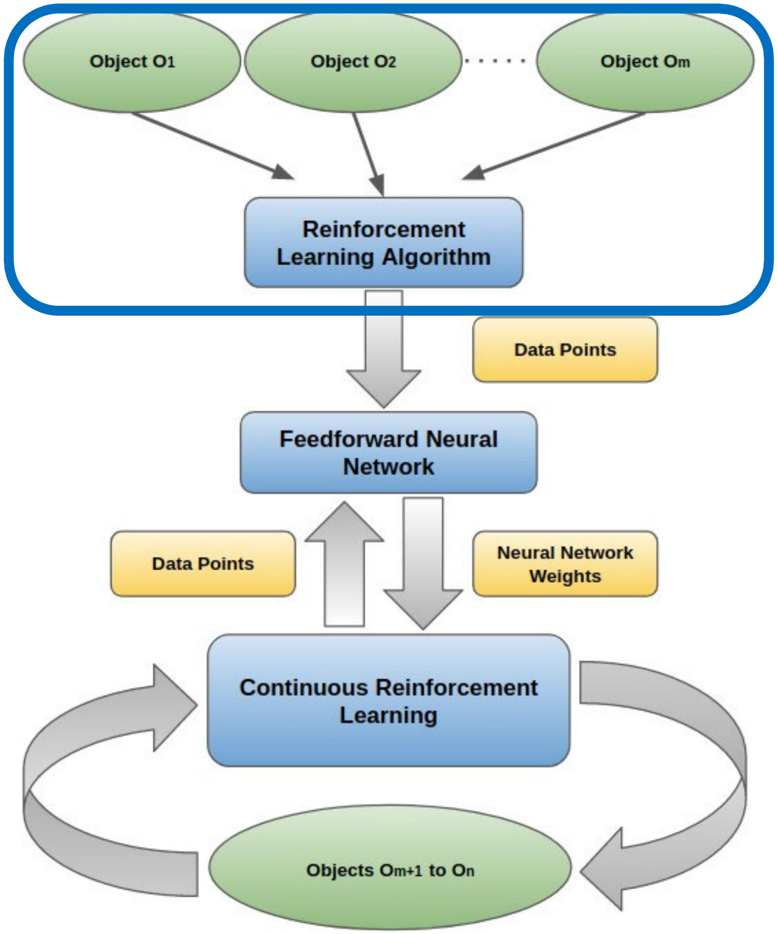




How?

Sean McGovern and Jing Xiao, "Learning and Predicting Center of Mass through Manipulation and Torque Sensing," *IEEE Inter Conf on Mechatronics and Robotics Engin, 2022.*

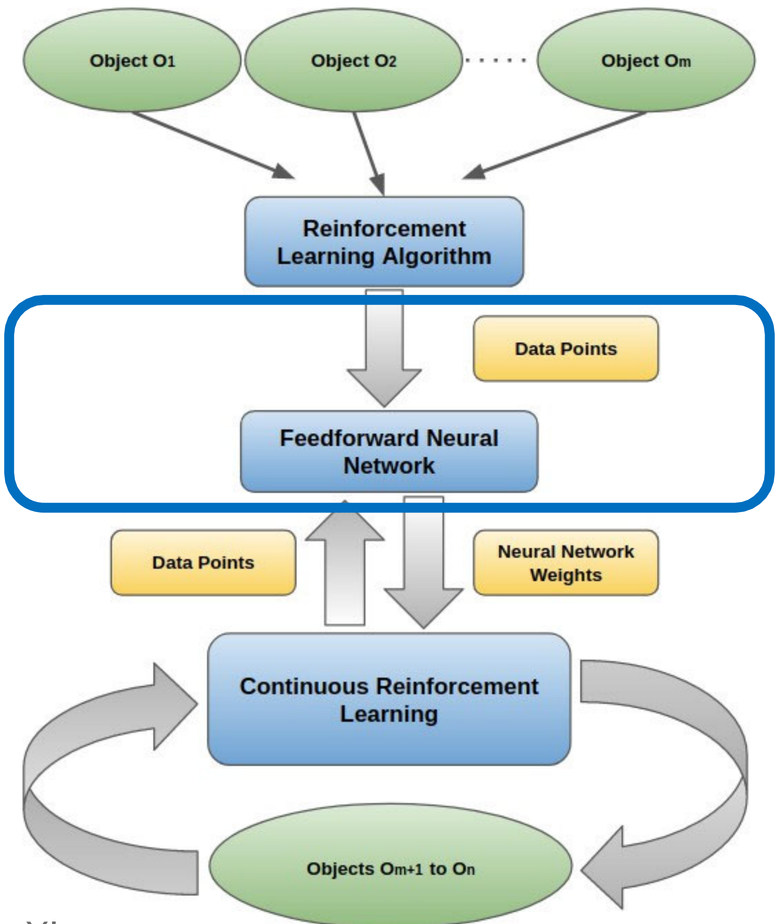
Approach



Reinforcement Learning Algorithm online:

- **State**: pick-up location on object and bounding box dimensions
- **Reward**: based on torque sensing
- Learns CoM along primary axis over several pick-up steps
- Robust to uncertainties in sensors/hardware

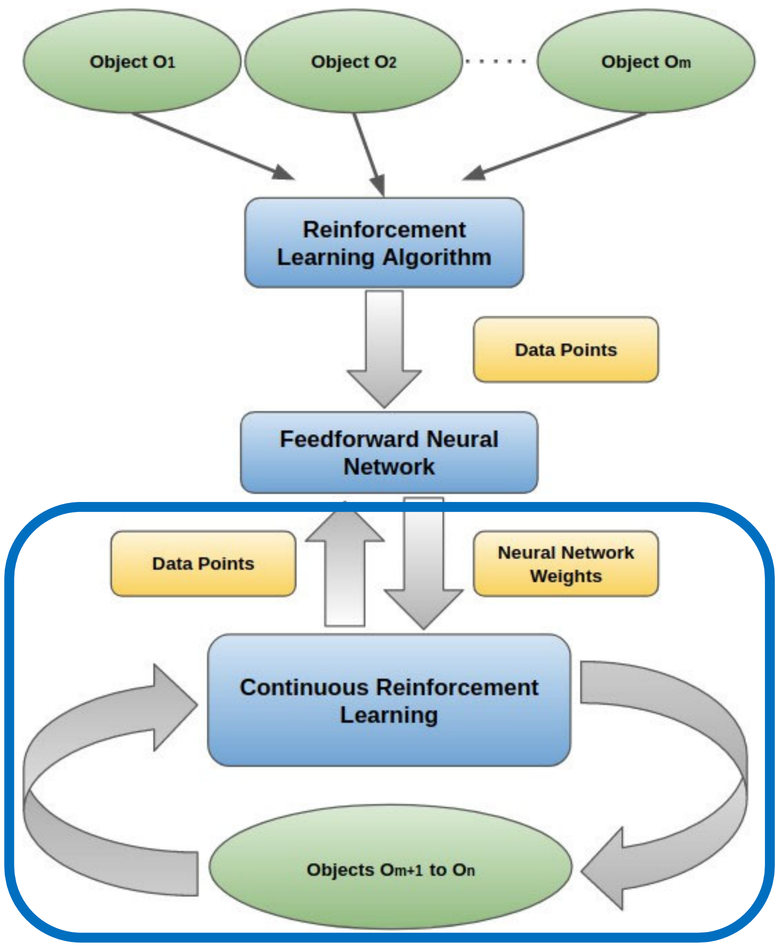
Approach



Feedforward Neural Network:

- Use <input, CoM> data points from RL for training.
- Trained FNN is used for CoM predictions of new objects.

Approach

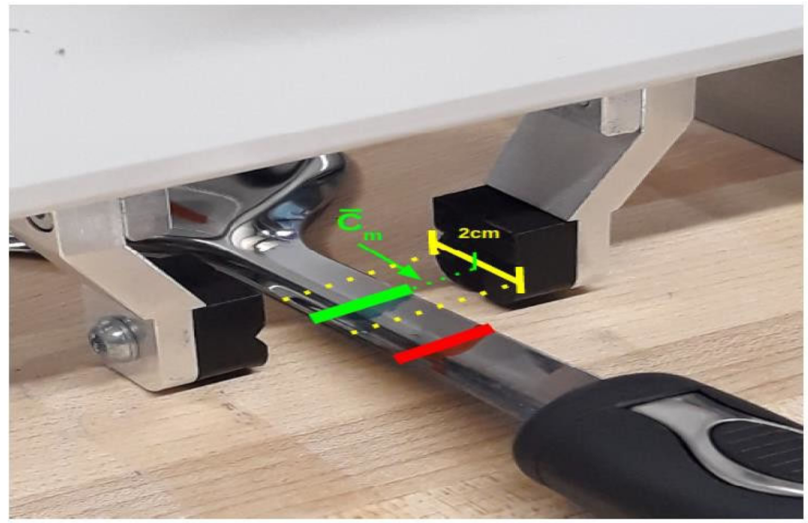


Continuous Reinforcement Learning (online):

- Use CoM prediction initially.
- Further RL to determine CoM.
- New data points from RL updates FNN.

