Meta-Learning: Challenges and Frontiers

Chelsea Finn

BAIR
BERKELEY ARTIFICIAL INTELLIGENCE RESEARCH
UC Berkeley

Google Brain

Stanford

To learn intelligent, **general-purpose** behavior, learning each new skill from scratch isn't going to cut it.

meta-learning $\longleftrightarrow$ learning priors, structure such that learning new tasks is fast

Can we learn a representation under which **SL** is fast and efficient?

Given 1 example of 5 classes:

Classify new examples

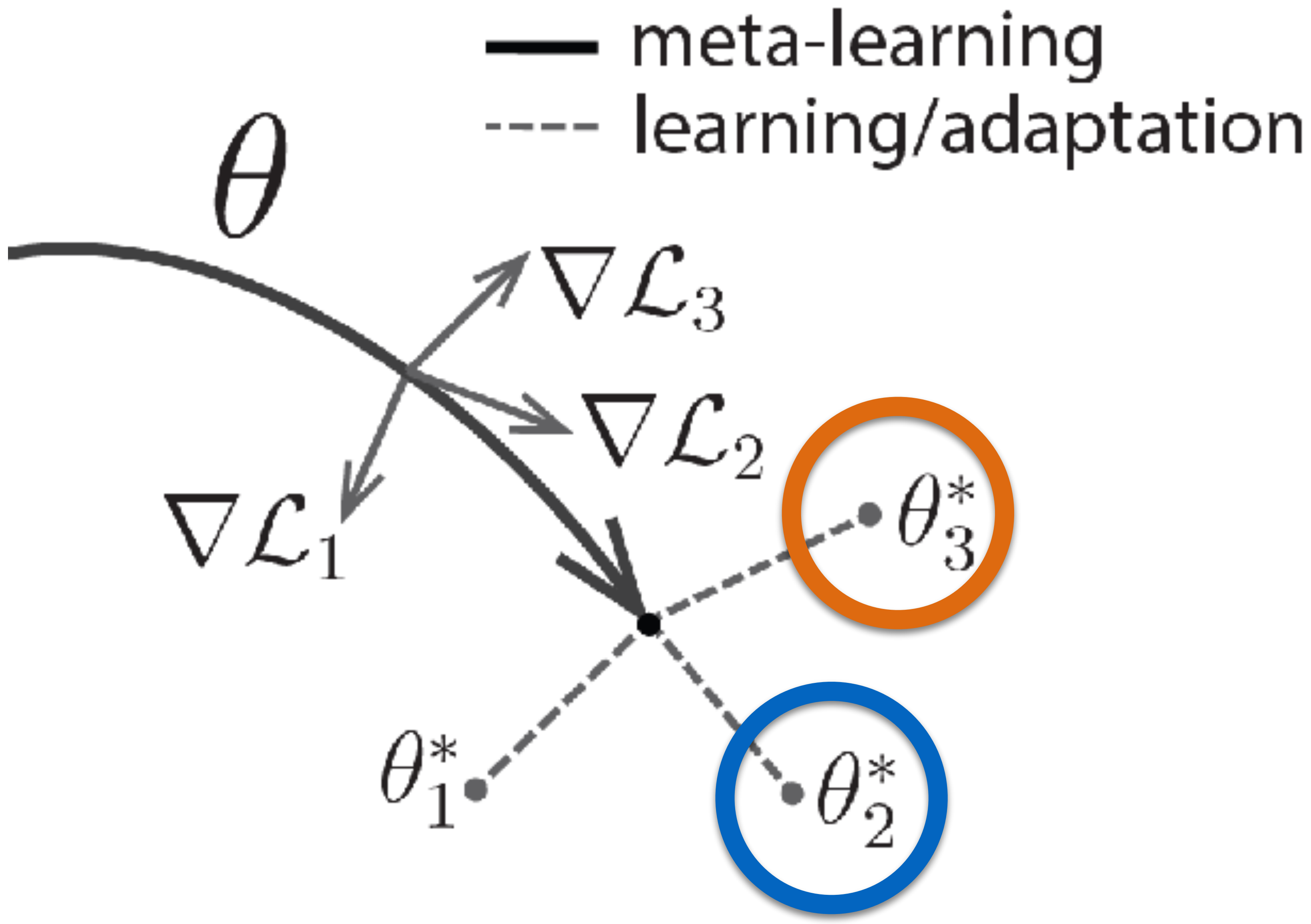
training data

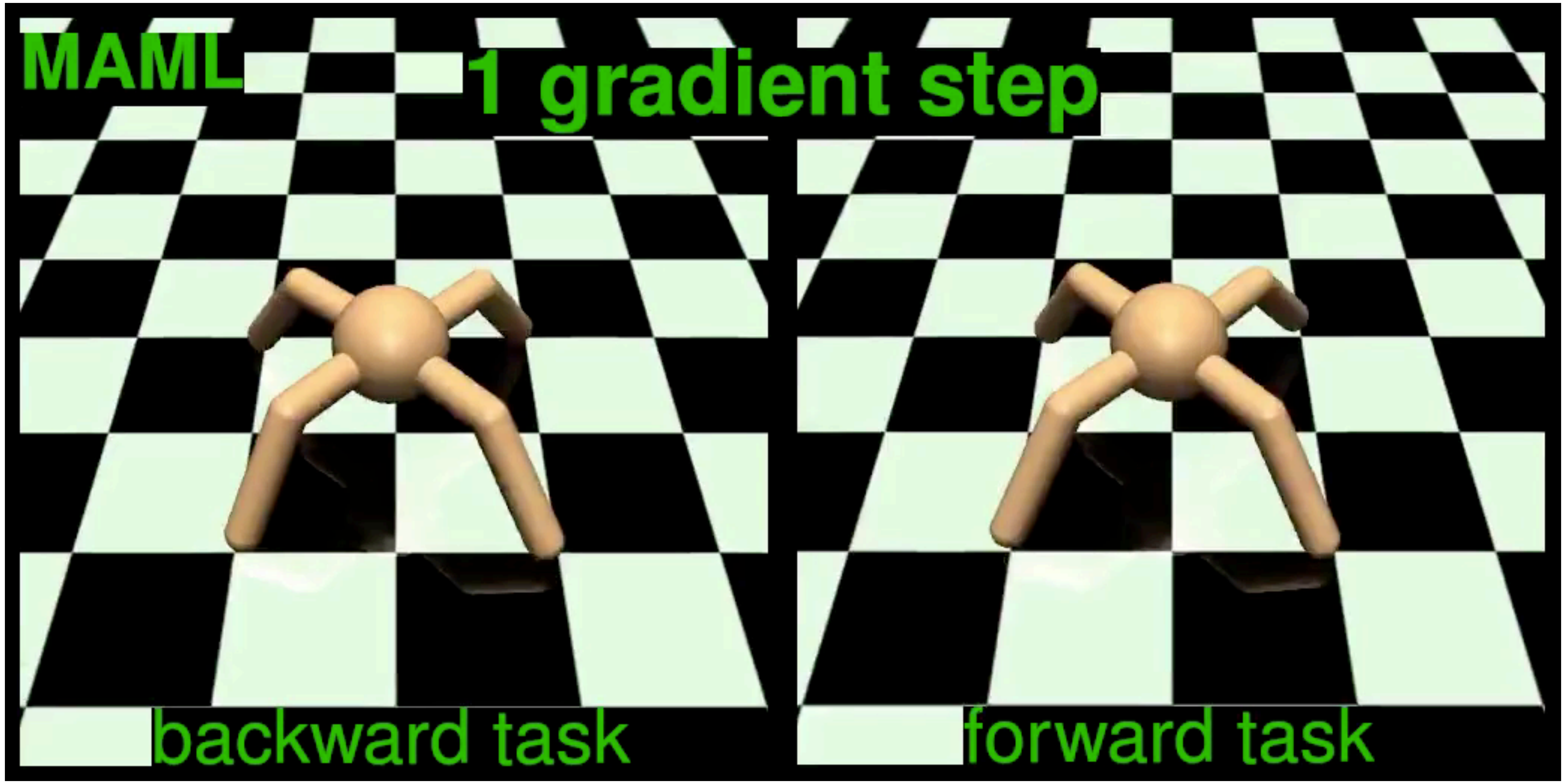
test datapoints

training classes

diagram adapted from Ravi & Larochelle '17

Can we learn a representation under which **RL** is fast and efficient?





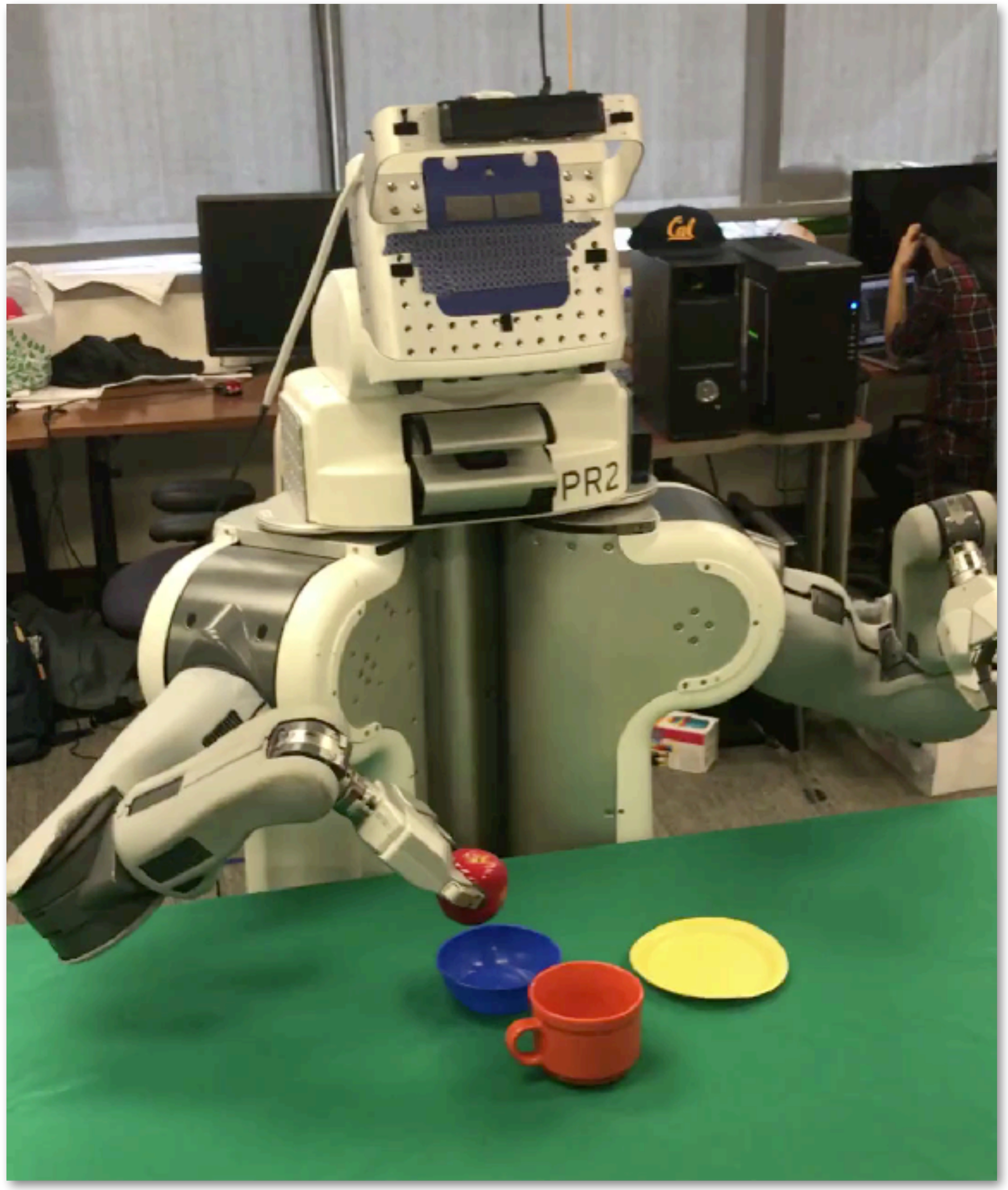
Finn, Abbeel, Levine. Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks. ICML '17

Can we learn a representation under which **imitation** is fast and efficient?

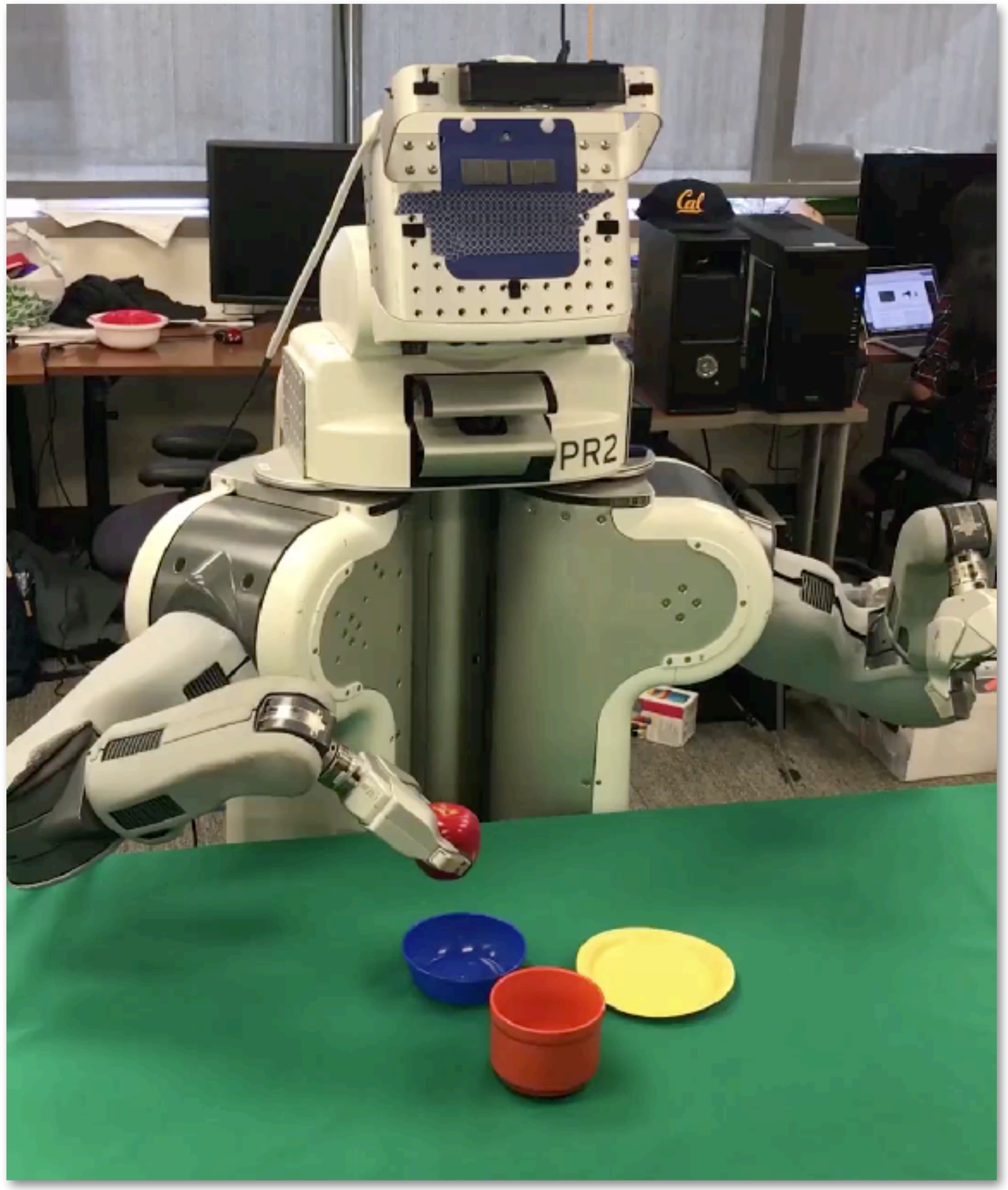
subset of
training objects

held-out test objects

input demo
(via teleoperation)



