Meta-Reinforcement Learning of Structured Exploration Strategies

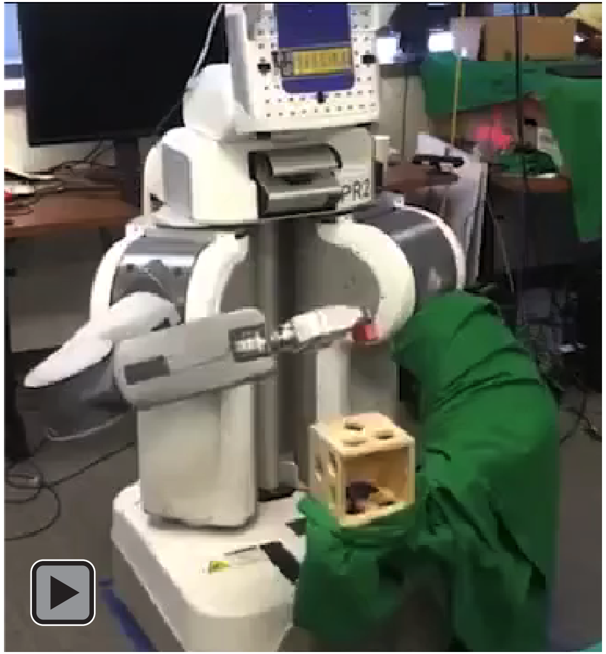
Abhishek Gupta, Russell Mendonca, YuXuan Liu, Pieter Abbeel, Sergey Levine

