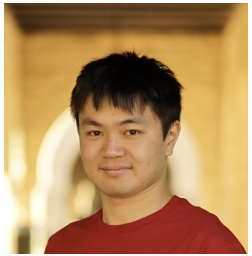
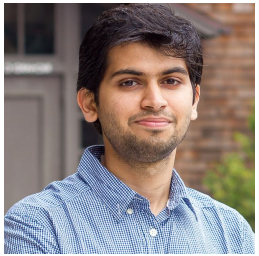
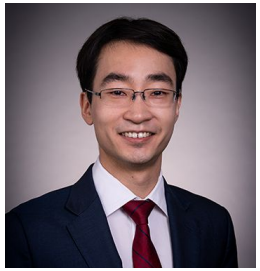
Meta-Evolve: Continuous Robot Evolution for One-to-many Policy Transfer

Xingyu Liu, Deepak Pathak, Ding Zhao

Carnegie Mellon University

ICLR

Carnegie Mellon University

Robotics Industry Created Many Successful Robots

How to train control policies on multiple different robots?

Train Control Policies on Multiple Different Robots

Train from Scratch

Train from Scratch

Train from Scratch

Train from Scratch

Train from Scratch

Train from Scratch

Naive Solution

Train Control Policies on Multiple Different Robots

Train from Scratch

Train from Scratch

Train from Scratch

Train from Scratch

Train from Scratch

Train from Scratch

Train from Scratch

Naive Solution

Proposed Solution: Policy Transfer

Train Control Policies on Multiple Different Robots

Train from Scratch

Train from Scratch

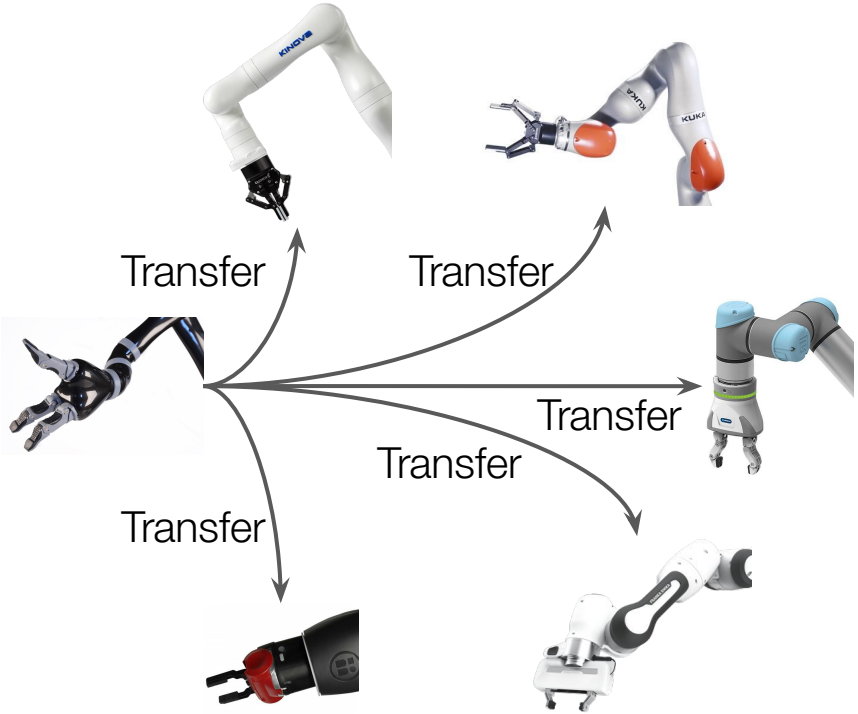
Train from Scratch

Train from Scratch

Train from Scratch

Train from Scratch

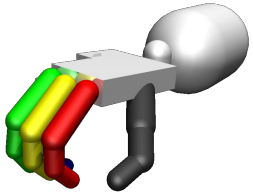
Naive Solution

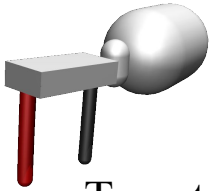


Proposed Solution: Policy Transfer

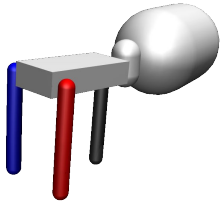
One-to-Many Robot-to-robot Policy Transfer: How to?

Source Robot

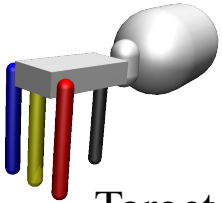




Target Robot 1



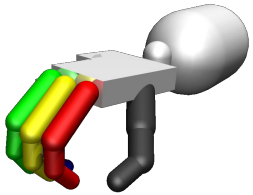
Target Robot 2



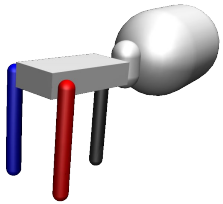
Target Robot *N*

One-to-One Robot-to-robot Policy Transfer: How to?

Source Robot



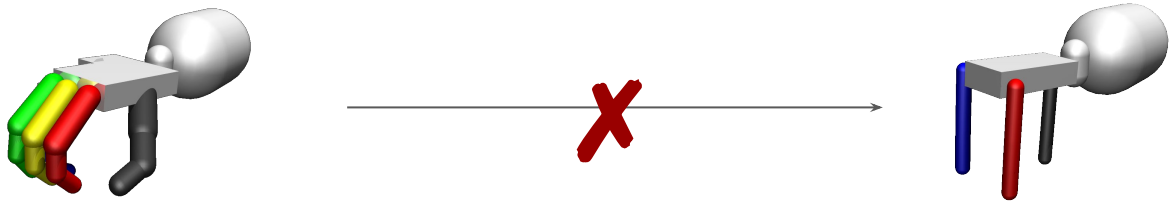
Target Robot



One-to-One Robot-to-robot Policy Transfer: Imitation Learning?

