Scalable Robot Autonomy through Remote Operator Assistance and Lifelong Learning

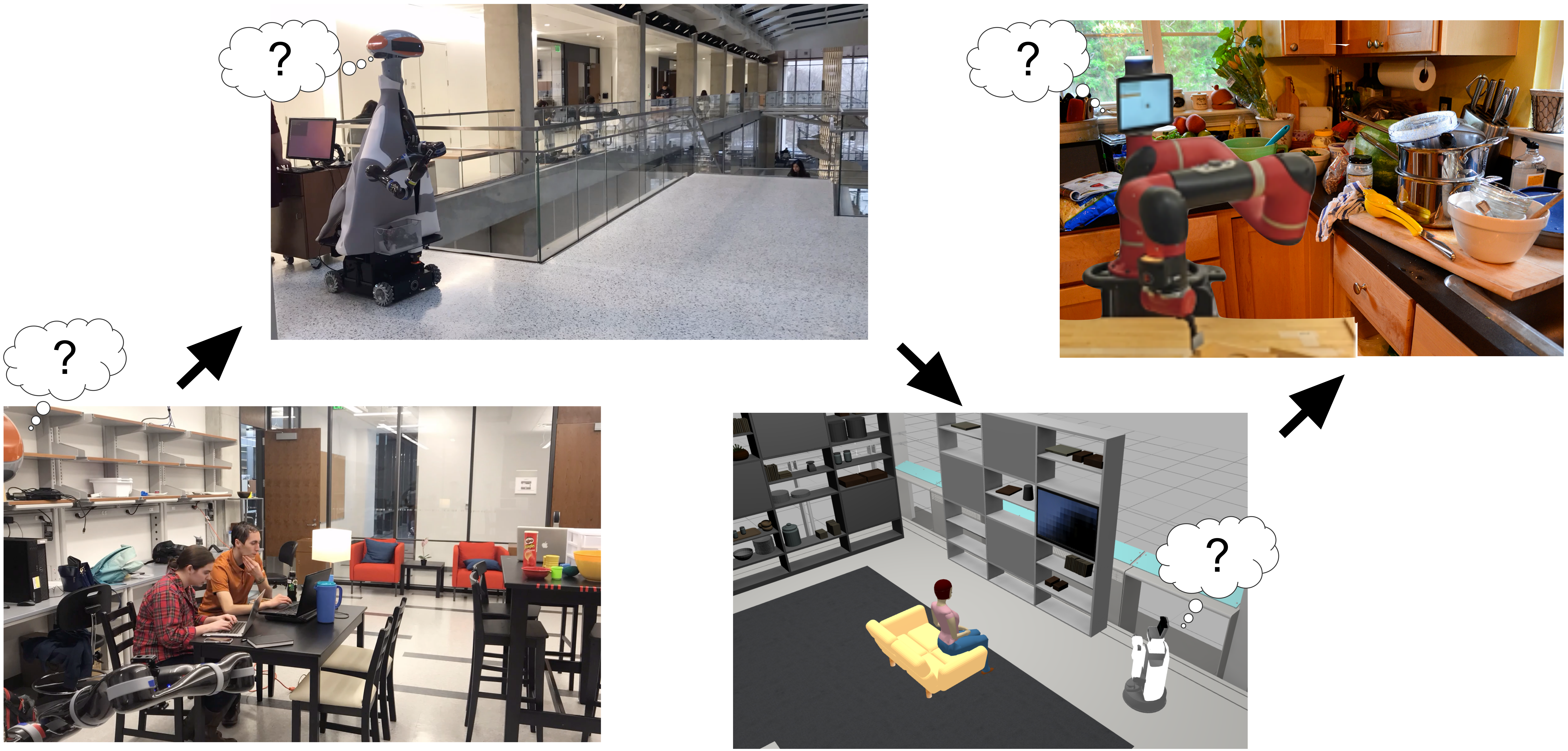
Scott Niekum
Andrea Thomaz
Elaine Schaertl Short

University of Texas at Austin

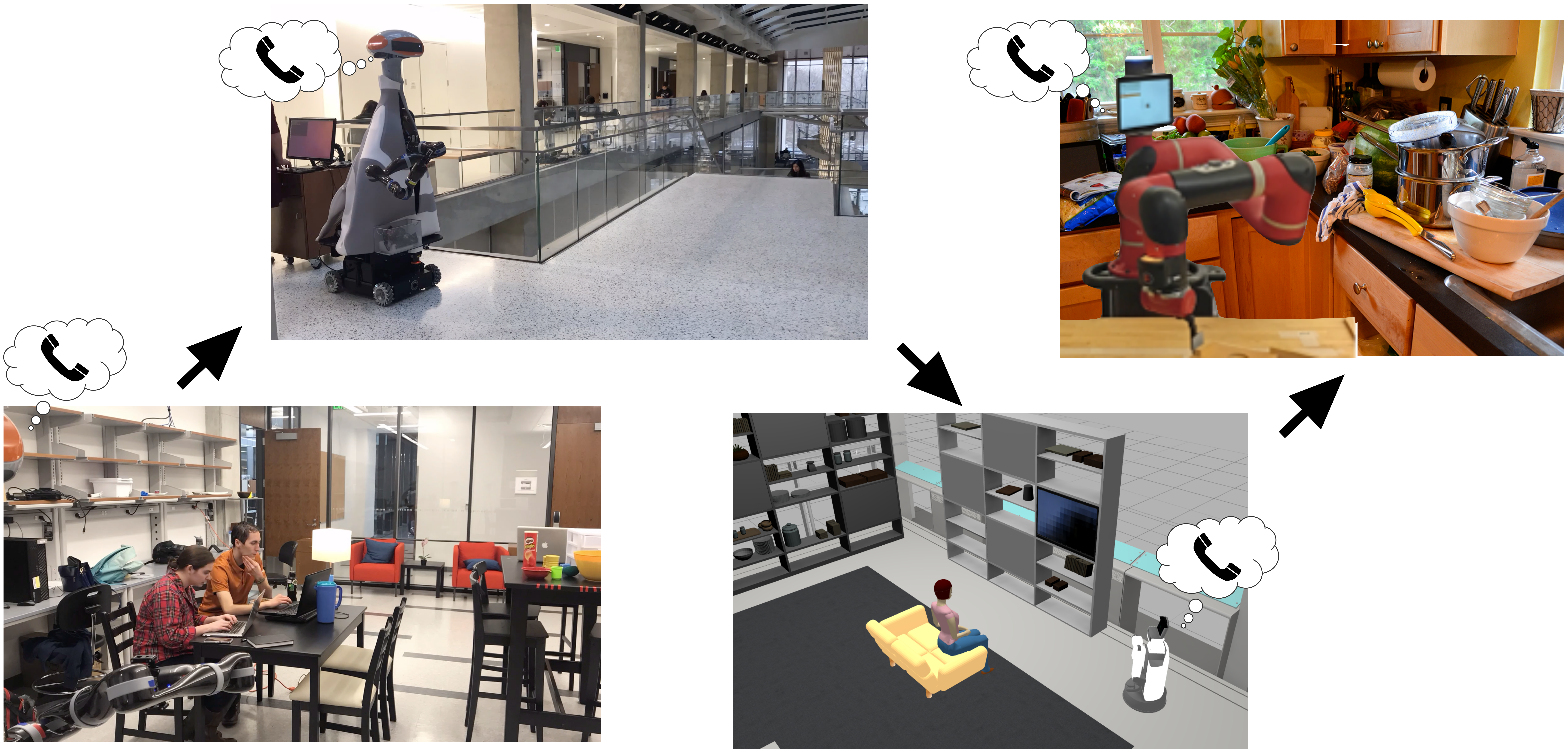
Sonia Chernova
Harish Ravichandar

Georgia Institute of Technology

Deploy robots in the real world **now**



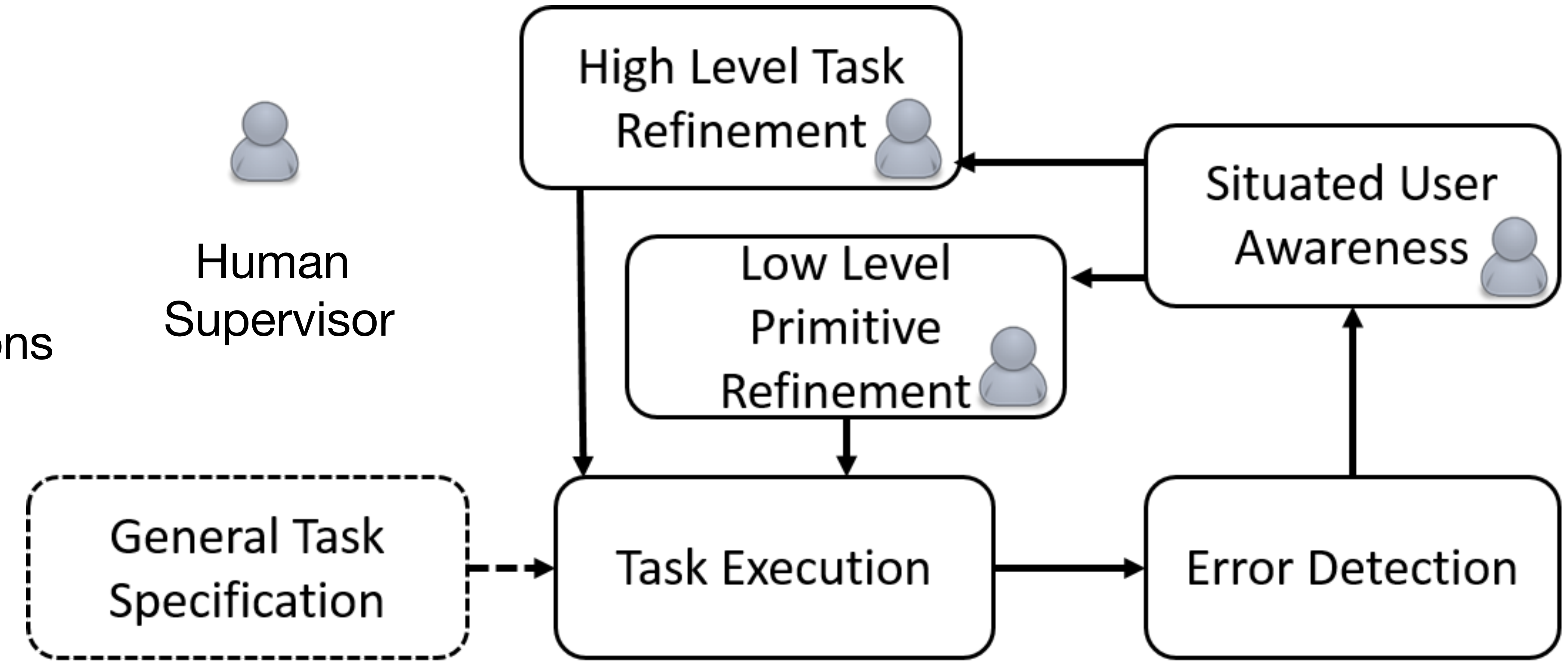
Robots will inevitably **run into problems**



Get **feedback** from a remote **supervisor** to **learn better models**

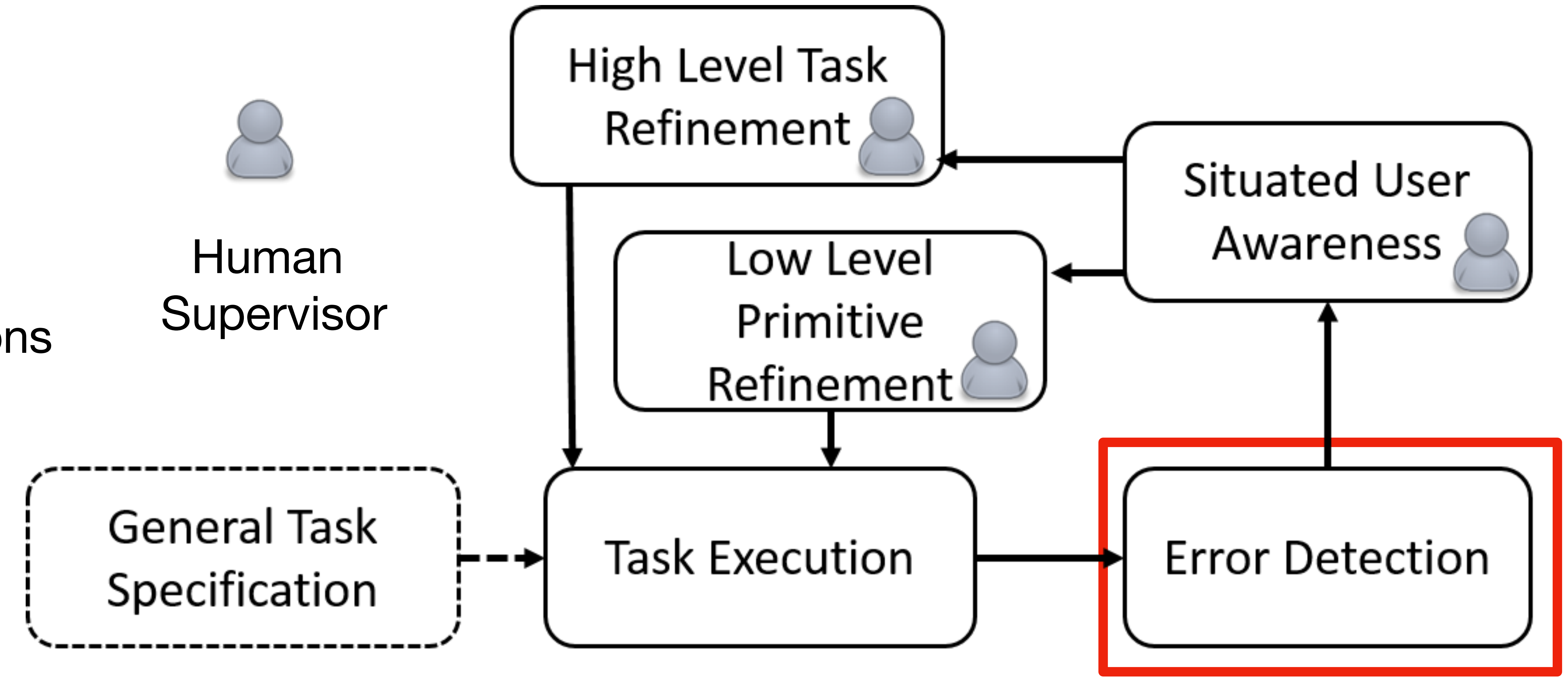
Overview

- UT Austin
  - Viewpoint Selection for Visual Failure Detection
  - Human Gaze Following for Human-Robot Interaction
  - Active Reward Learning from Critiques
  - Incremental Task Modification via Corrective Demonstrations
- Georgia Tech
  - Learning Generalizable Skills from Demonstrations
  - Skill Learning in Cluttered Environments
  - Automated Multi-Coordinate Cost Balancing
  - End User Evaluation of LfD Methods



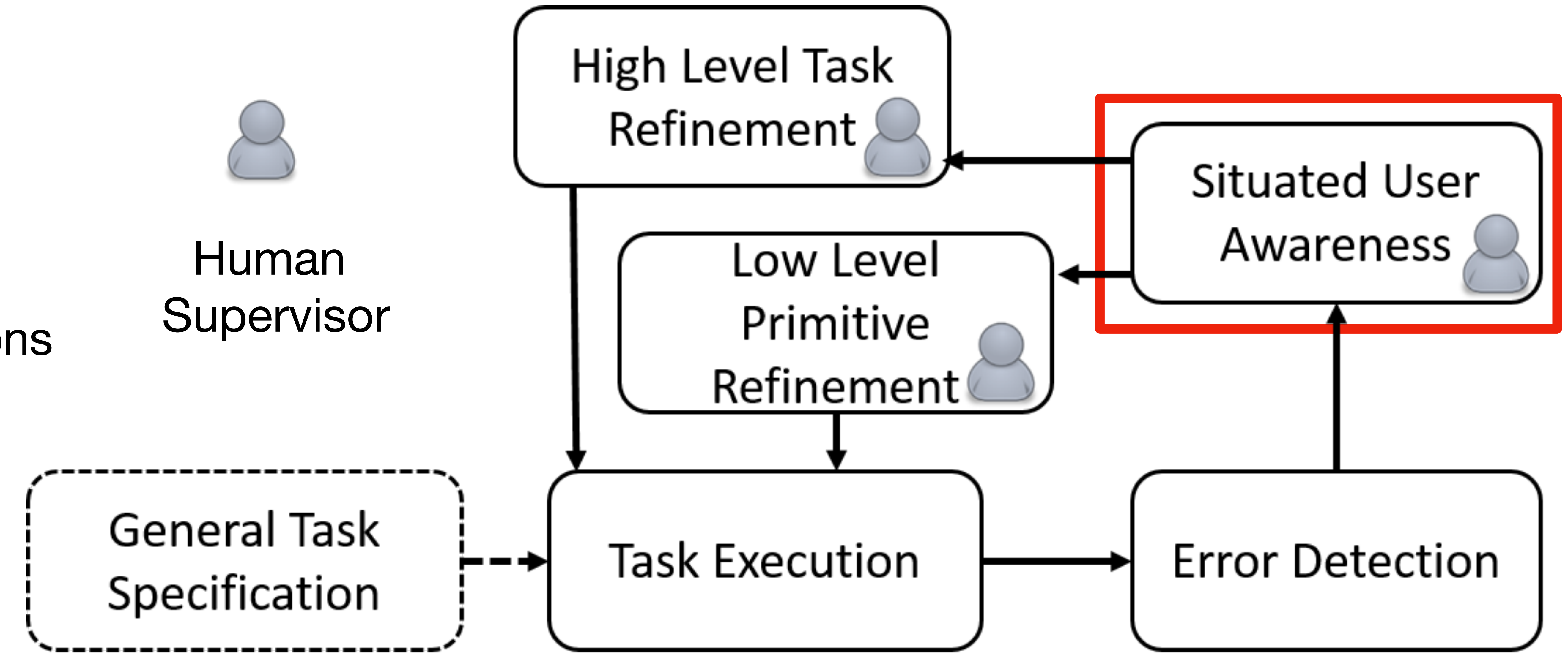
Overview

- UT Austin
  - **Viewpoint Selection for Visual Failure Detection**
  - Human Gaze Following for Human-Robot Interaction
  - Active Reward Learning from Critiques
  - Incremental Task Modification via Corrective Demonstrations
- Georgia Tech
  - Learning Generalizable Skills from Demonstrations
  - Skill Learning in Cluttered Environments
  - Automated Multi-Coordinate Cost Balancing
  - End User Evaluation of LfD Methods



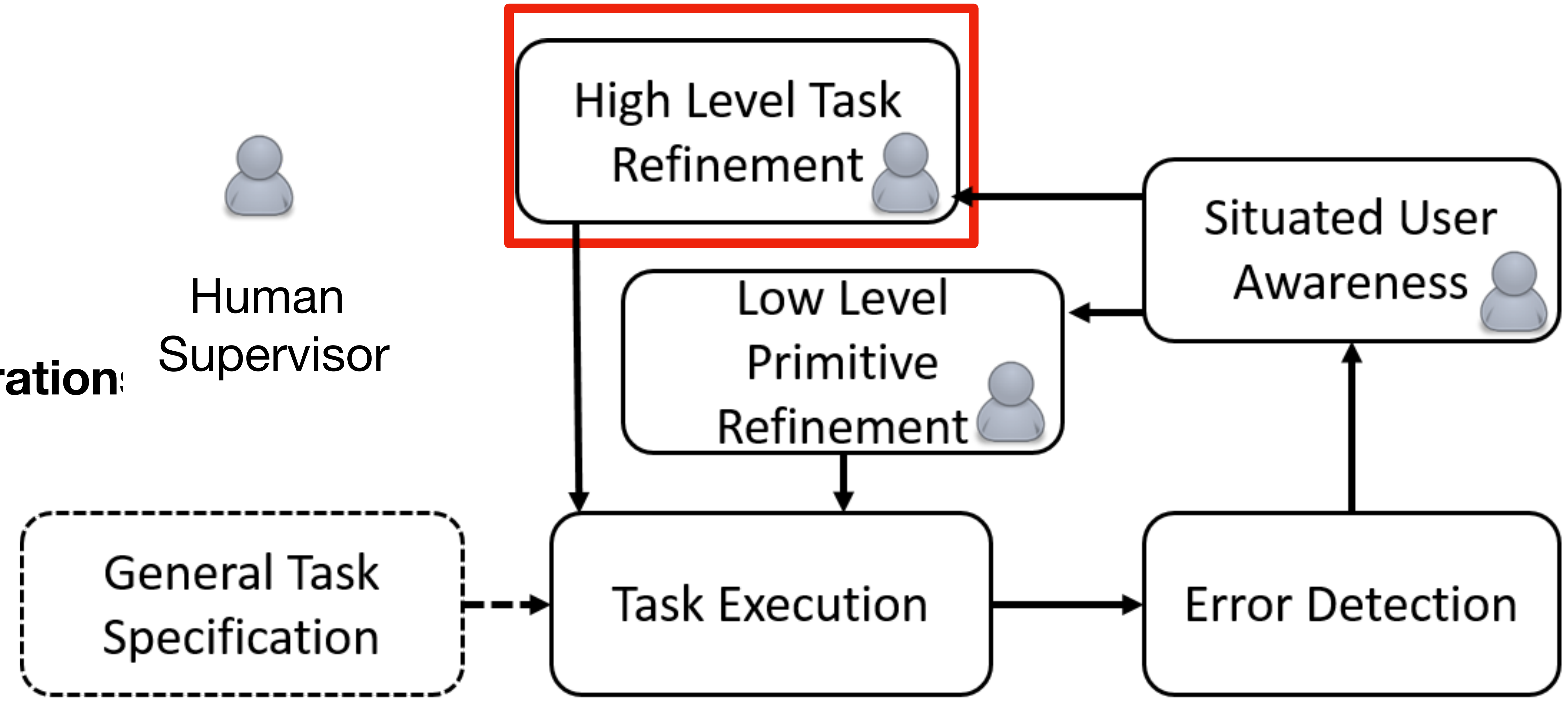
Overview

- UT Austin
  - Viewpoint Selection for Visual Failure Detection
  - **Human Gaze Following for Human-Robot Interaction**
  - Active Reward Learning from Critiques
  - Incremental Task Modification via Corrective Demonstrations
- Georgia Tech
  - Learning Generalizable Skills from Demonstrations
  - Skill Learning in Cluttered Environments
  - Automated Multi-Coordinate Cost Balancing
  - End User Evaluation of LfD Methods