BAIR

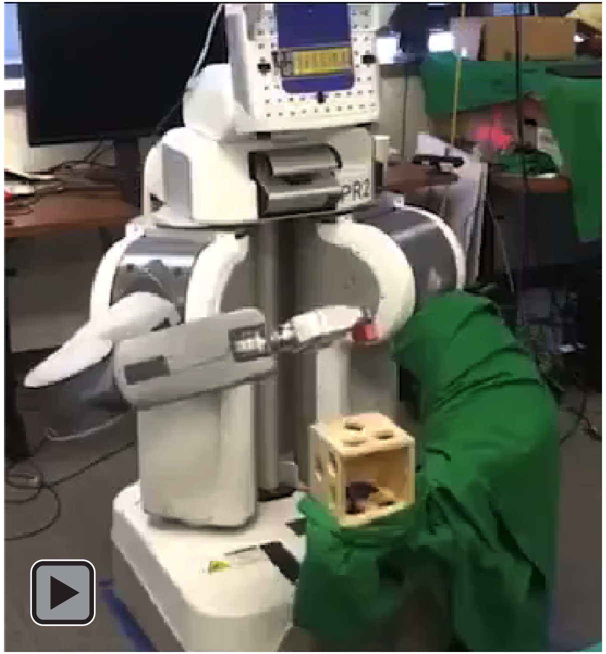
BERKELEY ARTIFICIAL INTELLIGENCE RESEARCH

Human Exploration vs Robot Exploration

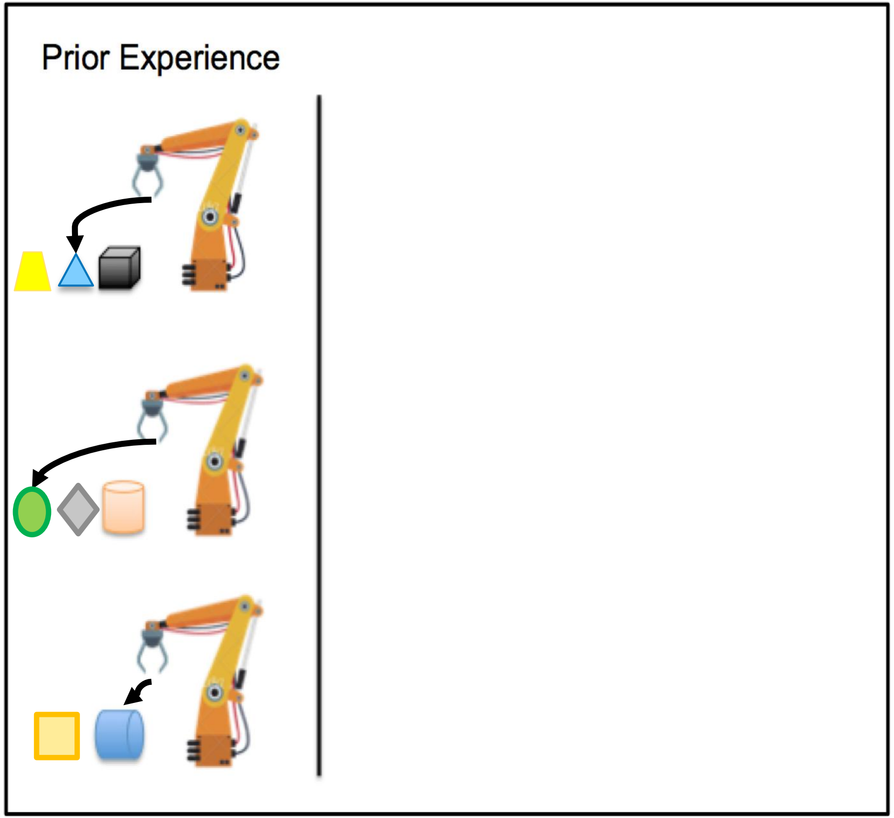
Human Exploration vs Robot Exploration



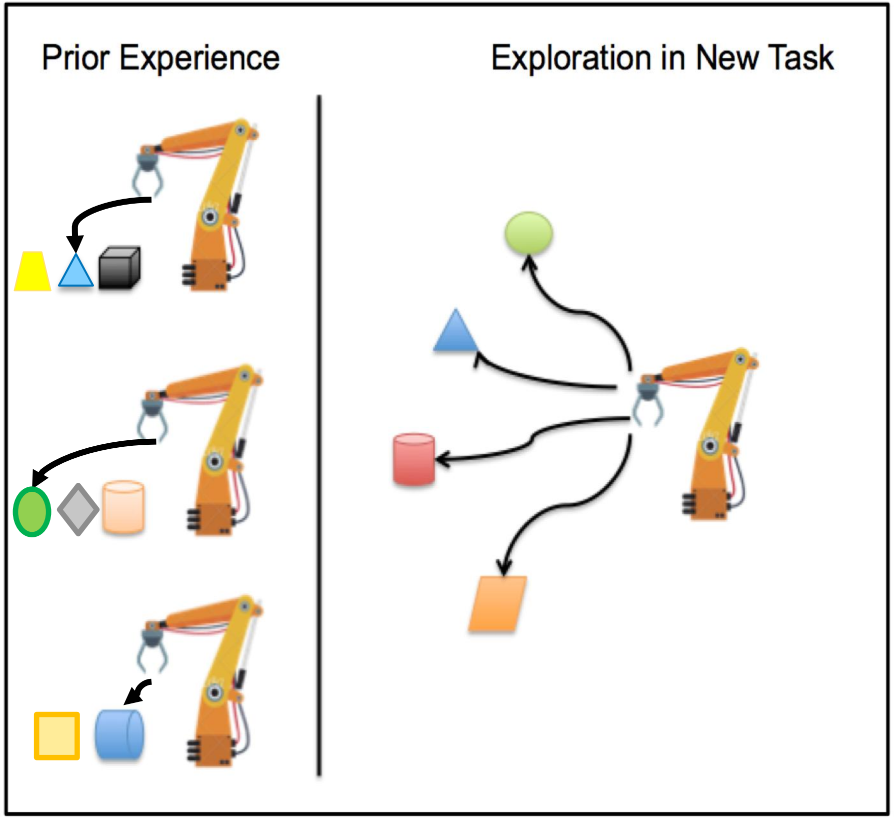
Human Exploration vs Robot Exploration



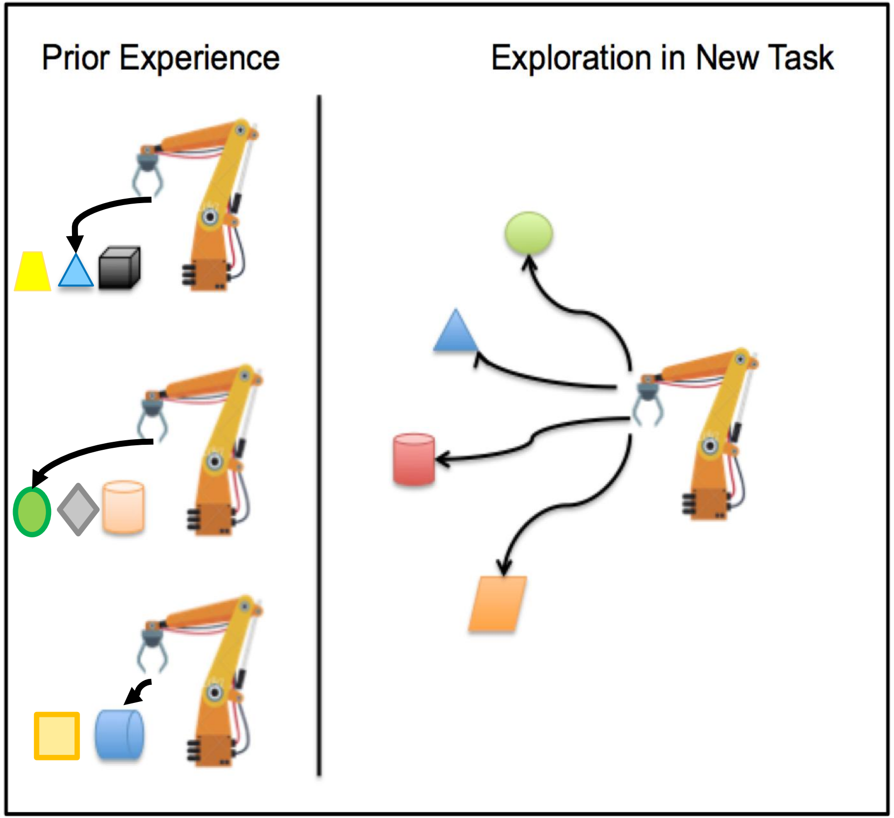
Exploration Informed by Prior Experience



Exploration Informed by Prior Experience



Exploration Informed by Prior Experience



Desired:

- Effective exploration for sparse rewards
- Quick adaptation for new tasks

Key Insights in MAESN

1. Explore with random but structured behaviors (exploration)

Key Insights in MAESN

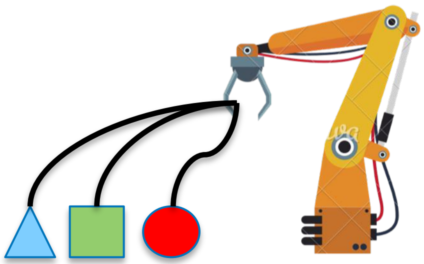
1. Explore with random but structured behaviors (exploration)

