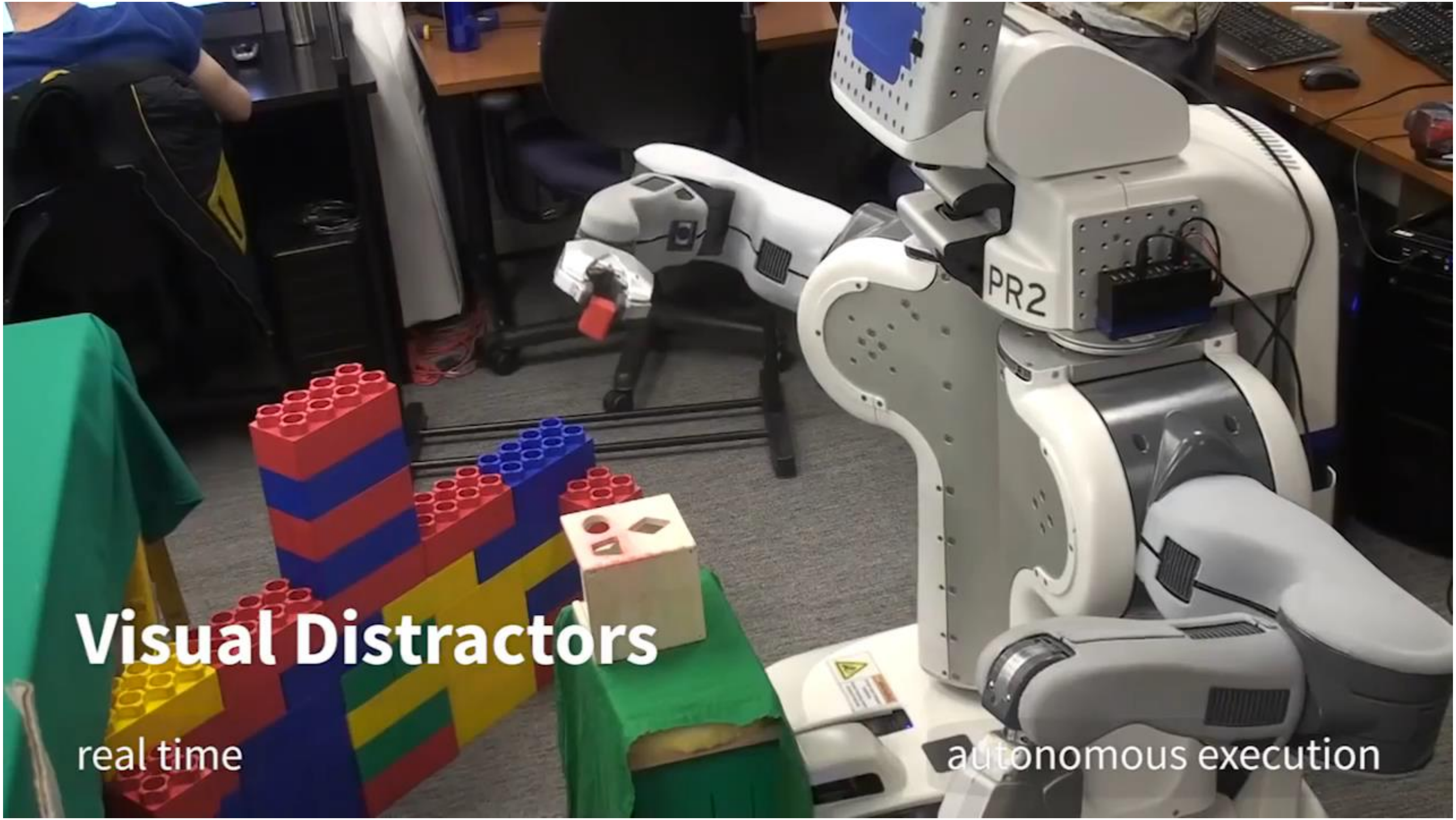
What’s Wrong with Meta-Learning
(and how we might fix it)

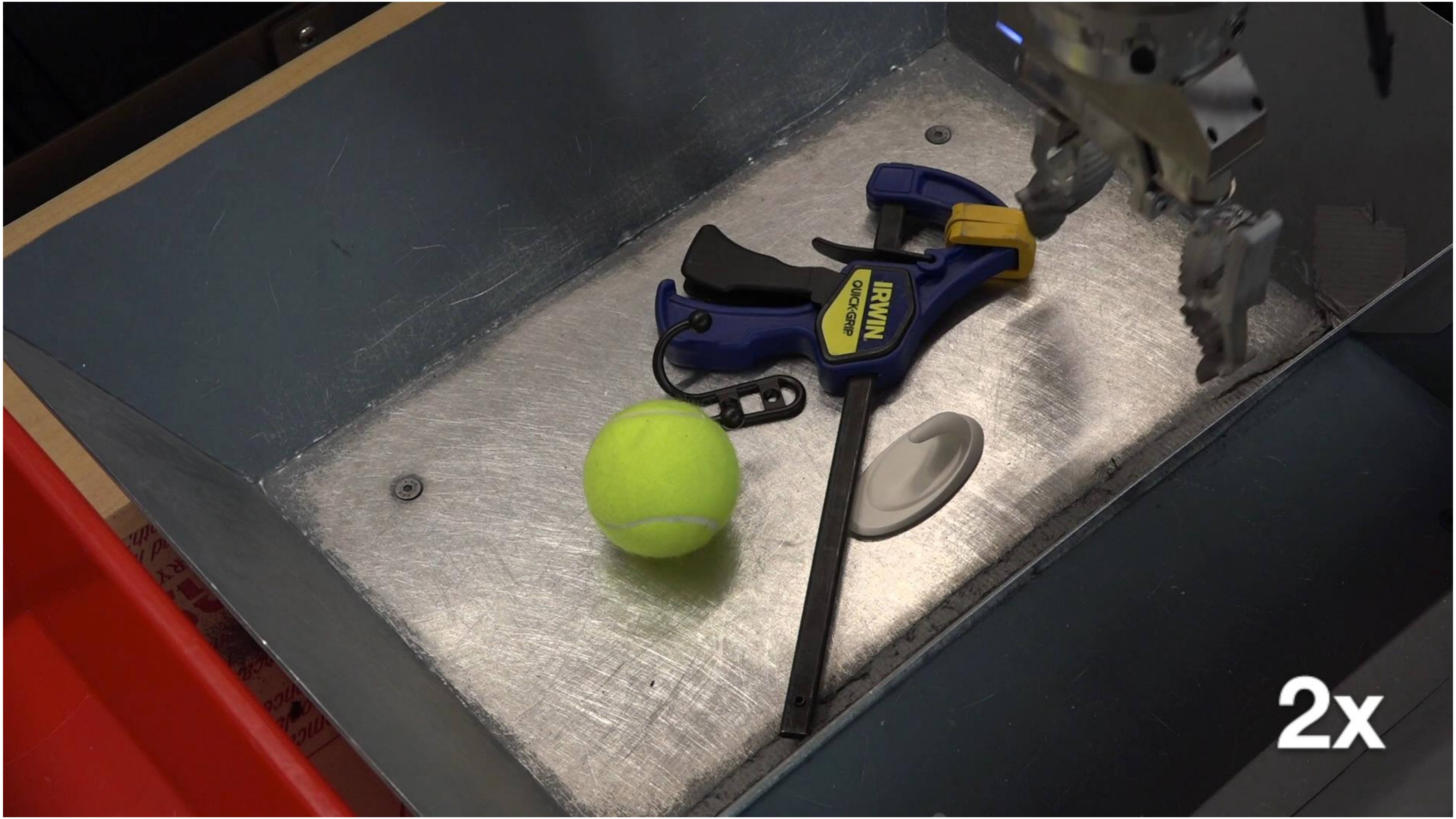
Sergey Levine

UC Berkeley **Google Brain**


PR2
Visual Distractors
real time
autonomous execution

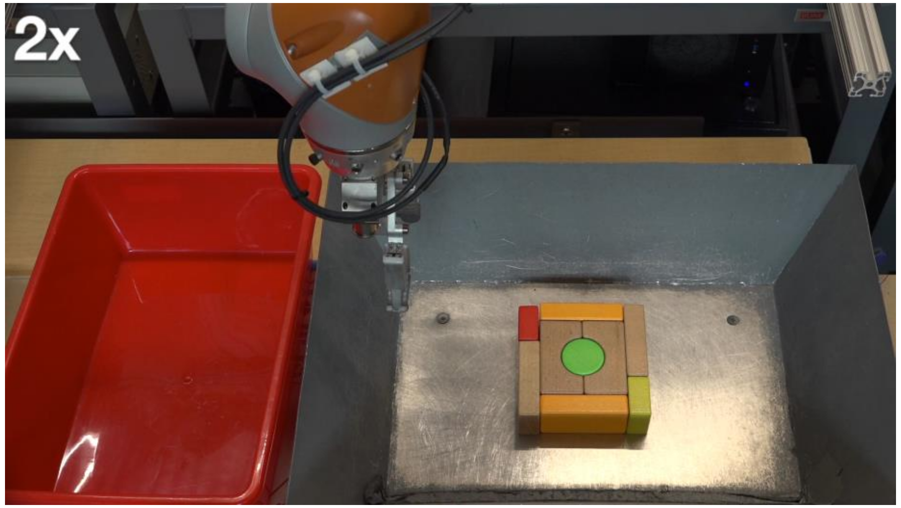
Yahya, Li, Kalakrishnan, Chebotar, Levine, ‘16

Kalashnikov, Irpan, Pastor, Ibarz, Herzong, Jang, Quillen, Holly, Kalakrishnan, Vanhoucke, Levine. **QT-Opt: Scalable Deep Reinforcement Learning of Vision-Based Robotic Manipulation Skills**



Kalashnikov, Irpan, Pastor, Ibarz, Herzong, Jang, Quillen, Holly, Kalakrishnan, Vanhoucke, Levine. **QT-Opt: Scalable Deep Reinforcement Learning of Vision-Based Robotic Manipulation Skills**

about four hours



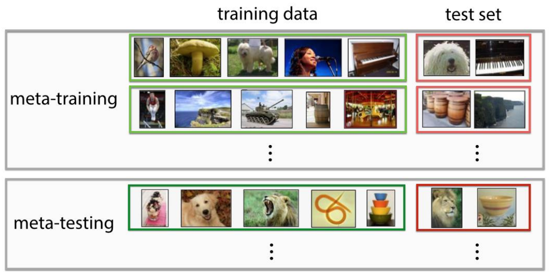
about four weeks, nonstop

people can learn new skills **extremely** quickly

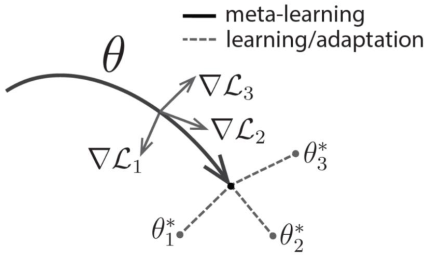
how?

we never learn from scratch!