Source Robot

Target Robot

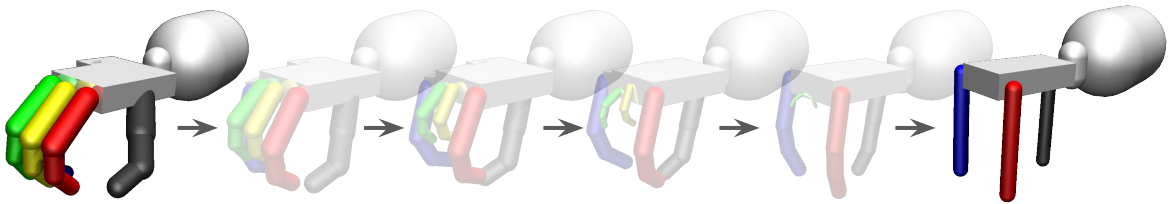


Different MDP dynamics, cannot directly transfer

One-to-One Robot-to-robot Policy Transfer: Continuous Robot Evolution

Source Robot

Target Robot



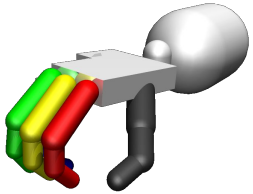
✔

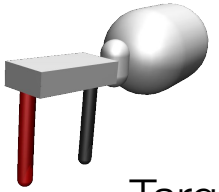
Interpolate robot morphology and transfer policy

Liu et al., “REvolveR: Continuous Evolutionary Models for Robot-to-robot Policy Transfer”, ICML 2022 **(Long Oral)**
Liu et al., “HERD: Continuous Human-to-Robot Evolution for Learning from Human Demonstration”, CoRL 2022

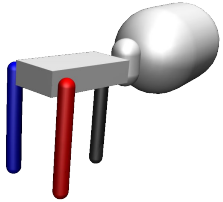
One-to-Many Robot-to-robot Policy Transfer