resulting policy

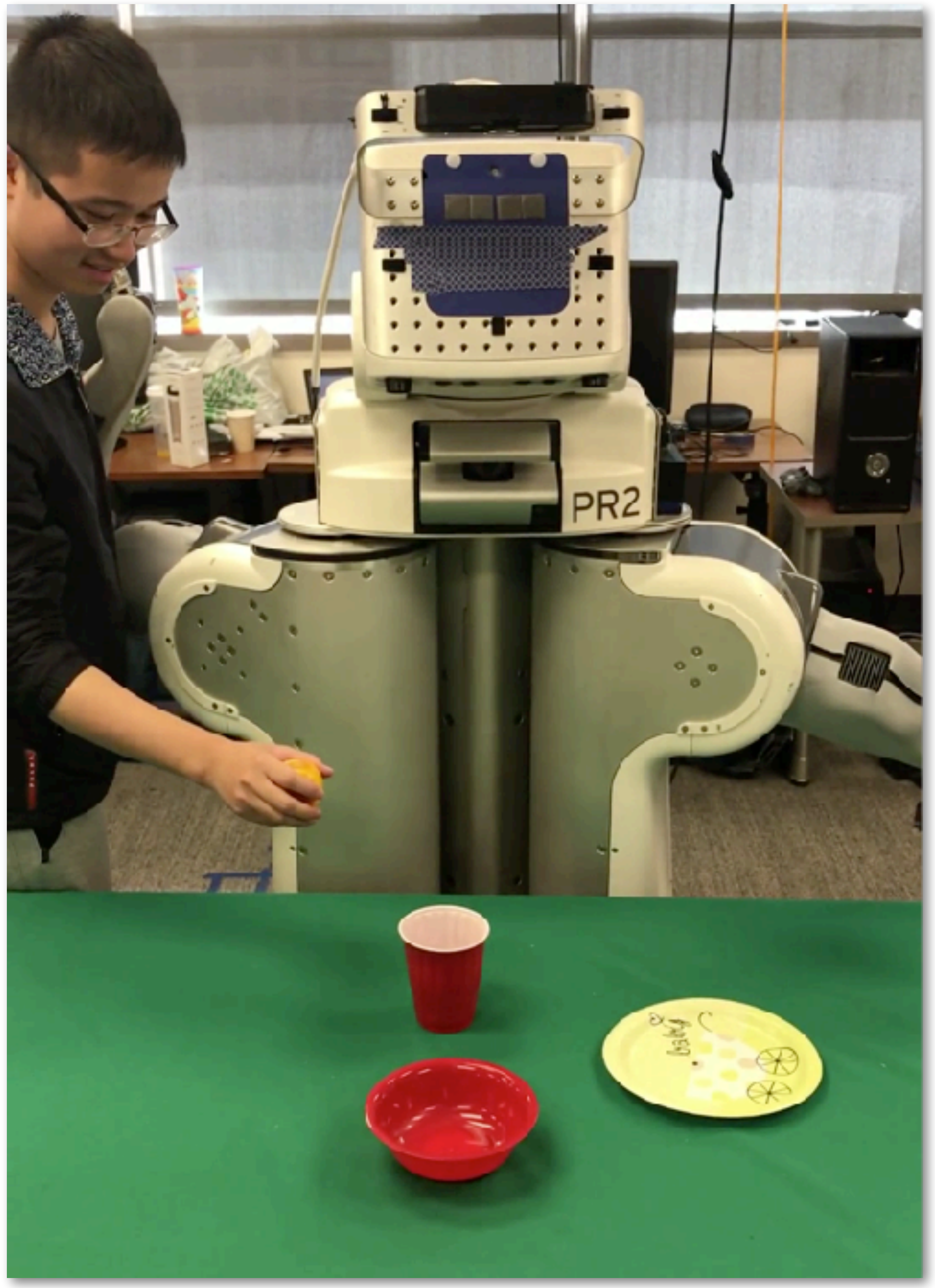


[real-time execution]

Finn*, Yu*, Zhang, Abbeel, Levine. One-Shot Visual Imitation Learning via Meta-Learning. CoRL '17

Can we learn a representation under which **human imitation** is fast and efficient?

input human demo

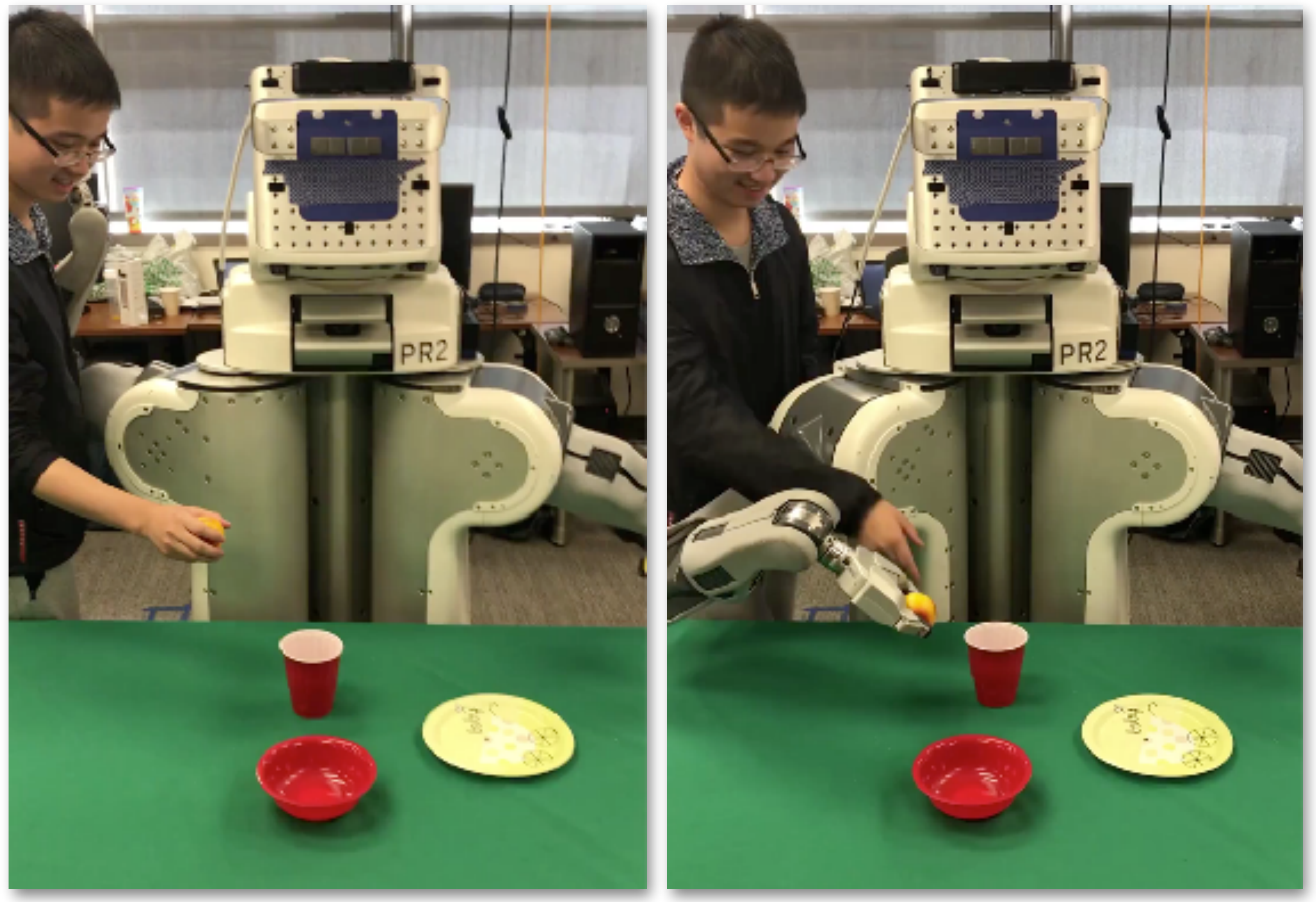


resulting policy

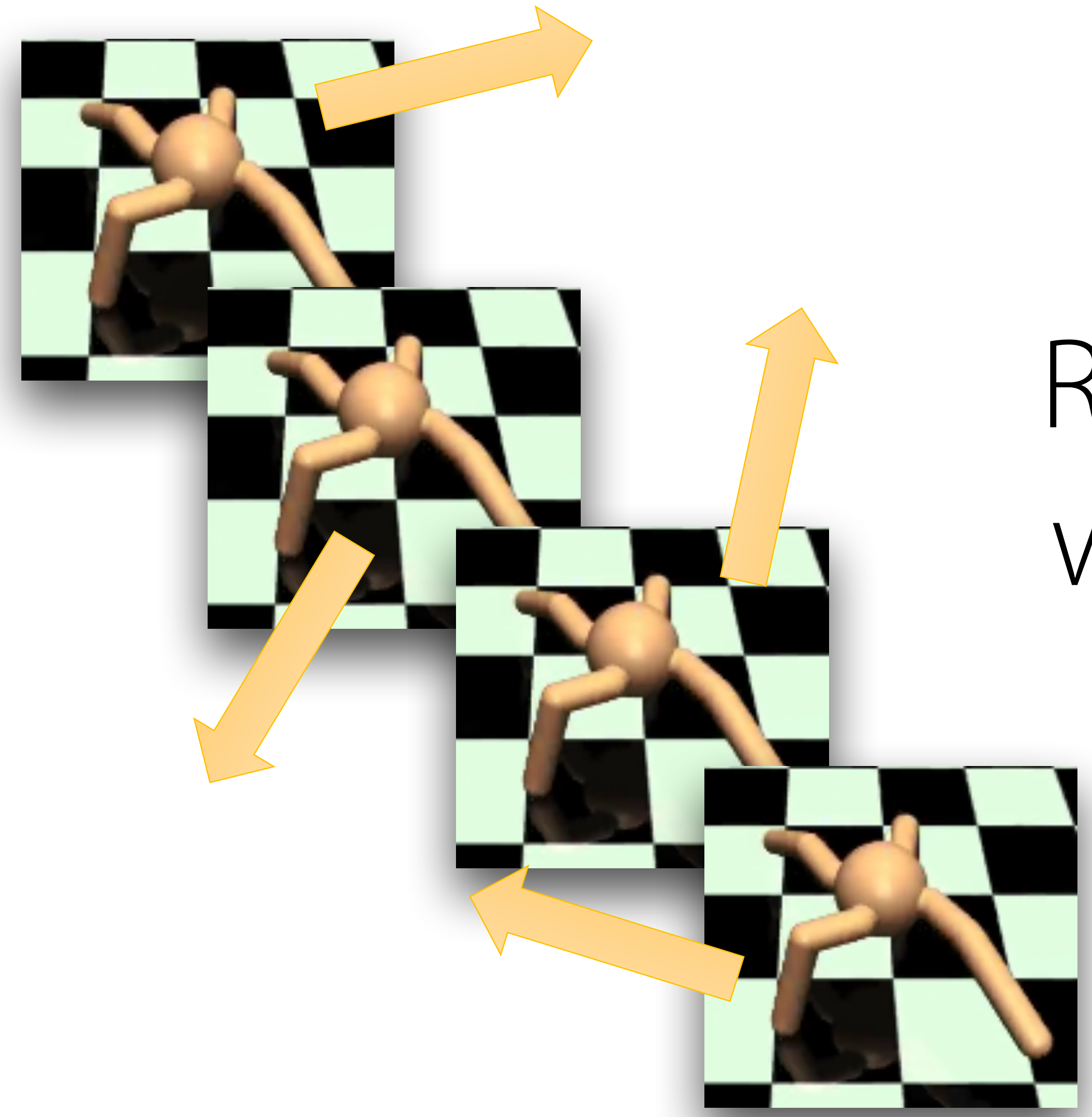
Yu*, Finn*, Dasari, Zhang, Abbeel, Levine. One-Shot Imitation from Observing Humans via Domain Adaptive Meta-Learning. RSS '18

Where do the tasks come from?

Requires tasks constructed from labeled data



Requires demos for many previous tasks



Requires many tasks with corresponding reward functions

Meta-learning: manual algorithm design —> manual task distribution design

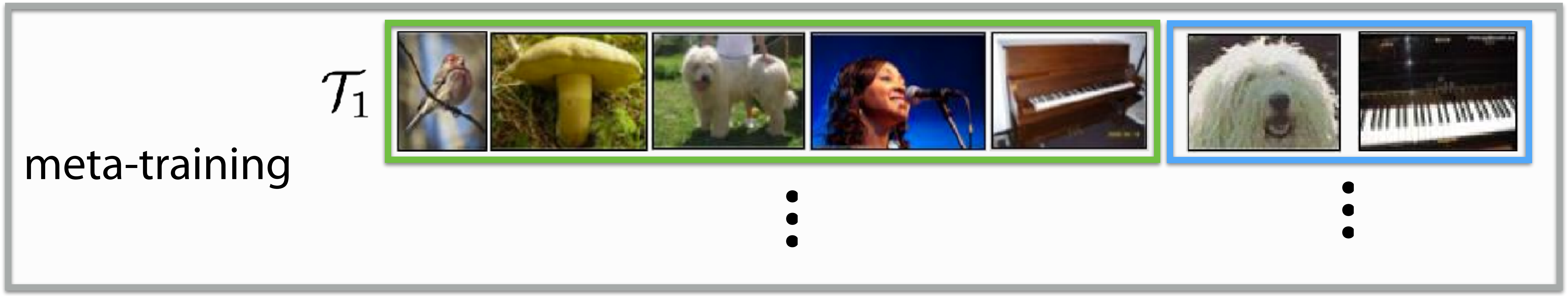
Are we simply kicking the can down the road?

