Hierarchical Learning Approaches for Long Horizon Robotics Tasks

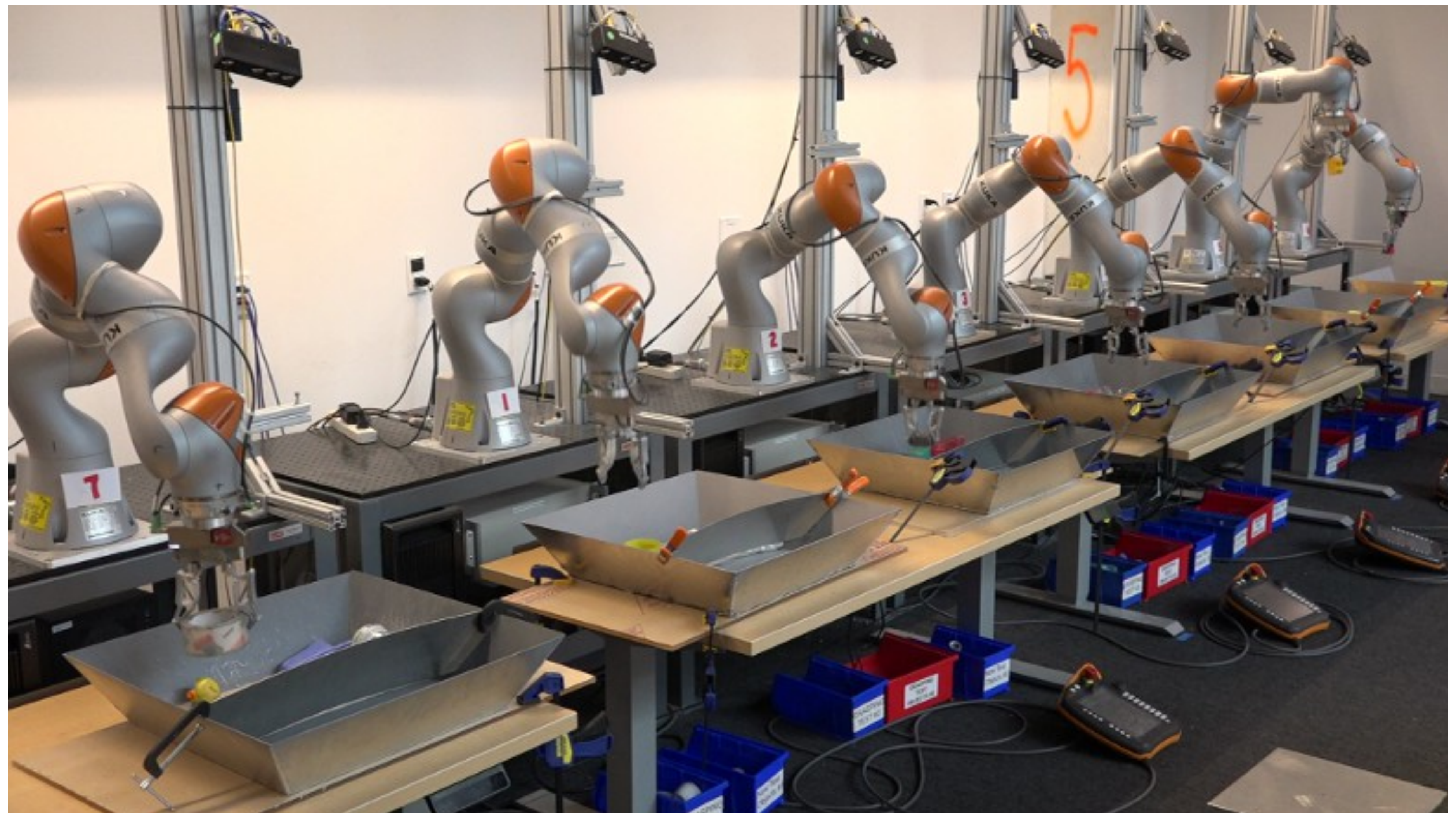
Fei Xia, Research Scientist, Google

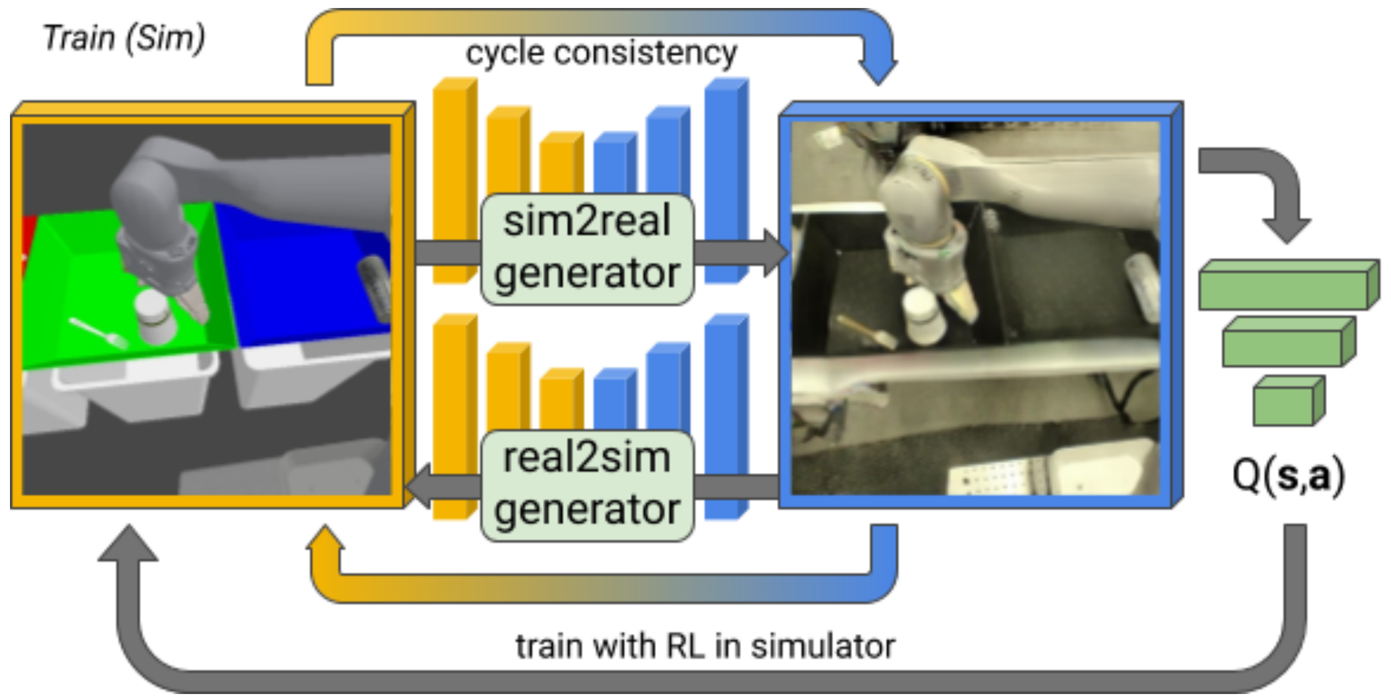
Who are we?

- Robotics at Google
- g.co/robotics

Goals:

- Improve robotics via machine learning, and improve machine learning via robotics
- Enable learning at scale on real and simulated robotic systems

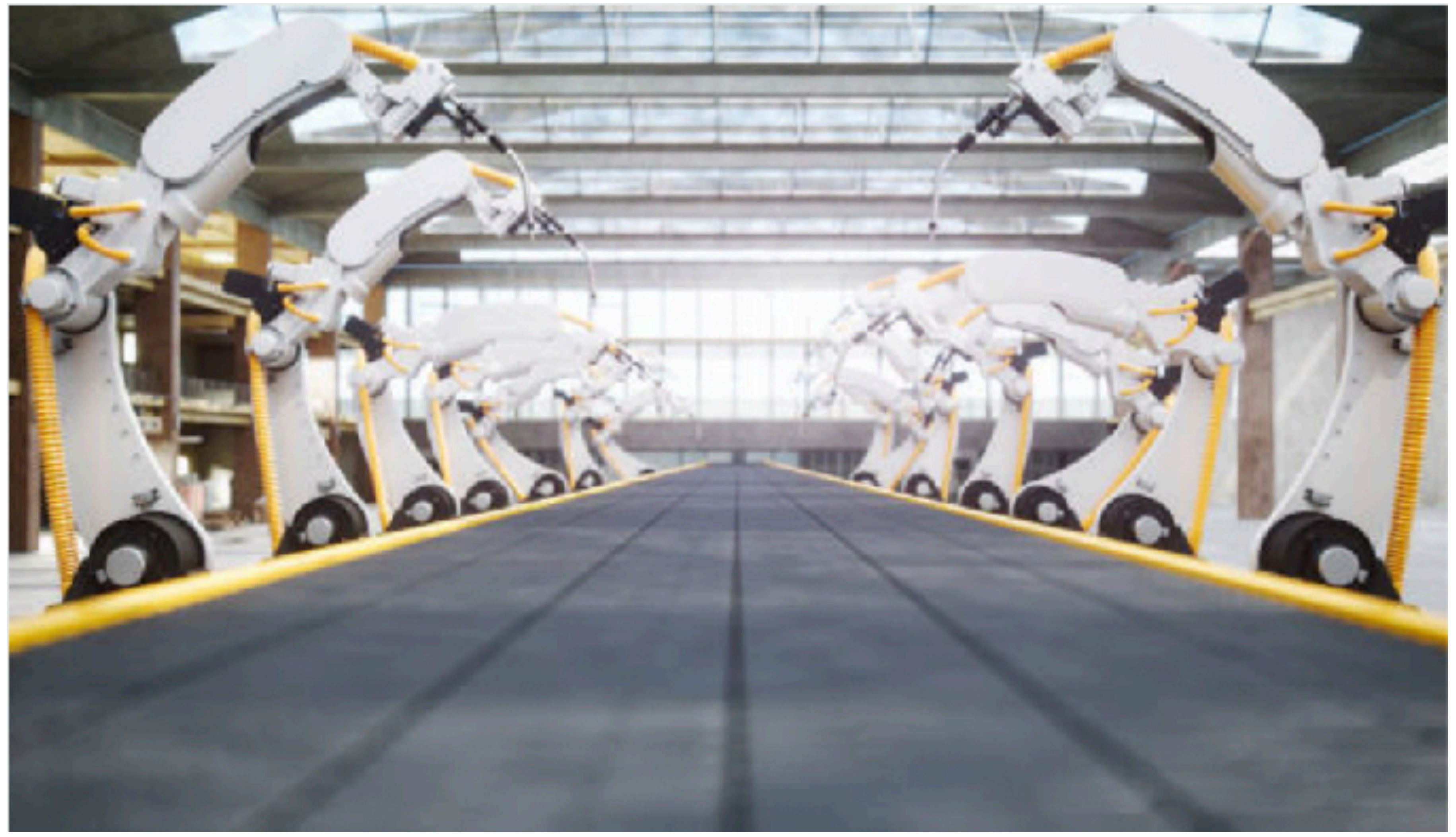




Robot @ Home

Image source: iStock photos, Standard license

Towards complex and unstructured environments

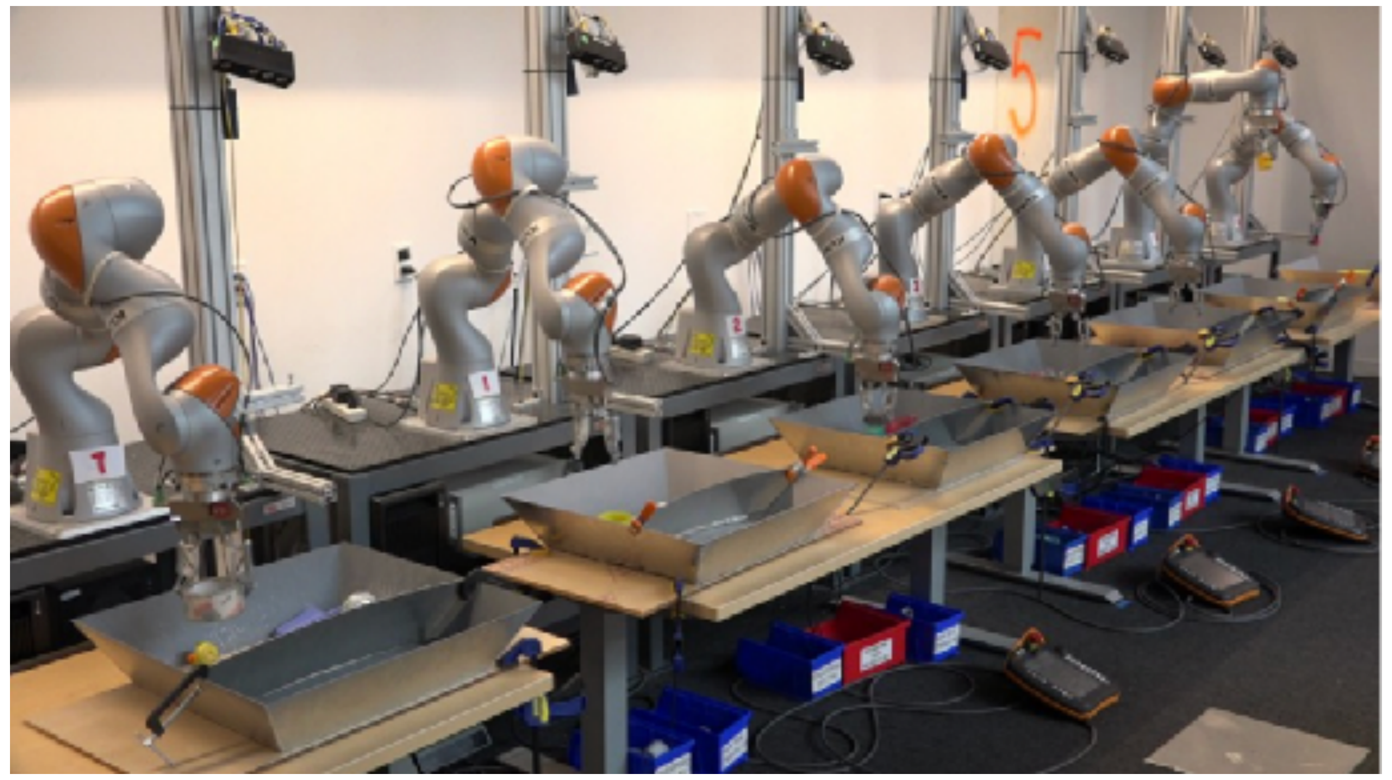


"Classical" robotics

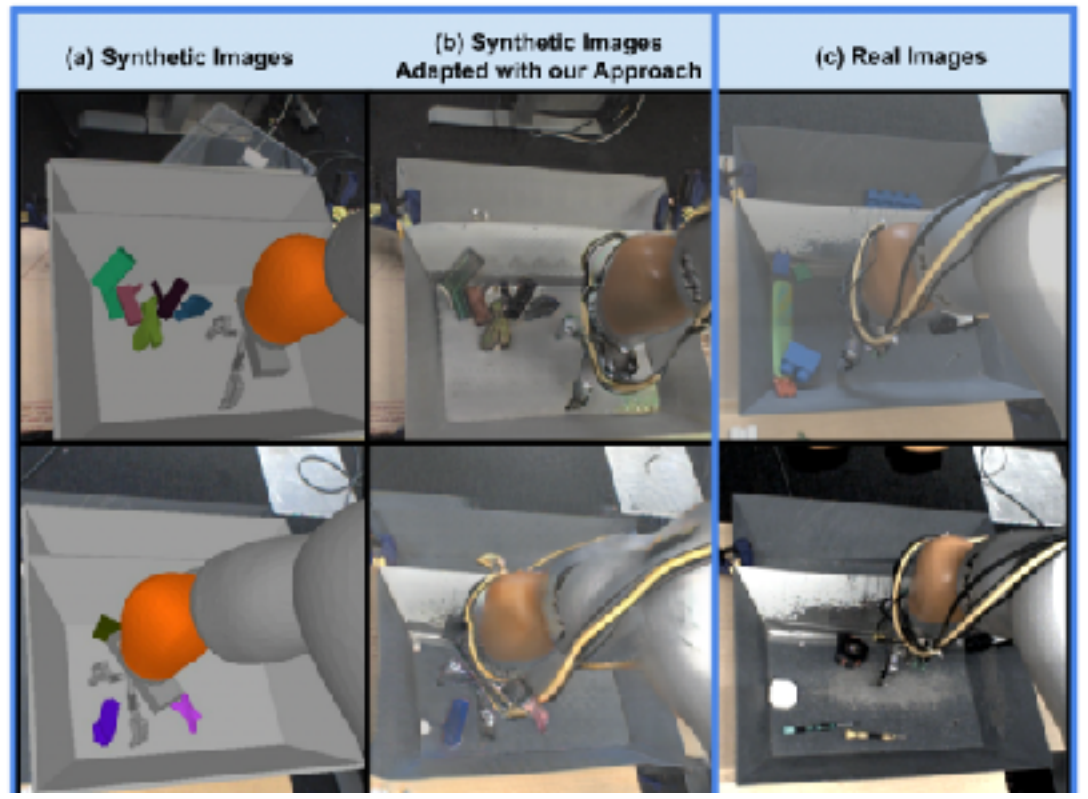
?

Learning based robotics

Levine et al. 2016.



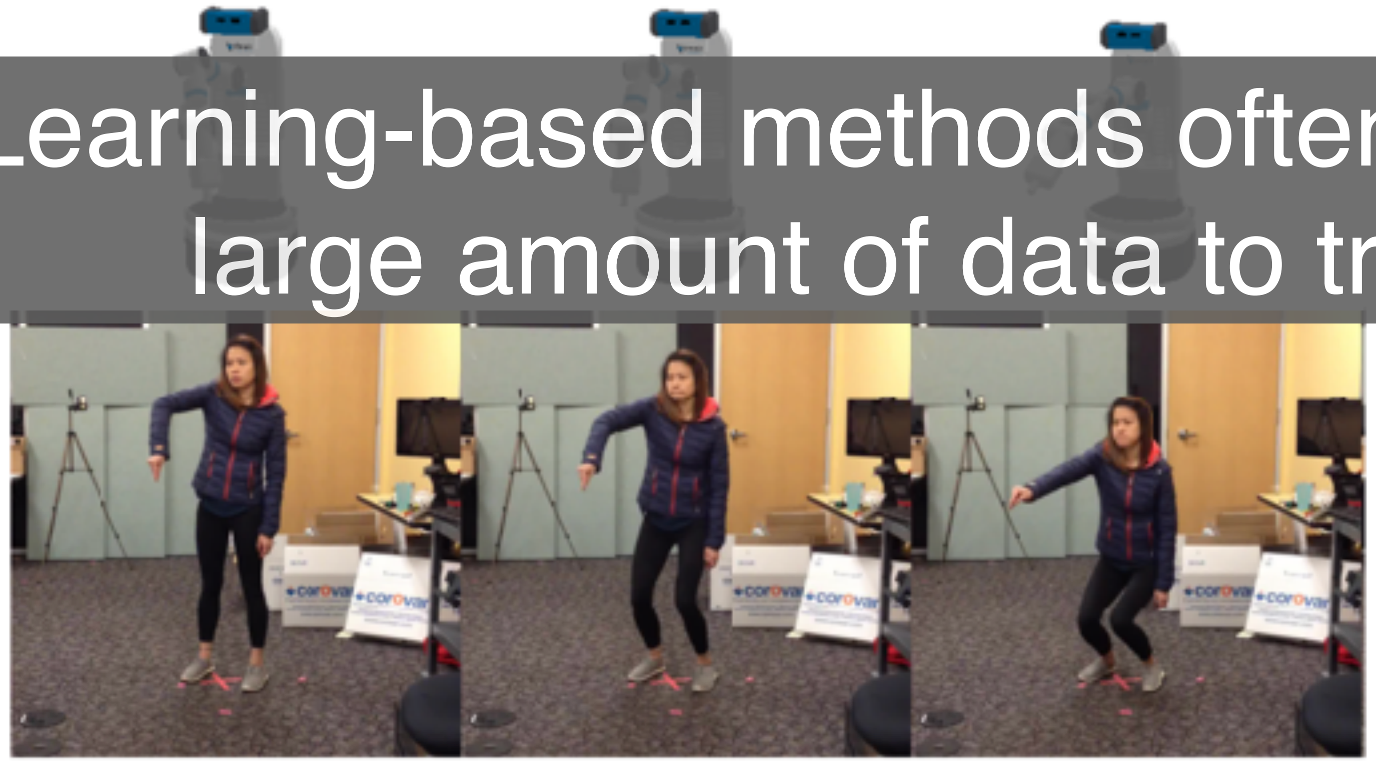
Kalashnikov et al. 2018.



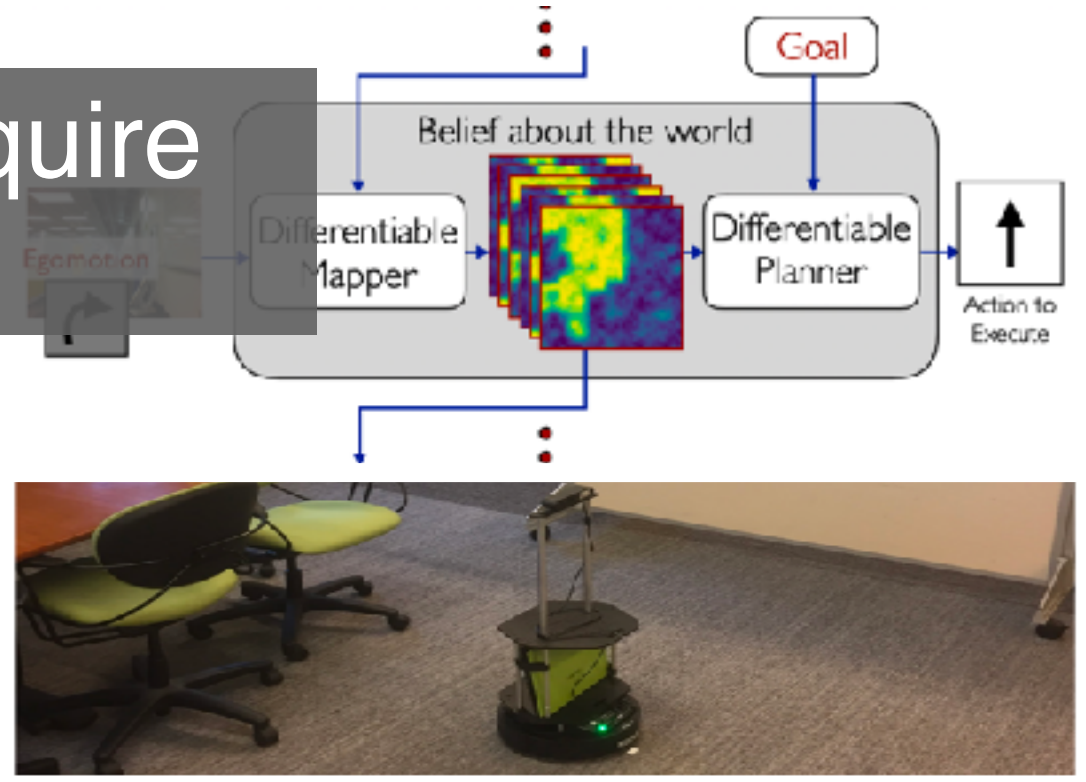
Bousmalis et al. 2018.

Gupta et al. 2018.

Learning-based methods often require large amount of data to train.



Sermanet et al. 2018.