Source Robot

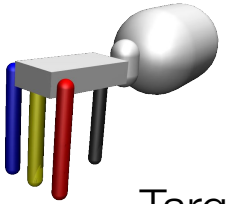




Target Robot 1

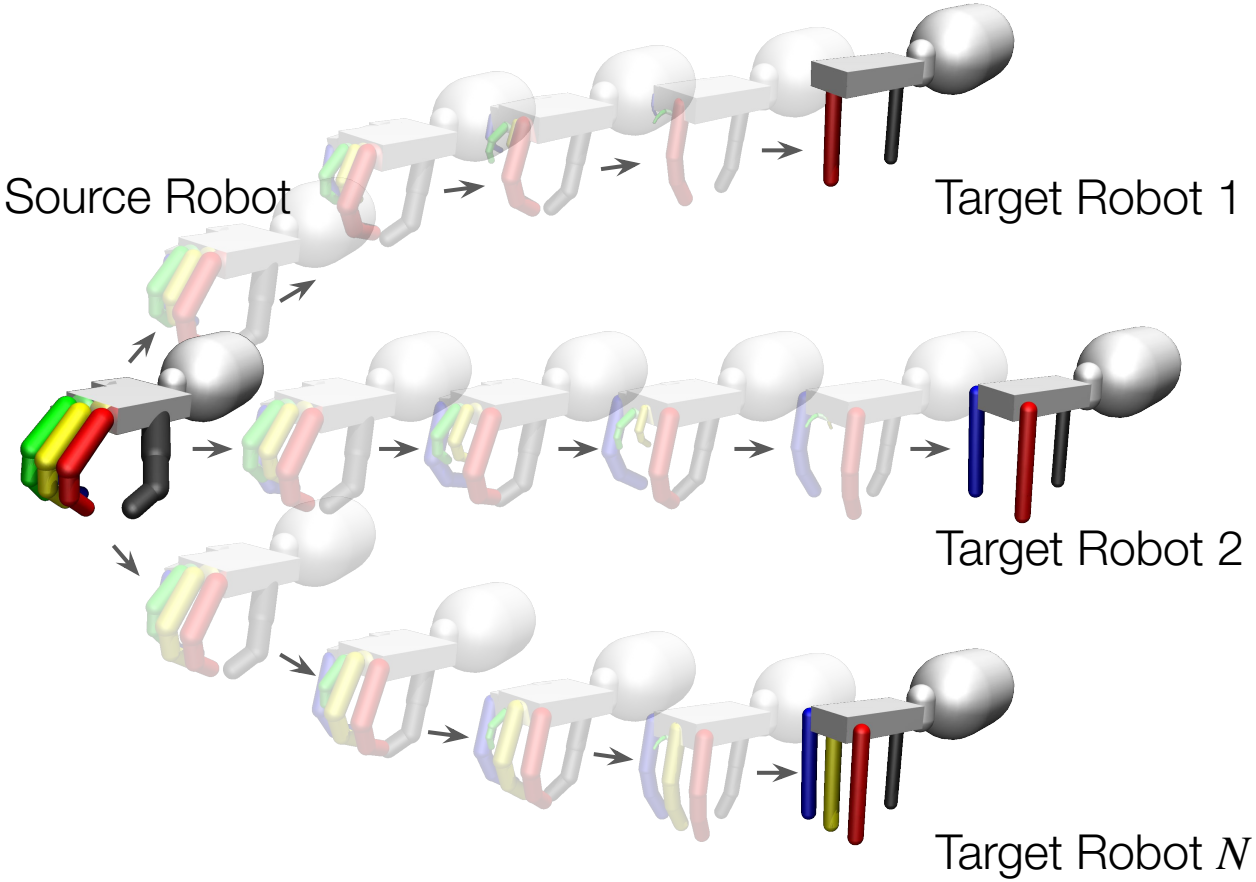


Target Robot 2



Target Robot N

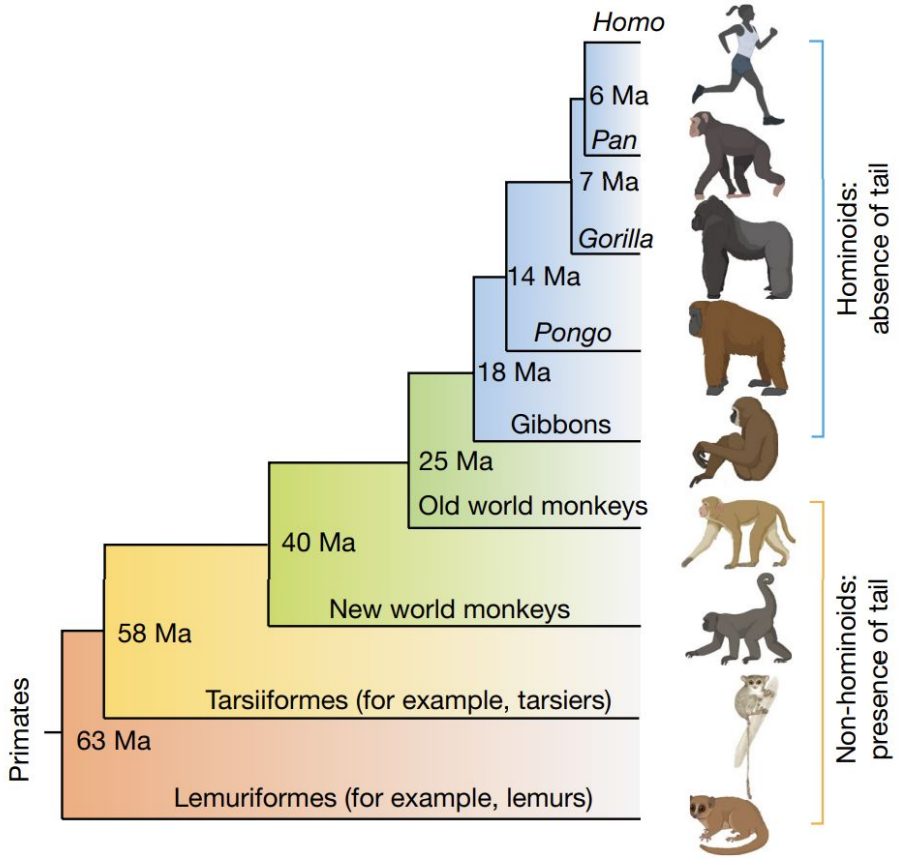
One-to-Many Policy Transfer via Continuous Evolution: Vanilla Solution



One-to-Many Policy Transfer via Continuous Evolution: Vanilla Solution

Can we do better than that?

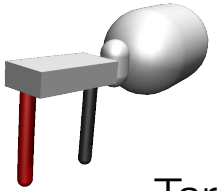
Evolution History: Creatures with Similar Morphology Share Same Ancestors



Robot evolution paths may also be shared!

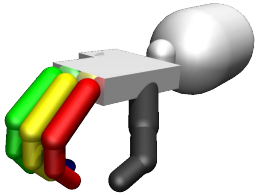
Xia et al., “The genetic basis of tail-loss evolution in humans and apes”, Nature, 2024

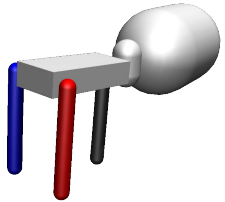
One-to-Many Policy Transfer via Continuous Evolution: **Evolution Tree**



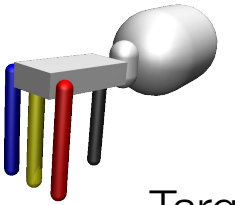
Target Robot 1

Source Robot





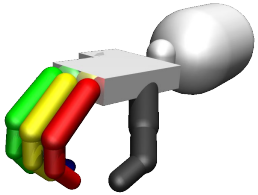
Target Robot 2



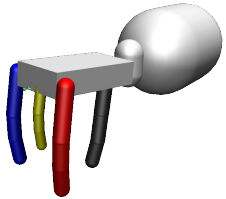
Target Robot N

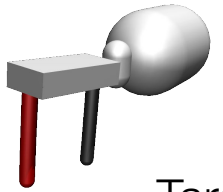
One-to-Many Policy Transfer via Continuous Evolution: **Evolution Tree**