Example

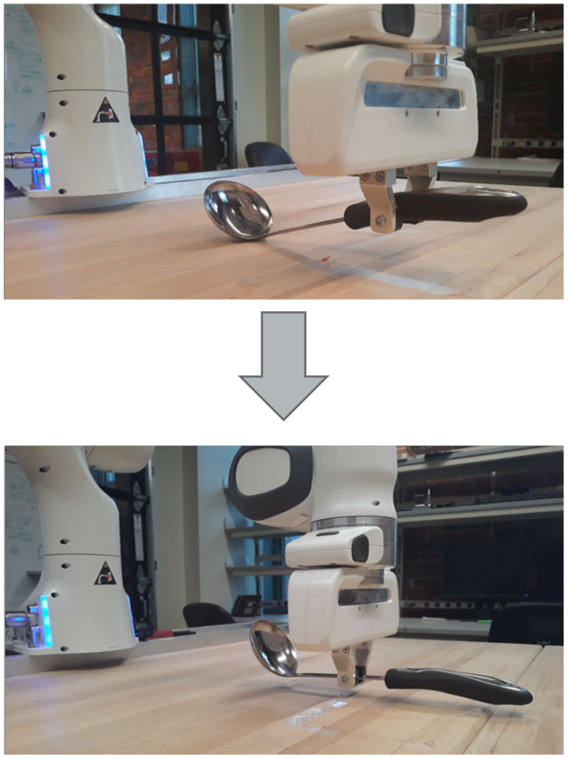
Red: center
Green: center of mass (CoM)



End-effector pick-up resolution

Reinforcement Learning Results (Stage 1)

Utensil #	Learned C_m (cm)	Range $\overline{C_m} \pm 1$(cm)	Learning time in # of Steps
#1	-1.8	$[-3.5, -1.5]$	51 Steps
#2	-0.4	$[-0.5, 1.5]$	24 Steps
#3	3.7	$[2.0, 4.0]$	41 Steps
#4	-1.7	$[-2.5, -0.5]$	30 Steps
#5	0.5	$[-1.5, 0.5]$	22 Steps
#6	0.0	$[-2.0, 0.0]$	20 Steps
#7	0.7	$[0.5, 2.5]$	24 Steps
#8	0.5	$[-1.0, 1.0]$	16 Steps

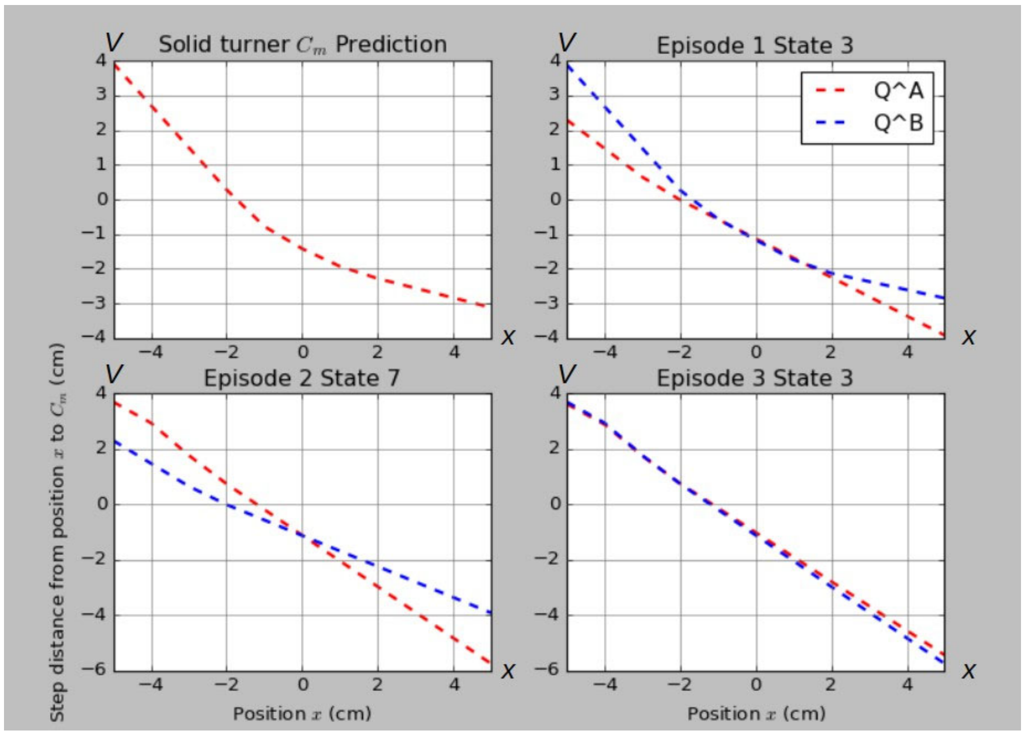


Ladle (#1)

FNN Prediction Results
(Stage 2, trained by data from #1-#3)

Utensil #	Prediction C_m (cm)	Range $\overline{C}_m \pm 1$(cm)	In range?
#4	-1.4	$[-2.5, -0.5]$	Y
#5	-2.7	$[-1.5, 0.5]$	N
#6	-1.4	$[-2.0, 0.0]$	Y
#7	2.3	$[0.5, 2.5]$	Y
#8	-2.1	$[-1.0, 1.0]$	N

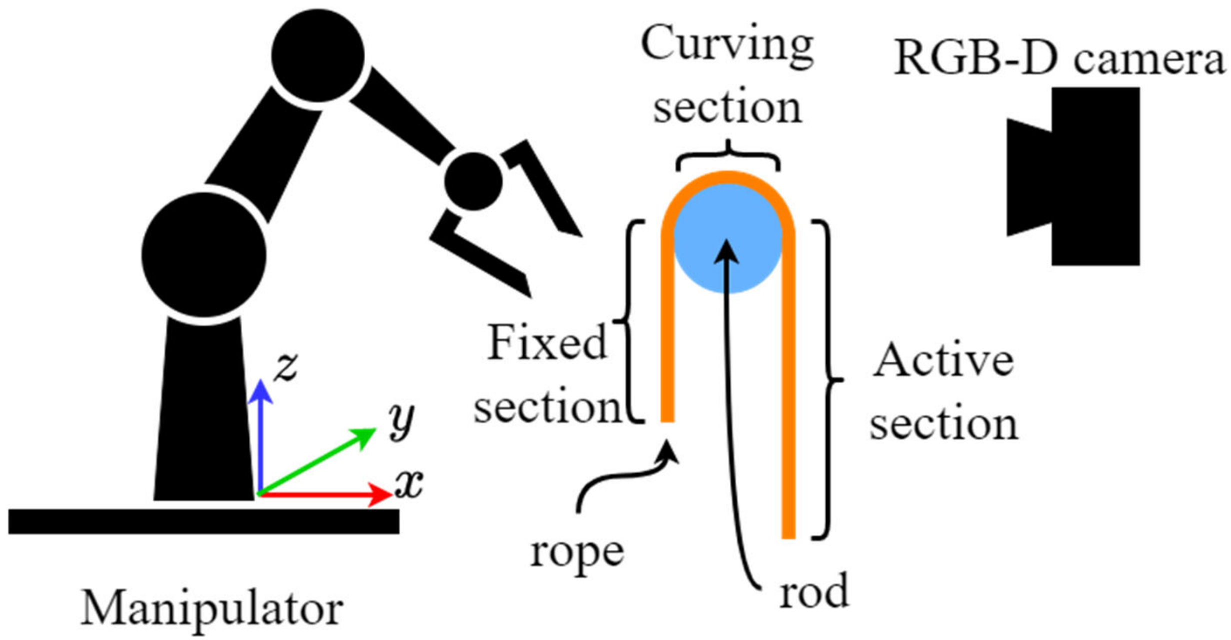
Continuous RL Results (Stage 3)



9 total steps over 3 episodes

WPI

Novel Rope Wrapping

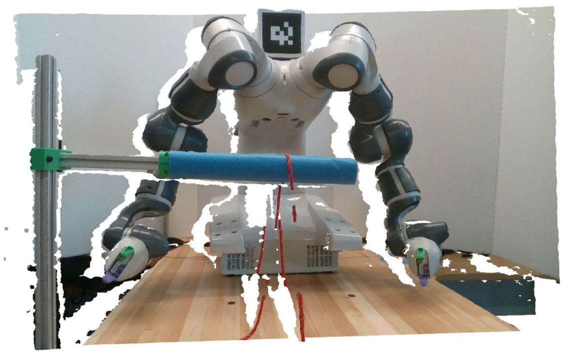


No prior information of rod and rope!