2. Explicitly train for quick learning on new tasks (adaptation)

Key Insights in MAESN

1. Explore with random but structured behaviors (exploration)

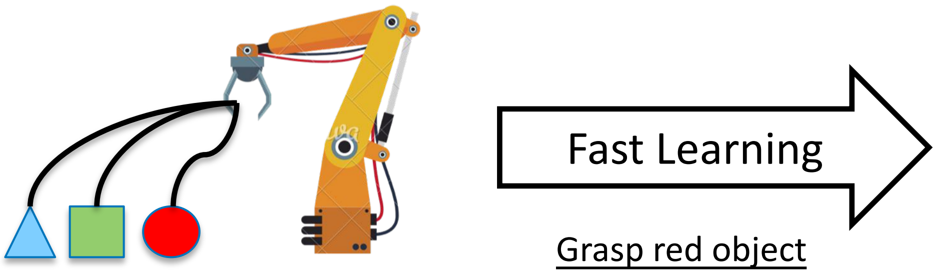
2. Explicitly train for quick learning on new tasks (adaptation)



Key Insights in MAESN

1. Explore with random but structured behaviors (exploration)

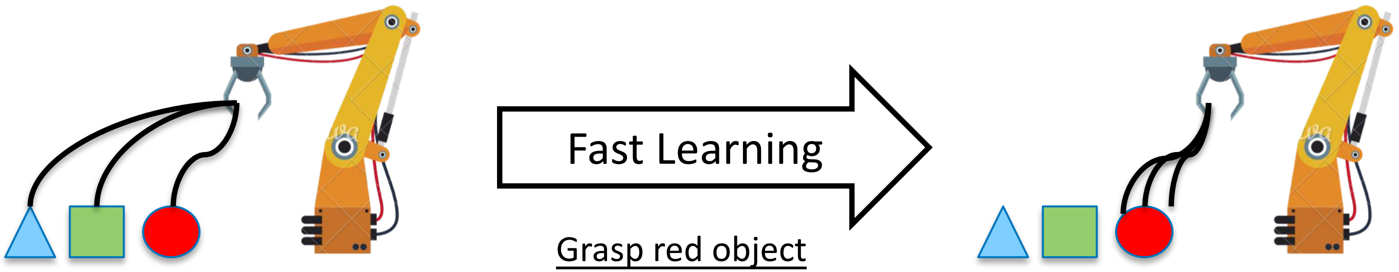
2. Explicitly train for quick learning on new tasks (adaptation)



Key Insights in MAESN

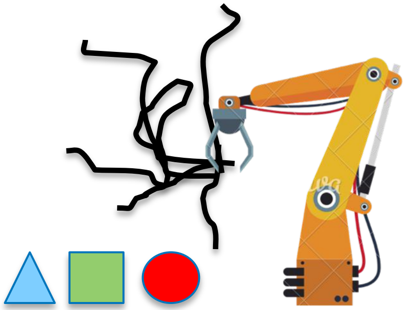
1. Explore with random but structured behaviors (exploration)

2. Explicitly train for quick learning on new tasks (adaptation)

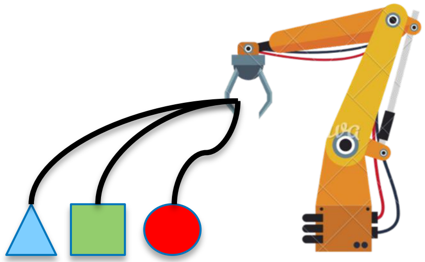


Using Structured Stochasticity

Per-timestep Exploration



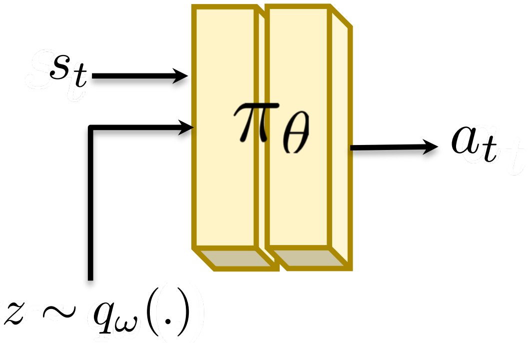
Structured Exploration



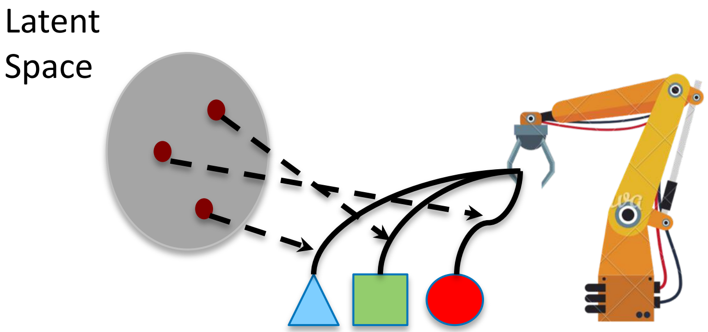
Structured exploration: pick an intention, execute for entire episode.
Explore across different intentions