Source Robot

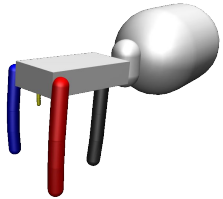


Meta Robot 1

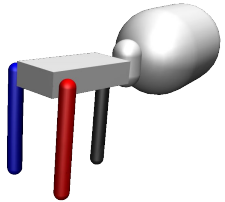




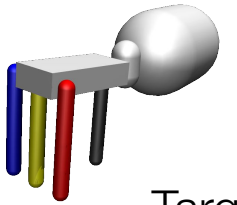
Target Robot 1



Meta Robot 2

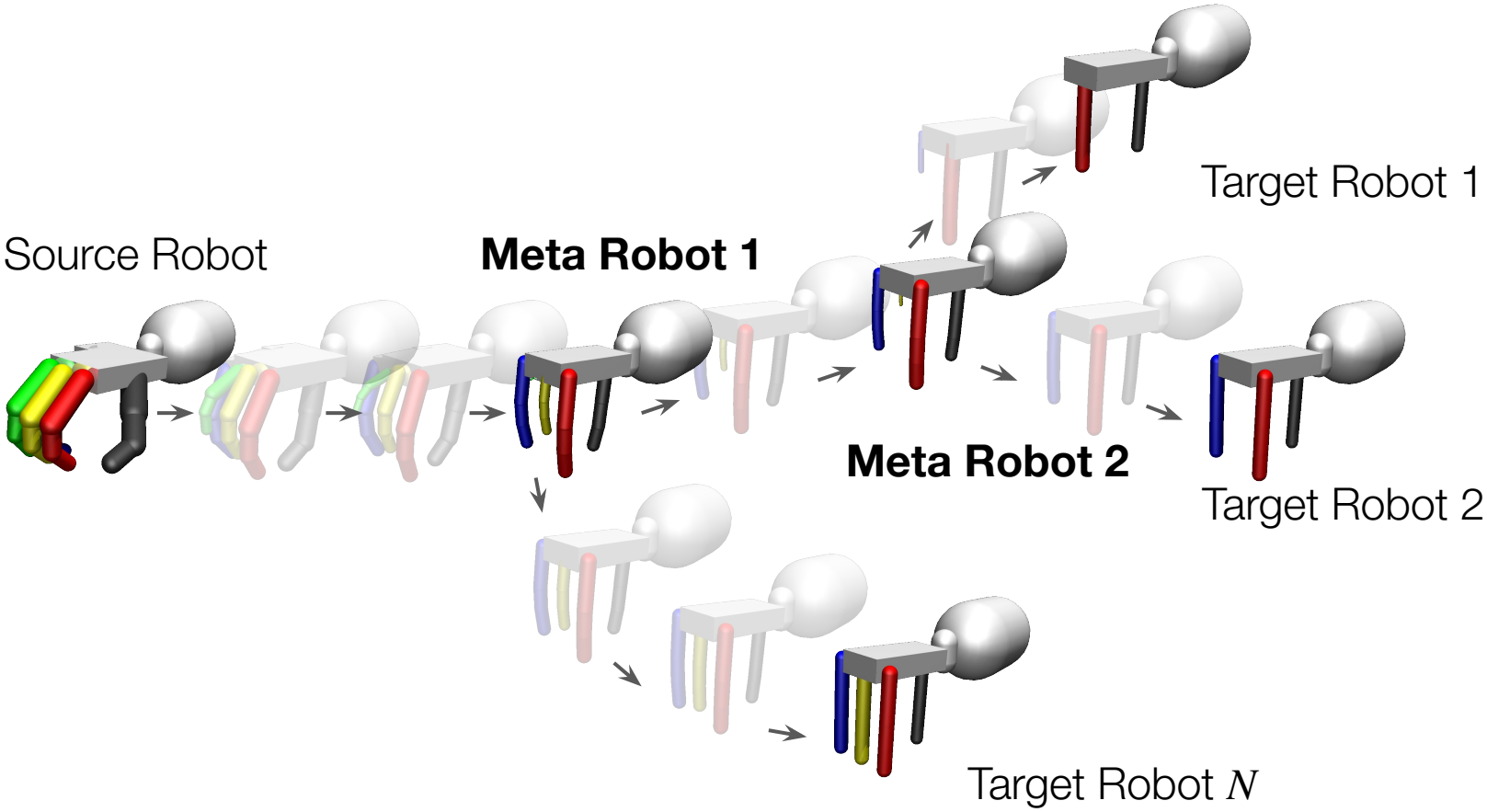


Target Robot 2

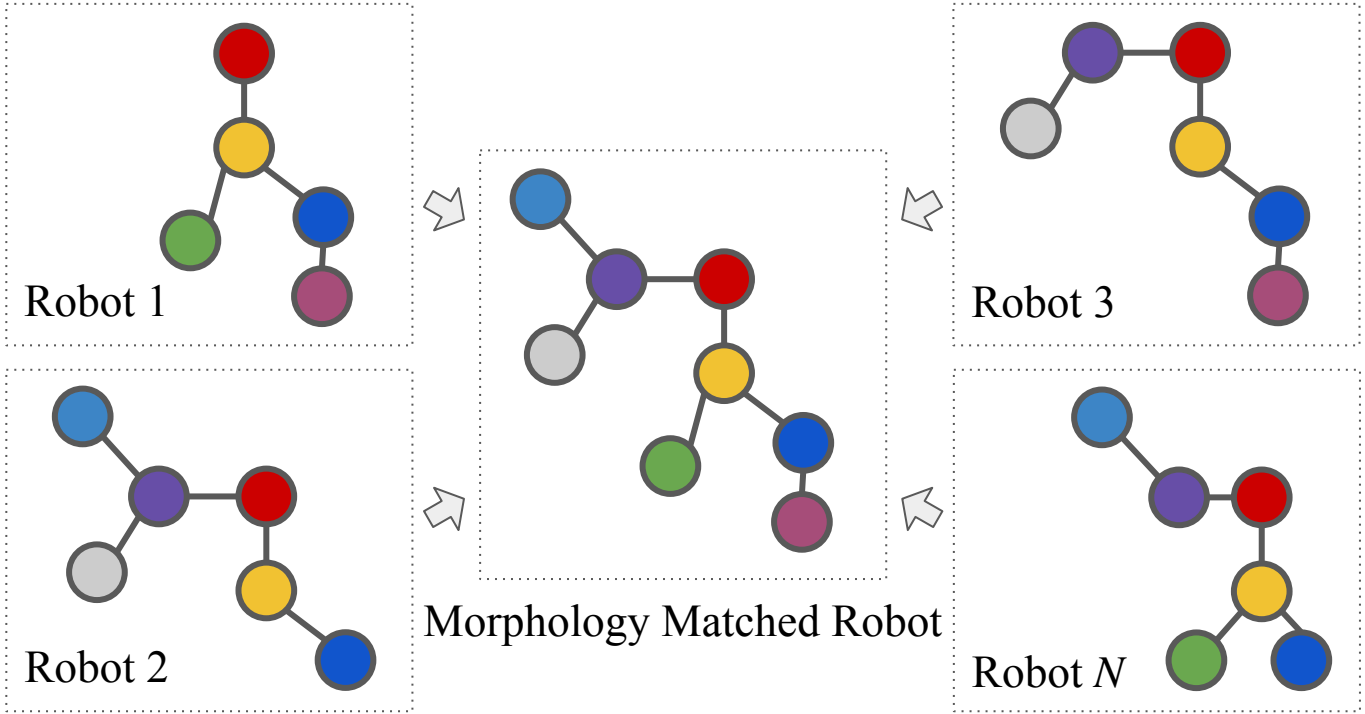


Target Robot N

