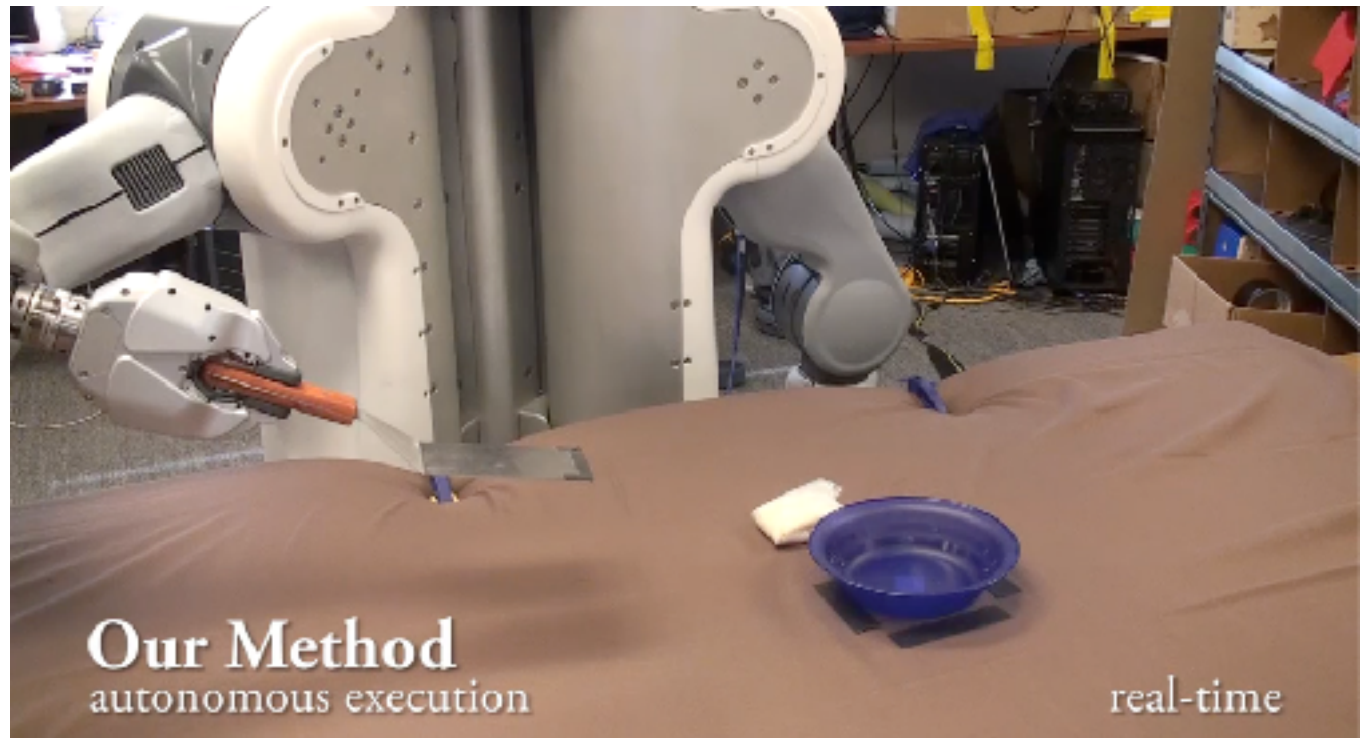
Flexible Neural Networks and the Frontiers of Meta-Learning

Chelsea Finn

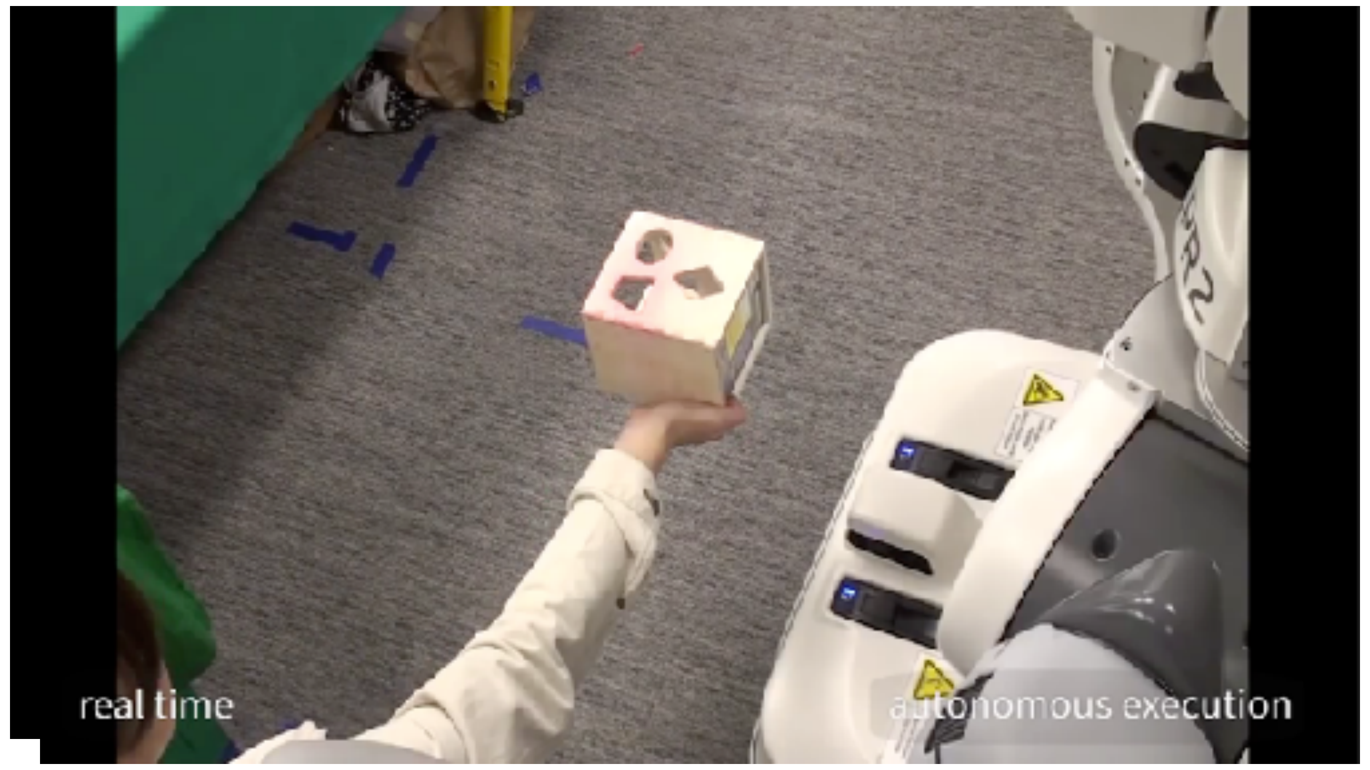
BERKELEY ARTIFICIAL INTELLIGENCE RESEARCH

How can we enable agents to learn skills in the real world?

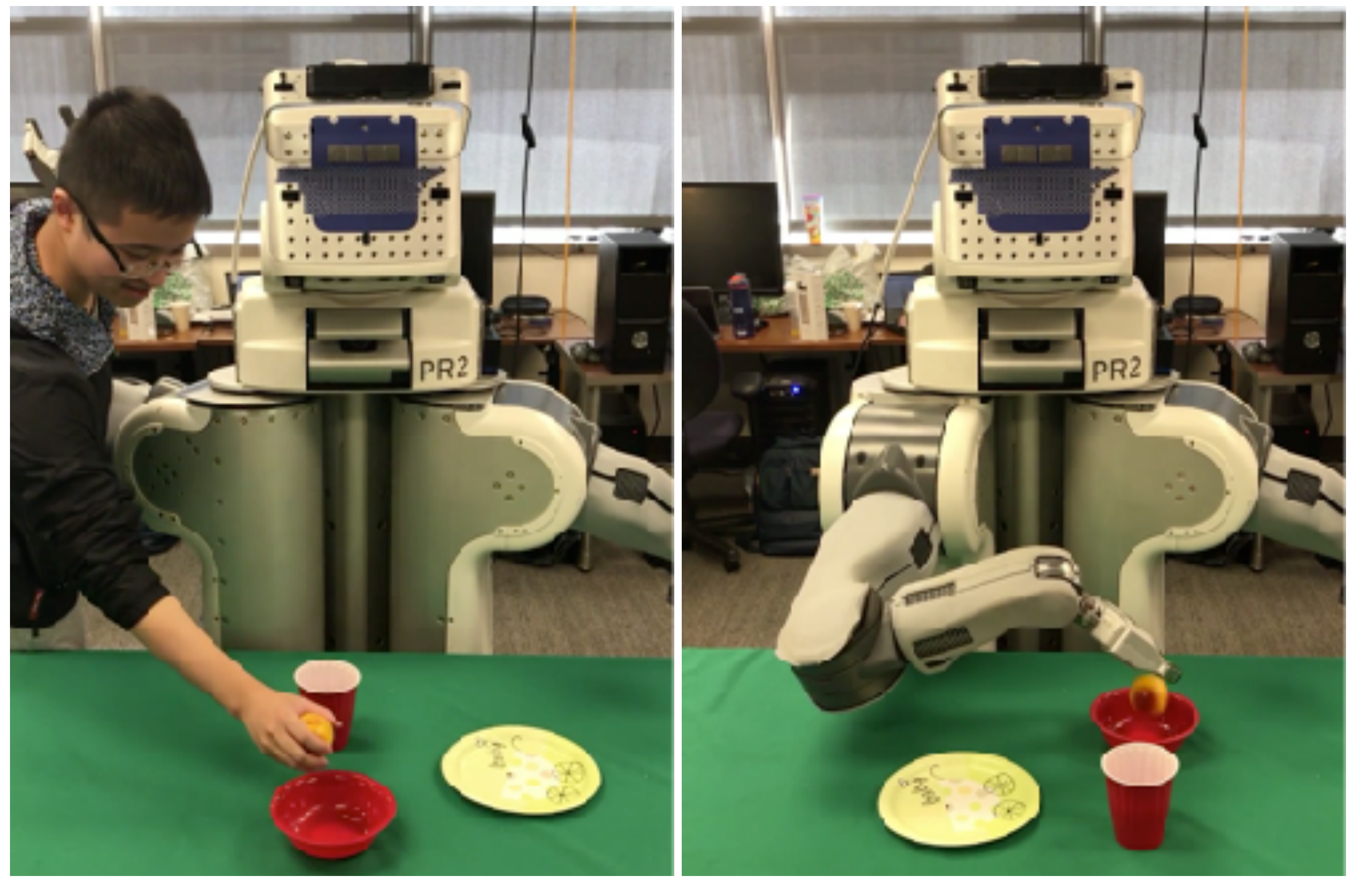
Robots.



Finn, Tan, Duan, Darrell, Levine, Abbeel. ICRA '16



Levine*, Finn*, Darrell, Abbeel. JMLR '16



Yu*, Finn*, Xie, Dasari, Zhang, Abbeel, Levine, RSS '18

Why robots? Robots can teach us things about intelligence.

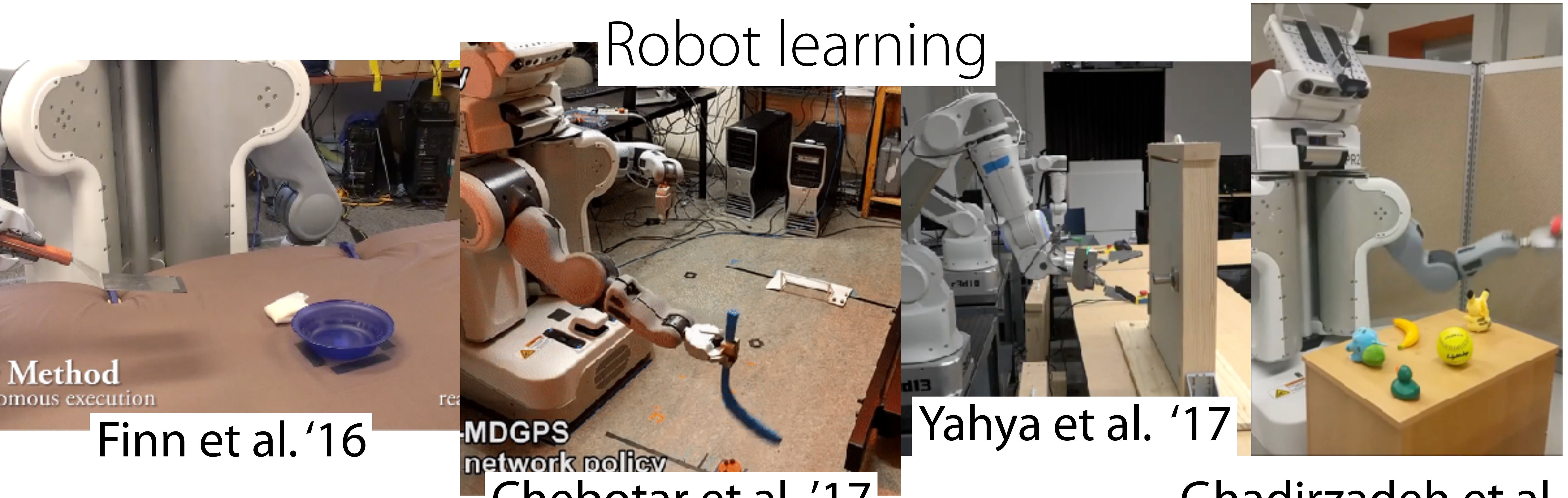
faced with the **real world**

must **generalize** across tasks, objects, environments, etc

need some **common sense understanding** to do well

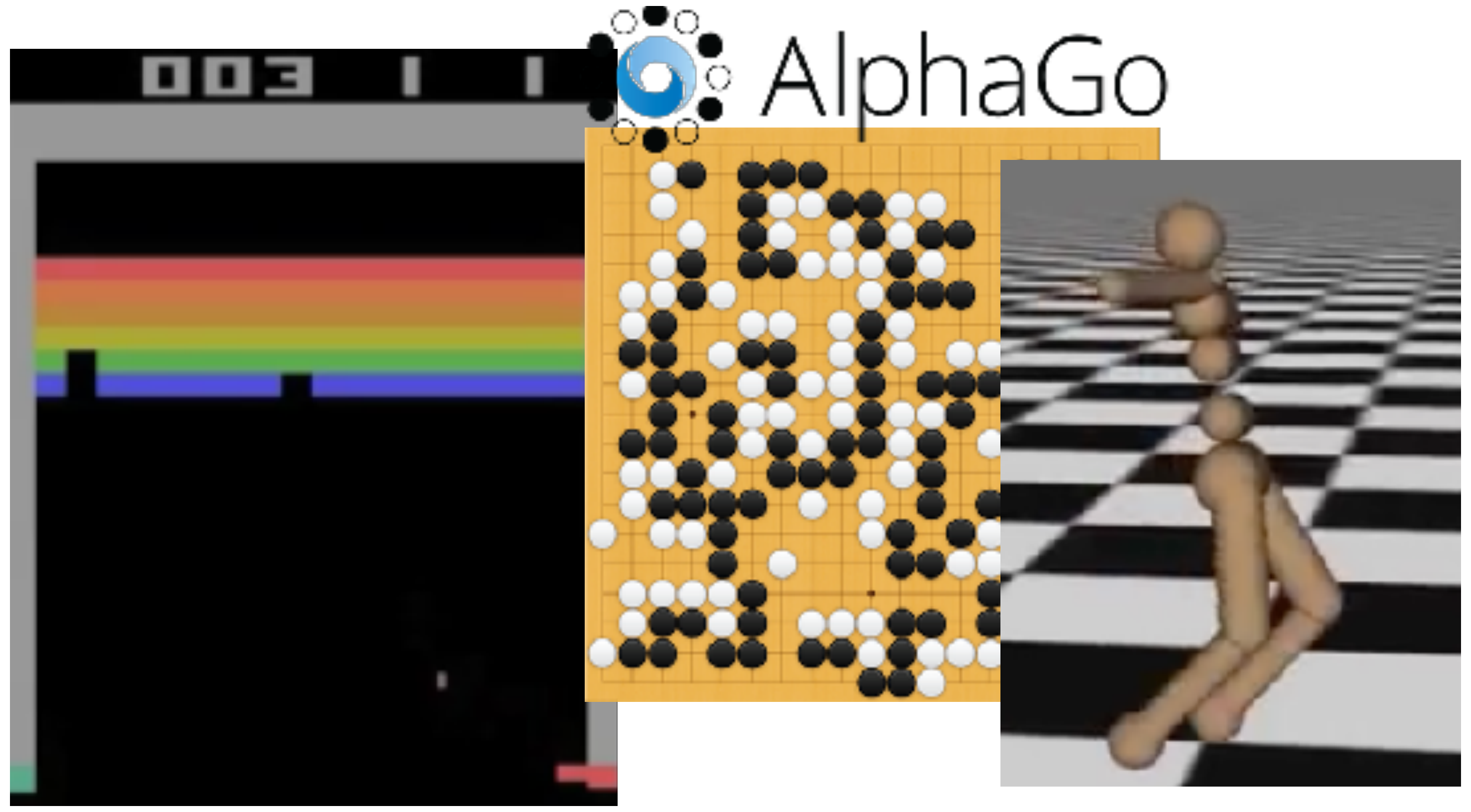
supervision can't be taken for granted

Robot learning



Finn et al. '16   Chebotar et al. '17   Yahya et al. '17   Ghadirzadeh et al. '17

Reinforcement learning



Atari   locomotion

Learn **one task** in **one environment**, **starting from scratch**

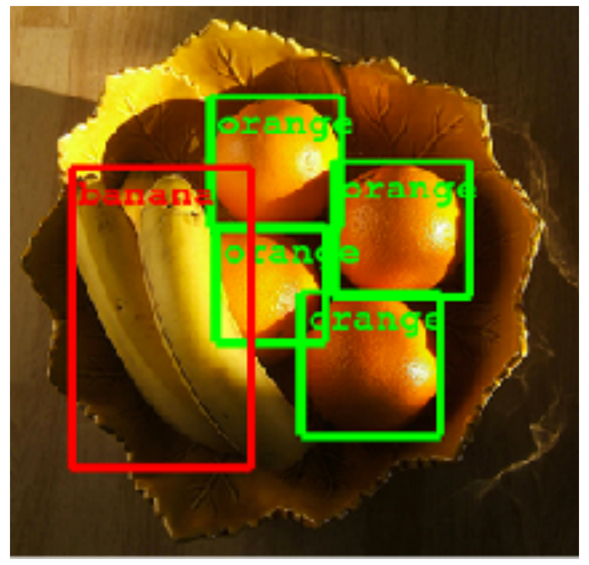
rely on **detailed reward feedback**.

Not just a problem with reinforcement learning & robotics.

specialists

machine translation

speech recognition



object detection

More diverse, yet still **one task**, **from scratch**, with **detailed supervision**

Can we enable systems to **accumulate experiences** and **acquire general-purpose priors** that enable fast learning and reasoning?

training data

Braque Cezanne

test datapoint

By Braque or Cezanne?

How did you accomplish this?

Through previous experience.

How should we incorporate prior experience into ML systems?

Modeling image formation

Geometry

SIFT features, HOG features + SVM

Fine-tuning from ImageNet features

Domain adaptation from other painters

???

Fewer human priors,
more data -driven priors

Greater success.

Can we explicitly learn priors from previous experience
that lead to efficient downstream learning?

Can we learn to learn?

How should we incorporate **prior experience** into ML systems?