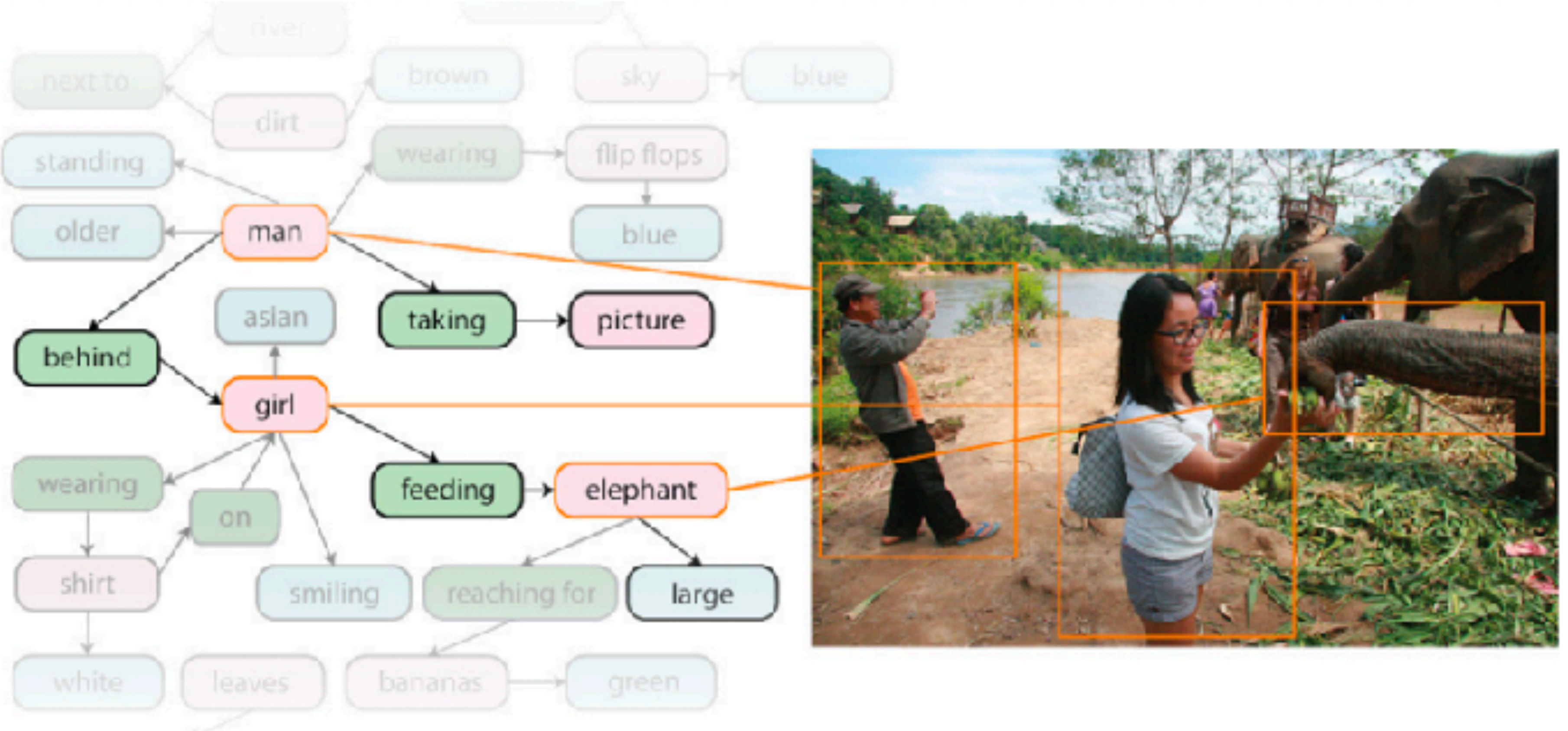


Gupta et al. 2019.

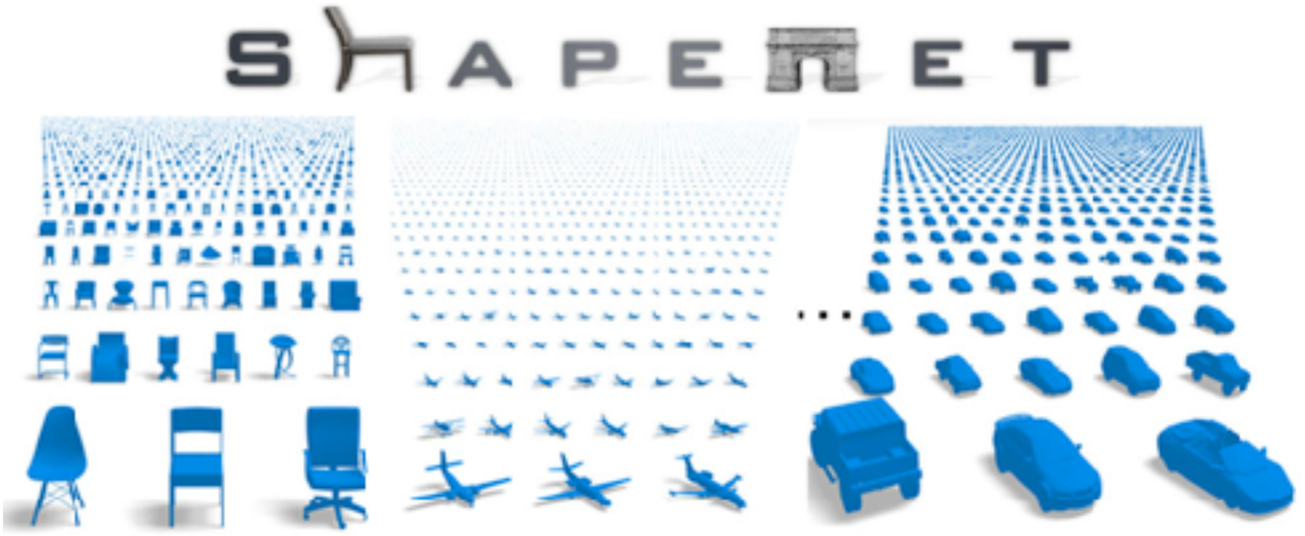
[Rajeswaran et al. 2018.] [Zhu et al 2018] [Kurenkov et al 2019] [Mandlekar et al 2020] [Andrychowicz et al 2019] [Mahler et al 2017] [Mousavian et al 2019] [Rao et al 2020] [James et al 2019] [Tobin et al 2017] [Peng et al 2018] [Kaspar et al 2020] [Chebotar et al 2019] [Choi et al 2021] [Puig et al 2018] [Erickson et al 2020] [Tolani et al 2021] [Wu et al 2019] [Peng et al 2018] [Yang et al 2020]

Datasets for Computer Vision

ImageNet, Deng et al 2009.



Visual Genome, Krishna et al 2017.



ShapeNet, Chang et al 2015.

MS COCO, Lin et al 2014.

Pascal VOC, Everingham et al 2012.

OpenImage,
Krasin et al 2016.

[Xiao et al 2014] [Fei-Fei et al 2003] [Daimler AG et al 2016] [A. Krizhevsky et al 2009] [A. Krizhevsky et al 2009] [A Geiger et al. 2012] [A. Janoch et al. 2014] [Mo et al 2019]

From Perception to Interaction

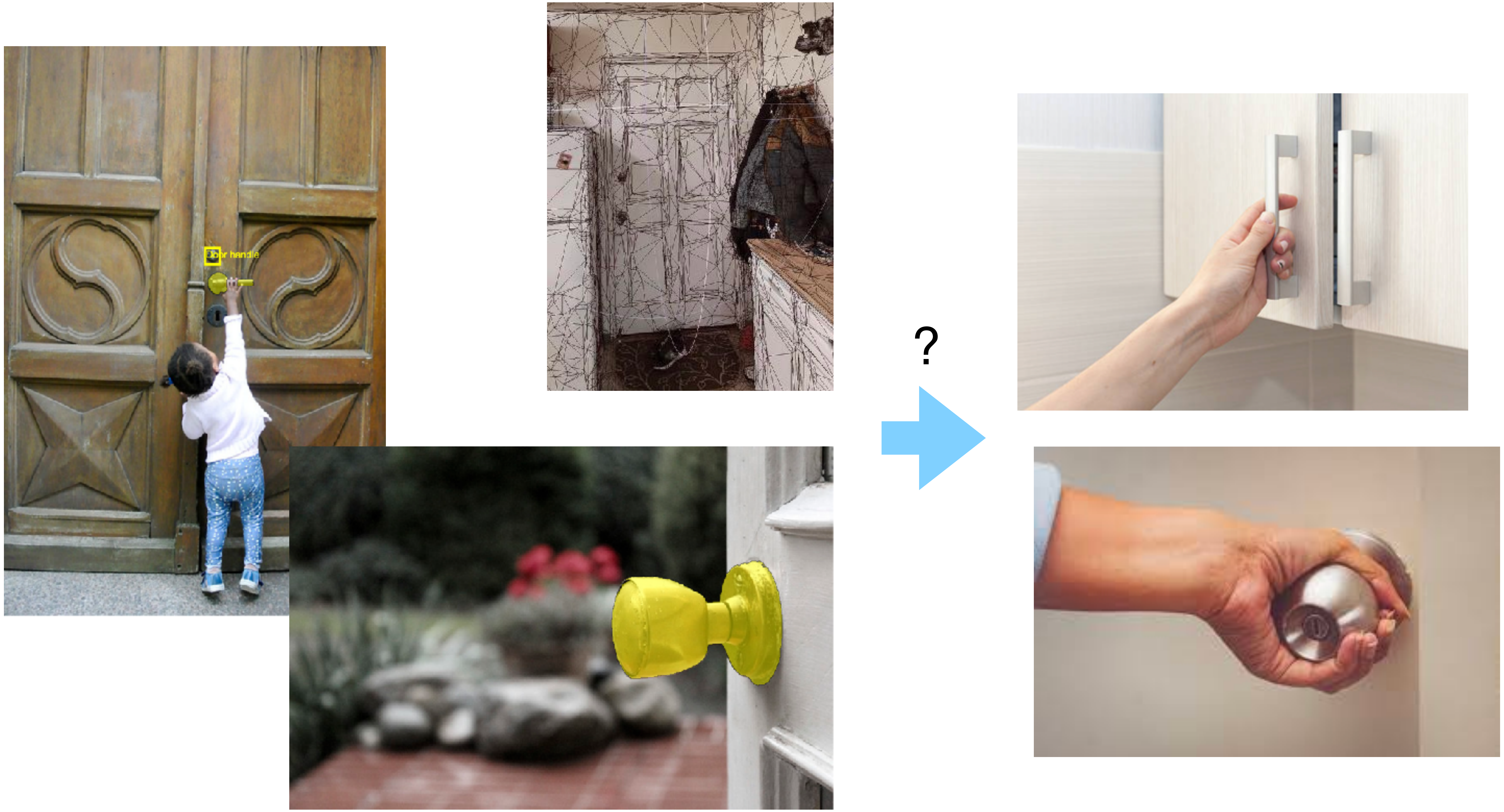


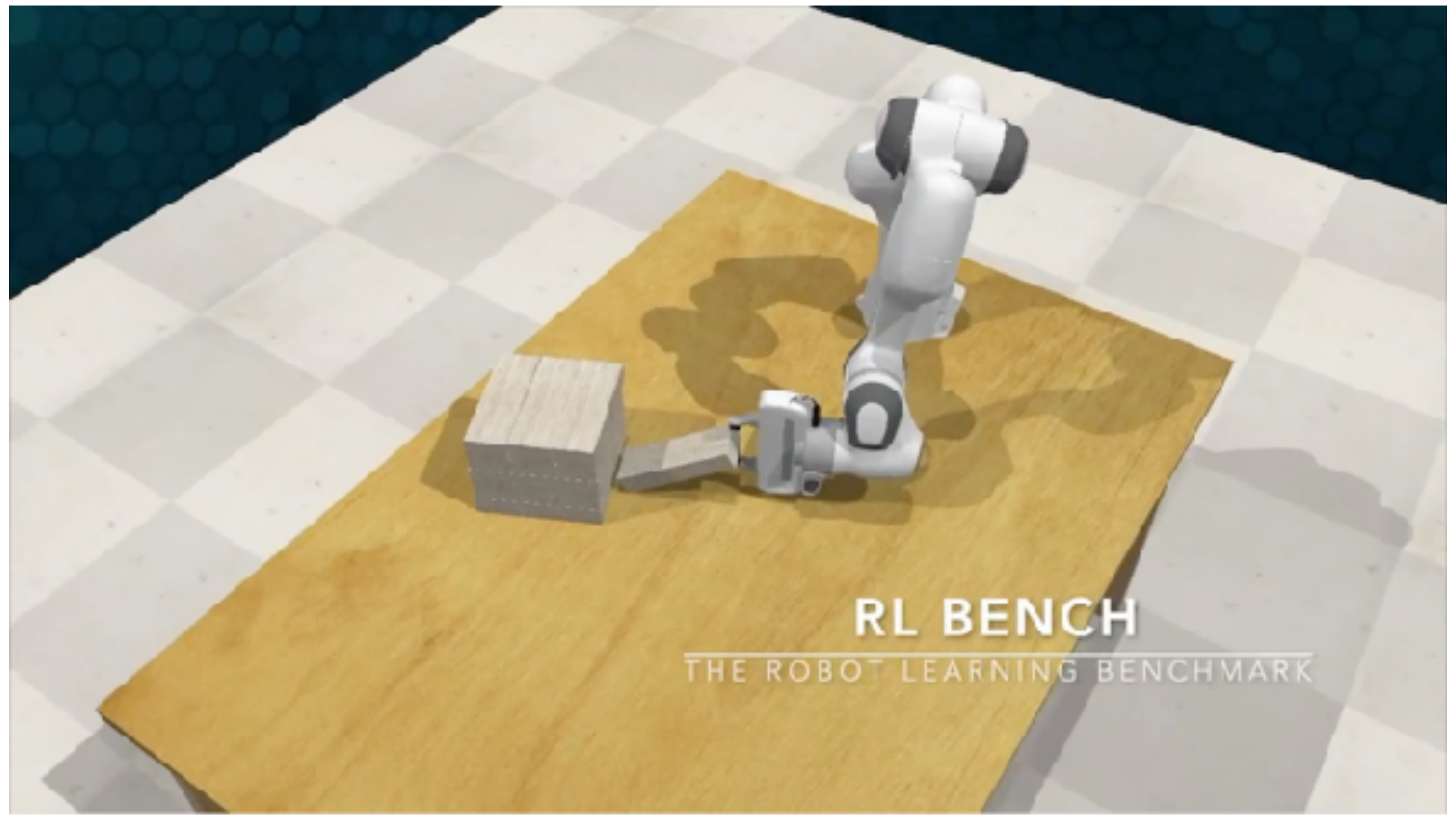
Image source: iStock photos, Standard license, Open Image dataset, iGibson dataset

Learning from interactions with the environment

Human learn from interacting with the environment.

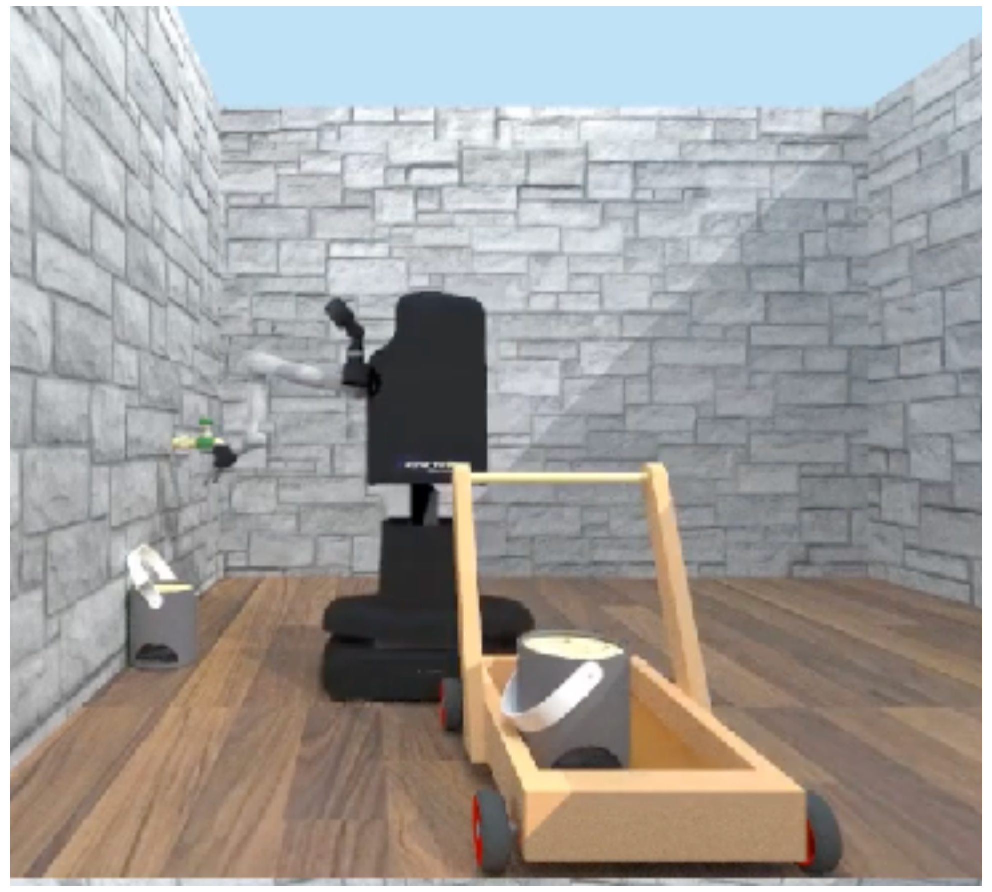
Image source: iStock photos, Standard license [Gopnik et al 1999] [Goldstein et al 1994]

Learning from simulation



RLBench, James et al 2020.

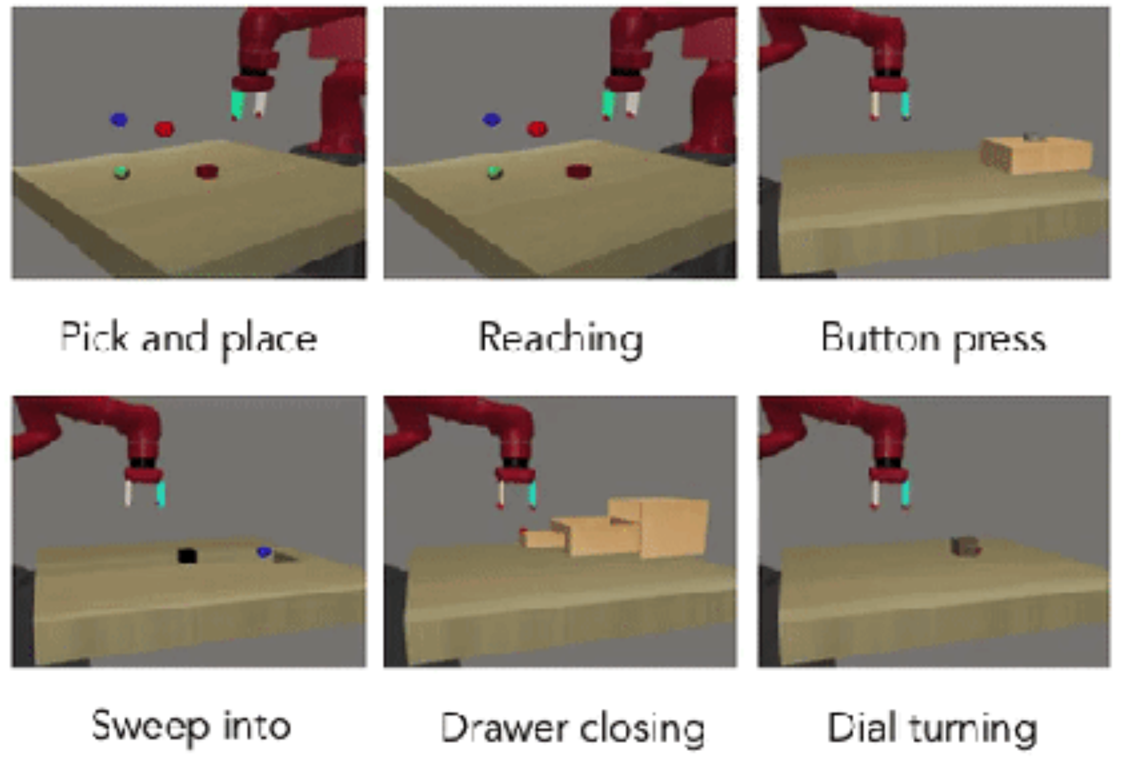
AI2Thor, Kolve et al 2017.



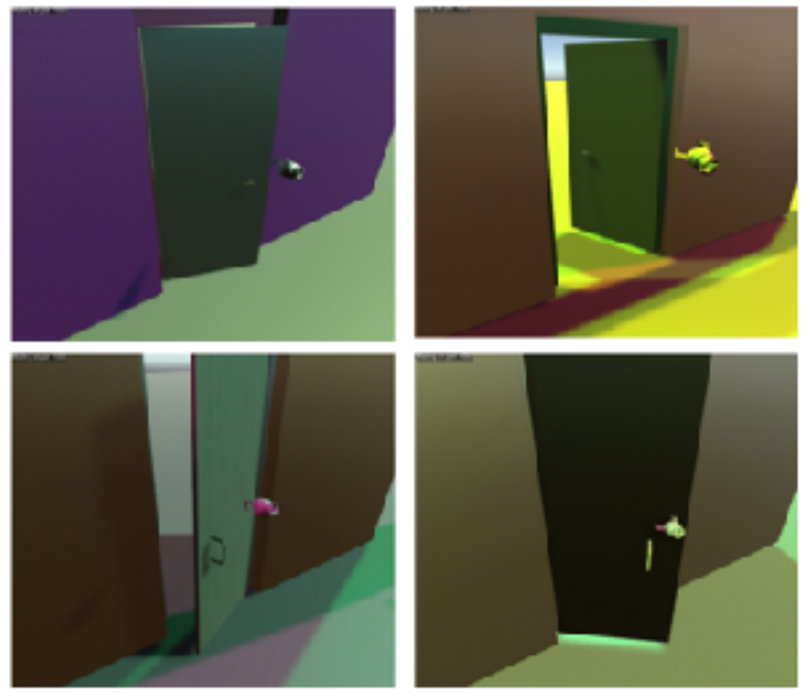
SAPIEN, Xiang et al 2020.

Ikea assembly, Lee et al 2019.

TDW Gan et al 2020.



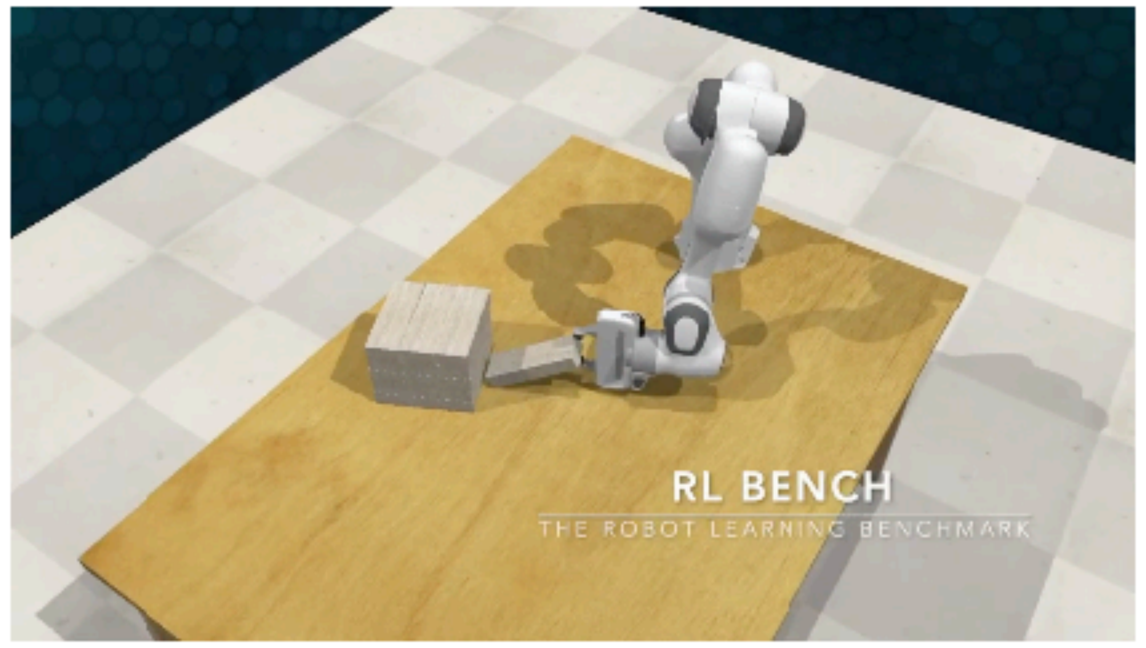
Meta World, Yu et al 2020.



DoorGym, Urakami et al 2019.

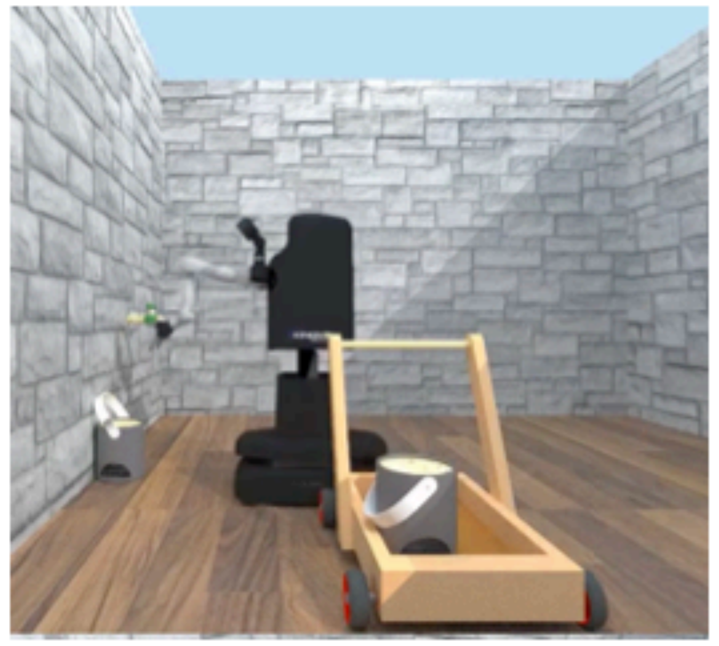
[Savva et al 2019] [Puig et al 2018] [Wu et al 2018] [Ehsani et al 2021] [Xiang et al 2020] [Gan et al 2020]

Learning from simulation



RLBench, James et al 2020.