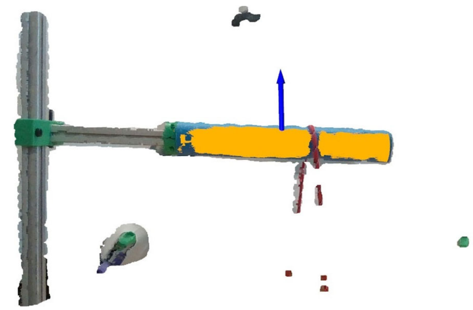
No simulation!

Zhaoyuan Ma and Jing Xiao, "Robotic Perception-motion Synergy for Novel Rope Wrapping Tasks," *IEEE Robotics and Automation Letters*, 2023.

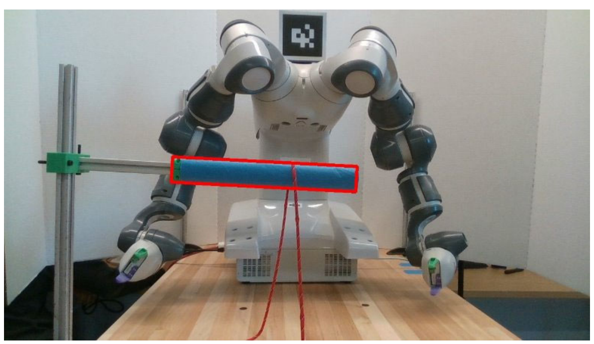
Rod and Rope Estimation



Point cloud

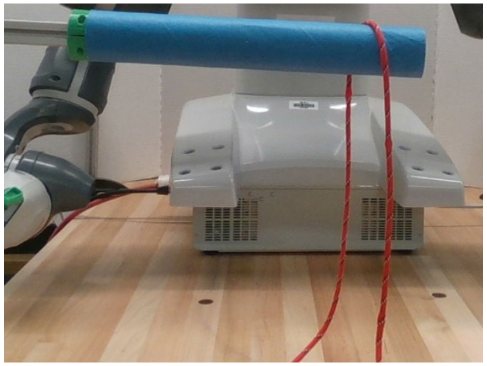


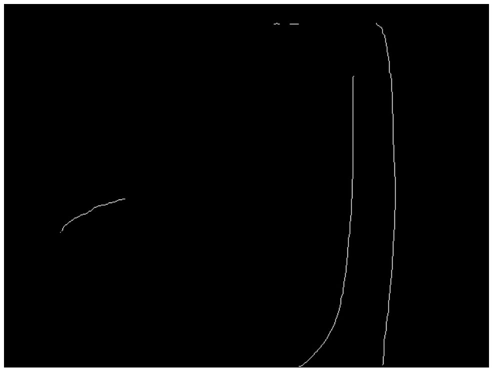
3D half cylinder matched (Rod)

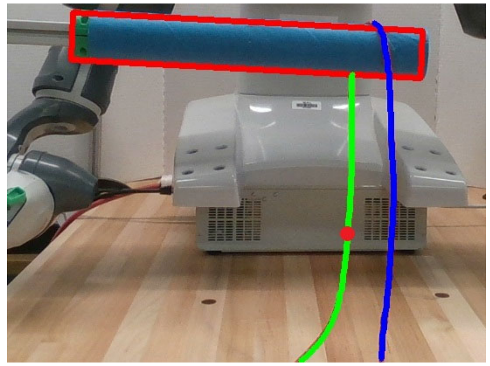


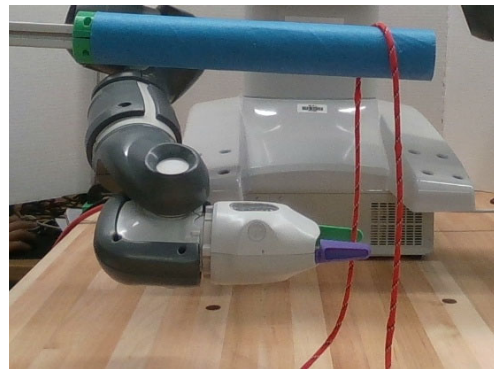
Contour found (Rope)

Single Wrap: Grasp Point Selection



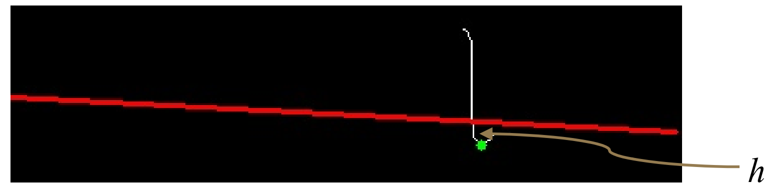


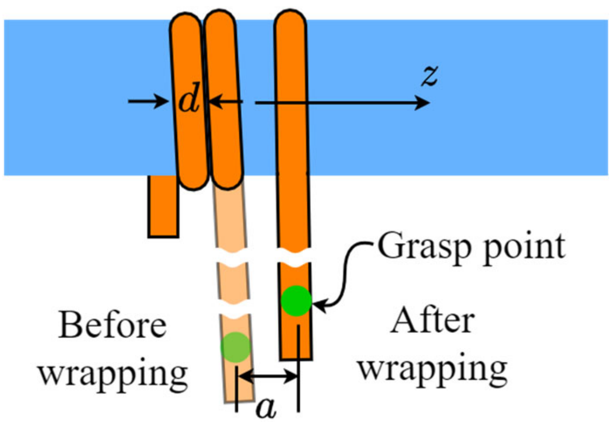




A. Caporali, R. Zanella, D. De Gregorio, and G. Palli, “Ariadne+: Deep Learning-based Augmented Framework for the Instance Segmentation of Wires,” IEEE Trans. Ind. Inform., pp. 1–1, 2022. (pre-trained DeepLabV3+)

Online Learning to Improve Wrapping Motion (perception & control)





Update height:

$$R_{n+1} = R_n - K_{PR} \cdot q_r, \text{ where } q_r = h - t_R$$

Update advance:

$$a_{n+1} = a_n - K_{Pa} \cdot q_a, \text{ where } q_a = S_g/(S_g + S_r)$$

Training Session (rod1, R=21mm)

Training session with Rod1:
adjust values of R, L', and a.

Testing (rod1)