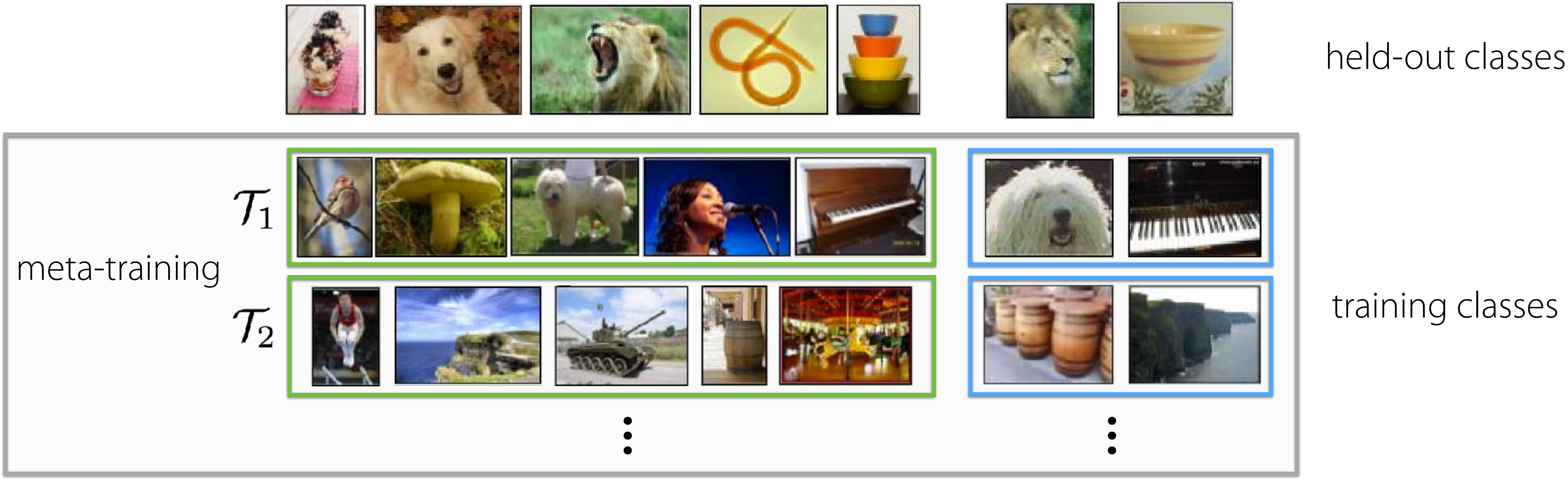
First: a primer on meta-learning & what it can accomplish

Second: challenges & frontiers

Example: Few-Shot Image Classification

5-way, 1-shot image classification (MiniImagenet)

Given 1 example of 5 classes: Classify new examples



Can replace image classification with: **regression, language generation, skill learning, any ML problem**

The Meta-Learning Problem: The Mechanistic View

^
typical

Supervised Learning:

Inputs: $\mathbf{x}$ Outputs: $\mathbf{y}$ Data: $\{(\mathbf{x}, \mathbf{y})_i\}$

$$\mathbf{y} = f(\mathbf{x}; \theta)$$

Meta-Supervised Learning:

Inputs: $\mathcal{D}^{\text{tr}}$ $\mathbf{x}_{\text{test}}$ Outputs: $\mathbf{y}_{\text{test}}$ Data: $\{\mathcal{D}_i\}$

$$\overbrace{\{(\mathbf{x}, \mathbf{y})_{1:K}\}}^{\mathcal{D}^{\text{tr}}}$$

$$\mathcal{D}_i : \{(\mathbf{x}, \mathbf{y})_j\}$$

$$\mathbf{y}_{\text{test}} = f(\mathcal{D}^{\text{tr}}, \mathbf{x}_{\text{test}}; \theta)$$

Why is this view useful?

Reduces the problem to the design & optimization of *f*.

Finn. Learning to Learn with Gradients. PhD thesis 2018

The Meta-Learning Problem: The Probabilistic View

Supervised Learning:

Inputs: $\mathbf{x}$ Outputs: $\mathbf{y}$ Data: $\{(\mathbf{x}, \mathbf{y})_i\}$

$$\mathbf{y} = f(\mathbf{x}; \theta)$$

As inference: $p(\theta|\mathcal{D})$

Meta-Supervised Learning:

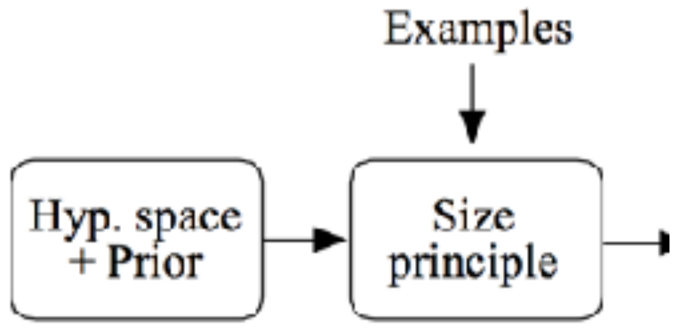
Inputs: $\underbrace{\mathcal{D}^{\mathrm{tr}}}_{\{(\mathbf{x}, \mathbf{y})_{1:K}\}}$ $\mathbf{x}_{\mathrm{test}}$ Outputs: $\mathbf{y}_{\mathrm{test}}$ Data: $\{\mathcal{D}_i\}$

$\mathcal{D}_i : \{(\mathbf{x}, \mathbf{y})_j\}$

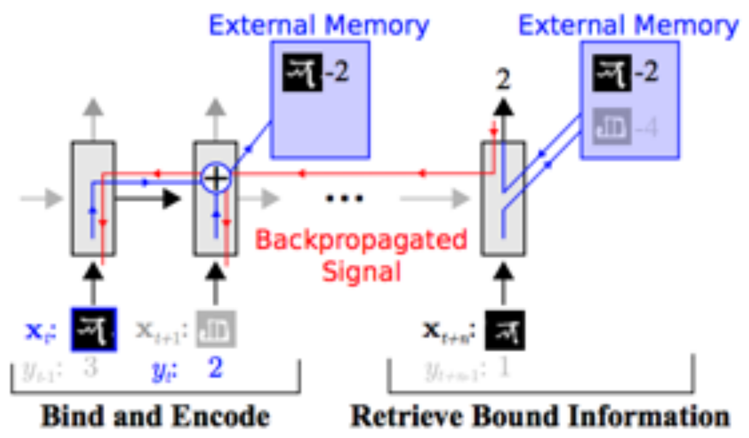
$$\mathbf{y}_{\mathrm{test}} = f(\mathcal{D}^{\mathrm{tr}}, \mathbf{x}_{\mathrm{test}}; \theta)$$

As inference: $p(\phi_i|\mathcal{D}_i^{\mathrm{tr}}, \theta)$ $\max_{\theta} \sum_i \log p(\phi_i|\mathcal{D}_i^{\mathrm{ts}})$

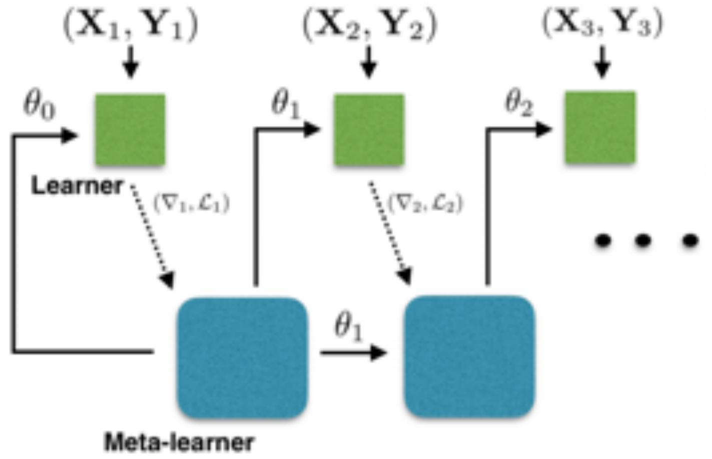
Few-Shot Learning



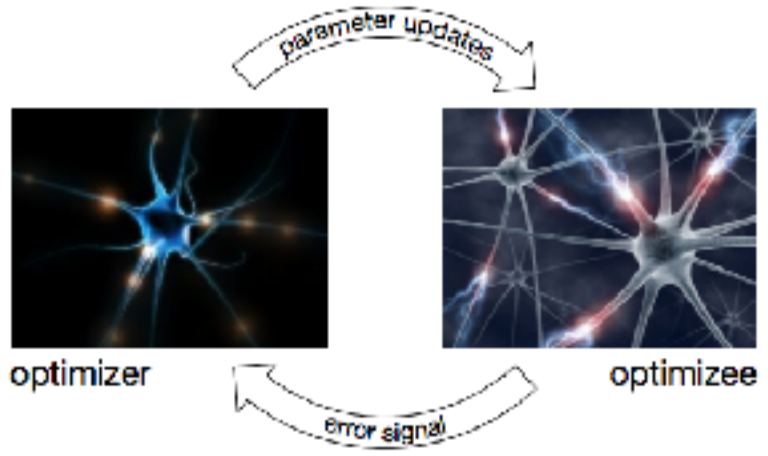
Tenenbaum '99
Fei-Fei et al. '05
Lake et al. '11



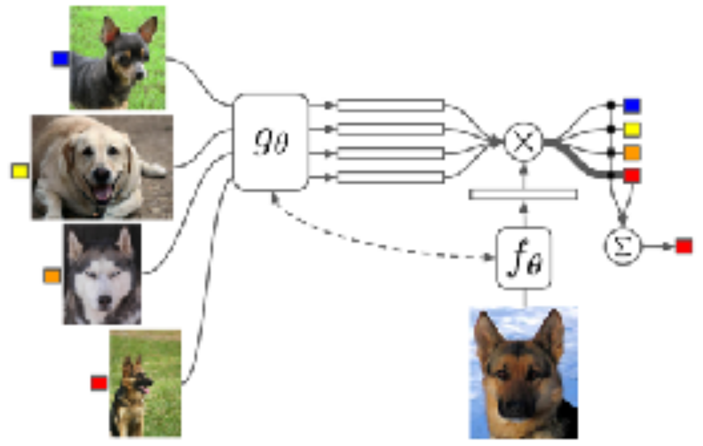
Santoro et al. '16



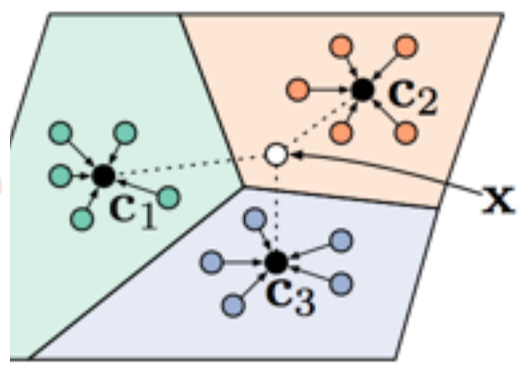
Ravi & Larochelle '17



Hochreiter et al. '01
Andrychowicz et al. '16
Li & Malik '16



Vinyals et al. '16



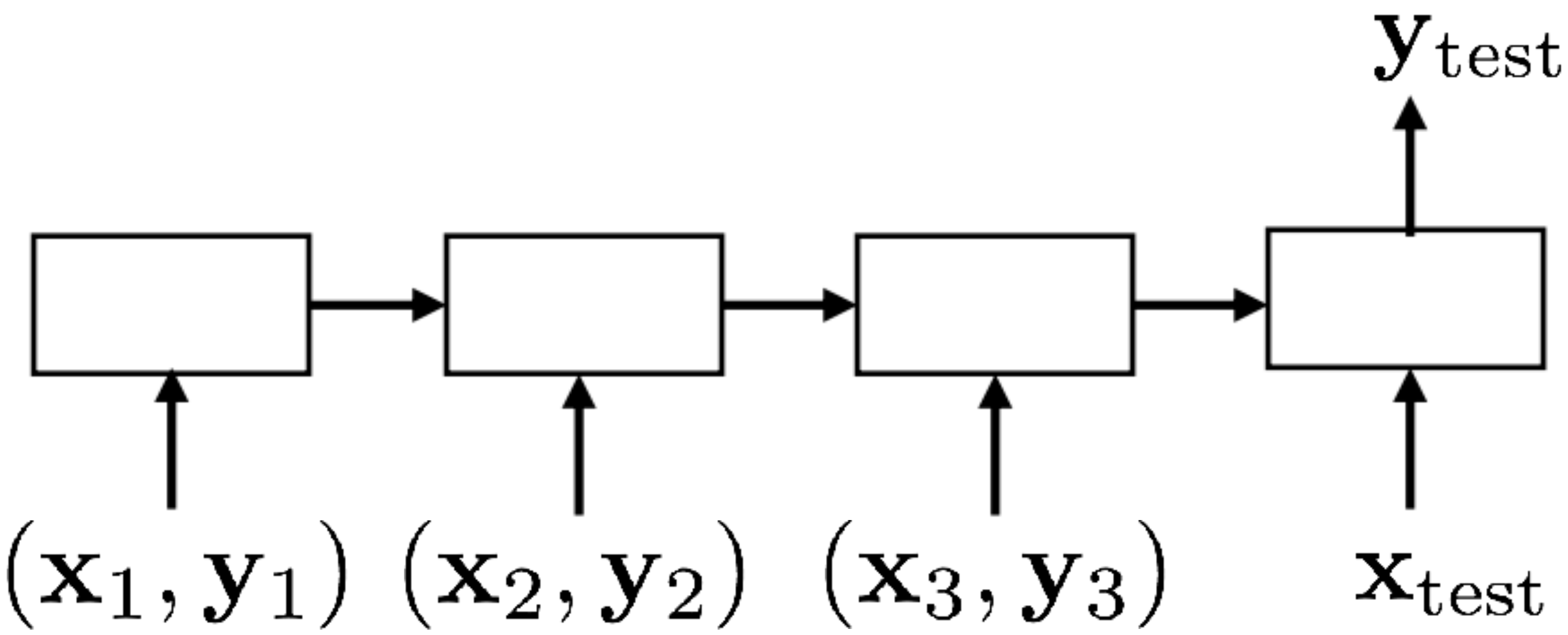
Snell et al. '17

and many *many* more approaches

Recurrent network
(LSTM, NTM, Conv)

$$\mathbf{y}_{\text{test}} = f(\mathcal{D}_{\text{train}}, \mathbf{x}_{\text{test}}; \theta)$$

Santoro et al. '16, Duan et al. '17, Wang et al. '17, Munkhdalai & Yu '17, Mishra et al. '17, …



+ expressive, general
+ applicable to range of problems

- complex model for complex task of learning
- often large data requirements

Optimization-Based Inference

pre-trained parameters

Fine-tuning

[test-time]

$$\phi \leftarrow \theta - \alpha \nabla_{\theta} \mathcal{L}(\theta, \mathcal{D}^{\mathrm{tr}})$$

training data
for new task

Meta-learning

$$\min_{\theta} \sum_{\text{task } i} \mathcal{L}(\theta - \alpha \nabla_{\theta} \mathcal{L}(\theta, \mathcal{D}_i^{\mathrm{tr}}), \mathcal{D}_i^{\mathrm{ts}})$$

Key idea: Over many tasks, learn parameter vector θ that transfers via fine-tuning

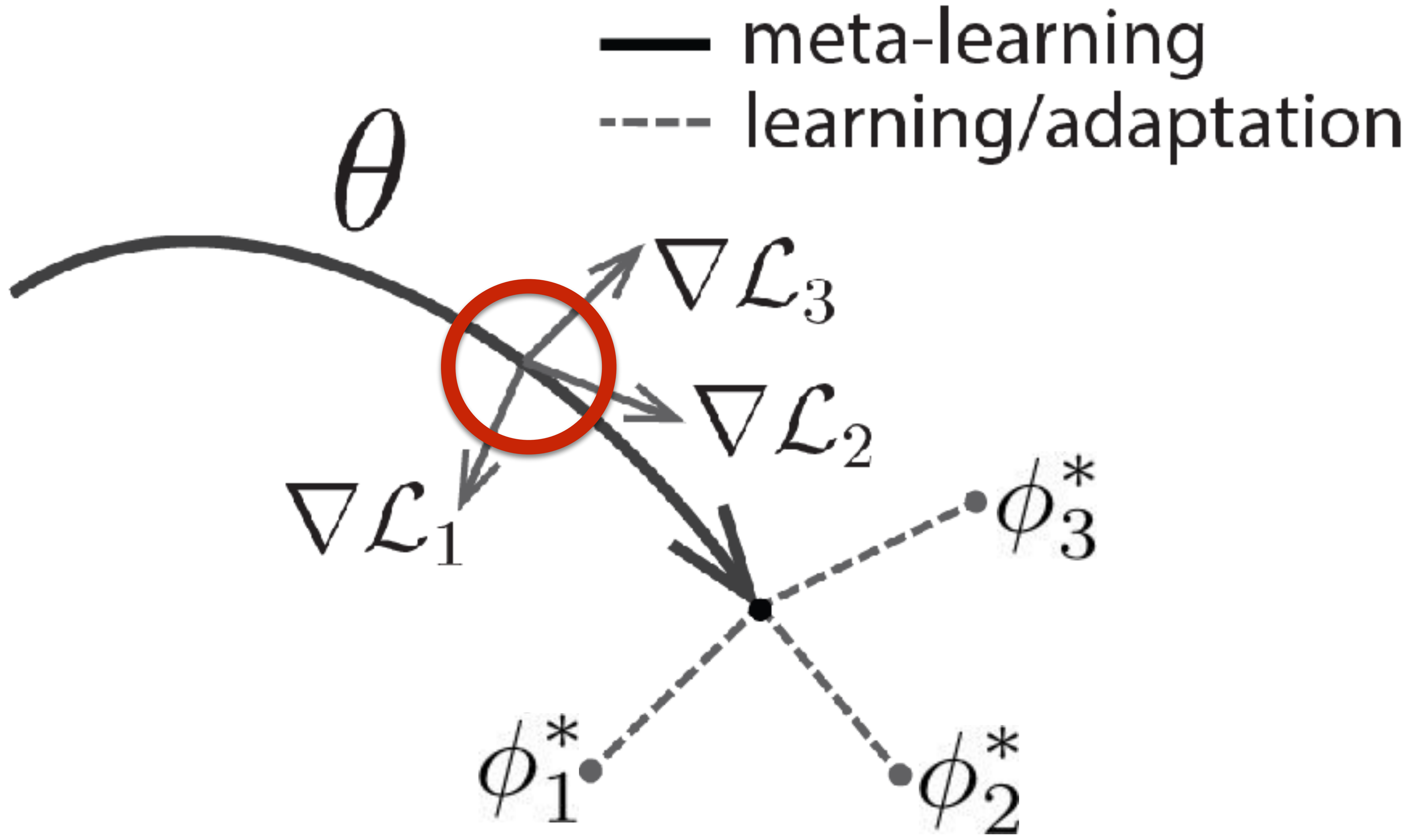
Finn, Abbeel, Levine. Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks. ICML '17

Optimization-Based Inference

$$\min_{\theta} \sum_{\text{task } i} \mathcal{L}(\theta - \alpha \nabla_{\theta} \mathcal{L}(\theta, \mathcal{D}_i^{\text{tr}}), \mathcal{D}_i^{\text{ts}})$$