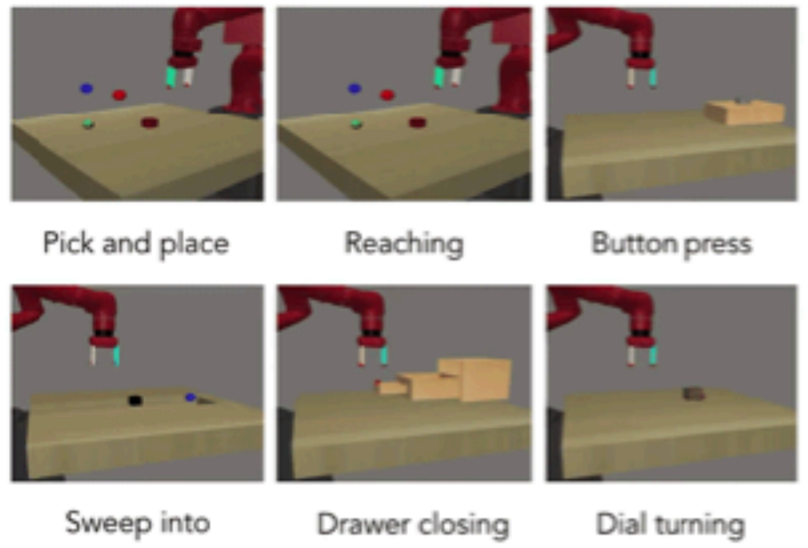
AI2Thor, Kolve et al 2017.



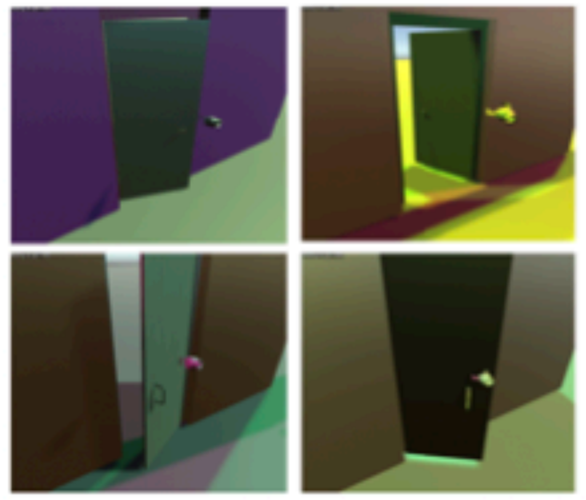
SAPIEN, Xiang et al 2020.

Ikea assembly, Lee et al 2019.

TDW Gan et al 2020.



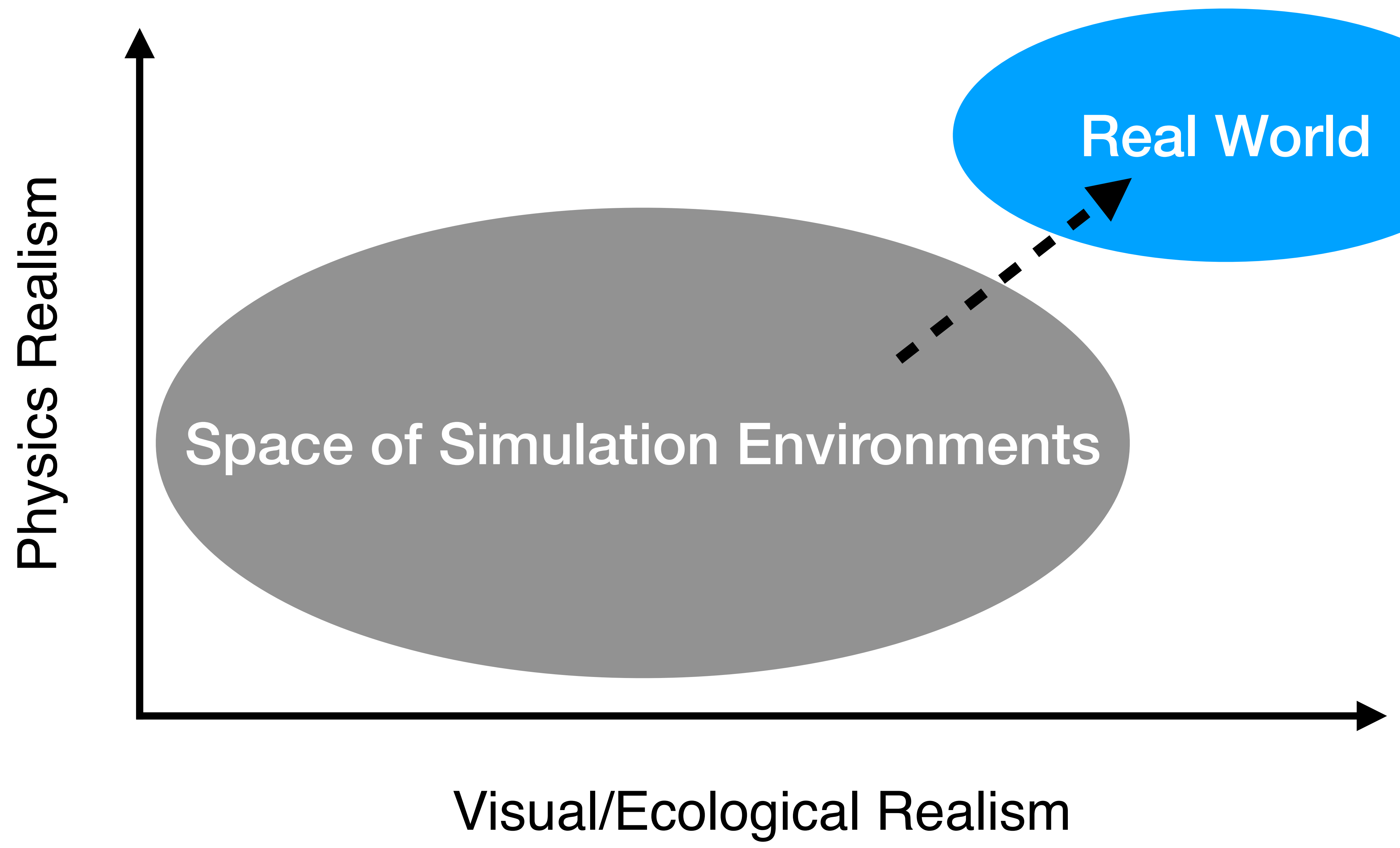
Meta World, Yu et al 2020.

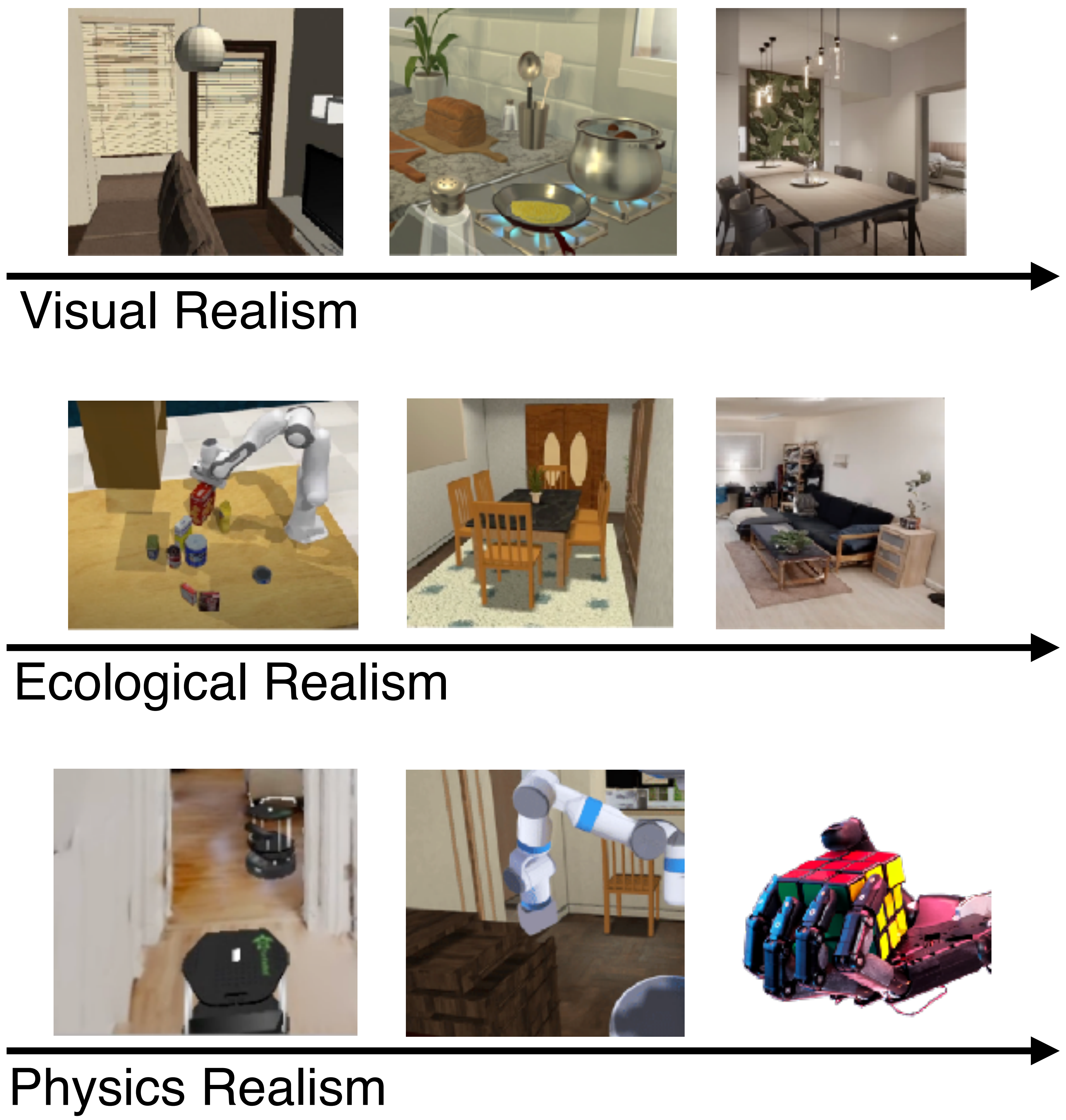


DoorGym, Urakami et al 2019.

Not based on real-world scenes
Small scale
Lacks sim2real transfer
…

Learning from simulation



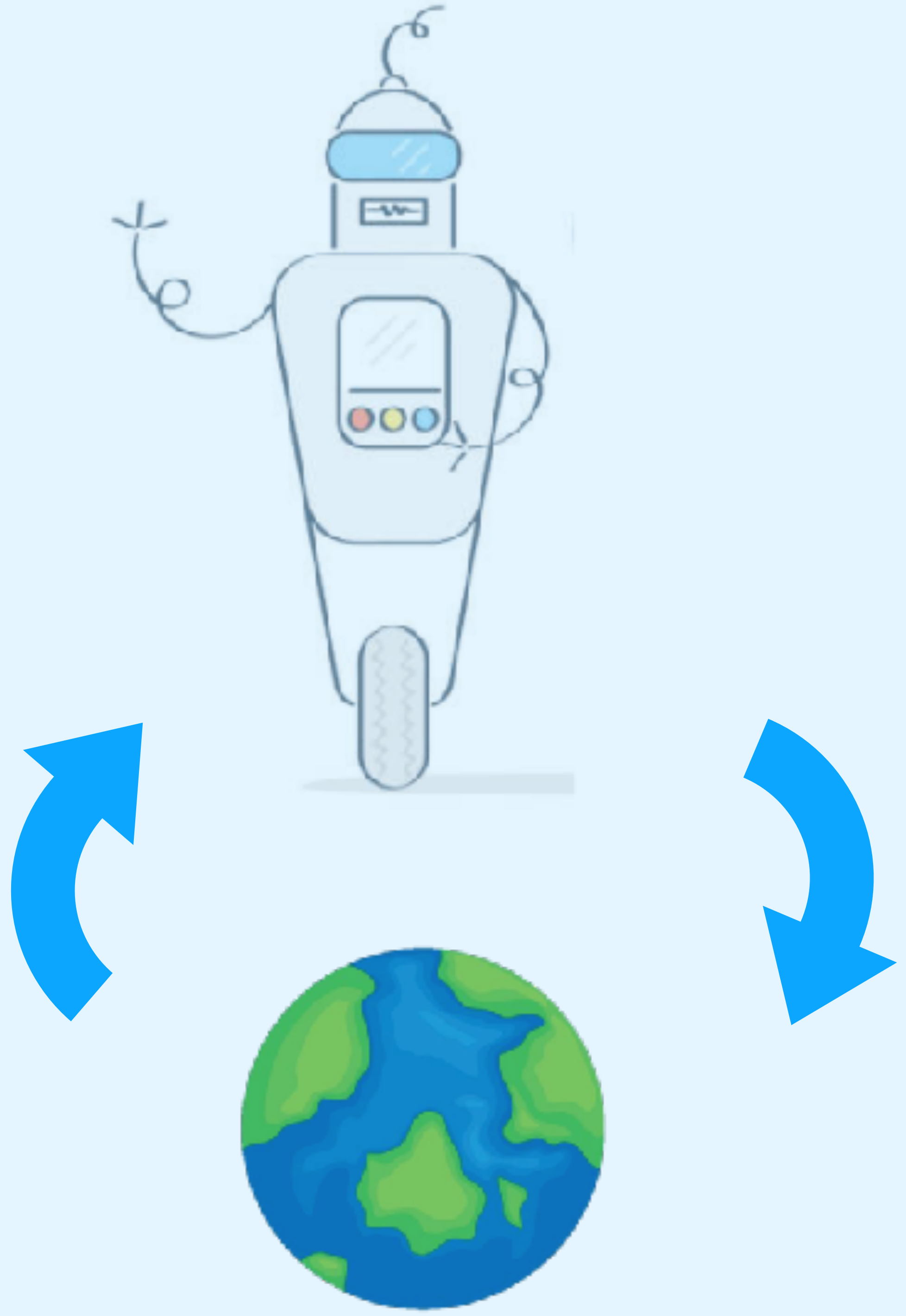


Goal: Create simulation environments that reflect **complexity** of the **real world**, and develop intelligent agents in those environments.

Large Scale Simulation for Embodied Perception and Robot Learning

Goal: Create simulation environments that reflect **complexity** of the **real world**, and develop intelligent agents in those environments.

Creating a digital playground that replicates the complexity of the real world



Learning algorithms for interactive and long-horizon tasks

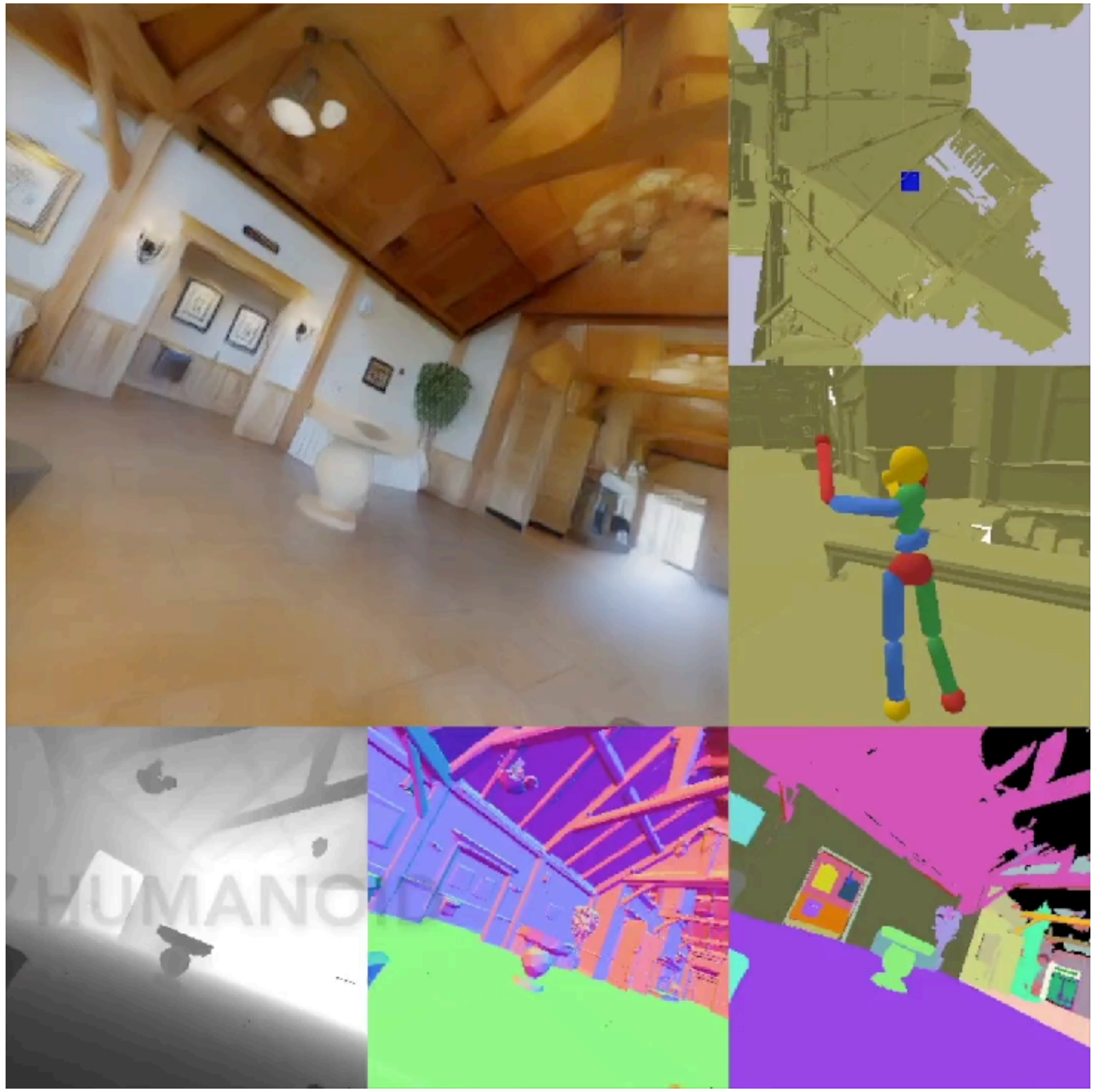
Image source: iStock photos, Standard license

