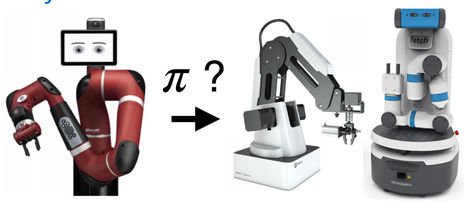
Hardware Conditioned Policies for Multi-Robot Transfer Learning

Tao Chen Adithyavairavan Murali Abhinav Gupta
The Robotics Institute, Carnegie Mellon University

1. Introduction

This work learns generalizable policies over robots with:

- different kinematic structure
- different degrees of freedom
- different dynamics



2. Approach

We propose ***Hardware Conditioned Policies*** (**HCP**) which are conditioned on encoding vectors representing the hardware properties.

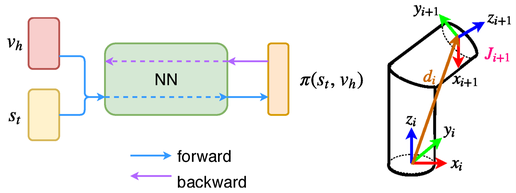
Such vectors can be constructed either explicitly (HCP-E) or implicitly (HCP-I)

Project Website

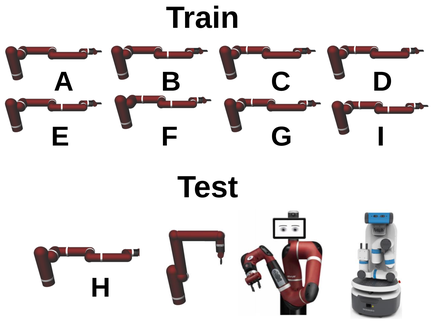
- https://sites.google.com/view/robot-transfer-hcp

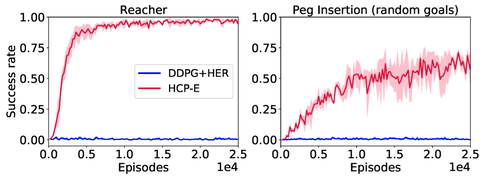
3. HCP-E

- $v_h = P_{-1} \oplus P_0 \oplus \cdots \oplus P_{n-1}$
- P_i is the relative pose between J_i and J_{i+1}

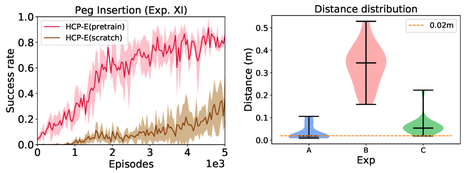


4. Experiments





(a) (b)



(c) (d)

(a) and **(b)**: Learning curves

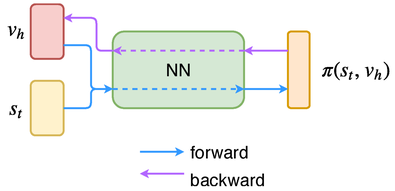
(c): Finetuning on zero-shot transfer failure tasks

(d): Zero-shot Sim-to-Real transfer

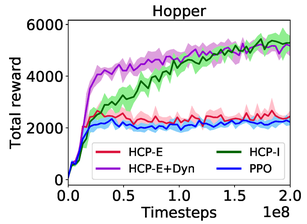
Test Performance (%)	Reacher	Peg Insertion (fixed goal)	Peg Insertion (random goals)
HCP-E	92.50 ± 1.96	92.20 ± 2.75	4.10 ± 1.50
DDPG+HER	0.20 ± 0.40	0.00 ± 0.00	0.00 ± 0.00

5. HCP-I

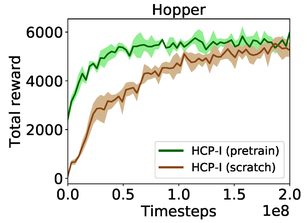
- Learns the dynamics implicitly
- v_h is randomly initialized



6. Experiments



(a)

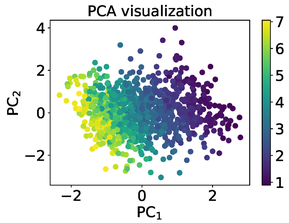


(b)

(a): Learning curves

(b): Transfer learning on new agents

(c): Embedding smoothness



(c)

7. Conclusion

- HCP-E is able to zero-shot transfer to new robots with different kinematics and dynamics as well as sim-to-real transfer
- HCP-I learns a vector representation of robot hardware