θ parameter vector being meta-learned

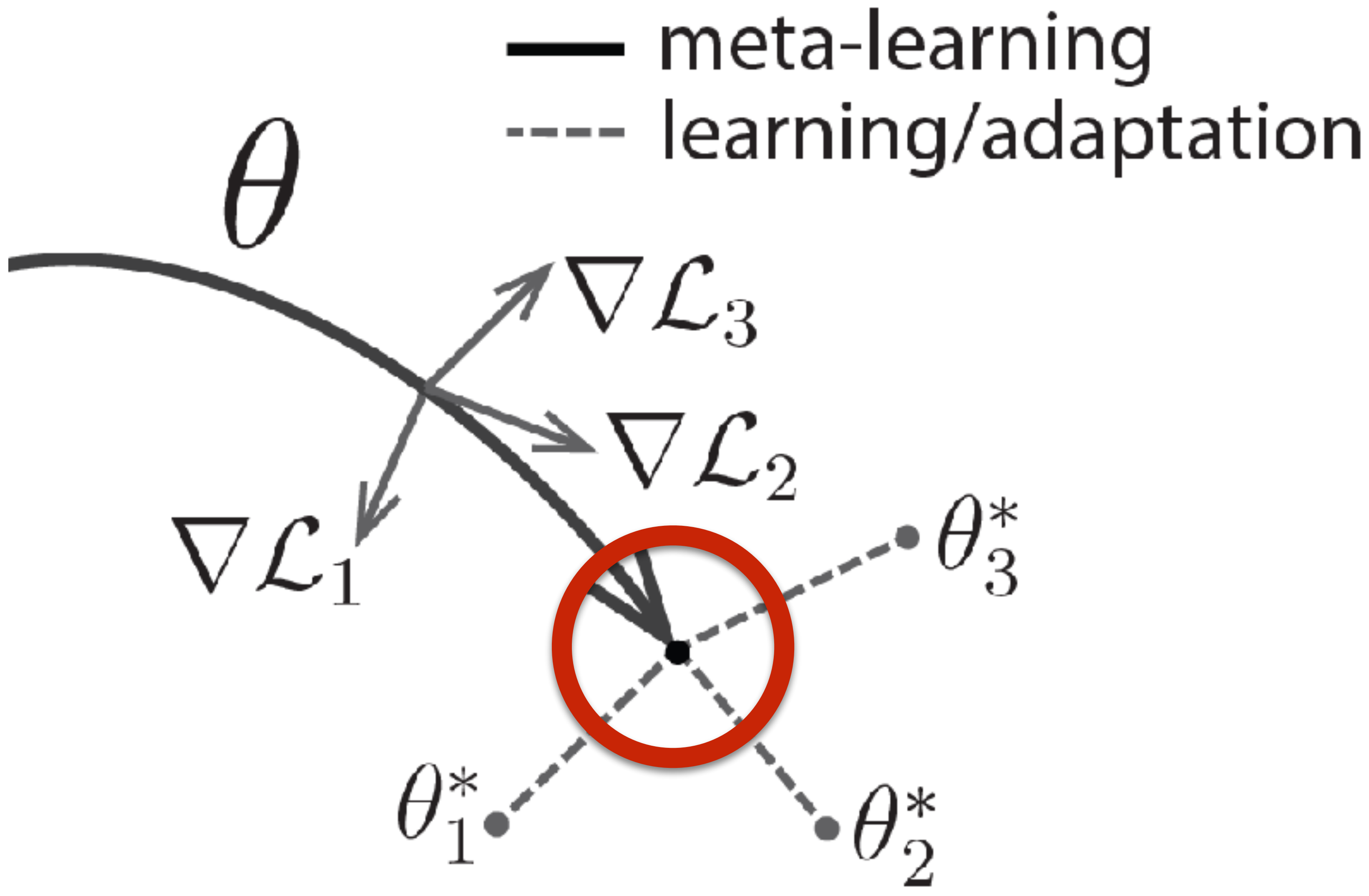
ϕ_i^* optimal parameter vector for task i

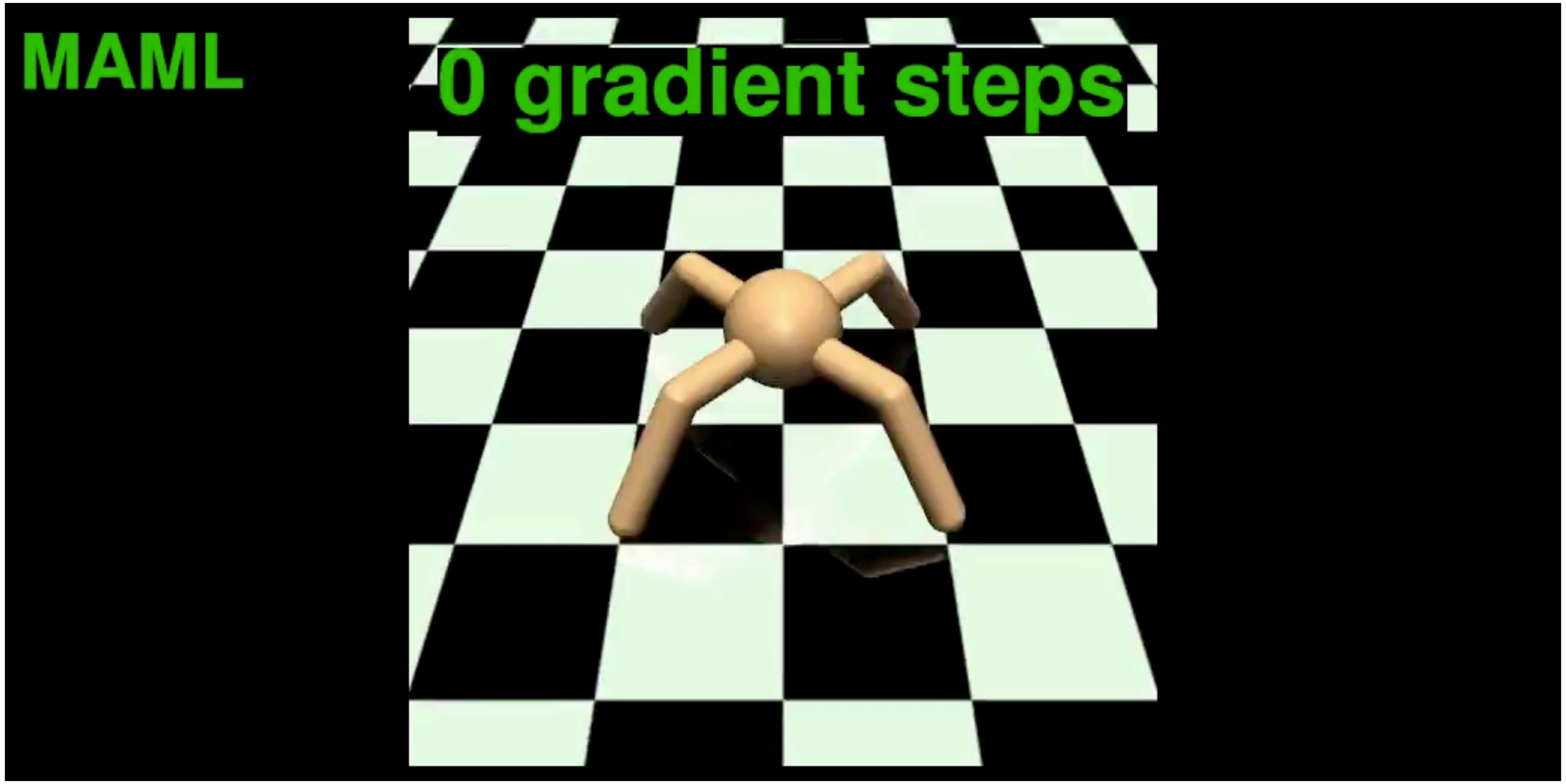


Model-**A**gnostic **M**eta-**L**earning

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

To give some intuition…

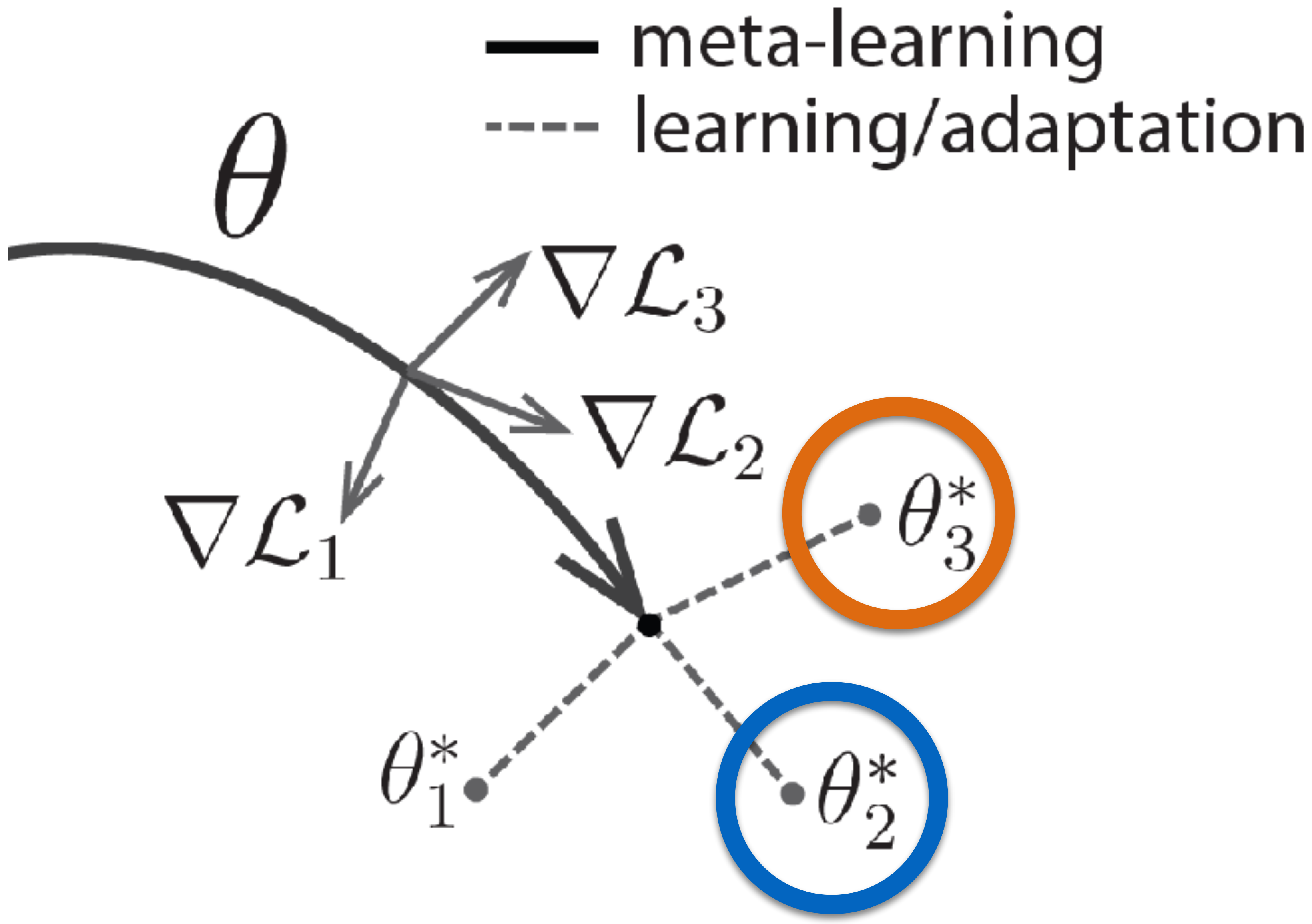


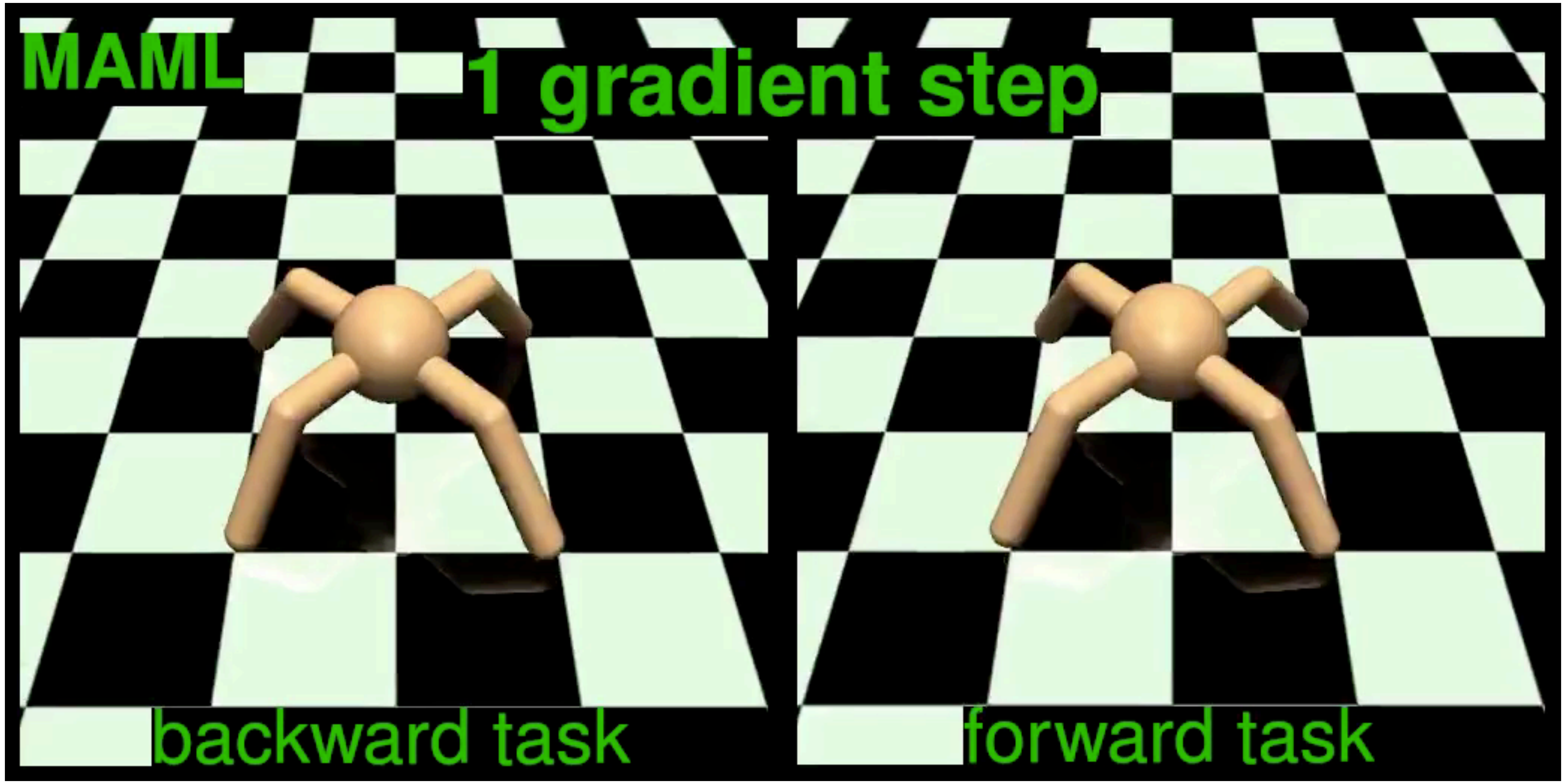


two tasks: running backward, running forward

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

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





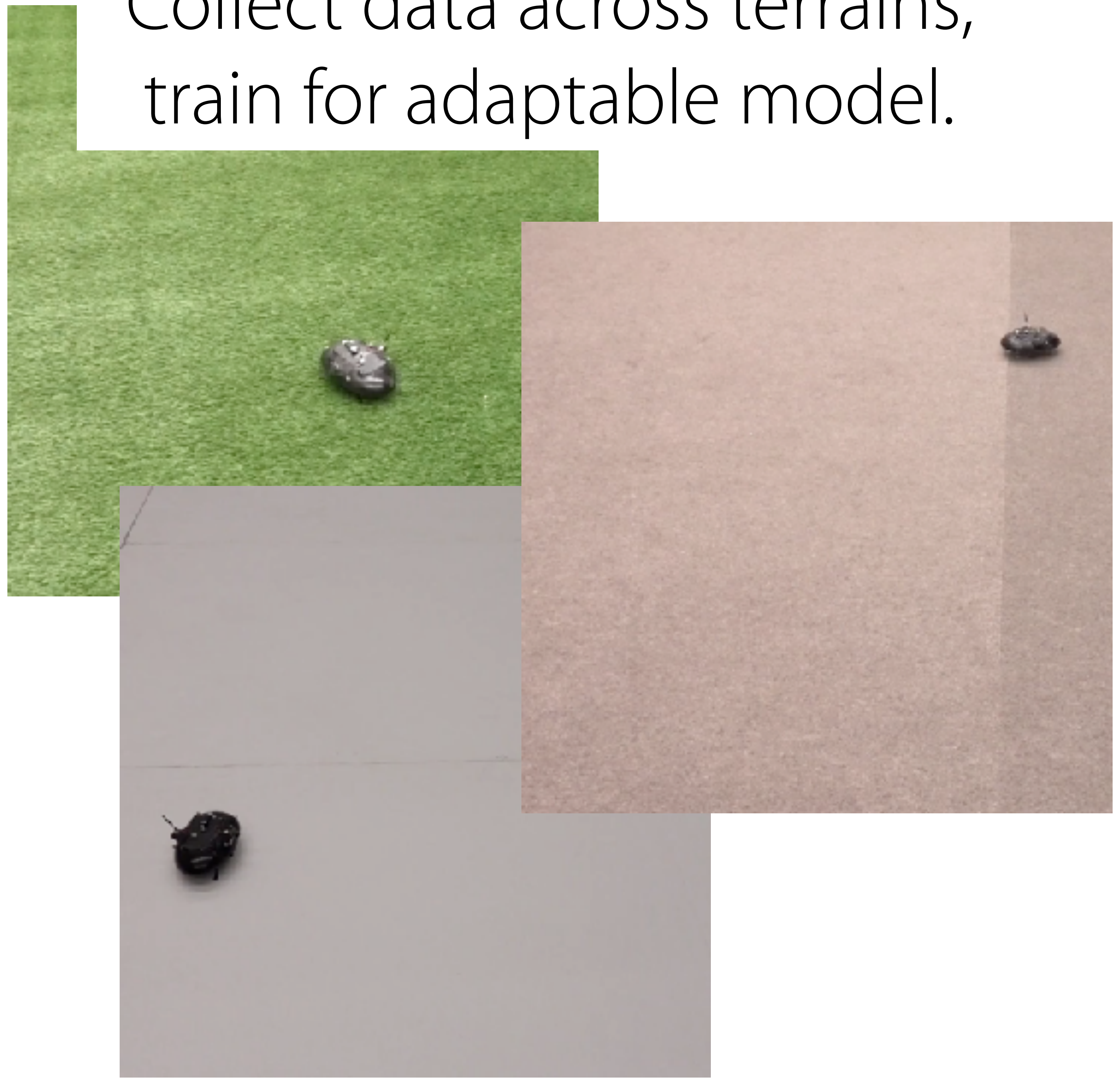
two tasks: running backward, running forward