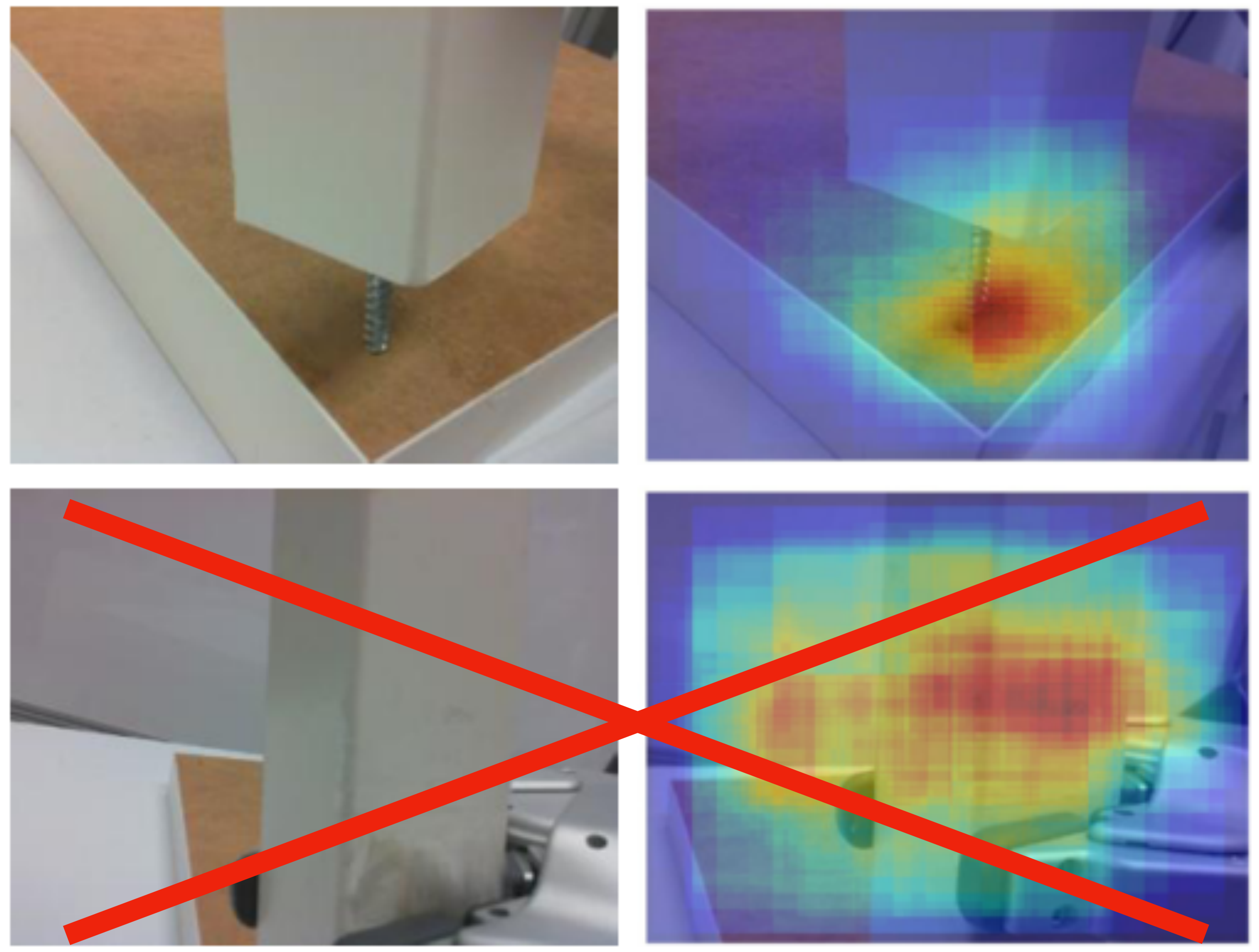


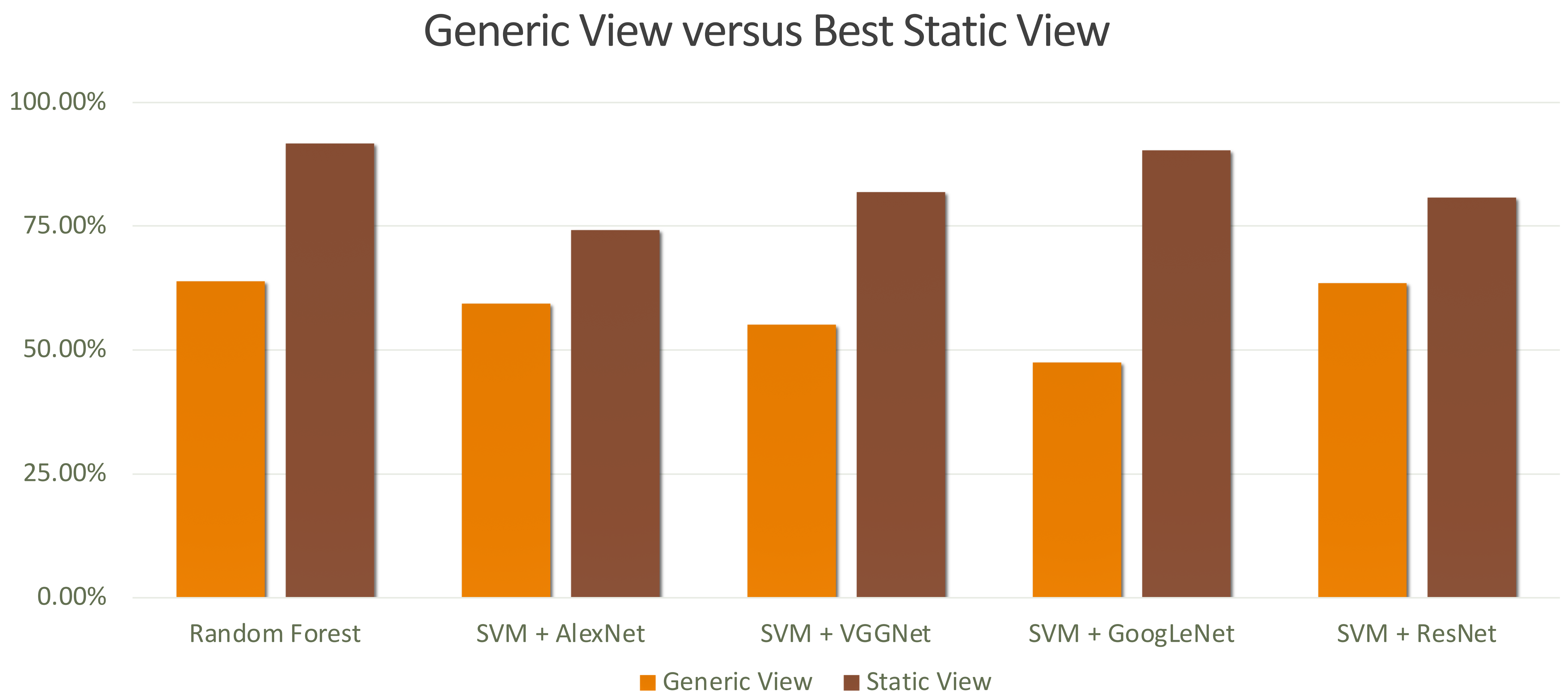
Overview

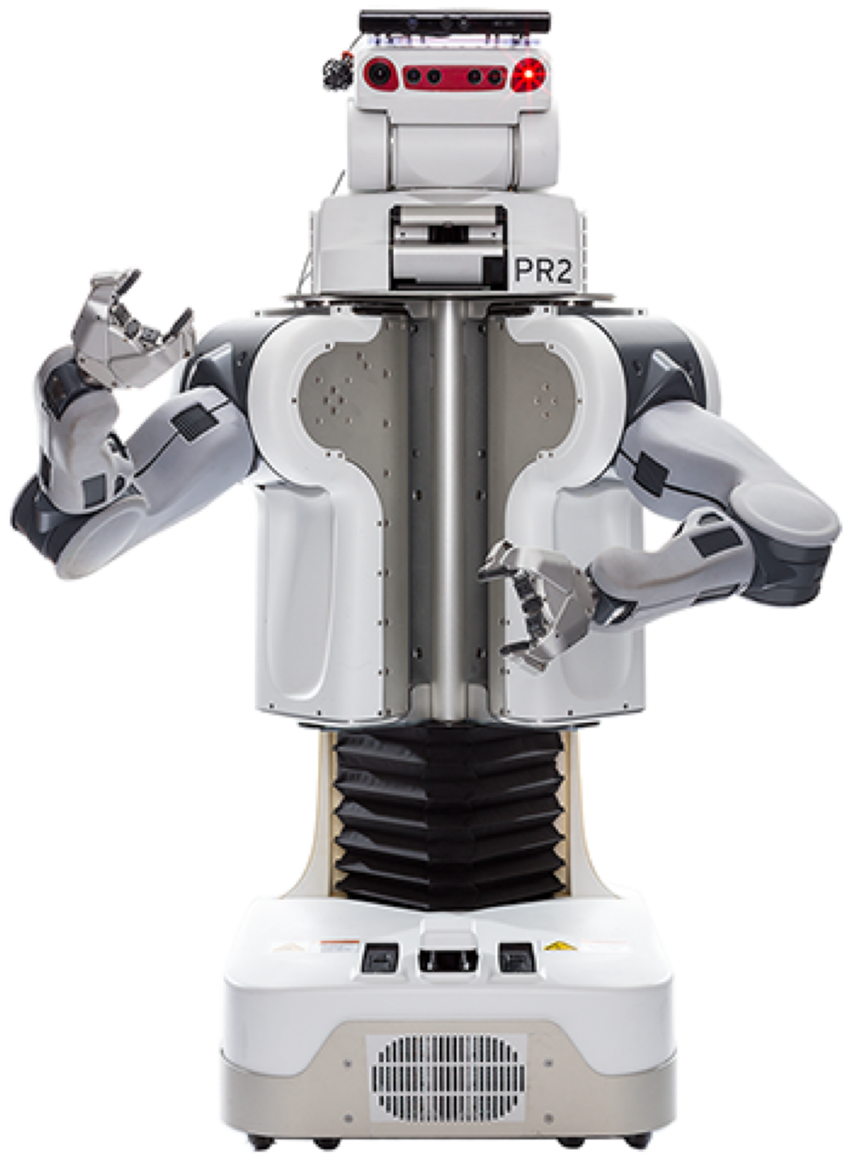
- UT Austin
  - Viewpoint Selection for Visual Failure Detection
  - Human Gaze Following for Human-Robot Interaction
  - **Active Reward Learning from Critiques**
  - **Incremental Task Modification via Corrective Demonstrations**
- Georgia Tech
  - Learning Generalizable Skills from Demonstrations
  - Skill Learning in Cluttered Environments
  - Automated Multi-Coordinate Cost Balancing
  - End User Evaluation of LfD Methods

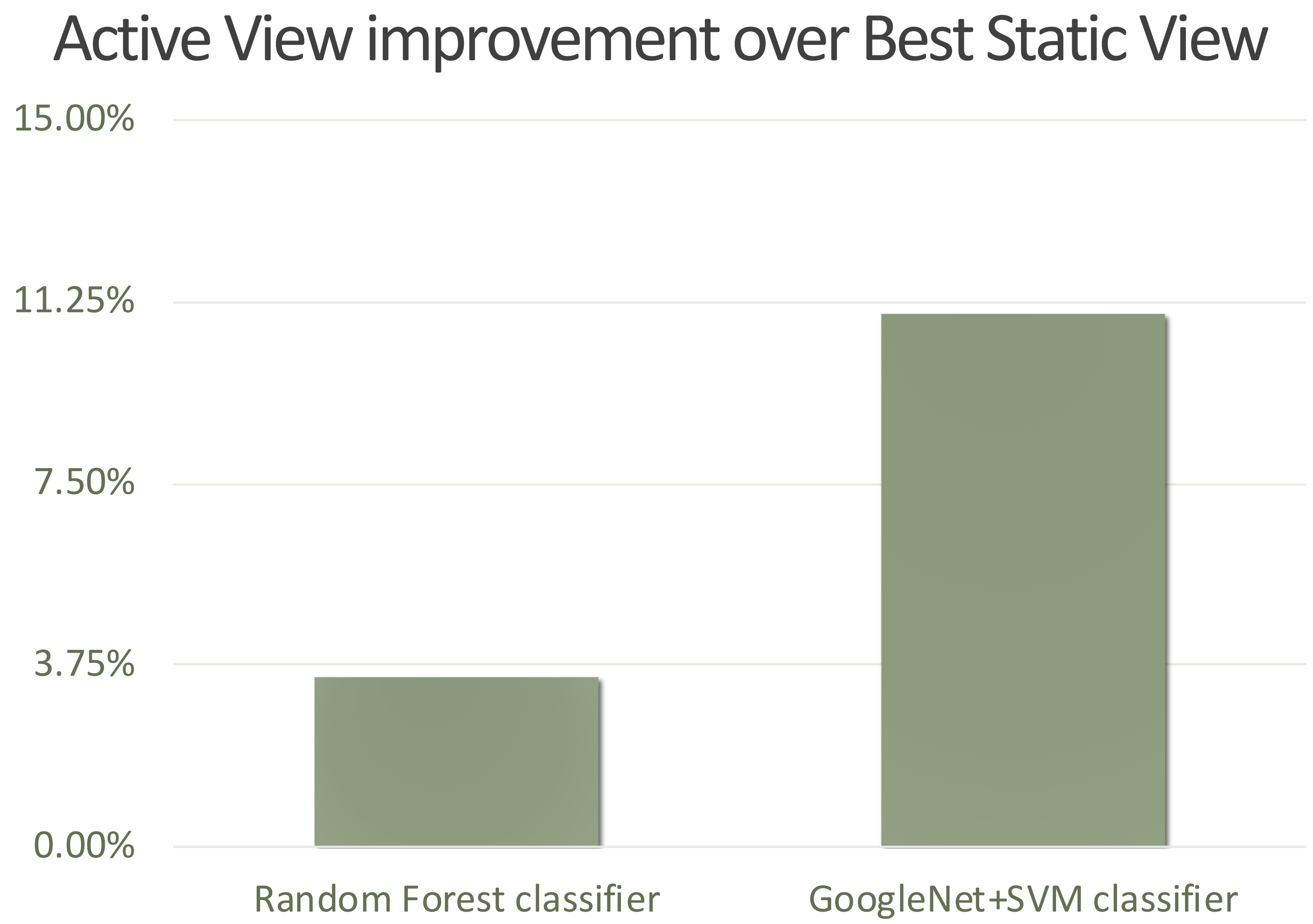


Viewpoint Selection for Visual Failure Detection



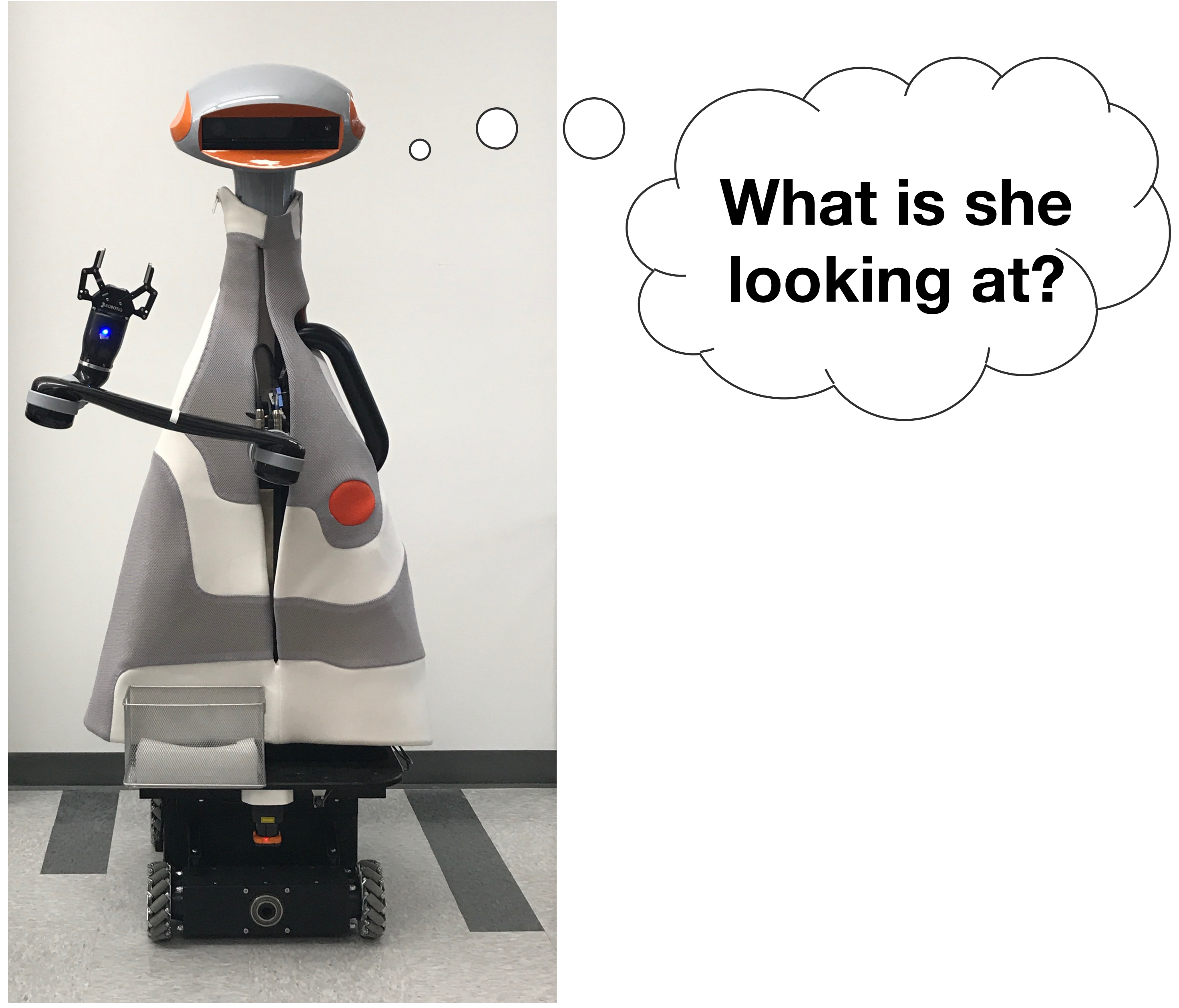






[Saran et al., IROS, 2017]

Human Gaze Following for Human-Robot Interaction

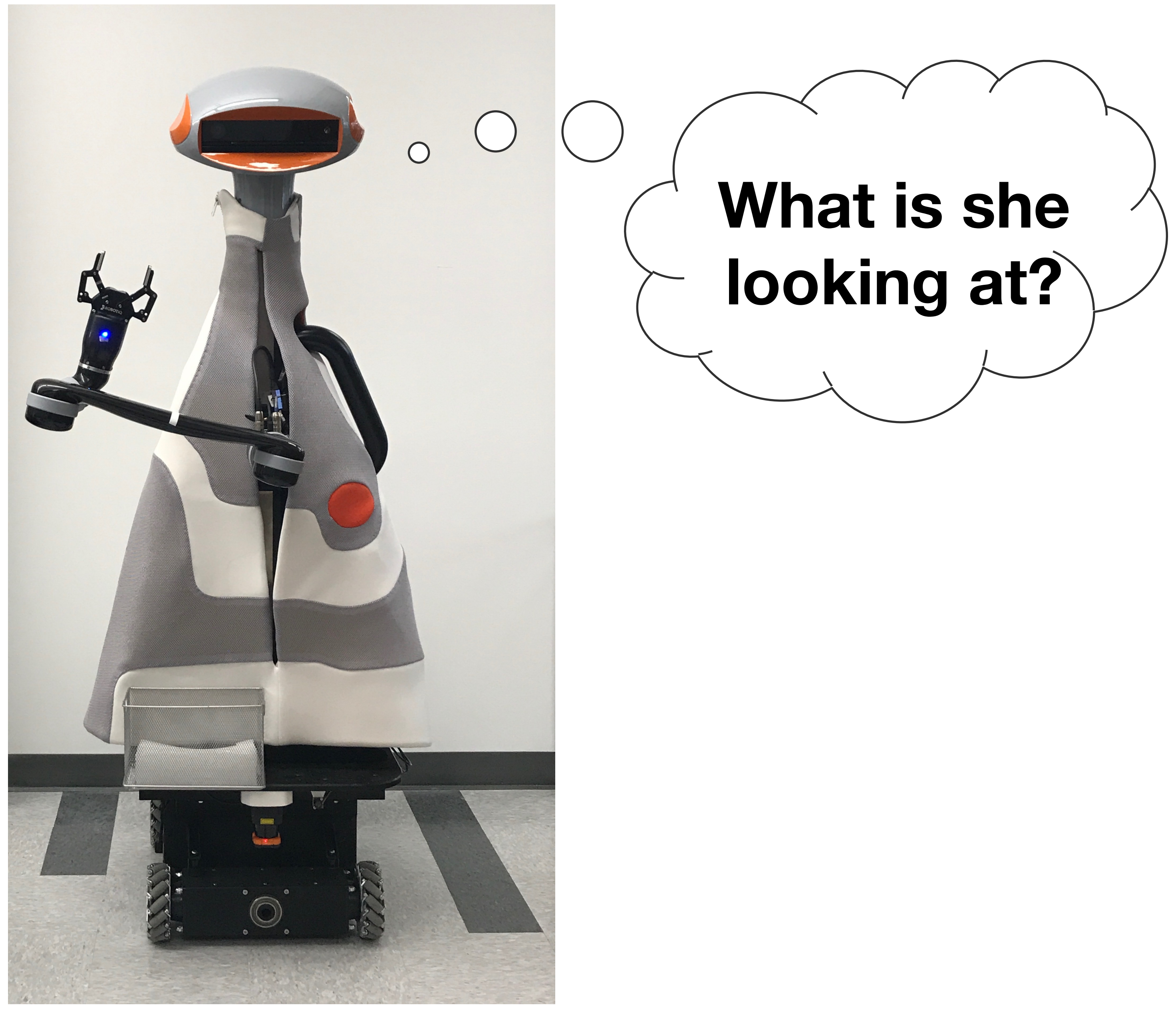


Referential Gaze

Mutual Gaze

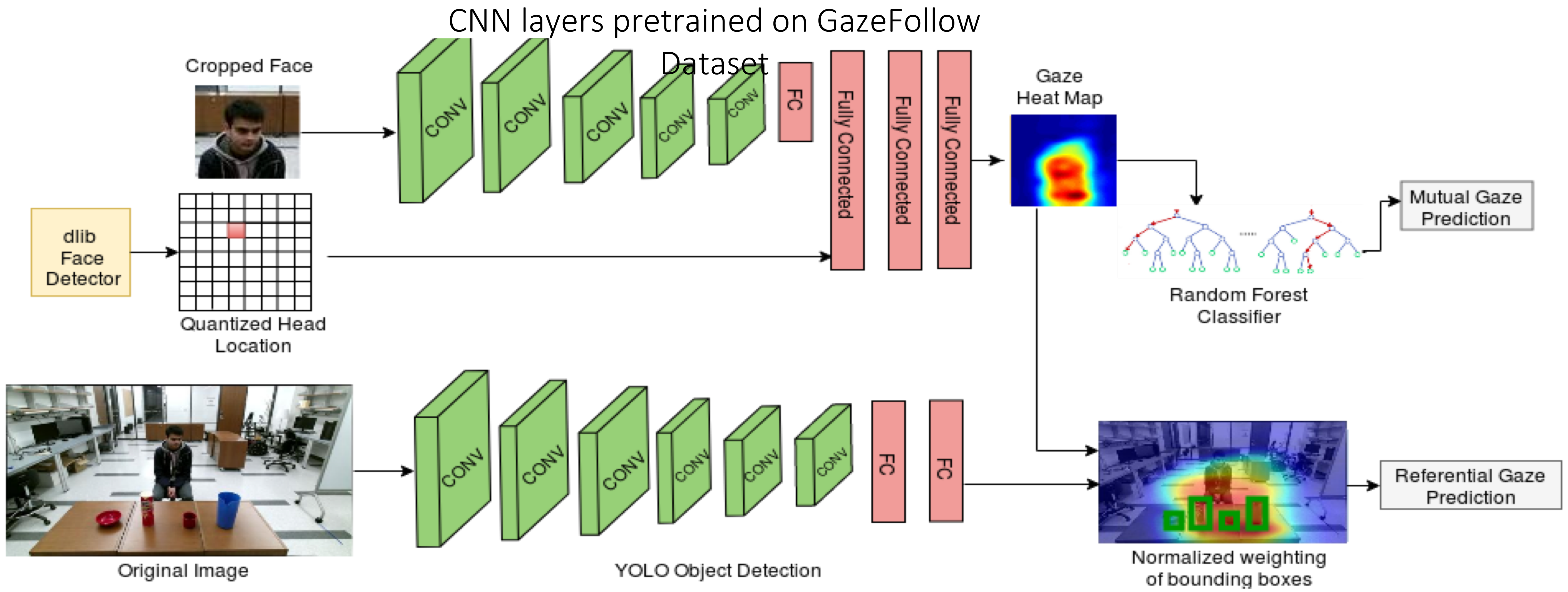
[Saran et al., IROS, 2018]