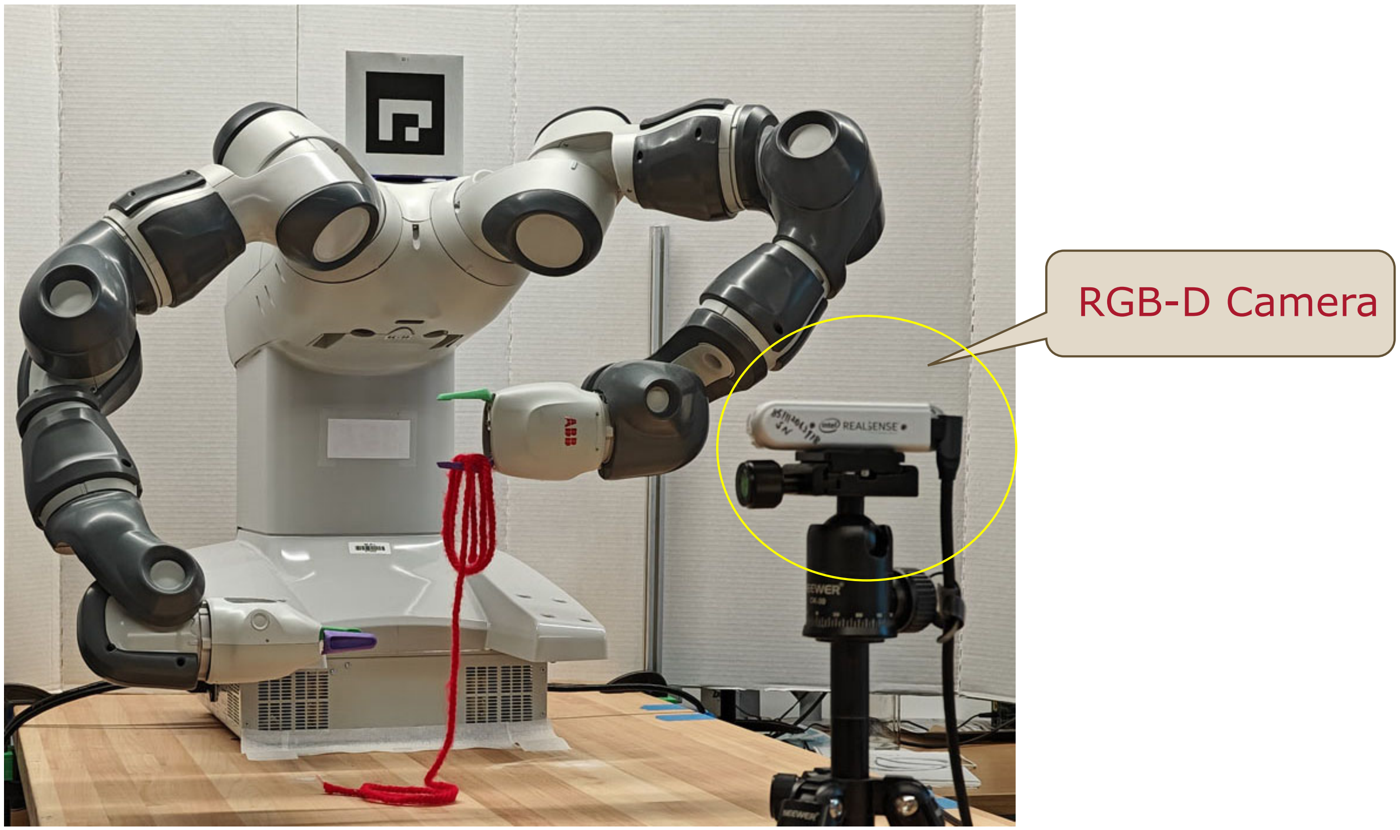
Testing session with Rod1:
apply adjusted R, L', and a.

WPI

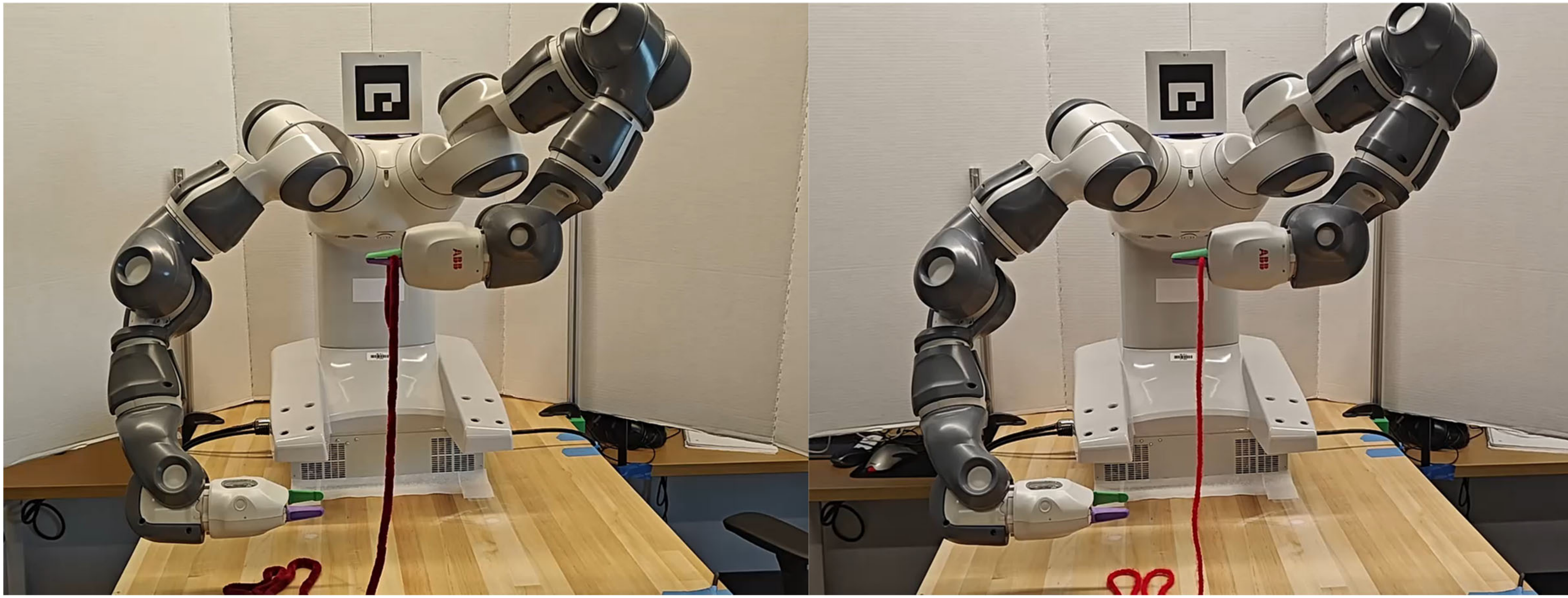
Testing (rod2, R=16.9mm)

Testing session with Rod2: apply adjusted R, L', and a.