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

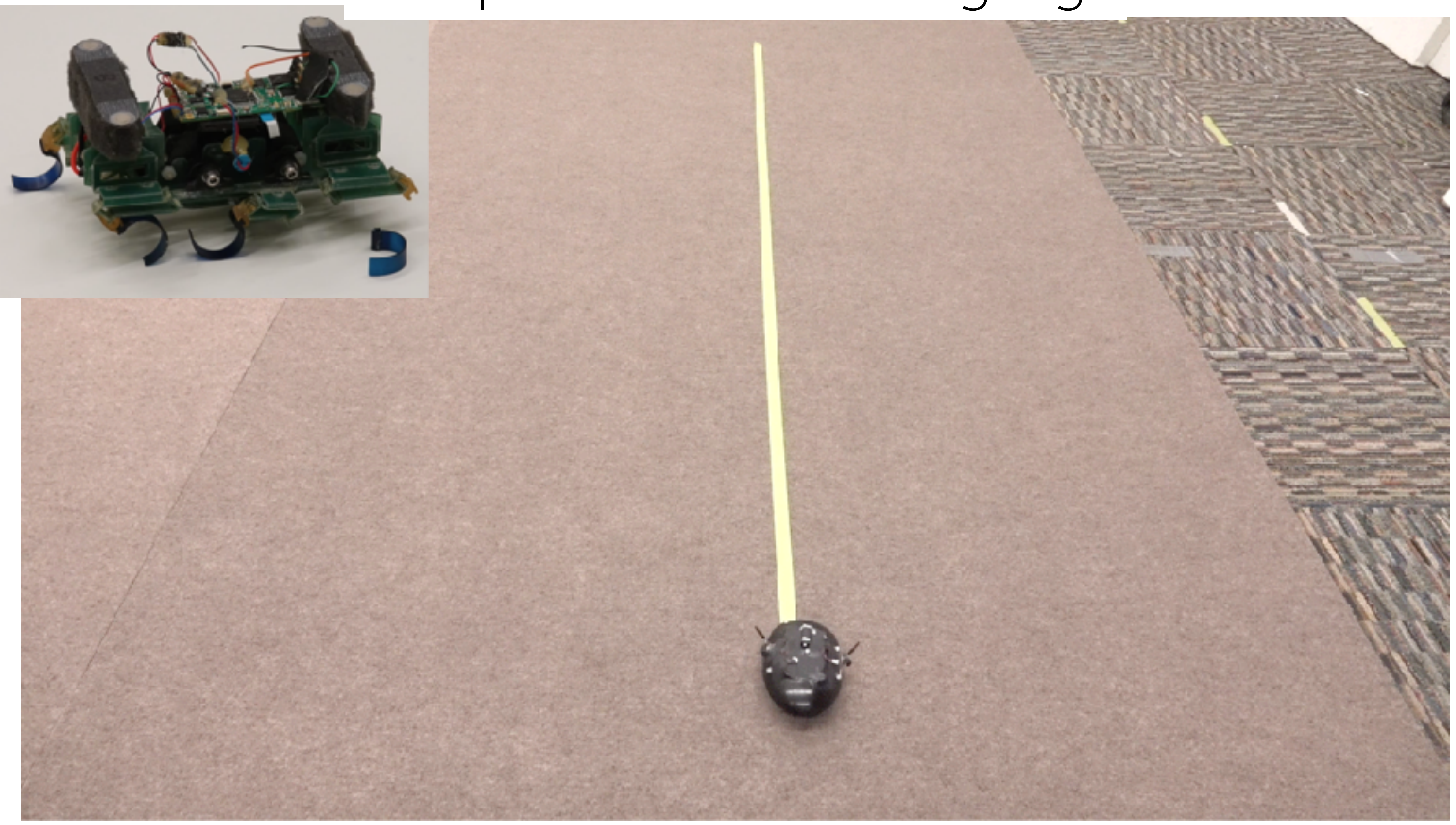
What can we do with meta-learning?

Leverage data with previous **environments** to quickly adapt to new ones?

Collect data across terrains,
train for adaptable model.



Adapt online to missing leg.



Nagabandi*, Clavera*, Liu, Fearing, Abbeel, Levine, Finn. Learning to Adapt in Dynamic Real World Environments. ICLR 2019

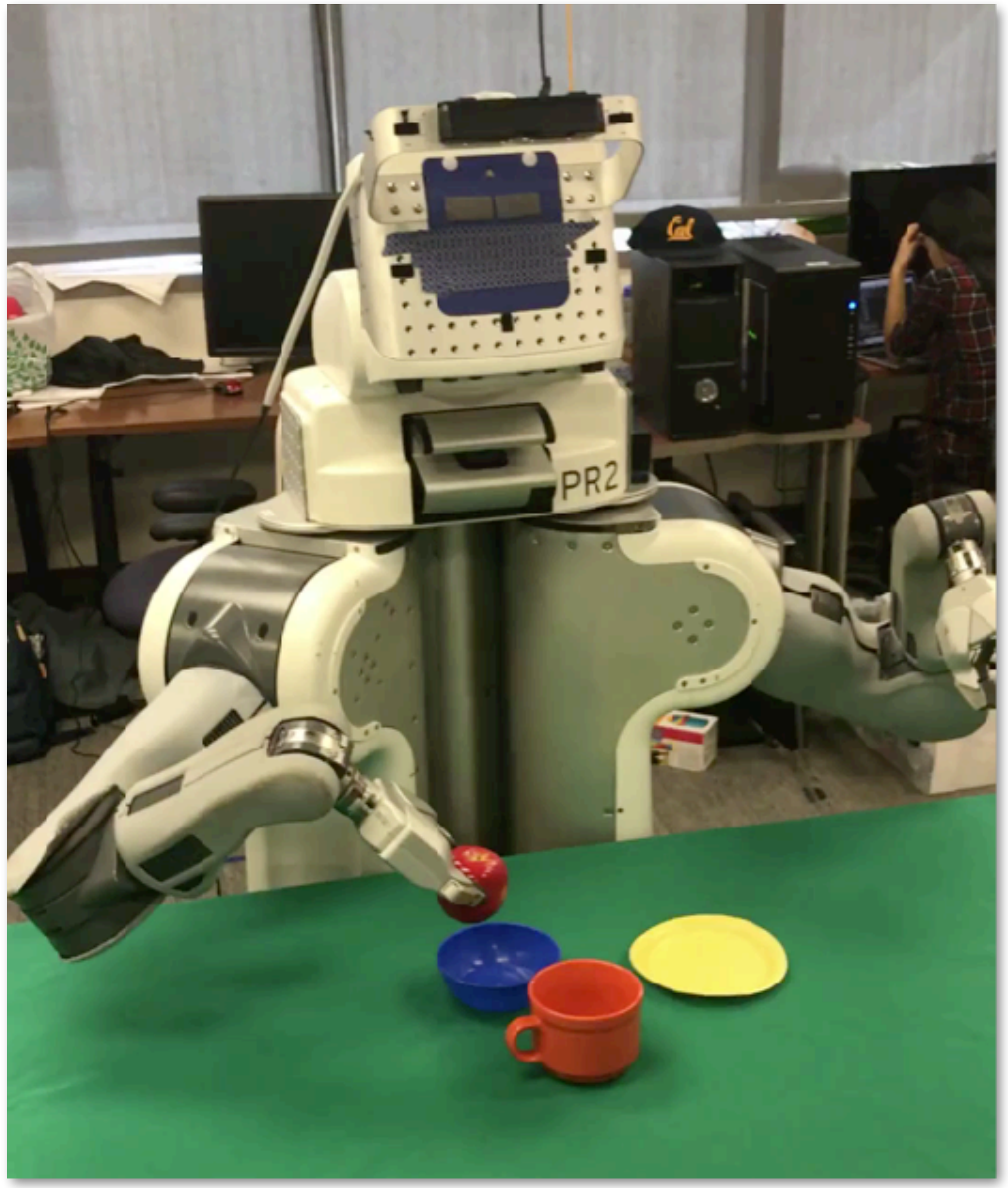
Leverage data with previous **objects** to quickly adapt to new ones?

Previous demo data
from *training objects*

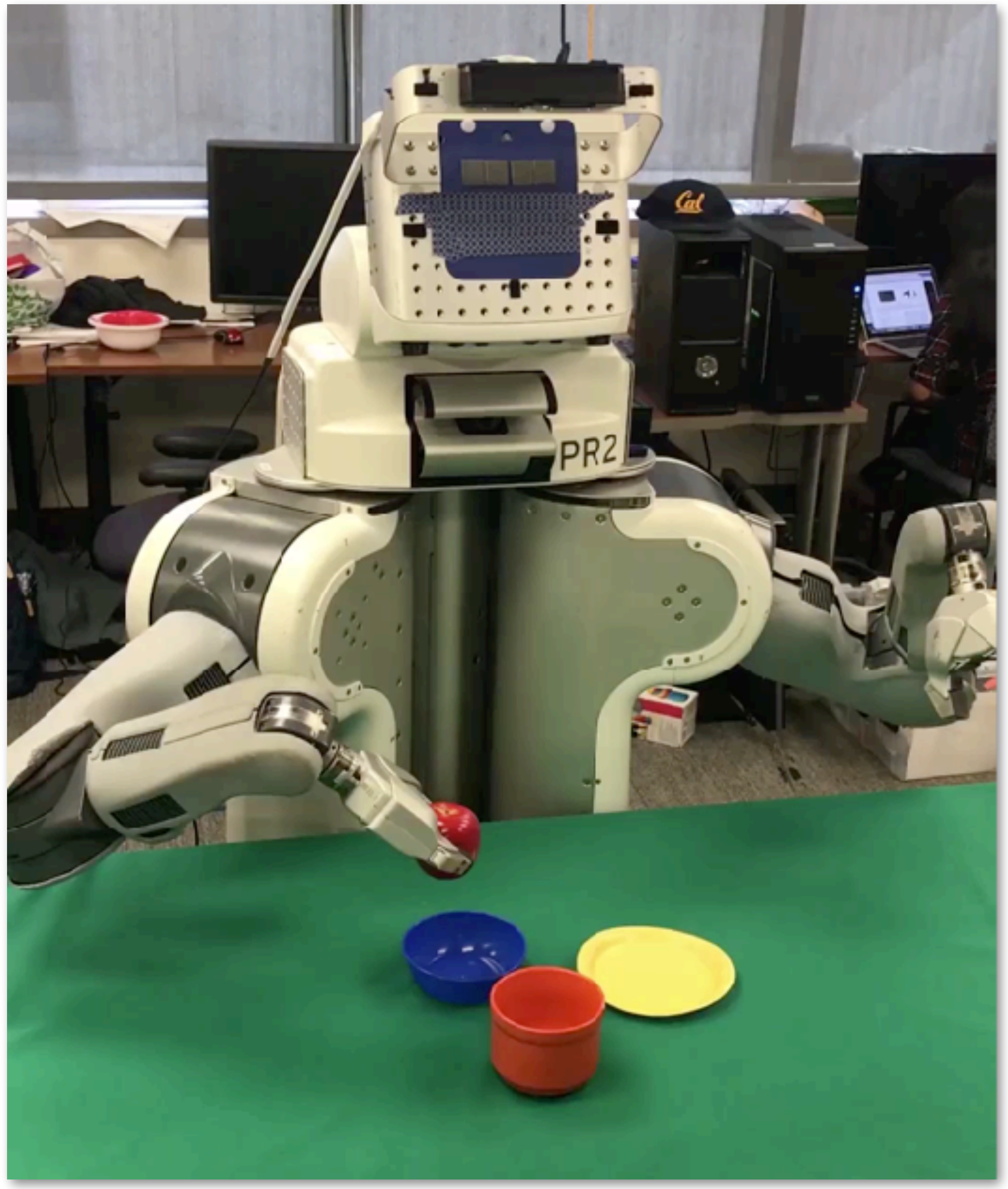
subset of
training objects

tasks: placing object
into target container

input demo
(via teleoperation)



resulting policy



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

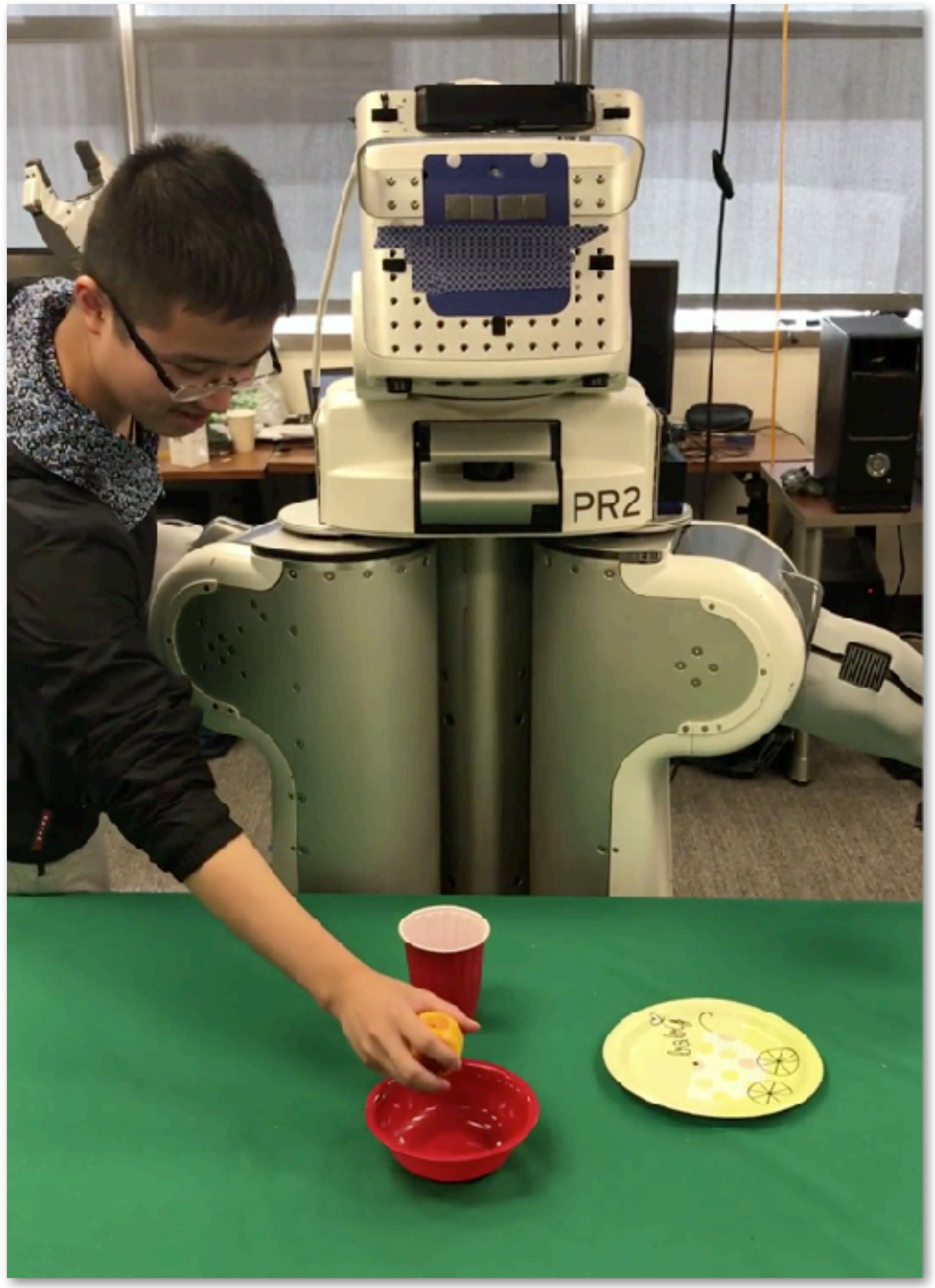
Leverage data with previous **objects** to quickly adapt to new ones?

Previous demo data
+ *human video data*
from training objects

subset of
training objects

tasks: placing object
into target container

input *human* demo

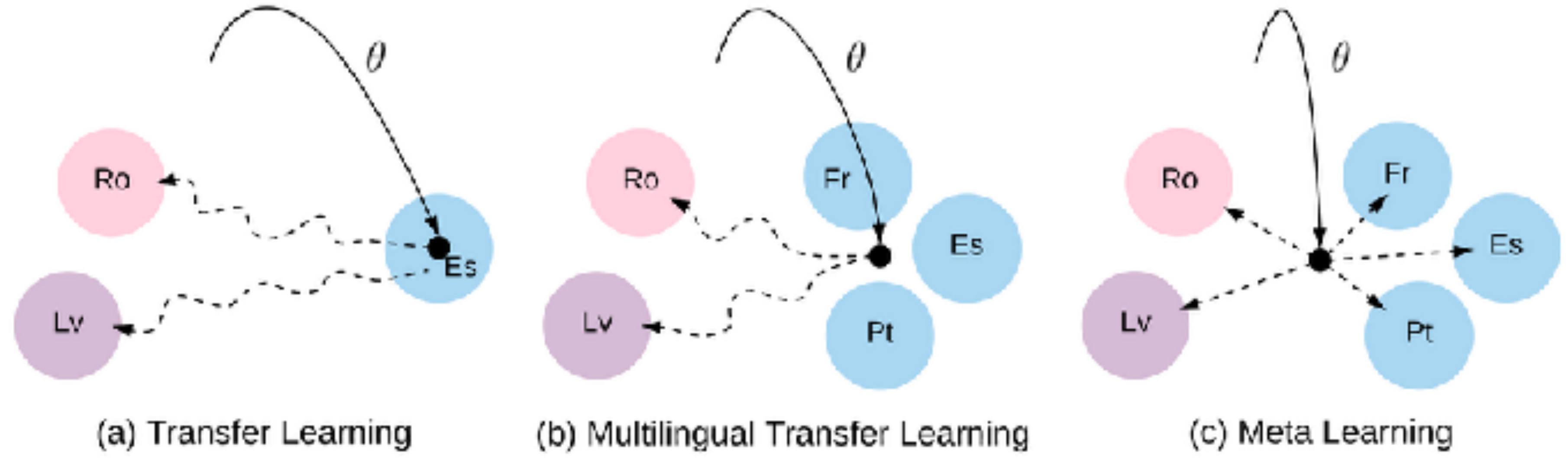


resulting policy

Yu*, Finn*, Xie, Dasari, Abbeel, Levine. One-Shot Imitation from Observing Humans via Domain-Adaptive Meta-Learning. RSS 2018

Leverage previous **language** experience

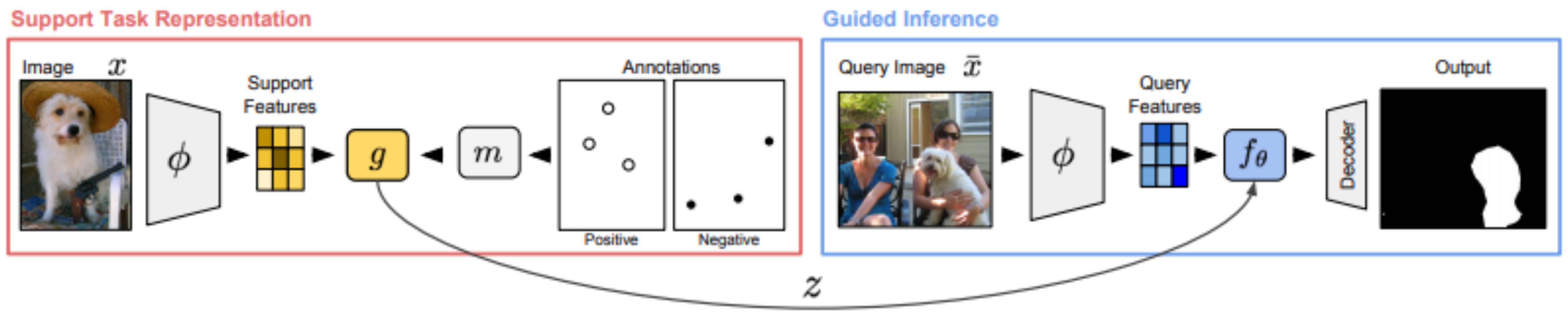
Low-resource neural machine translation



see, e.g.: Gu et al. EMNLP '18

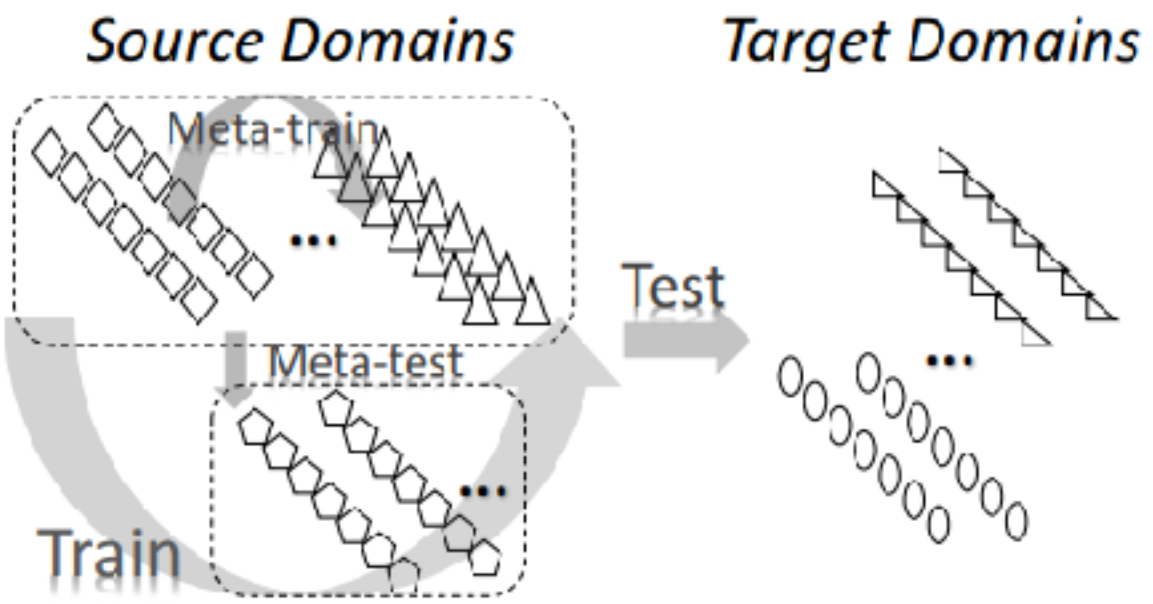
Leverage previous **segmentation** experience

Image segmentation from a few pixel labels



see, e.g.: Shaban, et al. One-Shot Learning for Semantic Segmentation.
Rakelly, et al. Few-Shot Segmentation Propagation with Guided Networks.
Dong & Xing. Few-Shot Semantic Segmentation with Prototype Learning.