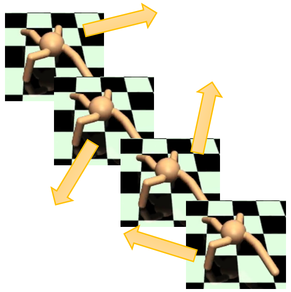
can we transfer past experience in order to *learn how to learn?*



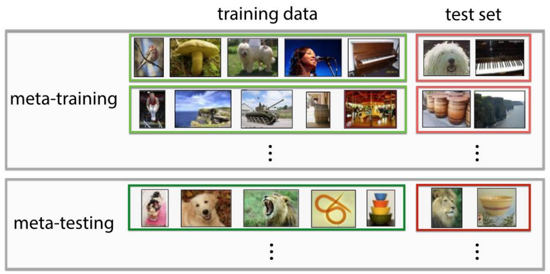
The meta-learning/few-shot learning problem



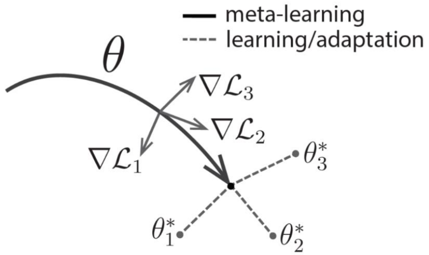
A simpler, *model-agnostic*, meta-learning method



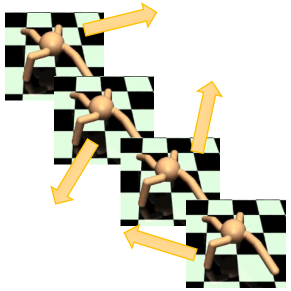
Unsupervised meta-learning



The meta-learning/few-shot learning problem



A simpler, *model-agnostic*, meta-learning method



Unsupervised meta-learning

Few-shot learning: problem formulation in pictures

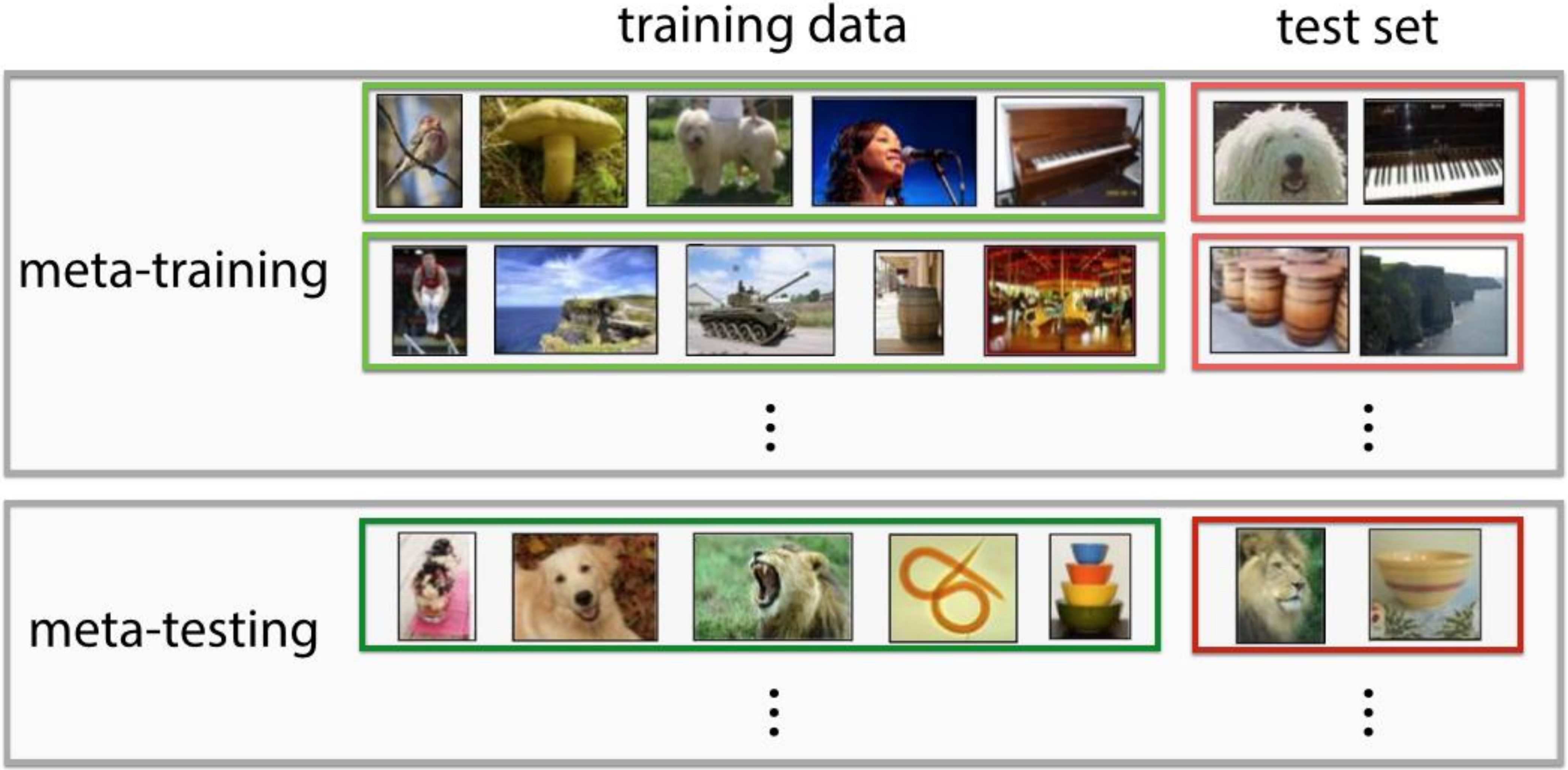
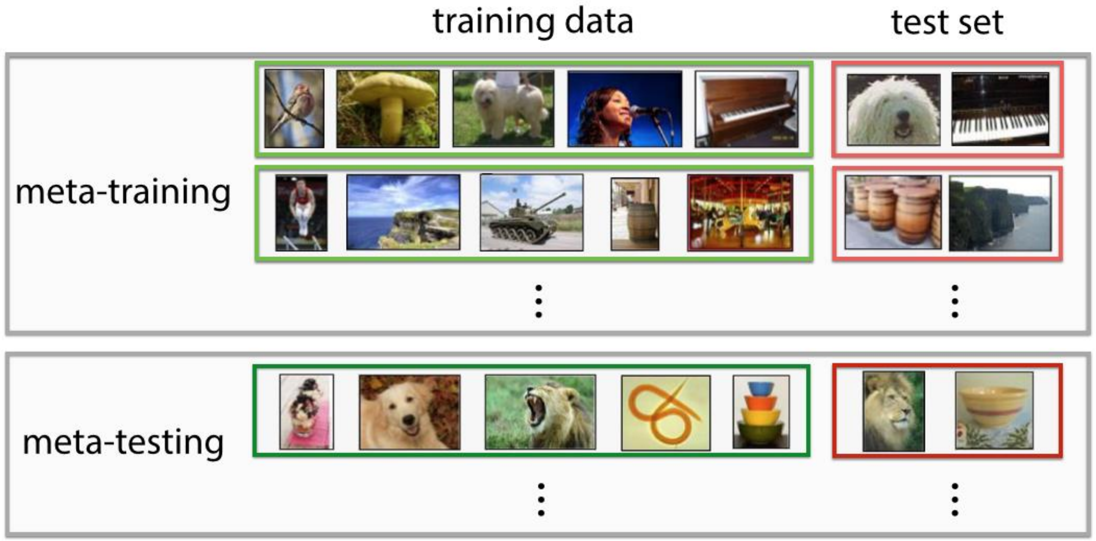


image credit: Ravi & Larochelle '17

Few-shot learning: problem formulation in equations

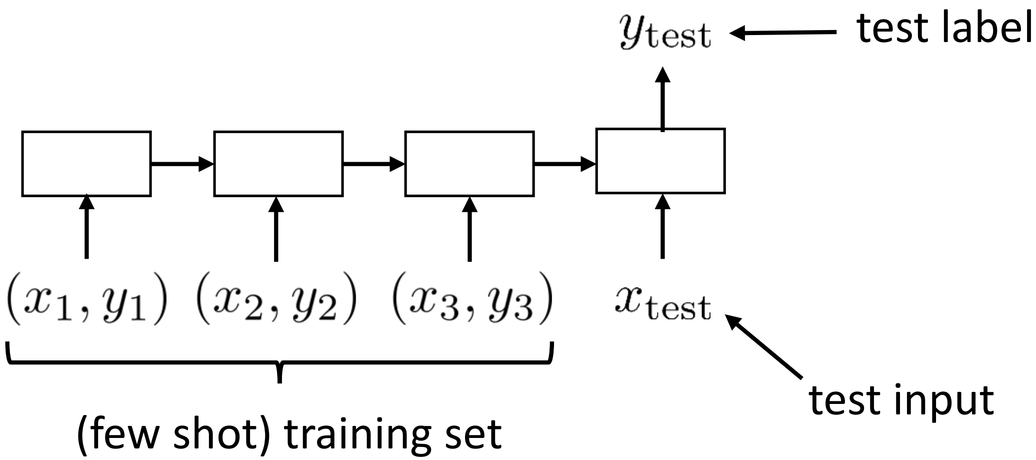


$$\text{supervised learning:}\quad f(x) \to y$$

input (e.g., image) output (e.g., label)