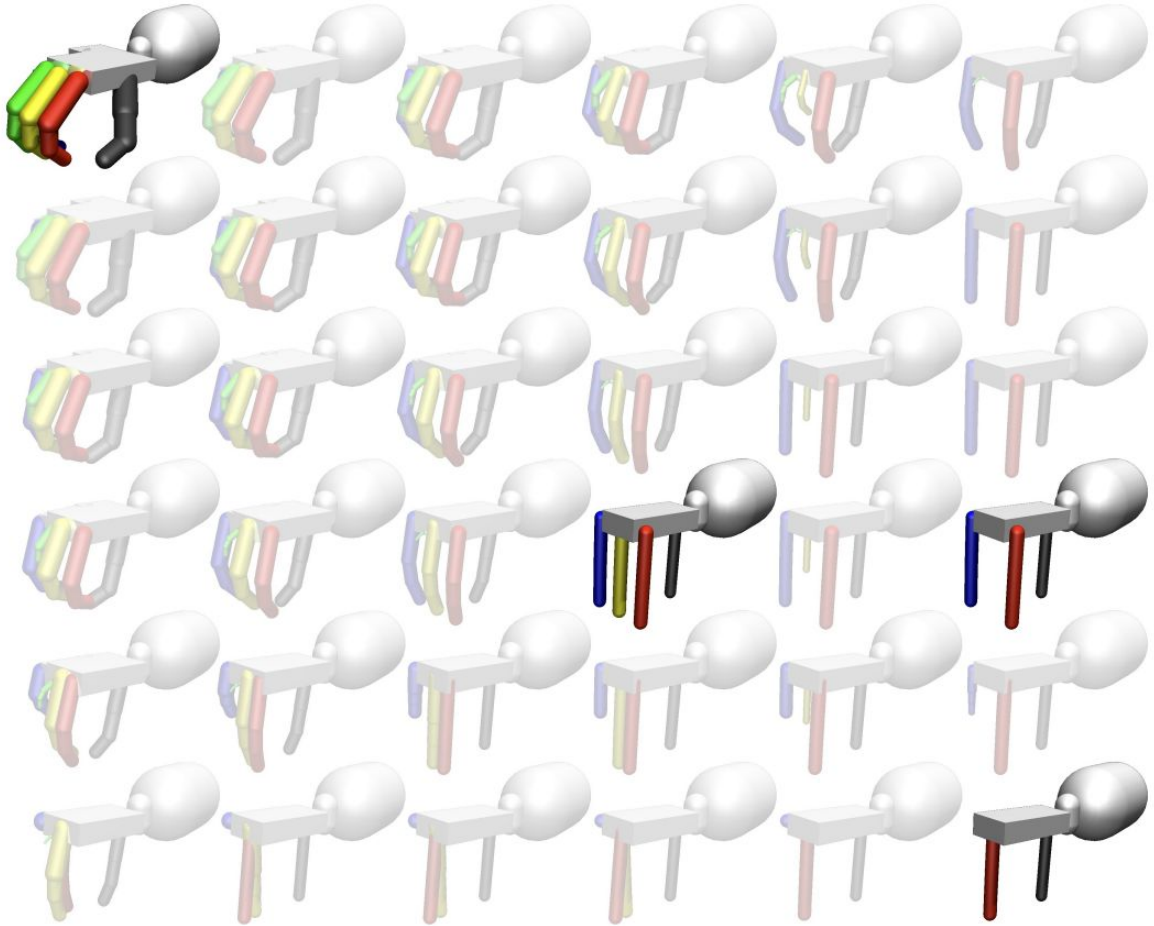
One-to-Many Policy Transfer via Continuous Evolution: **Evolution Tree**



Evolution Tree Implementation Step 1: Kinematic Structure Matching



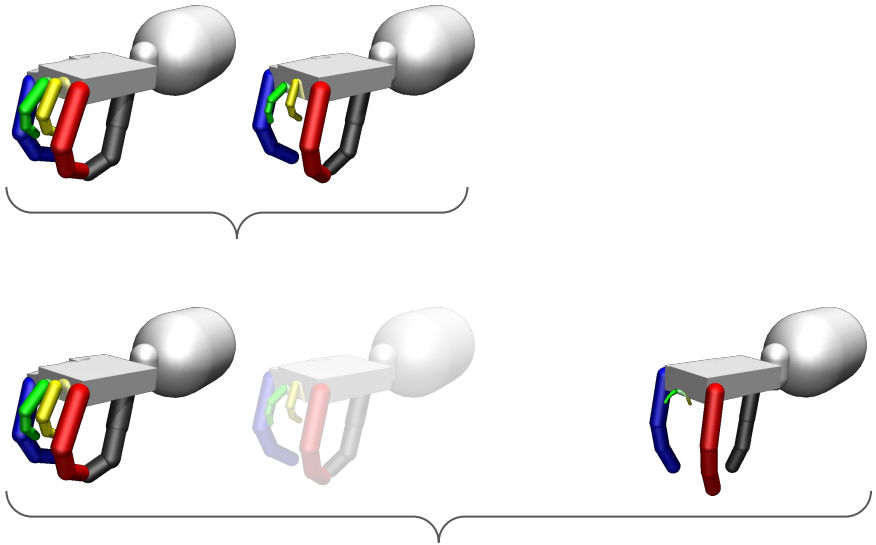
Evolution Tree Implementation Step 2: Physical Parameter Interpolation