GibsonEnv

Rendered RGB

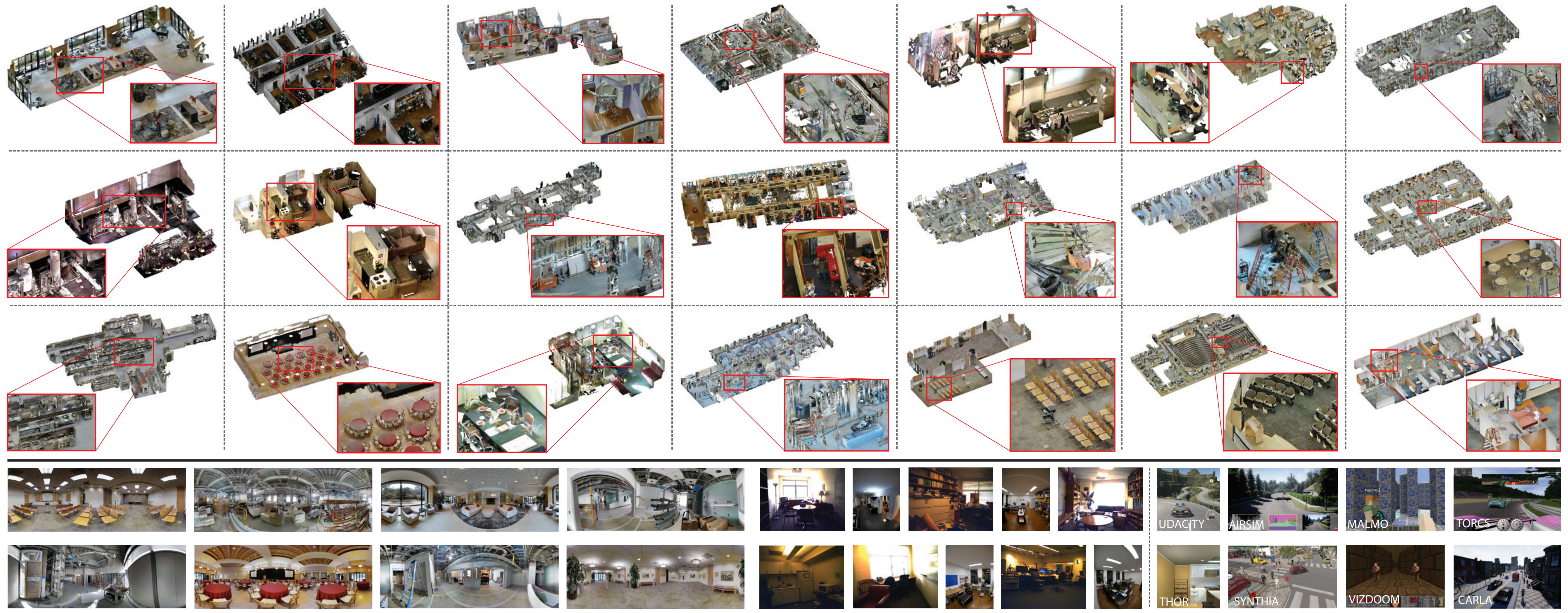
Physics Simulation

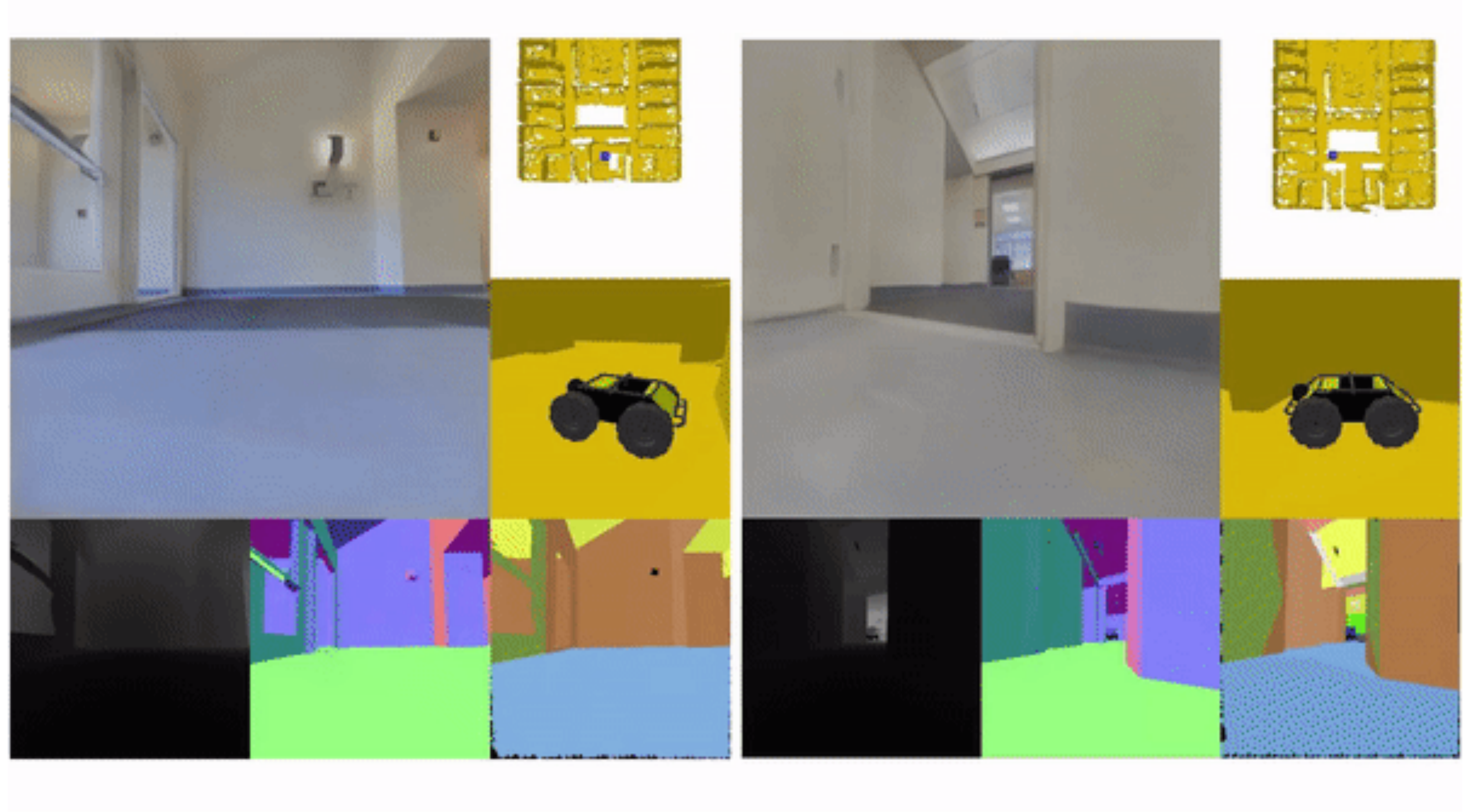
Rendered sensor signals

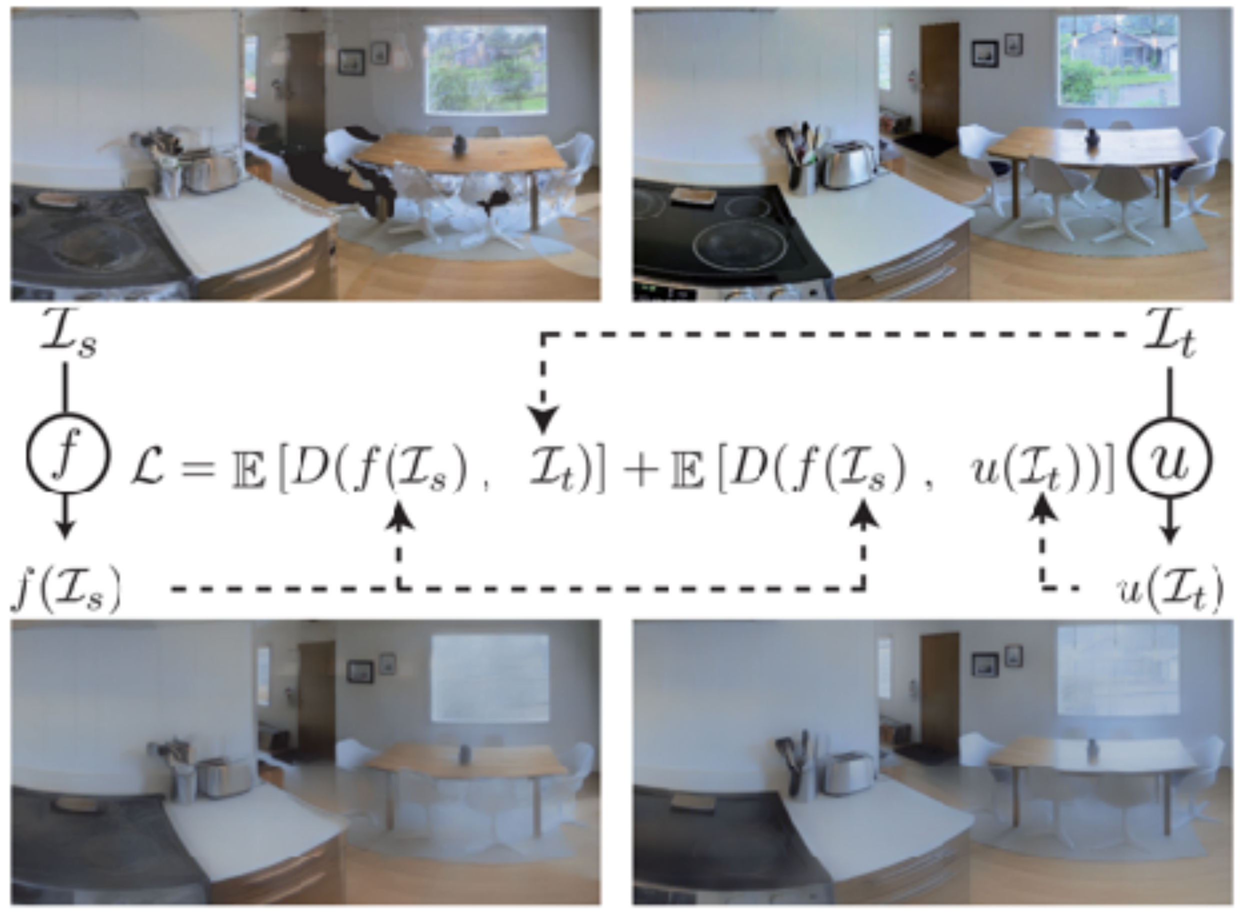


X*Z*H*SMS, GibsonEnv, CVPR 2018.

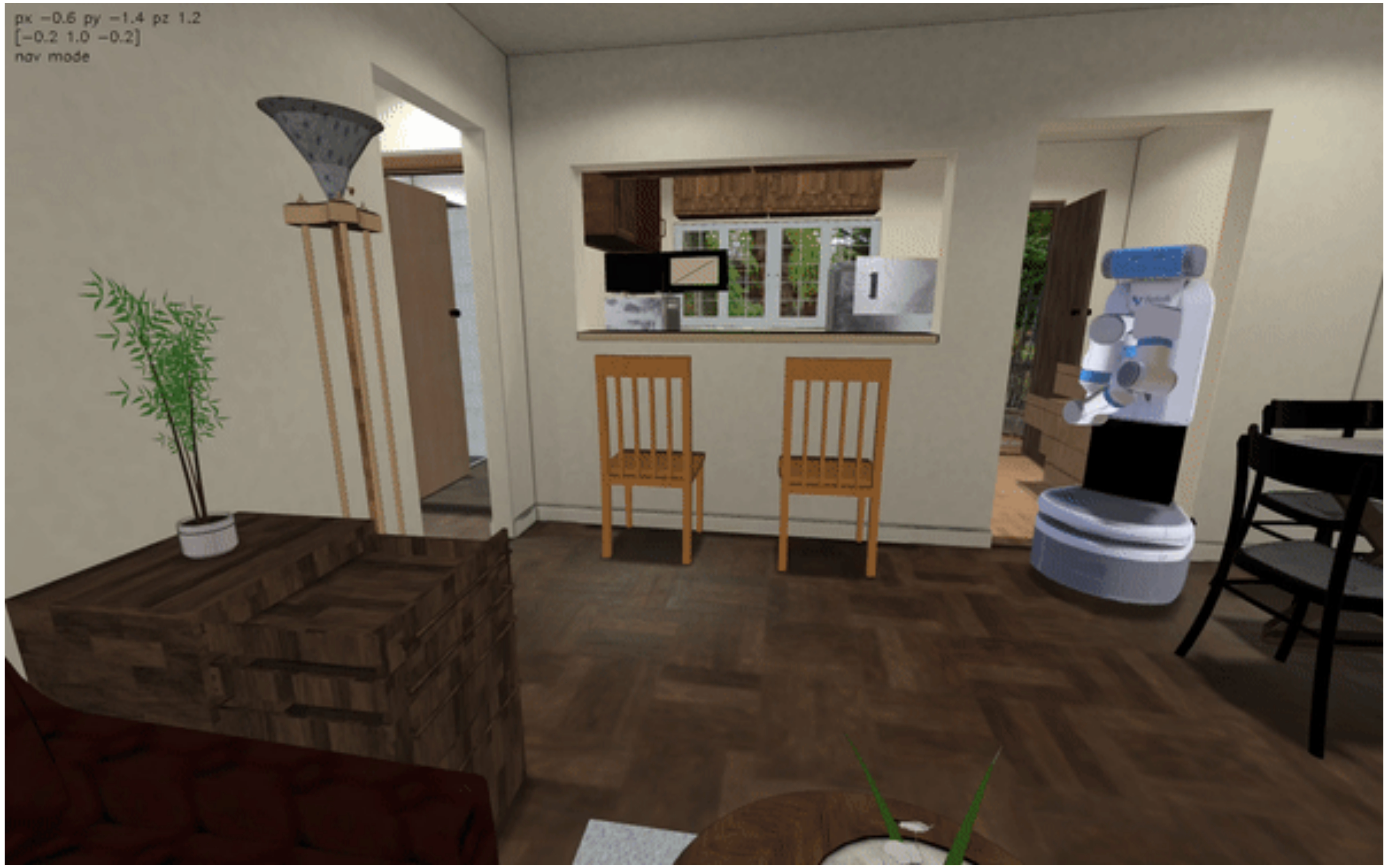
GibsonEnv - Overview



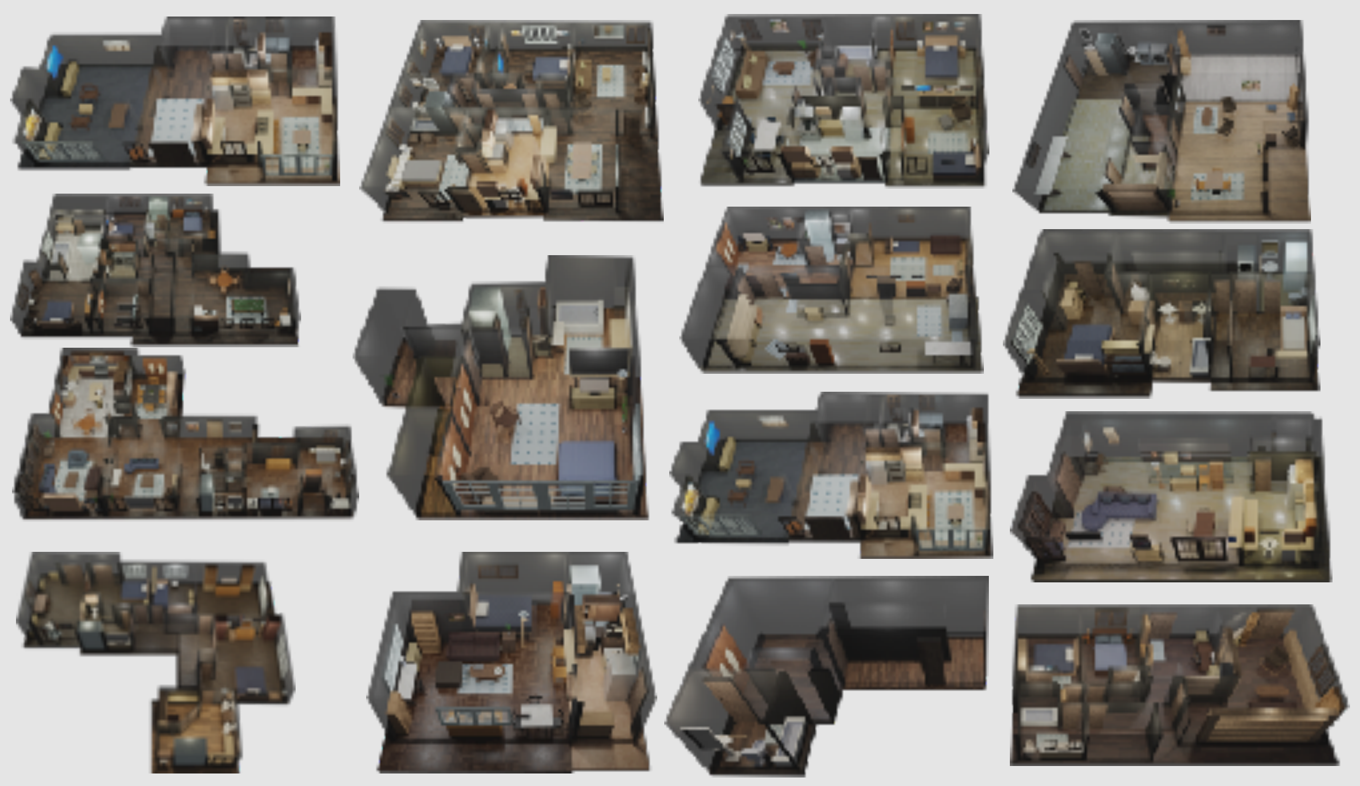




iGibson: A Simulation Environment for Interactive Tasks in Large Realistic Scenes



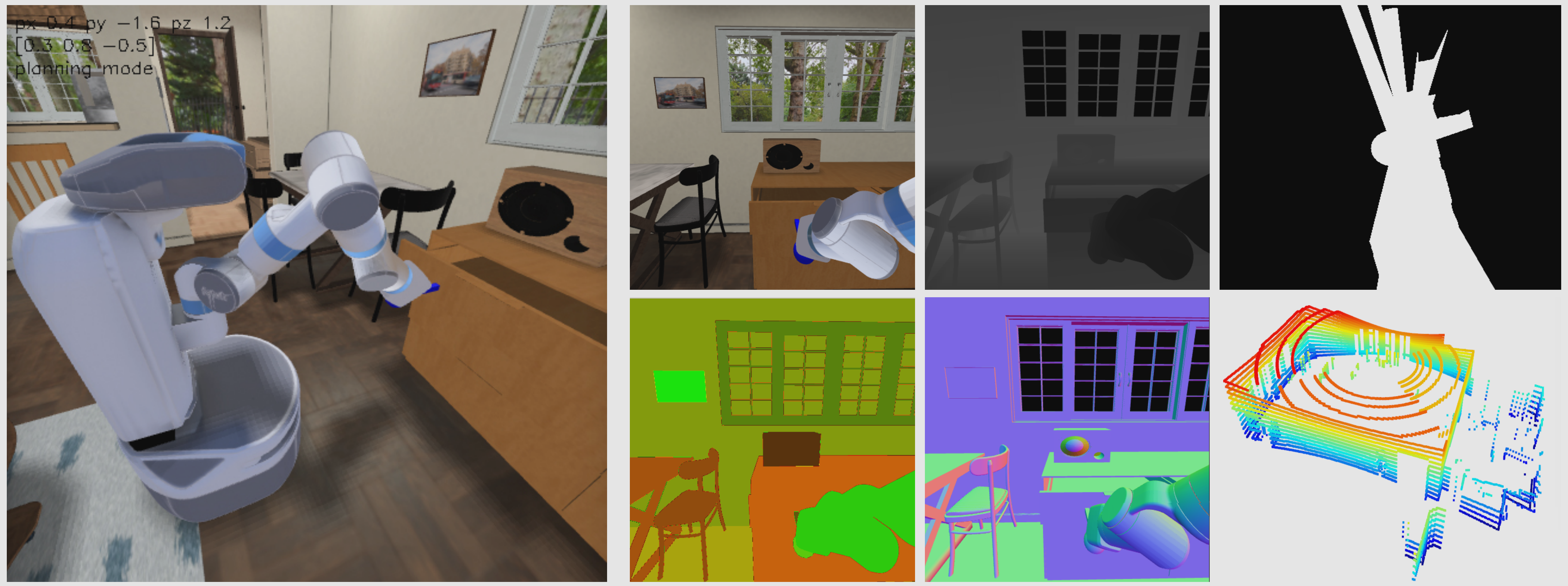
How do these features facilitate developing embodied agents?



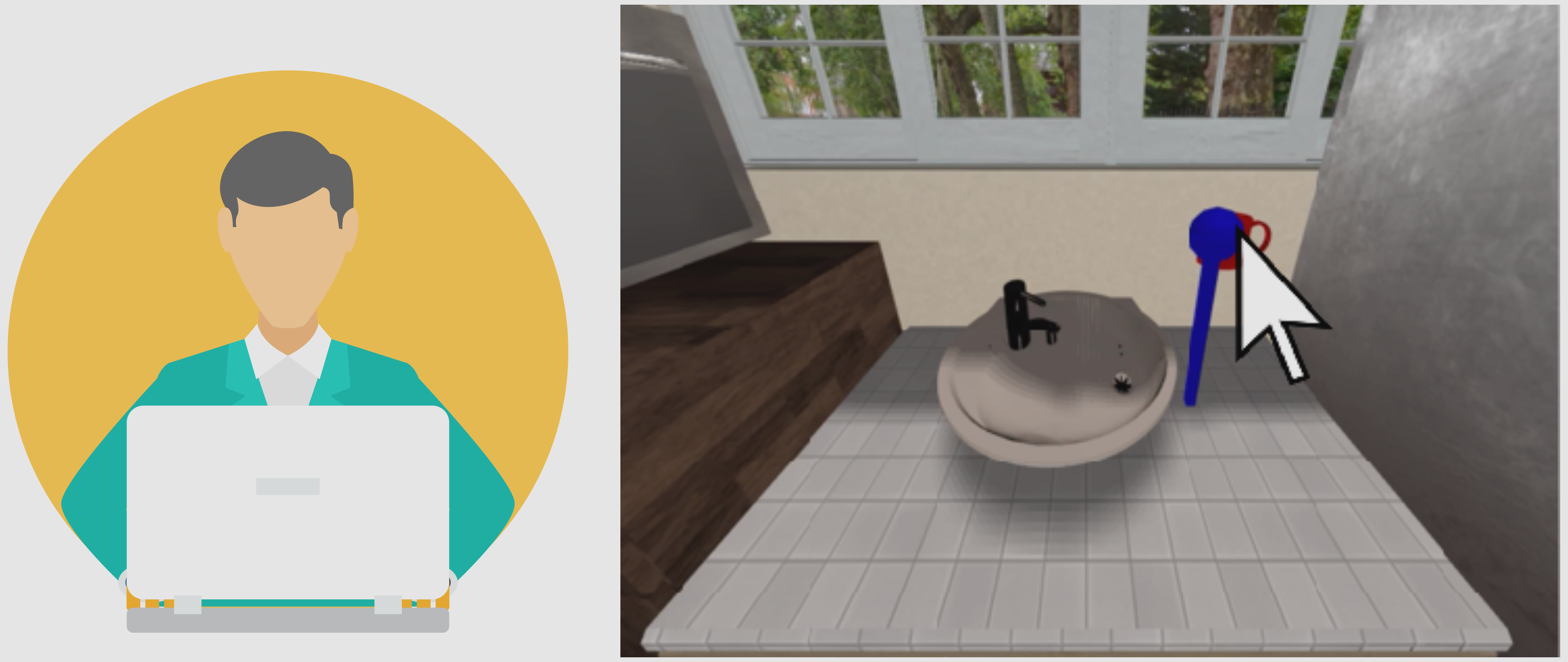
Repository of large scale scenes

Fully interactive scenes

Domain Randomization



Large variety of high quality virtual sensor signals



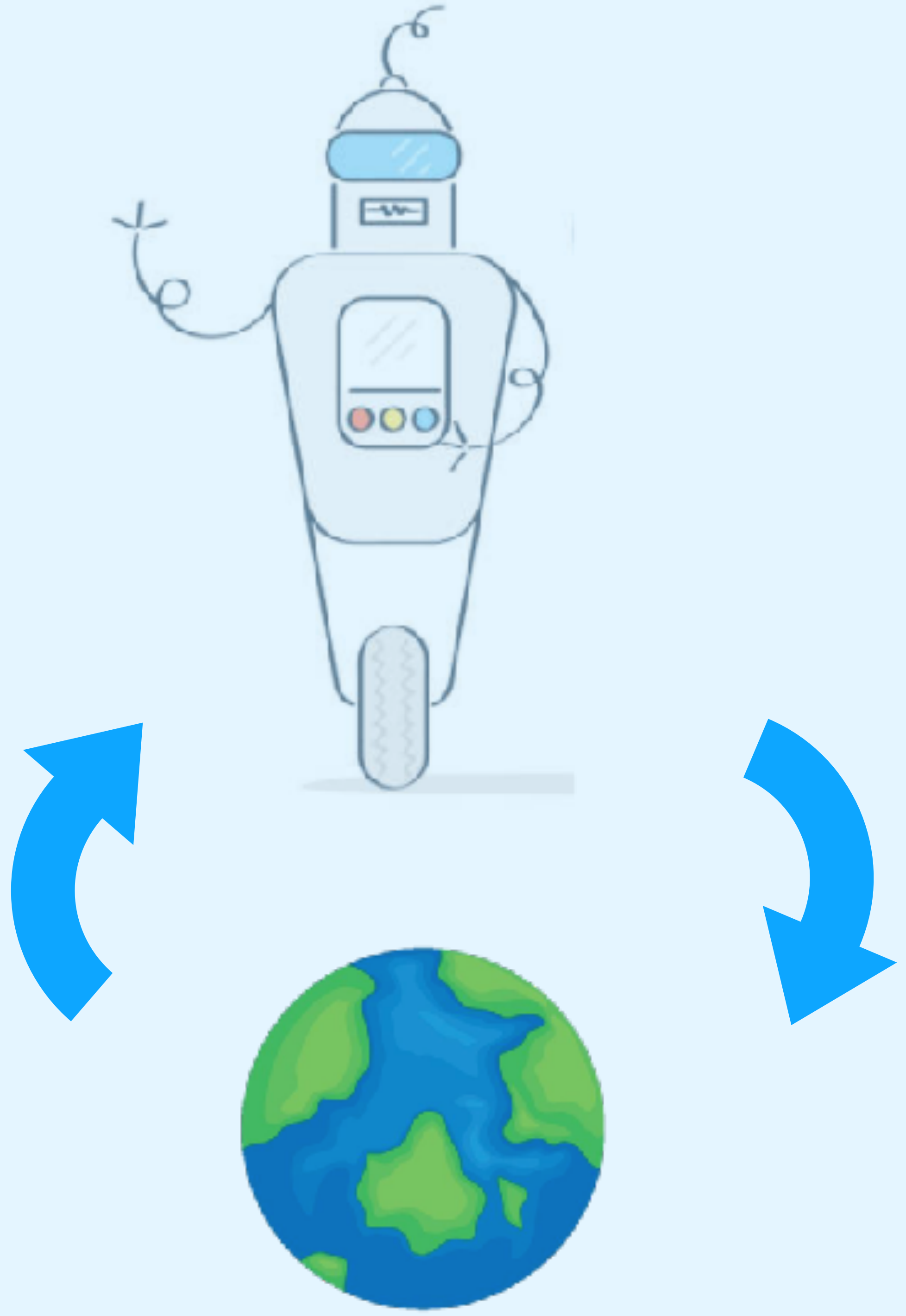
Easy to use Human-iGibson interface

S*X*L*M*…FS, iGIbson, *Submitted to IROS'21*

Large Scale Simulation for Embodied Perception and Robot Learning

Goal: Create simulation environments that reflect **complexity** of the **real world**, and develop intelligent agents in those environments.