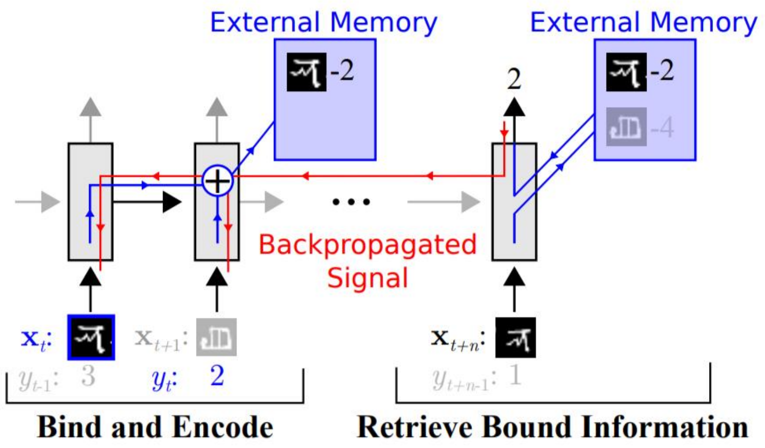
$$\text{supervised meta-learning:}\quad f(\mathcal{D}_{\text{train}}, x) \to y$$

training set

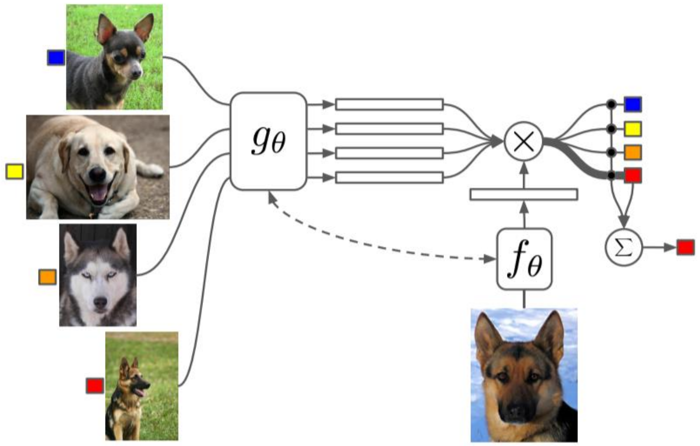


- How to read in training set?
  - Many options, RNNs can work

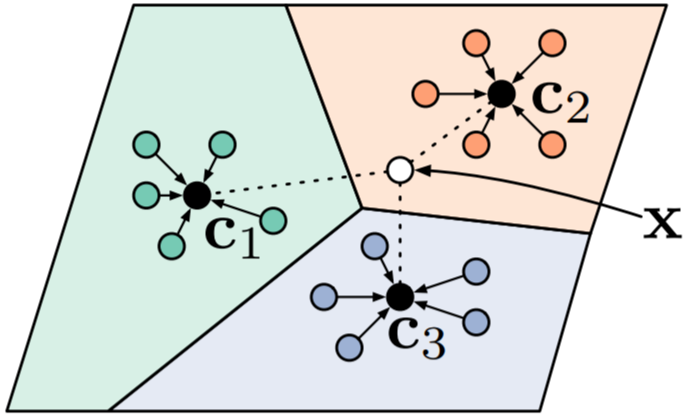
Some examples of representations



Santoro et al. "Meta-Learning with Memory-Augmented Neural Networks."



Vinyals et al. "Matching Networks for One-Shot Learning"

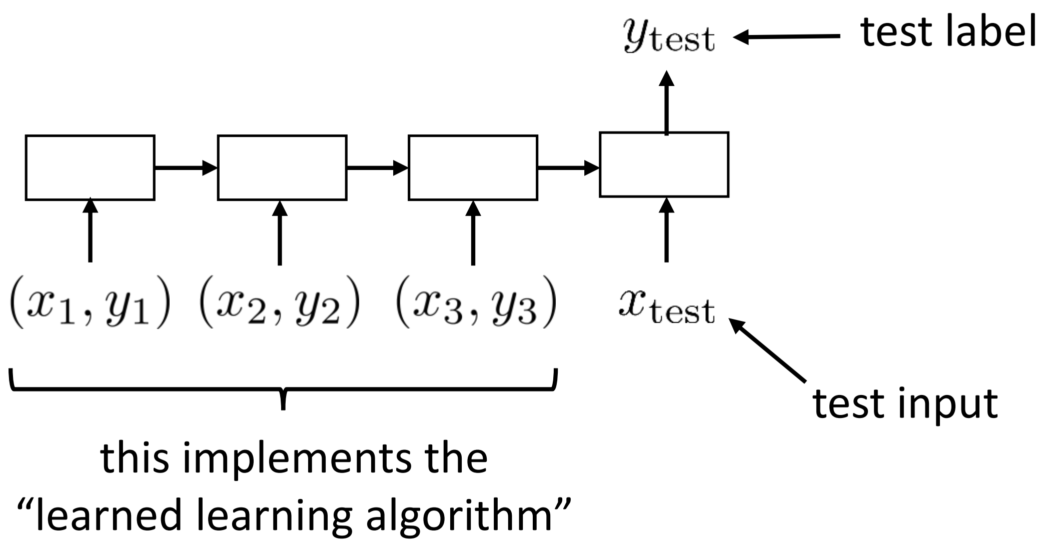


Snell et al. "Prototyping Networks for Few-Shot Learning"

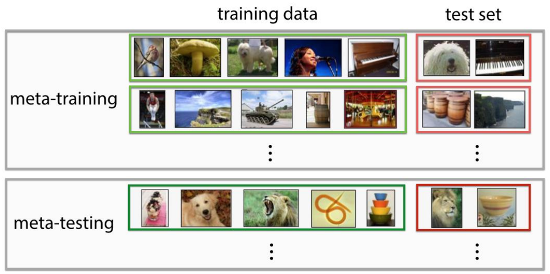
...and many *many* **many** others!

What kind of *algorithm* is learned?

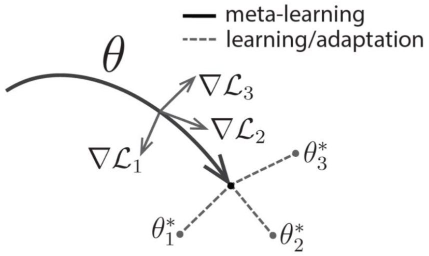
RNN-based meta-learning



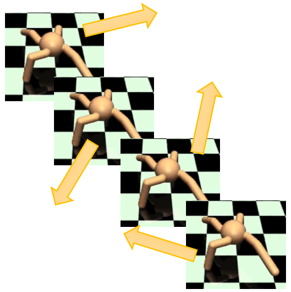
- Does it converge?
  - Kind of?
- What does it converge to?
  - Who knows…
- What to do if it's not good enough?
  - Nothing…



The meta-learning/few-shot learning problem

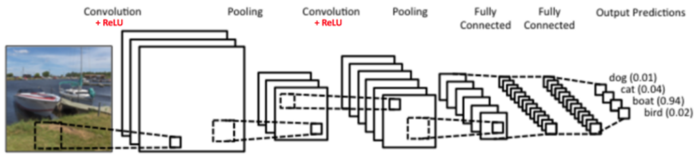


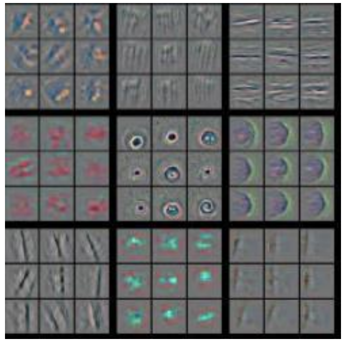
A simpler, *model-agnostic*, meta-learning method

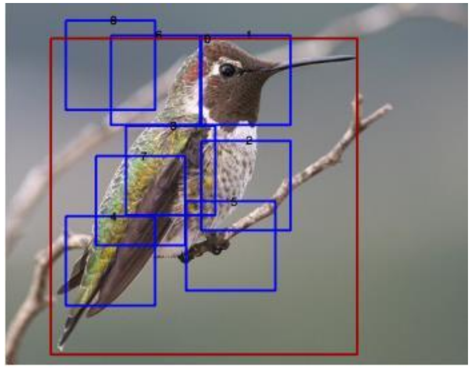


Unsupervised meta-learning

Let’s step back a bit…







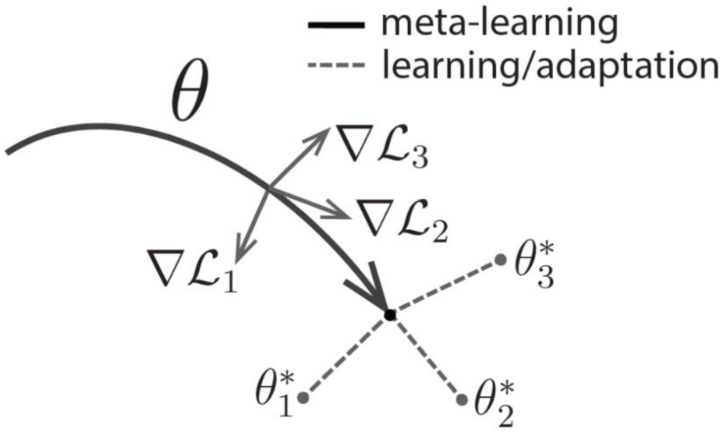
is pretraining a *type* of meta-learning?

better features = faster learning of new task!

Model-agnostic meta-learning

a general recipe:

$\mathcal{D}^i_{\text{train}}$ $\mathcal{D}^i_{\text{test}}$



$$\theta \leftarrow \theta - \beta \sum_i \nabla_\theta \underbrace{\mathcal{L}(\theta - \alpha \nabla_\theta \mathcal{L}(\theta, \mathcal{D}^i_{\text{train}}), \mathcal{D}^i_{\text{test}})}_{\text{“meta-loss” for task } i}$$

* in general, can take more than one gradient step here

** we often use 4 – 10 steps

Chelsea Finn

Finn et al., “Model-Agnostic Meta-Learning”

What did we just do?

supervised learning: $f(x) \to y$

supervised meta-learning: $f(\mathcal{D}_{\text{train}}, x) \to y$

model-agnostic meta-learning: $f_{\text{MAML}}(\mathcal{D}_{\text{train}}, x) \to y$

$$f_{\text{MAML}}(\mathcal{D}_{\text{train}}, x) = f_{\theta'}(x)$$

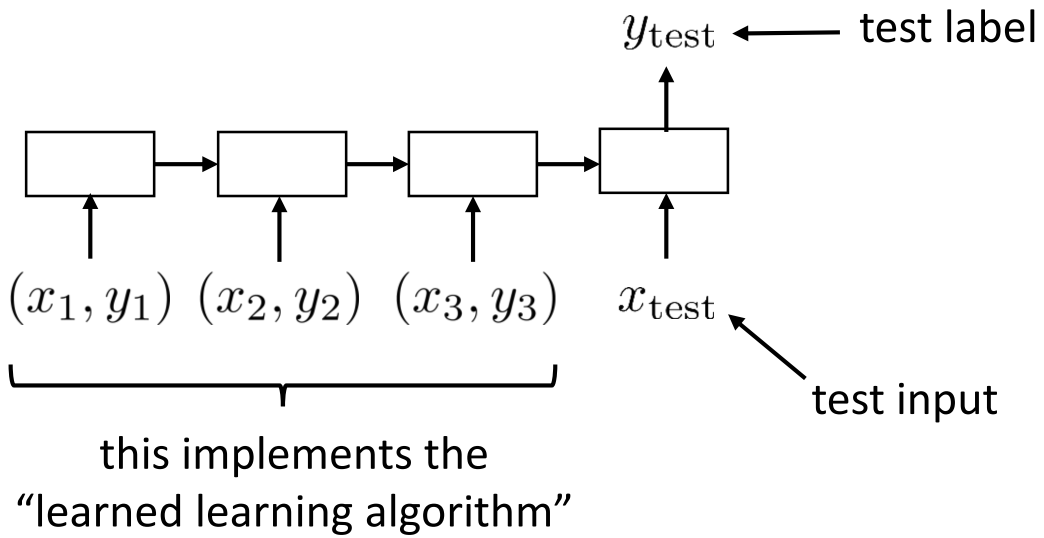
$$\theta' = \theta - \alpha \sum_{(x,y) \in \mathcal{D}_{\text{train}}} \nabla_\theta \mathcal{L}(f_\theta(x), y)$$