General Deformable Linear Object Manipulation



6X

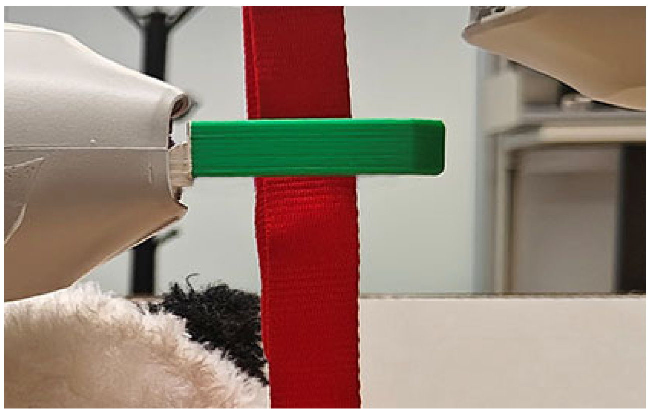


No simulation

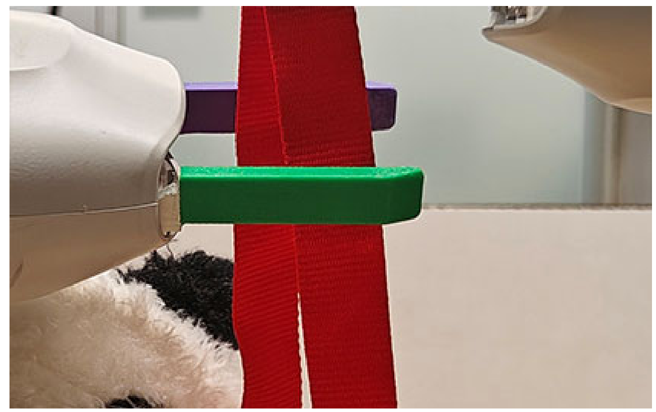
No object modeling

Perception ⇄ Spatial Relation ⇄ Action

Motion for Perception

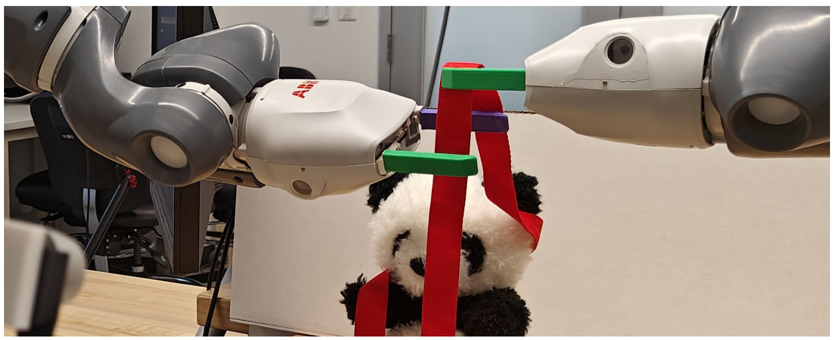


Rear ribbon & finger occluded

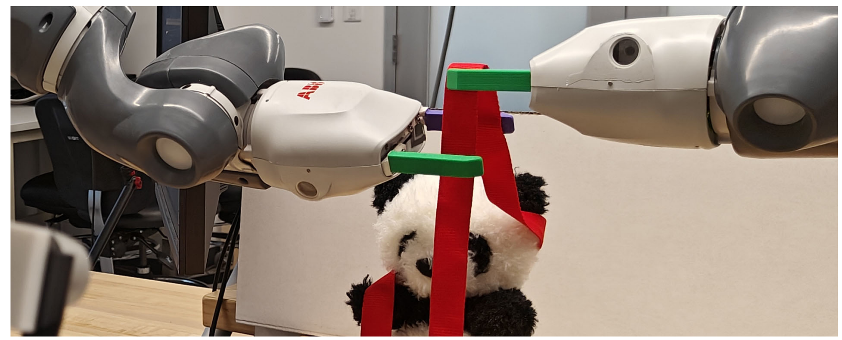


Rear ribbon & finger visible

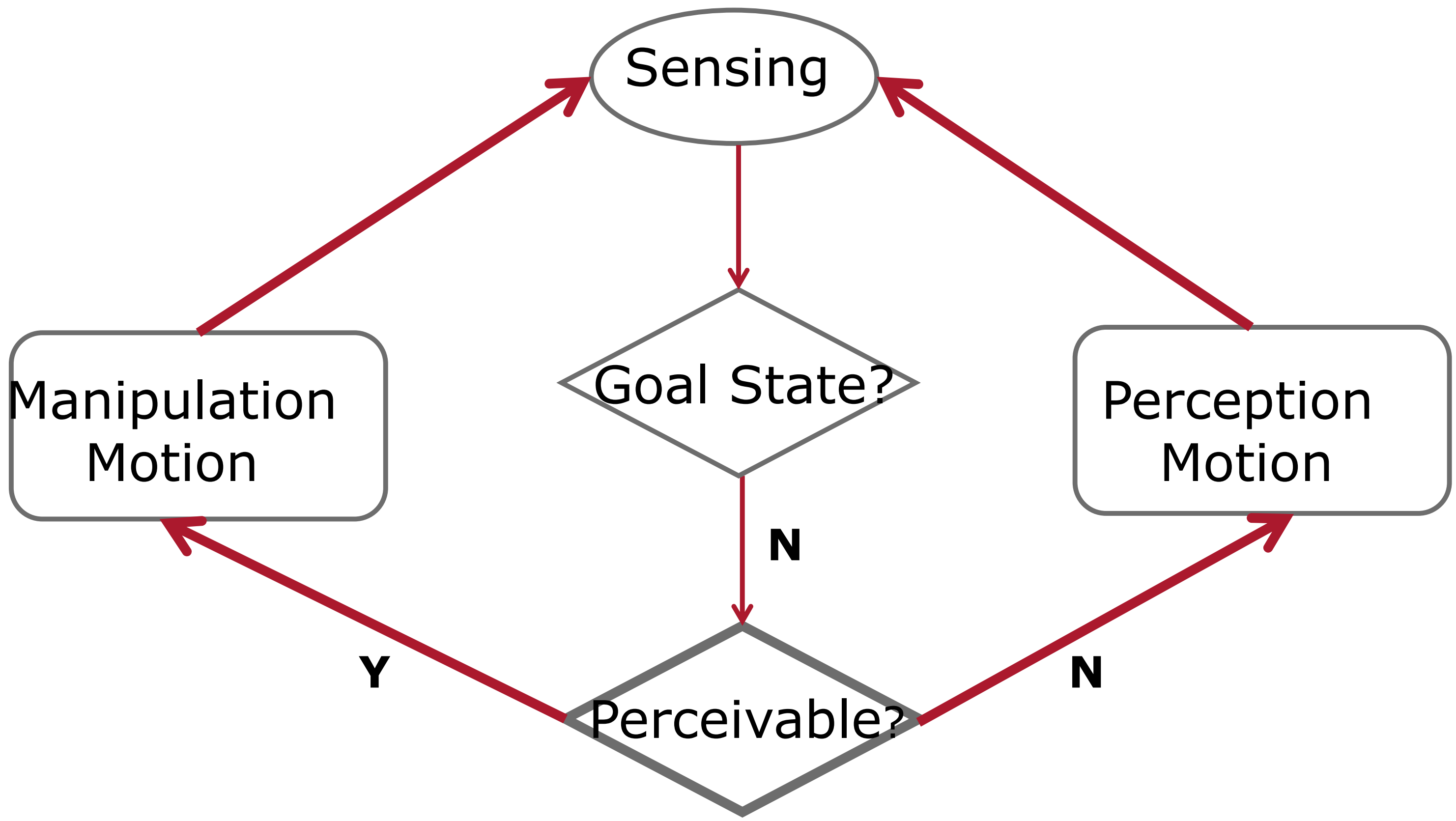
Motion for Manipulation



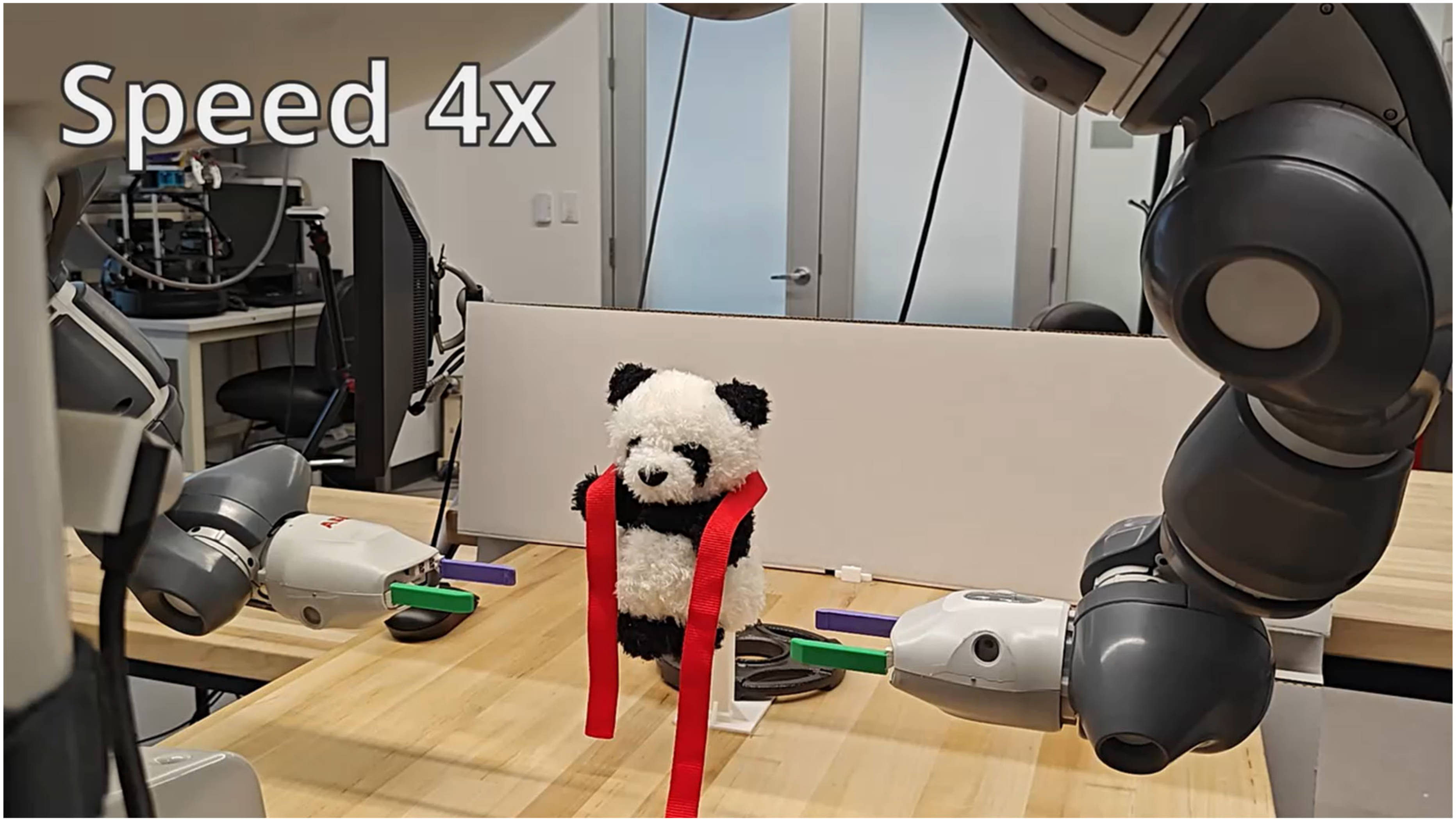
Spatial relation 1:
one ribbon between the left fingers