Where do the tasks come from?

self-driven: propose your own tasks

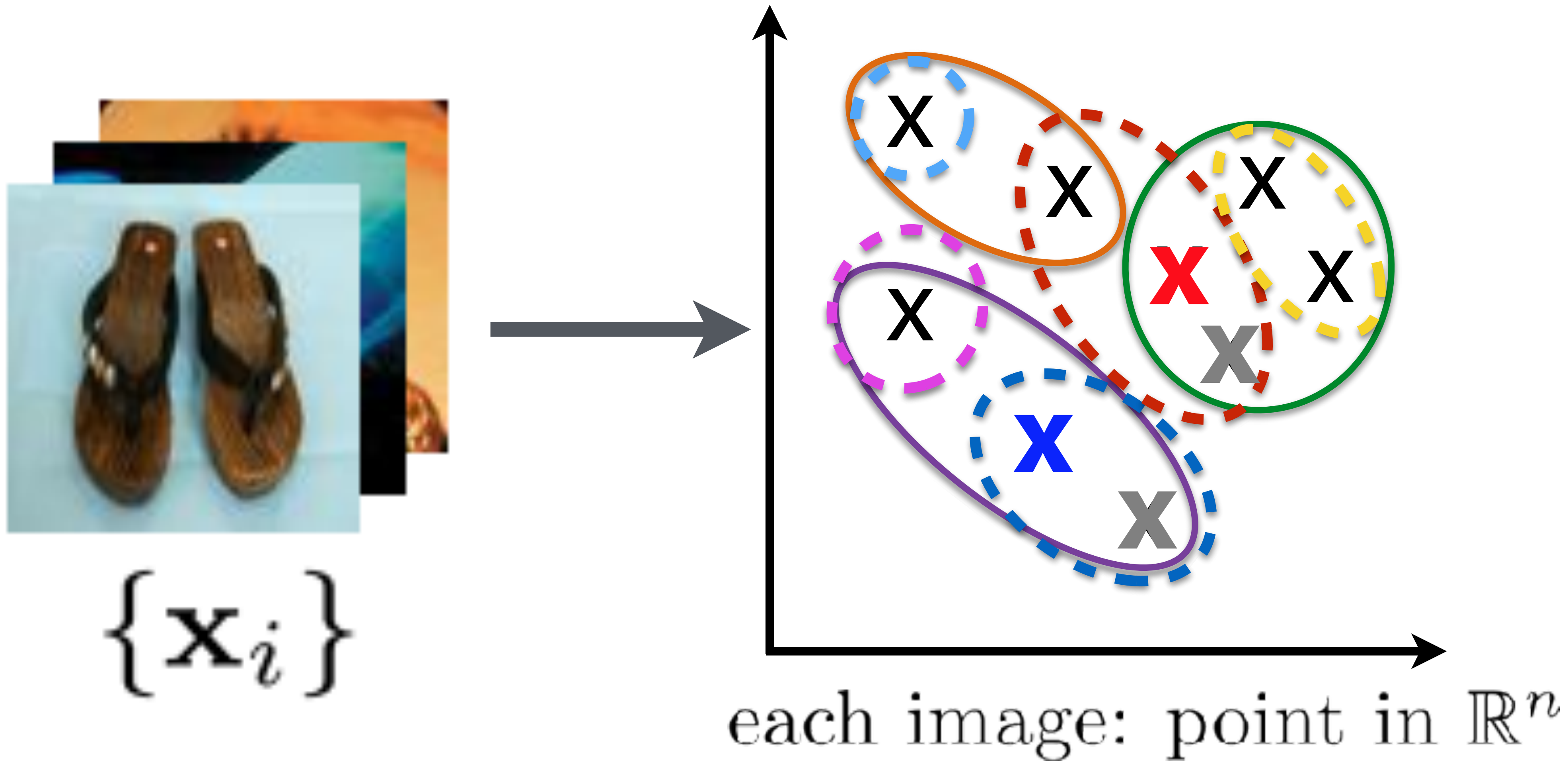
environment-driven: dynamic, real-world environment

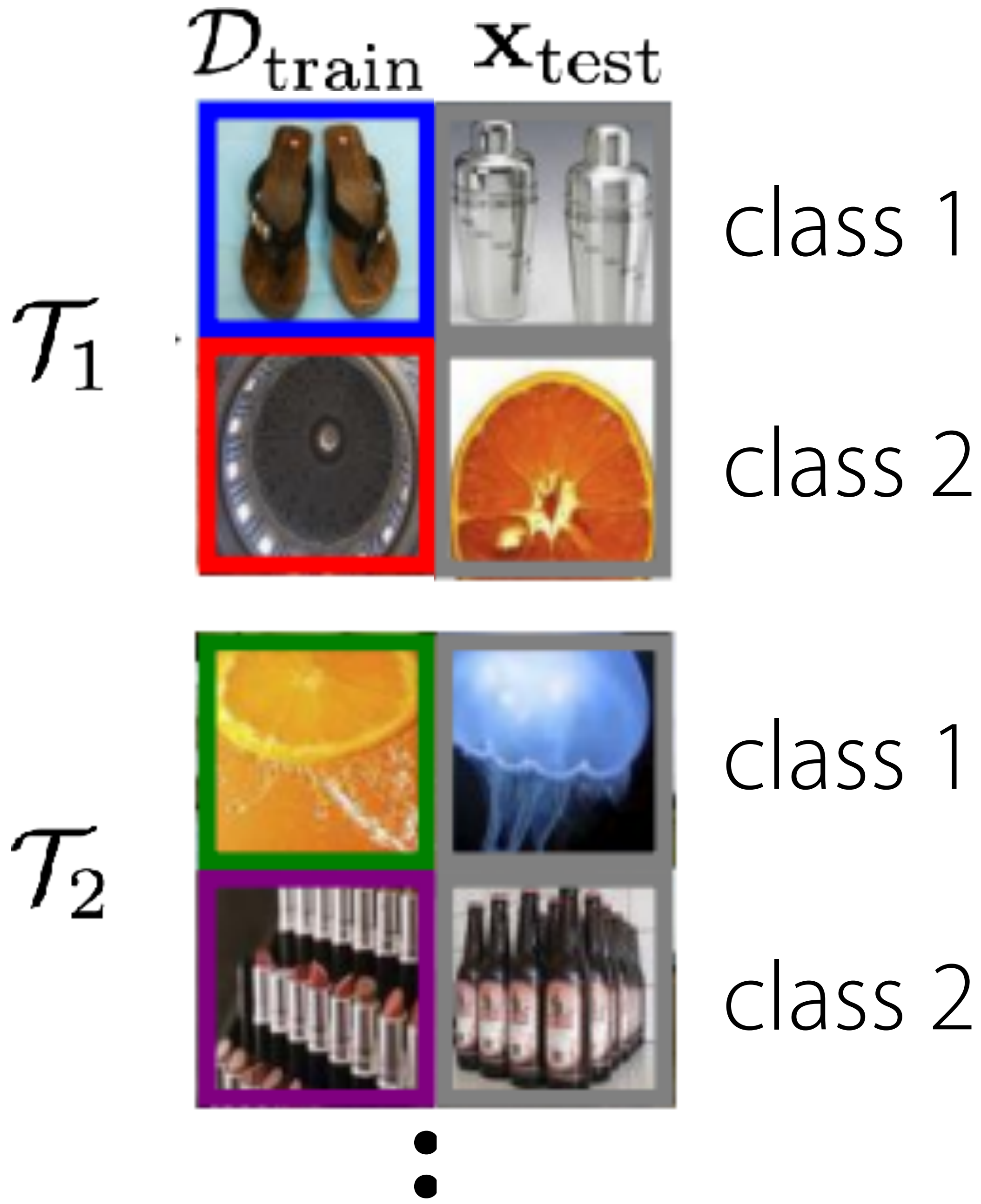
Meta-learn with only **unlabeled** images?

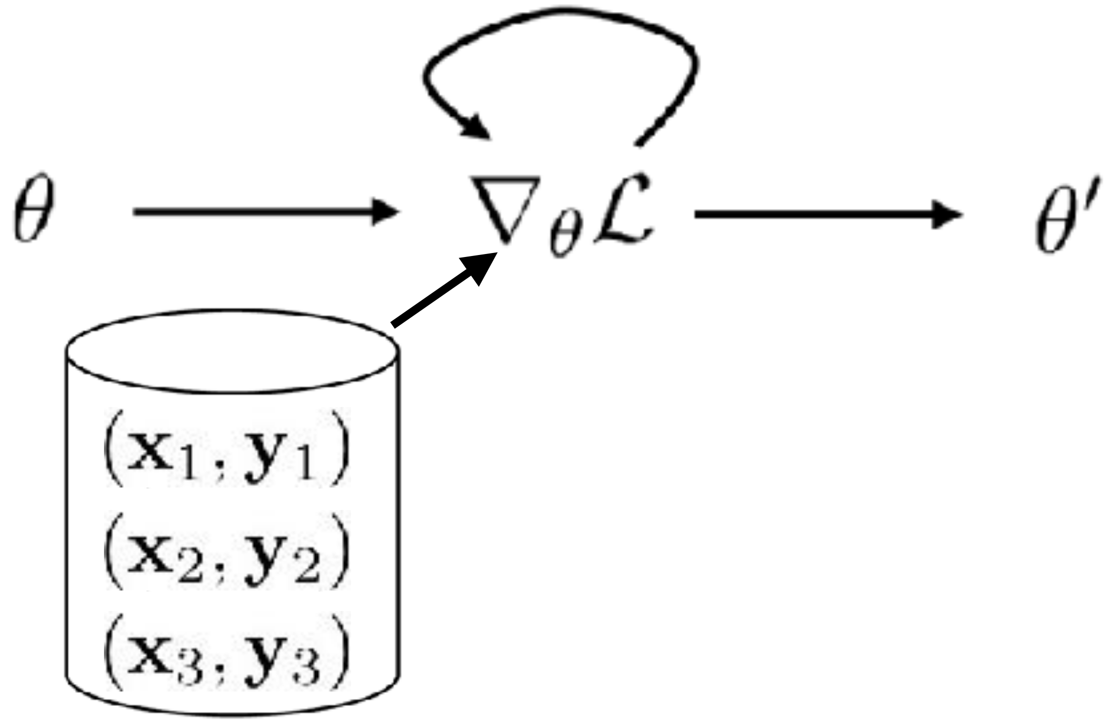


Construct tasks without labeled data?

Unsupervised learning (to get an embedding space) → **Propose tasks** → **Run meta-learning**







Result: representation suitable for learning downstream tasks


Kyle Hsu


Sergey Levine

Hsu, Levine, Finn. Unsupervised Learning via Meta-Learning. ICLR '19

Can we meta-learn with only **unlabeled** images?

Unsupervised learning (to get an embedding space) → **Propose tasks** → **Run meta-learning**

A few options:
BiGAN — Donahue et al. '17
DeepCluster — Caron et al. '18

Clustering to Automatically Construct Tasks for Unsupervised Meta-Learning (CACTUs)

MAML — Finn et al. '17
ProtoNets — Snell et al. '17

CACTUs MAML

miniImageNet 5-way 5-shot

method	accuracy
MAML with labels	62.13%
BiGAN kNN	31.10%
BiGAN logistic	33.91%
BiGAN MLP + dropout	29.06%
BiGAN cluster matching	29.49%
BiGAN CACTUs MAML	51.28%
DeepCluster CACTUs MAML	**53.97%**

Same story for:
- **4 different embedding methods**
- **4 datasets** (Omniglot, CelebA, miniImageNet, MNIST)
- **2 meta-learning methods** (*)
- Test tasks with **larger datasets**

*ProtoNets underperforms in some cases.

Hsu, Levine, Finn. Unsupervised Learning via Meta-Learning. ICLR'19

What about unsupervised meta-RL?

Environment → **Propose tasks** → **Run meta-RL**

Result: Environment-specific RL algorithm


Abhishek Gupta


Ben Eysenbach


Sergey Levine

Gupta, Eysenbach, Finn, Levine. Unsupervised Meta-Learning for Reinforcement Learning. arXiv '18

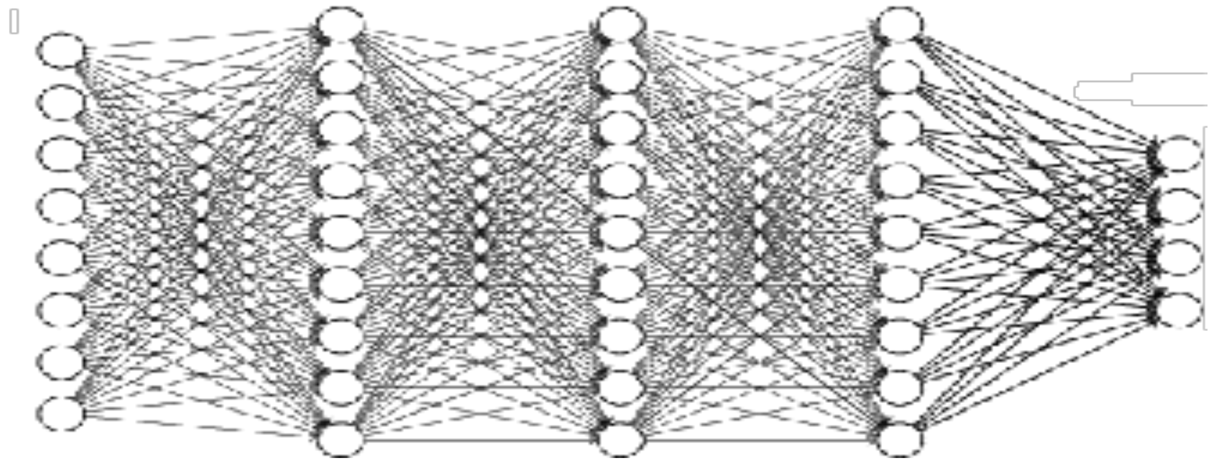
What about unsupervised meta-RL?