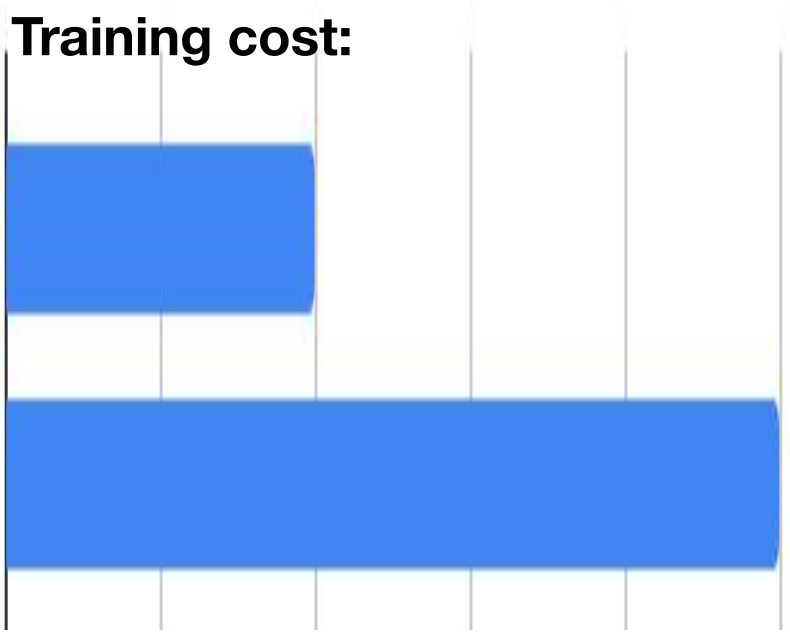
How to Compute Evolution Tree?

Assumption: policy transfer training cost locally proportional to distribution difference of the MDP transition dynamics, and locally proportional to the robot hardware difference measured in vector L_p distance

Hardware difference:



Training cost:

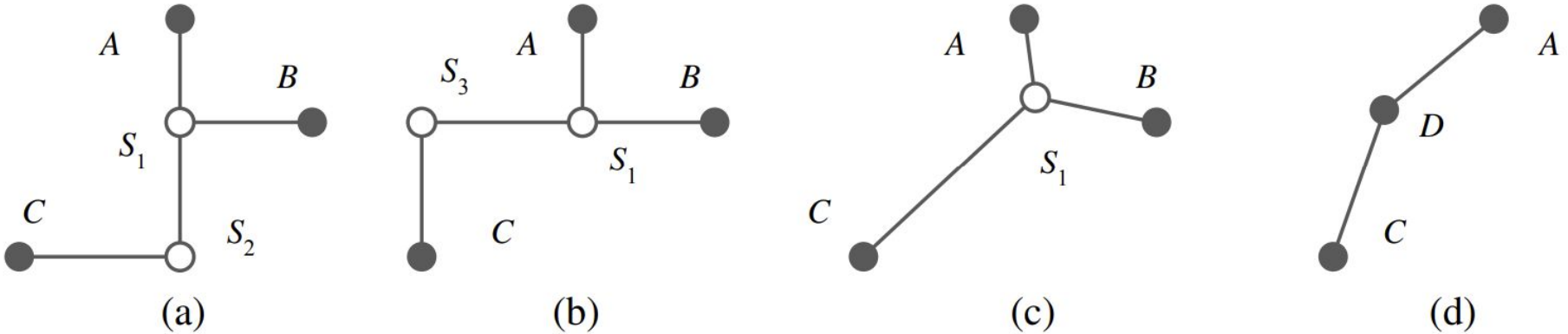


Evolution Tree Implementation

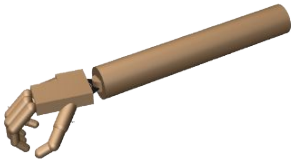
Heuristics: aim to minimize the total L_p distance in robot evolution parameter space

Mathematically: undirected graph that interconnects a set of points and minimizes total L_p travel distance is the p**-Steiner tree**

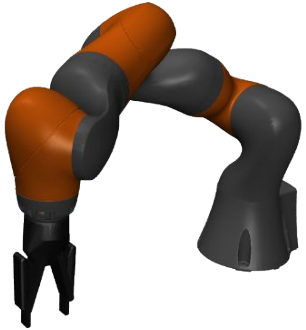
$$(\boldsymbol{V}_{\mathrm{ST}}, \boldsymbol{E}_{\mathrm{ST}}) = \underset{(\boldsymbol{V},\boldsymbol{E}):\kappa((\boldsymbol{V},\boldsymbol{E}))=1,\{\boldsymbol{\beta}_0,\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_N\}\subseteq\boldsymbol{V}}{\arg\min} \sum_{(\boldsymbol{v}_1,\boldsymbol{v}_2)\in\boldsymbol{E}} ||\boldsymbol{v}_1 - \boldsymbol{v}_2||_p$$



Realistic Implementation on **Real** Commercial Robots



Source Robot



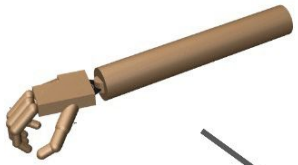
Target Robot 2

Target Robot 1

Target Robot 3

Liu et al., “Meta-Evolve: Continuous Robot Evolution for One-to-many Policy Transfer”, ICLR 2024