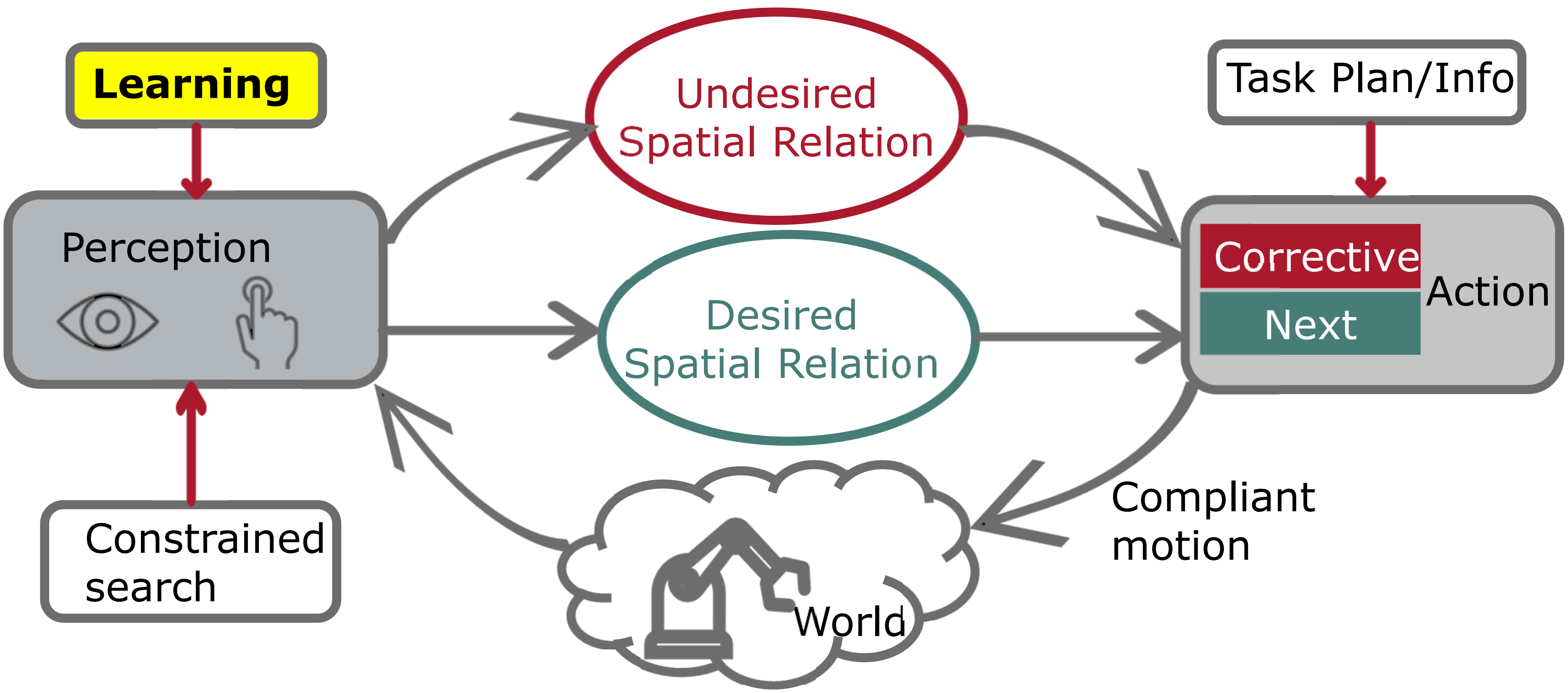
Spatial relation 2:
both ribbons between the left fingers



Sensing
Goal State?
Manipulation Motion
Perception Motion
N
Y
Perceivable?
N

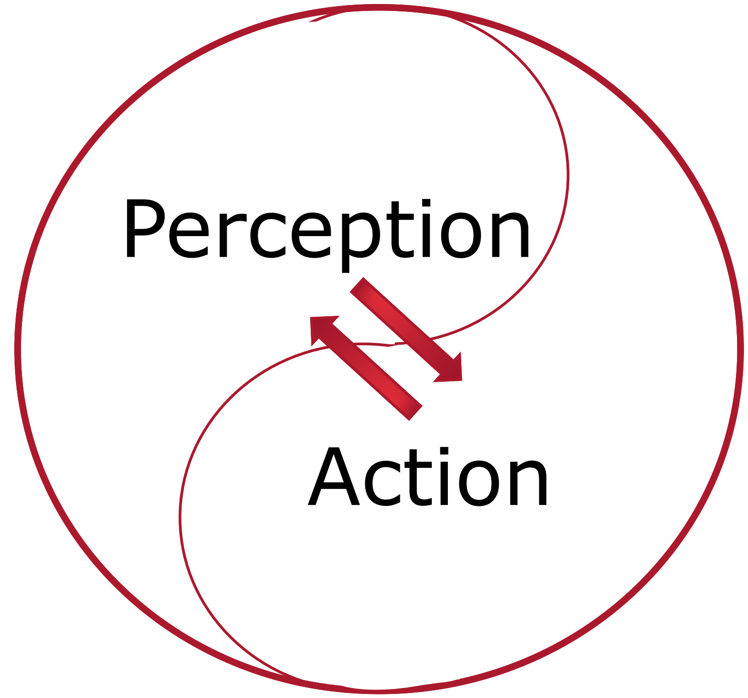

Speed 4x

Learning is Crucial for Handling Uncertainties



Learning is Crucial for Handling Uncertainties

- Learning for better perception under uncertainty
  - ✓ How to be efficient and general?
- Learning unknowns through automatic manipulation
  - ✓ RL or feedback control
- Perception-action feedback loop
  - ✓ Closed-loop to minimize undesired effects



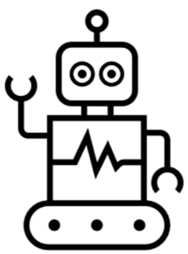
High Touch and
High Precision

Acknowledgement

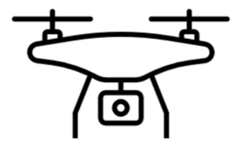
NSF I/UCRC on Robots and Sensors for Human Well-being (ROSE-HUB)