Just another computation graph…

Can implement with any autodiff package (e.g., TensorFlow)

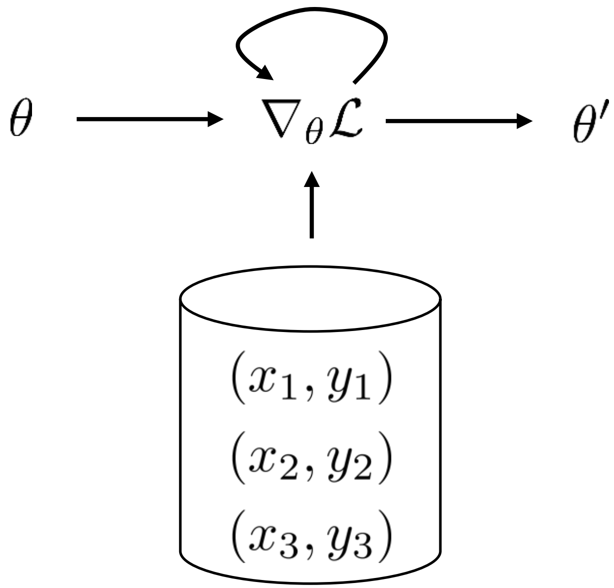
Why does it work?

RNN-based meta-learning



- Does it converge?
  - Kind of?
- What does it converge to?
  - Who knows…
- What to do if it's not good enough?
  - Nothing…

MAML



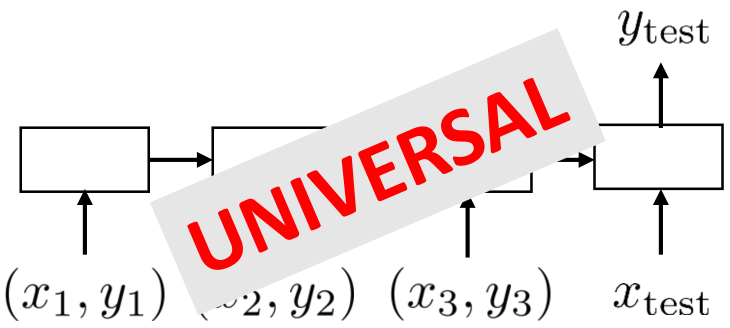
- Does it converge?
  - Yes (it's gradient descent…)
- What does it converge to?
  - A local optimum (it's gradient descent…)
- What to do if it's not good enough?
  - Keep taking gradient steps (it's gradient descent…)

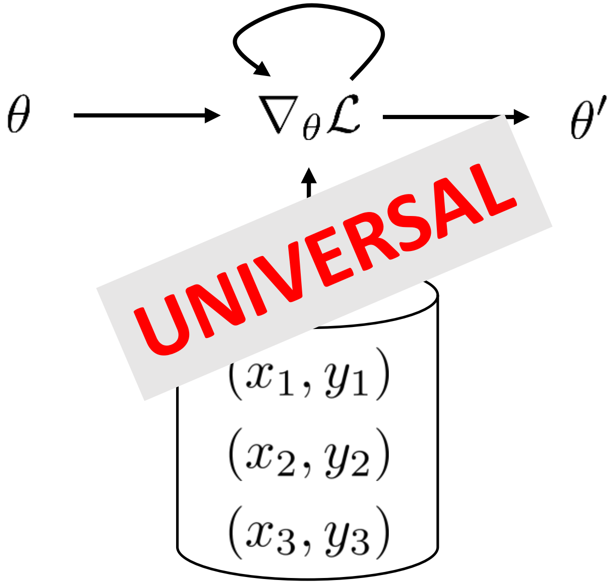
Universality

Did we lose anything?

Universality: meta-learning can learn any "algorithm"

more precisely, can represent any function $f(\mathcal{D}_{\text{train}}, x)$

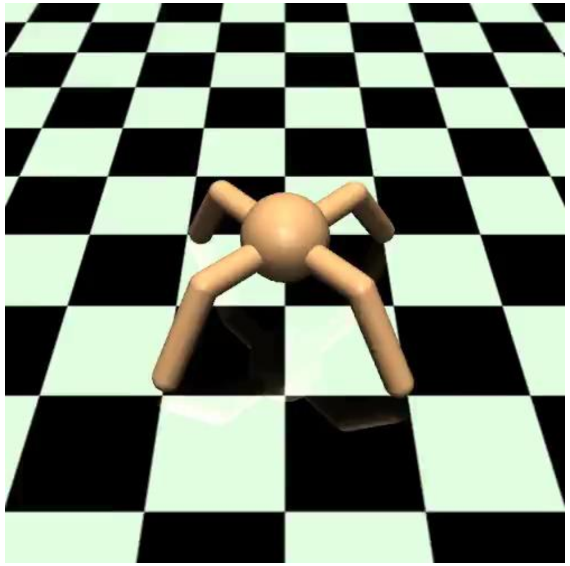


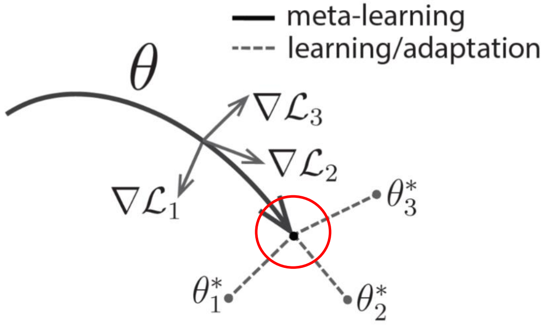


Finn & Levine. "Meta-Learning and Universality"

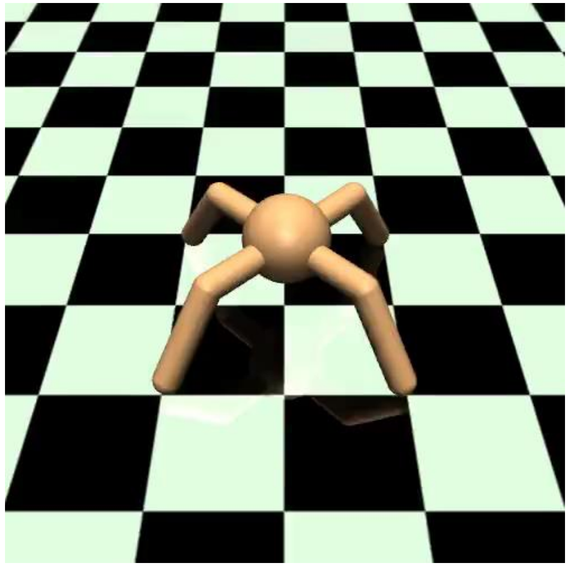
Model-agnostic meta-learning: forward/backward locomotion

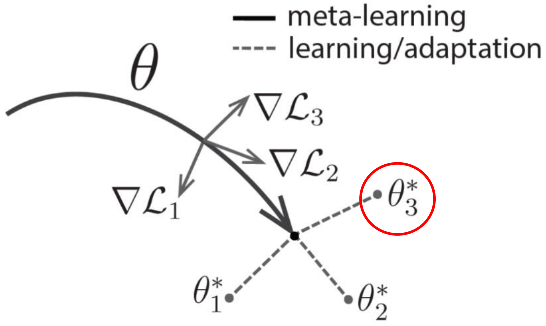
after MAML training



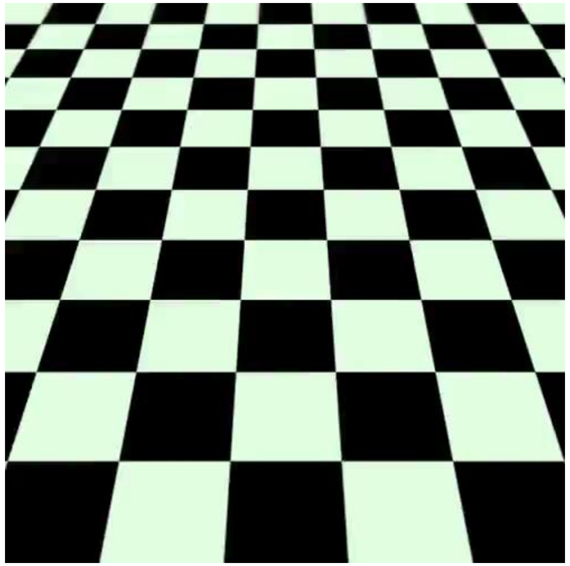


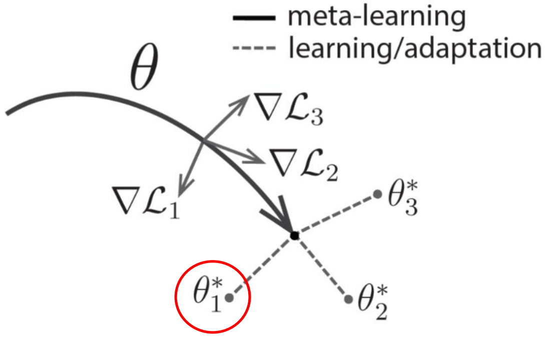
after 1 gradient step (forward reward)



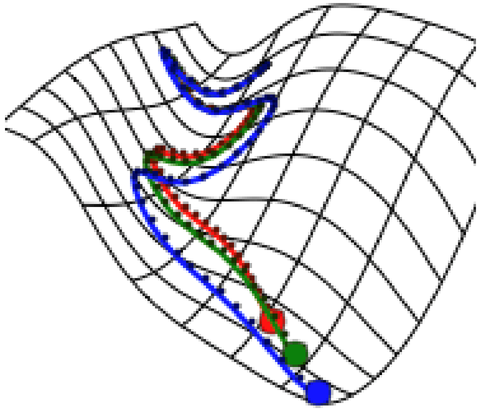


after 1 gradient step (backward reward)