Human Gaze Following for Human-Robot Interaction

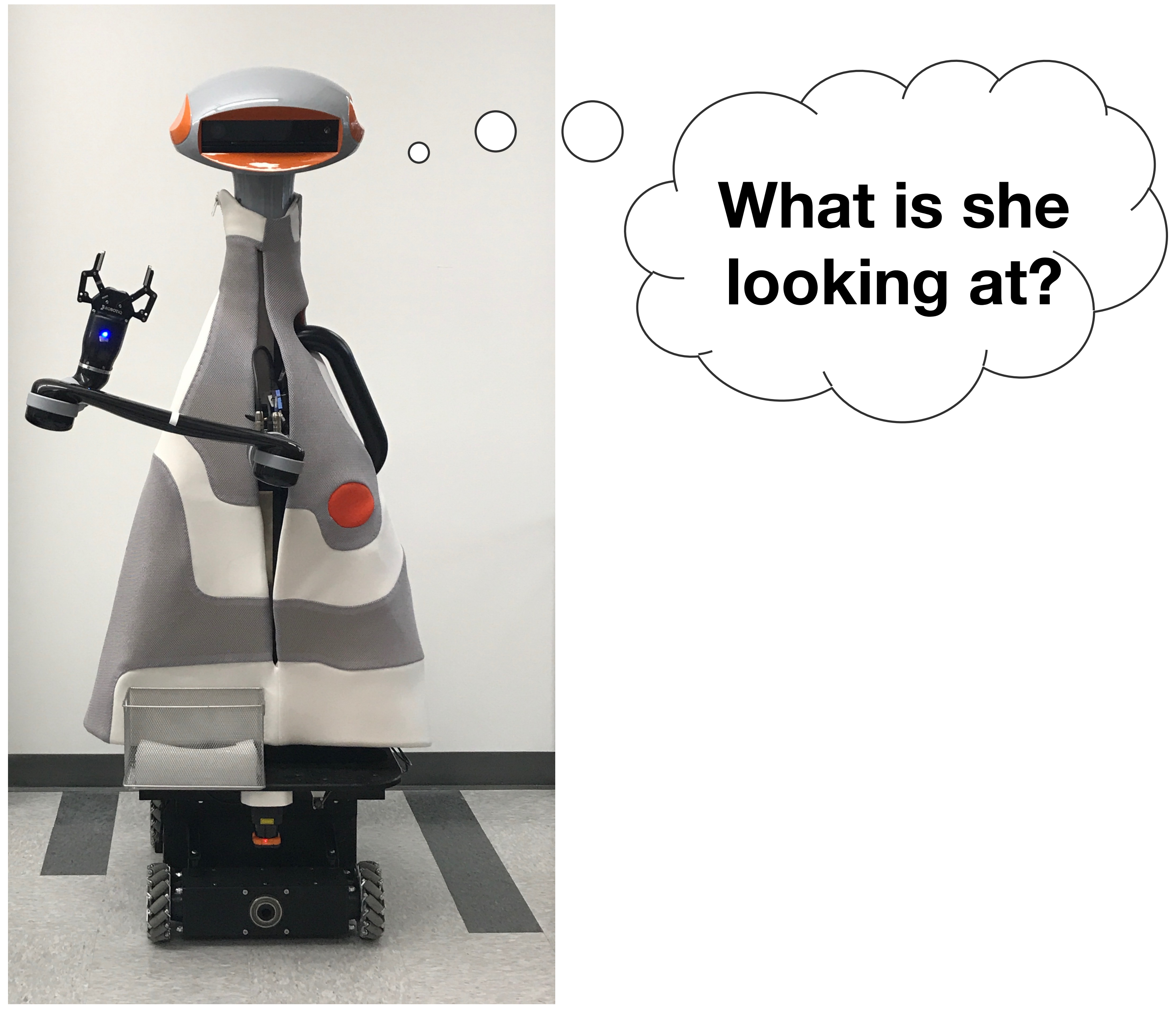


Referential Gaze

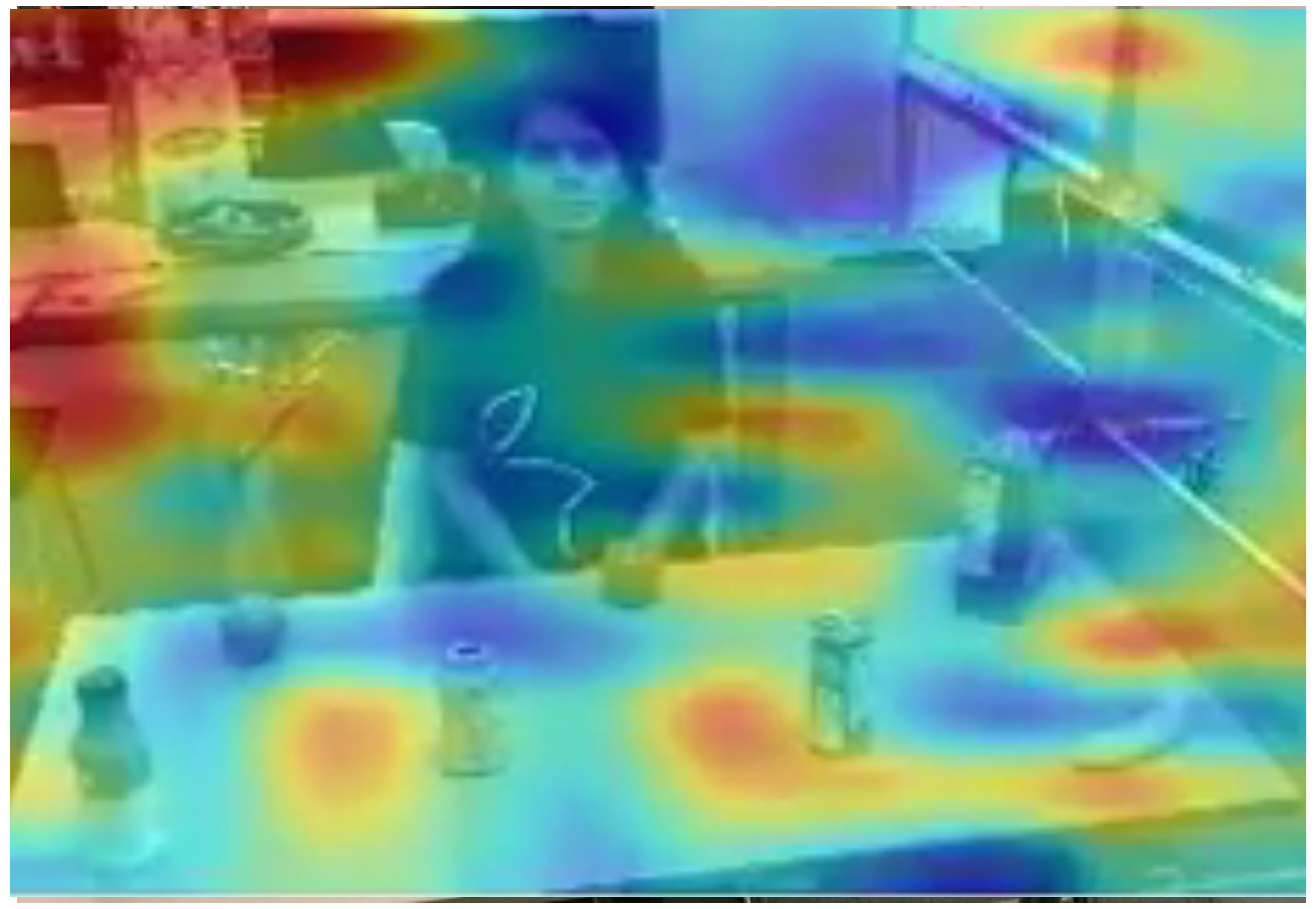
Mutual Gaze



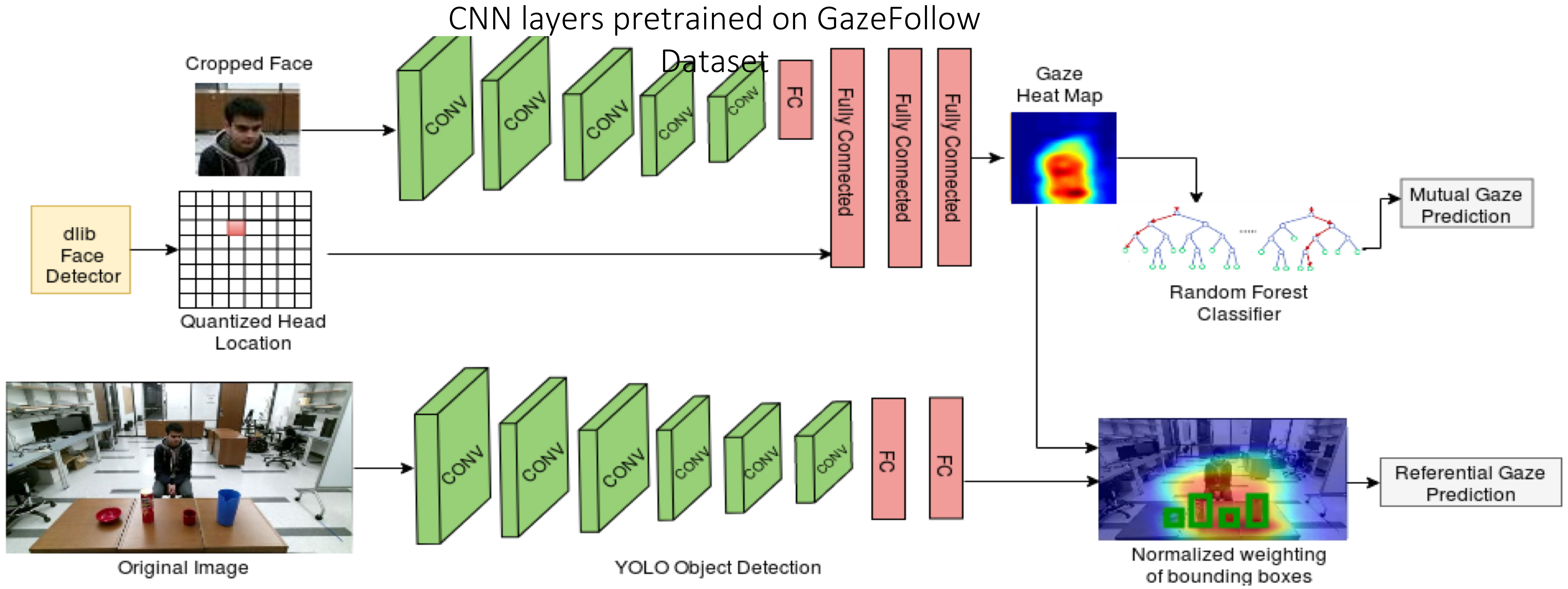
Human Gaze Following for Human-Robot Interaction



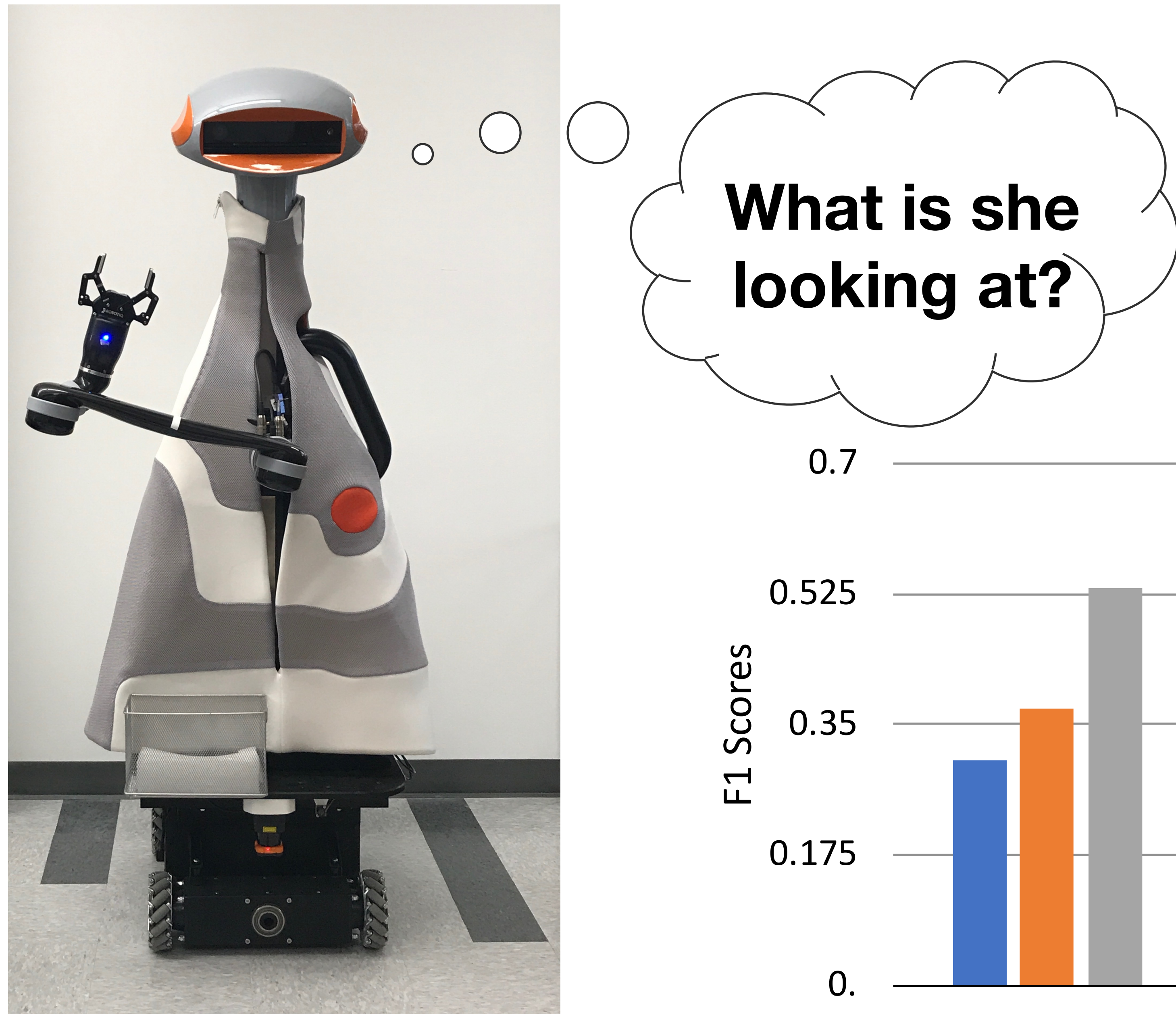
Referential Gaze



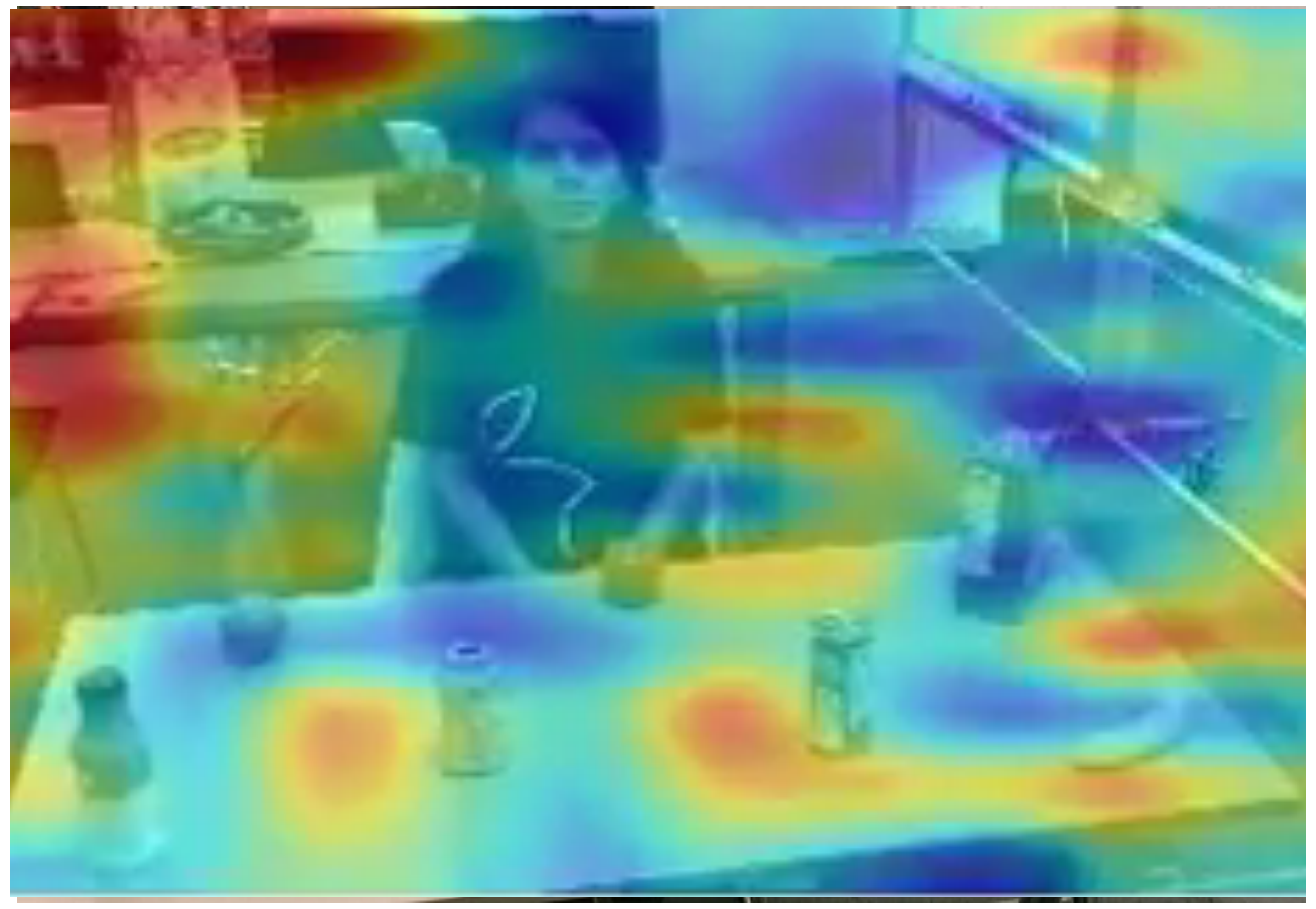
Mutual Gaze



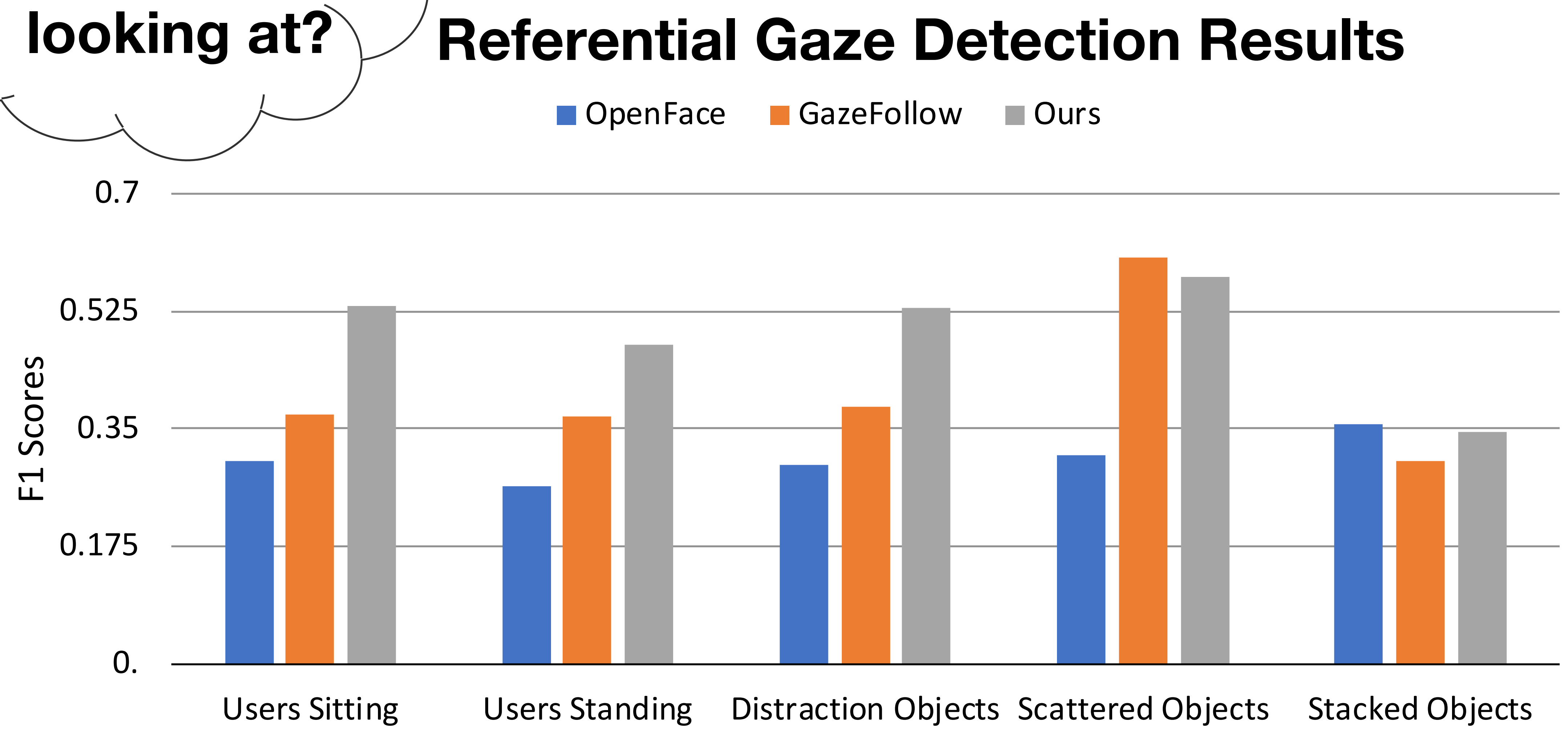
Human Gaze Following for Human-Robot Interaction

