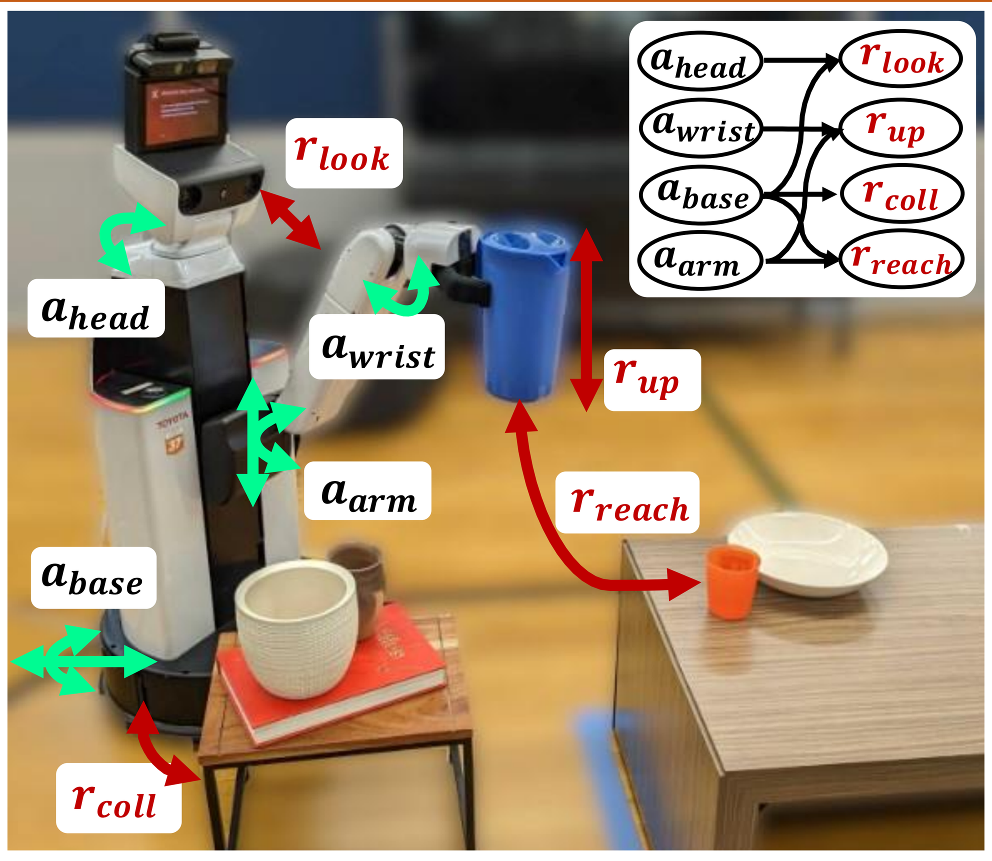
Causal Policy Gradient for Whole-Body Mobile Manipulation

Jiaheng Hu, Peter Stone, Roberto Martín-Martín

Introduction



Reinforcement Learning:

Powerful tool for autonomously learning sensor-motor controls

Mobile Manipulators:

- Locomotion
- Manipulation
- Observation

Problem: learning end-to-end policies for whole-body mobile manipulation is hard

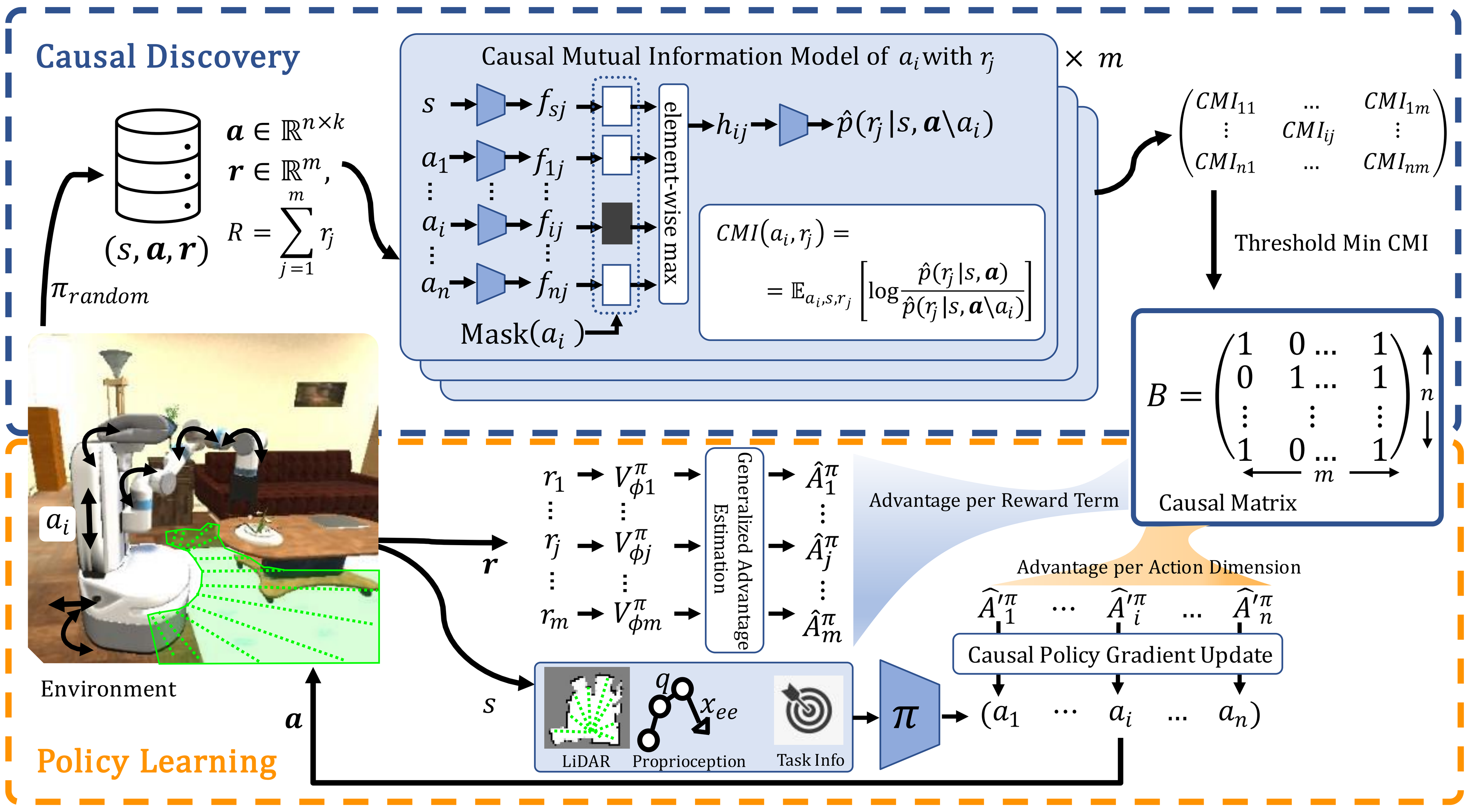
- ➢ Large action space (head, arm, base, …)
- ➢ Composite reward function (grasp, collision avoidance, navigation, …)

Key insight:

- ➢ In mobile manipulation tasks, the causal dependencies between the controllable embodiment (i.e., dimensions of the action space) and the sub-objectives (i.e., elements of the reward signal) are often sparse.

We improve **Reinforcement Learning** for **whole-body** Mobile Manipulation by identifying and making use of causal dependencies between the robot's **action dimensions** and the **reward terms**

Training Pipeline : Causal MoMa



Two-step procedure:

Step 1: Infers the causal dependencies existing between reward terms and action dimensions

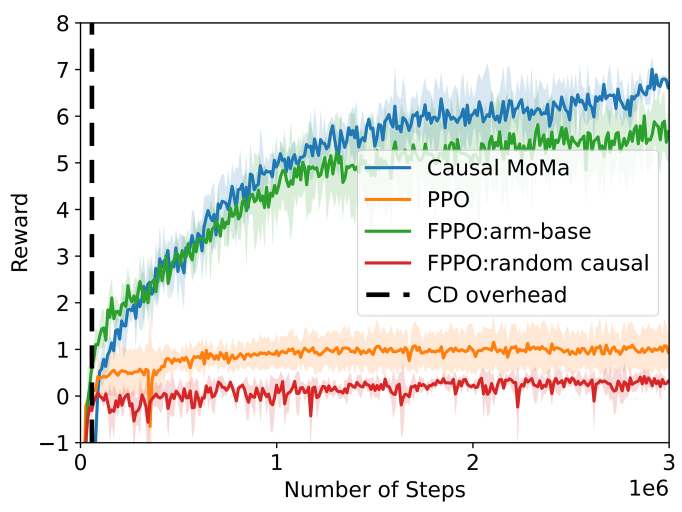
Step 2: Exploiting the discovered Causal Matrix through factored policy gradient to learn whole-body policies

Read more on our project webpage!

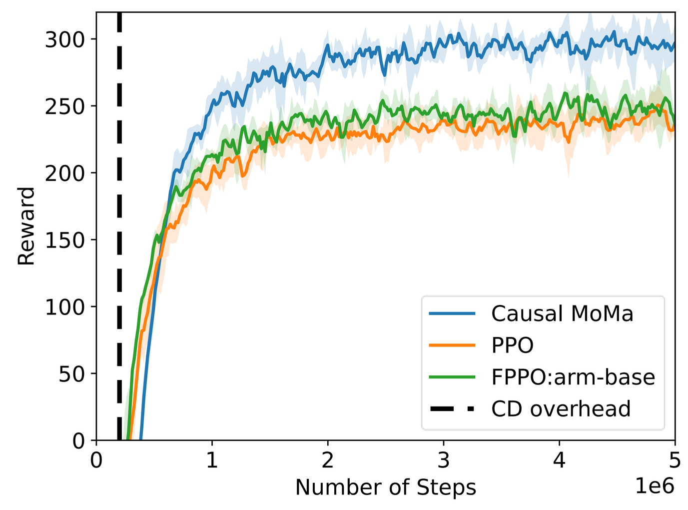
Experiments

Simulation Domains:

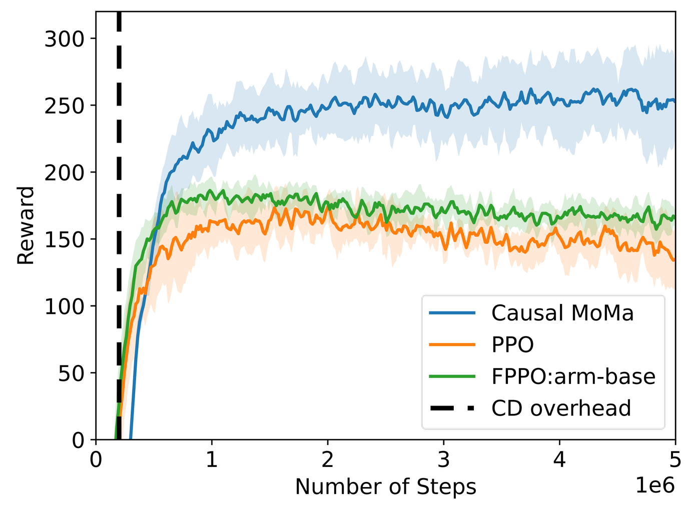
Compared to no-factorization or predefined factorization policy learning, Causal MoMa leads to more robust training and higher return in all domains



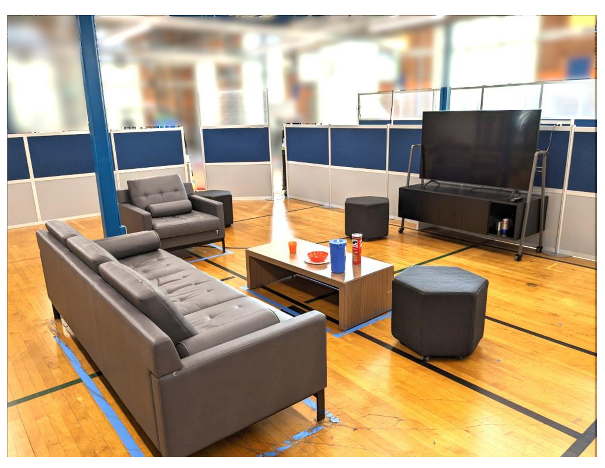
Minigrid

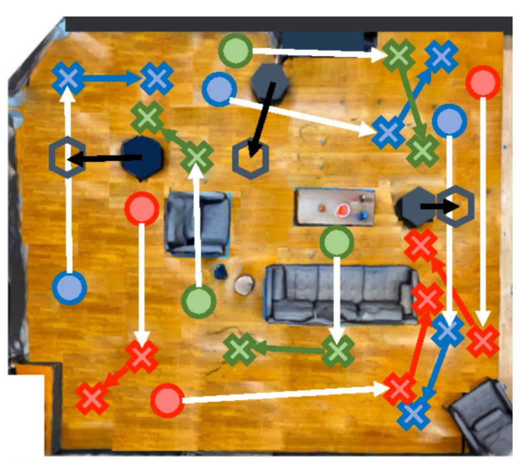


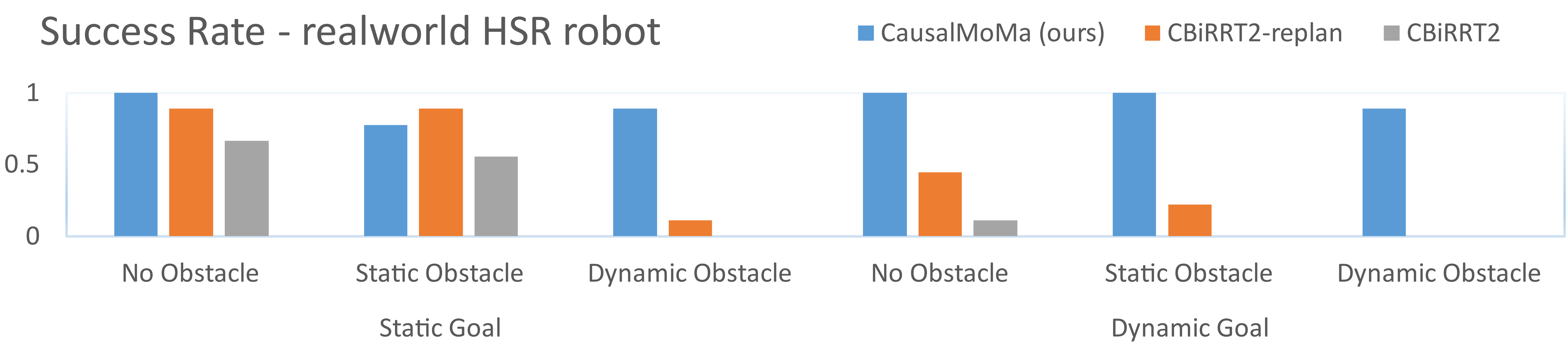
Fetch



HSR







Causal MoMa learned policies transfer sim2real to previously unseen scenes without the need of a model of the environment