Latent Conditioned Policies

Structured stochasticity introduced through latent conditioned policy

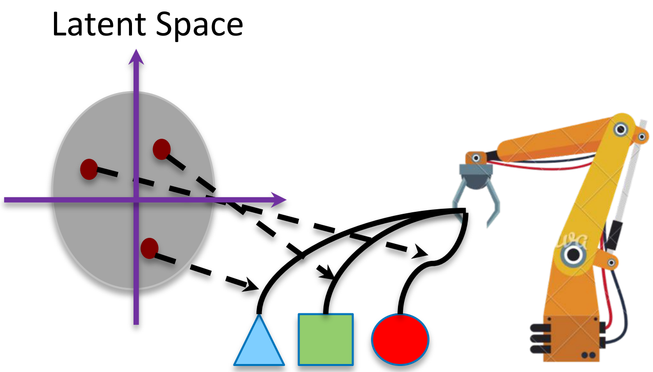


Train latent space to capture prior task distribution



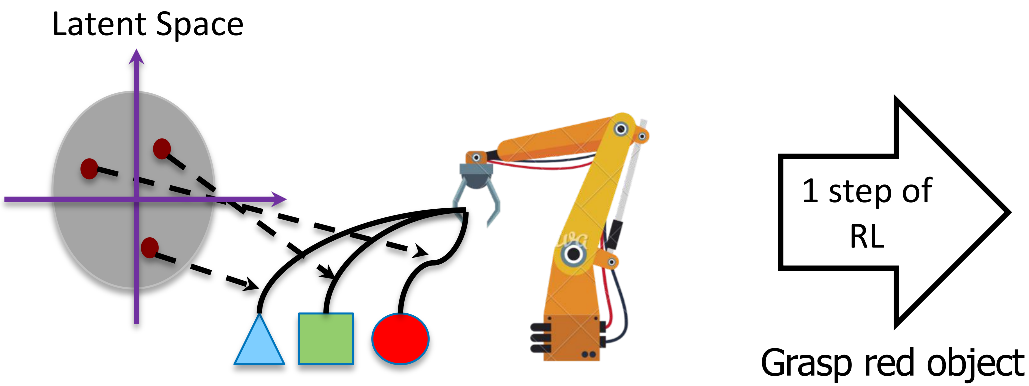
Meta-Training Latent Spaces

Beyond capturing task distribution, train for quick adaptation via meta-learning



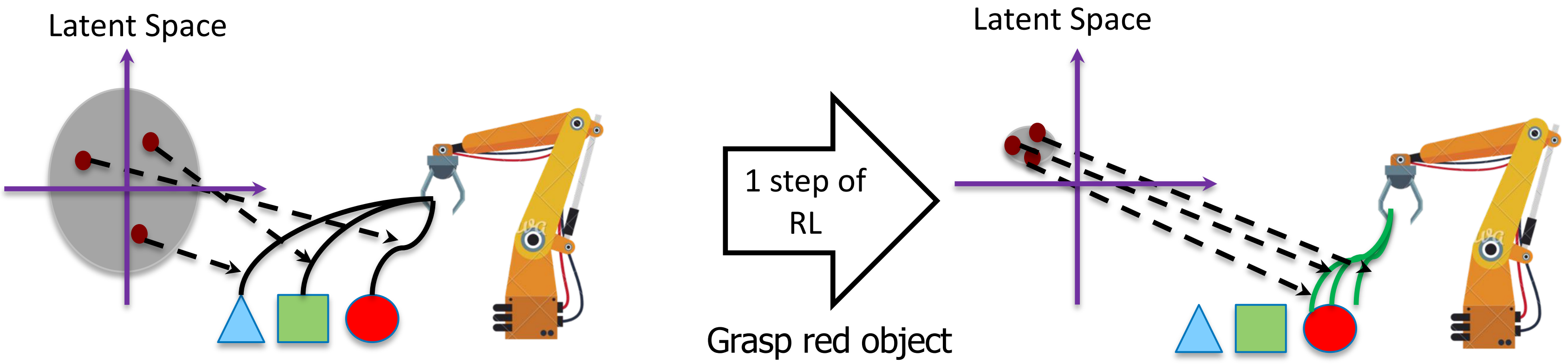
Meta-Training Latent Spaces

Beyond capturing task distribution, train for quick adaptation via meta-learning