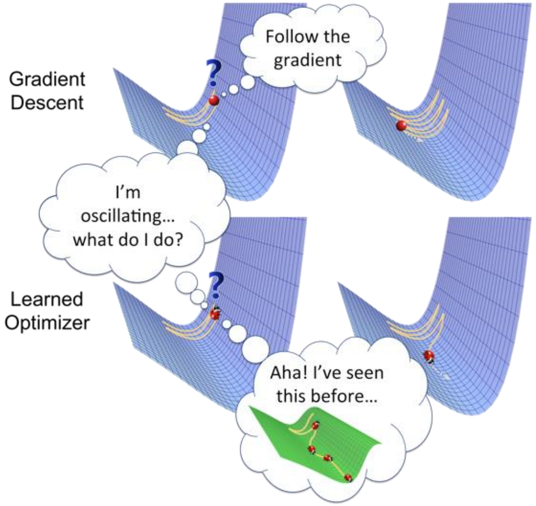




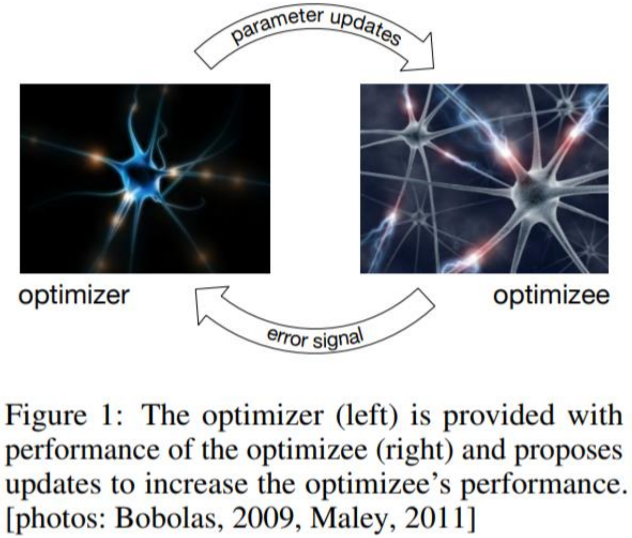
Related work



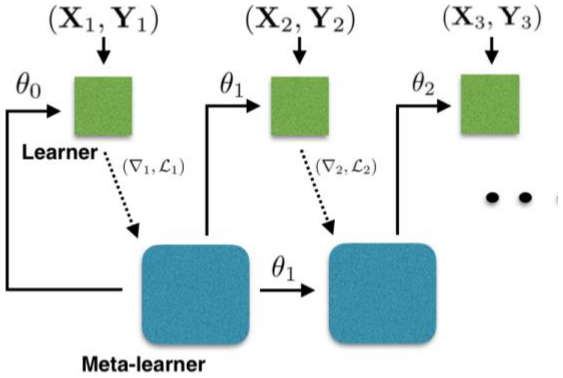
Maclaurin et al. “Gradient-based hyperparameter optimization”



Li & Malik. “Learning to optimize”



Andrychowicz et al. “Learning to learn by gradient descent by gradient descent.”

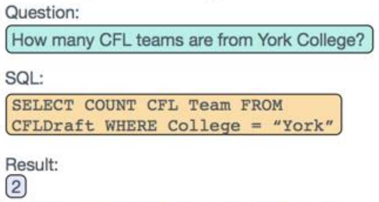


Ravi & Larochelle. “Optimization as a model for few-shot learning”

...and many *many* **many** others!

Follow-up work

Program Synthesis

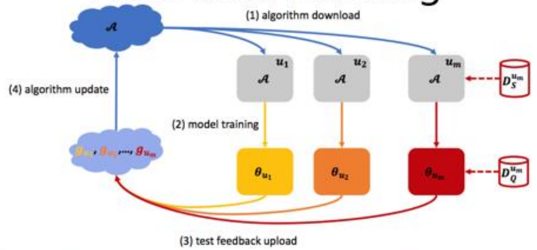


Huang, Wang, Singh, Yih, He NAACL '18

Learning to Learn Distributions

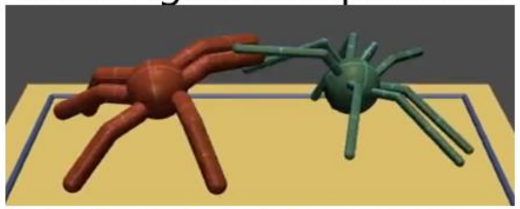
Reed, Chen, Paine, van den Oord, Eslami, Rezende, Vinyals, de Freitas ICLR '18

Federated Learning



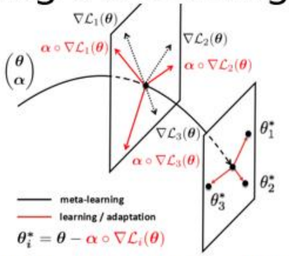
Chen, Dong, Li, He arXiv '18

Multi-Agent Competitions



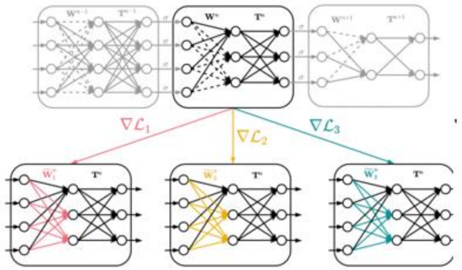
Al-Shedivat, Bansal, Burda, Sutskever Mordatch, Abbeel ICLR '18

Learning the learning rate



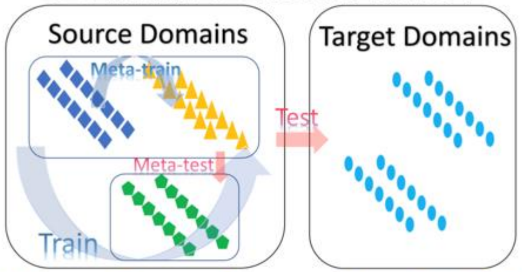
Li, Zhou, Chen, Li arXiv '17

Masked Transformations



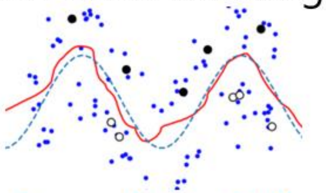
Lee & Choi arXiv '18

Domain Generalization



Li, Yang, Song, Hospedales AAAI '18

Semi-Supervised Few-Shot Learning



Boney & Ilin ICLR workshop track '18

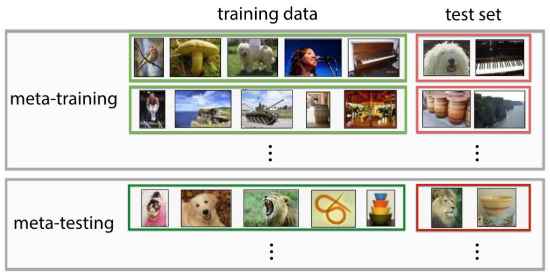
…and the results keep getting better

MiniImagenet few-shot benchmark: 5-shot 5-way

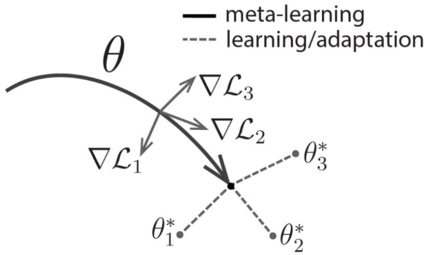
Finn et al. '17: 63.11%

Li et al. '17: 64.03%

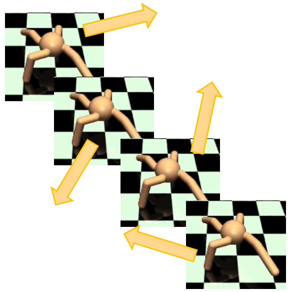
Kim et al. '18 (AutoMeta): 76.29%



The meta-learning/few-shot learning problem



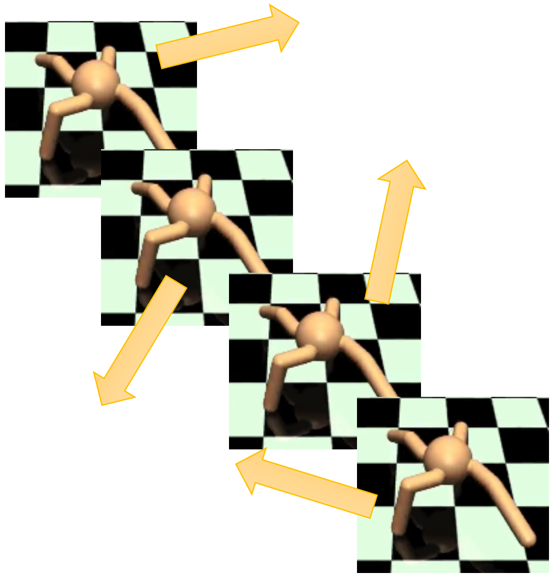
A simpler, *model-agnostic*, meta-learning method



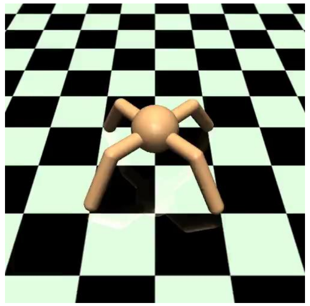
Unsupervised meta-learning

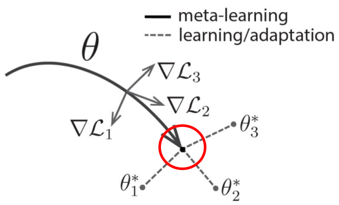
Let’s Talk about Meta-Overfitting

- Meta learning requires task distributions
- When there are too few meta-training tasks, we can *meta-overfit*
- Specifying task distributions is hard, especially for meta-RL!
- Can we propose tasks *automatically*?

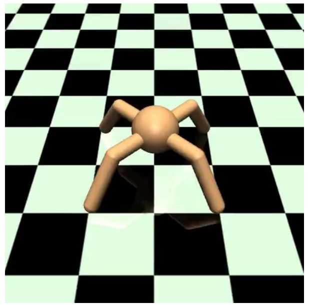


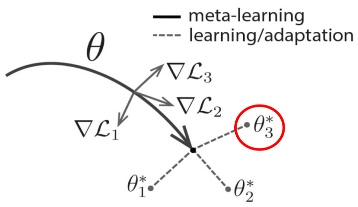
after MAML training



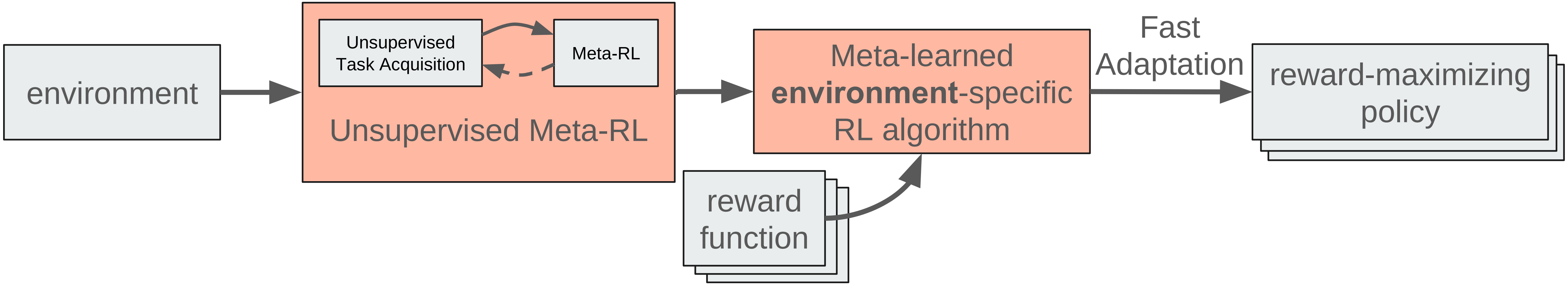


after 1 gradient step





A General Recipe for Unsupervised Meta-RL



Abhishek Gupta

Ben Eysenbach

Chelsea Finn

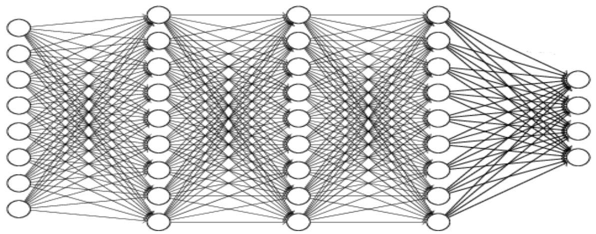
Gupta, Eysenbach, Finn, Levine. Unsupervised Meta-Learning for Reinforcement Learning.

