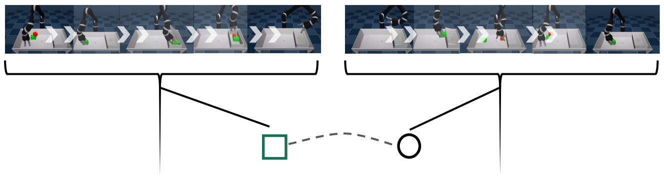
Skill Representation and Supervision in Multi-Task RL

Karol Hausman

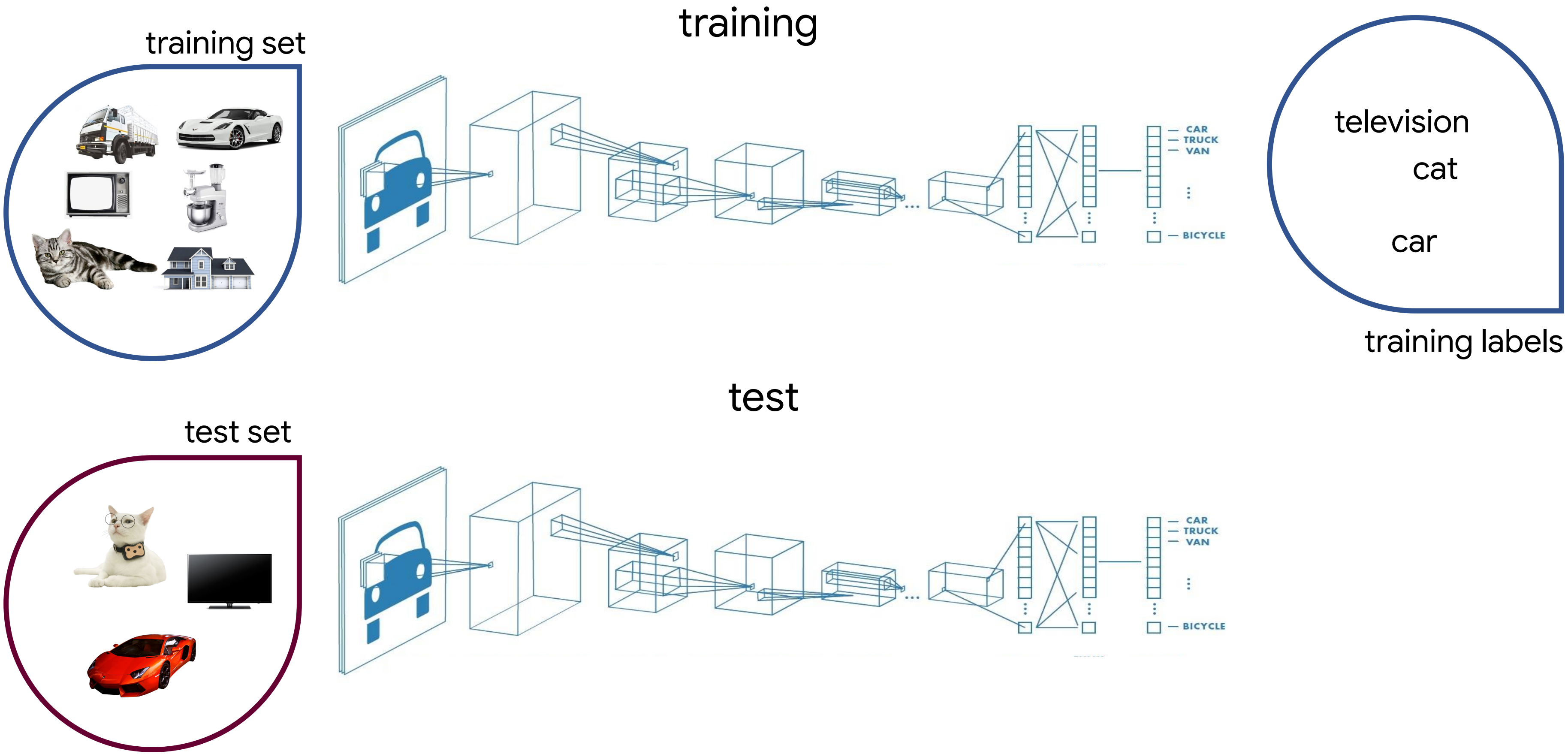