Amazon Robotics

Search and Exploration Tasks

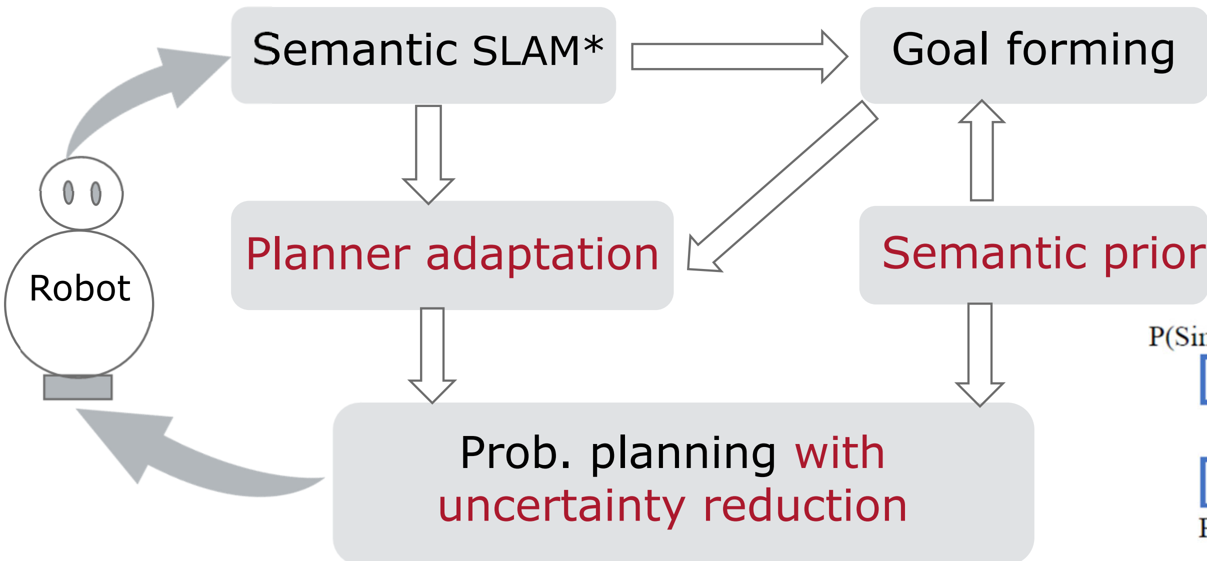


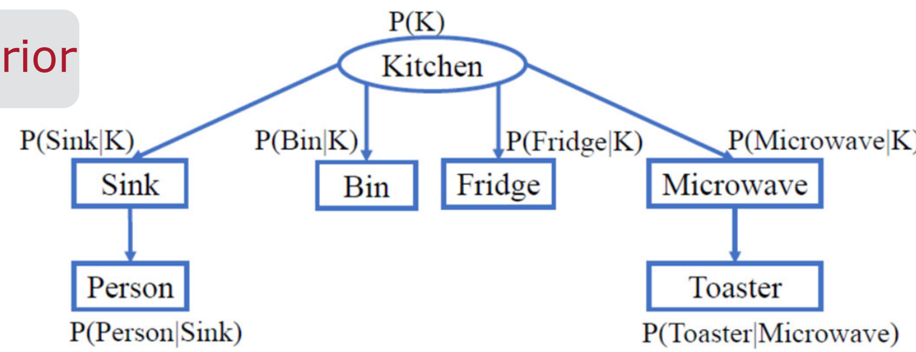
Search of
semantic objects



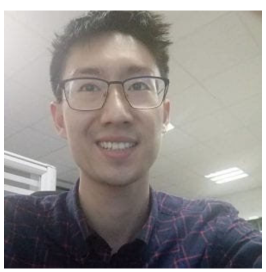
Visual coverage
of structure

Autonomous Search of Semantic Objects





*[Z. Qian, K. Patath, J. Fu, and J. Xiao, ICRA 2021][Z. Qian. J. Fu, and J. Xiao. RAL 2022]



Zhentian Qian, Jie Fu, and Jing Xiao, "Simultaneously Search and Localize Semantic Objects in Unknown Environments," *IEEE Robotics and Automation Letters*, October 2024.

WPI

Visual Coverage for Unknown Structure

Large complex unknown environment

Real-time

- perception
- mapping
- planning
- Motion

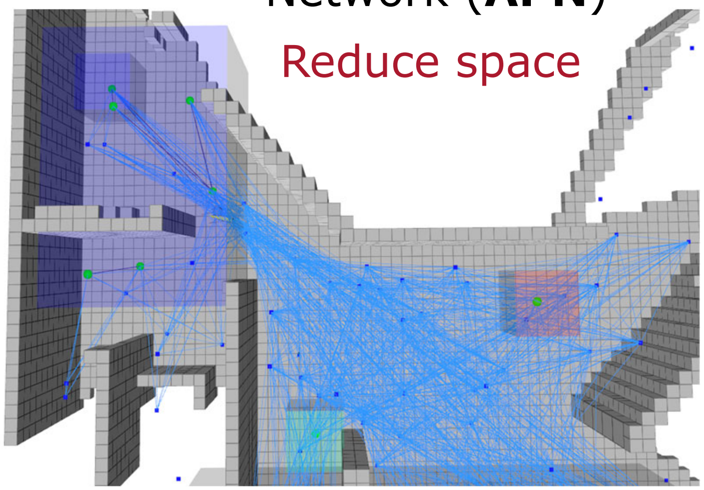
with limited resources

Autonomous robot

Approach

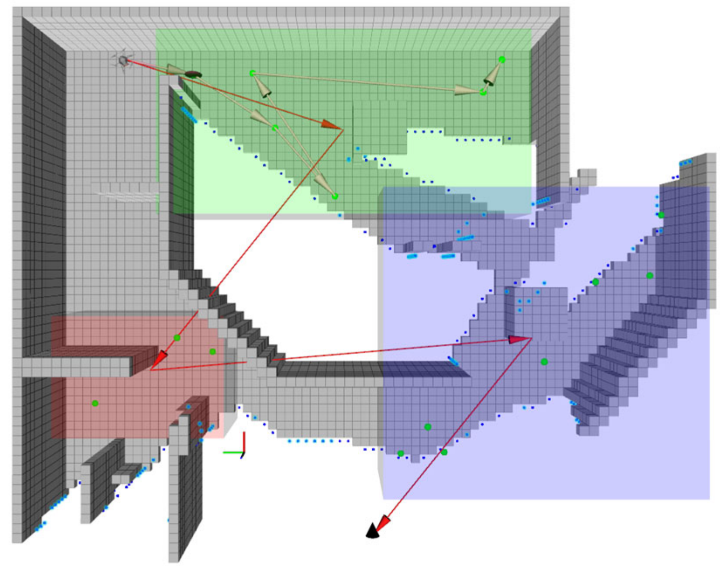
Representation
+ Adaptation

Active Perception Network (**APN**)
Reduce space



Differential Regulation (**DFR**) Reduce time

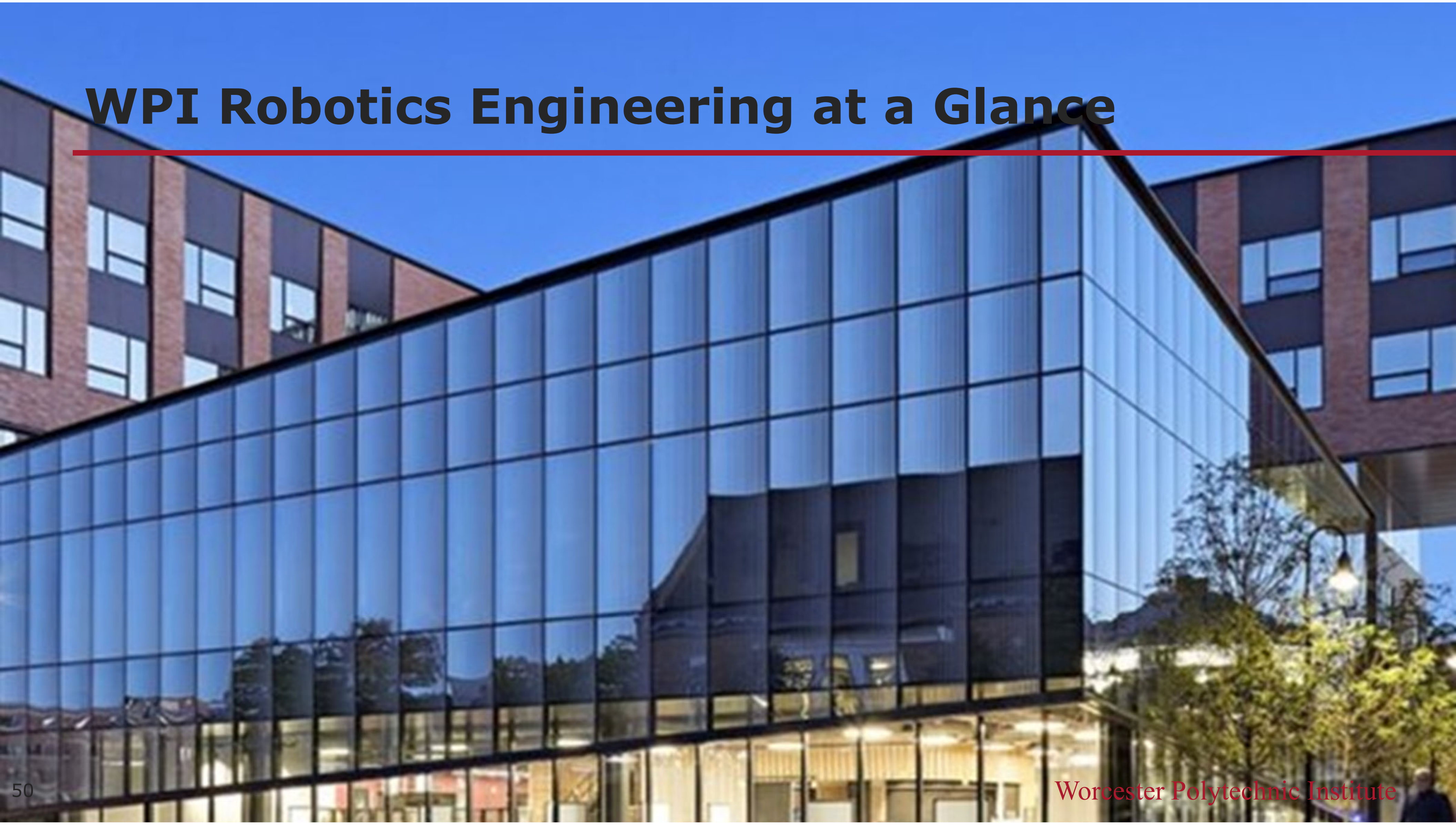
Non-myopic APN Planner (**APN-P**)
Enable real-time visual coverage tour planning