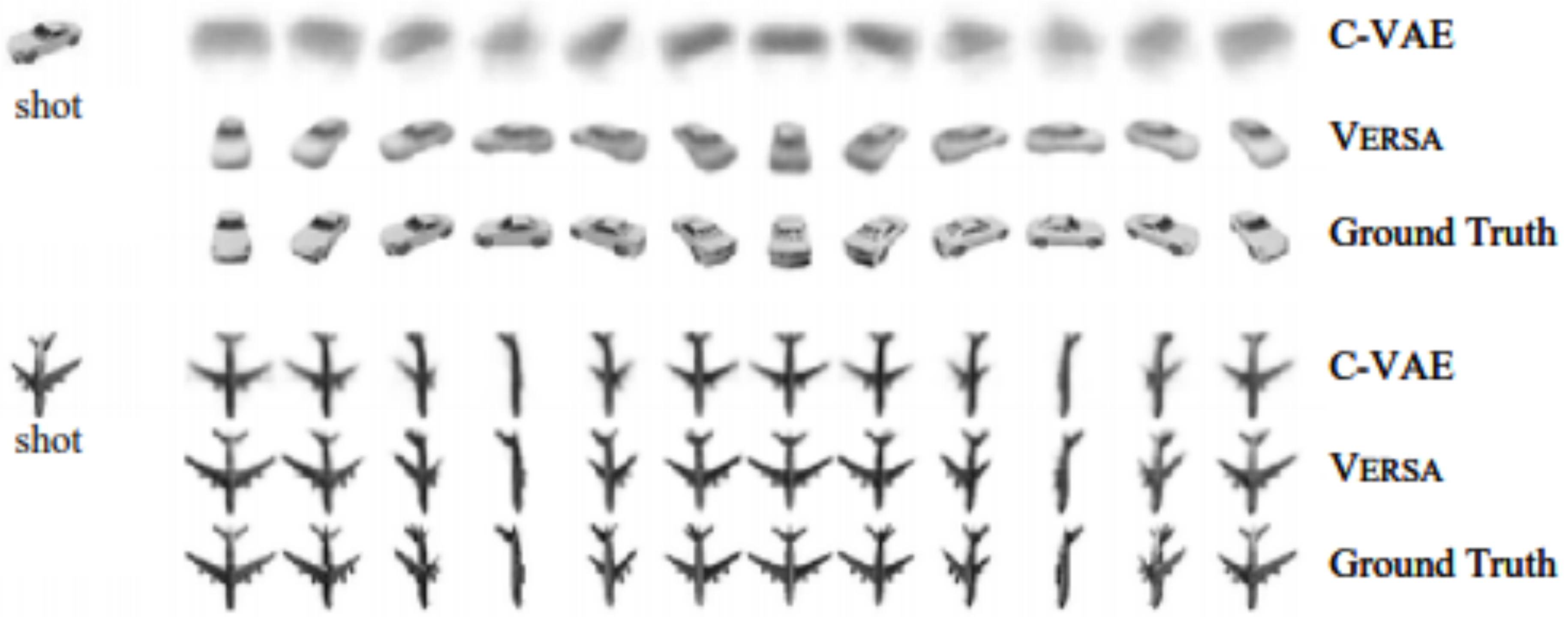
Leverage experience with previous **domains**



see, e.g.: Li et al. Learning to Generalize: Meta-Learning for Domain Adaptation.

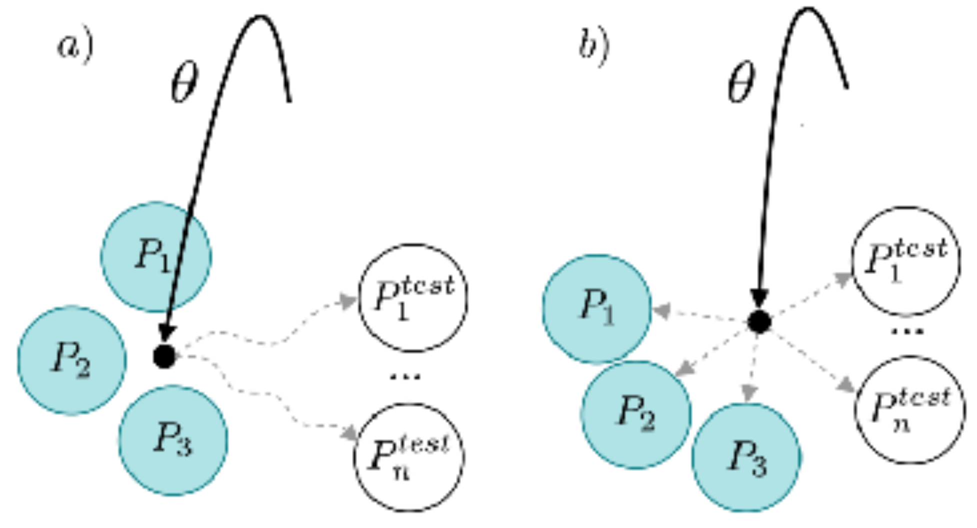
Leverage previous experience with **objects**

Few-shot image generation



see, e.g.: Gordon et al. VERSA: Versatile and Efficient Few-Shot Learning.

Leverage previous experience with **people**



Personalize dialog to a persona

see, e.g.: Lin*, Madotto* et al. ACL '19

And many many others…

How should we incorporate **prior experience** into ML systems?

The algorithms work pretty well.

But is the problem statement what we want?

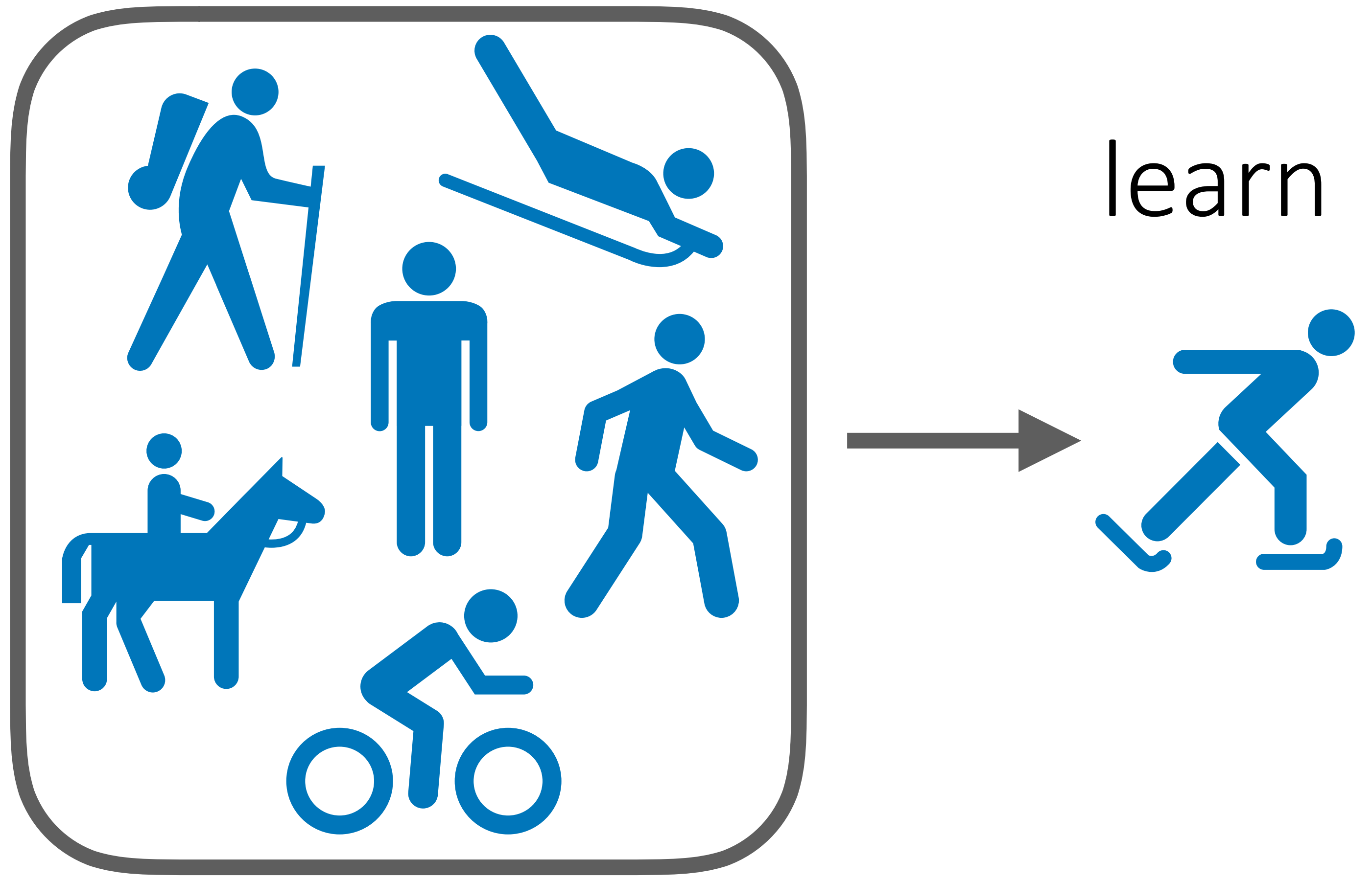
Prior experience doesn't typically come all at once.

What are the tasks and where do they come from?

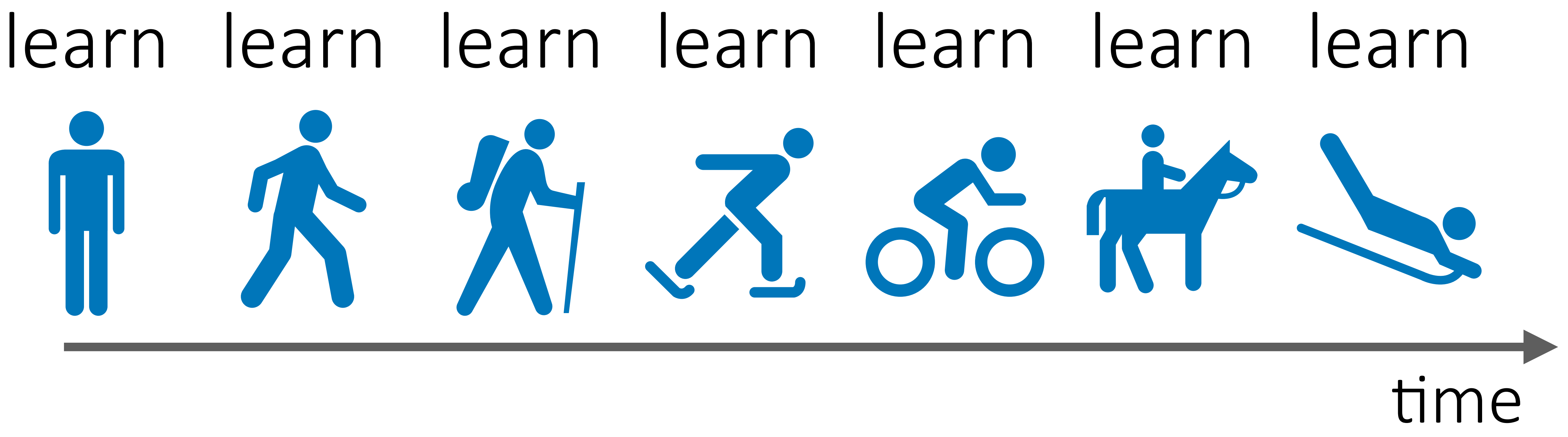
Meta-Learning

(Schmidhuber et al. ’87, Bengio et al. ’92)

Given i.i.d. task distribution,
learn a new task efficiently



More realistically:

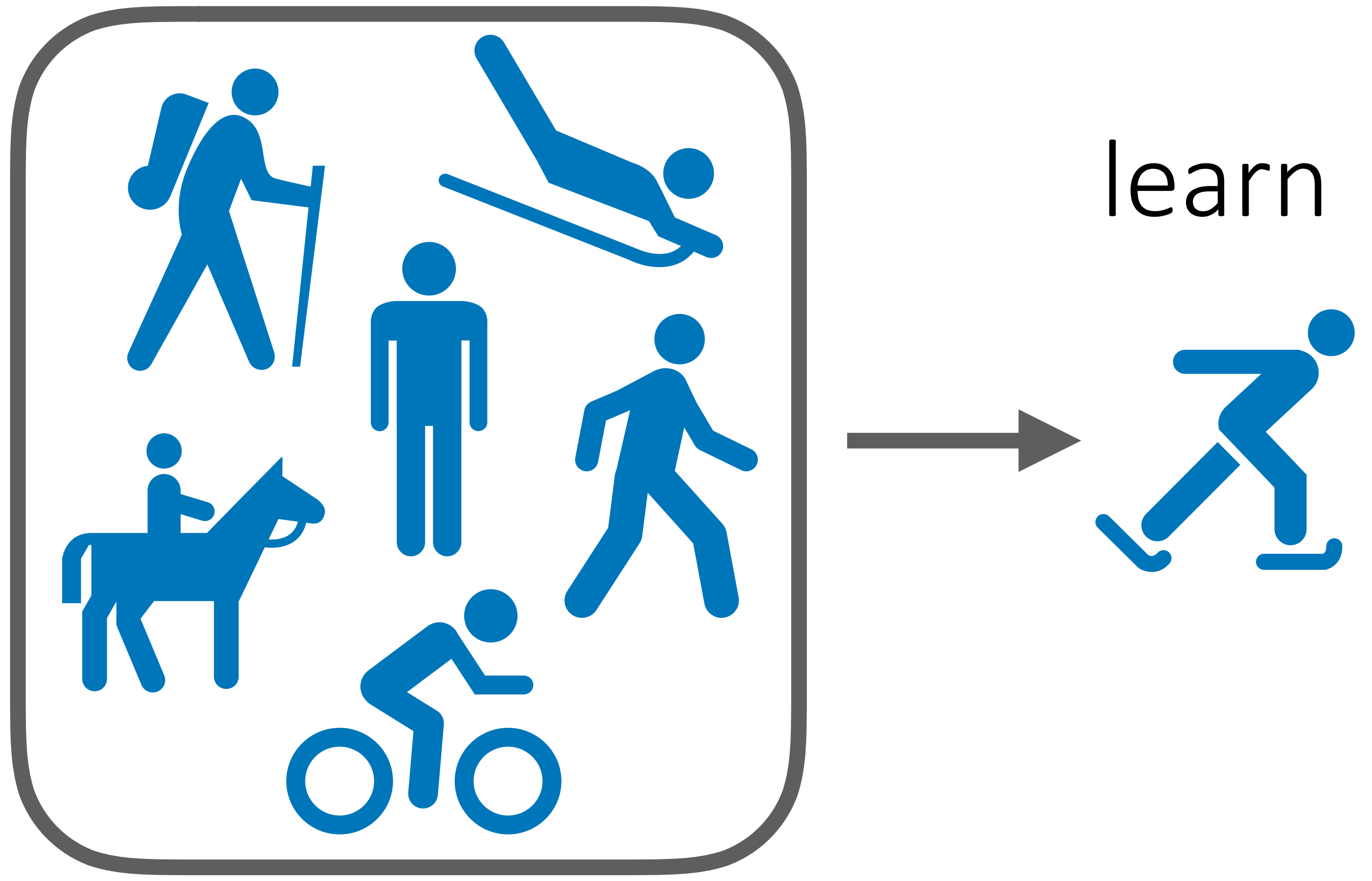


slow learning → rapid learning

Meta-Learning

(Schmidhuber et al. ’87, Bengio et al. ’92)

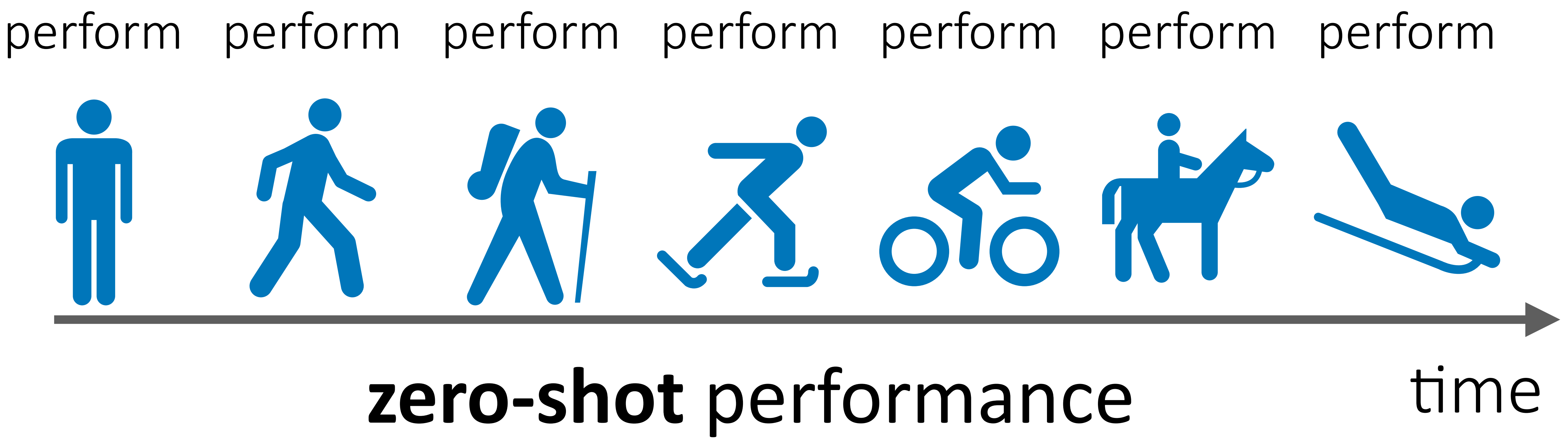
Given i.i.d. task distribution,
learn a new task efficiently



Online Learning

(Hannan ’57, Zinkevich ’03)

Perform sequence of tasks
while minimizing static regret.