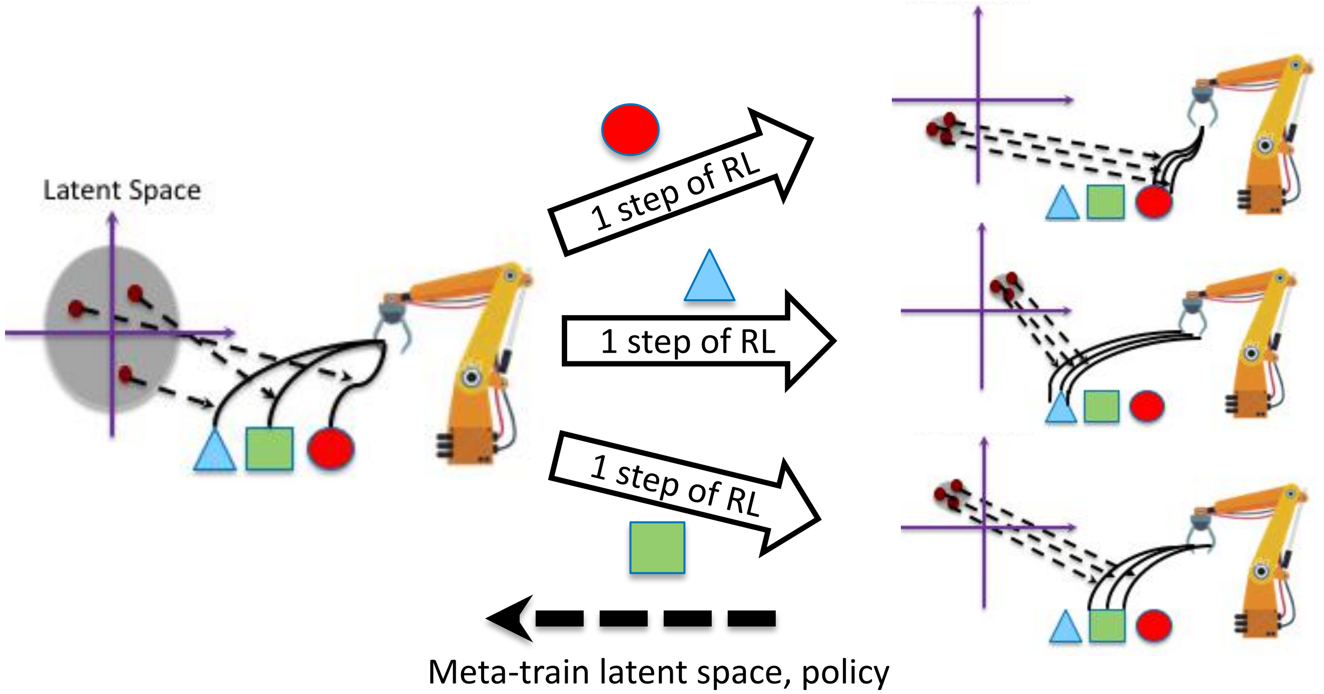
Meta-Training Latent Spaces

Beyond capturing task distribution, train for quick adaptation via meta-learning



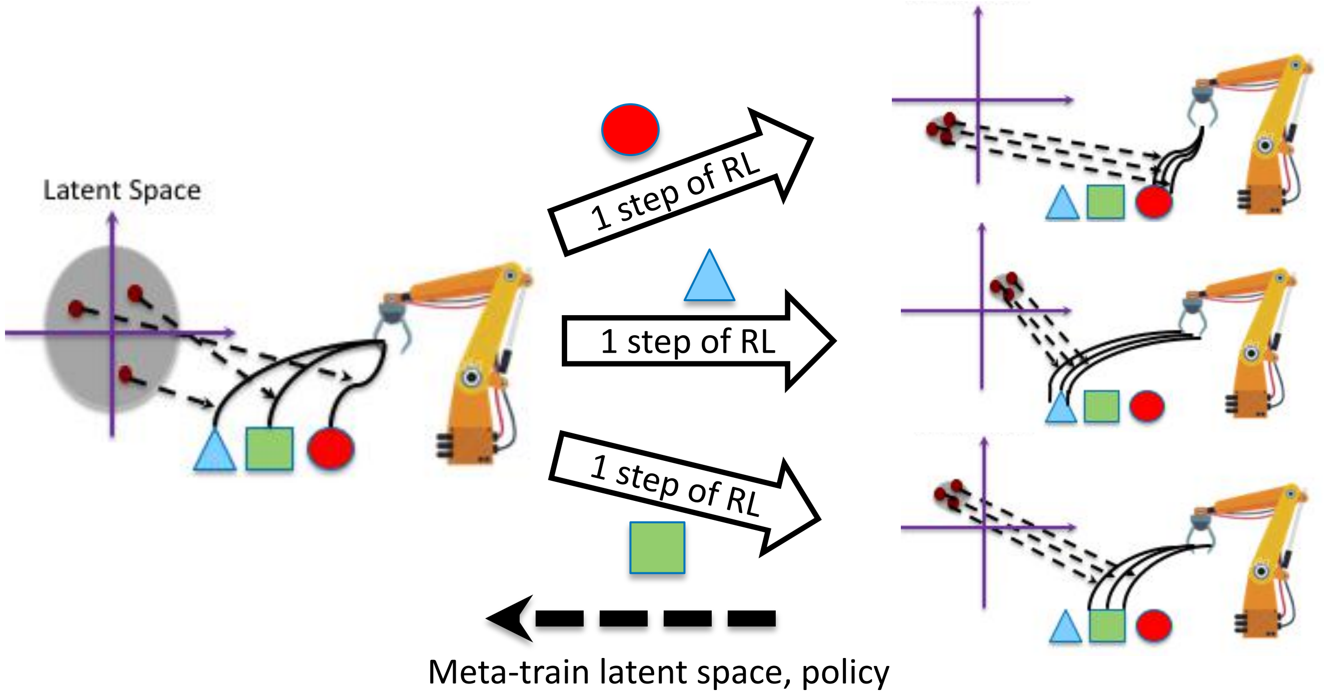
Meta-Training Latent Spaces

Beyond capturing task distribution, train for quick adaptation via meta-learning



Meta-Training Latent Spaces

Beyond capturing task distribution, train for quick adaptation via meta-learning

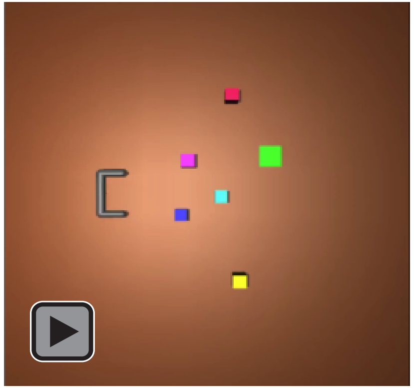


Train with algorithm based on Model Agnostic Meta-Learning[1]

Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks, Finn et al ICML 2017

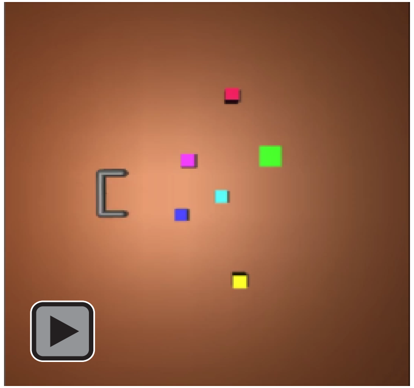
Experiments: Robotic Manipulation

Random Exploration

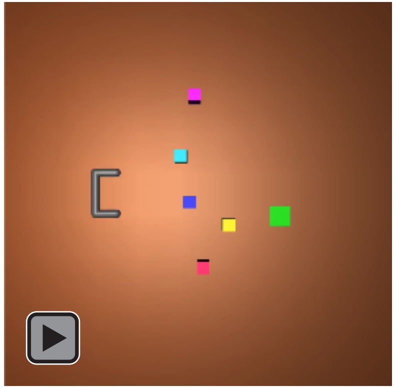


Experiments: Robotic Manipulation

Random Exploration

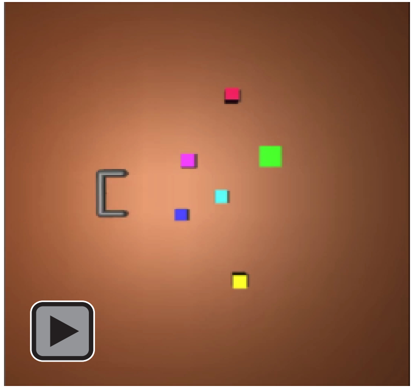


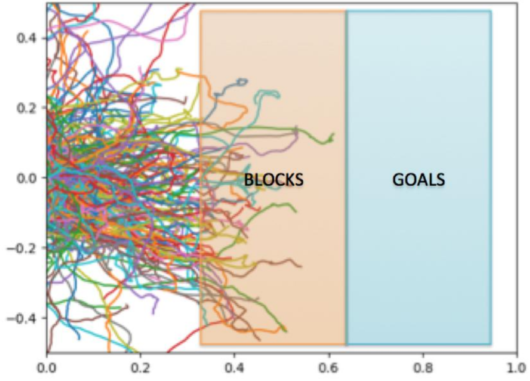
MAESN exploration



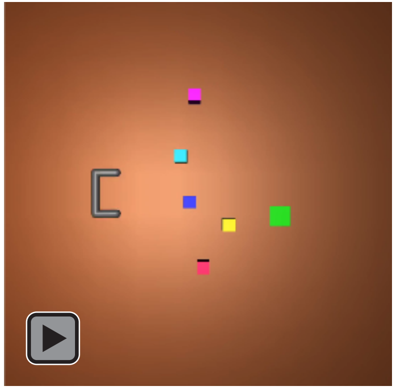
Experiments: Robotic Manipulation

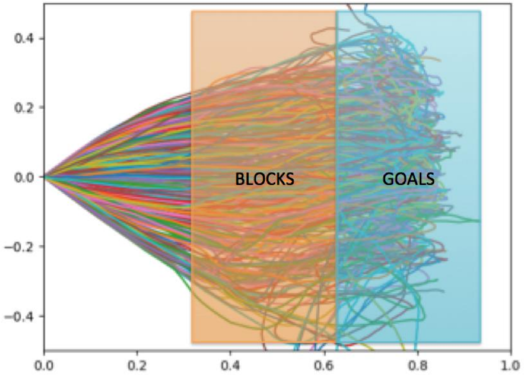
Random Exploration





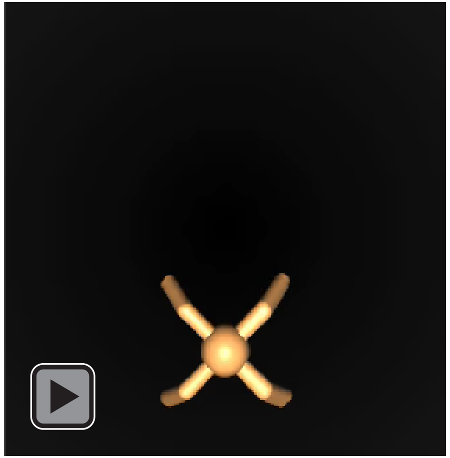
MAESN exploration





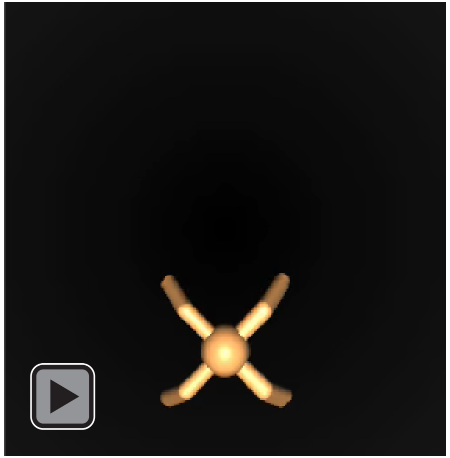
Experiments: Legged Locomotion