Environment $\longrightarrow$ **Propose tasks** $\longrightarrow$ **Run meta-RL**

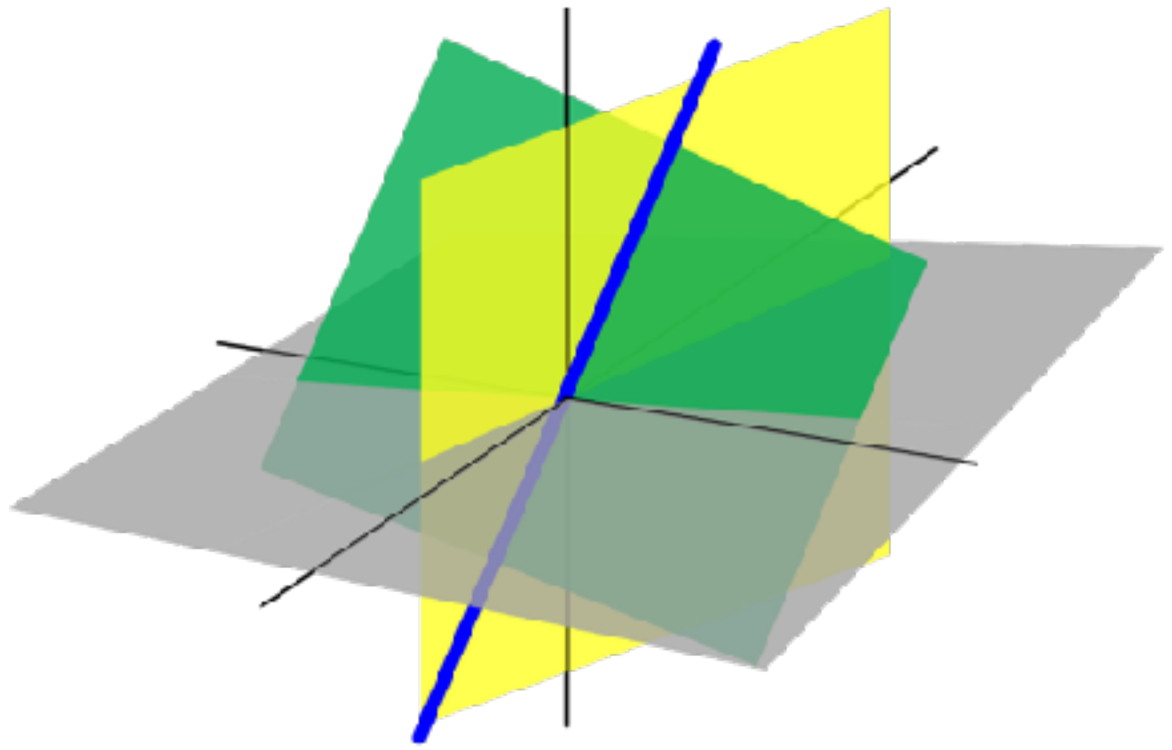
Result: Environment-specific RL algorithm

- Discrete set of random tasks

D → randomly initialized network



$$R(s, z) = \log p_D(z|s)$$



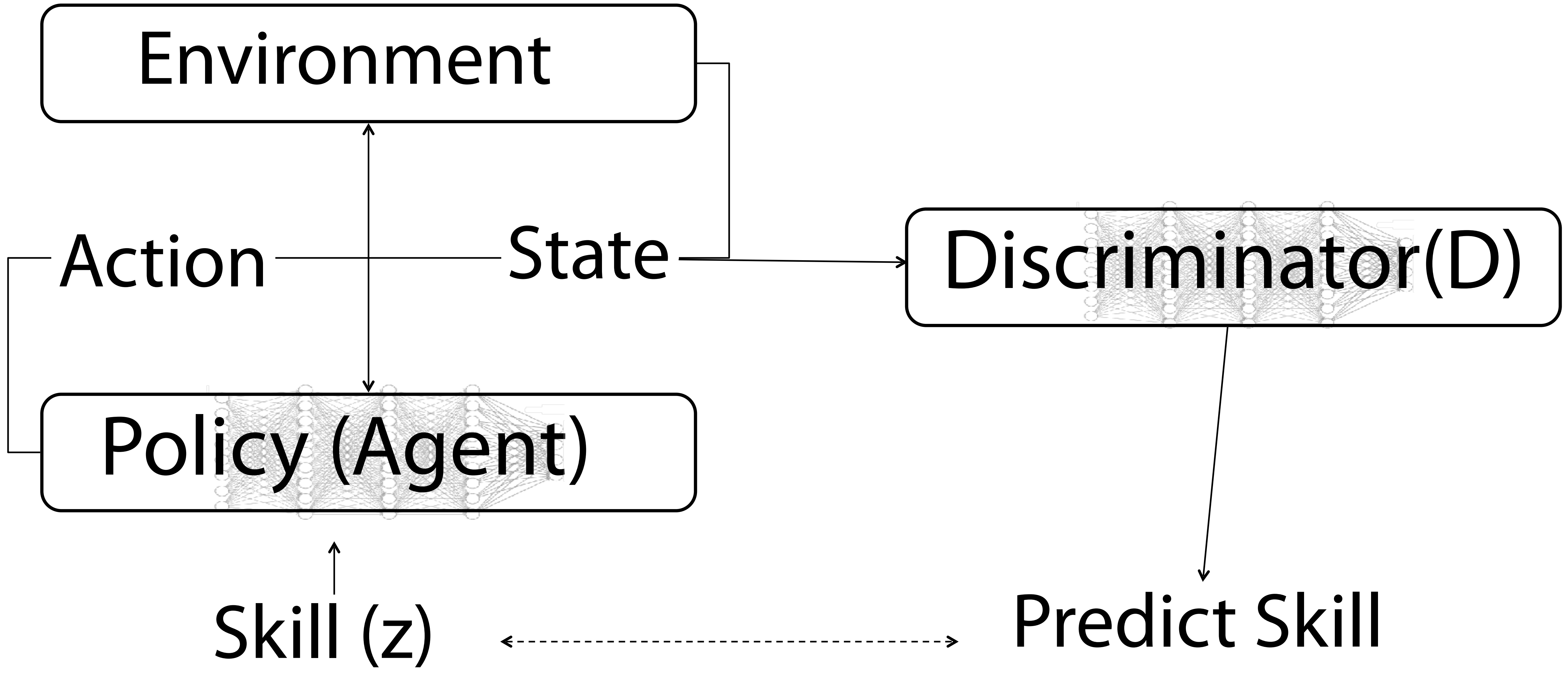
Gupta, Eysenbach, Finn, Levine. Unsupervised Meta-Learning for Reinforcement Learning. arXiv '18

What about unsupervised meta-RL?

Environment ⟶ **Propose tasks** ⟶ **Run meta-RL**

Result: Environment-specific RL algorithm

- Discrete set of random tasks
- Learned diverse set of tasks by pushing apart discriminator classes



- Policy -> visit states that are discriminable
- Discriminator -> predict skill from state

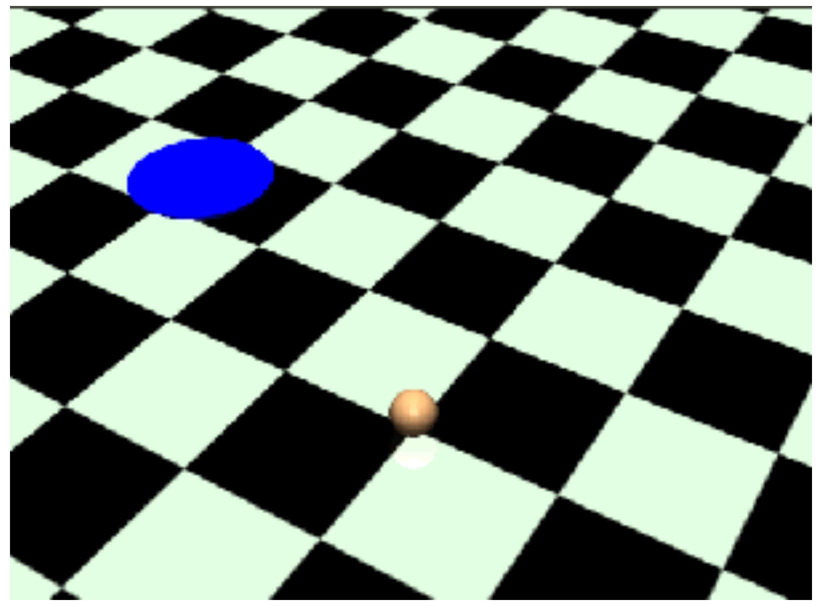
$$R(s, z) = \log p_D(z|s)$$

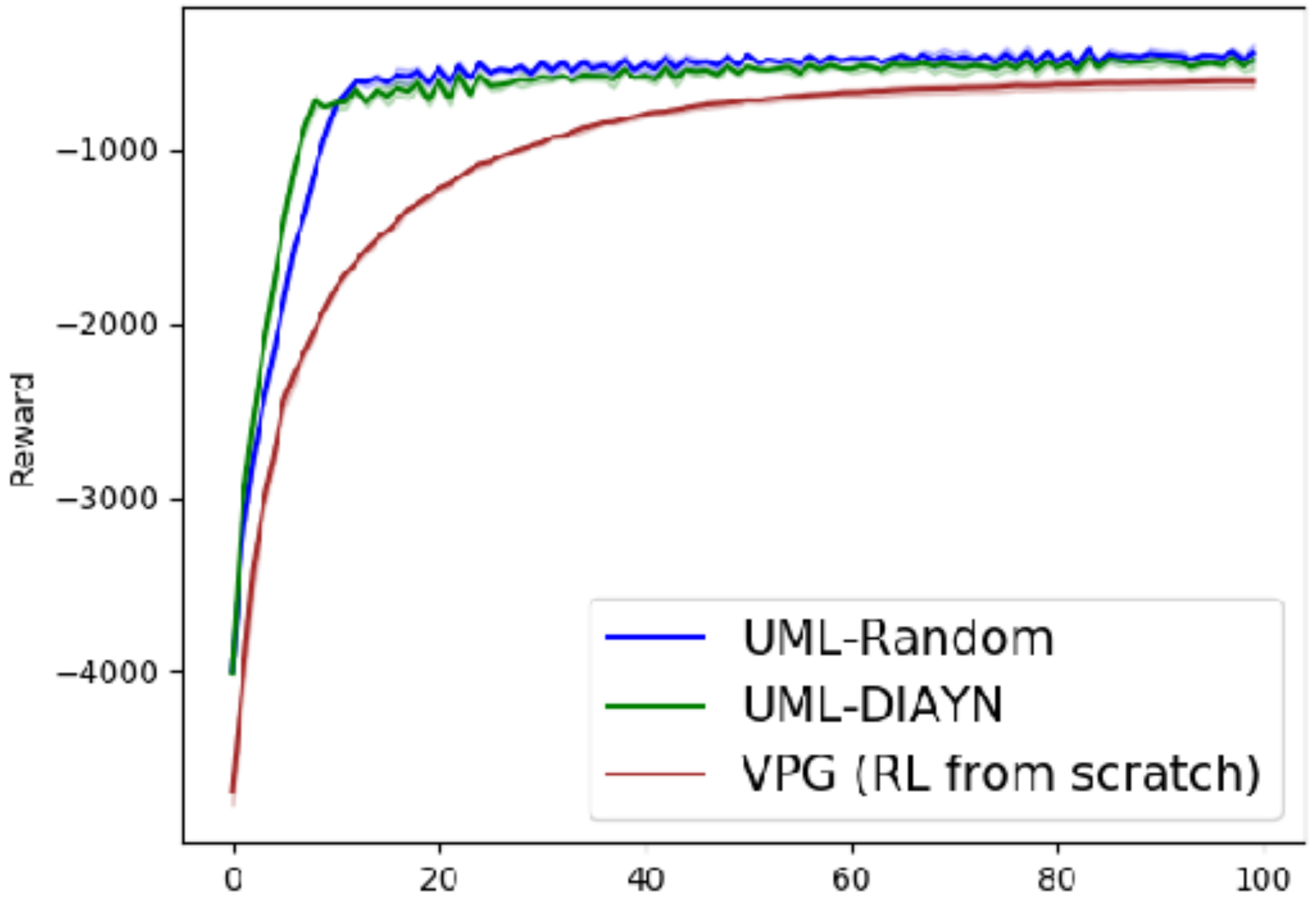
Eysenbach, Gupta, Ibarz, Levine. Diversity is All You Need. ICLR '19
Gupta, Eysenbach, Finn, Levine. Unsupervised Meta-Learning for Reinforcement Learning. arXiv '18

Does it work?

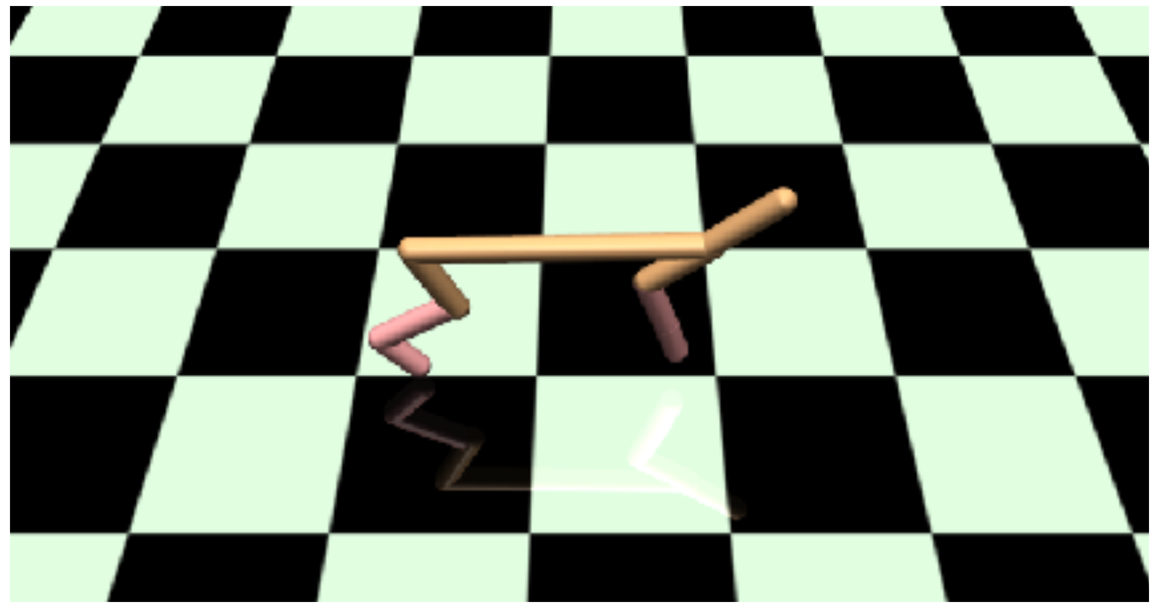
- compare **UML-Random**, **UML-DIAYN**, **RL from scratch**
- measure learning performance on **test tasks *with* rewards**