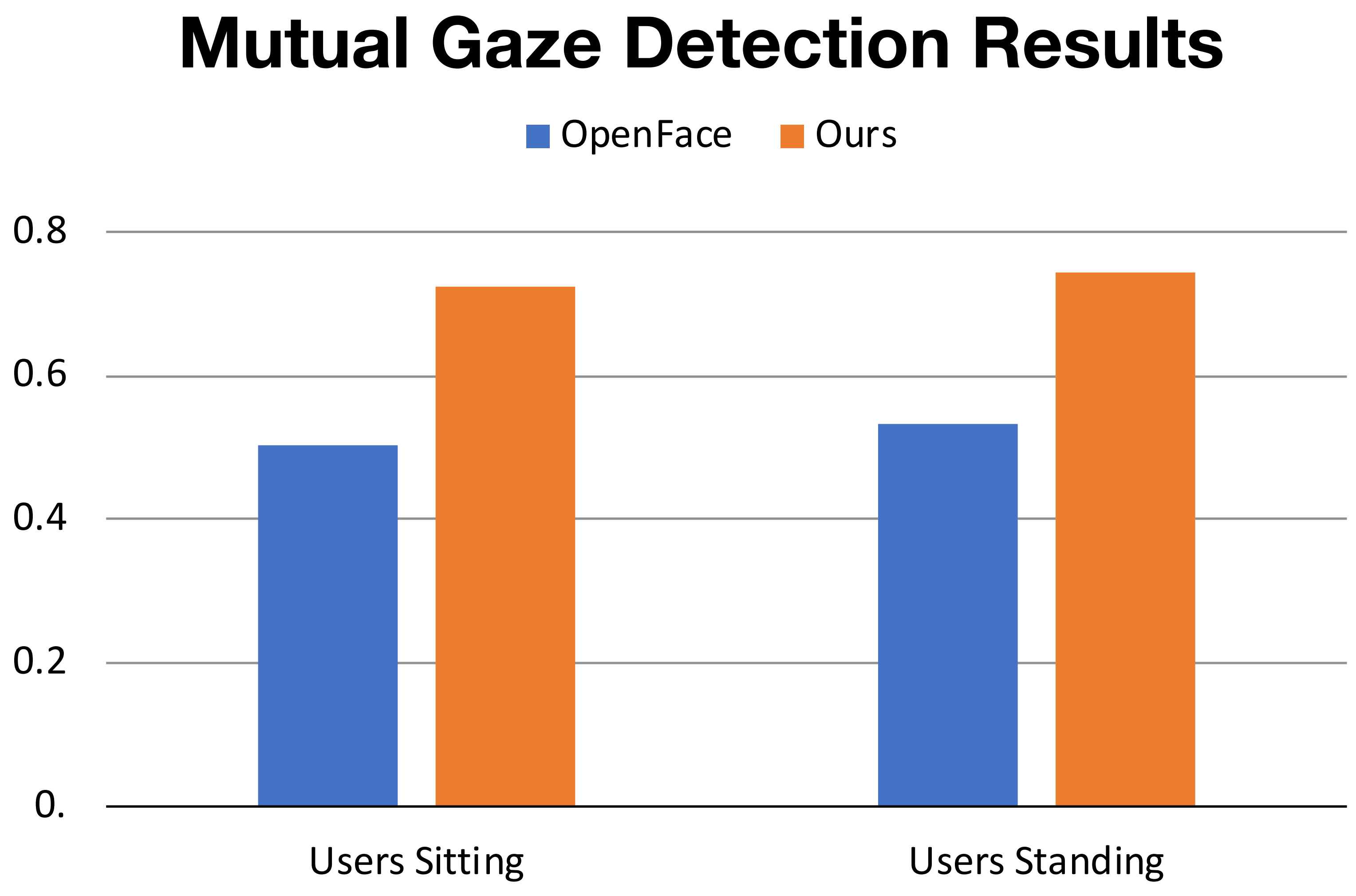


Referential Gaze

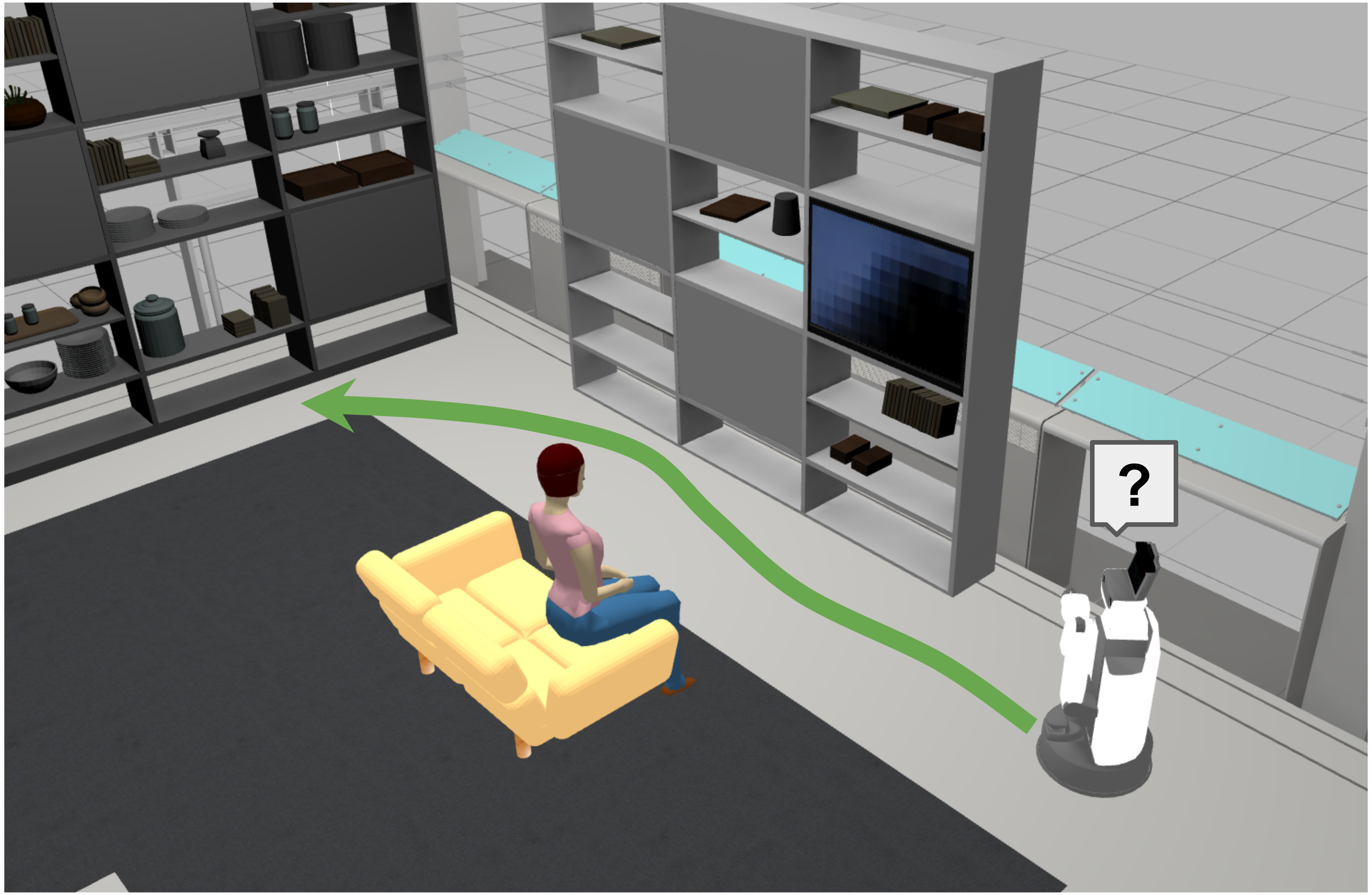


Mutual Gaze





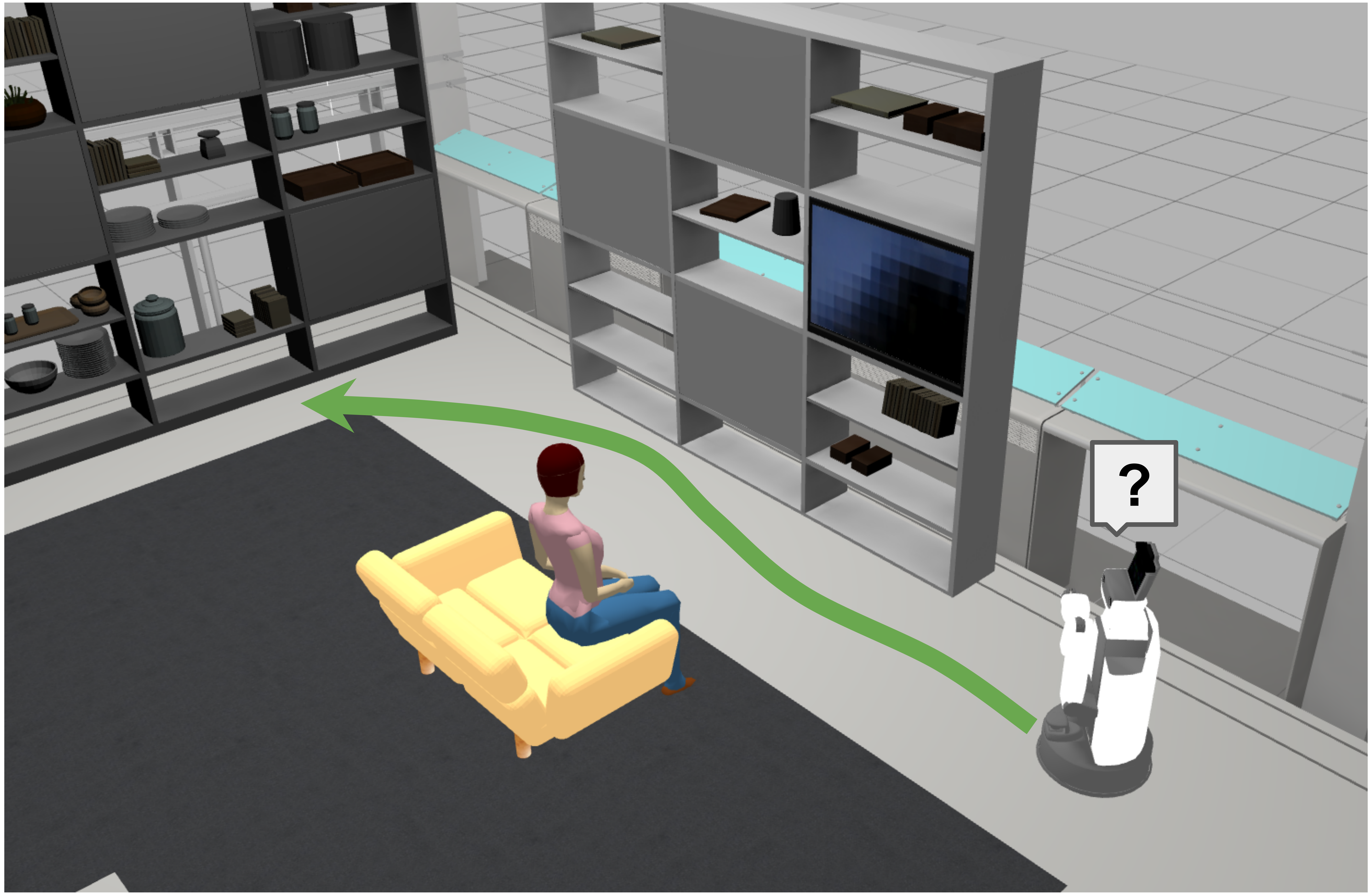
Active Reward Learning from Critiques



[Cui et al., ICRA, 2018]

Use active learning to reduce burden on supervisor

Active Reward Learning from Critiques



Propose maximum information gain path

$$G^+(s_i,a_i) = G(D^+ \cup (s_i,a_i) \mid Be(R)) = \boxed{Pr(a_i \in O(s_i) \mid Be(R))} D(Be'(R) || Be(R))$$
$$G^-(s_i,a_i) = G(D^- \cup (s_i,a_i) \mid Be(R)) = \boxed{Pr(a_i \notin O(s_i) \mid Be(R))} D(Be'(R) || Be(R))$$

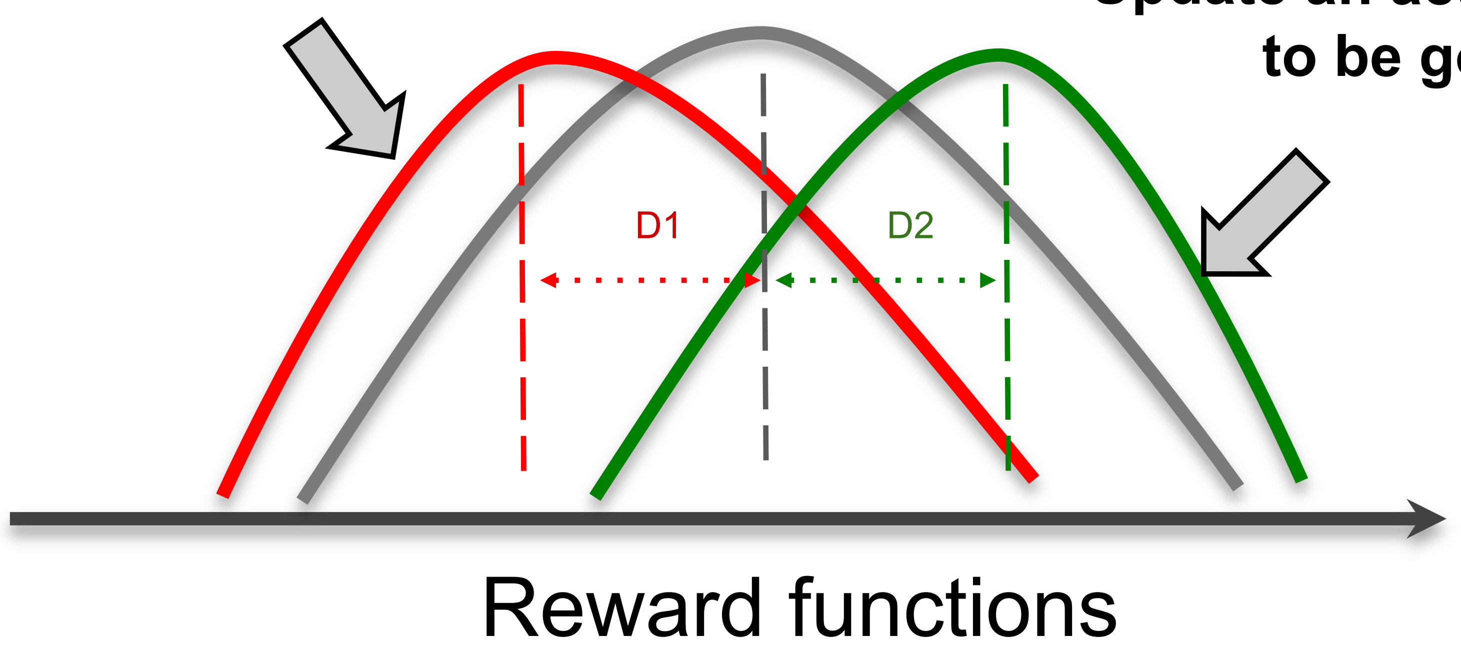
$$D_{KL}(P||Q) = \sum_i P(i) \log \frac{P(i)}{Q(i)}$$

$$\boxed{Pr(a_i \notin O(s_i) \mid R) = 1 - \frac{1}{Z_i} e^{\alpha Q(s_i,a_i,R)}}$$

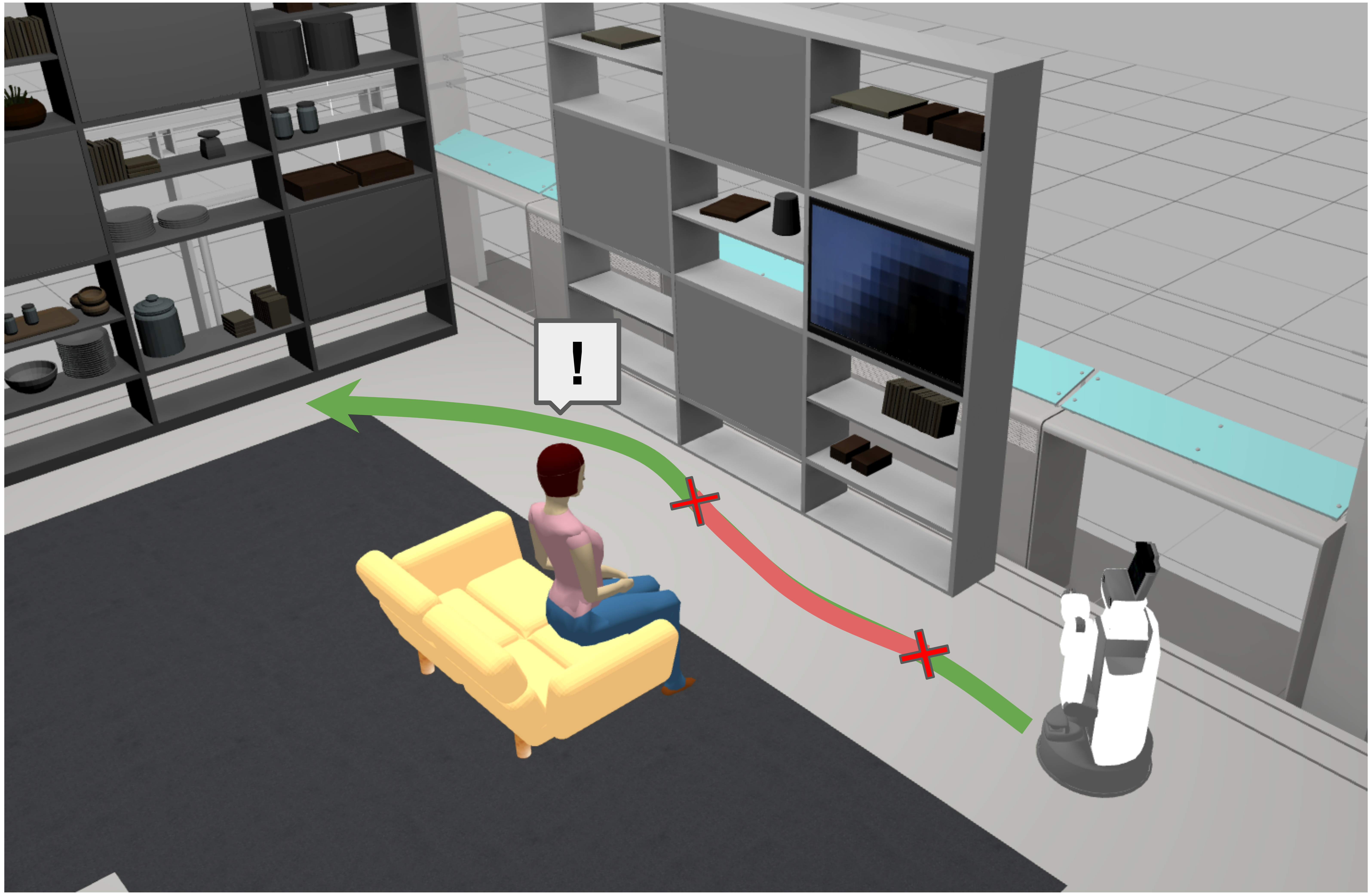
$$\boxed{Pr(a_i \in O(s_i) \mid R) = \frac{1}{Z_i} e^{\alpha Q(s_i,a_i,R)}}$$