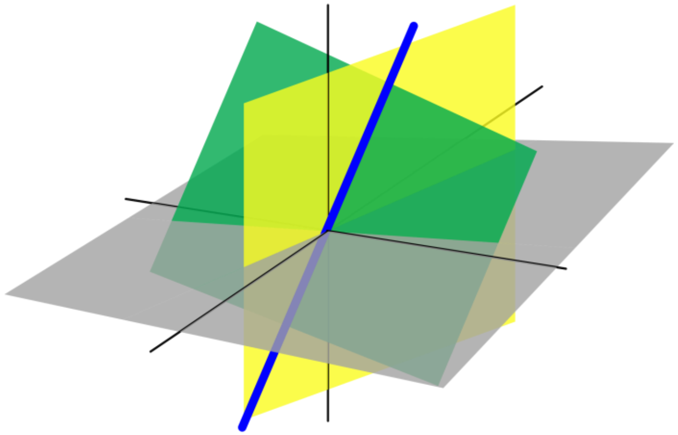
Random Task Proposals

- Use randomly initialize discriminators for reward functions

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

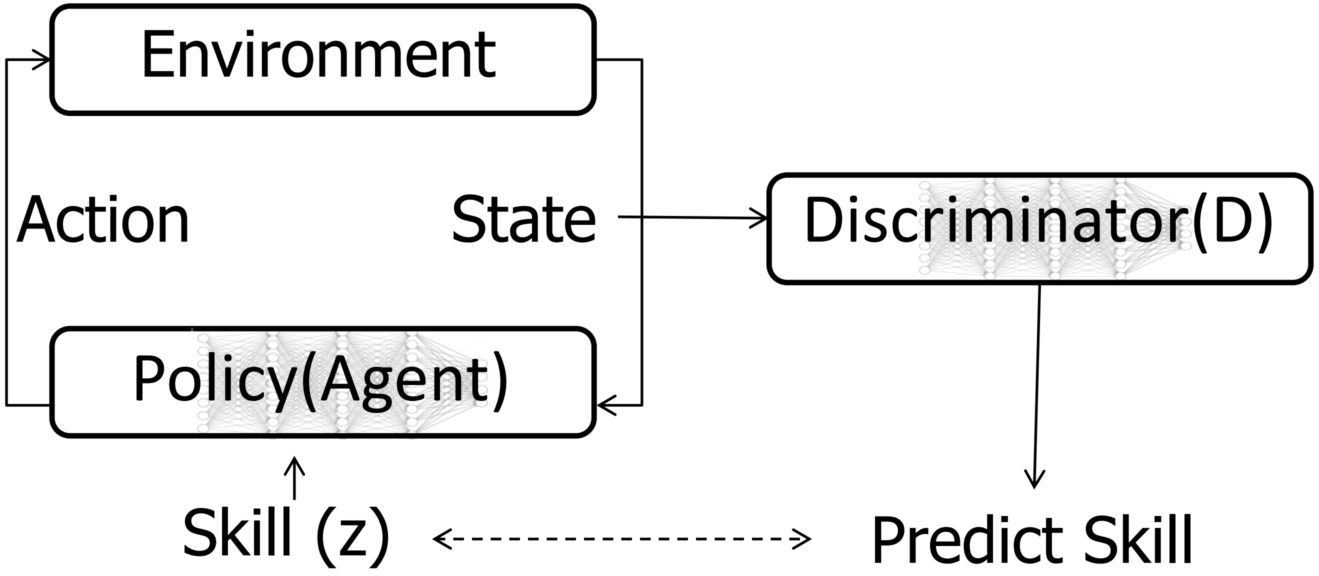
D → randomly initialized network





- Important: Random functions over state space, **not** random policies

Diversity-Driven Proposals

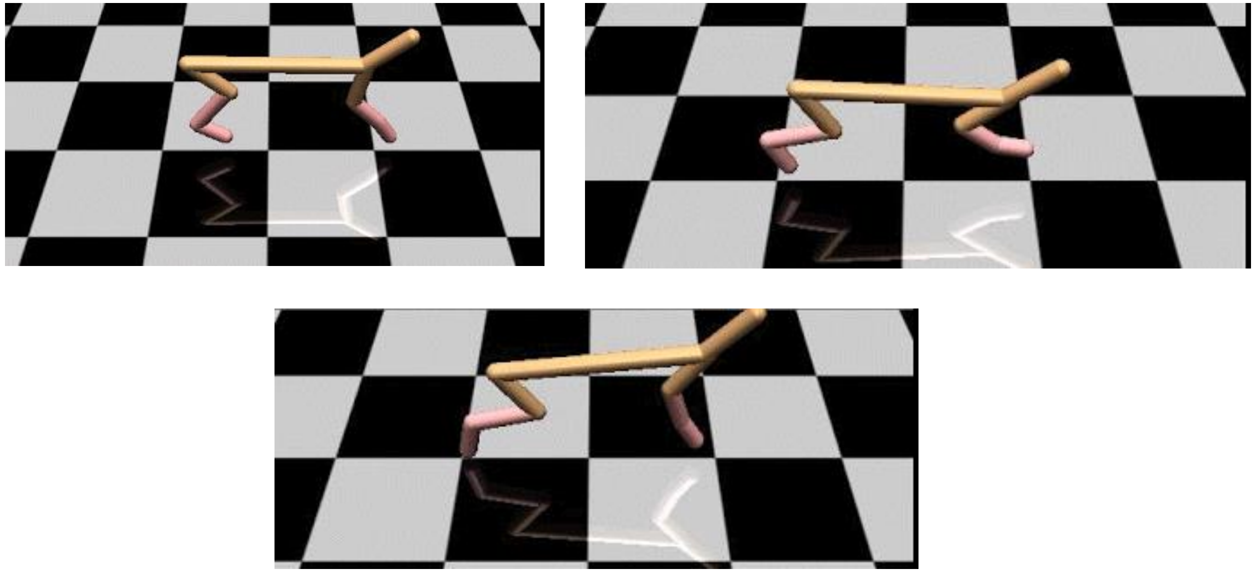


- Policy → visit states which are discriminable
- Discriminator → predict skill from state

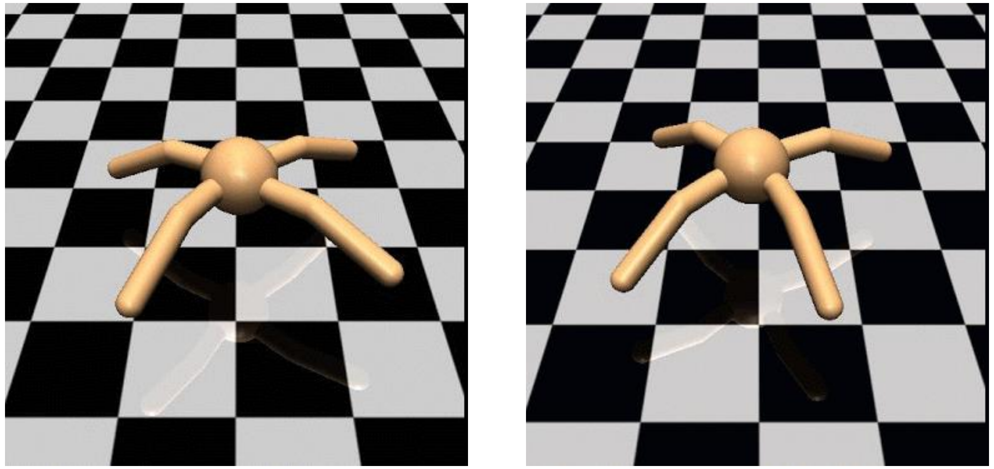
Task Reward for UML: $R(s, z) = \log p_D(z|s)$

Eysenbach, Gupta, Ibarz, Levine. Diversity is All You Need.

Examples of Acquired Tasks



Cheetah

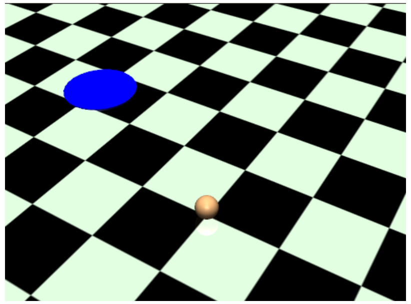


Ant

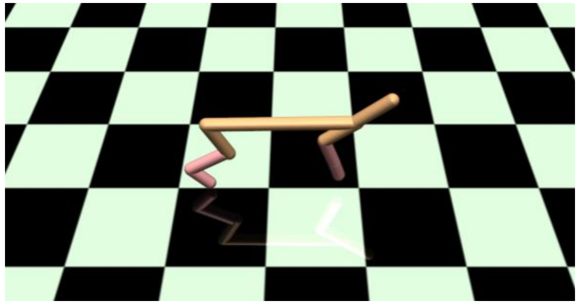
Eysenbach, Gupta, Ibarz, Levine. Diversity is All You Need.

Does it work?