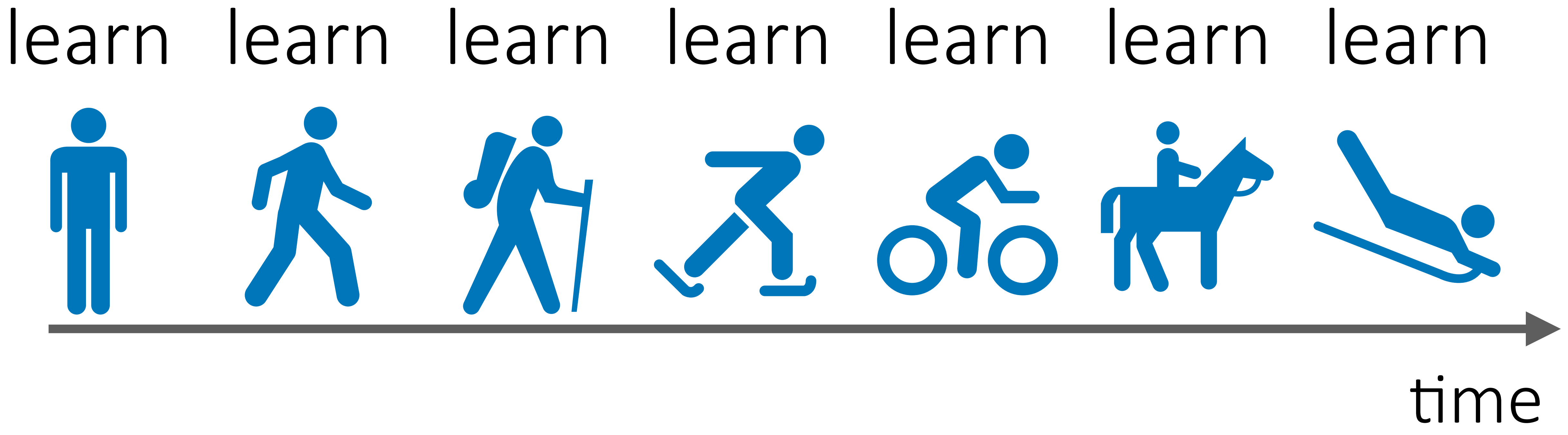


zero-shot performance

Online Meta-Learning

(this work)

Efficiently learn a sequence of tasks
from a non-stationary distribution.



performance after seeing a small amount of data

Finn*, Rajeswaran*, Kakade, Levine. Online Meta-Learning. ICML 2019

The Online Meta-Learning Problem Setting

For round $t \in \{1, 2, \ldots \infty\}$:

$$\tilde{\theta}_t = \theta_t - \alpha \nabla \hat{\ell}_t(\theta_t)$$

1. World picks a loss function $\ell_t(\cdot)$
2. Agent should pick θ_t without knowledge of ℓ_t
3. Agent uses update procedure $\Phi_t : \Theta \to \Theta$, and obtains $\tilde{\theta}_t = \Phi_t(\theta_t)$
4. Agent suffers $\ell_t(\tilde{\theta}_t)$ for the round

Goal: Learning algorithm with sub-linear

$$\mathrm{Regret}_T := \underbrace{\sum_{t=1}^{T} \ell_t(\Phi_t(\theta_t))}_{\text{Loss of algorithm}} - \underbrace{\min_{\theta \in \Theta} \sum_{t=1}^{T} \ell_t(\Phi_t(\theta))}_{\text{Loss of best algorithm in hindsight}}$$

Finn*, Rajeswaran*, Kakade, Levine. Online Meta-Learning. ICML 2019

Follow the Meta-Leader (FTML) : $$\theta_{t+1} = \arg\min_{\theta} \sum_{t=1}^{T} \ell_t(\Phi_t(\theta))$$

Can be implemented with MAML

Theorem (Informal): If $\{\ell_t(\cdot), \hat{\ell}_t(\cdot)\}$ $\forall t$ are C^2-smooth and strongly convex, the sequence of models $\{\theta_1, \theta_2, \ldots, \theta_T\}$ returned by FTML has the property:

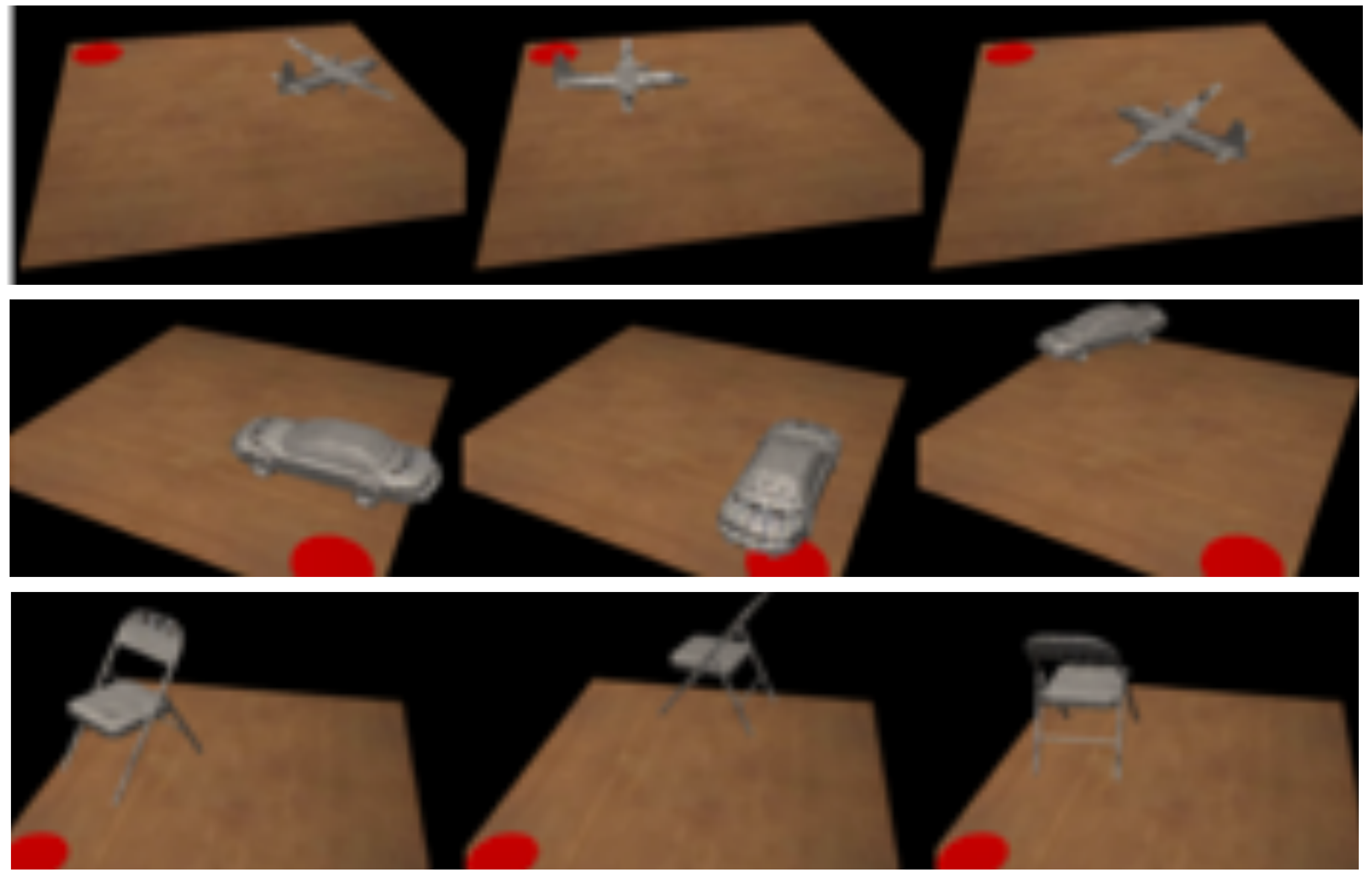
$$\text{Regret}_T := \sum_{t=1}^{T} \ell_t(\Phi_t(\theta_t)) - \min_{\theta \in \Theta} \sum_{t=1}^{T} \ell_t(\Phi_t(\theta)) = O(\log T)$$

Finn*, Rajeswaran*, Kakade, Levine. Online Meta-Learning. ICML 2019

Practical instantiation of FTML: meta-train with MAML on all data so far, fine-tune on current task

Experiment with **sequences of tasks**:
- Colored, rotated, scaled **MNIST**
- **3D object pose prediction**
- **CIFAR-100** classification

Example pose prediction tasks



plane

car

chair

Compare to:

- **TOE** (train on everything): train on all data so far
- **FTL** (follow the leader): train on all data so far, fine-tune on current task
- **From Scratch:** train from scratch on each task

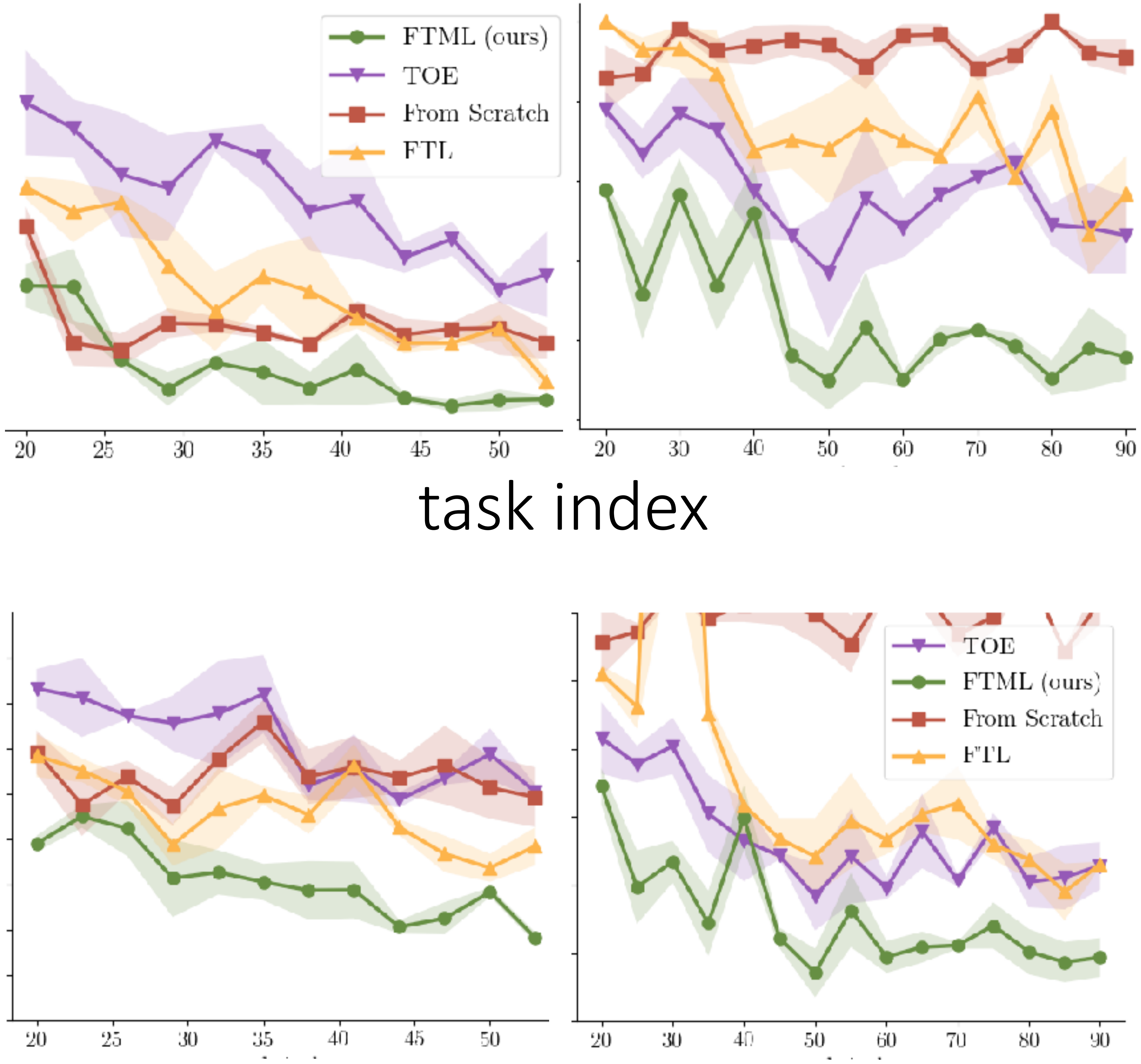
Finn*, Rajeswaran*, Kakade, Levine. Online Meta-Learning. ICML 2019

Experiments

Rainbow MNIST   Pose Prediction

Learning efficiency (# datapoints)

Learning proficiency (error)



task index

TOE (train on everything): improves over time, but **prone to negative transfer**

FTL (follow the leader): **consistent forward transfer**, sometimes **overfits**

FTML (ours): **learns each new task faster** & **with greater proficiency**, approaches **few-shot learning** regime

Finn*, Rajeswaran*, Kakade, Levine. Online Meta-Learning. ICML 2019

How should we incorporate **prior experience** into ML systems?

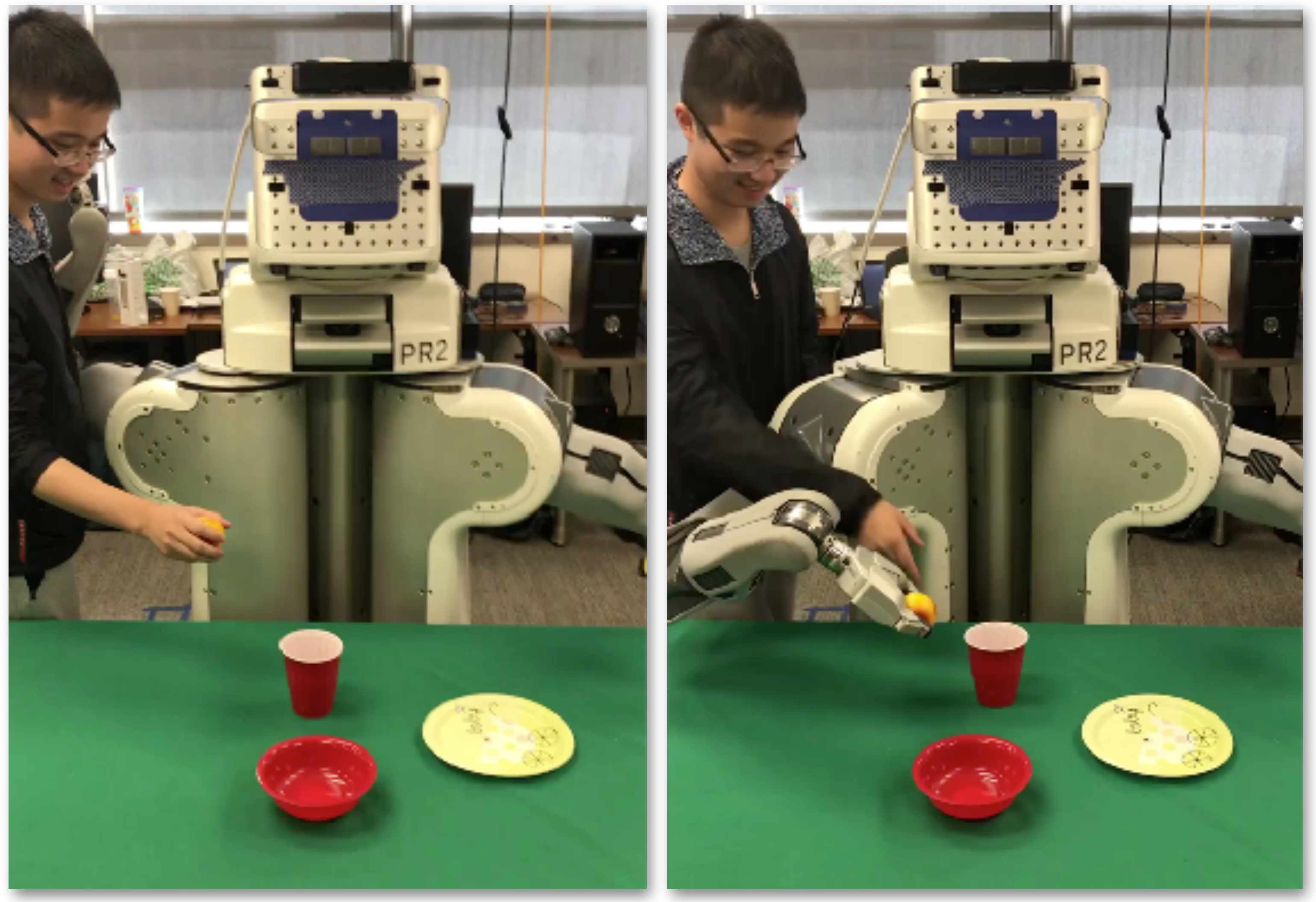
The algorithms work pretty well.

But is the problem statement what we want?

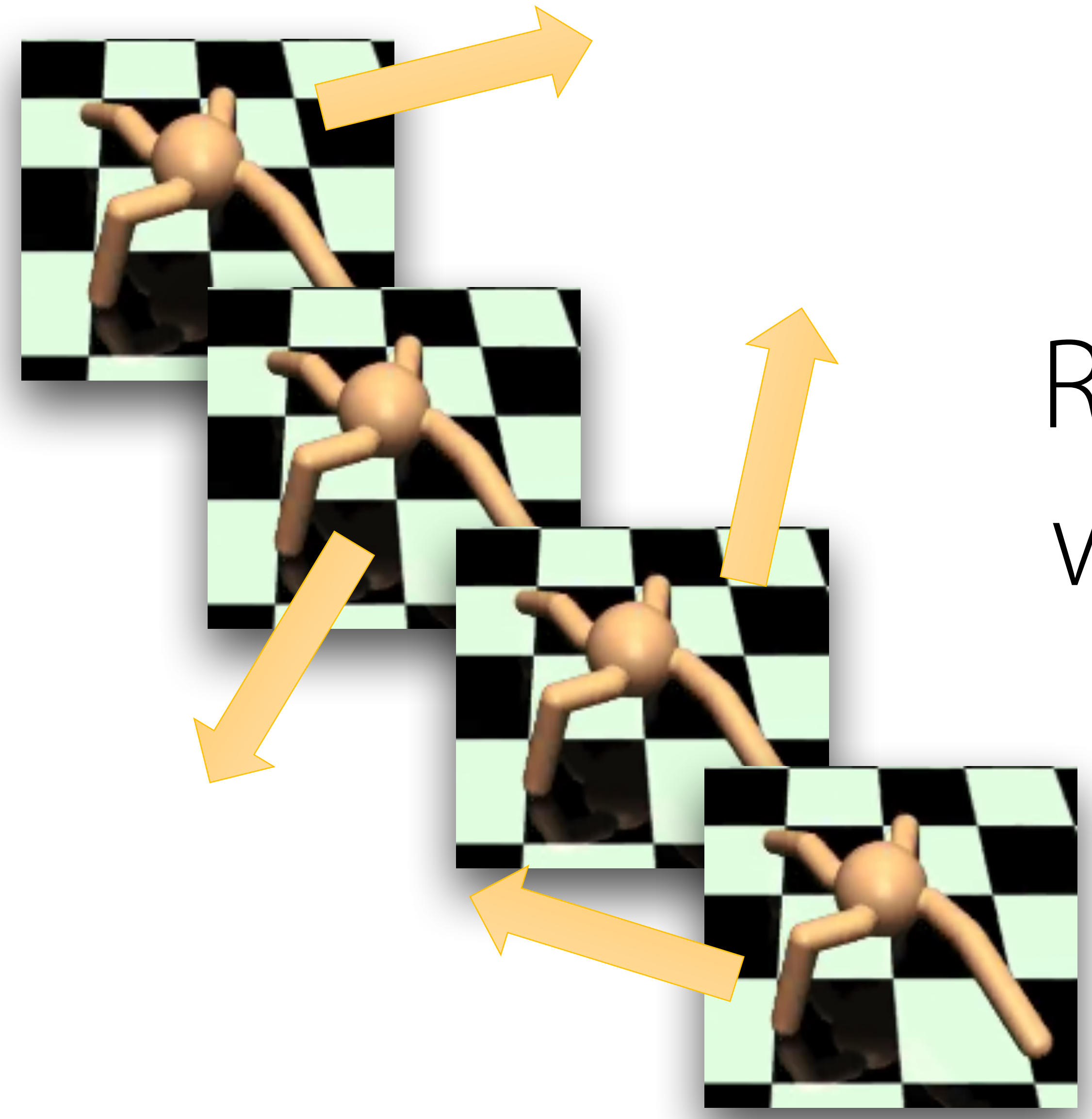
Prior experience doesn't typically come all at once.