Update an action to be bad

Update an action to be good

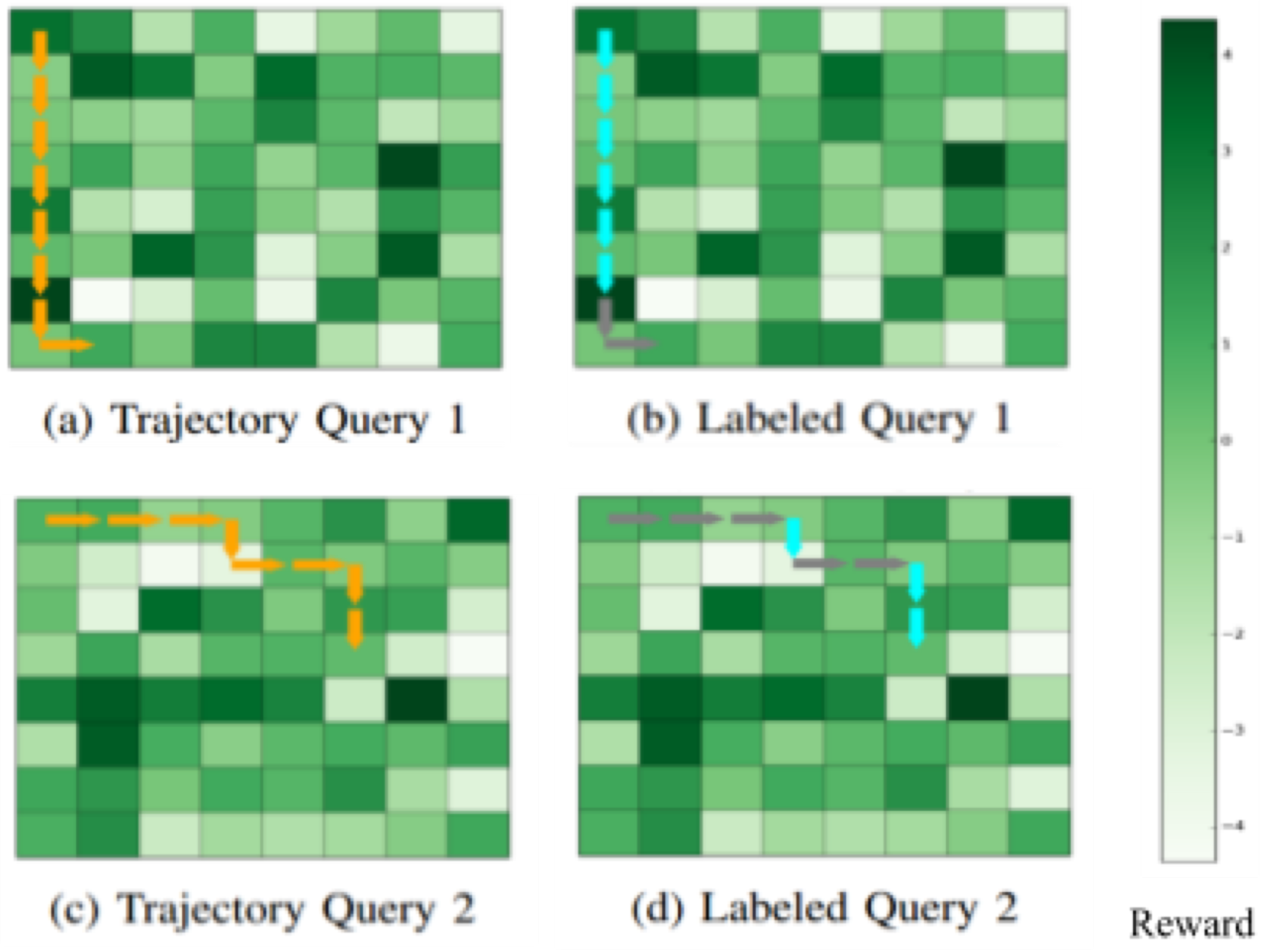


Active Reward Learning from Critiques

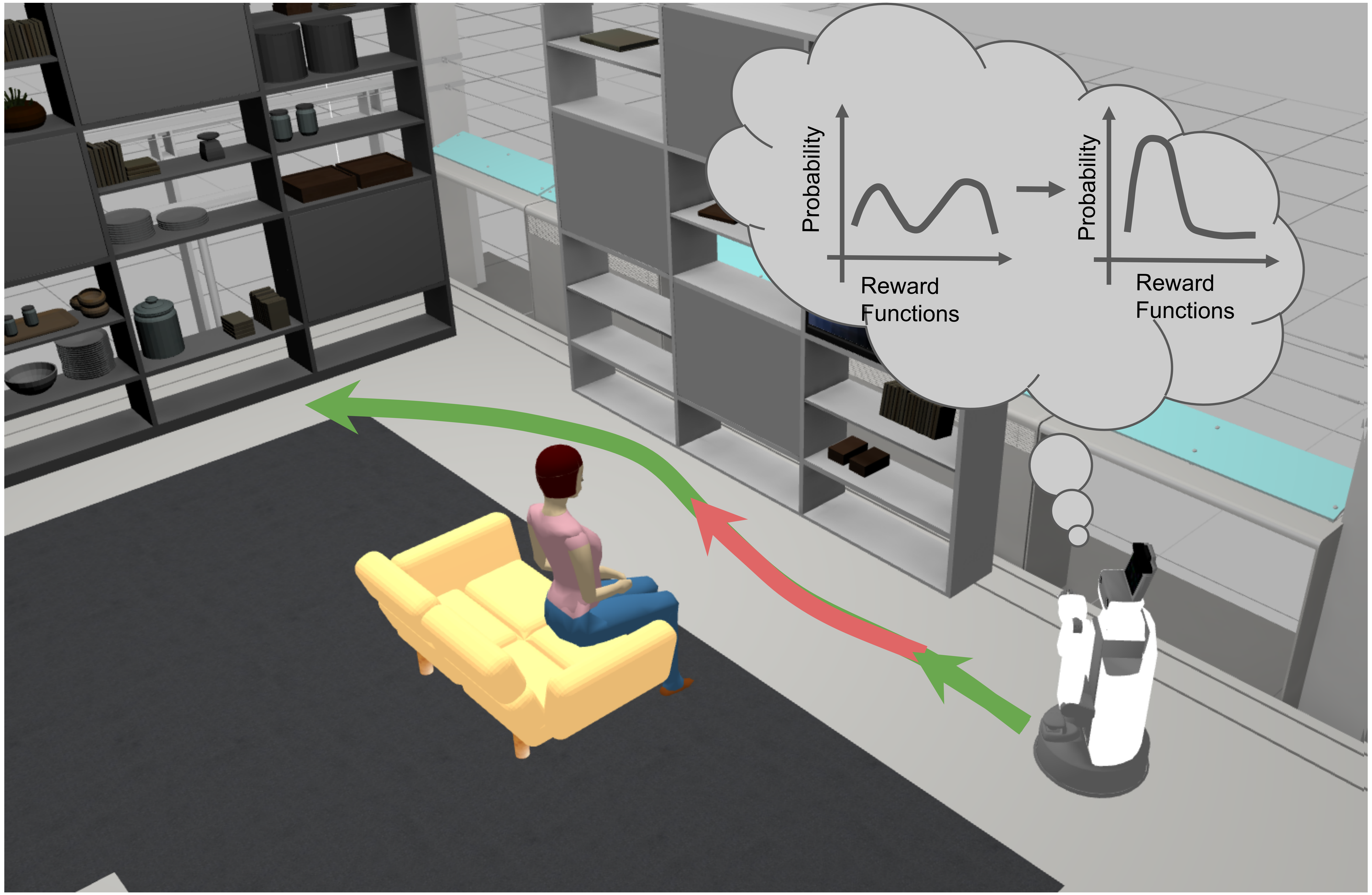


Receive human segmentation of good and bad subpaths

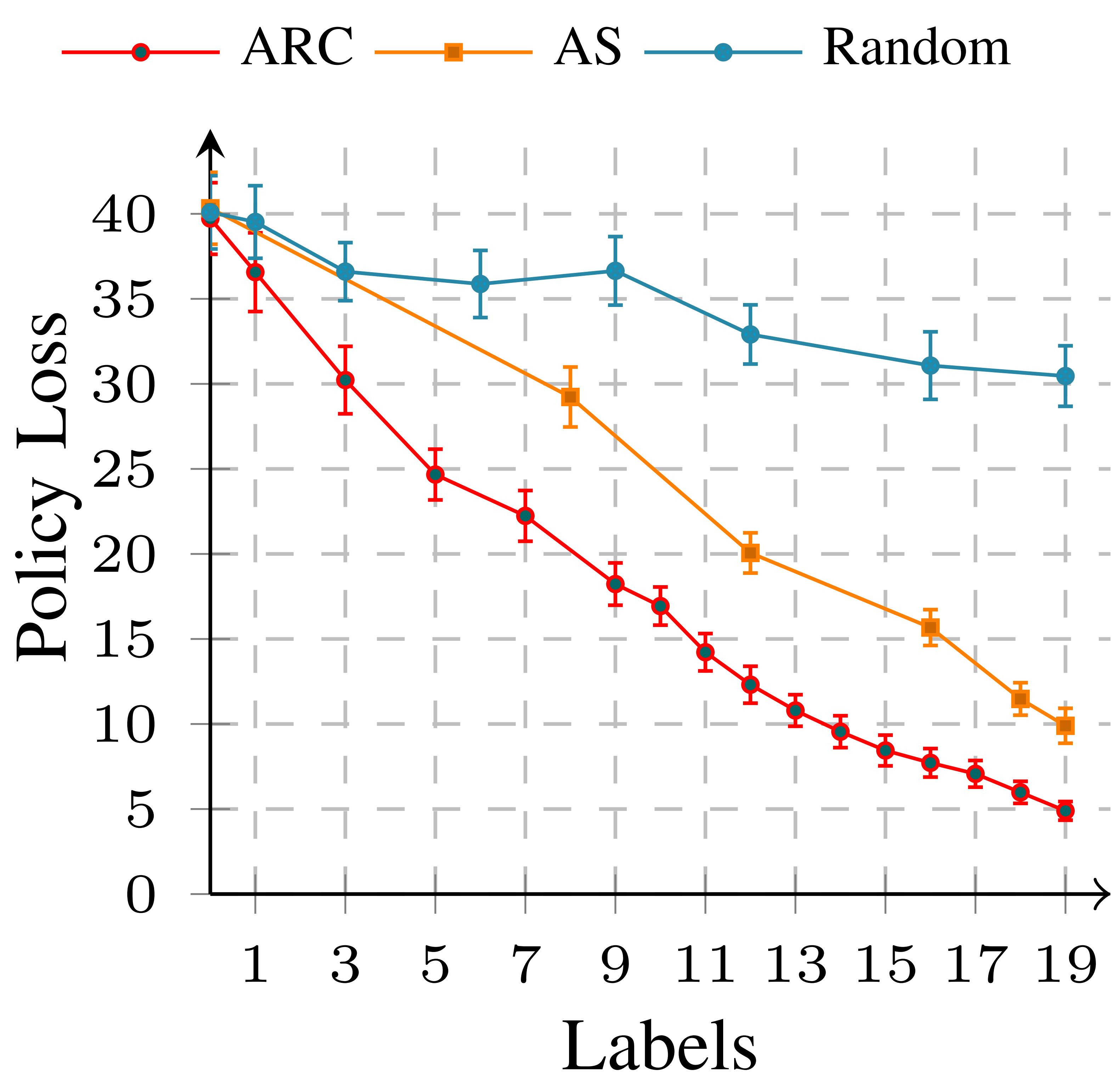
Gridworld Navigation Task



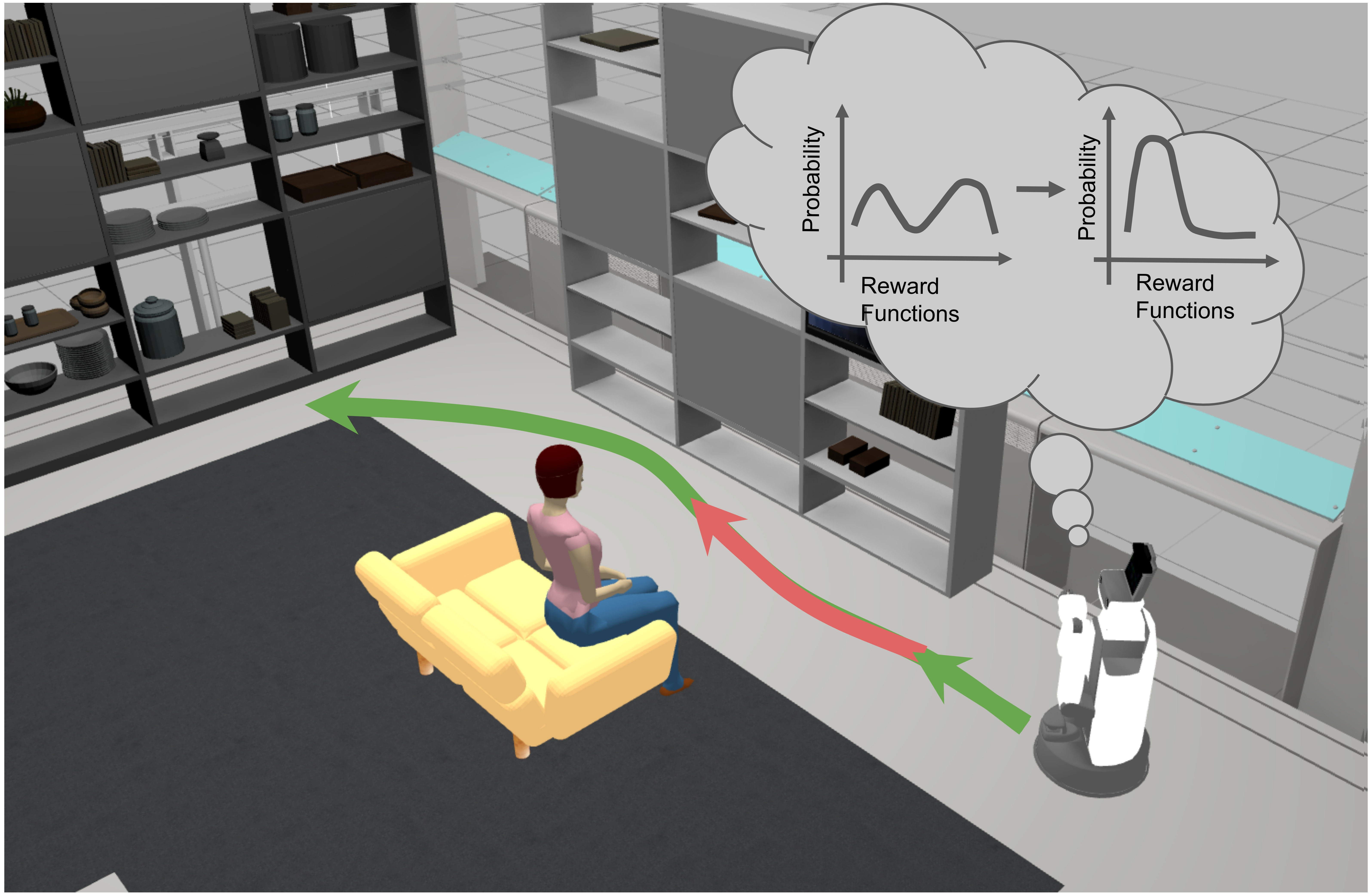
Active Reward Learning from Critiques



Update models

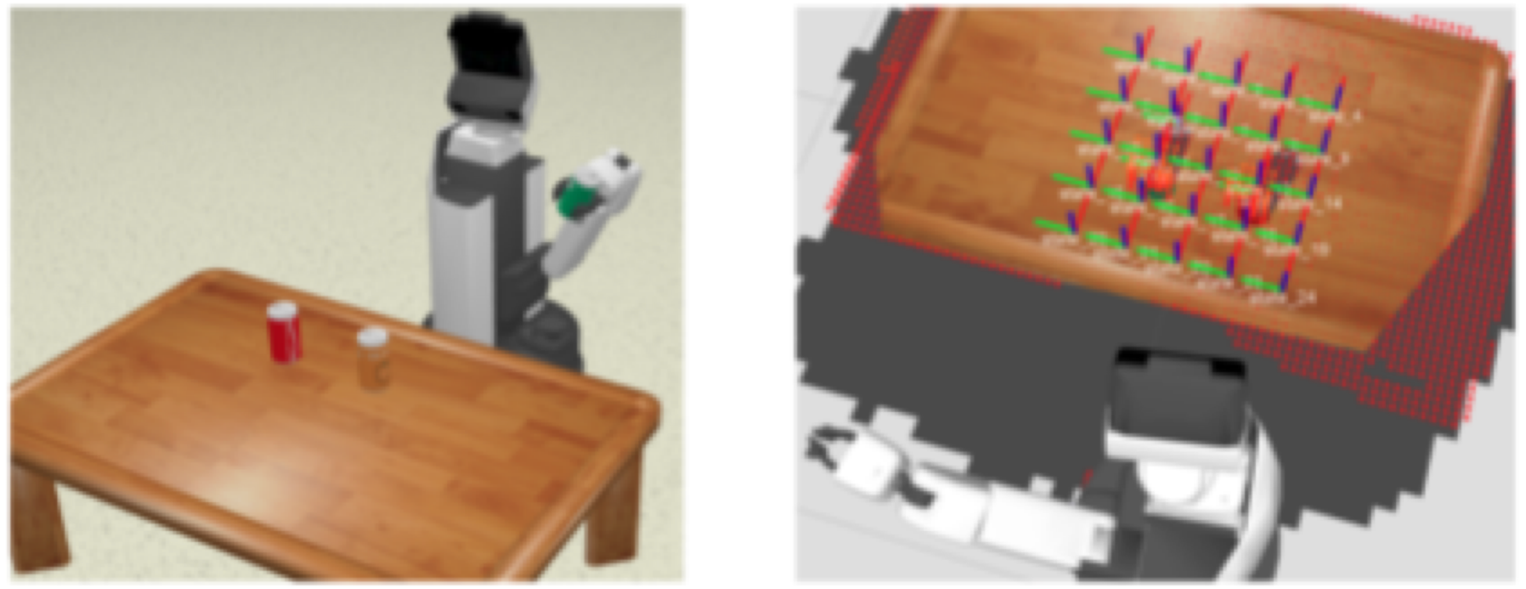


Active Reward Learning from Critiques



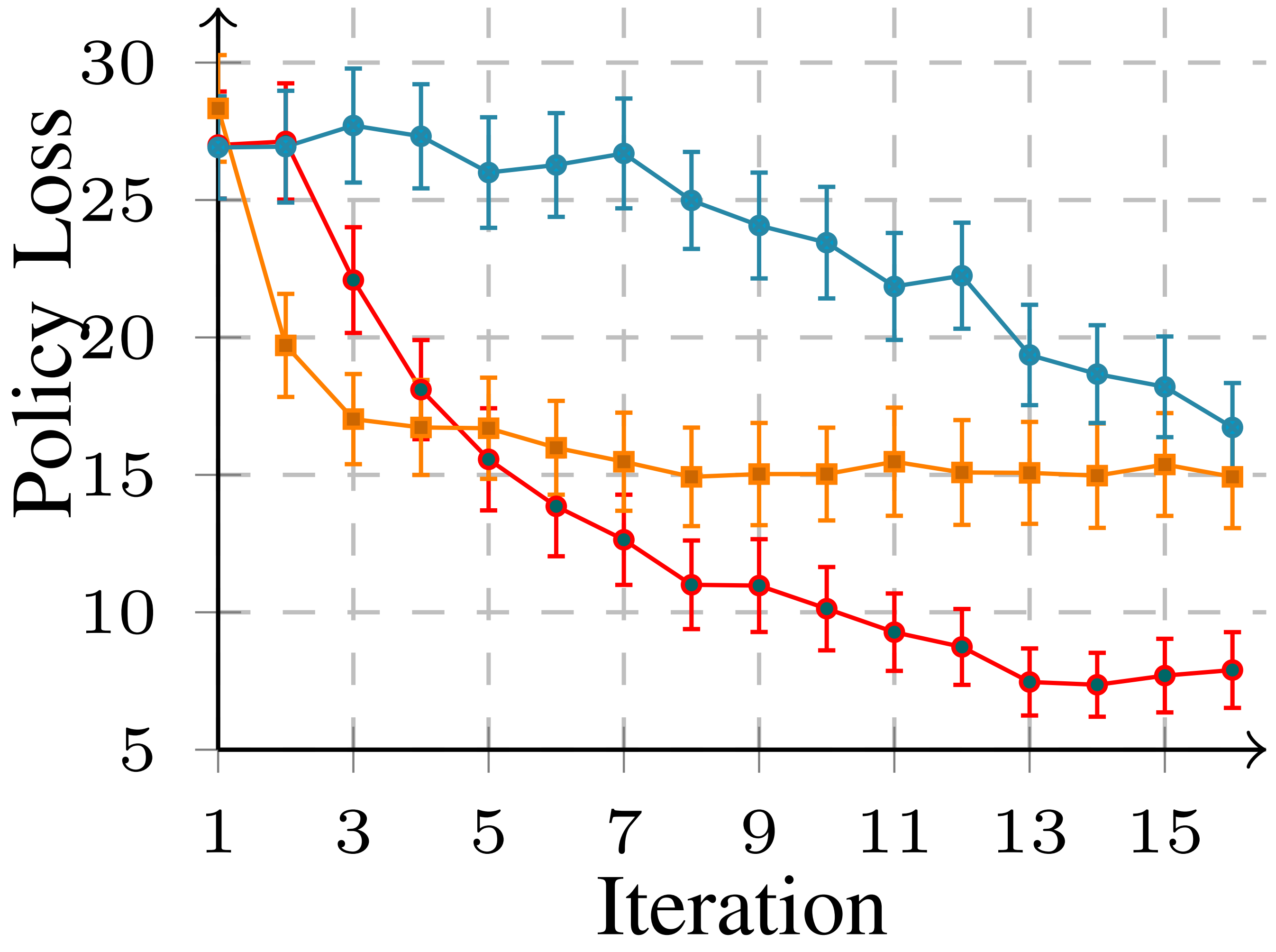
Update models

Object Placement Task

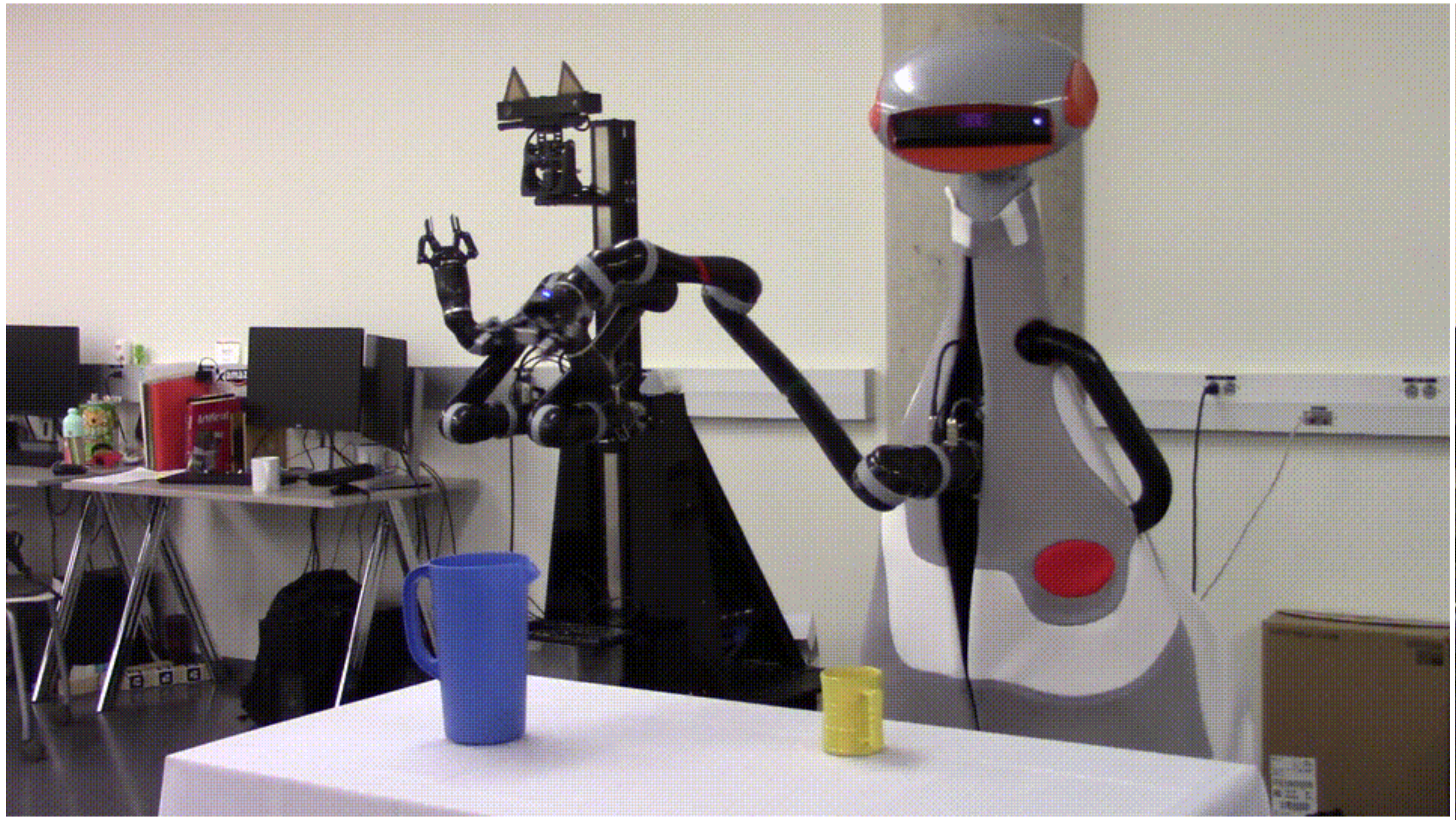


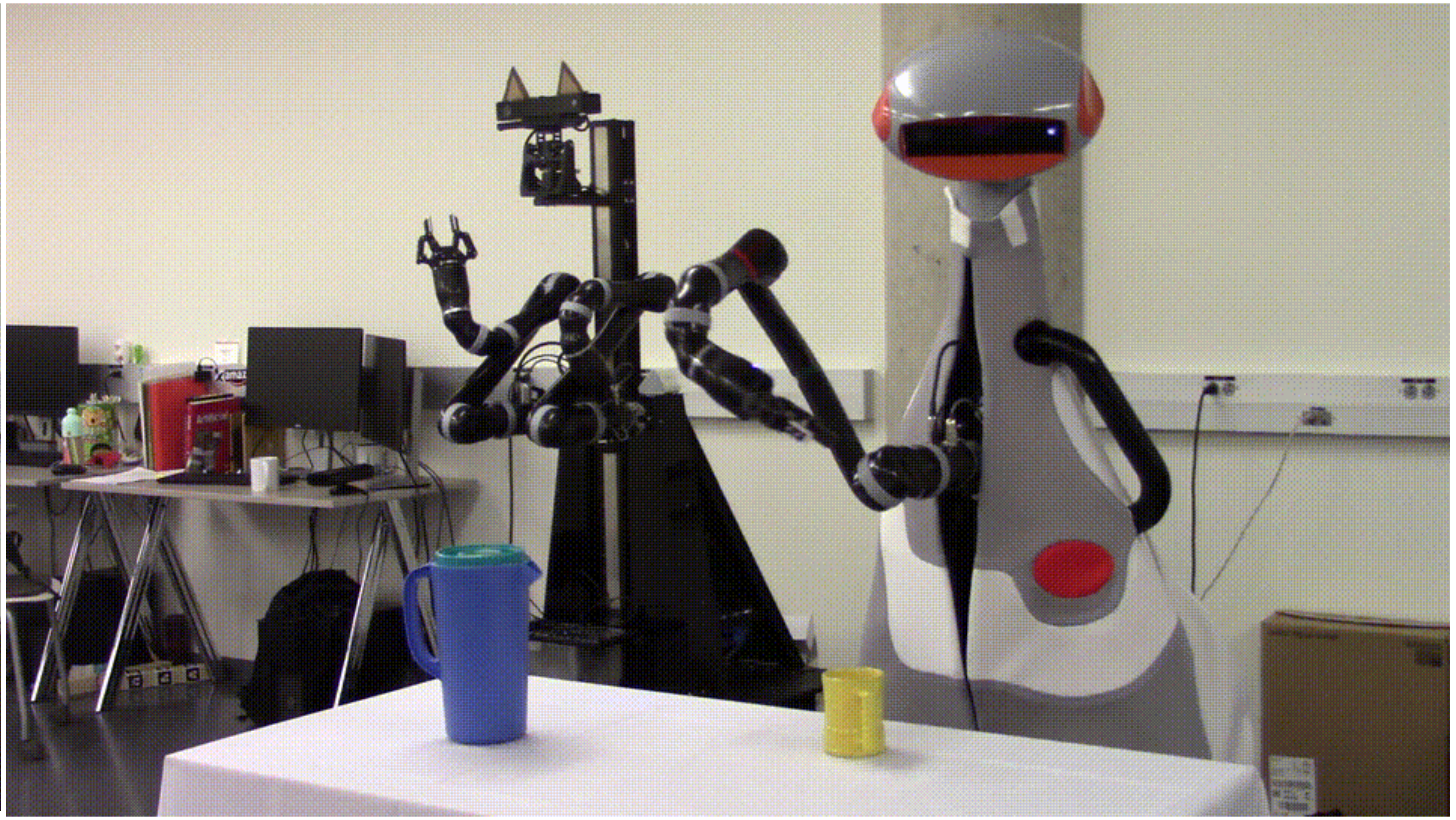
(a) Simulated Scene

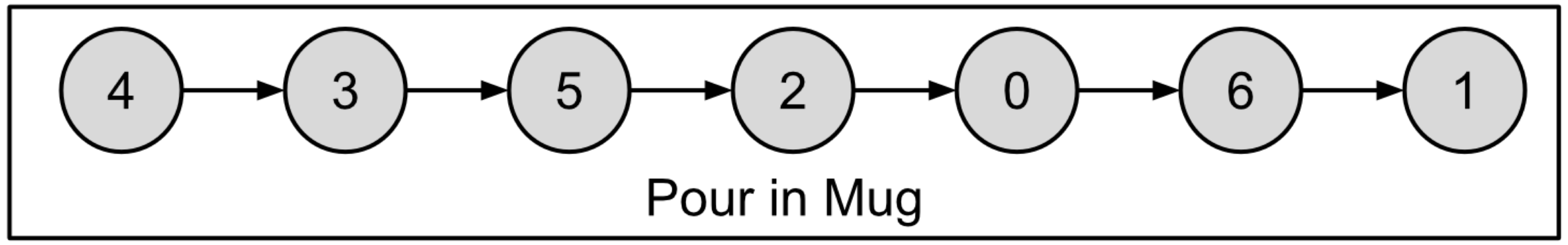
(b) Detected States

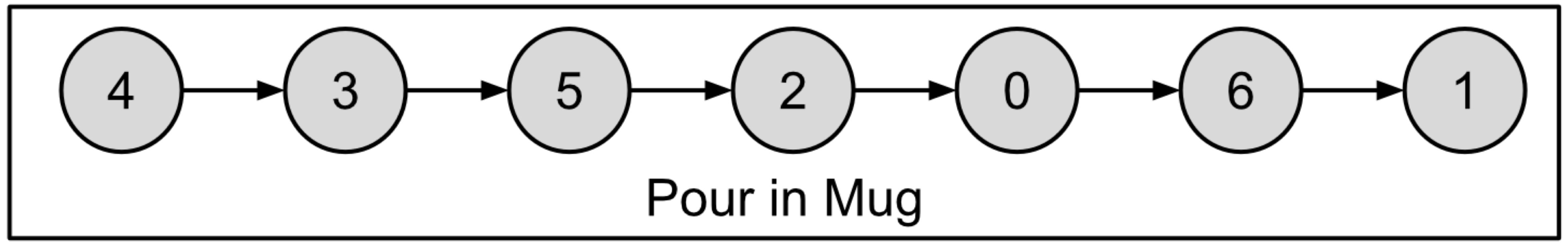


Incremental Task Modification via Corrective Demonstrations