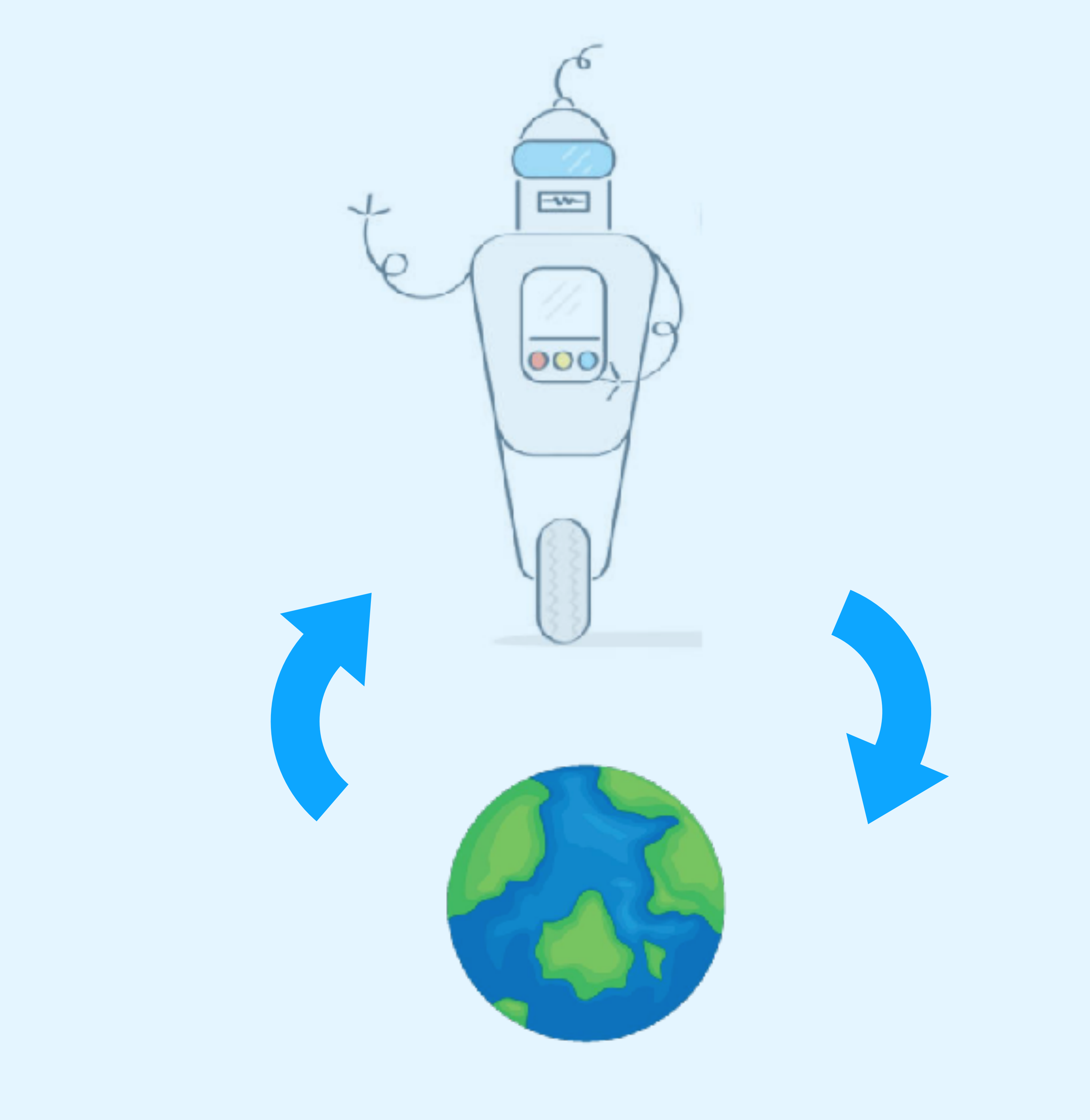
Creating a digital playground that replicates the complexity of the real world



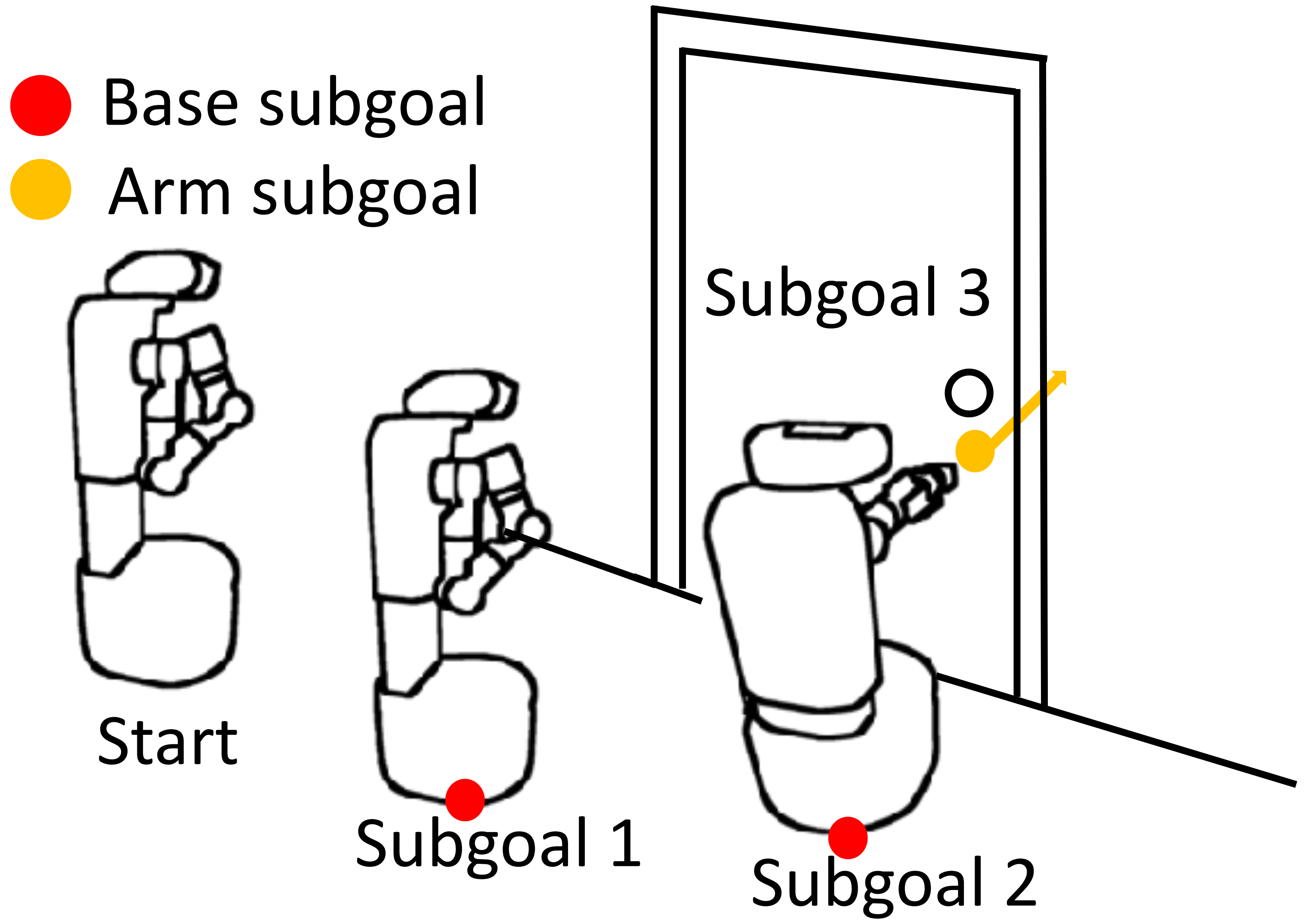
Learning algorithms for interactive and long-horizon tasks

Large Scale Simulation for Embodied Perception and Robot Learning

Goal: Develop intelligent agents for long horizon mobile manipulation tasks



Learning algorithms for interactive and long-horizon tasks



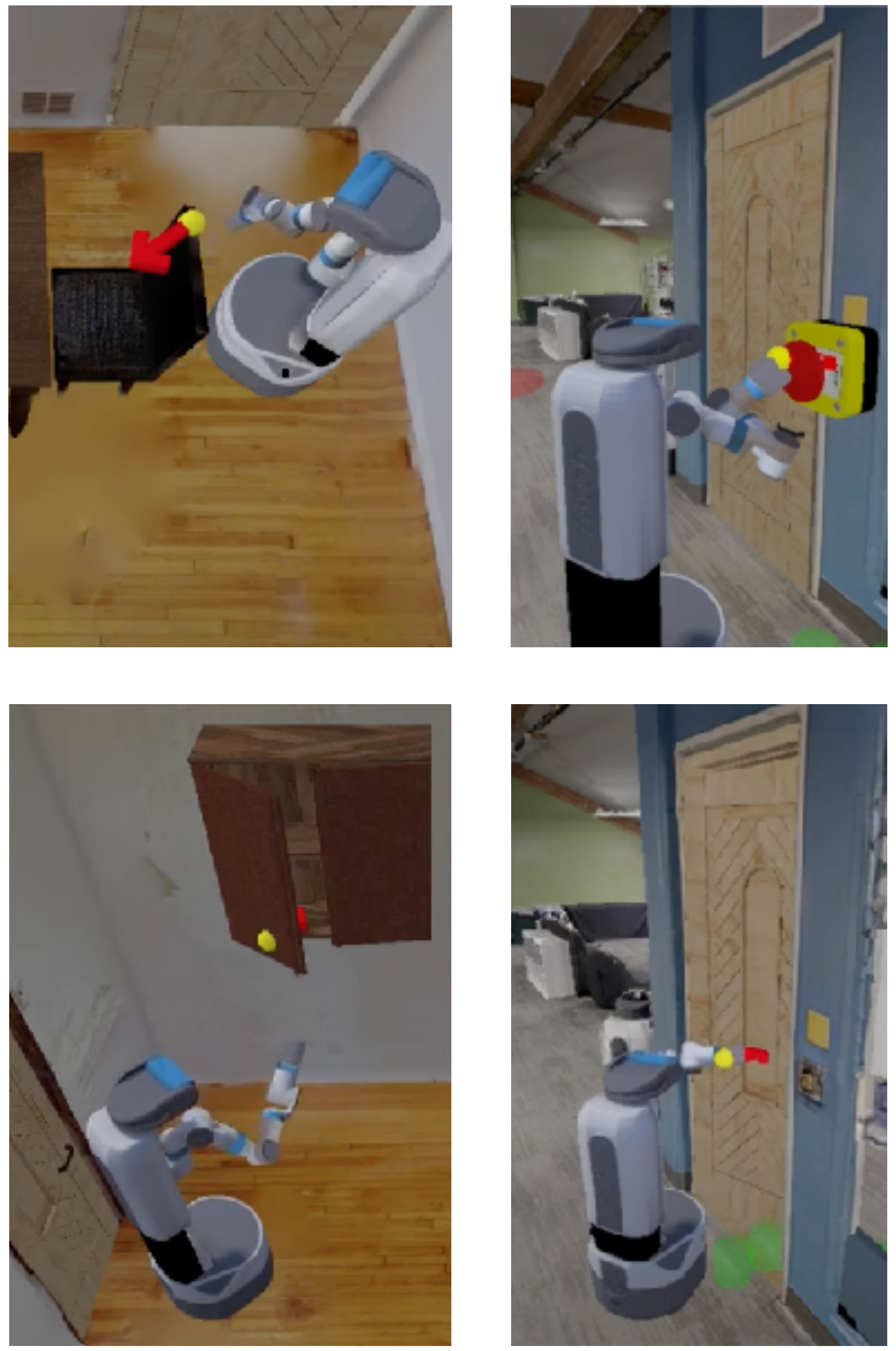
ReLMoGen. Xia et al 2021

SayCan. Google 2022

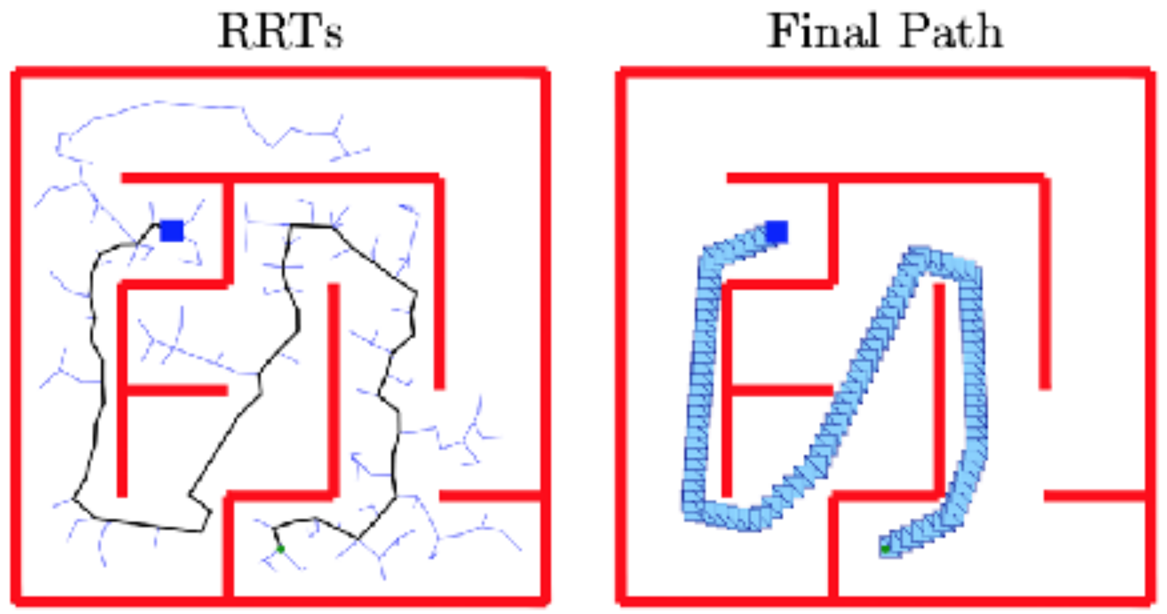
Image source: iStock photos, Standard license

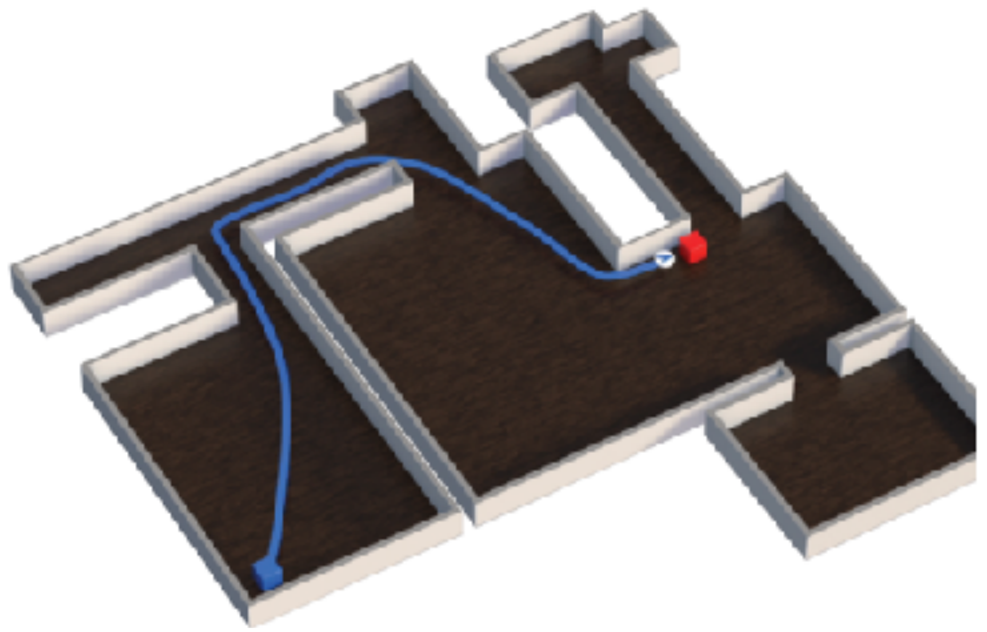
Introduction

- Mobile manipulation tasks -> a sequence of base and arm subgoals
- Subgoals -> points of interests in the environment, e.g. doors, chairs, cabinets, waypoints.



Introduction

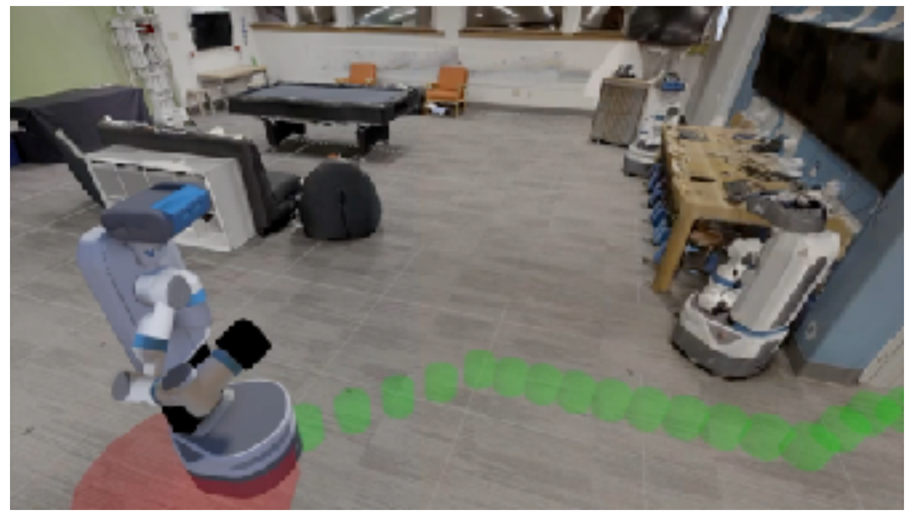




Motion Generation is good at solving "how to reach a point"

How do we combine the best of both worlds?

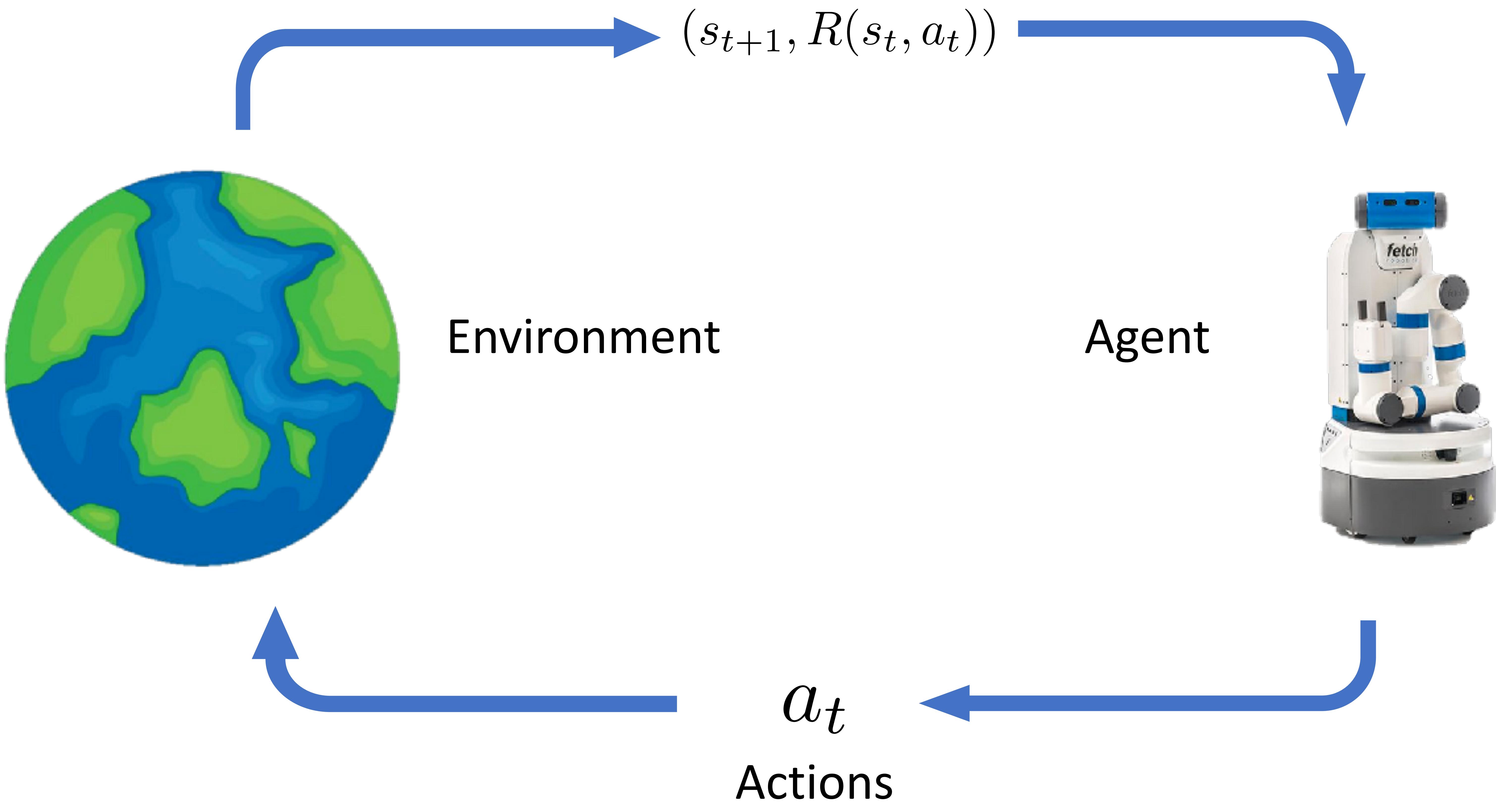
Reinforcement Learning is good at solving "where to go"



[Wijmans et al 2019] [Kuffner et al 2000]

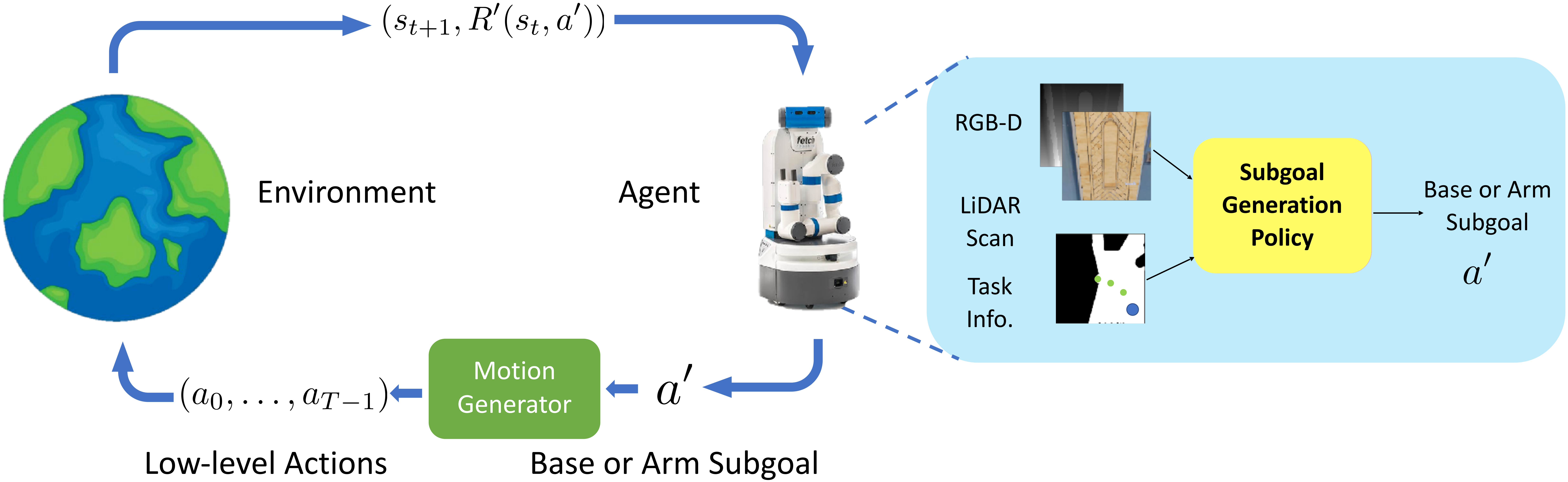
ReLMoGen

- A framework that leverages a Motion Generator in an RL loop, and lifts the action space from joint commands to subgoals.

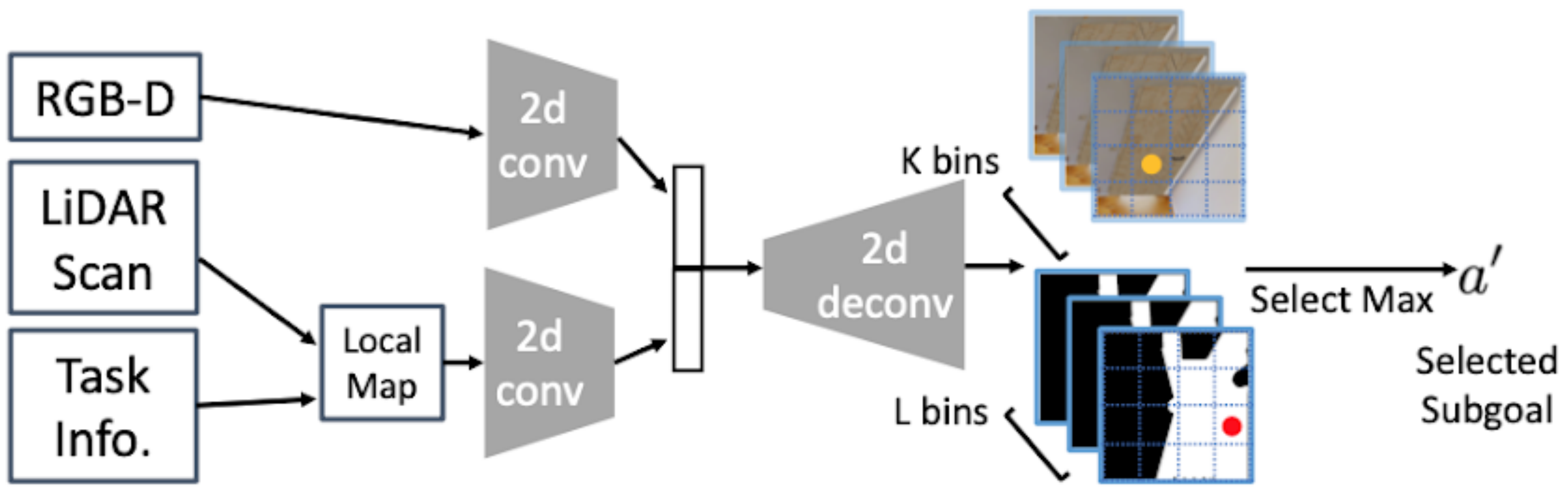


ReLMoGen

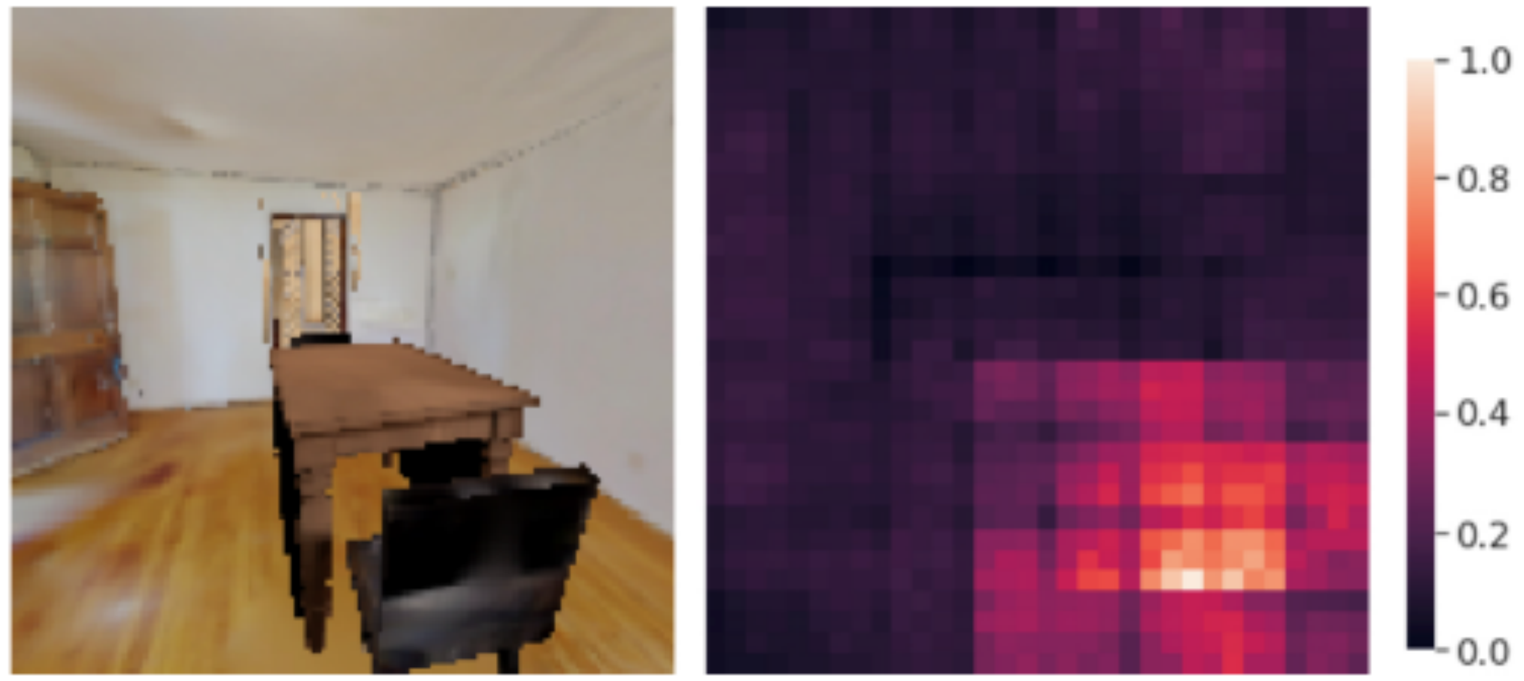
- A framework that leverages a Motion Generator in an RL loop, and lifts the action space from joint commands to subgoals.