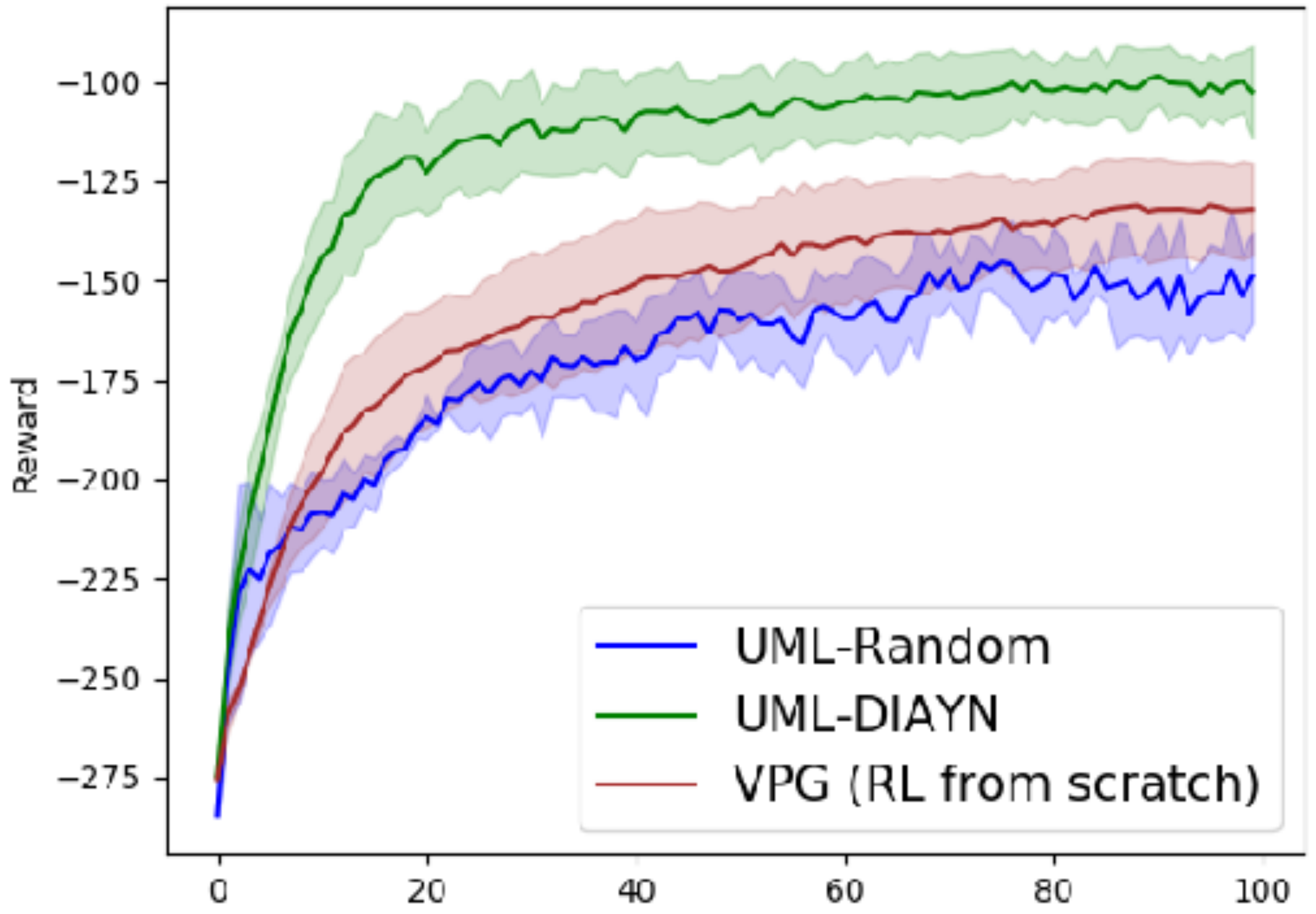
2D Navigation



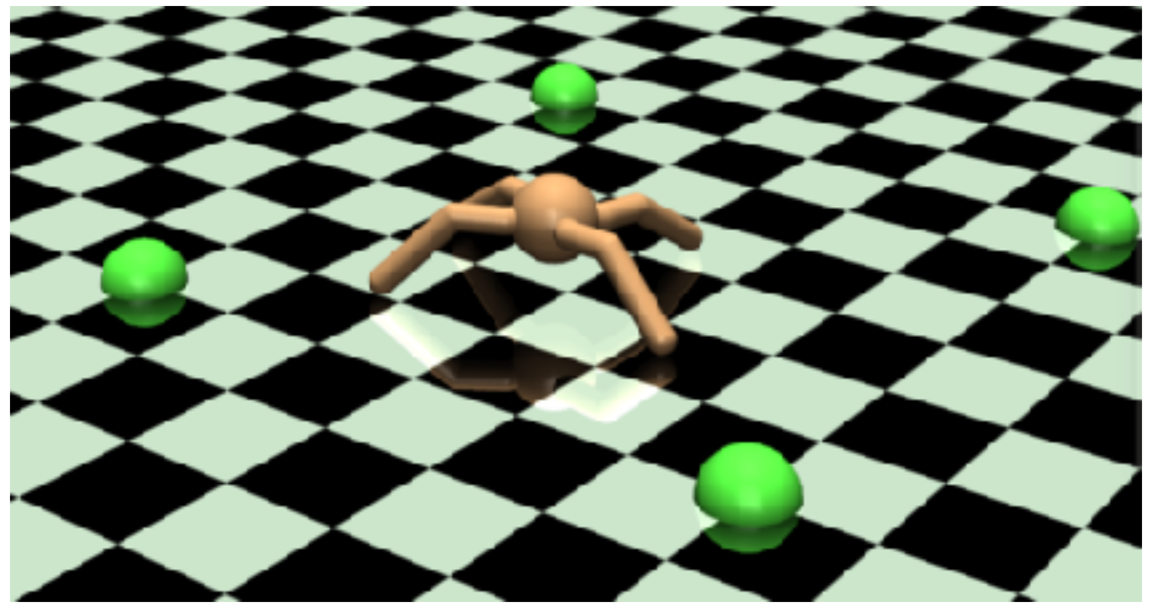


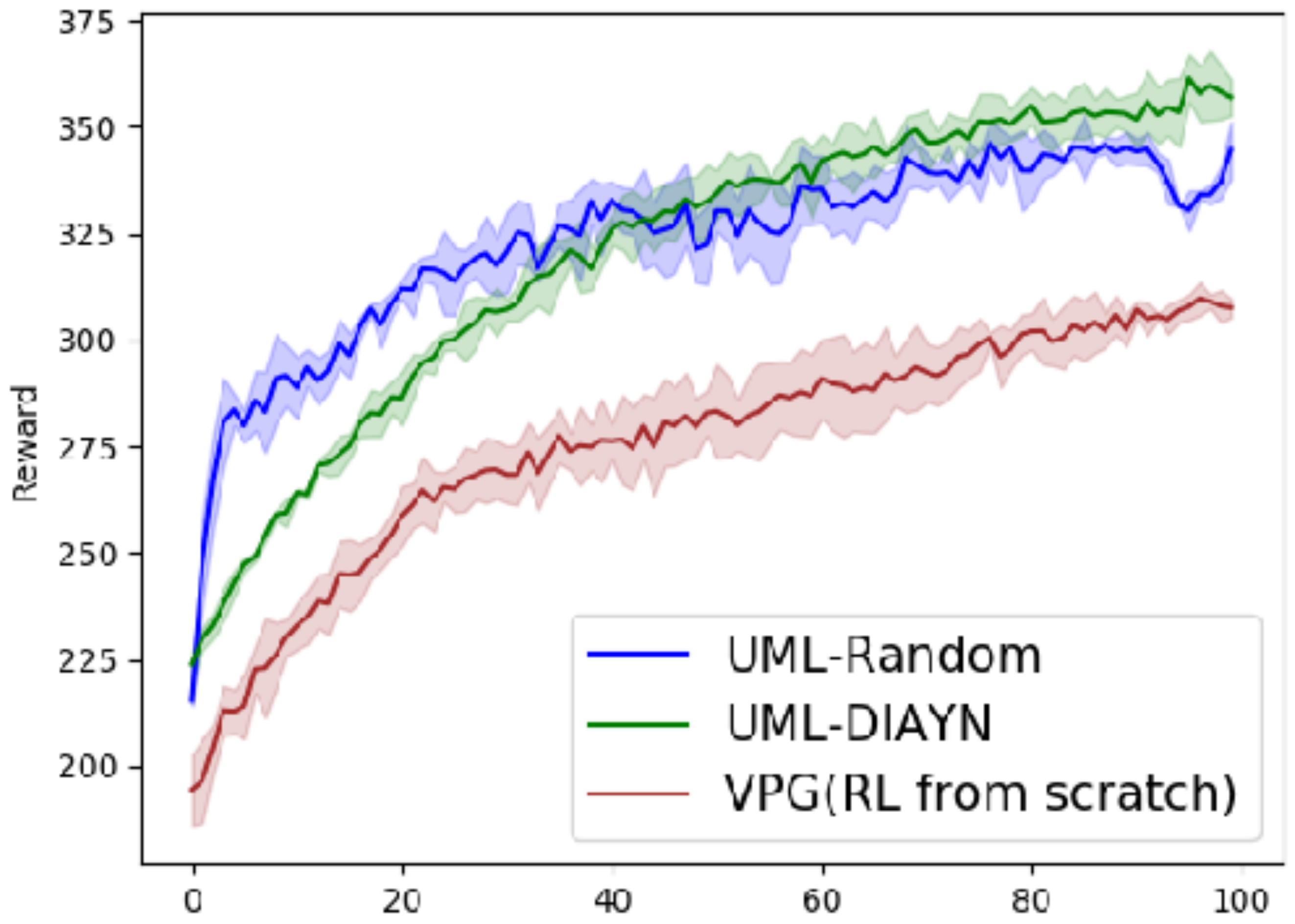
Cheetah





Ant





Takeaway: Relatively **simple** mechanisms for proposing tasks work surprisingly well.

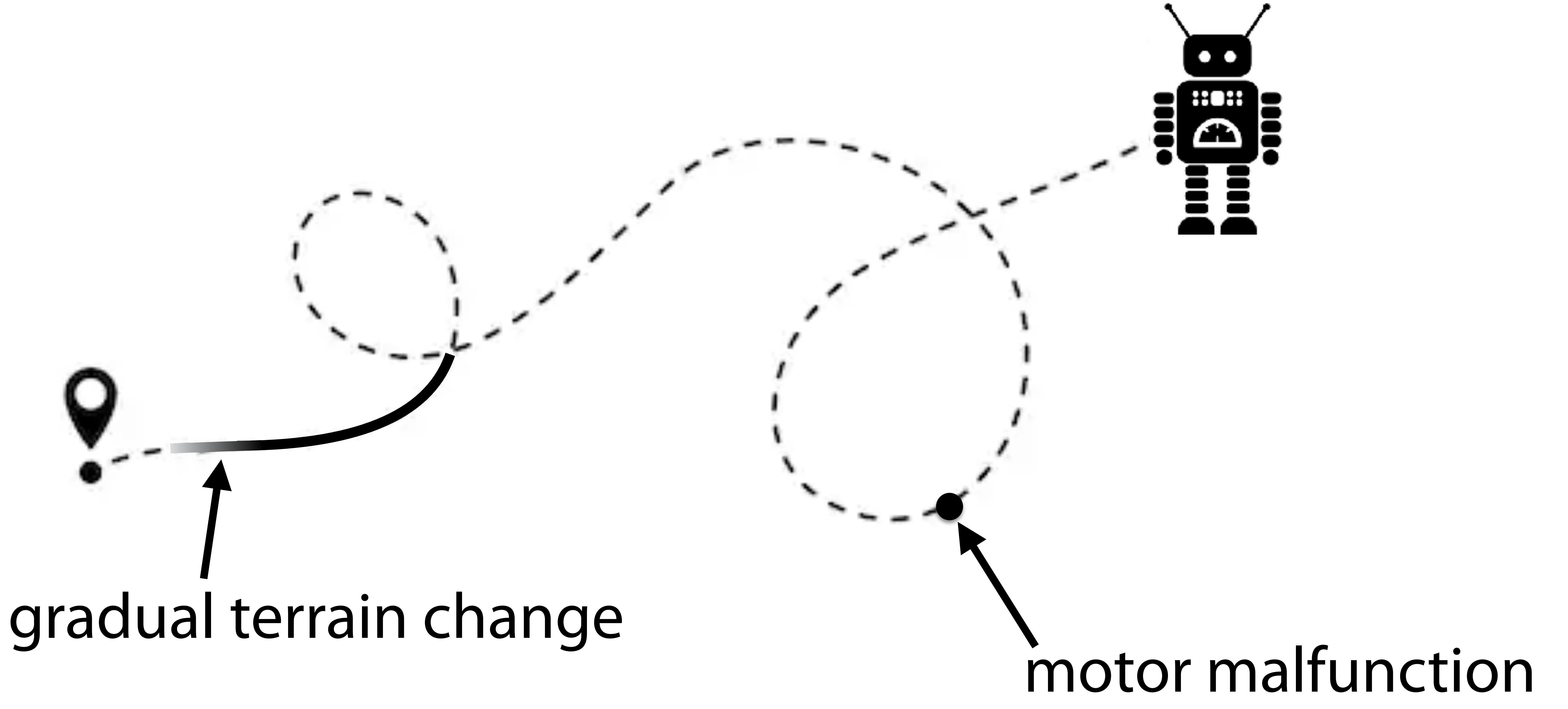
Gupta, Eysenbach, Finn, Levine. Unsupervised Meta-Learning for Reinforcement Learning. arXiv '18

Where do the tasks come from?

self-driven: propose your own tasks

environment-driven: dynamic, real-world environment

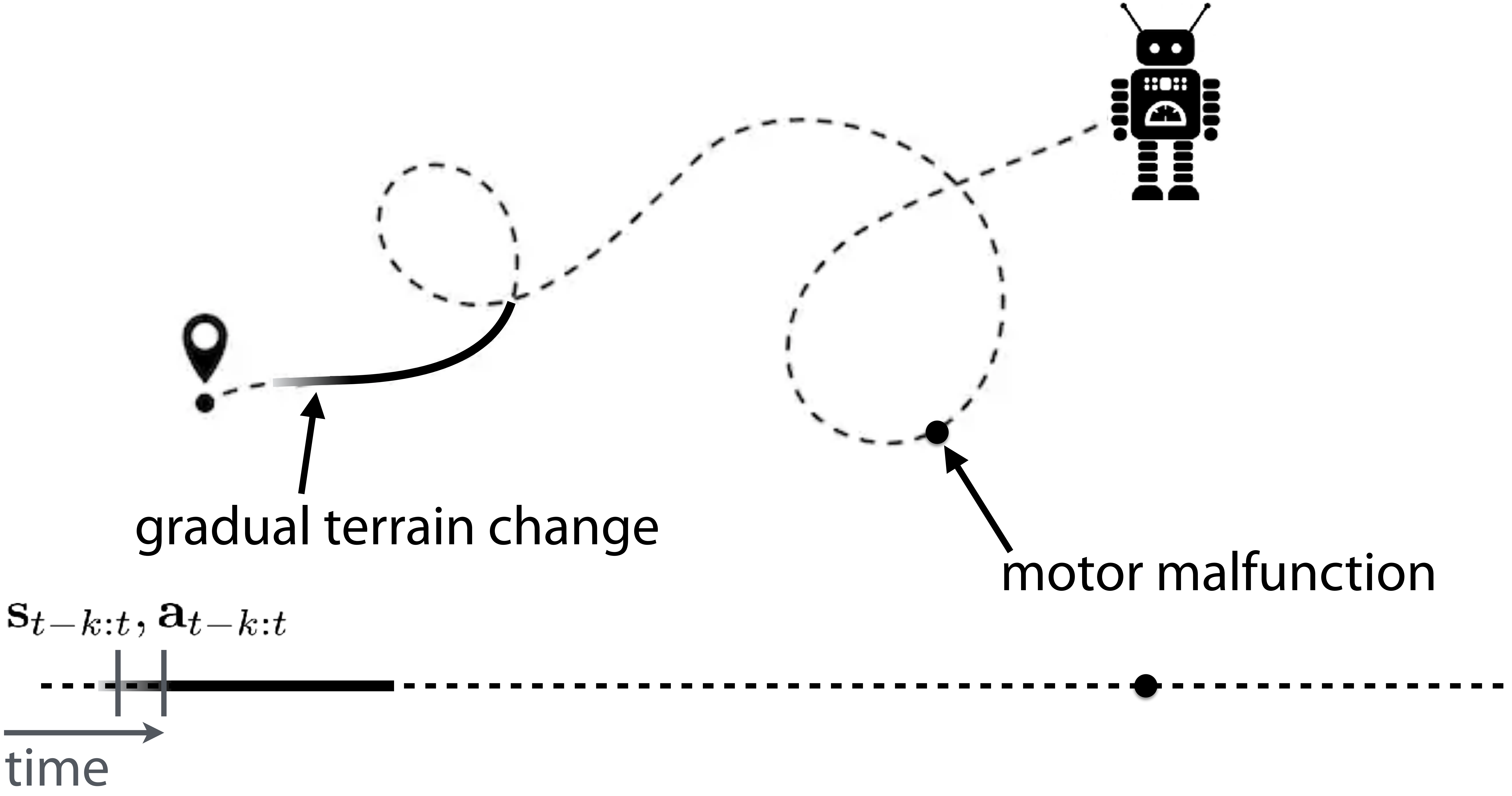
Deriving tasks from dynamic, real-world environments