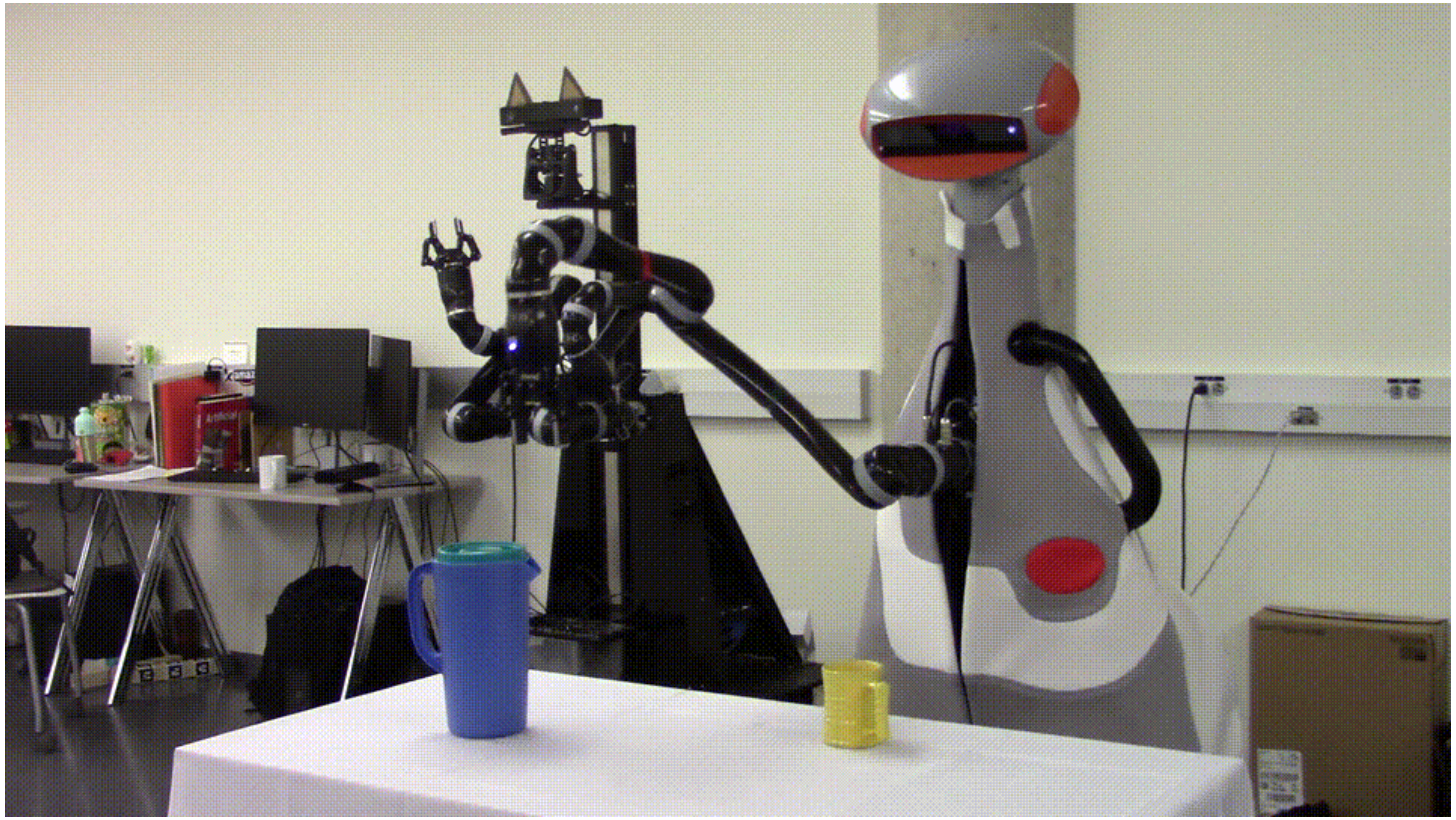


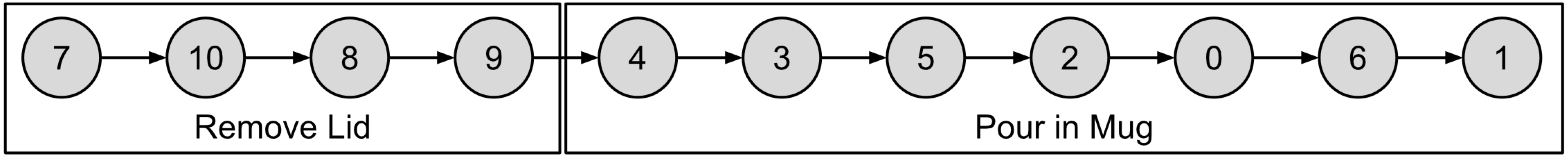
[Gutierrez et al., ICRA, 2018]

Need to update models with corrections

Incremental Task Modification via Corrective Demonstrations

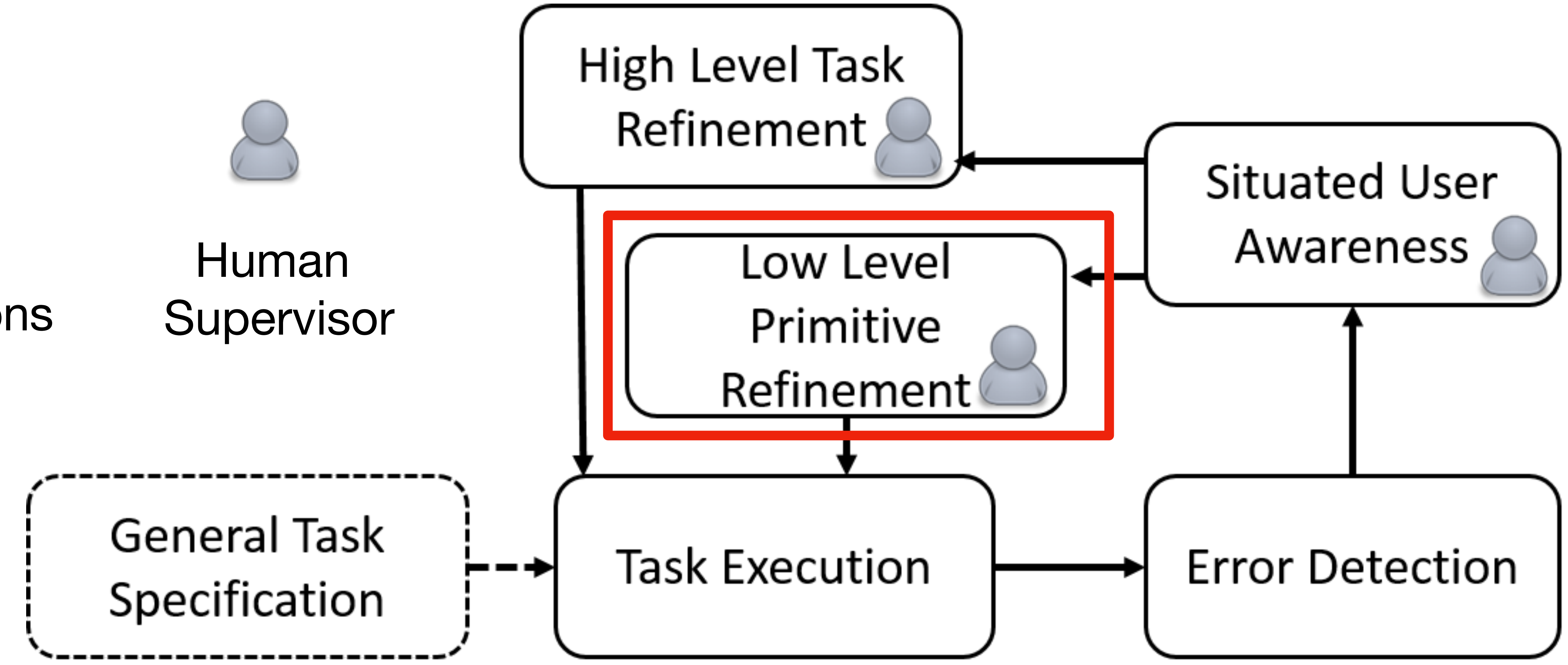
Incremental Task Modification via Corrective Demonstrations





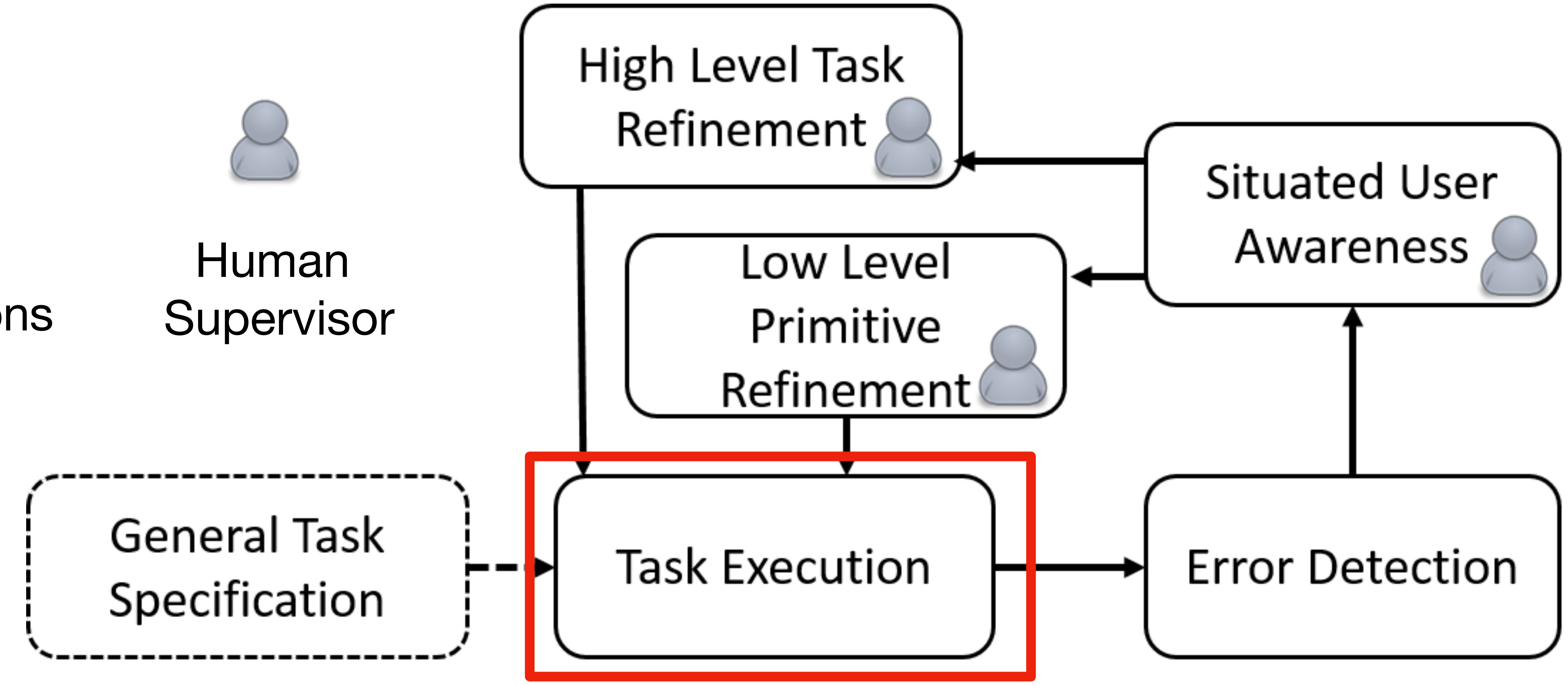
Overview

- UT Austin
  - Viewpoint Selection for Visual Failure Detection
  - Human Gaze Following for Human-Robot Interaction
  - Active Reward Learning from Critiques
  - Incremental Task Modification via Corrective Demonstrations
- **Georgia Tech**
  - **Learning Generalizable Skills from Demonstrations**
  - **Skill Learning in Cluttered Environments**
  - **Automated Multi-Coordinate Cost Balancing**
  - End User Evaluation of LfD Methods