Realistic Implementation on **Real** Commercial Robots

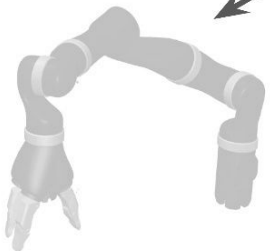


Source Robot

Target Robot 2

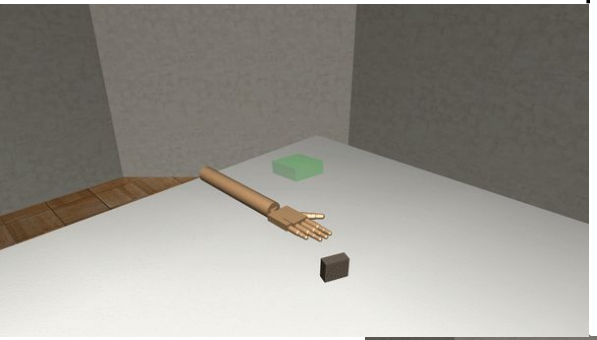
Target Robot 1

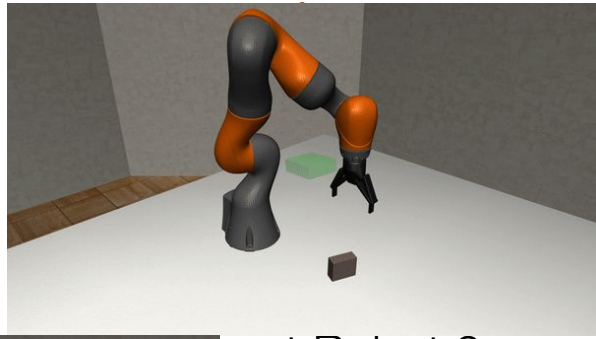
Target Robot 3



Liu et al., “Meta-Evolve: Continuous Robot Evolution for One-to-many Policy Transfer”, ICLR 2024

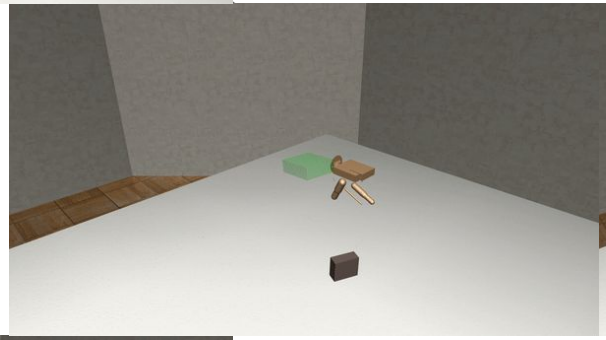
Realistic Implementation on **Real** Commercial Robots