Anusha Nagabandi Ignasi Clavera Simin Liu

Nagabandi*, Clavera*, Liu, Fearing, Abbeel, Levine, Finn. Learning to Adapt in Dynamic Environments through Meta-RL. ICLR'19

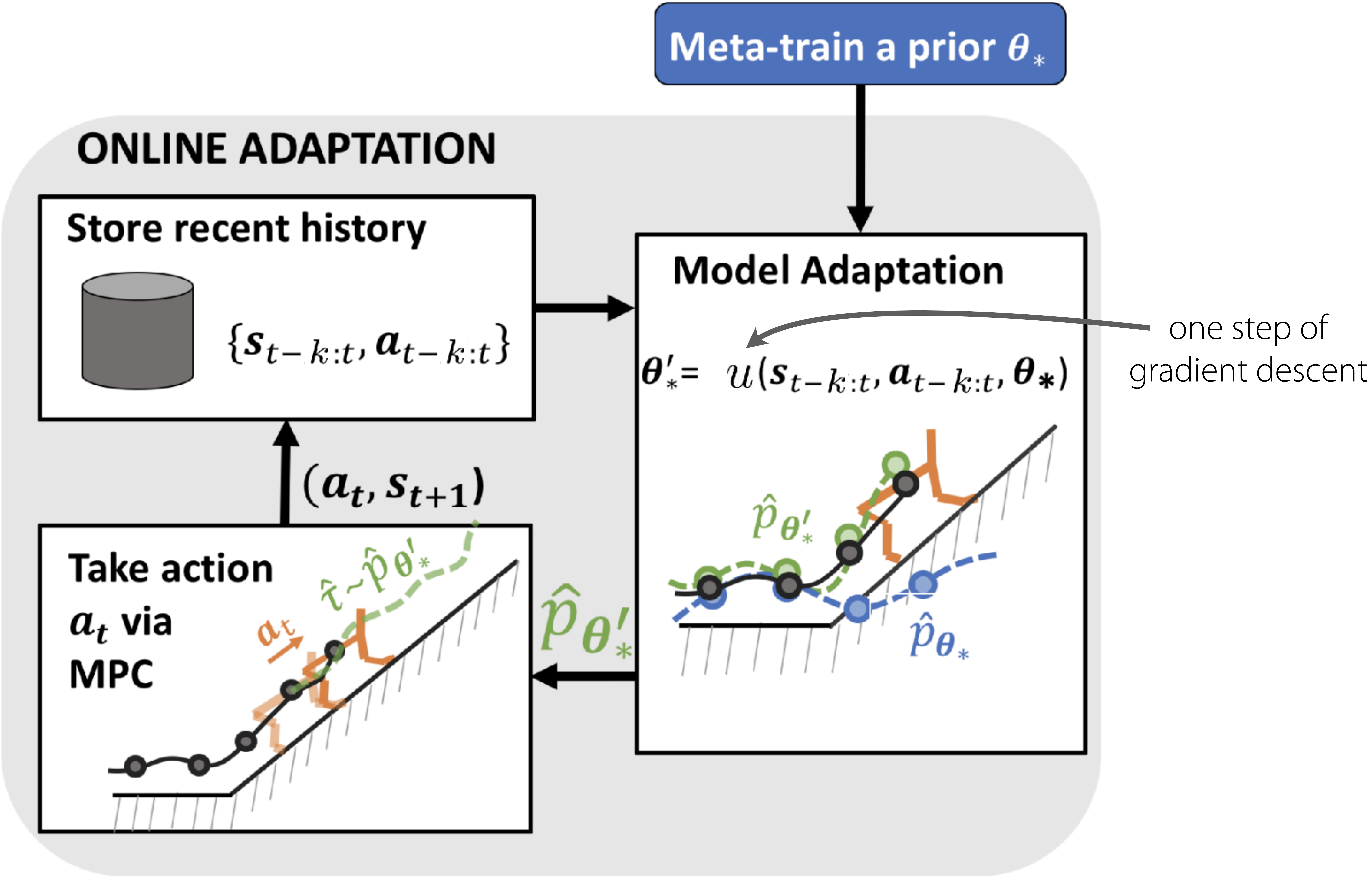
Deriving tasks from dynamic, real-world environments



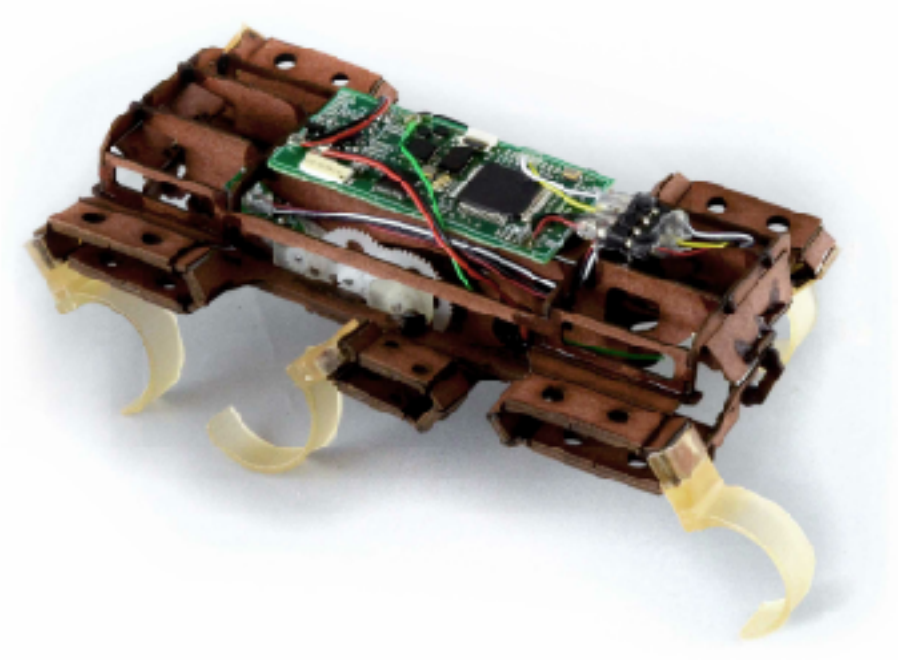
online adaptation = few-shot learning

tasks are **temporal slices** of experience

Nagabandi*, Clavera*, Liu, Fearing, Abbeel, Levine, Finn. Learning to Adapt in Dynamic Environments through Meta-RL. ICLR'19

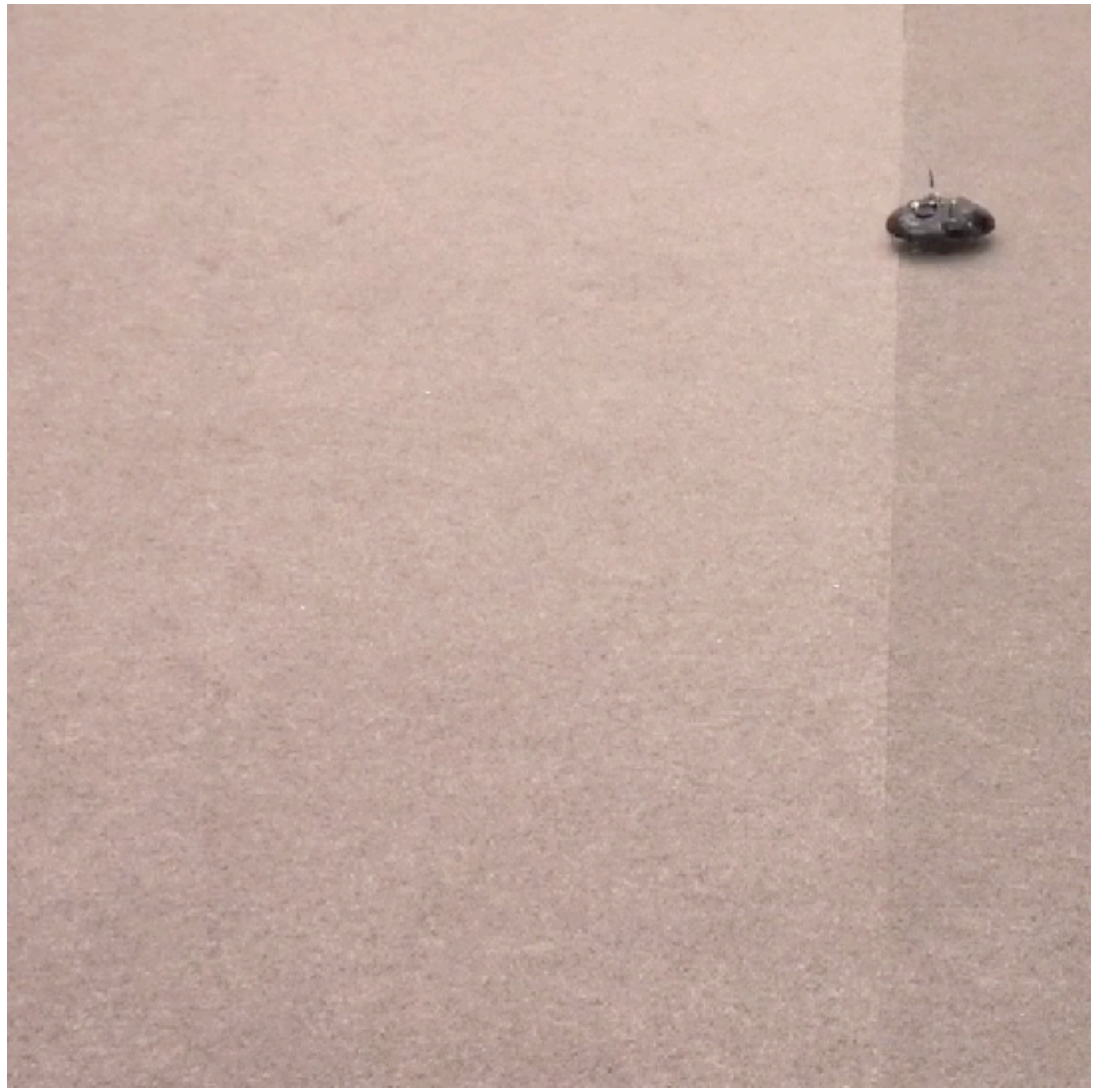


Nagabandi*, Clavera*, Liu, Fearing, Abbeel, Levine, Finn. Learning to Adapt in Dynamic Environments through Meta-RL. ICLR'19



VelociRoACH Robot

Meta-train on **variable terrains**



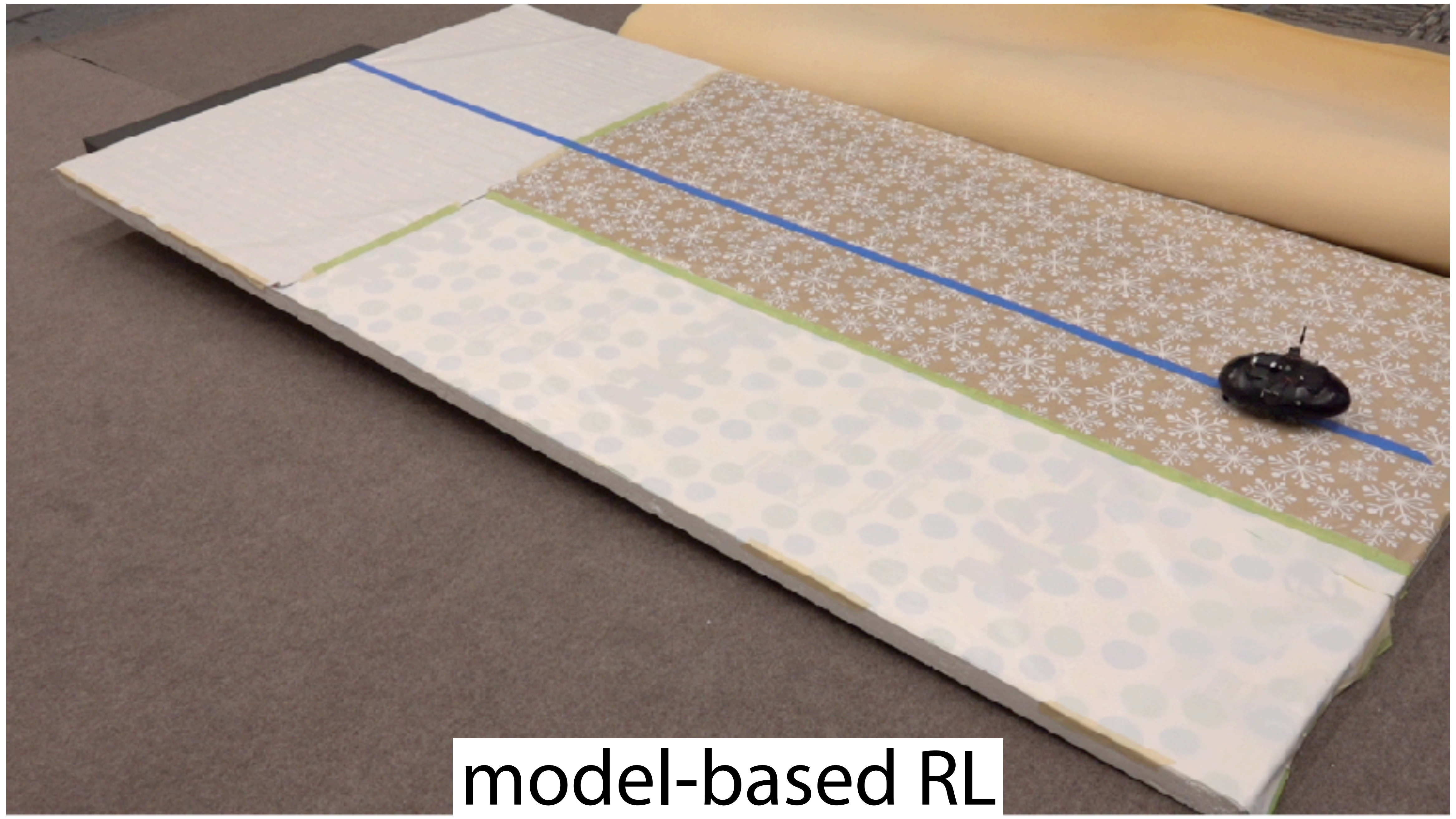
Meta-test with **slope, missing leg, payload, calibration errors**

Nagabandi*, Clavera*, Liu, Fearing, Abbeel, Levine, Finn. Learning to Adapt in Dynamic Environments through Meta-RL. ICLR '19