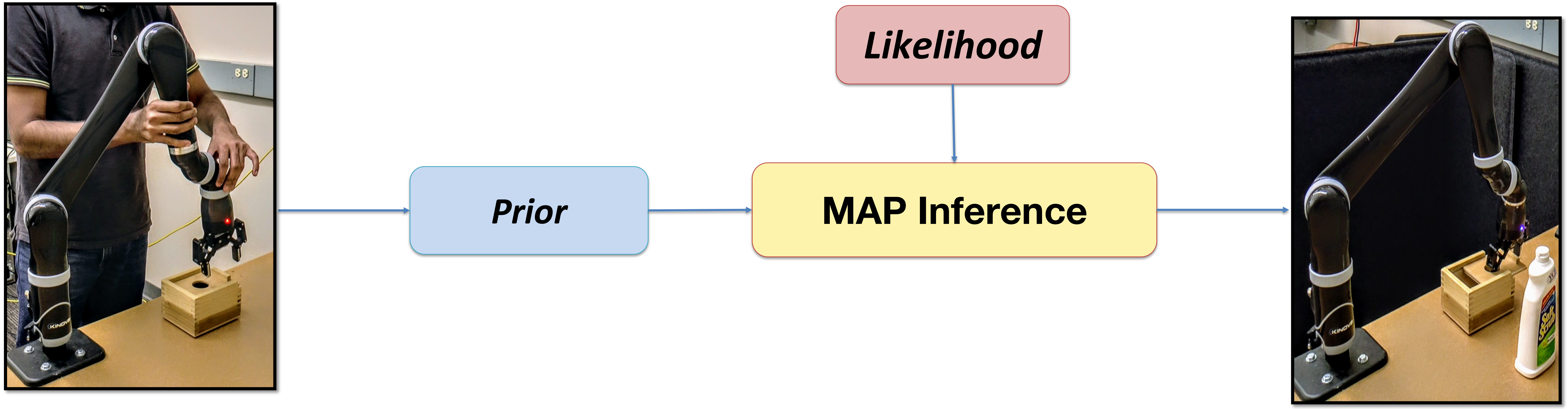
Overview

- UT Austin
  - Viewpoint Selection for Visual Failure Detection
  - Human Gaze Following for Human-Robot Interaction
  - Active Reward Learning from Critiques
  - Incremental Task Modification via Corrective Demonstrations
- **Georgia Tech**
  - **Learning Generalizable Skills from Demonstrations**
  - **Skill Learning in Cluttered Environments**
  - **Automated Multi-Coordinate Cost Balancing**
  - **End User Evaluation of LfD Methods**



CLAMP
(Combined Learning from demonstration And Motion Planning)

- **Unifies** LfD and motion planning
- ***Optimal*** according to the learned skill
- ***Generalizes*** to obstacles and positional constraints

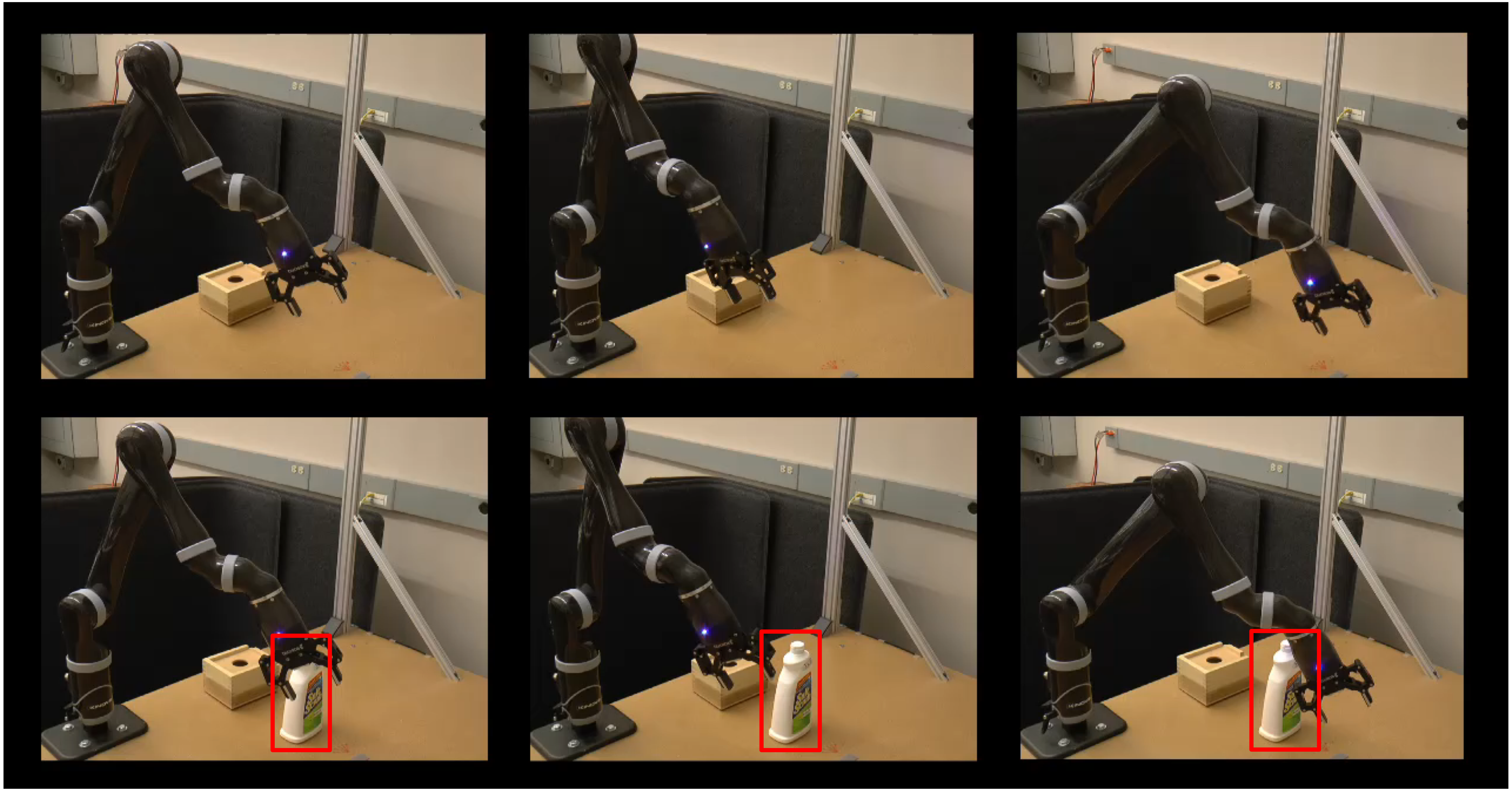


[Rana et al., CoRL, 2017]

Initial State 1 | Initial State 2 | Initial State 3

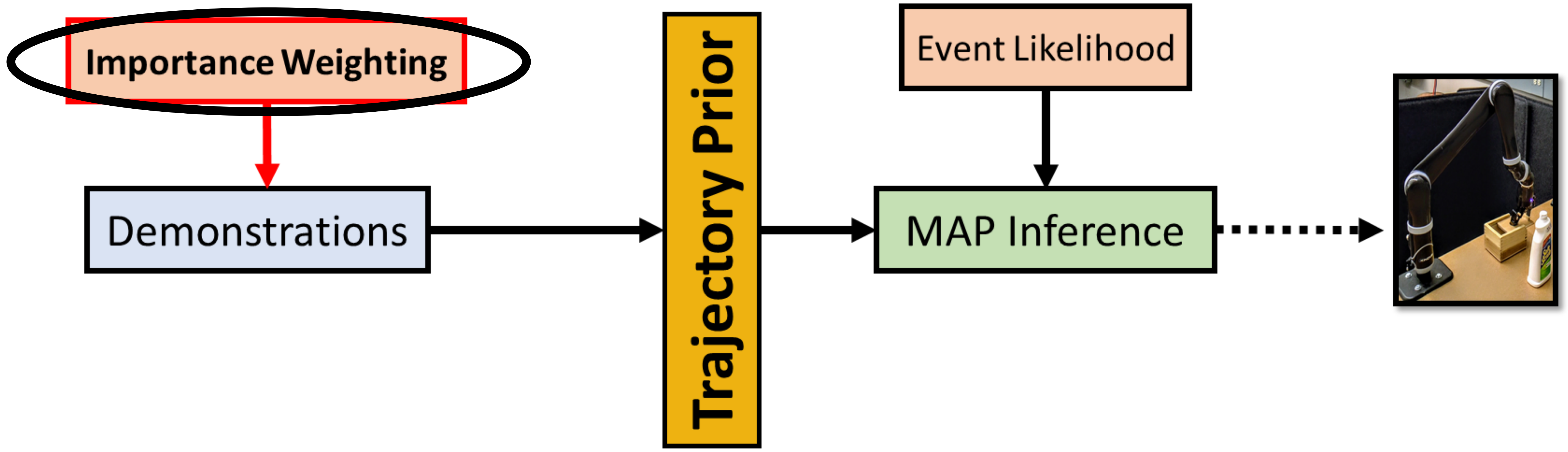
w/o obstacle

w/ obstacle



[Rana et al., CoRL, 2017]

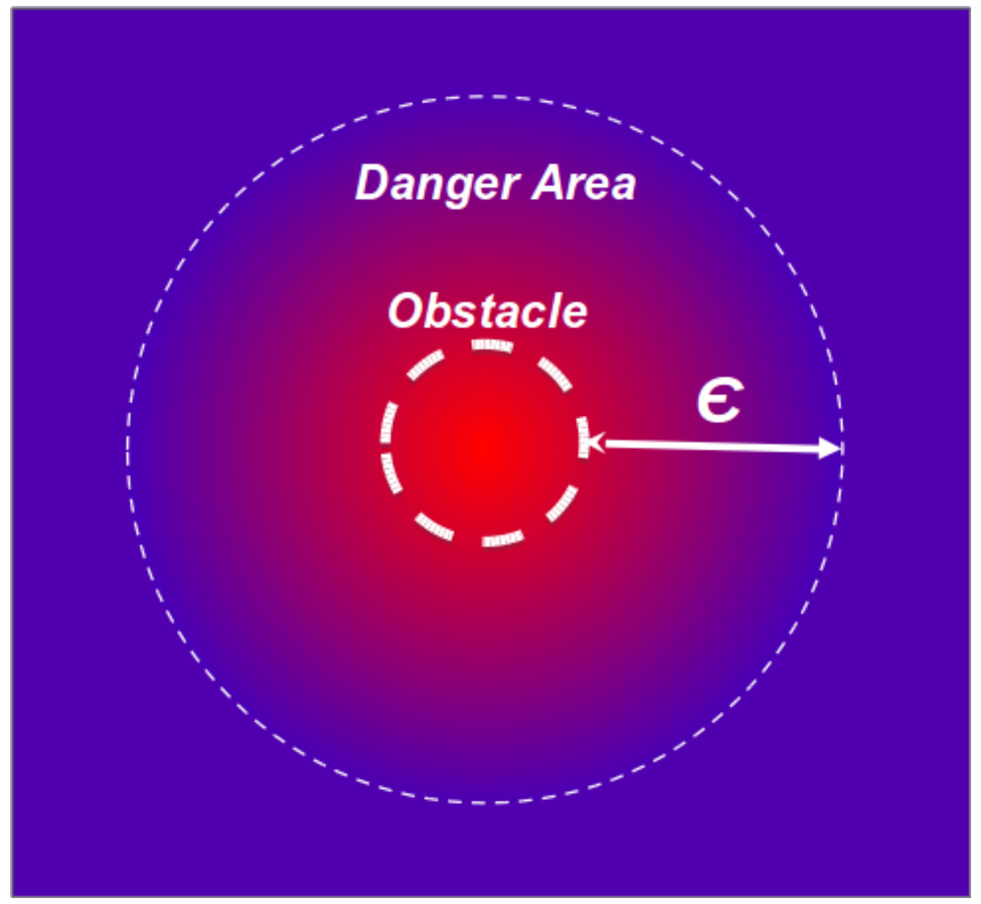
Skill Learning in Cluttered Environments



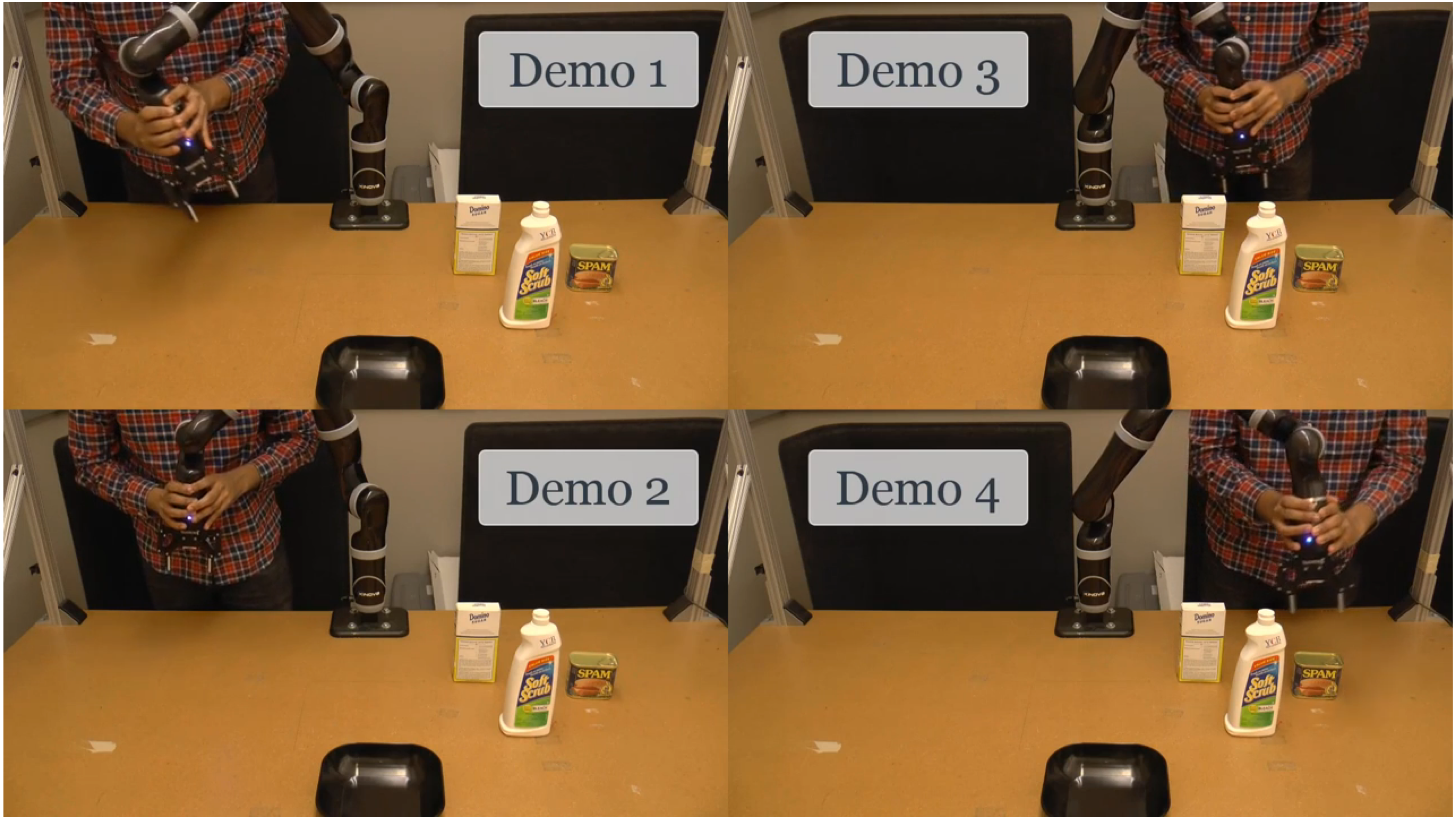
- **Remove external influences** (clutter) from the demonstrations

- Parts of demonstrations closer to obstacles are more likely to deviate from the desired skill constraints

- We use weighted linear regression

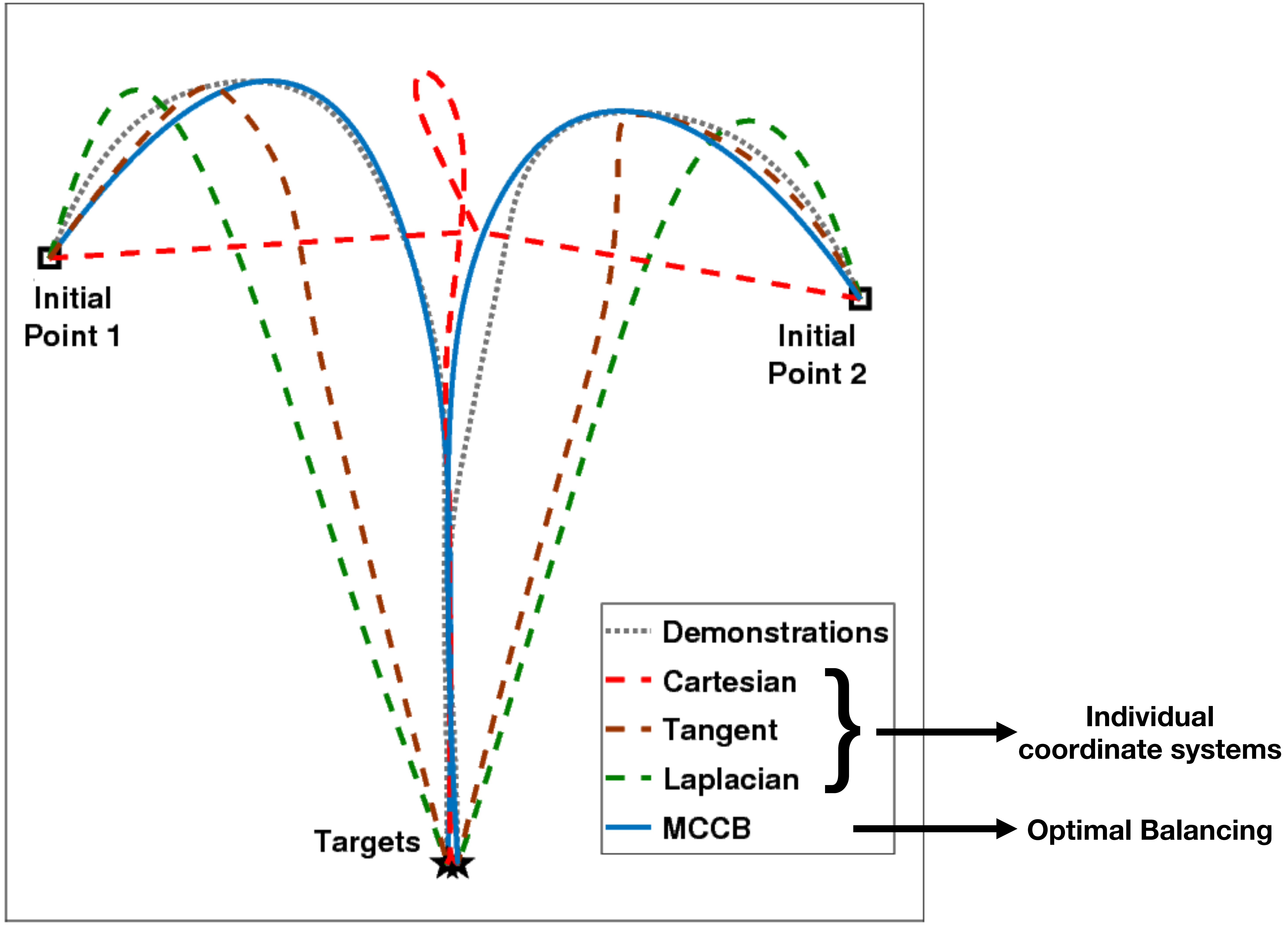


[Rana et al., IROS, 2018]

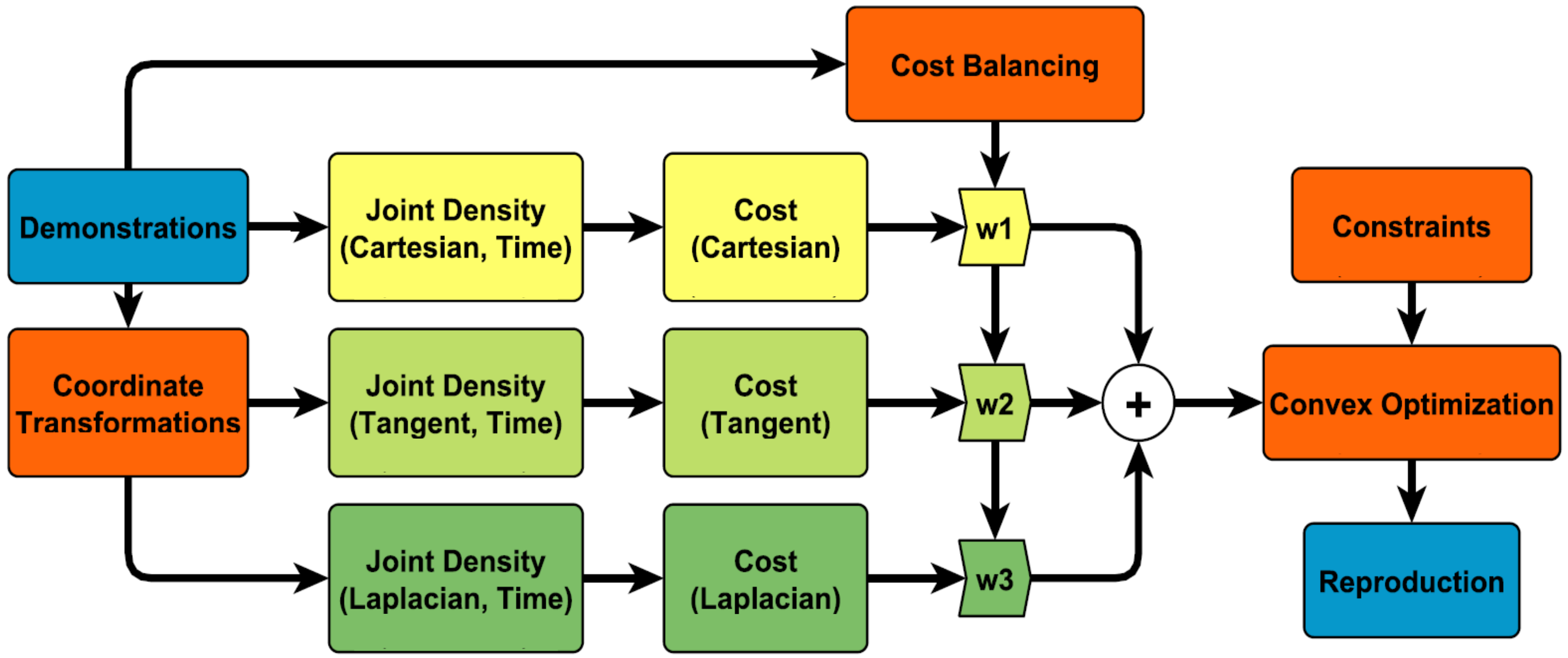


[Rana et al., IROS, 2018]