Random Exploration

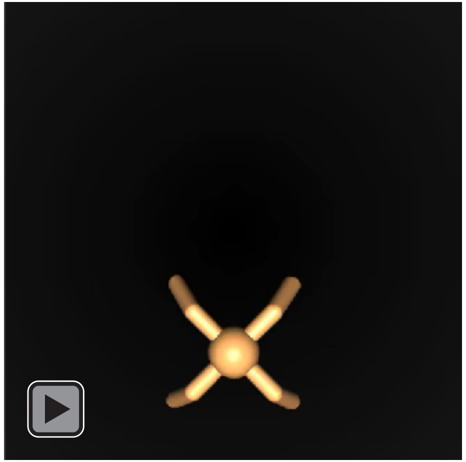


Experiments: Legged Locomotion

Random Exploration

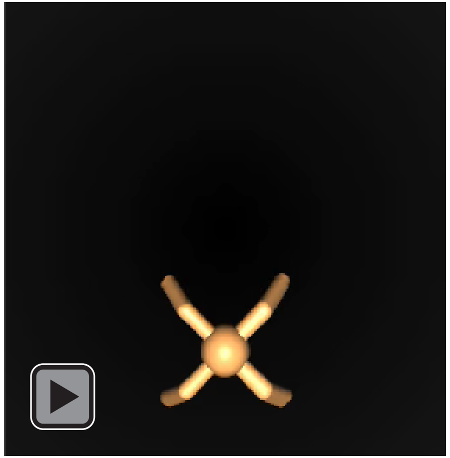


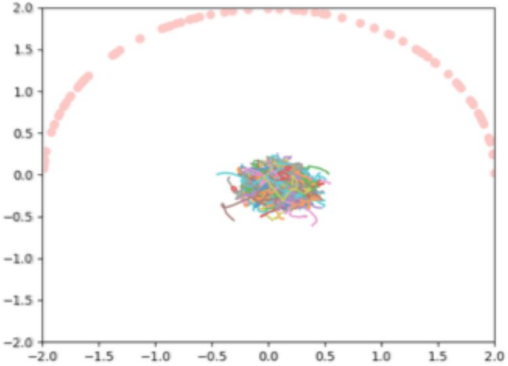
MAESN exploration



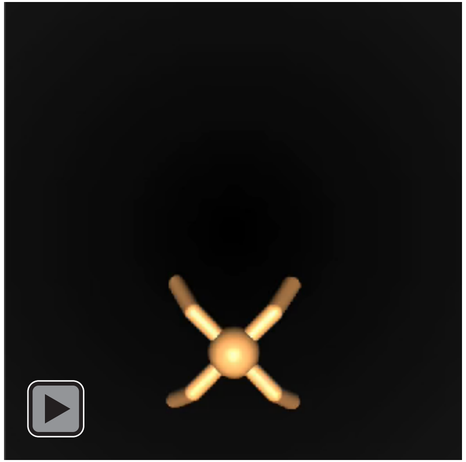
Experiments: Legged Locomotion

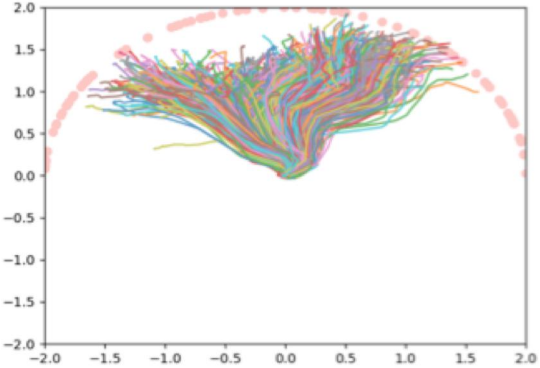
Random Exploration



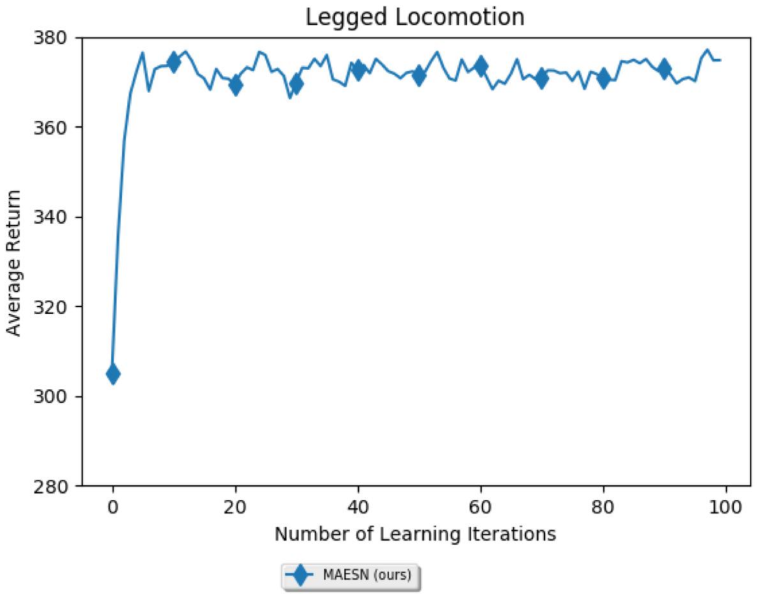


MAESN exploration



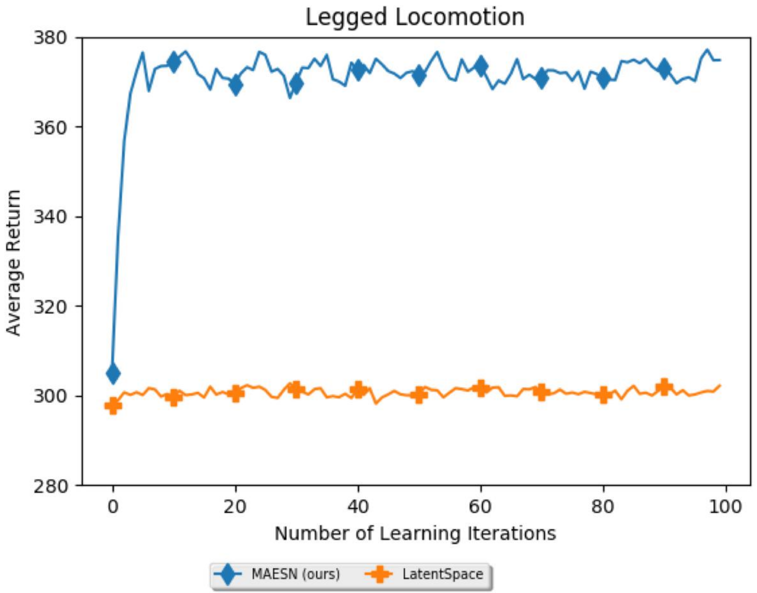


Quick Learning of New Tasks



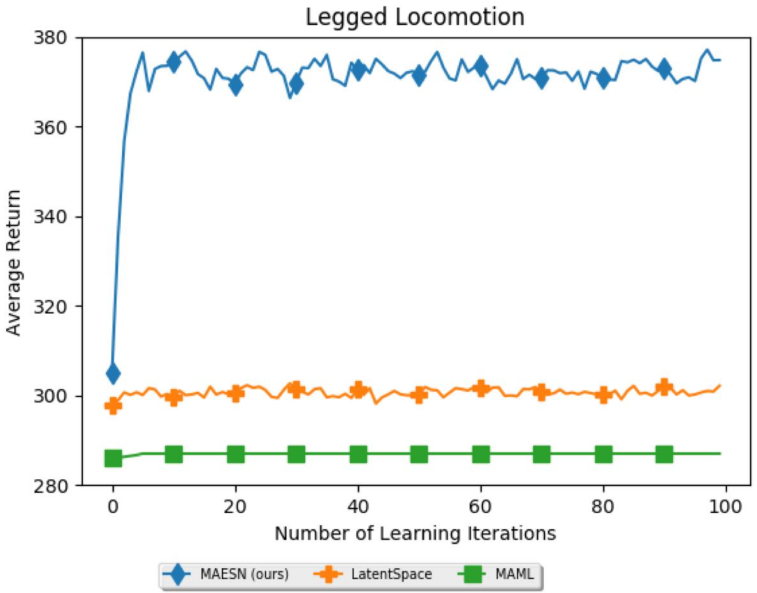
- Learns very quickly
- Higher asymptotic reward than prior methods
- Better exploration