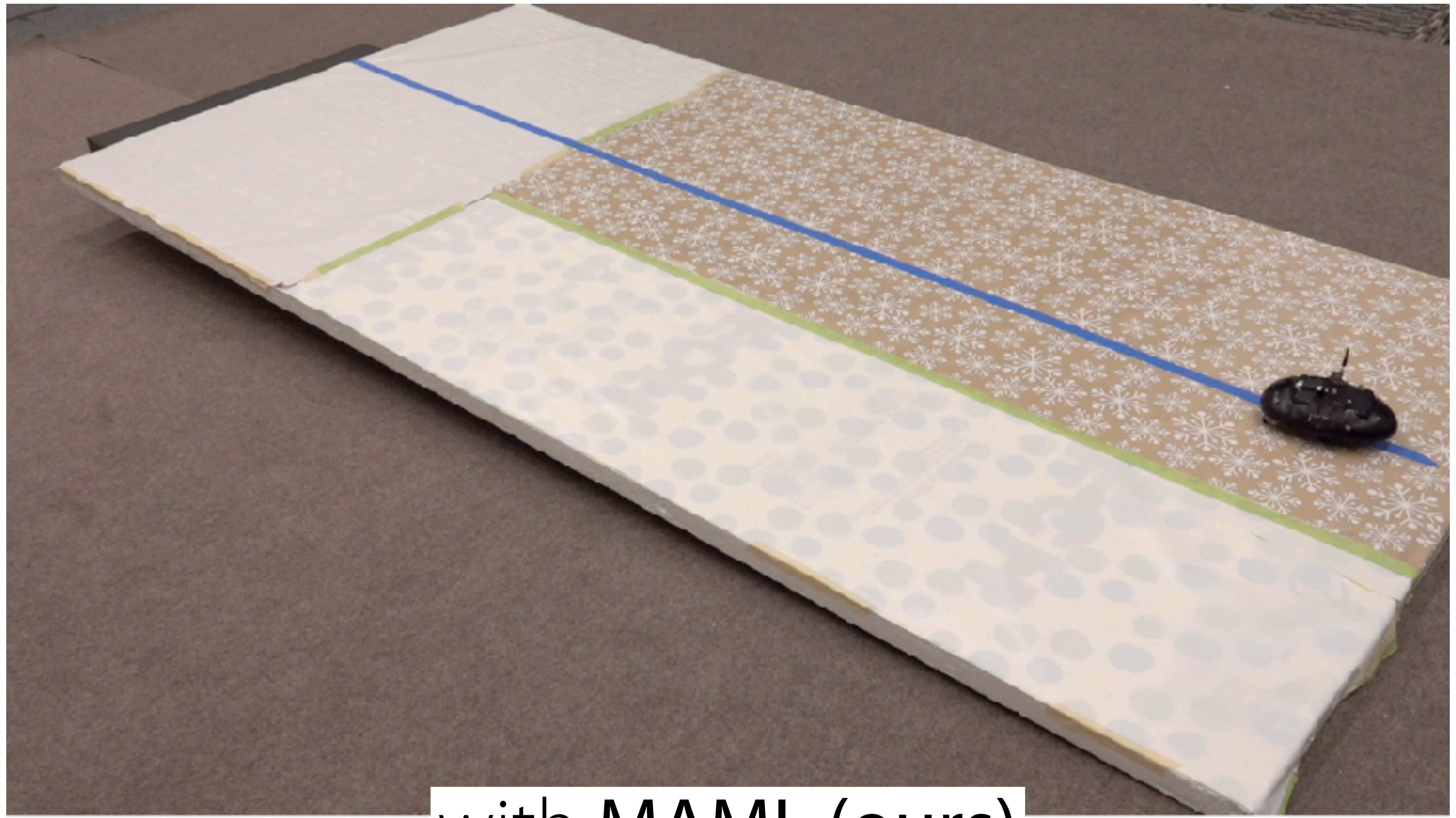
VelociRoACH Robot

Meta-train on **variable terrains**

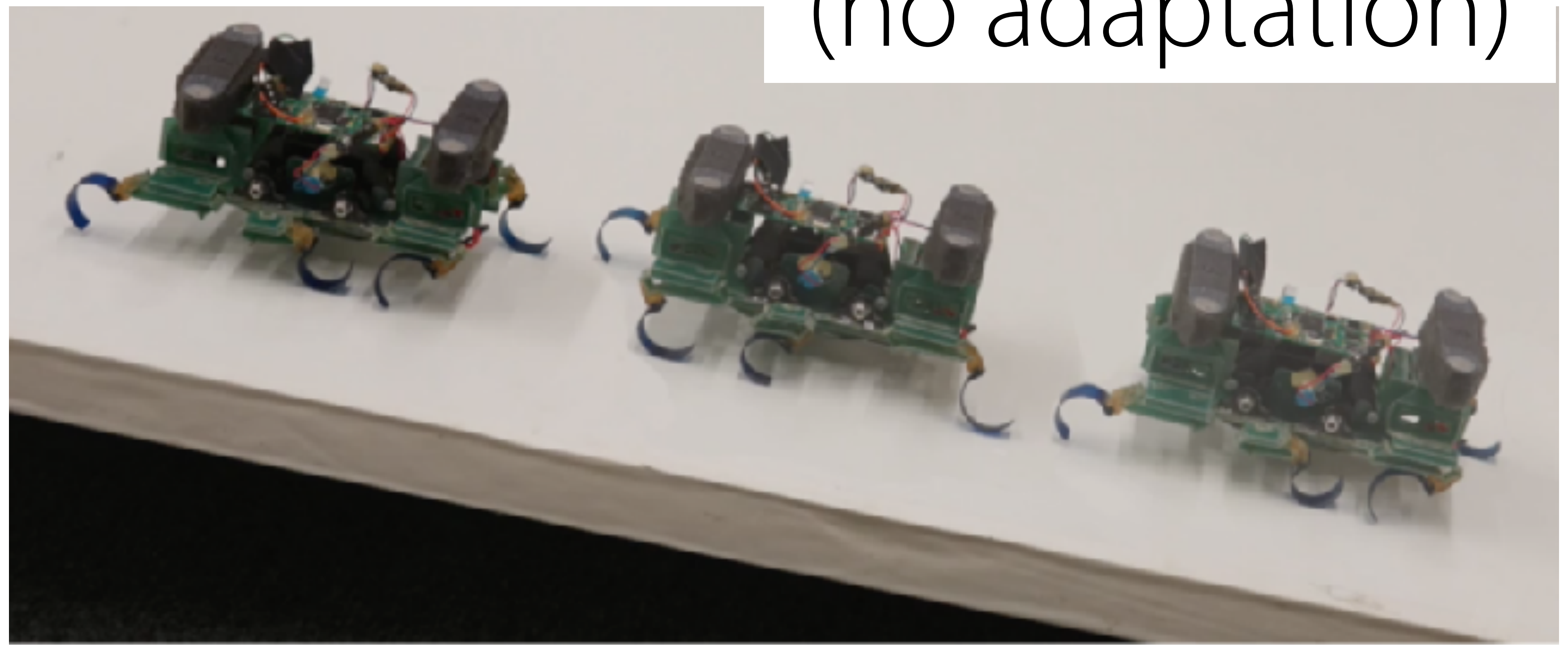
Meta-test with **slope**, **missing leg, payload, calibration errors**



model-based RL
(no adaptation)



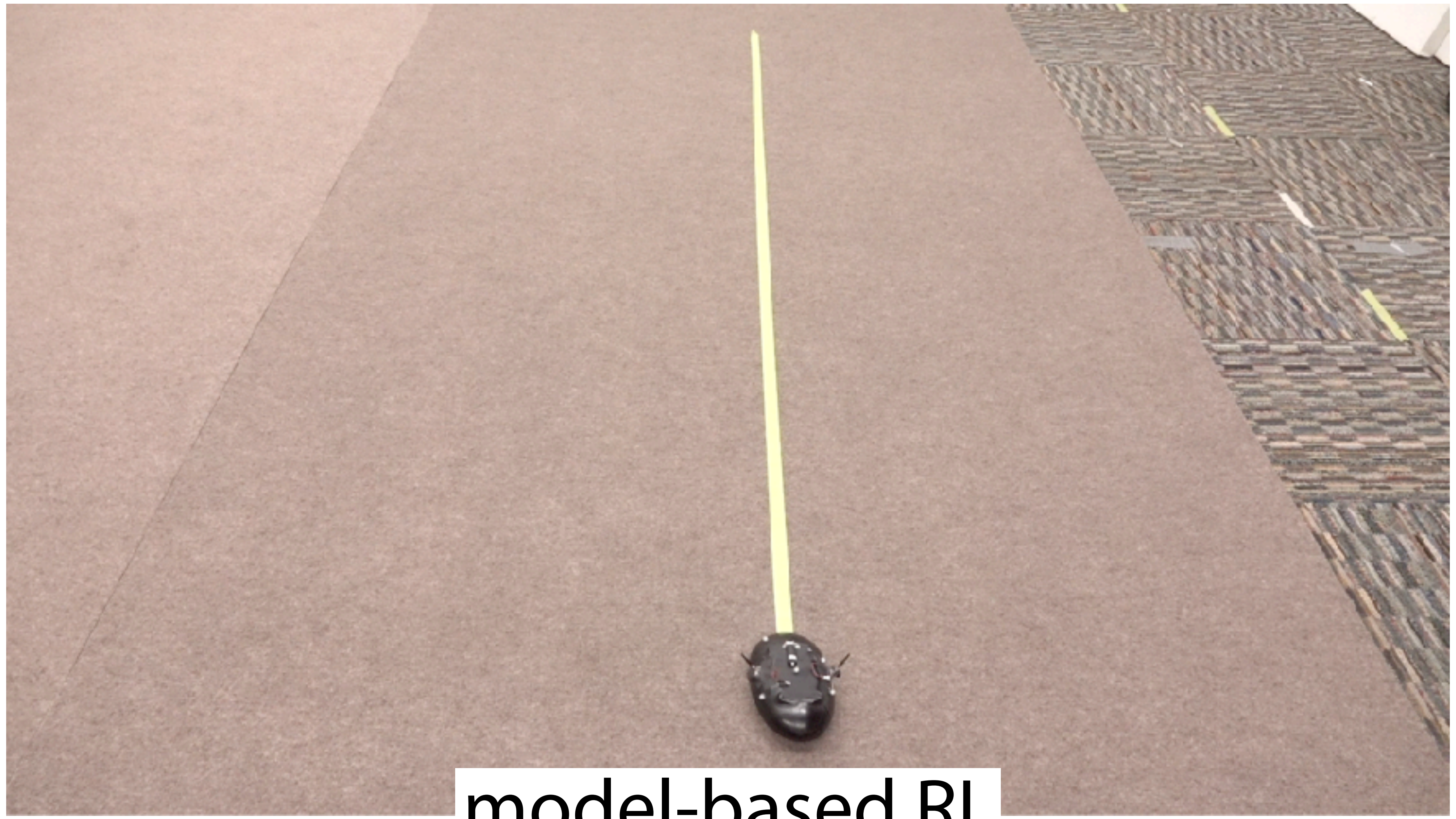
with **MAML (ours)**



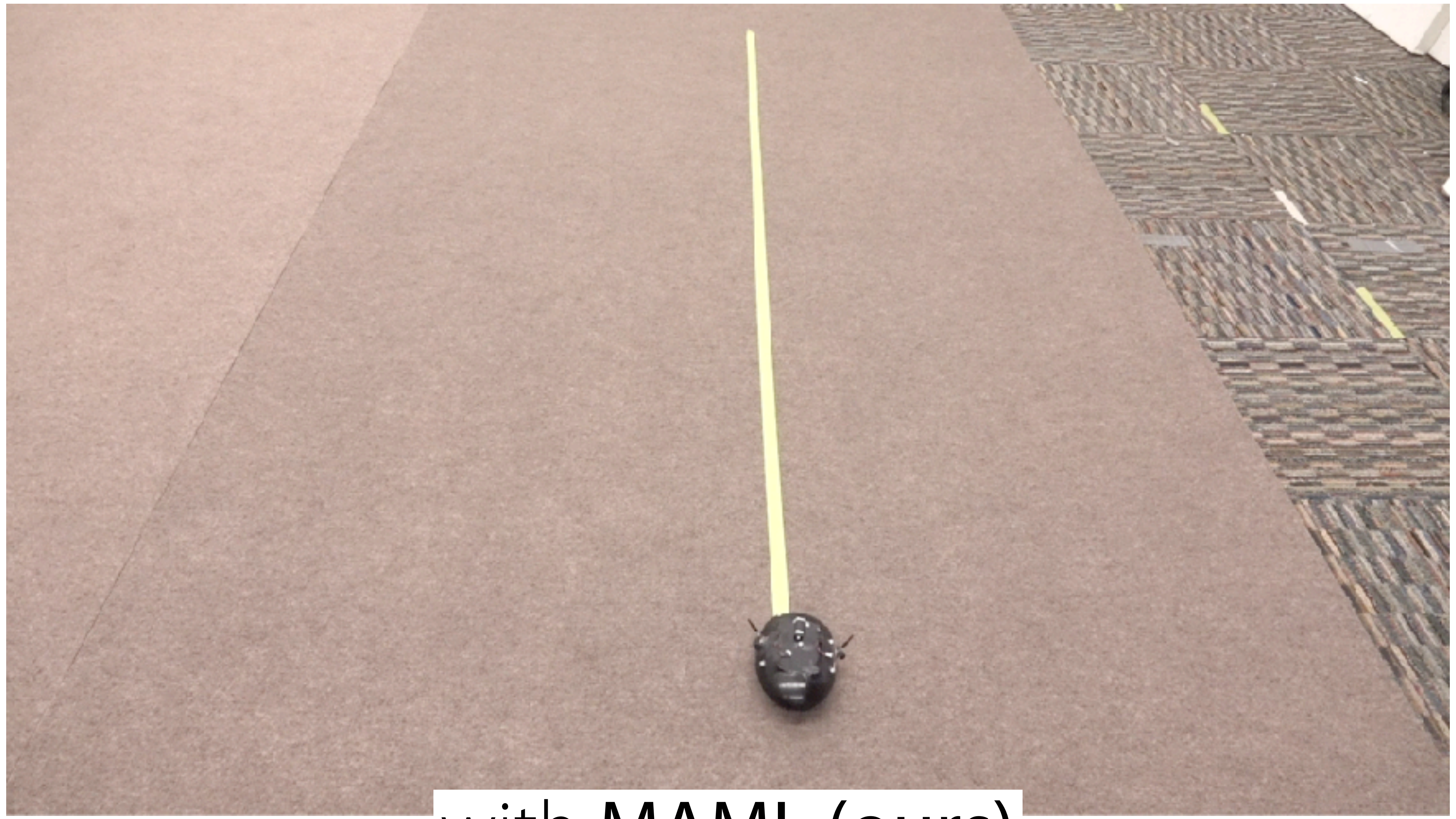
Nagabandi*, Clavera*, Liu, Fearing, Abbeel, Levine, Finn. Learning to Adapt in Dynamic Environments through Meta-RL. ICLR'19