Policy Networks

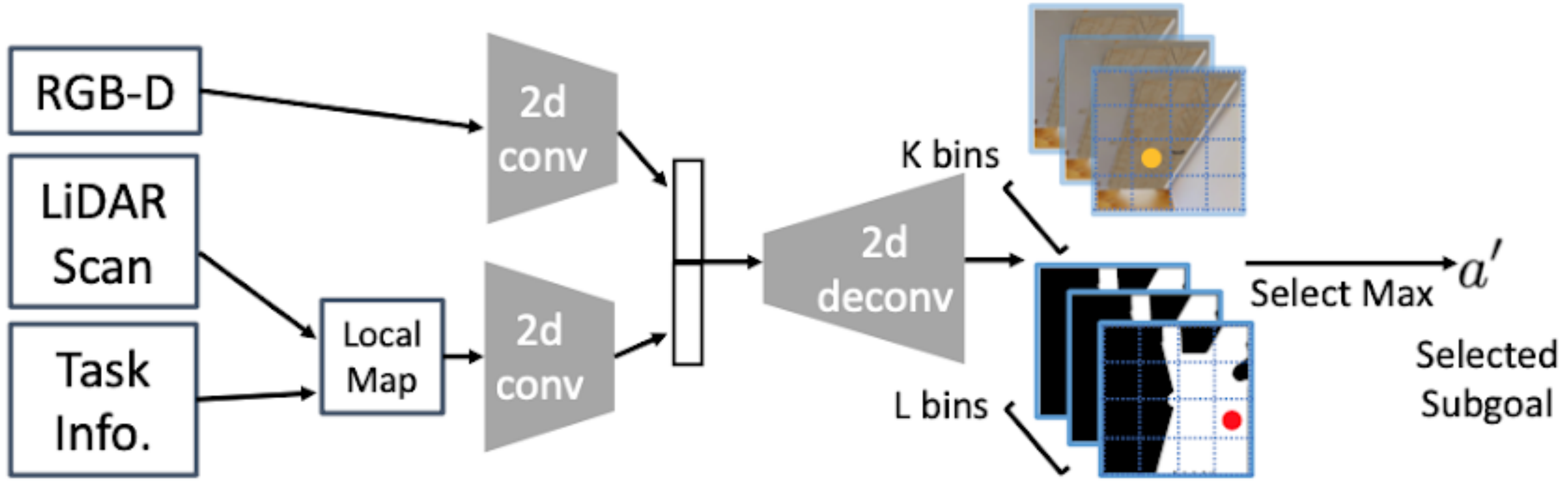


(a) Subgoal Generation Policy SGP-D

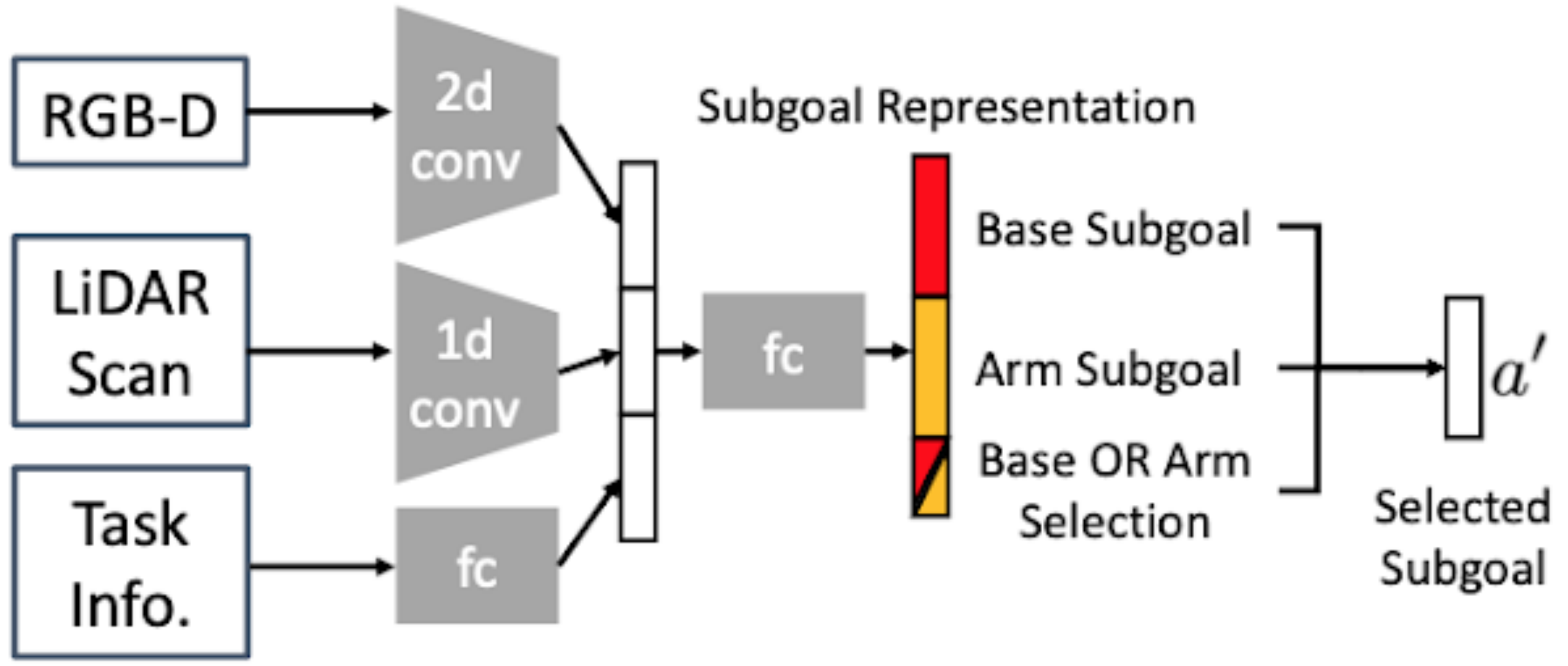


D indicates Dense

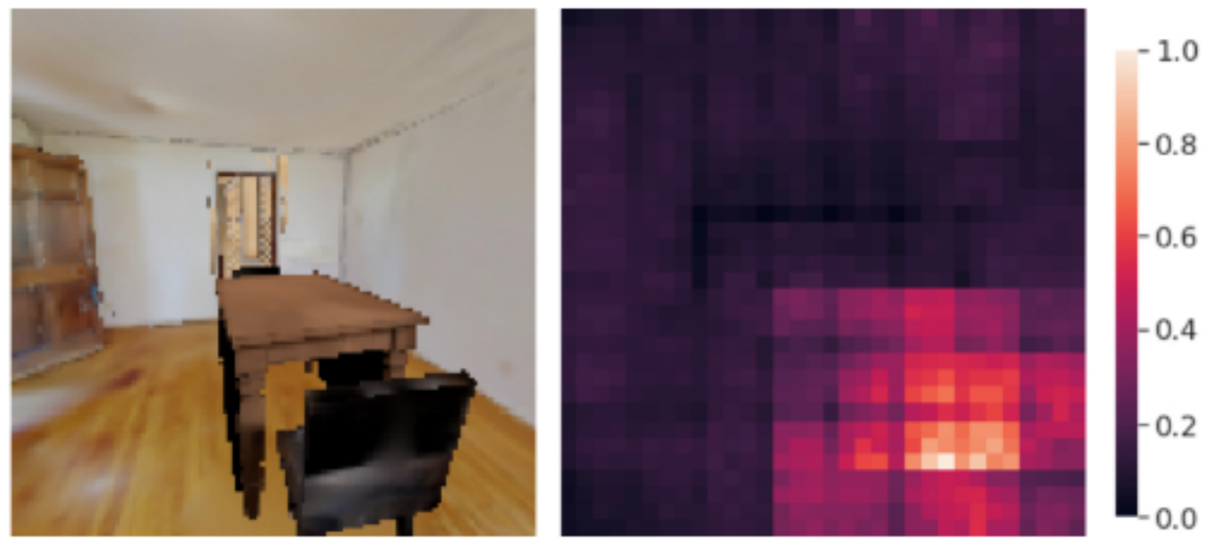
Policy Networks



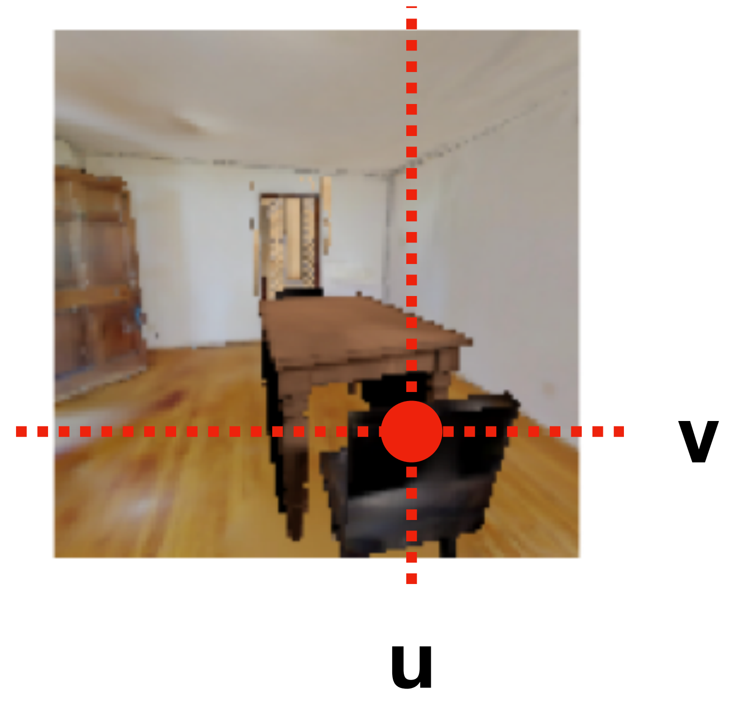
(a) Subgoal Generation Policy SGP-D



(b) Subgoal Generation Policy SGP-R



D indicates Dense



R indicates Regression

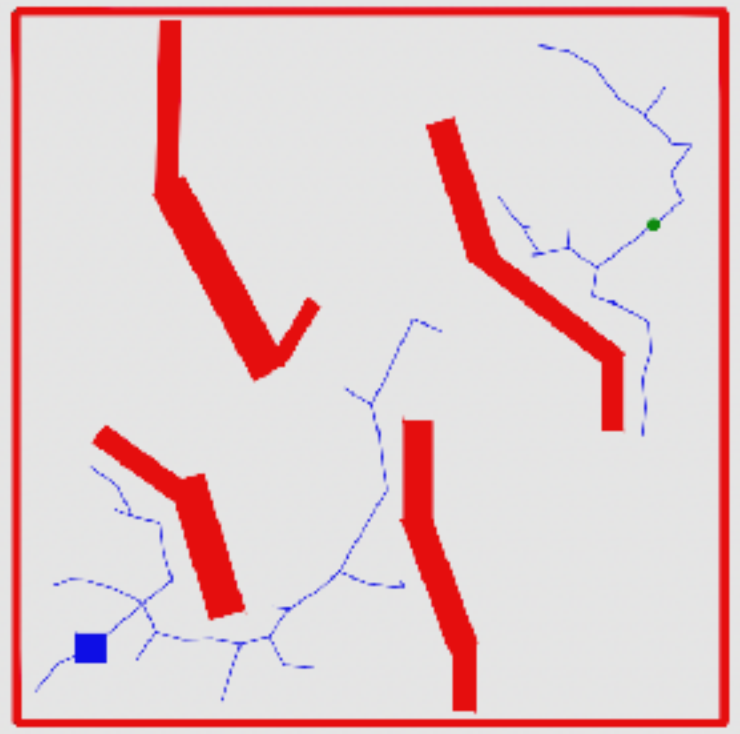
Motion Generation

Motion Generation

- A motion planner that searches for trajectories based on current sensor information
- A set of common low-level controllers that execute the planned trajectories

Motion Planners used:

- RRT-Connect
- LazyPRM

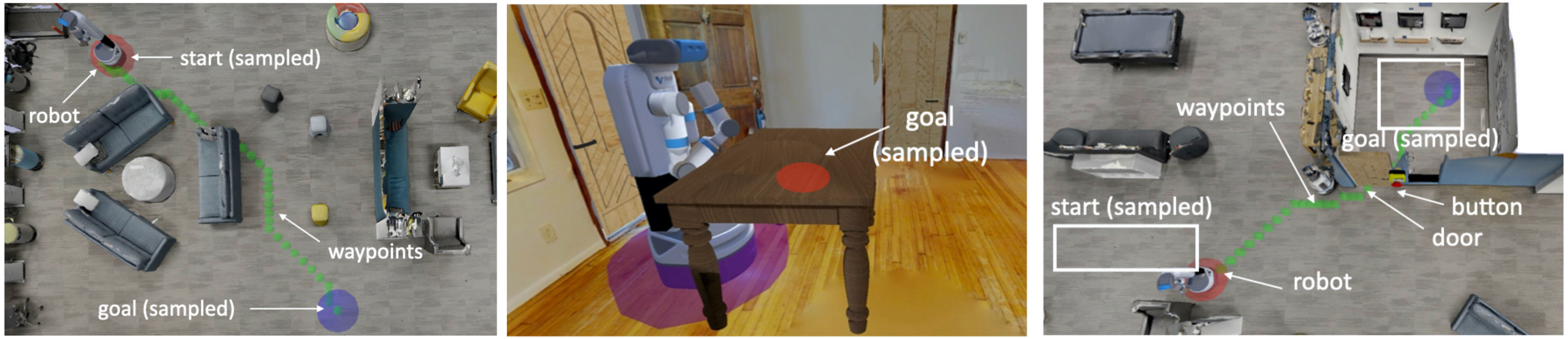


Training time optimization:

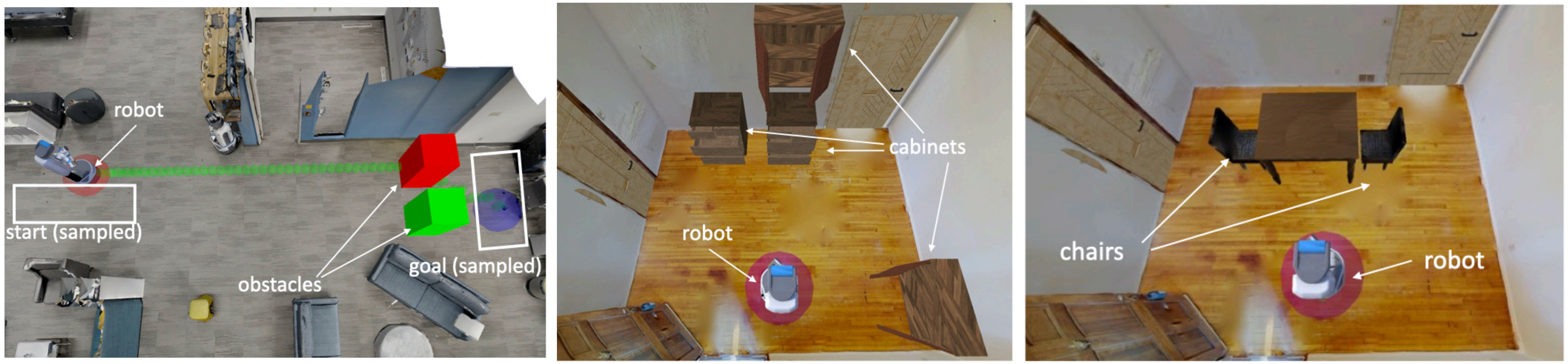
- Jump to the last state in the plan if plan is collision-free

[Kuffner et al 2000] [Bohlin et al 2000]