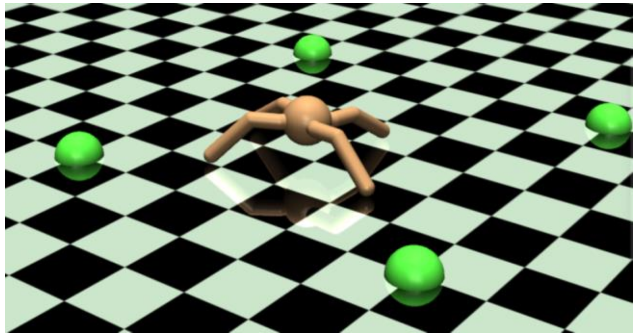
2D Navigation

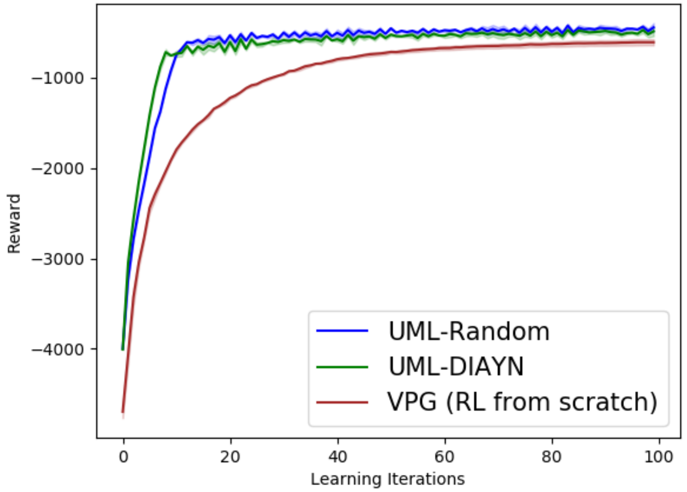


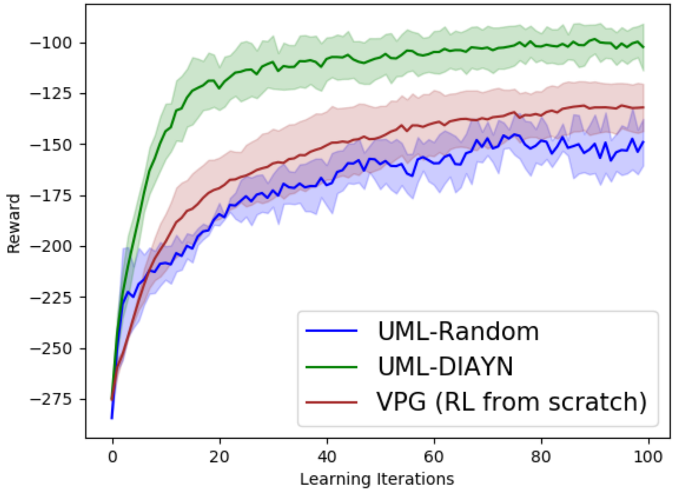
Cheetah

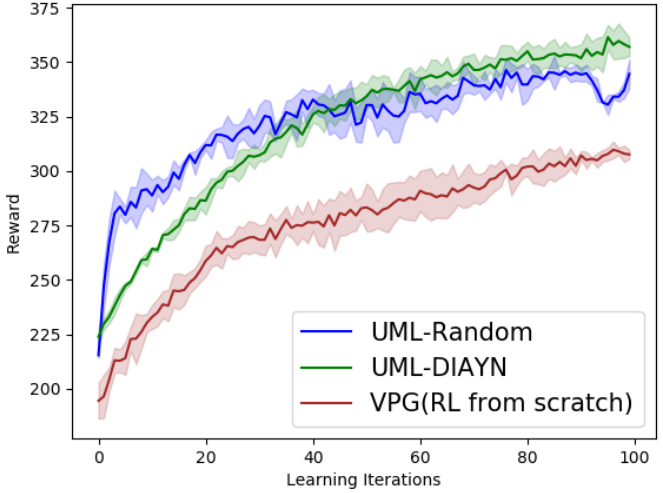


Ant





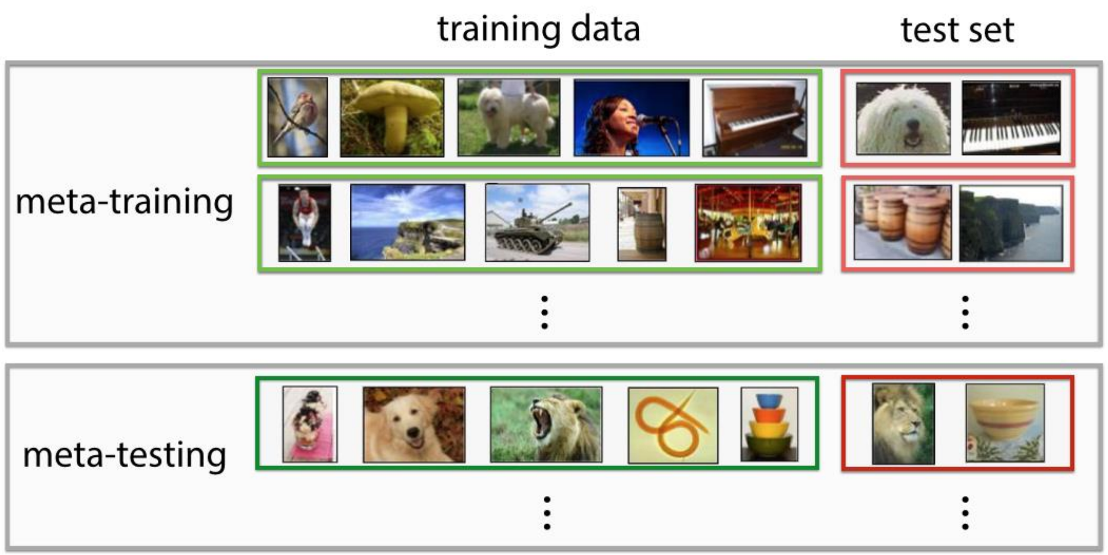


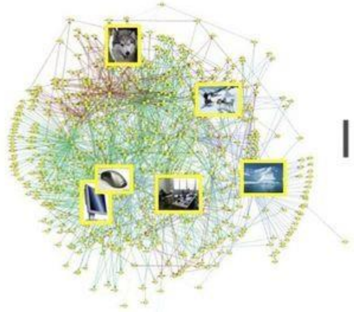


Meta-test performance with rewards

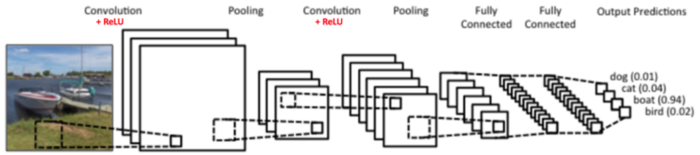
Gupta, Eysenbach, Finn, Levine. Unsupervised Meta-Learning for Reinforcement Learning.

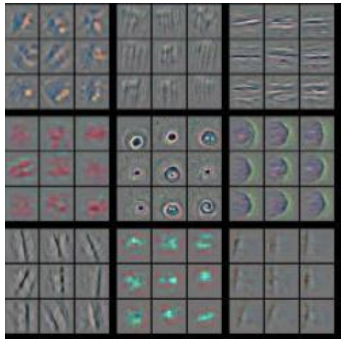
What about supervised learning?

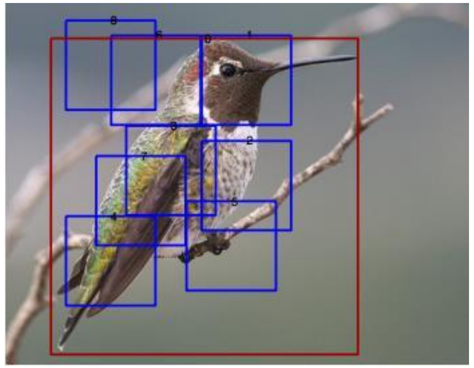




IMAGENET



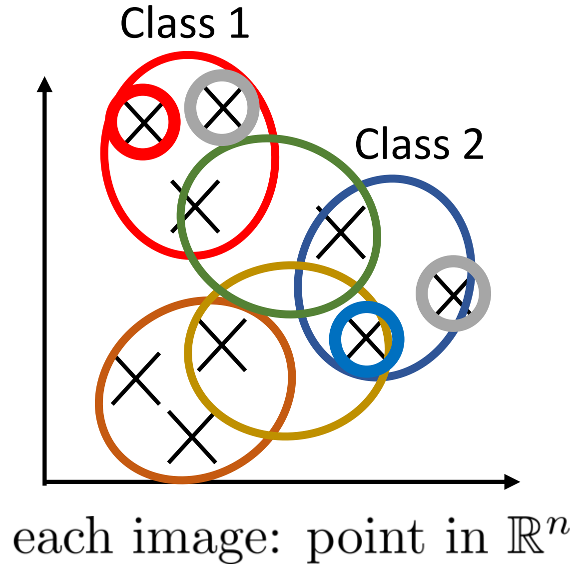


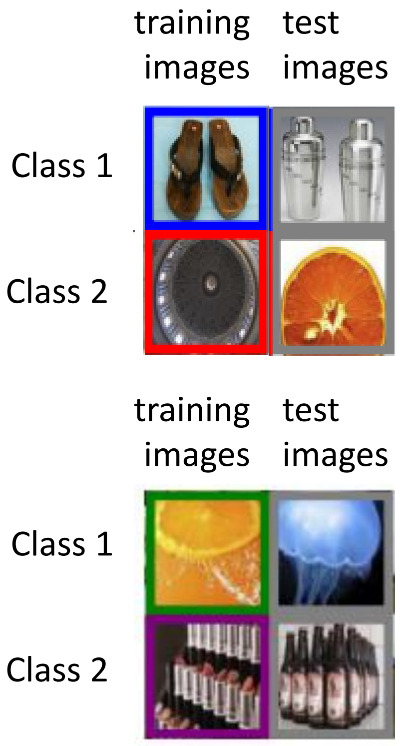


Can we meta-train on only *unlabeled* images?

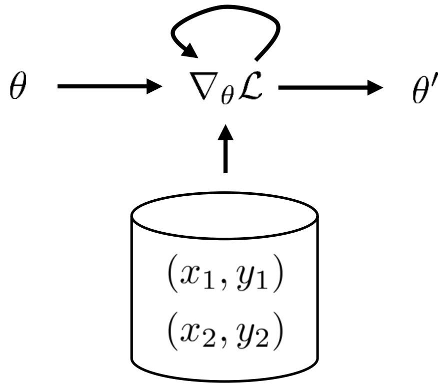
unsupervised learning → **task proposals** → **meta-learning**

$\{\mathbf{x}_i\}$



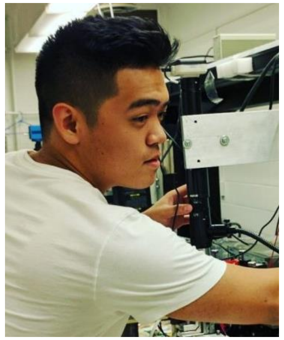


MAML



$$\theta \leftarrow \theta - \beta \sum_i \nabla_\theta \, \mathcal{L}(\theta - \alpha \nabla_\theta \mathcal{L}(\theta, \mathcal{D}^i_{\text{train}}), \mathcal{D}^i_{\text{test}})$$

Kyle Hsu



Chelsea Finn

But... does it outperform unsupervised learning?