What are the tasks and where do they come from?

Requires tasks constructed from labeled data



Requires demos for many previous tasks

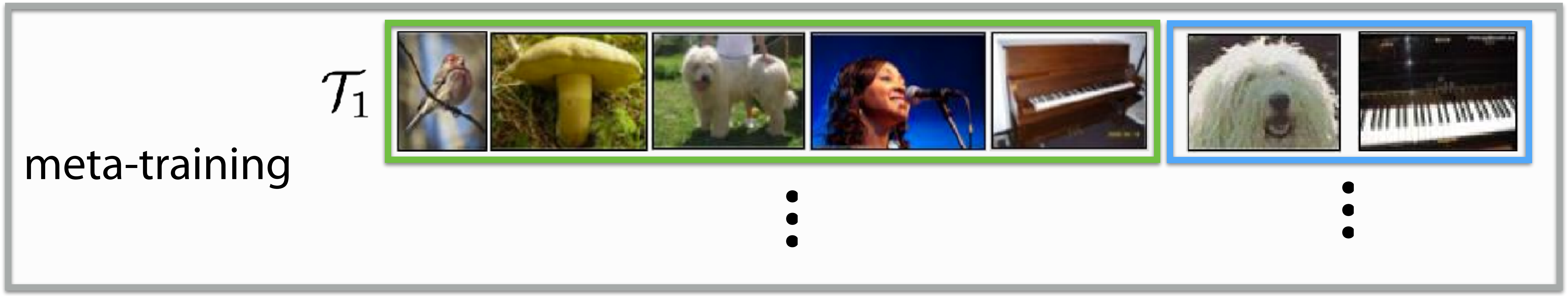


Requires many tasks with corresponding reward functions

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

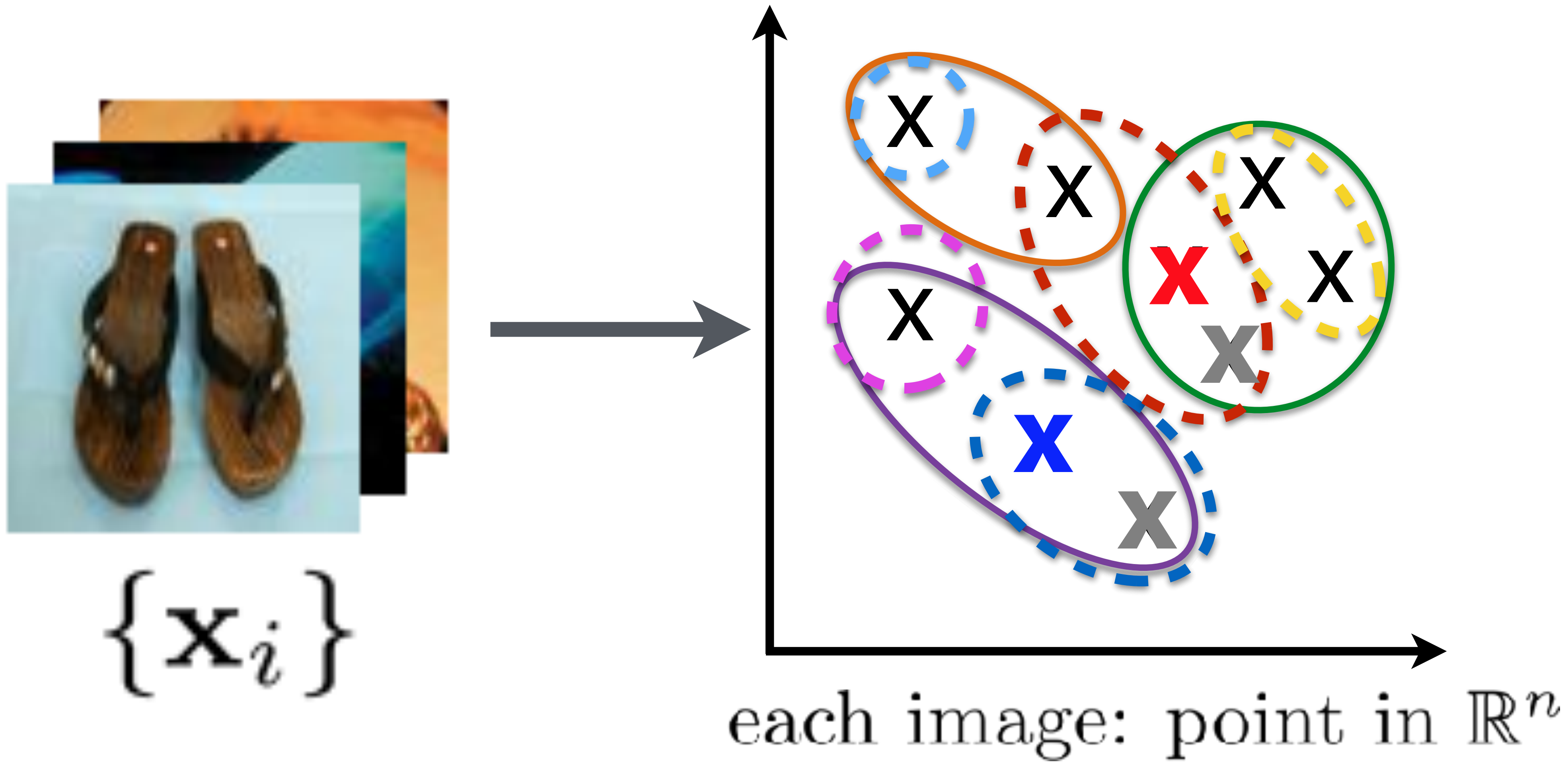
Can we also automate the task design process?

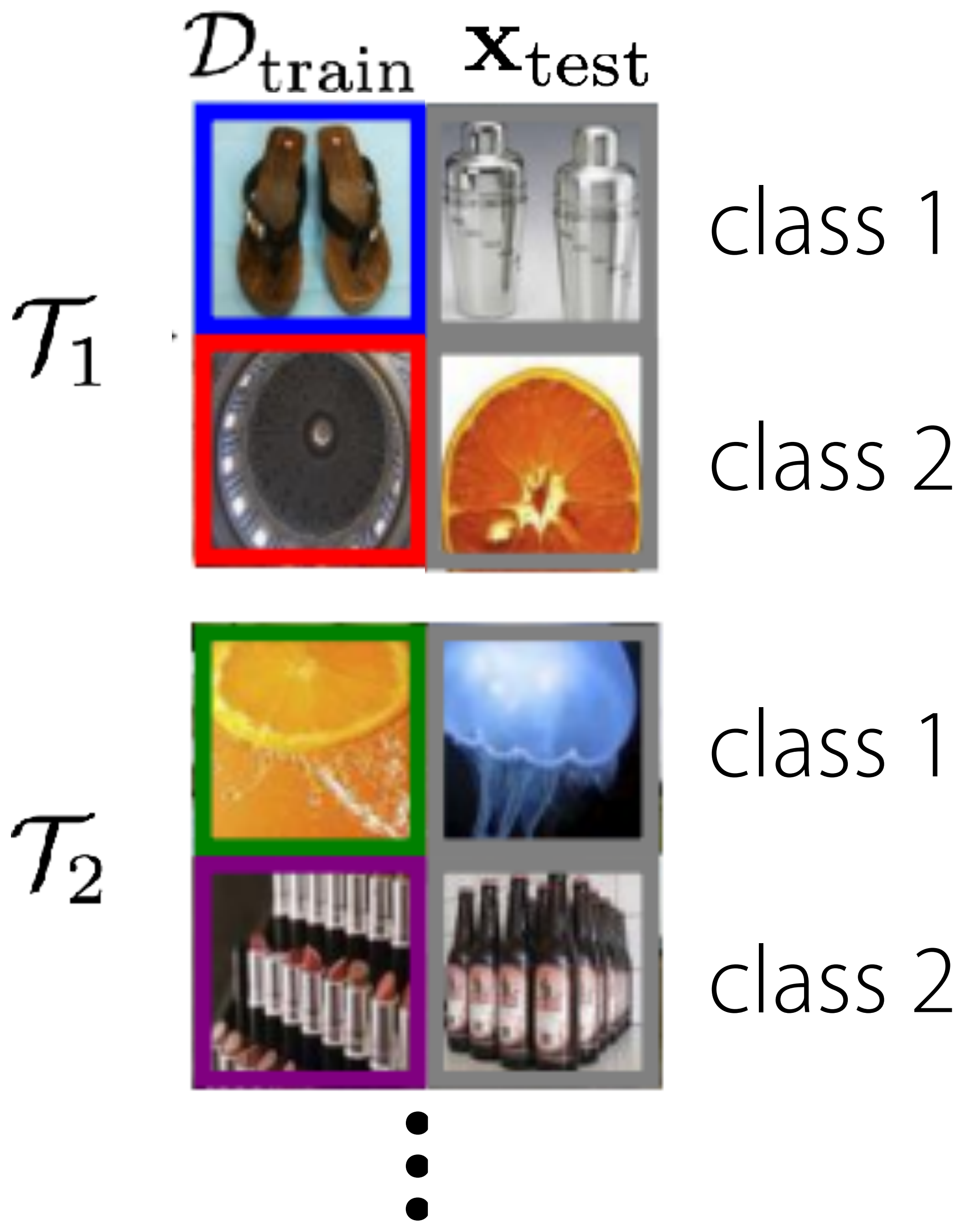
Propose tasks for meta-learning with only **unlabeled** images?

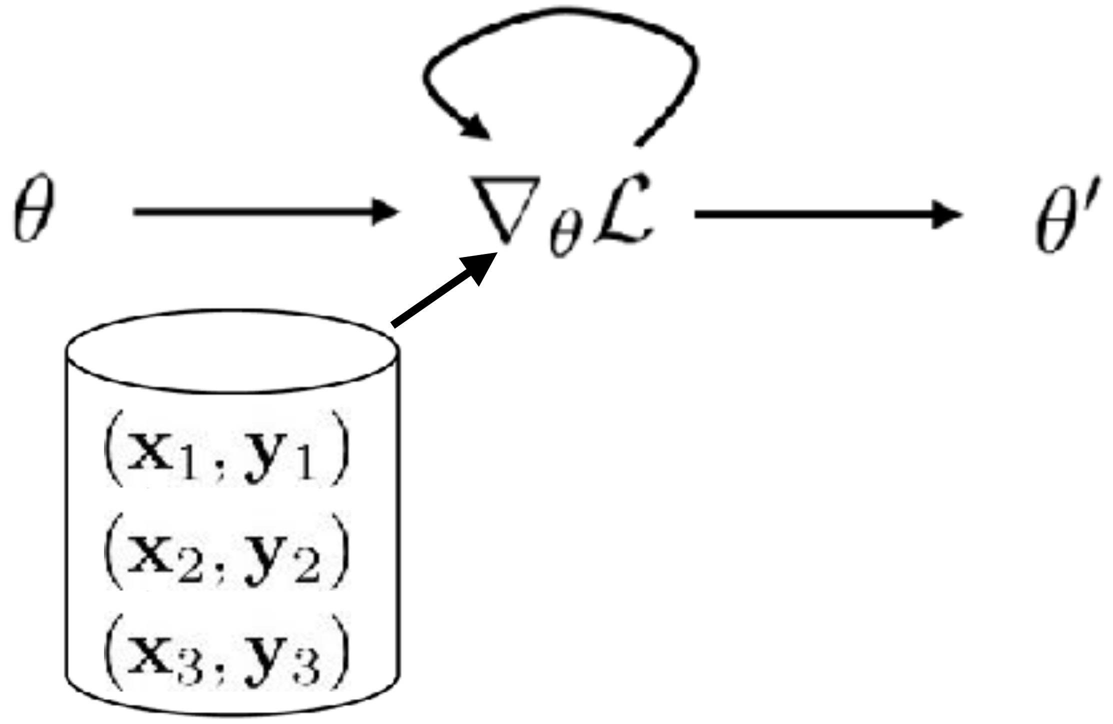


Construct tasks without labeled data?

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







Result: representation suitable for learning downstream tasks

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

Propose tasks for meta-learning 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

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

What about unsupervised meta-RL?

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

- Propose tasks using skill discovery methods (e.g. DIAYN)

latent skill: z
(discrete latent variable)

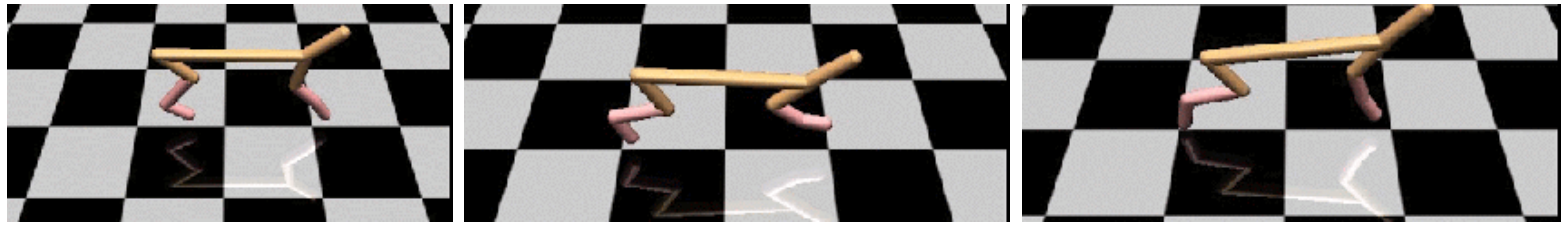
policy: $\pi(a|s,z)$

discriminator: $D(z|s)$

Goal: Maximize *mutual information* between s, z

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

Examples of acquired tasks:



Task reward for meta-learning:

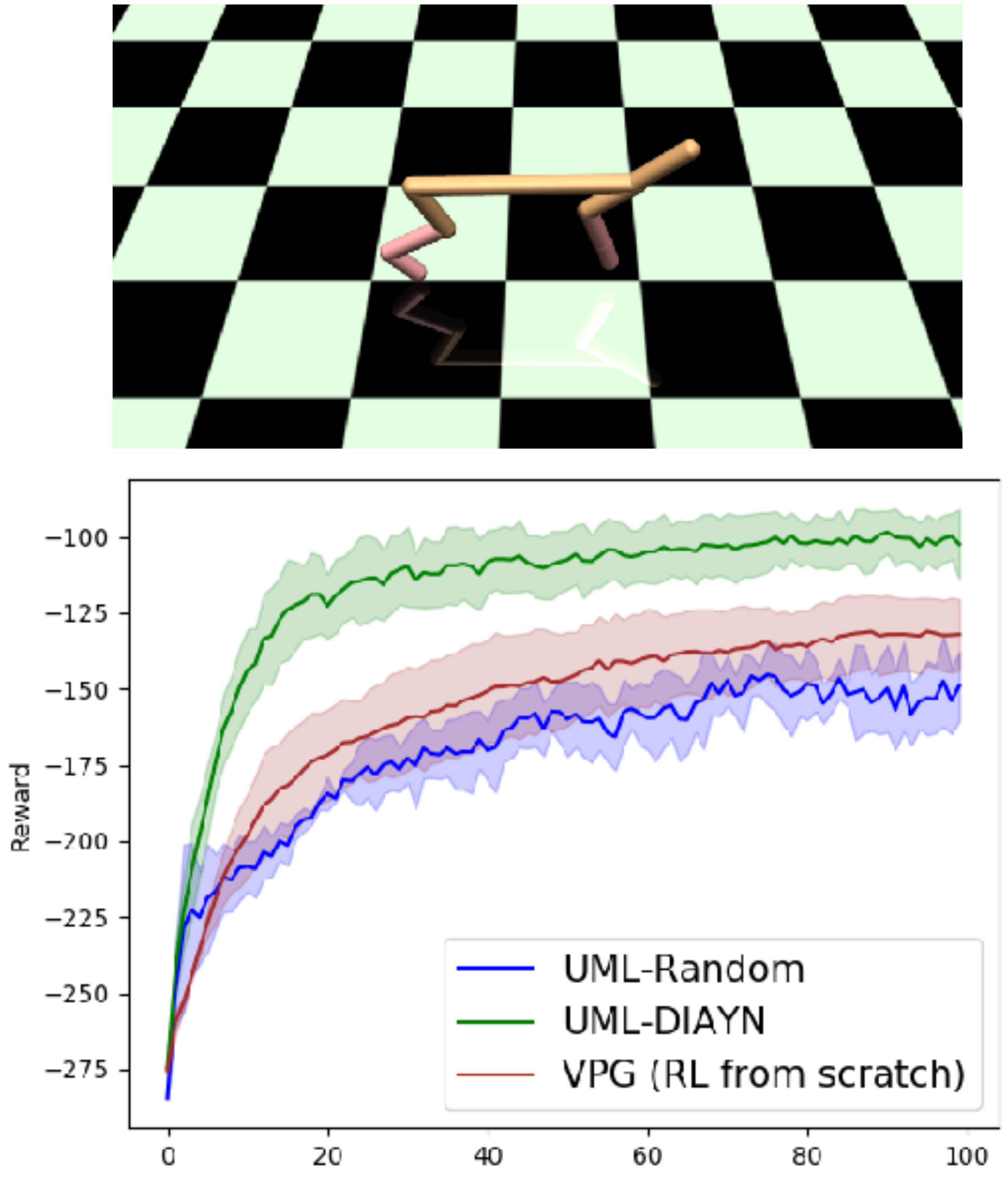
$$r(s,z) = \log D(z|s)$$

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

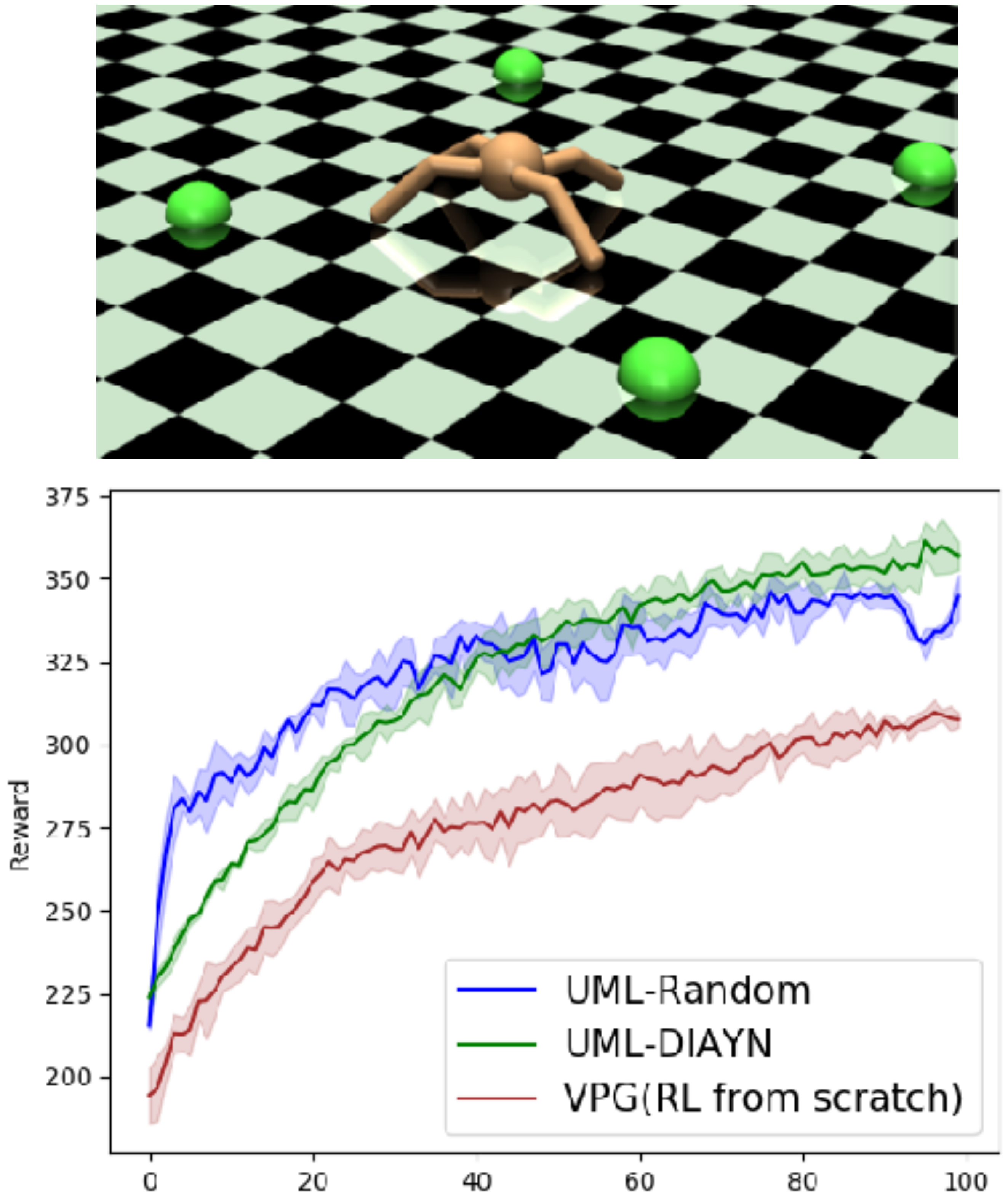
Does it work?

Measure learning performance on **test tasks *with* rewards**

Cheetah



Ant



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

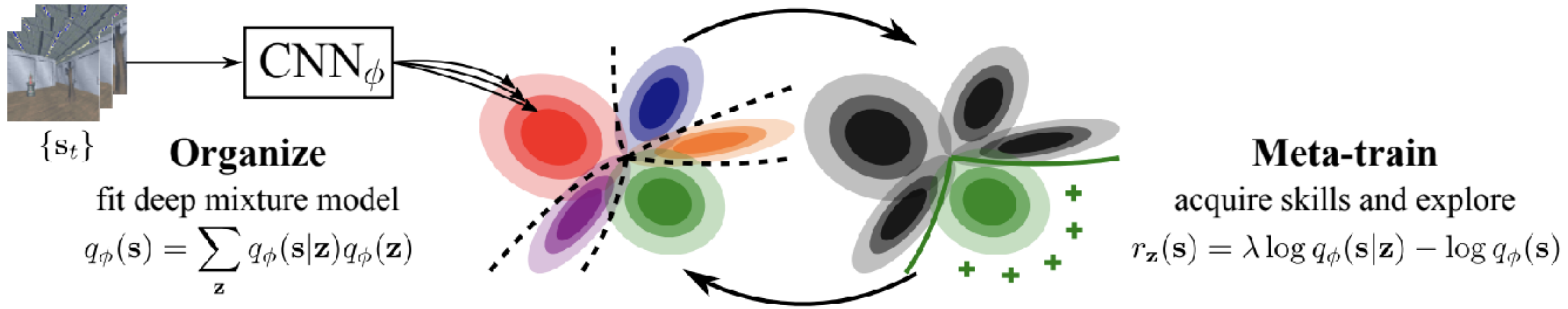
Gupta, Eysenbach, Finn, Levine. Unsupervised Meta-Learning for Reinforcement Learning. arXiv '18

What about unsupervised meta-RL?

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

Can we **adapt the task distribution** based on the meta-learner's current behavior?

Formulate task acquisition as an **information maximization problem**, optimized with EM



Learn latent representation **s** through deep trajectory-centric clustering.

Fit **generative mixture model** over **s**.

Meta-train w.r.t. **mutual information objective** under density model.

Jabri, Hsu, Eysenbach, Gupta, Efros, Levine, Finn. Unsupervised Curricula for Visual Meta-RL. '19

Natural to incorporate **density-based exploration**.

Meta-train

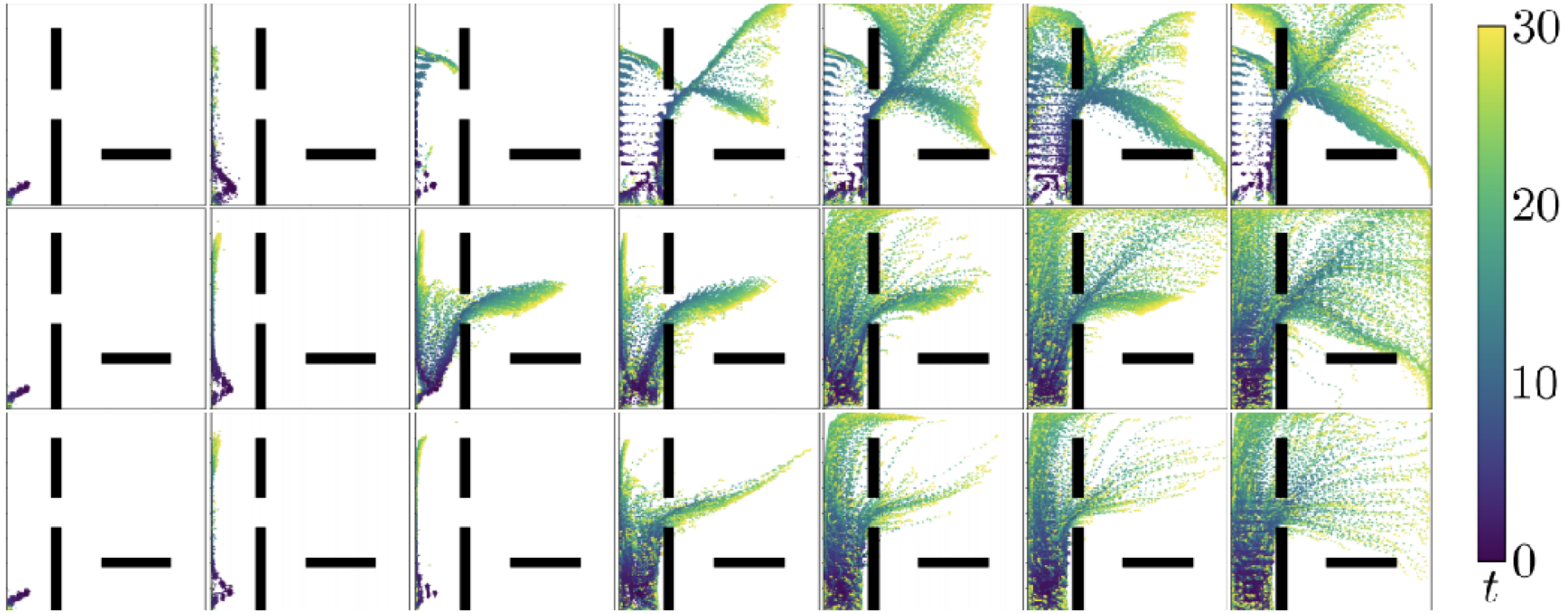
acquire skills and explore

$$r_{\mathbf{z}}(\mathbf{s}) = \lambda \log q_\phi(\mathbf{s}|\mathbf{z}) - \log q_\phi(\mathbf{s})$$

$\lambda = 0.2$

$\lambda = 0.5$

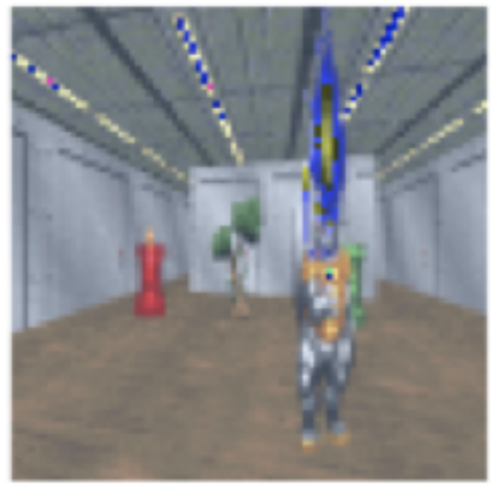
$\lambda = 0.8$



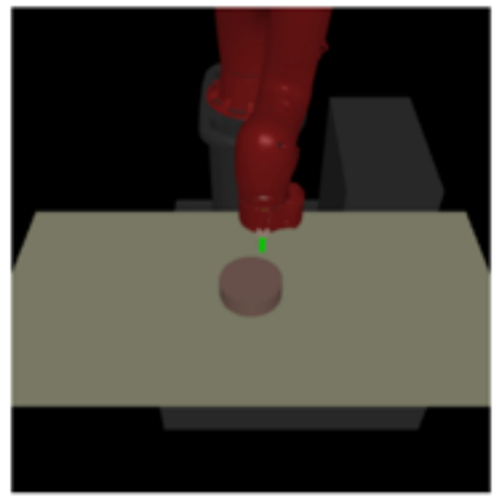
Scales naturally to **visual observations**.

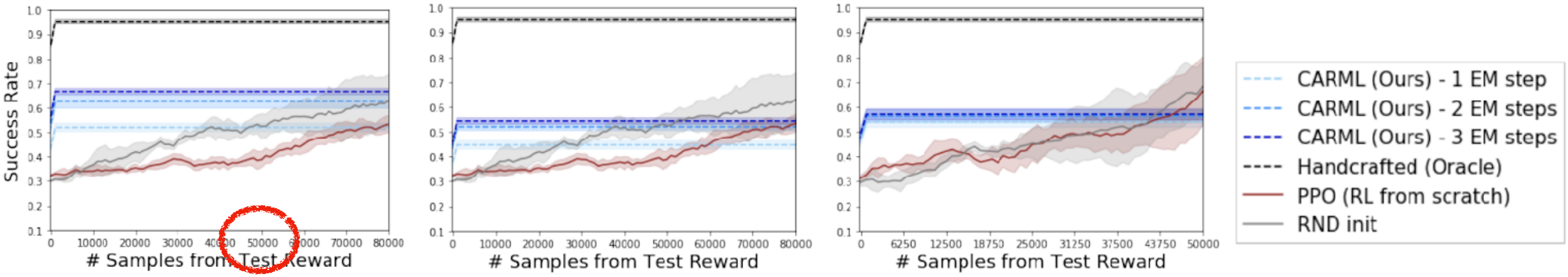
Directly transfers to **downstream** tasks.

VizDoom



Sawyer





(a) ViZDoom (fixed) (b) ViZDoom (random) (c) Sawyer

+ improves with further EM steps

Jabri, Hsu, Eysenbach, Gupta, Efros, Levine, Finn. Unsupervised Curricula for Visual Meta-RL. '19

How should we incorporate **prior experience** into ML systems?

Meta-learning provides a way to optimize for
priors that lead to few-shot learning

Prior experience doesn't
typically come all at once.

> **online meta-learning** setting

What are the tasks and
where do they come from?

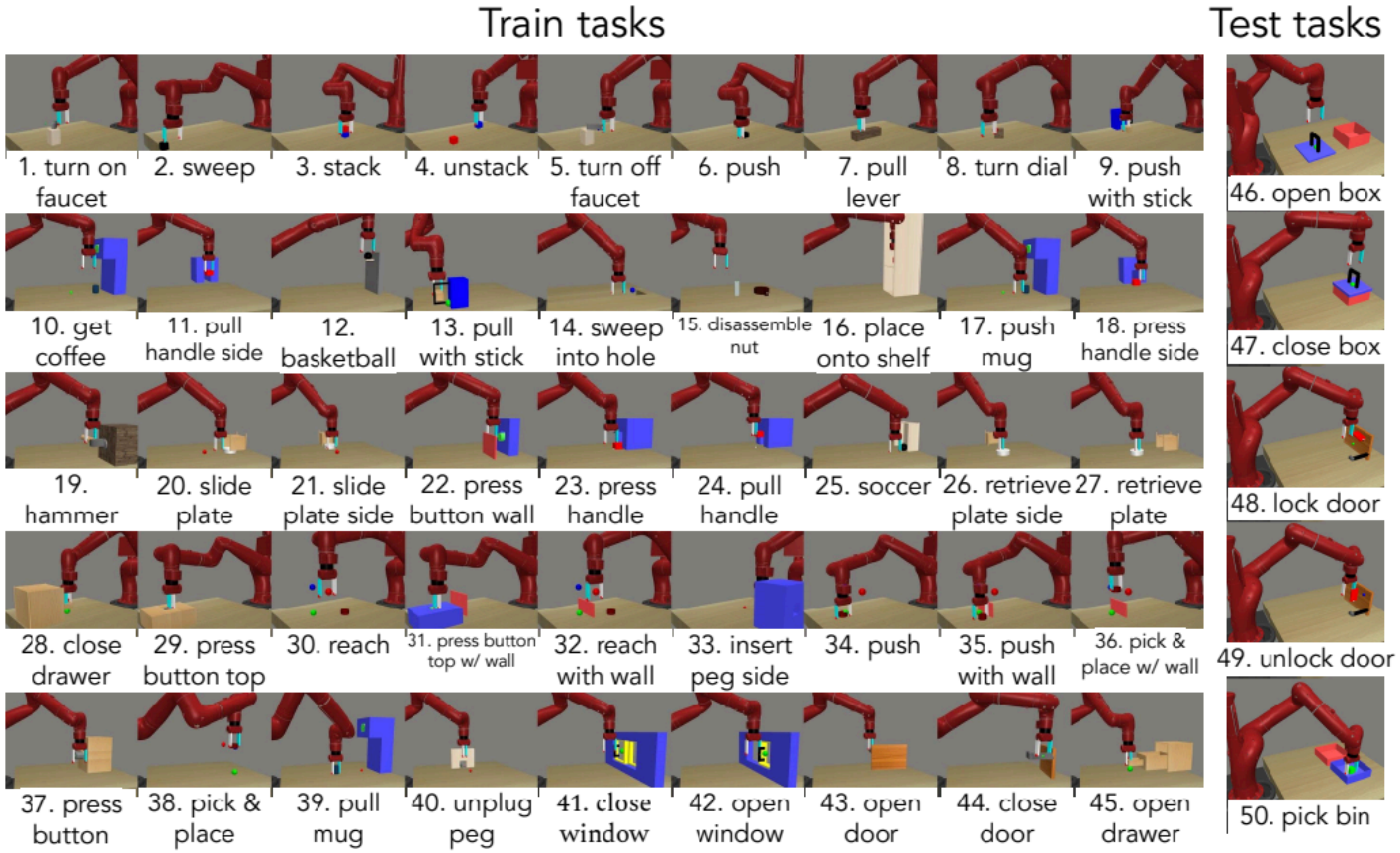
> can **propose tasks** from
unlabeled experience

Both of these bring us closer towards **realistic lifelong learning** scenarios.

Coming soon

Can we perform meta-RL across distinct task families?

Meta-World benchmark

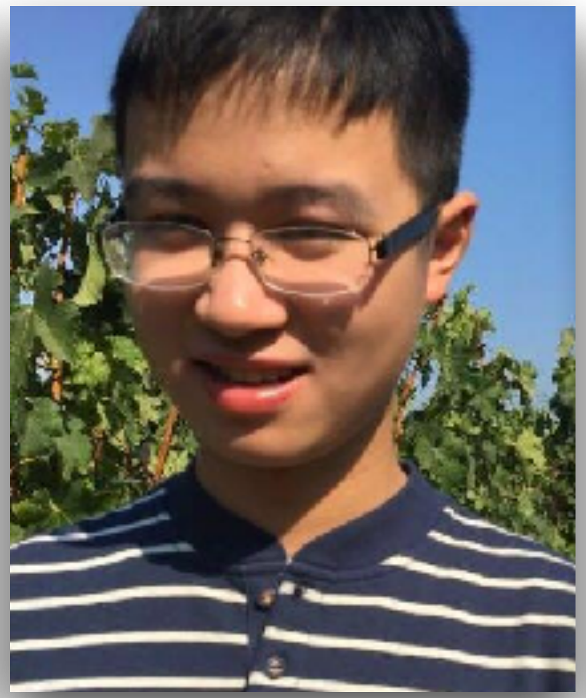


T Yu, D Quillen, Z He, R Julian, K Hausman, S Levine, C Finn

- 50 distinct control tasks
- shared workspace, action-space
- designed for studying **multi-task transfer**, **meta-learning**

Collaborators & Students

Sergey Levine


Tianhe Yu


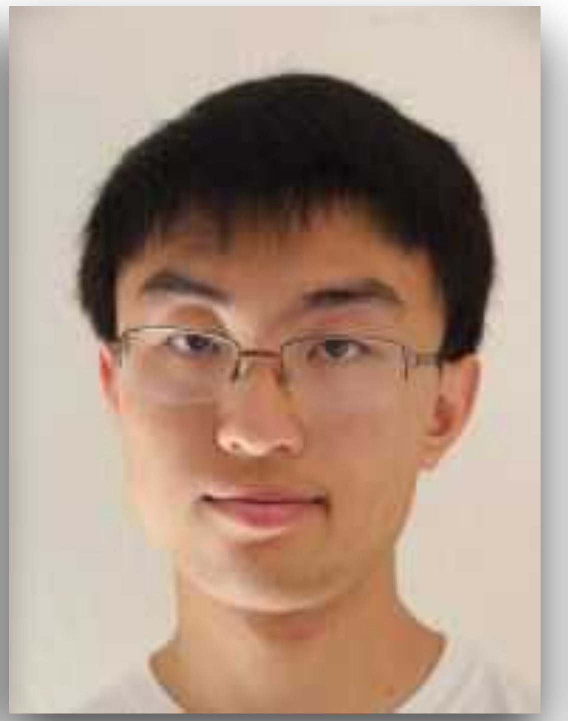
Anusha Nagabandi


Kyle Hsu


Ignasi Clavera


Pieter Abbeel


Kate Rakelly


Aurick Zhou


Russell Mendonca


Annie Xie


Sudeep Dasari


Simin Liu


Avi Singh


Abhishek Gupta


Ben Eysenbach


Allan Jabri


Questions?