VelociRoACH Robot

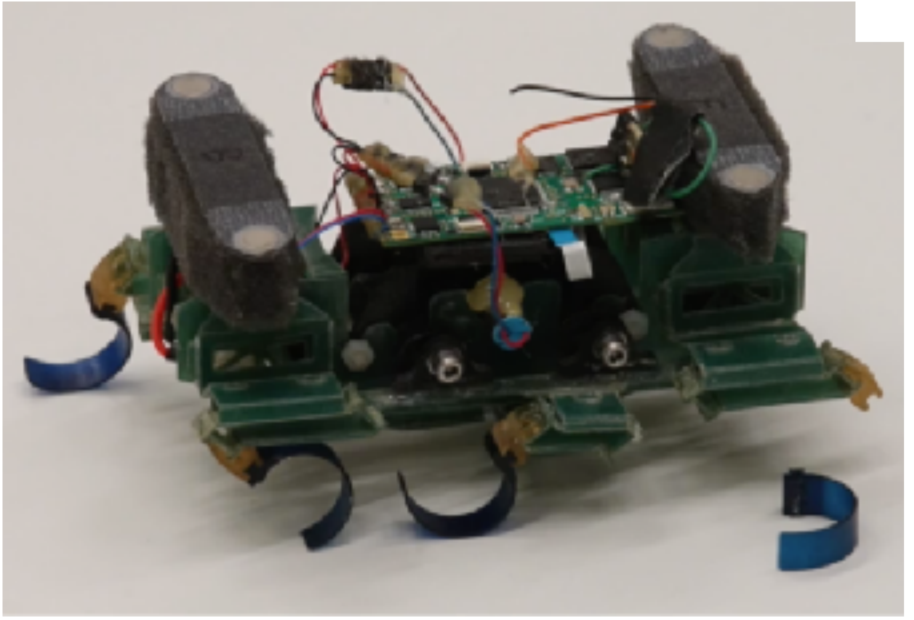
Meta-train on **variable terrains** **Meta-test** with **slope, <u>missing leg</u>, payload, calibration errors**



model-based RL
(no adaptation)



with **MAML (ours)**



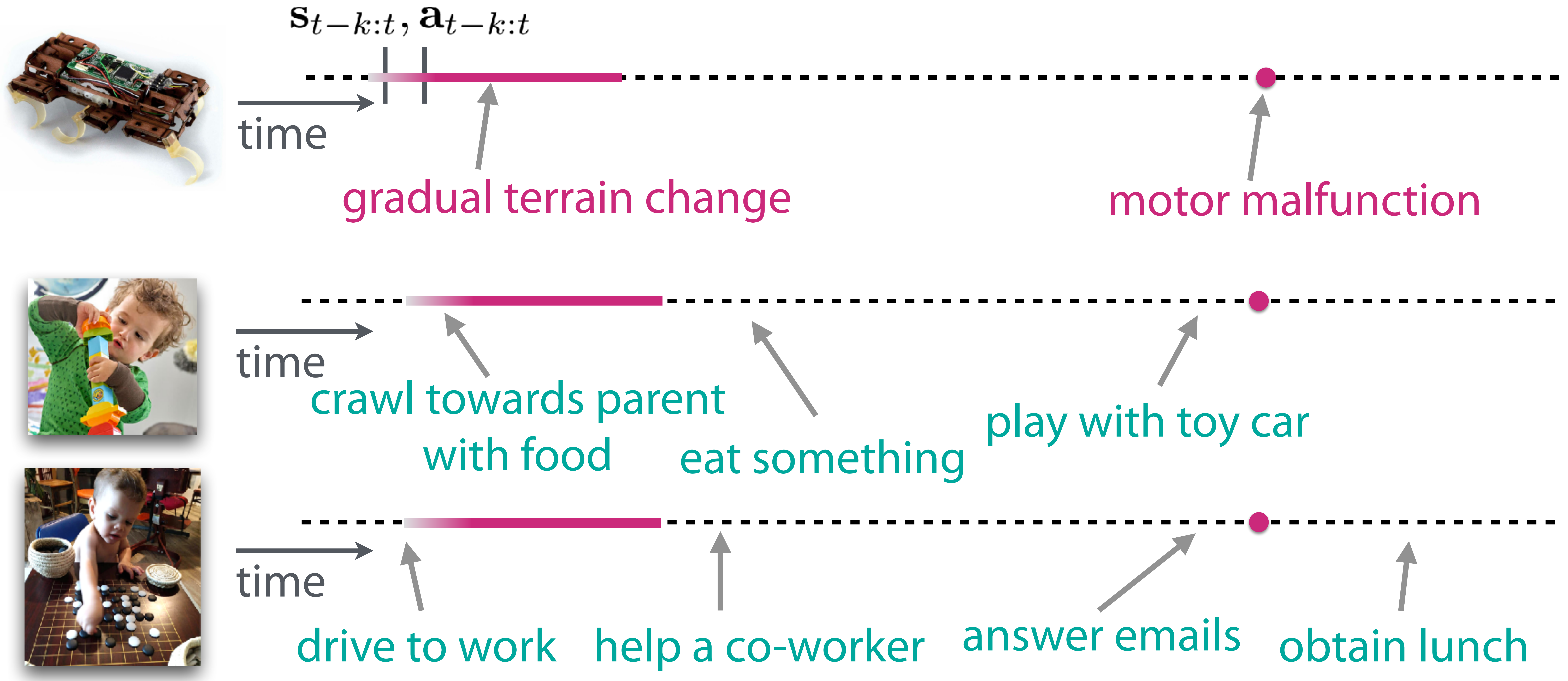
Nagabandi*, Clavera*, Liu, Fearing, Abbeel, Levine, Finn. Learning to Adapt in Dynamic Environments through Meta-RL. ICLR'19

Where do the tasks come from?

self-driven: propose your own tasks

environment-driven: dynamic, real-world environment

Future work: combine both.



sufficiently complex environment + *basic drives* → realistic & complex task distribution

e.g., the real world

survival, belonging, social interaction, curiosity, fulfillment

Open Problems in Meta-Learning

Data

Where do the tasks come from?
(defining one reward function is hard enough!)

Do we need all tasks to be available at once?
(i.e. moving away from iid sampling from task distribution)

In practice, data is available **incrementally**.
Can we learn priors when *only a few tasks are present*?

Preliminary work: *Online Meta-Learning*, Finn*, Rajeswaran*, Kakade, Levine. ICML '19

Algorithms

Open Problems in Meta-Learning

Data

Where do the tasks come from?

(defining one reward function is hard enough!)

Do we need all tasks to be available at once?

(i.e. moving away from iid sampling from task distribution)

Algorithms

How to reason over uncertainty w.r.t. prior?

(Bayesian DL is hard)

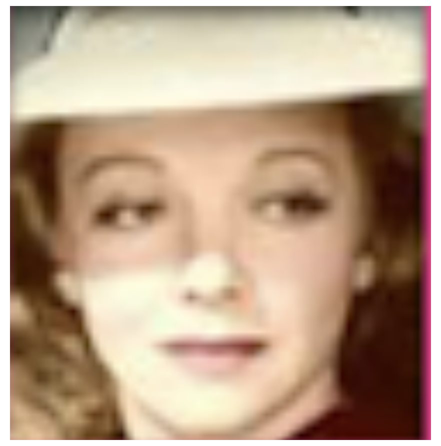
Ambiguity in Few-Shot Learning

\+ –

✓ Smiling,
✓ Wearing Hat,
✓ Young

?

✓ Smiling,
✓ Wearing Hat,
✗ Young



?

✗ Smiling,
✓ Wearing Hat,
✓ Young

Can we learn to generate hypotheses about the underlying function?

Existing meta-learning algorithms return the MAP, $\arg\max_{\phi_i} p(\phi_i|\mathbf{x}_i^{\text{train}}, \mathbf{y}_i^{\text{train}})$

Can we sample $\phi_i \sim p(\phi_i|x_i^{\text{train}}, y_i^{\text{train}})$?

Important for:

- learning to **actively learn**
- learning to **explore** in meta-RL
- trading off weight of prior & posterior in **variable-shot learning**

Variational Inference: Finn*, Xu*, Levine. Probabilistic Model-Agnostic Meta-Learning. NeurIPS '18
Gordon*, Bronskill*, et al. Meta-Learning Probabilistic Inference for Prediction. ICLR '19

Bayesian Ensembles: Kim*, Yoon* et al. Bayesian Model-Agnostic Meta-Learning. NeurIPS '18

Open Problems in Meta-Learning

Data

Where do the tasks come from?
(defining one reward function is hard enough!)

Do we need all tasks to be available at once?
(i.e. moving away from iid sampling from task distribution)

Algorithms

How to reason over uncertainty w.r.t. prior?
(Bayesian DL is hard)

How to learn from off-policy data in meta-RL?
(fundamental challenges to making it off-policy)

Meta-training:

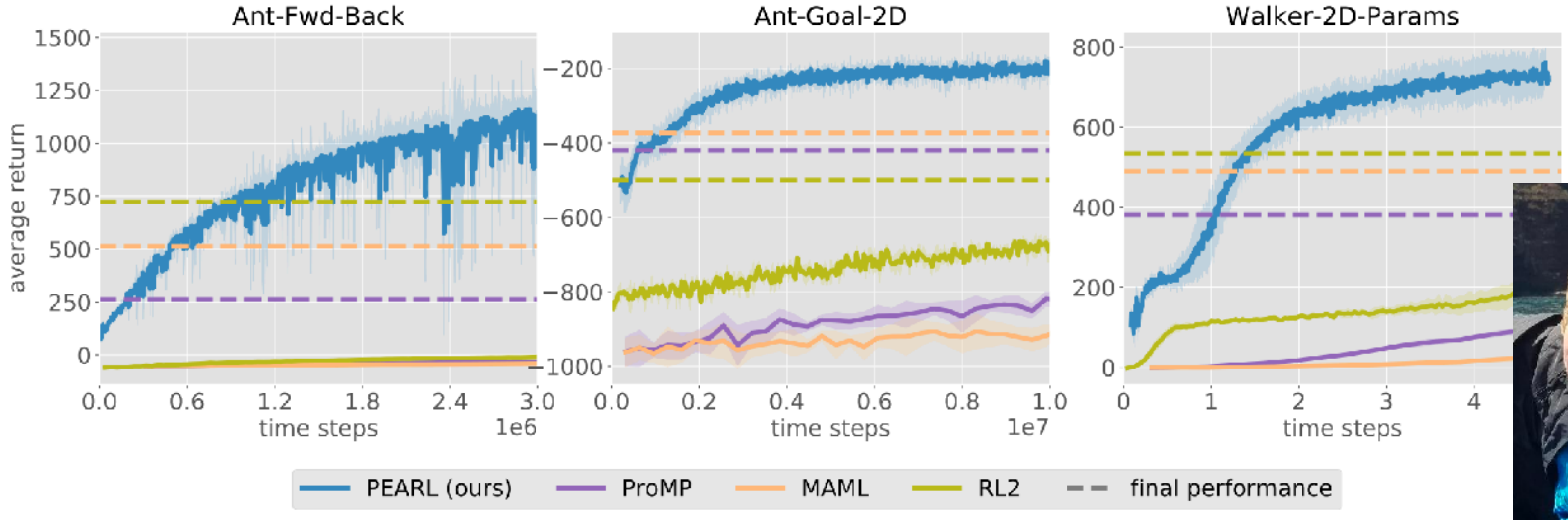
Learn from off-policy data

Meta-test time:

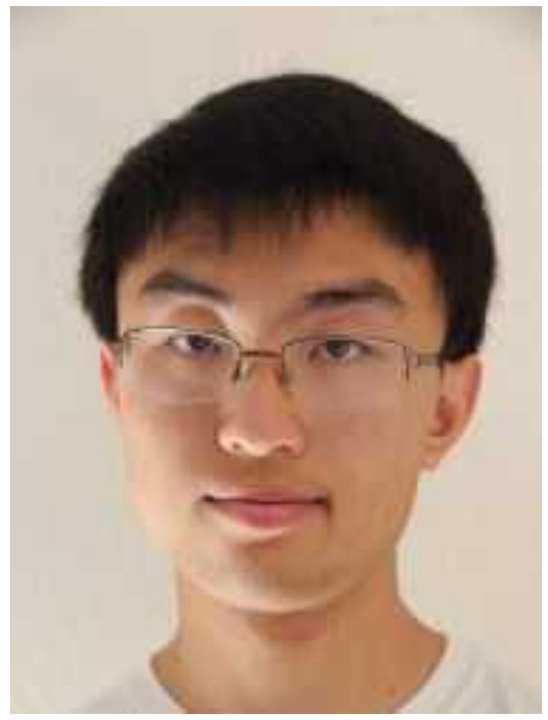
Iteratively explore, collect data, update policy ⟵ inherently on-policy!

We have a mismatch…

Proposed Solution: *pseudo on-policy data* for adaptation, *off-policy data* for meta-optimization
Thompson sampling from learned latent task variable to explore



Kate Rakelly



Aurick Zhou

Rakelly*, Zhou*, Quillen, Finn, Levine. Efficient Off-Policy Meta-Reinforcement Learning via Probabilistic Context Variables. ICML '19

Open Problems in Meta-Learning

Data

Where do the tasks come from?

(defining one reward function is hard enough!)

Do we need all tasks to be available at once?

(i.e. moving away from iid sampling from task distribution)

Algorithms

How to reason over uncertainty w.r.t. prior?

(Bayesian DL is hard)

How to learn from off-policy data in meta-RL?

(fundamental challenges to making it off-policy)

realistic & complex task distributions + powerful meta-learning algs

recover effective priors over intelligent behavior

Collaborators

Sergey Levine

Pieter Abbeel

Tianhe Yu

Anusha Nagabandi

Ignasi Clavera

Kate Rakelly

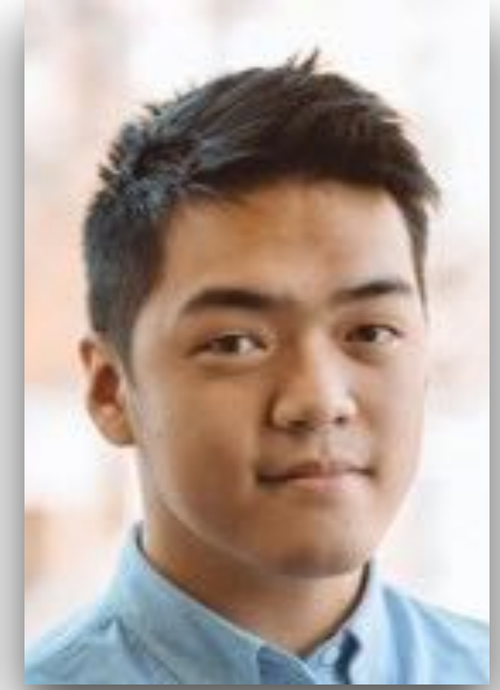
Aurick Zhou

Annie Xie

Sudeep Dasari

Simin Liu

Kyle Hsu



Abhishek Gupta

Ben Eysenbach

Kelvin Xu

Papers, **data** and **code** linked at: people.eecs.berkeley.edu/~cbfinn

Questions?