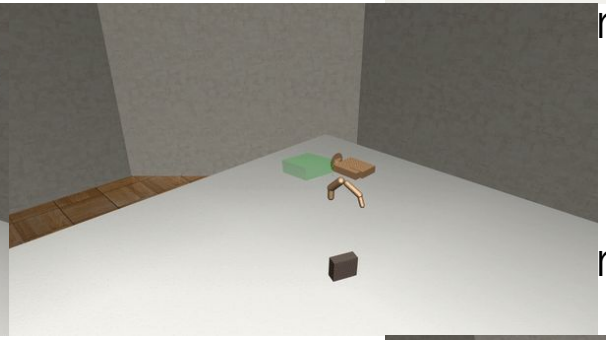




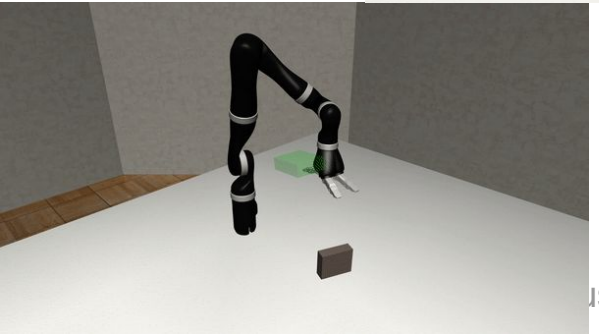
rget Robot 2

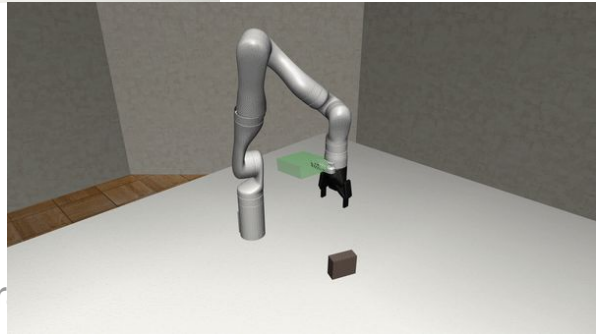
Target Robot



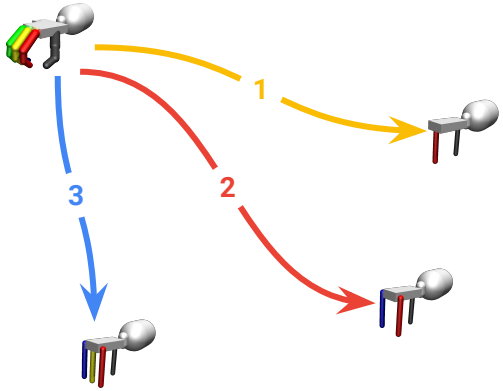


rget Robot 3

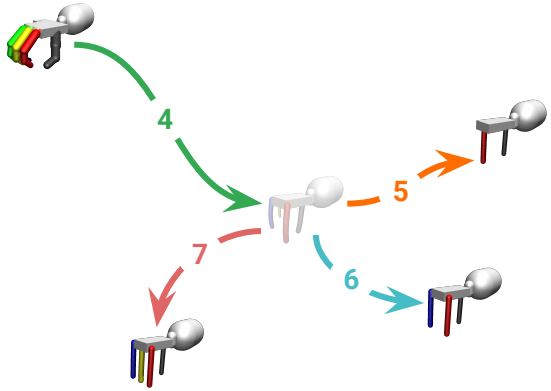




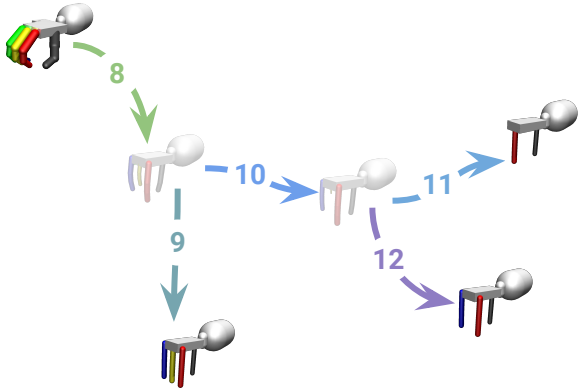
Quantitative Experiment Results



HERD

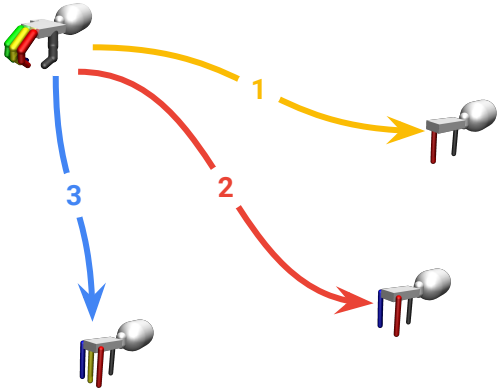


Geom-Med

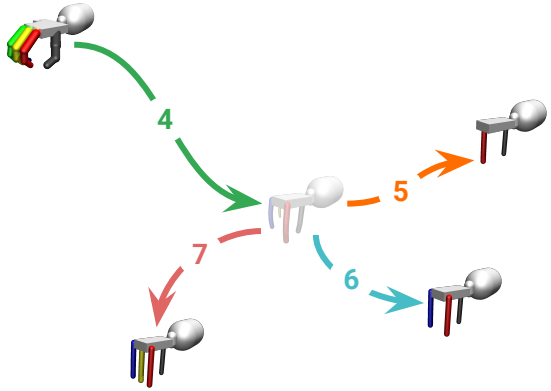


Steiner (Ours)

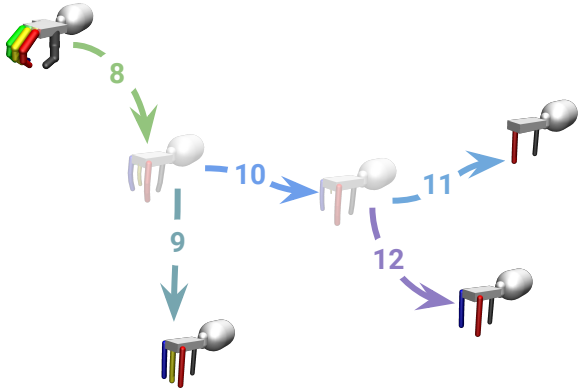
Quantitative Experiment Results



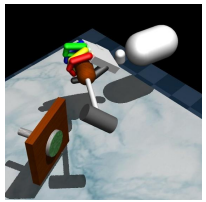
HERD



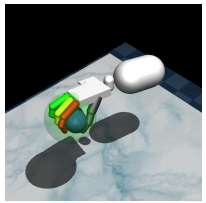
Geom-Med



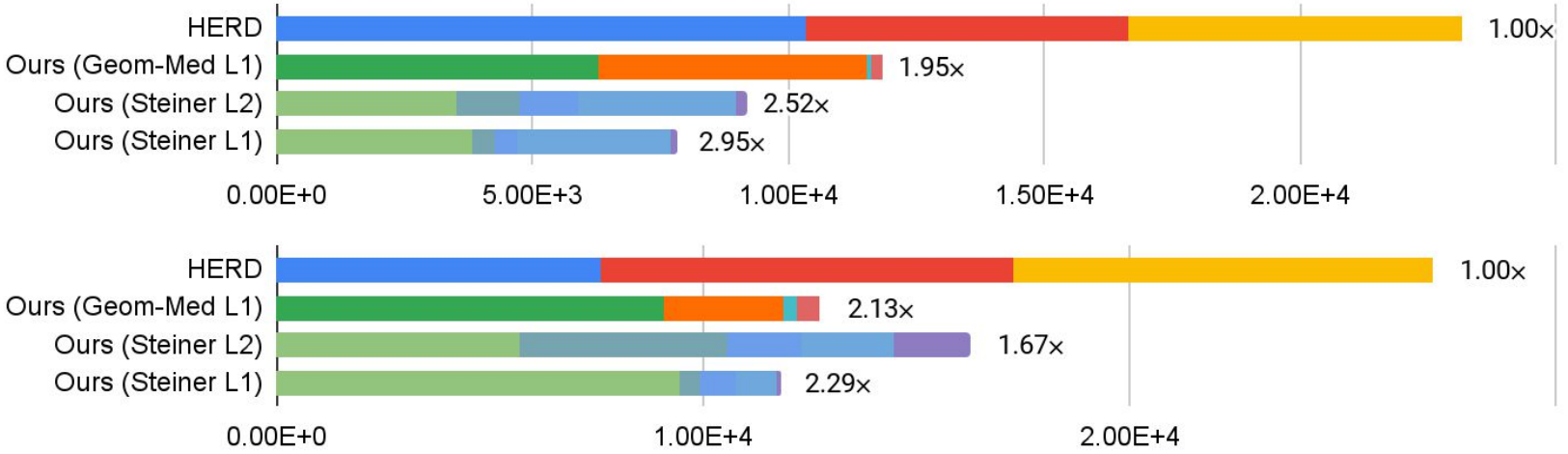
Steiner (Ours)



Hammer



Relocate



https://sites.google.com/view/meta-evolve