Experimental Setup



(a) `PointNav` (b) `TabletopReachM` (c) `PushDoorNav`, `ButtonDoorNav`



(d) `InteractiveObstaclesNav` (e) `ArrangeKitchenMM` (f) `ArrangeChairMM`

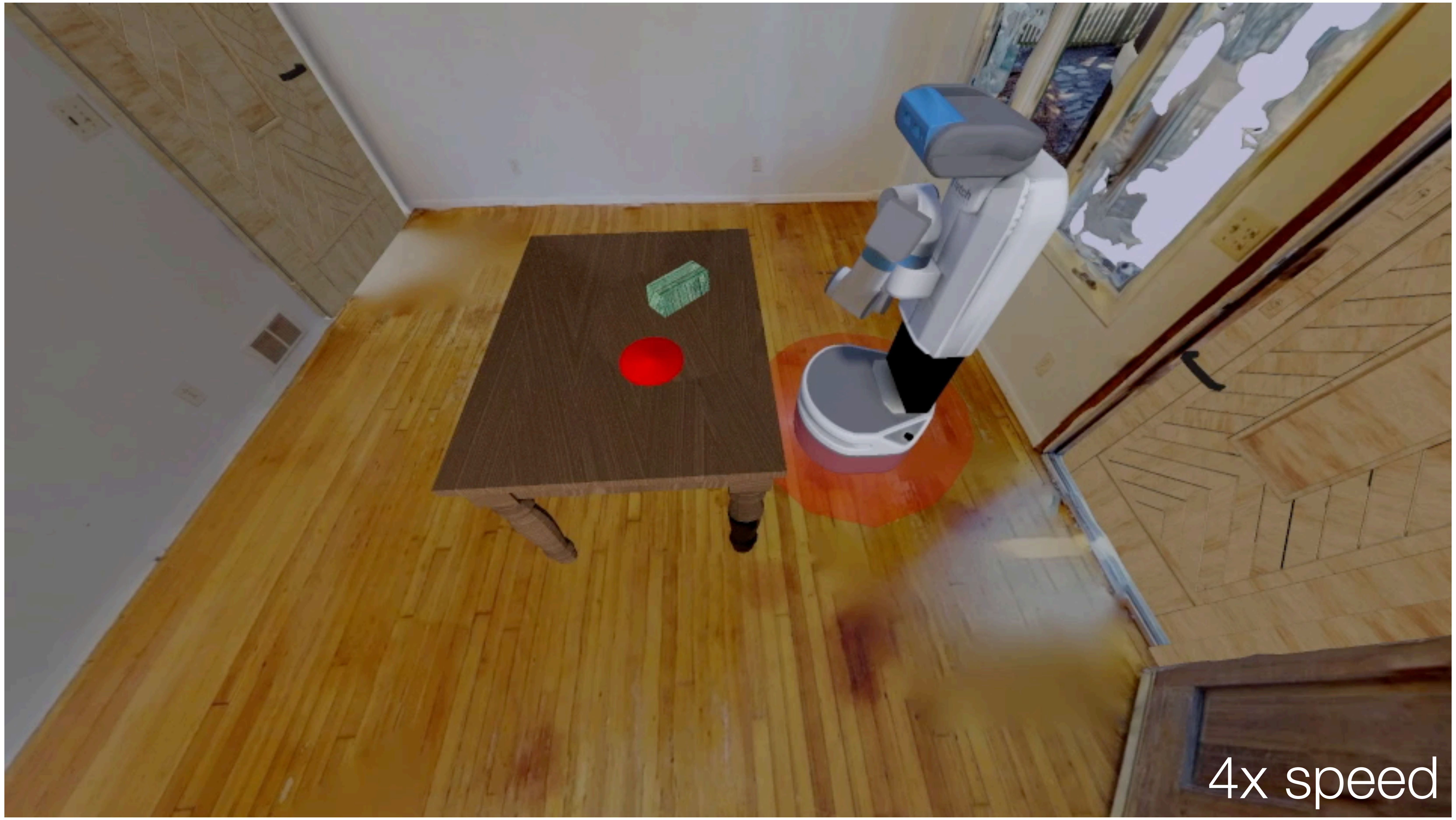
Policy Visualization

PointNav

Policy Visualization

TabletopReachM

Policy Visualization



TabletopManipM

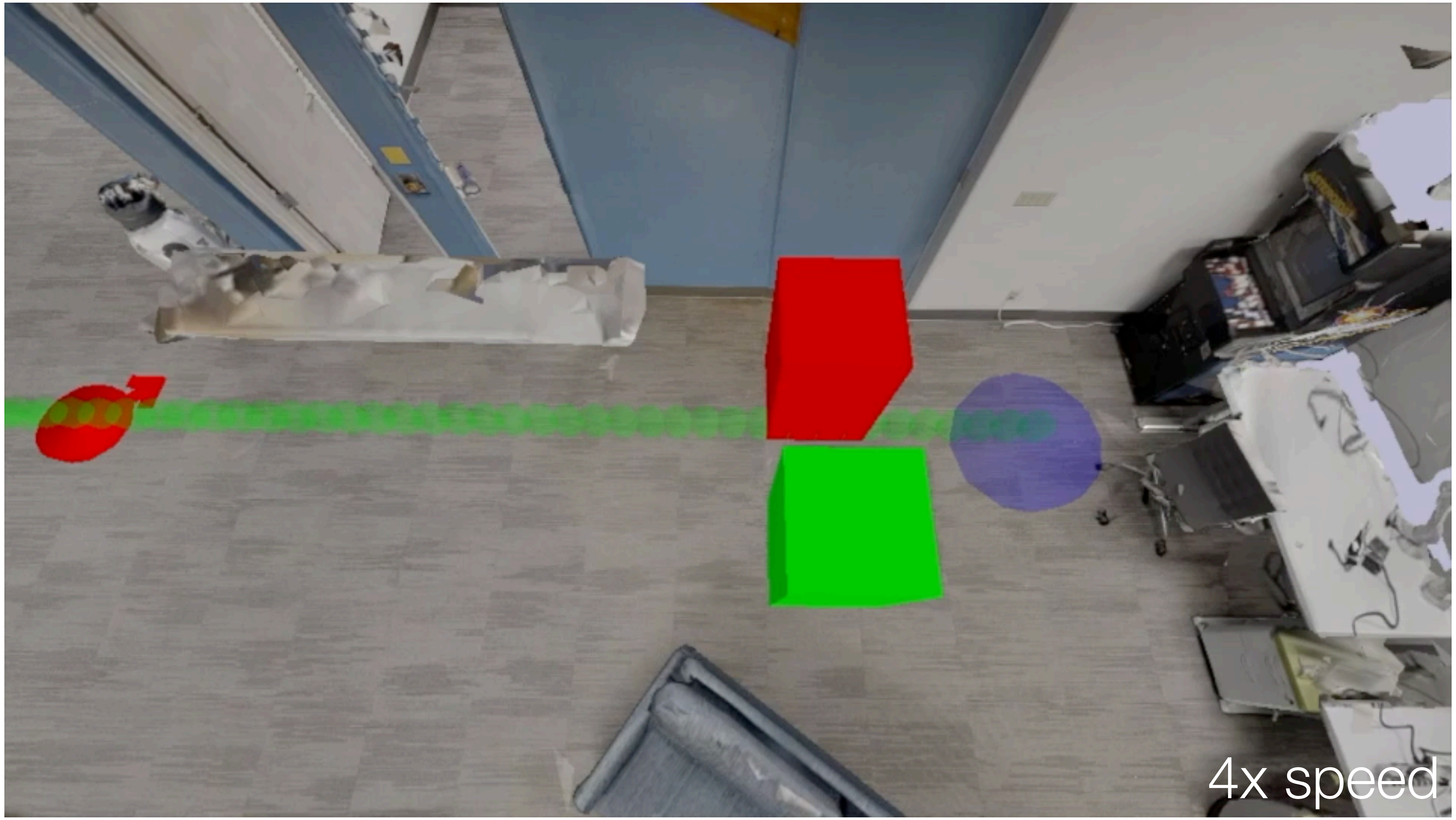
Policy Visualization

PushDoorNav

Policy Visualization

ButtonDoorNav

Policy Visualization



InteractiveObstaclesNav

Policy Visualization

ArrangeChairMM

Policy Visualization

ArrangeKitchenMM

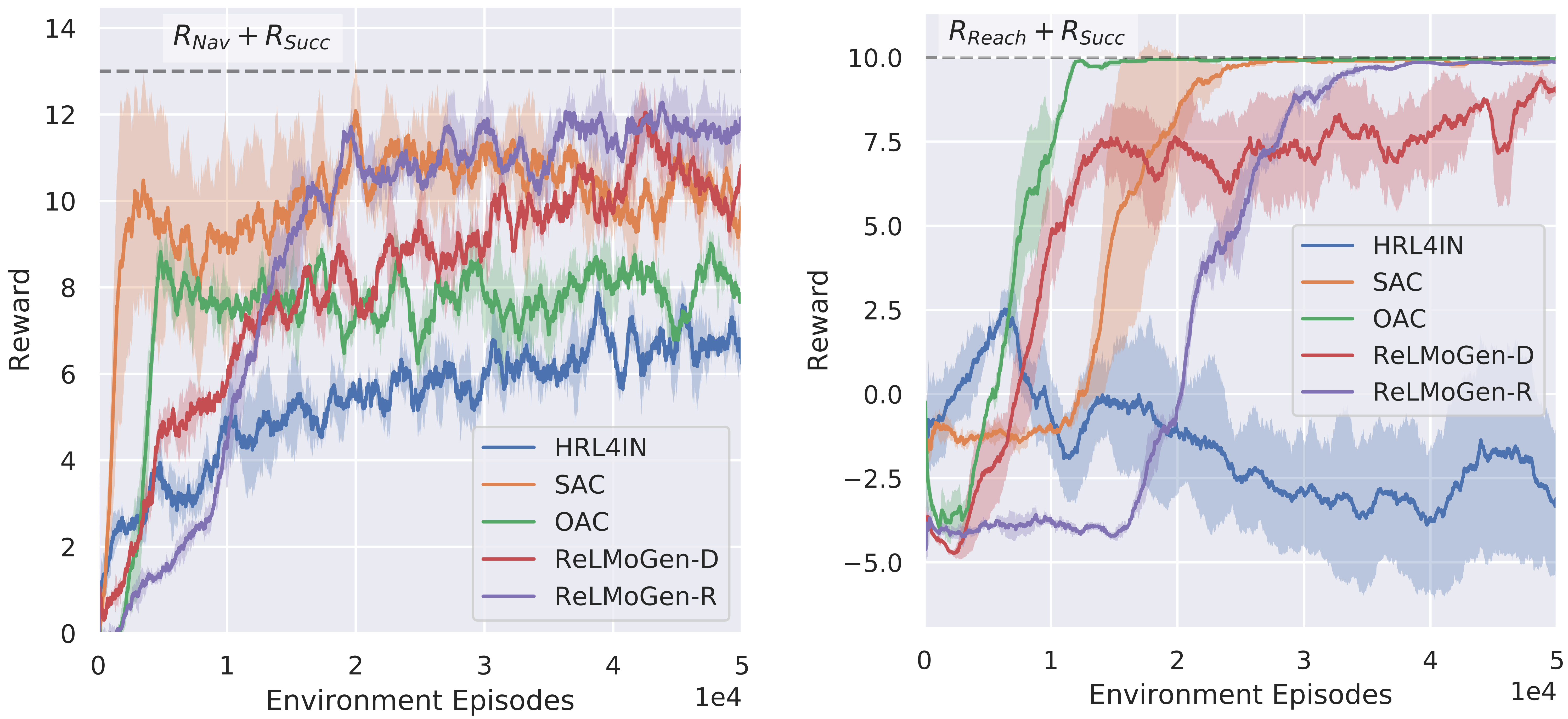
Baselines and Metrics

- Baselines
  - SAC on joint velocities
  - OAC on joint velocities
  - HRL4IN on joint velocities

- Metrics:
  - SPL (Success weighted by Path Length) for navigation tasks
  - Task completion (number of drawers/cabinets closed, chairs tucked within 10 degs / 10 cm and 5 degs / 5 cm) for mobile manipulation tasks

[SAC, Haarnoja et al 2018] [OAC, Ciosek et al 2019] [HRL4IN, Li et al 2020]

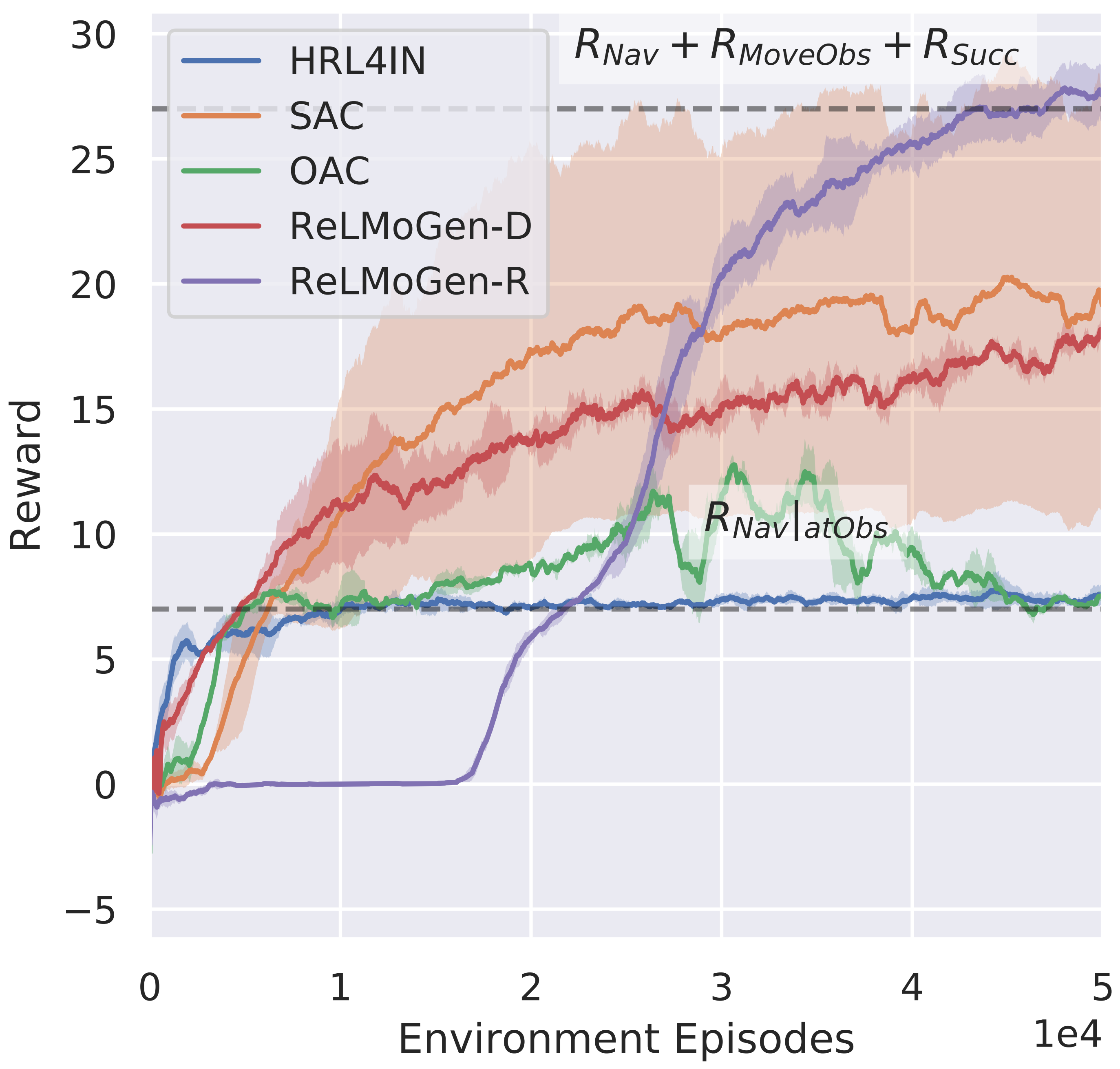
Quantitative Results



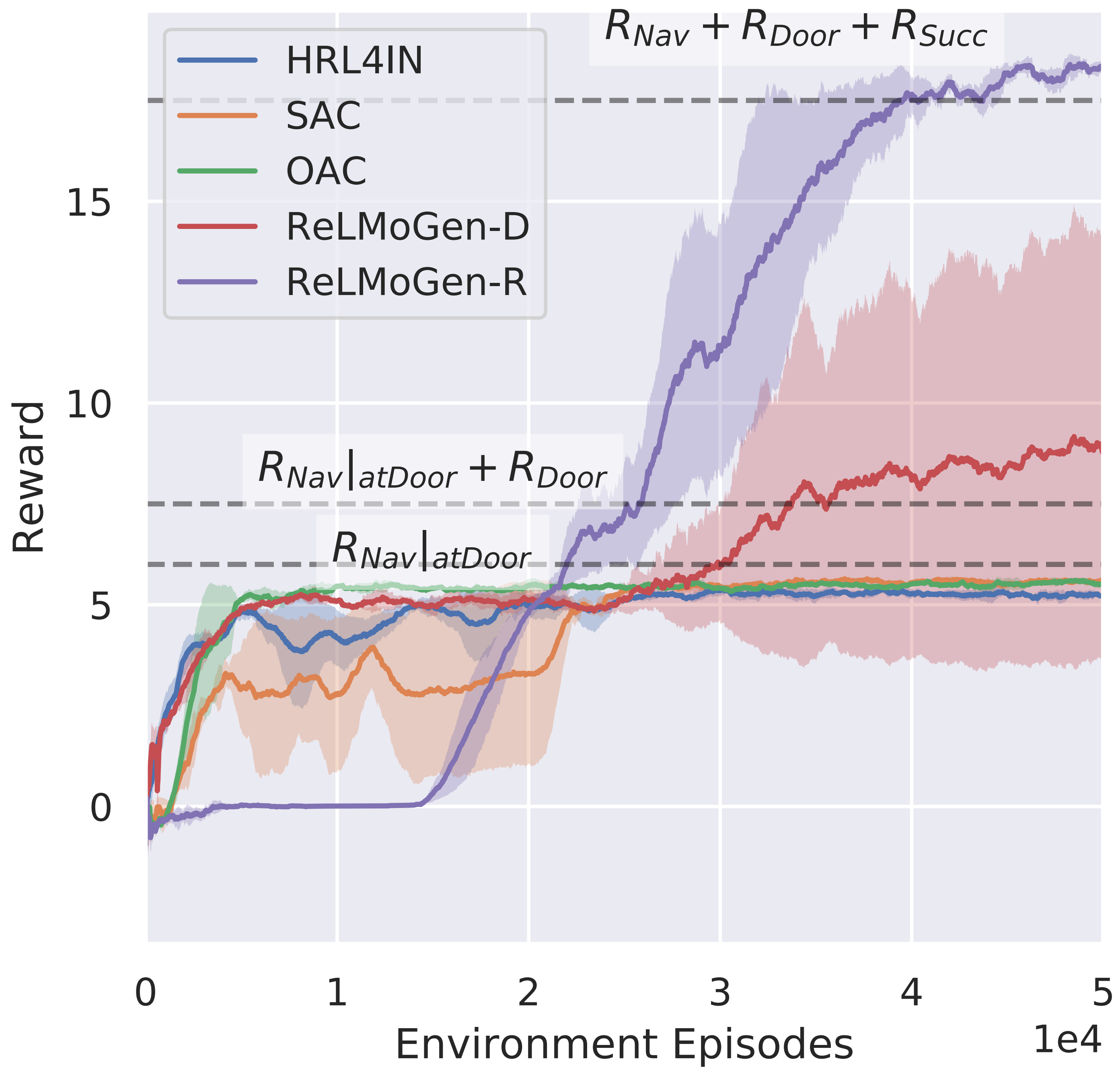
(a) PointNav (b) TabletopReachM

Short horizon tasks: similar performance as baselines

Quantitative Results



(c) `Int.ObstaclesNav`

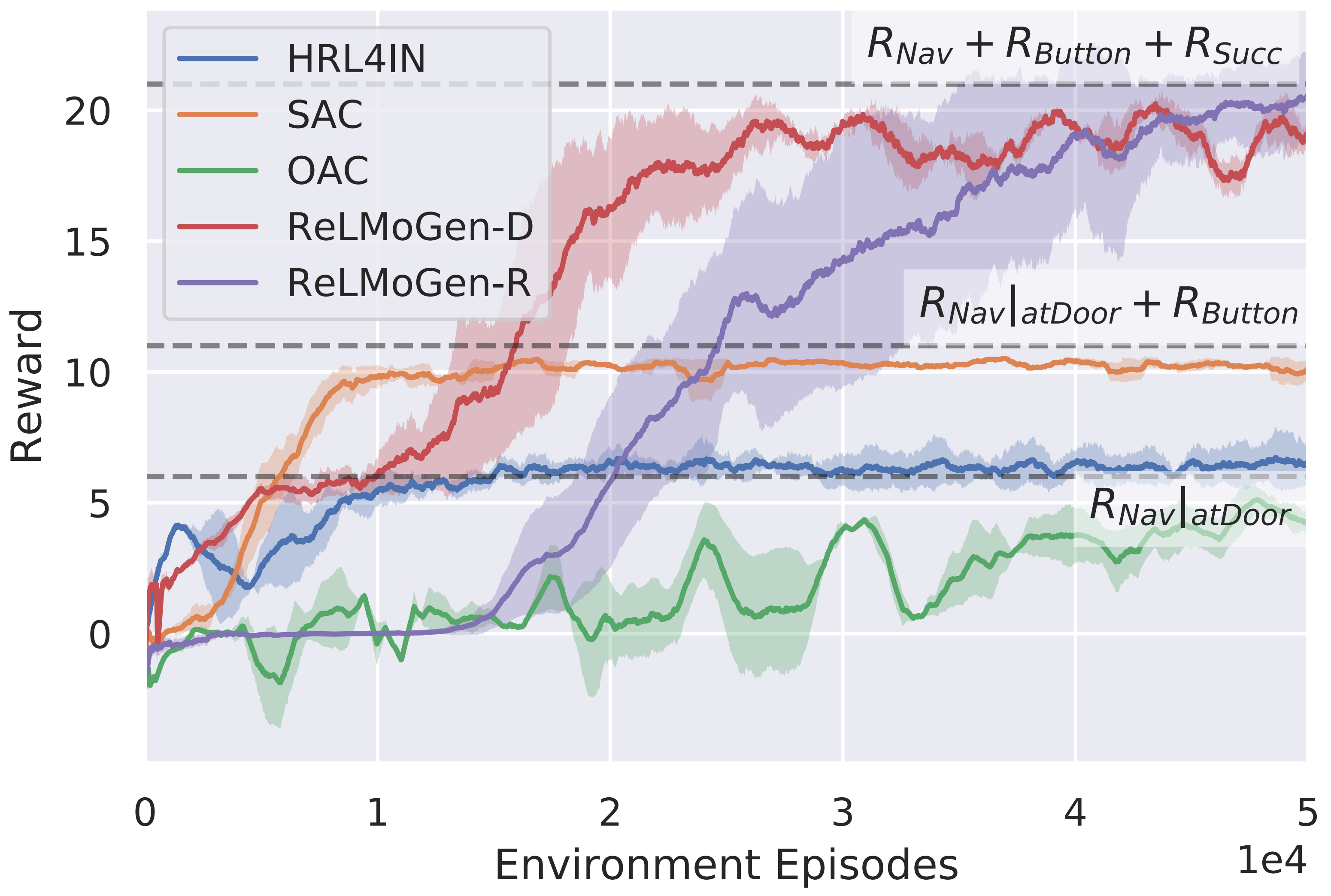


(d) `PushDoorNav`

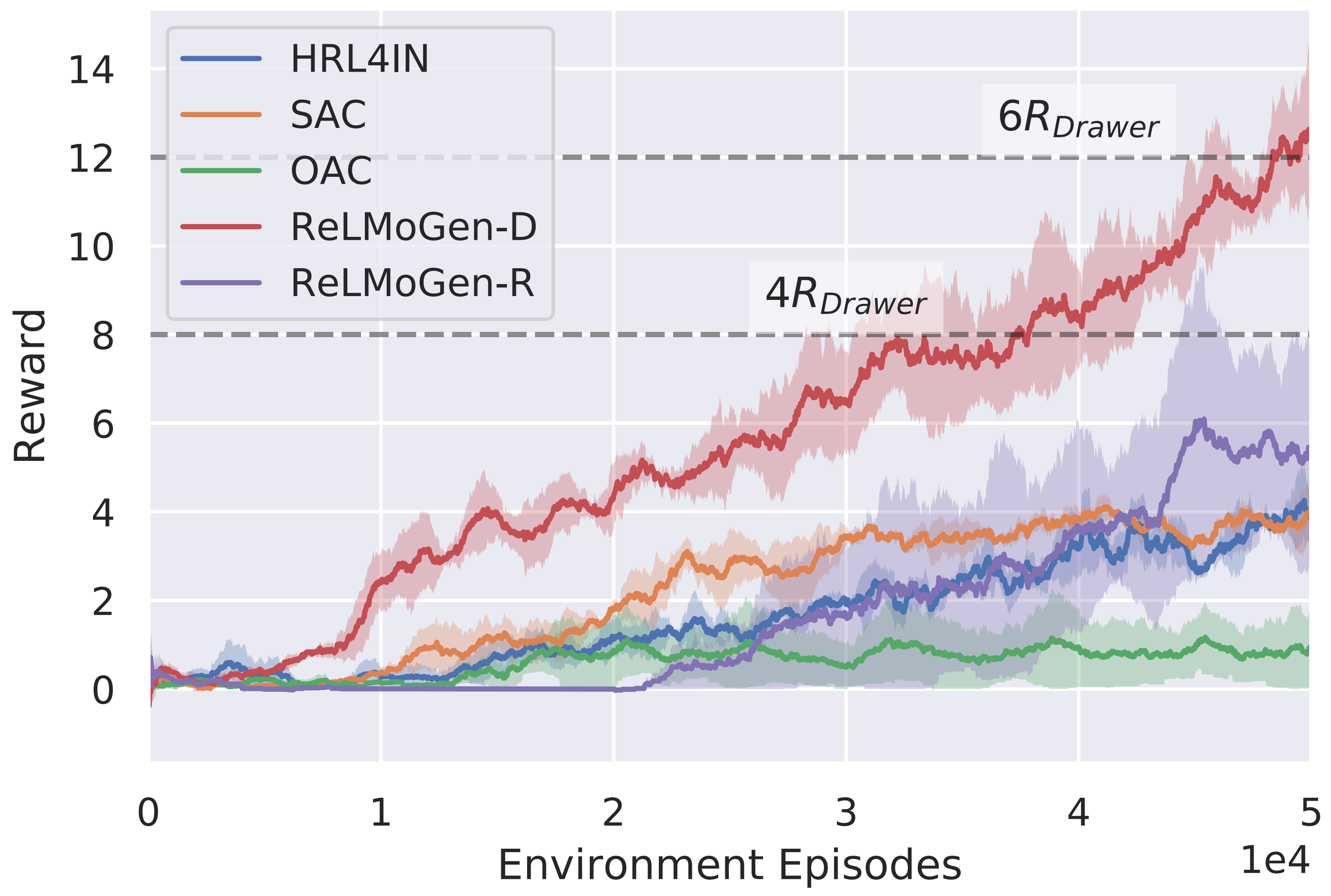
Interactive Navigation tasks:

- ReLMoGen outperforms baselines
- ReLMoGen-R has better sample efficiency

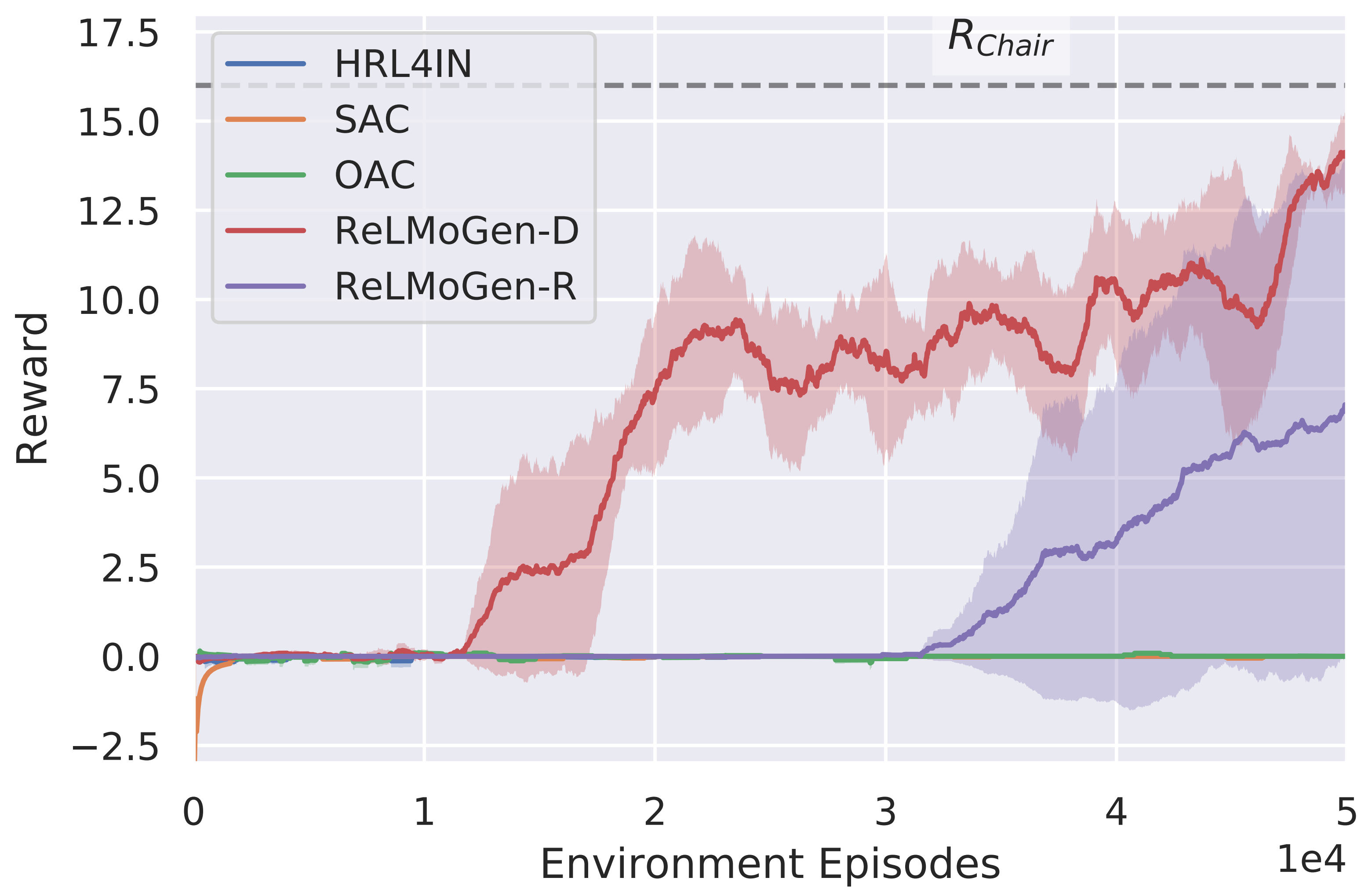
Quantitative Results



(e) ButtonDoorNav



(f) ArrangeKitchenMM



(g) ArrangeChairMM

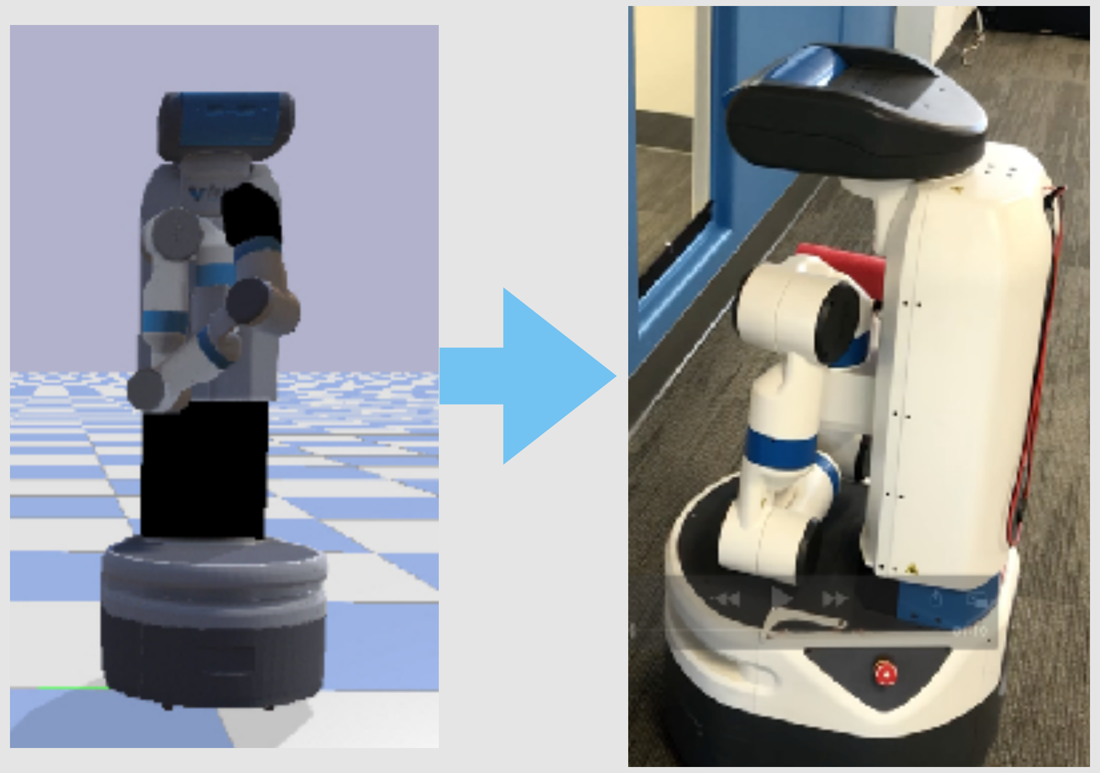
Mobile Manipulation Tasks

- ReLMoGen outperforms baselines
- ReLMoGen-D has better sample efficiency
- These tasks requires input-output image space alignment