Automated Multi-Coordinate Cost Balancing

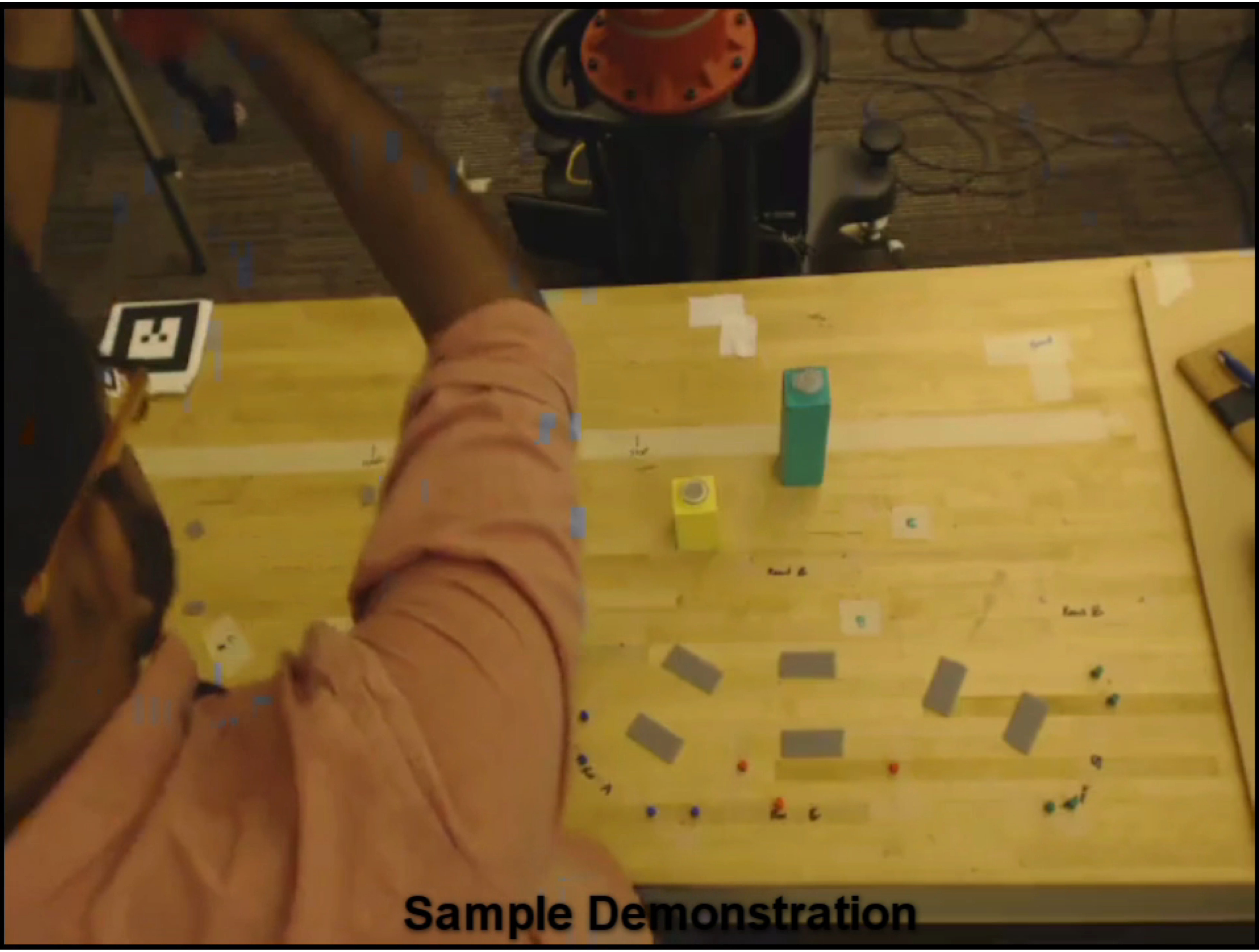


Automated Multi-Coordinate Cost Balancing



MCCB is a **task-independent** learning framework that

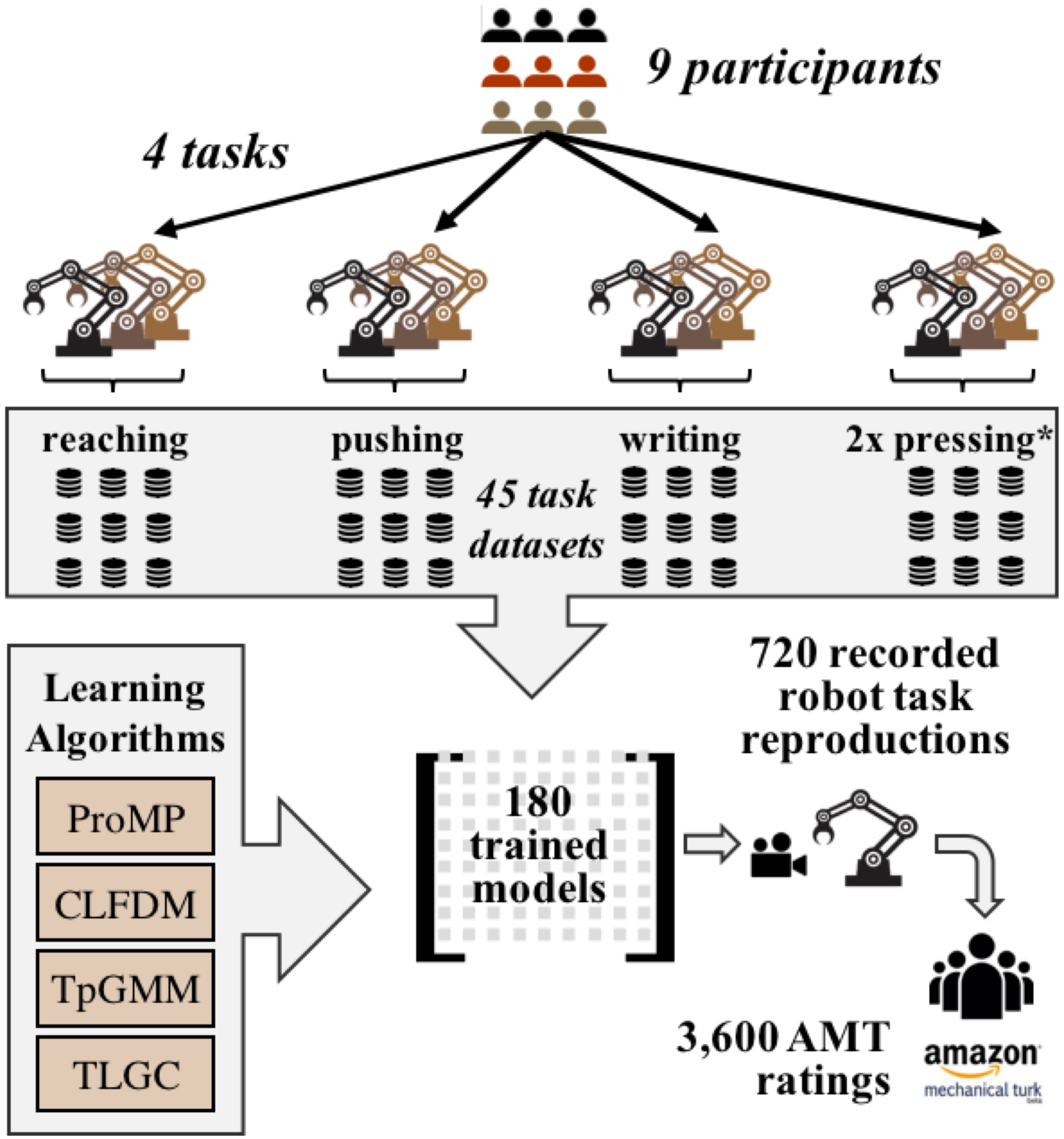
- Encode joint density with time ***simultaneously*** in multiple differential coordinates
- Defines a **blended cost function** that incentivizes *conformance to the norm* in each coordinate system while considering *expected variance*
- Learns *optimal weights* directly from the demonstrations to **balance the relative influence** of each coordinate system



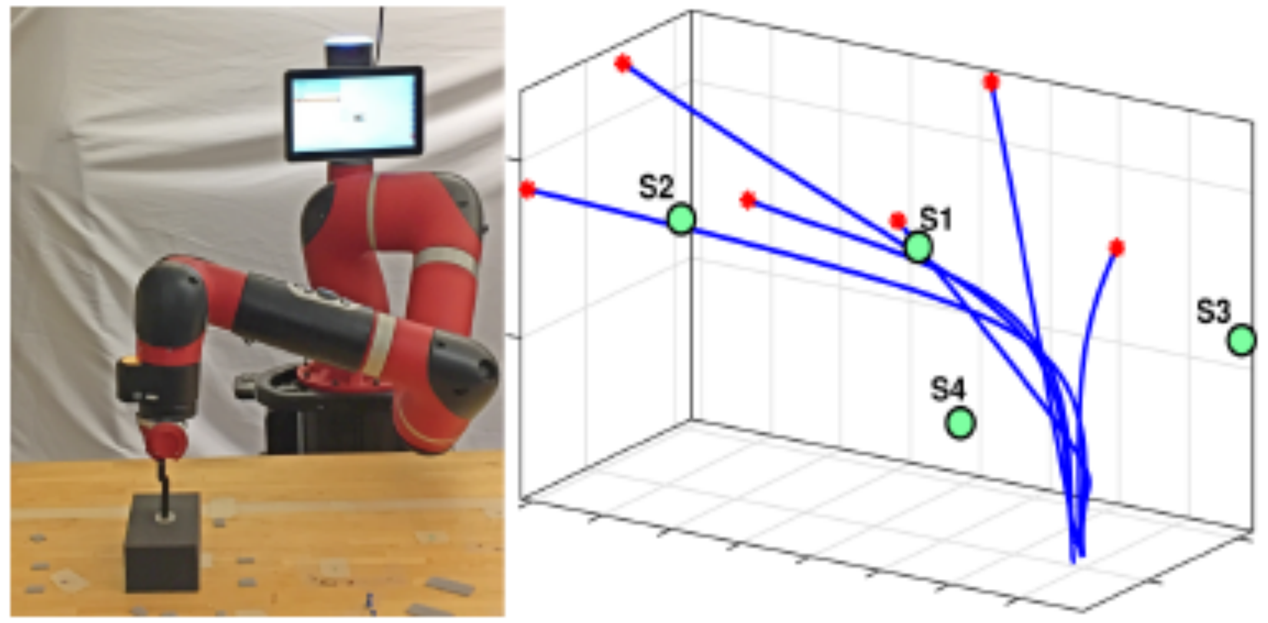
Sample Demonstration

End User Evaluation of LfD Methods

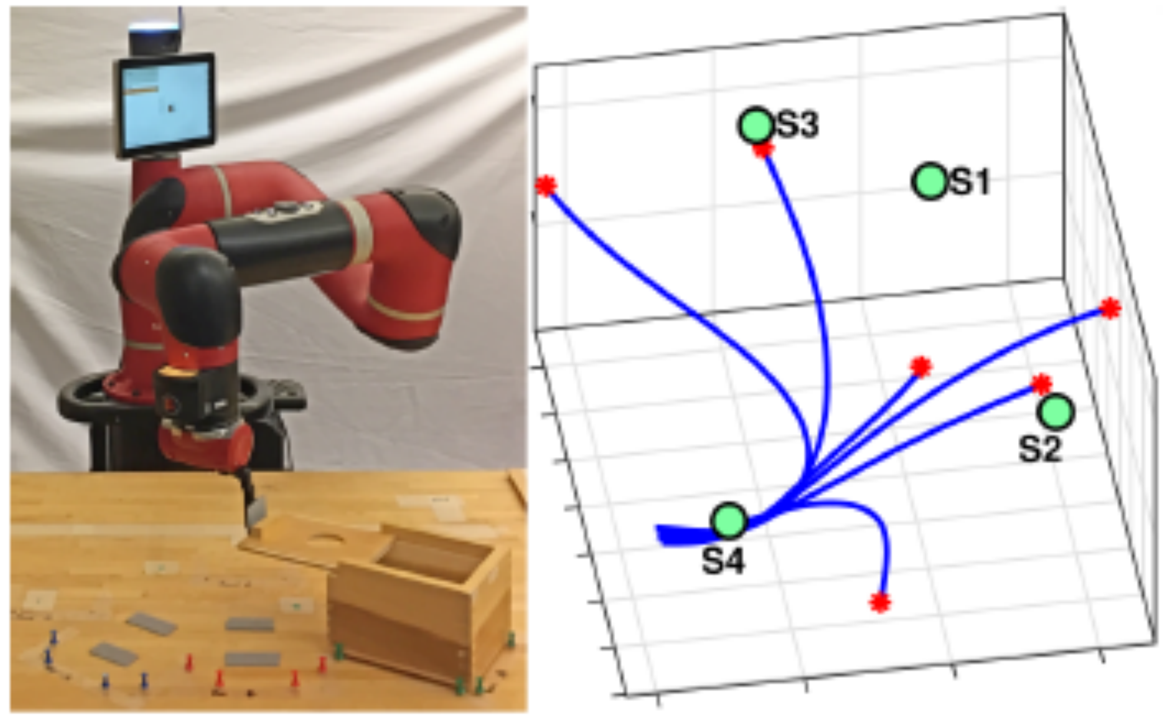
(probabilistic inference)
[Paraschos et al., NIPS, 2013]
(dynamical system)
[Khansari-Zadeh et al., RAS, 2014]
(statistical)
[Ahmadzadeh et al., RSS, 2017]
(geometric)
[Calinon et al., ISR, 2016]



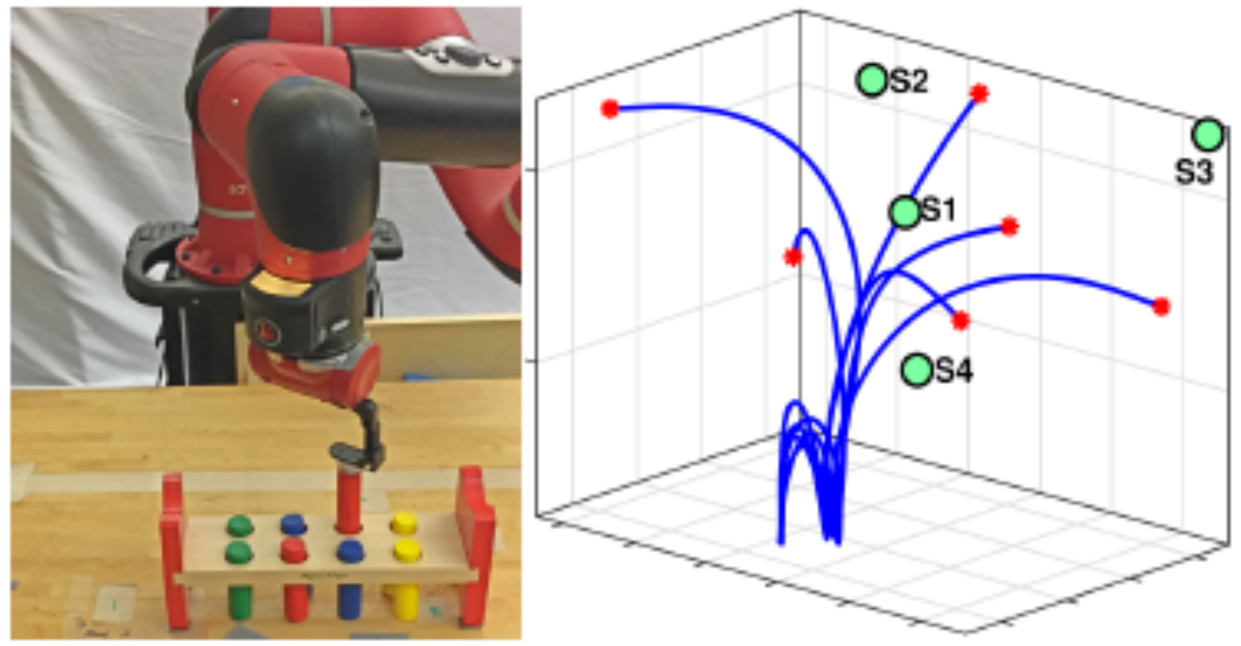
Tasks and Generalization Scenarios



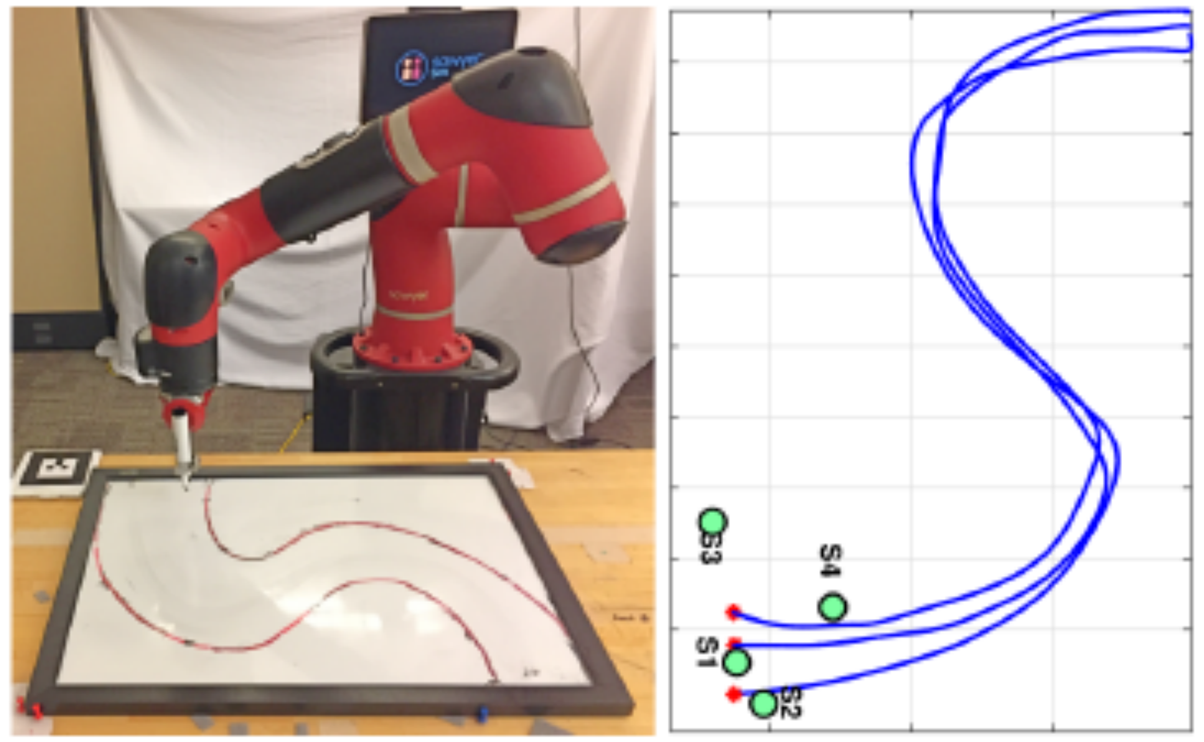
Reaching



Pushing



Pressing



Writing


User Ratings
CLFDM
ProMP
TLGC
TpGMM
Incomplete
Inefficient
Unsafe
Reaching
Pushing
Writing
Pressing (segmented)
Pressing
S1
S2
S3
S4

User Ratings: Takeaways

- Tasks with **constrained direction of the motion** (e.g., writing): **TLGC** (geometric)
- Tasks with **positional constraints** (e.g. reaching): **ProMP** (statistical) and **TpGMM** (probabilistic)
- Generalization to **starting locations closer to the target: CLF-DM** and **TLGC** (time-invariant)
- No one algorithm consistently yielded successful executions across all generalization scenarios for any given task
- **Demonstrator Experience level** positively correlated with performance

Publications

- A. Saran, B. Lakic, S. Majumdar, J. Hess, and S. Niekum. "Viewpoint Selection for Visual Failure Detection." in IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), September 2017.

- A. Saran, S. Majumdar, E.S. Short, A.L. Thomaz, and S. Niekum. "Human Gaze Following for Human-Robot Interaction." in IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), October 2018.

- Y. Cui and S. Niekum. "Active Reward Learning from Critiques." in IEEE International Conference on Robotics and Automation (ICRA), May 2018.

- R.A. Gutierrez, V. Chu, A.L. Thomaz, and S. Niekum. "Incremental Task Modification via Corrective Demonstrations." in IEEE International Conference on Robotics and Automation (ICRA), May 2018.

- Rana, Muhammad Asif; Mukadam, Mustafa; Ahmadzadeh, Reza S; Chernova, Sonia; Boots, Byron; “, Towards Robust Skill Generalization: Unifying Learning from Demonstration and Motion Planning," Conference on Robot Learning, 2017.

- Rana, Muhammad Asif; Mukadam, Mustafa; Ahmadzadeh, Reza S; Chernova, Sonia; Boots, Byron; "Learning Generalizable Robot Skills from Demonstrations in Cluttered Environments," IEEE/RSJ International Conference on Intelligent Robots and Systems, 2018.

- Ravichandar, Harish; Ahmadzadeh, Reza S; Rana, Muhammad Asif; Chernova, Sonia; "Skill Acquisition via Automated Multi-Coordinate Cost Balancing," IEEE International Conference on Robotics and Automation, 2019 – *under review*.

- Rana, Muhammad Asif; Ahmadzadeh, Reza S; Chernova, Sonia; "Benchmarking Skill Learning from Demonstration: Impact of User Experience, Task Complexity, and Start Configuration on Performance," IEEE International Conference on Robotics and Automation, 2019 – *under review*.

Goals for Year 3

- Correct errors or ask for help
- Request remote critiques or additional demonstrations
- Transfer learning
- Refine reward functions
- Human studies of the full system

Thank You!