Hsu, Levine, Finn. Unsupervised Learning via Meta-Learning.

Results: unsupervised meta-learning

unsupervised learning → **task proposals** → **meta-learning**

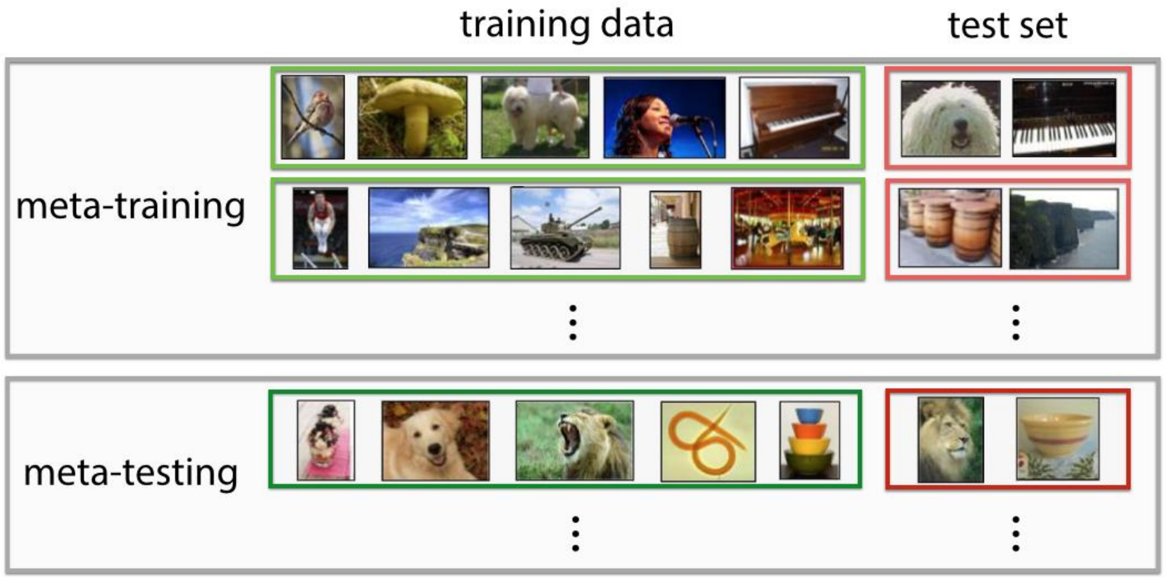
a few choices:

BiGAN – Donahue et al. ’17
DeepCluster – Caron et al. ‘18

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

*mini*ImageNet: 5 shot, 5 way

method	accuracy

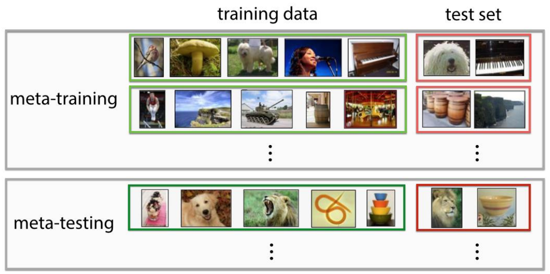


no true labels *at all*!

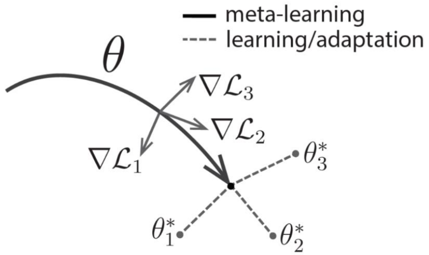
Same story across:

- 3 different embedding methods
- 4 datasets (Omniglot, miniImageNet, CelebA, MNIST)

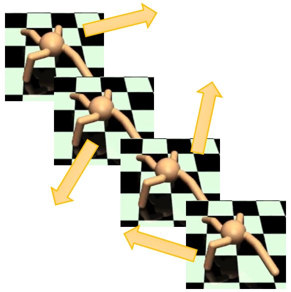
Hsu, Levine, Finn. Unsupervised Learning via Meta-Learning.



The meta-learning/few-shot learning problem

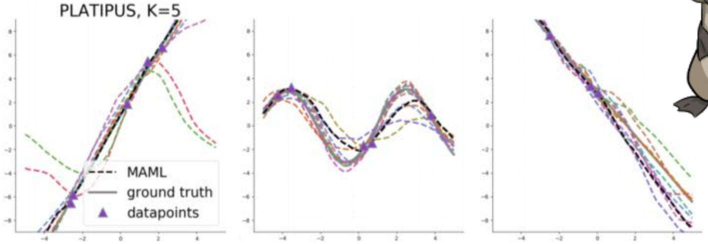


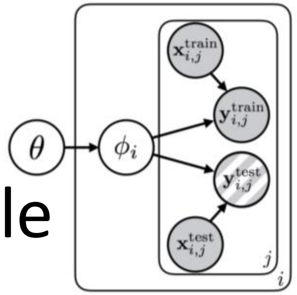
A simpler, *model-agnostic*, meta-learning method



Unsupervised meta-learning

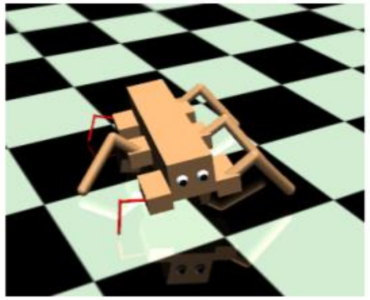
What's next?





Probabilistic meta-learning: learn to sample *multiple hypotheses*

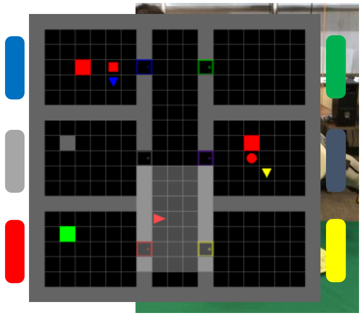
Finn*, Xu*, Levine. **Probabilistic Model-Agnostic Meta-Learning.** 2018.



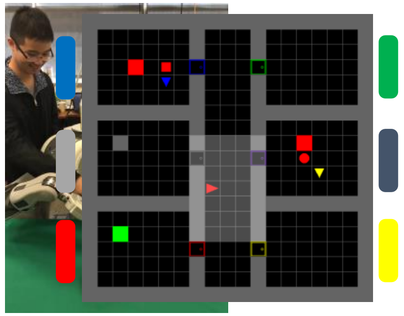
Meta-learning online learning & continual learning

Nagabandi, Finn, Levine. **Deep Online Learning via Meta-Learning: Continual Adaptation via Model-Based RL.** 2018.

Instruction: Move blue triangle to green goal.



Correction 1: Enter the blue room.

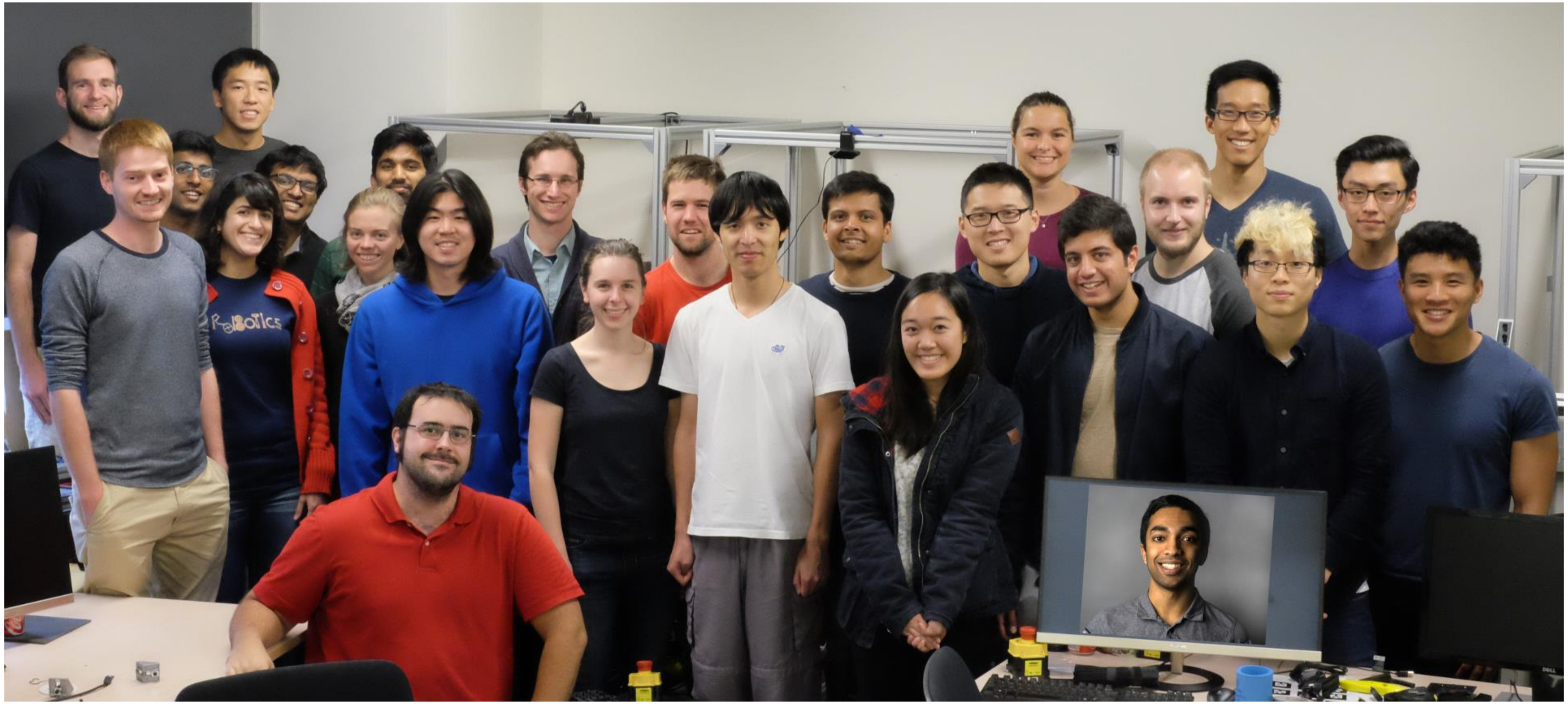


Correction 2: Enter the red room.

Meta-learning to interpret weak supervision and natural language

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

Co-Reyes, Gupta, Sanjeev, Altieri, DeNero, Abbeel, Levine. **Meta-Learning Language-Guided Policy Learning.** 2018.



RAIL
Robotic AI & Learning Lab

website: http://rail.eecs.berkeley.edu
source code: http://rail.eecs.berkeley.edu/code.html