Generalization

New Scenes

New Embodiment



Sim2real

Policy Visualization - Fine tuning

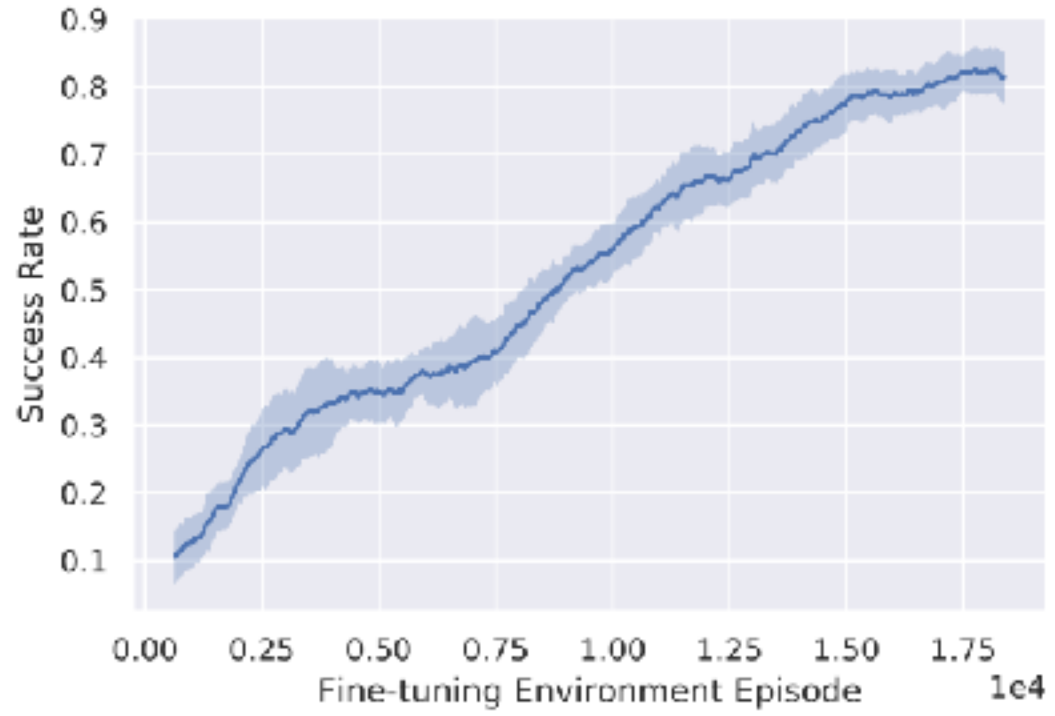
PushDoorNav - Fine tuning on novel environments

Transfer to new embodiment type

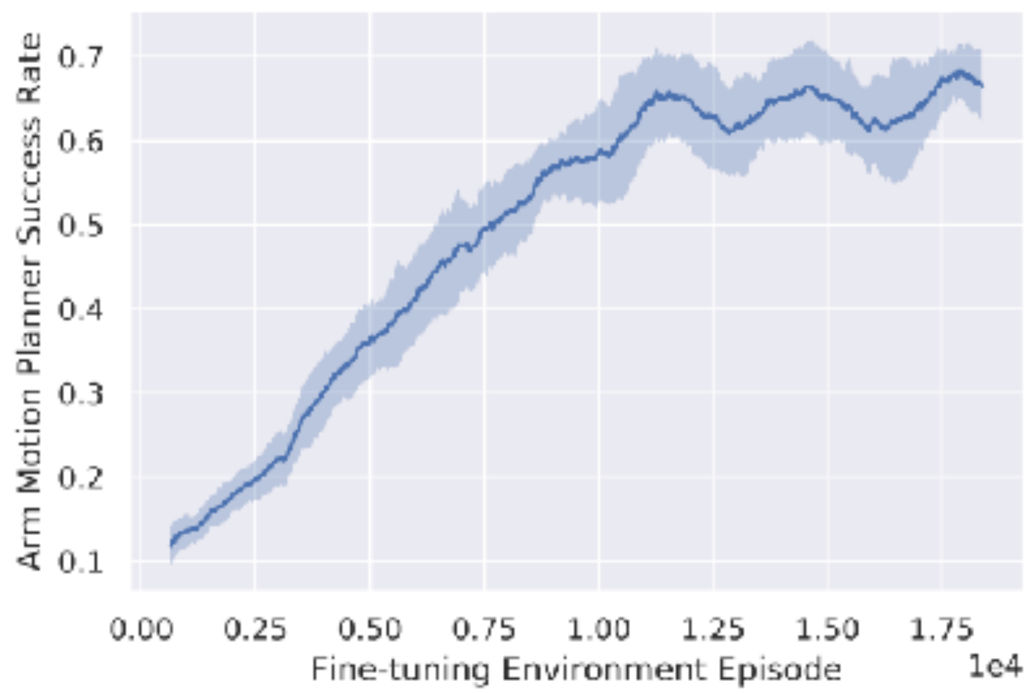
Fetch -> Movo

Movo on PushDoor Task

Transfer to new embodiment type



(b) Task Success Rate



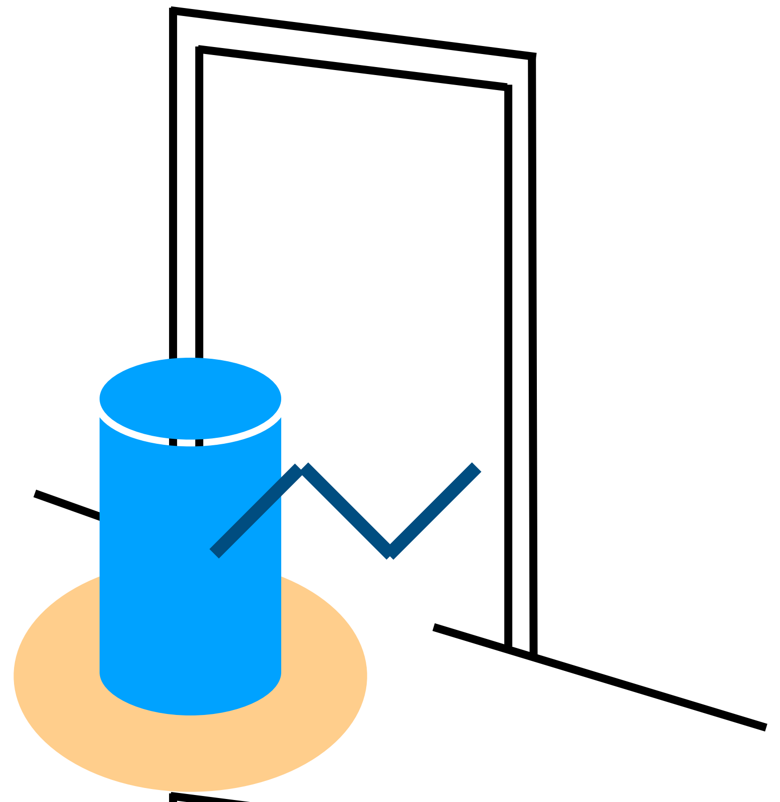
(c) Arm MP Success Rate

Convergence is 60% faster than from scratch

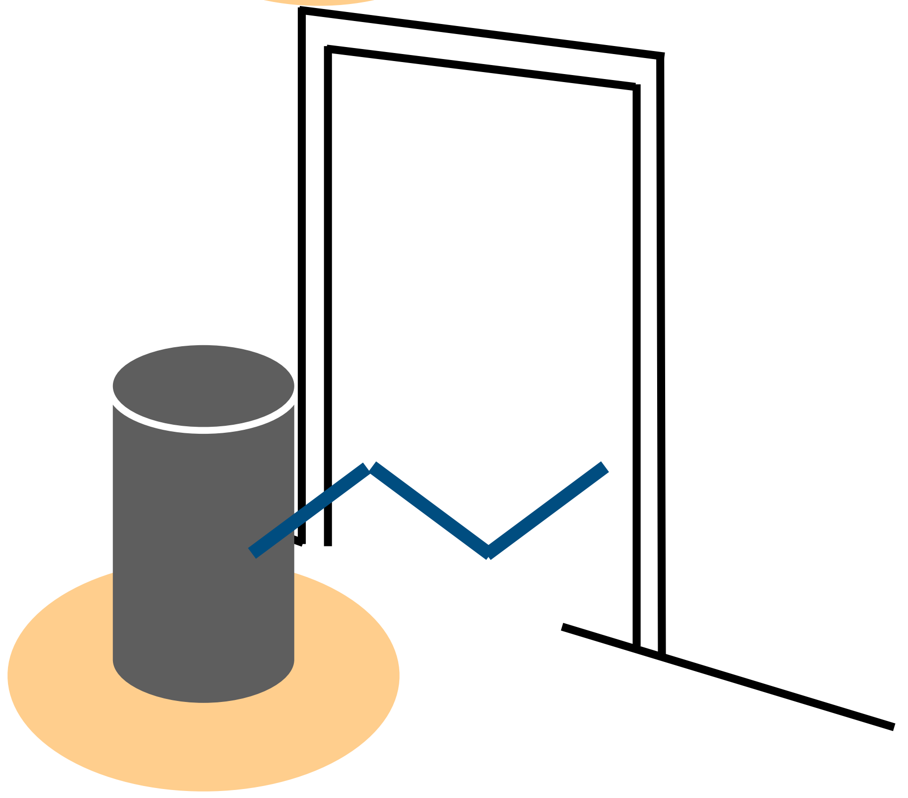
SAC optimization Objective

$$J(\theta) = \sum_{t=1}^{T} \mathbb{E}_{(s_t,a_t)\sim\rho_{\pi_\theta}} [r(s_t, a_t) + \boxed{\alpha\mathcal{H}(\pi_\theta(.\,|s_t))]}$$

Ideal base position for Fetch

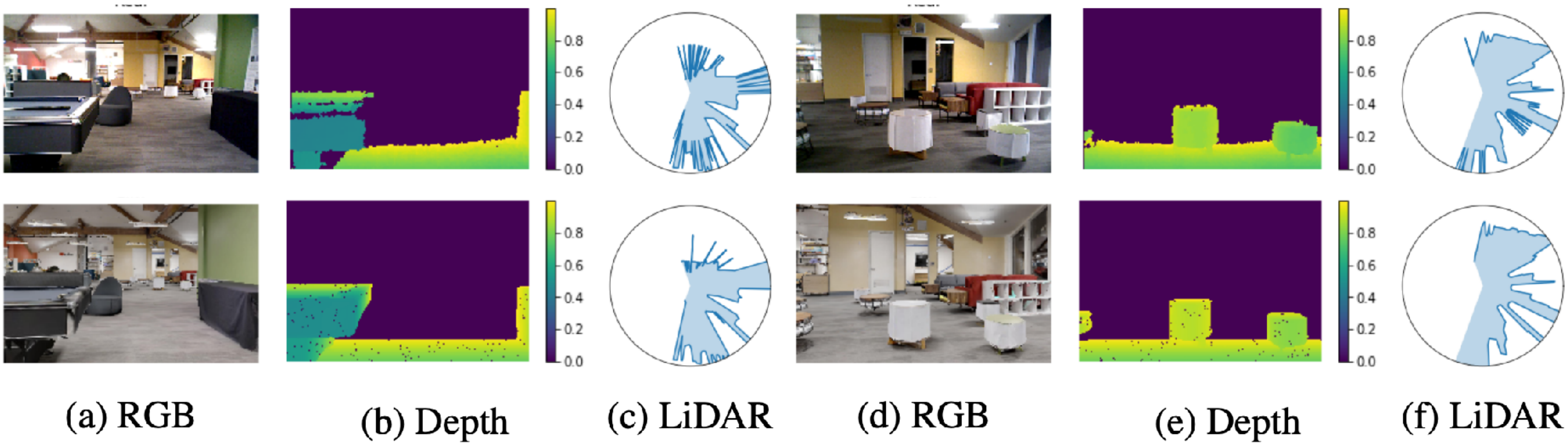


Ideal base position for Movo



Sim2Real transfer potential

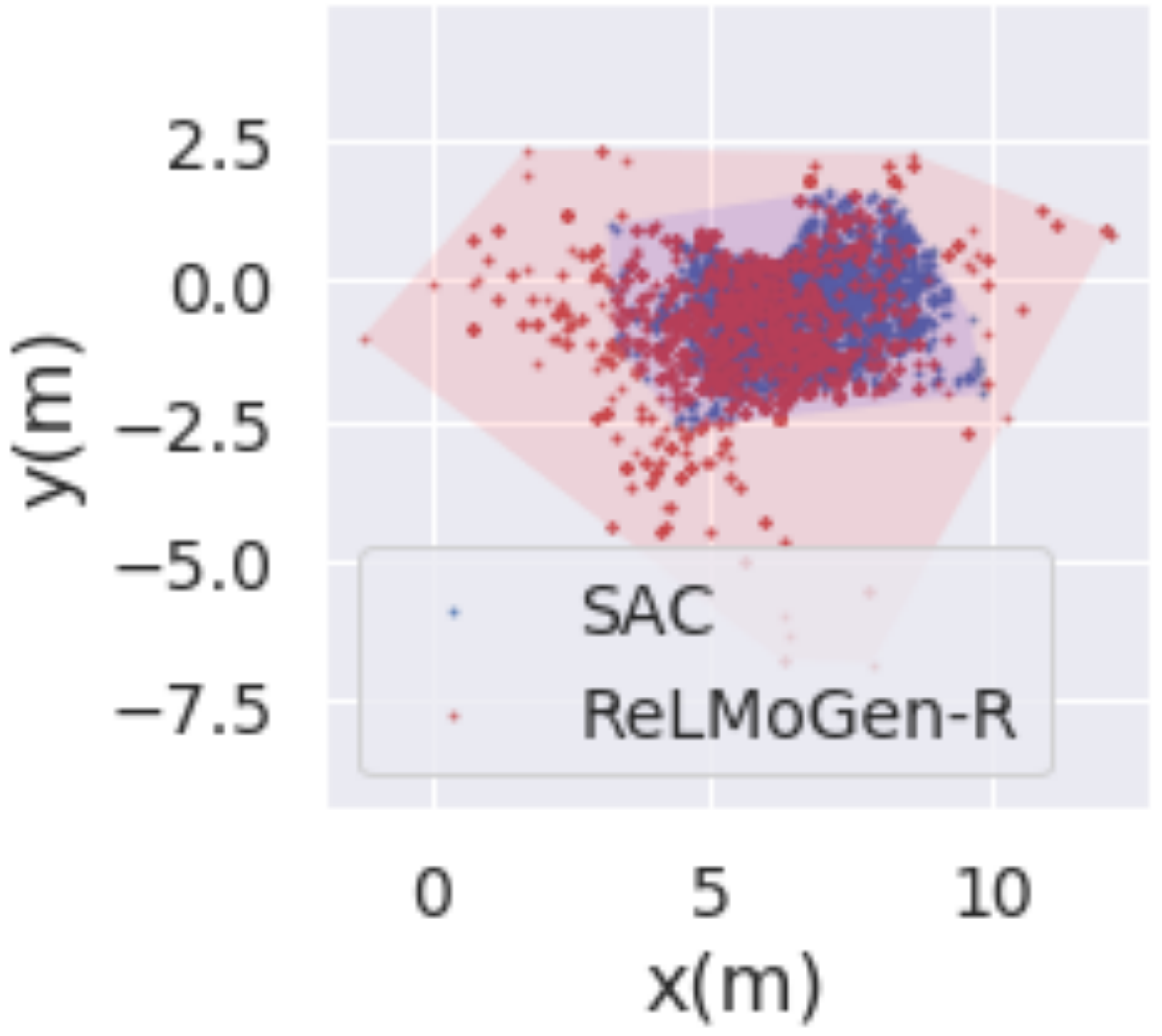
Real sensor observations

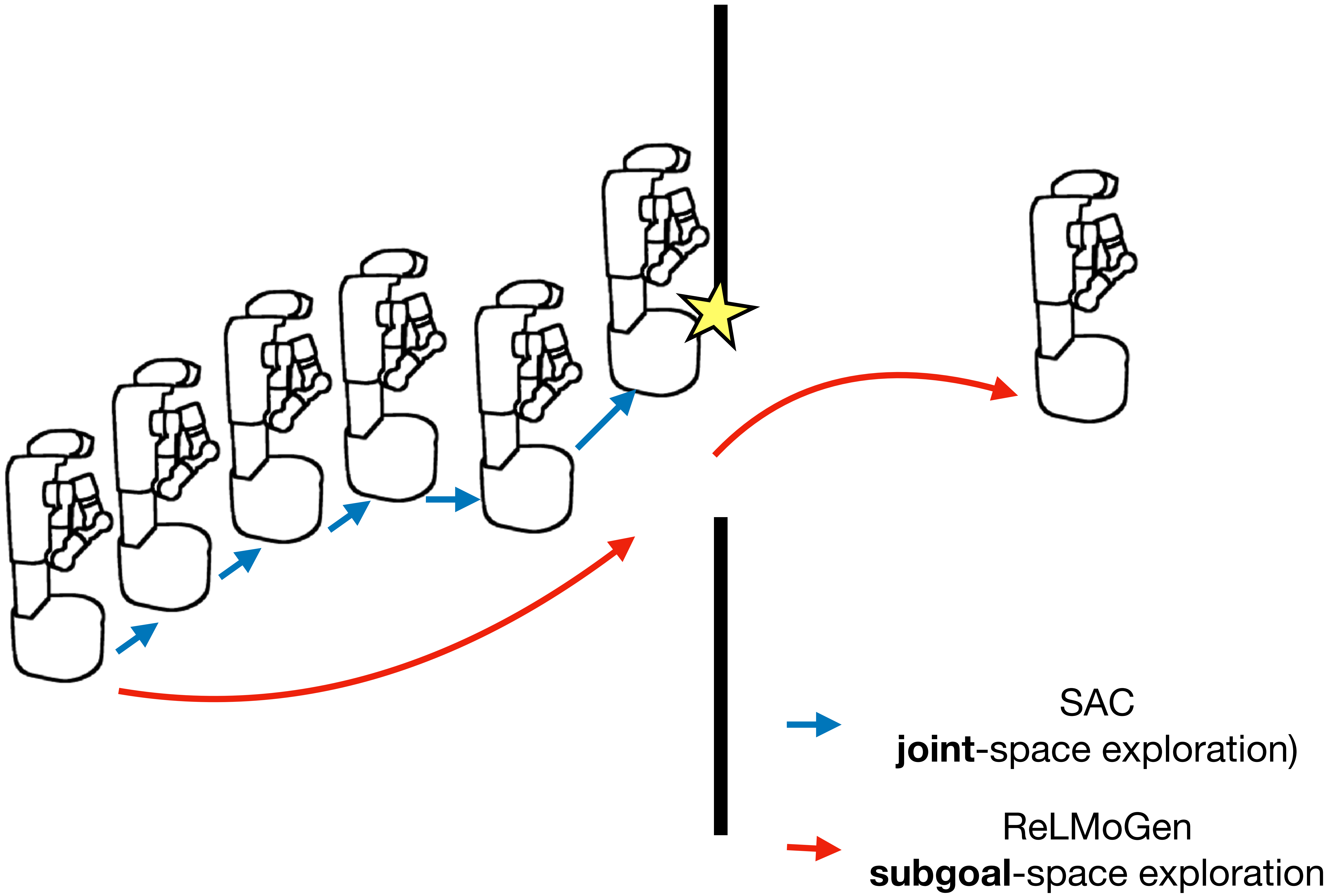


Sim sensor observations

Sim2Real transfer potential

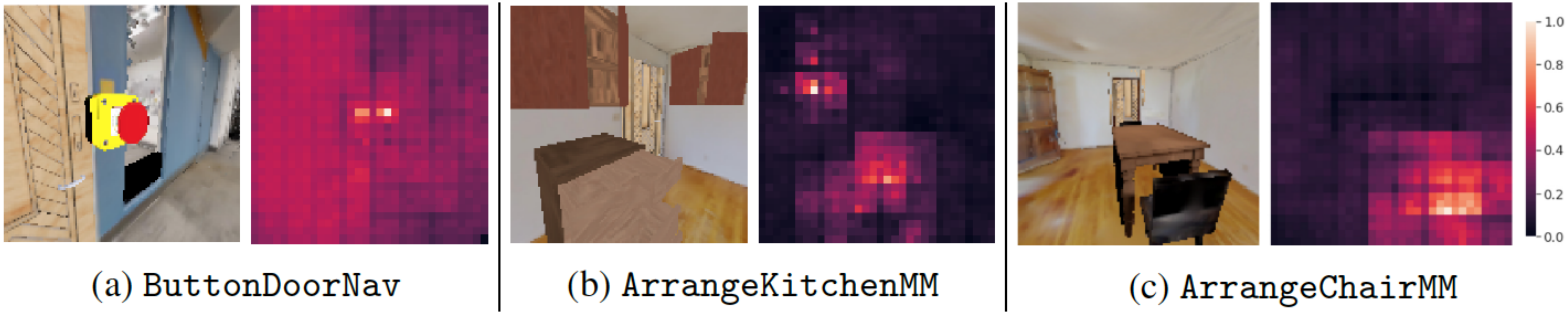
Analysis - Exploration





ReLMoGen has better exploration

Analysis - Interpretability



Visualization of ReLMoGen-D action maps during task execution

Main Contributions

- Proposed ReLMoGen, a framework that combines the strengths of RL and MG.
  - RL: maps observations to subgoals
  - MG: plans for and executes trajectories for subgoals
- Instantiated ReLMoGen with two different RL algorithms: SAC and DQN
- Outperformed baselines across a variety of tasks: (Interactive) Navigation, Mobile Manipulation
- Transfer to new motion planners, potential for real deployment

Do as I Can, Not as I Say (SayCan): Grounding Language In Robotic Affordances

Say-Can.github.io

Presenter: Fei Xia, Research Scientist, Robotics at Google

Robotics at Google

Everyday Robots

Authors

Michael Ahn[*], Anthony Brohan[*], Noah Brown[*], Yevgen Chebotar[*], Omar Cortes[*], Byron David[*], Chelsea Finn[*], Keerthana Gopalakrishnan[*], Karol Hausman[*], Alex Herzog[+], Daniel Ho[+], Jasmine Hsu[*], Julian Ibarz[*], Brian Ichter[*], Alex Irpan[*], Eric Jang[*], Rosario Jauregui Ruano[*], Kyle Jeffrey[*], Sally Jesmonth[*], Nikhil J Joshi[*], Ryan Julian[*], Dmitry Kalashnikov[*], Yuheng Kuang[*], Kuang-Huei Lee[*], Sergey Levine[*], Yao Lu[*], Linda Luu[*], Carolina Parada[*], Peter Pastor[+], Jornell Quiambao[*], Kanishka Rao[*], Jarek Rettinghouse[*], Diego Reyes[*], Pierre Sermanet[*], Nicolas Sievers[*], Clayton Tan[*], Alexander Toshev[*], Vincent Vanhoucke[*], Fei Xia[*], Ted Xiao[*], Peng Xu[*], Sichun Xu[*], Mengyuan Yan[+]

[*]Robotics at Google

[+]Everyday Robot