Quick Learning of New Tasks



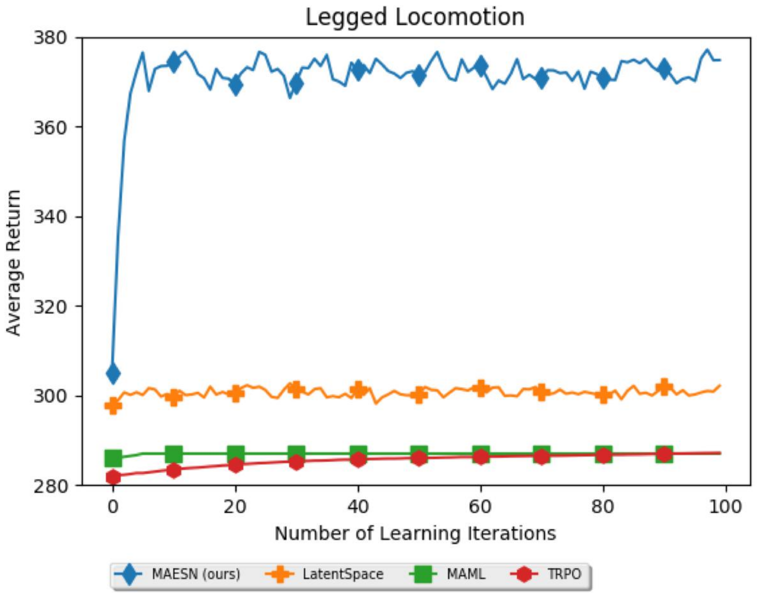
- Learns very quickly
- Higher asymptotic reward than prior methods
- Better exploration

Quick Learning of New Tasks



- Learns very quickly
- Higher asymptotic reward than prior methods
- Better exploration

Quick Learning of New Tasks

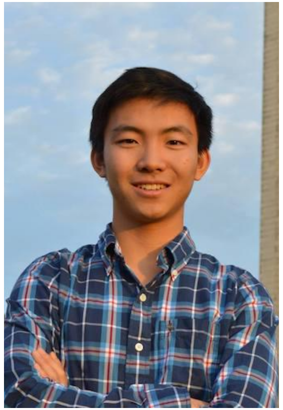


- Learns very quickly
- Higher asymptotic reward than prior methods
- Better exploration

Thank You!

Russell Mendonca

YuXuan Liu



Pieter Abbeel

Sergey Levine

Please come visit our poster at
Room 210 and 230, AB #134

Find code and paper online at https://sites.google.com/view/meta-explore/