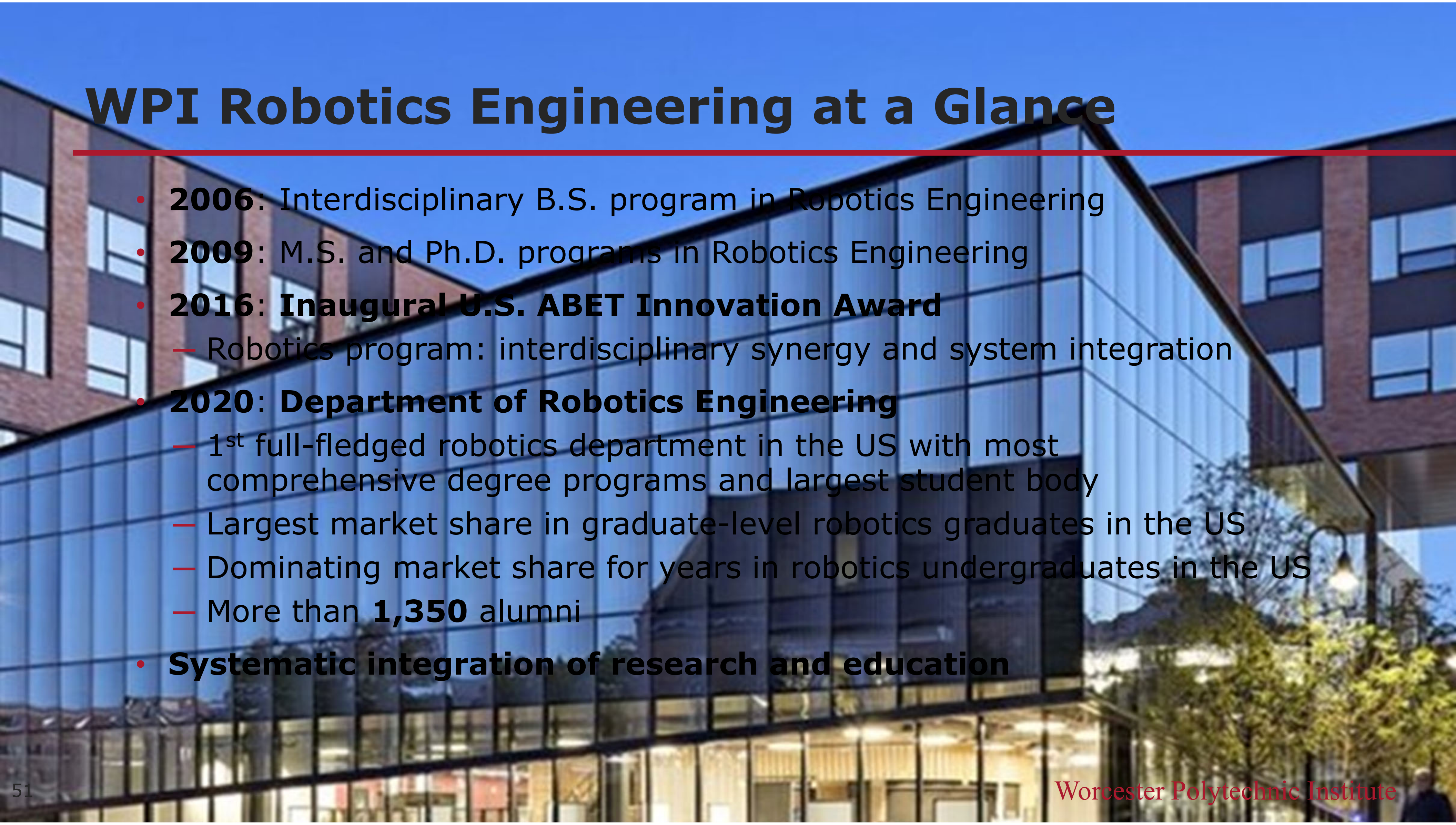


David Vutetakis and Jing Xiao, "Active Perception Network for Non-Myopic Online Exploration and Visual Surface Coverage," *International Journal of Robotics Research*, June 2024

WPI

WPI Robotics Engineering at a Glance





WPI Robotics Engineering at a Glance

- **2006**: Interdisciplinary B.S. program in Robotics Engineering
- **2009**: M.S. and Ph.D. programs in Robotics Engineering
- **2016**: **Inaugural U.S. ABET Innovation Award**
  - Robotics program: interdisciplinary synergy and system integration
- **2020**: **Department of Robotics Engineering**
  - 1st full-fledged robotics department in the US with most comprehensive degree programs and largest student body
  - Largest market share in graduate-level robotics graduates in the US
  - Dominating market share for years in robotics undergraduates in the US
  - More than **1,350** alumni
- **Systematic integration of research and education**

Research Groups

https://www.wpi.edu/academics/departments/robotics-engineering/research/groups

CONTACT

Location: Unity Hall

Phone: 508-831-6665

rbe@wpi.edu

Adaptive and Intelligent Robotics (AIR) Lab

Prof. Jing Xiao

UH 200E

Real-time perception-based robotic assembly; Manipulation; Navigation in uncertain or unknown environments

Aerial-robot Control and Perception (ACP) Lab

Prof. Guanrui Li

UH 100A

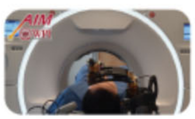
Aerial Robots; Multi-Robots System; Aerial Manipulation; Optimal Control

Automata Lab

Prof. Kevin Leahy

UH 100C

Formal methods; Artificial Intelligence; Autonomous Systems; Planning and Control



Automation and Interventional Medicine (AIM) Lab

Prof. Gregory Fischer

50P 4th Floor

Medical robotics; Image-guided surgery; MRI-compatible mechatronics; Intelligent robotic surgery; Assistive technology

Cognitive Medical Technology and Robotics (COMET) Lab

Prof. Loris Fichera

50P 4th Floor

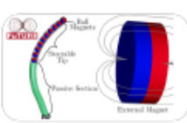
Surgical robotics; Continuum surgical robots; Energy-based surgical instruments; Lifesaving technologies

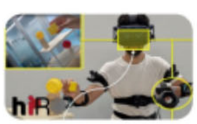
Efficient Learning and Planning for Intelligent Systems (ELPIS) Lab

Prof. Constantinos Chamzas

UH 200B

Learning for Efficient Planning; Planning under Uncertainty; Visual-Based Planning; Learning Abstractions for Planning





Prof. Giovanni Pittiglio

50P 4th Floor

Medical Robots; Medical Devices; Minimally Invasive Medicine

Multimodal multilateral human-robot interaction and interfaces; Shared autonomous nursing robots; Workload-adaptive human-robot collaboration

Prof. Berk Calli

UH 200D

Robotics Manipulation; Environmental Robotics; Vision-based Control; Robot Dexterity

Medical Frontier Ultrasound Imaging and Robotic Instrumentation (FUSION) Lab

Prof. Haichong Zhang

50P 4th Floor

Robotics for healthcare; Ultrasound and photoacoustic imaging; Image-guided intervention; Human-robot interface

Novel Engineering of Swarm Technologies (NEST) Lab

Prof. Carlo Pinciroli

UH 200C

Swarm robotics; Multi-robot systems; Artificial intelligence; Software engineering

Perception and Autonomous (PeAR) Group

Prof. Nitin J. Sanket

UH 100B

Artificial Intelligence; Aerial Robotics; Computer Vision; Deep Learning; Learning For Autonomy

PracticePoint

Prof. Gregory Fischer

50P 4th Floor

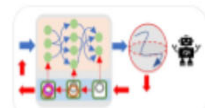
Surigcal robotics; Image guided therapy; Nursing Robotics; Rehabilitation robotics;

Principled Rigid-Soft Mechanisms (PRISM) Lab

Prof. Connor McCann

UH 100F Floor

Mechanism design; Soft robotics; Wearable robotics; Robotic manipulation; Bioinspired robotics



Safe Autonomy and Intelligence Lab (SAIL)

Prof. Wei Xiao

UH 200F

Safety-Critical Control Theory; Machine Learning; Multi-Agent Systems; AI-Enabled Robotics

Socially Intelligent Robotics for Healthcare (RoboCare) Lab

Prof. Fengpei (Fiona) Yuan

UH 100D

Soft Robotics Lab

Prof. Cagdas Onal

UH 100E

Soft continuum robots; Origami robots;


A 412
ICRA™
IEEE INTERNATIONAL CONFERENCE
ON ROBOTICS AND AUTOMATION
ATLANTA2025
lastcall
BY NEIMAN MARCUS

Thank you!

[Robotics Engineering | Worcester Polytechnic Institute (rbe.wpi.edu)](rbe.wpi.edu)