Google Brain



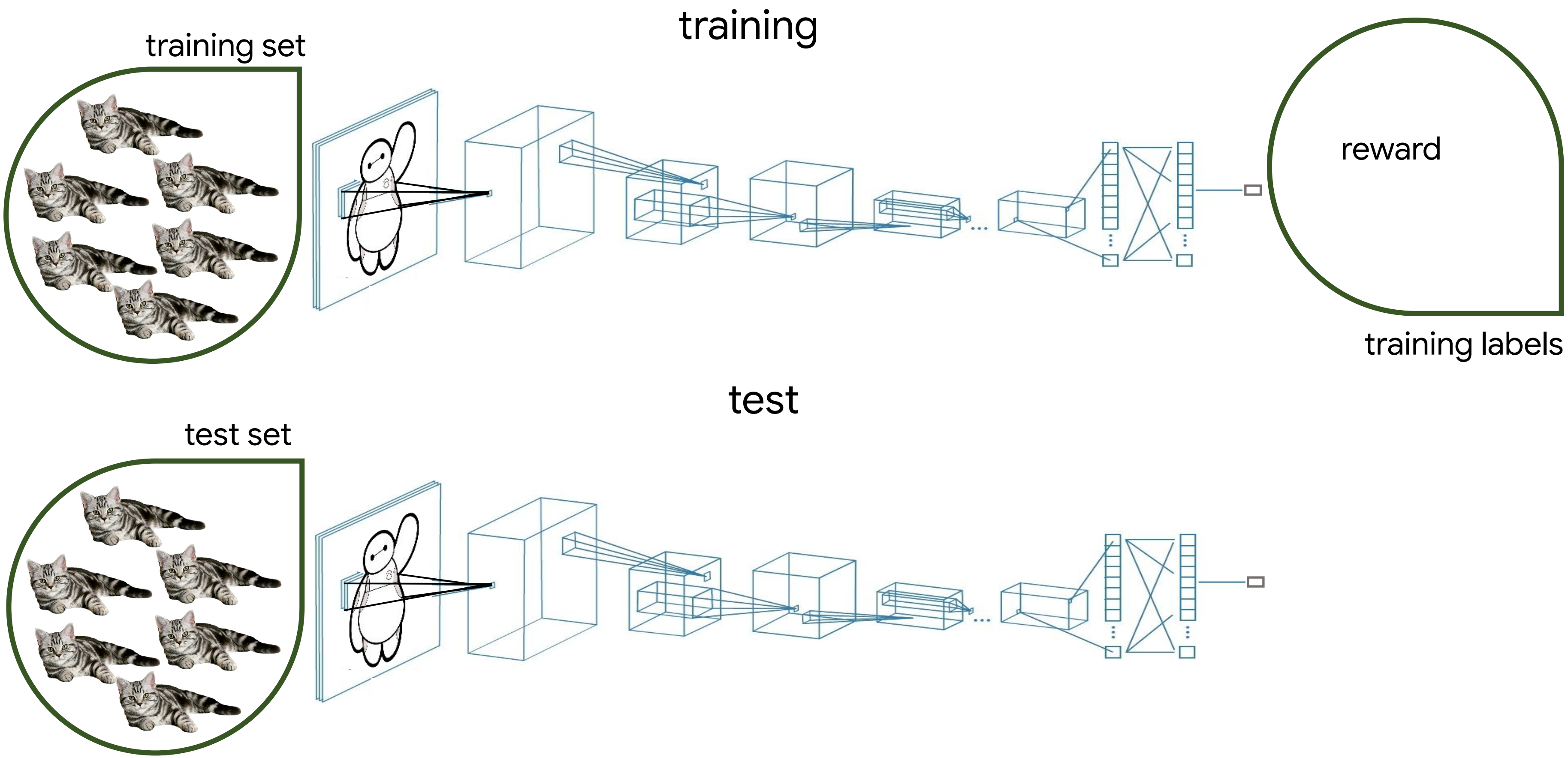
In collaboration with USC

Why multi-task reinforcement learning?

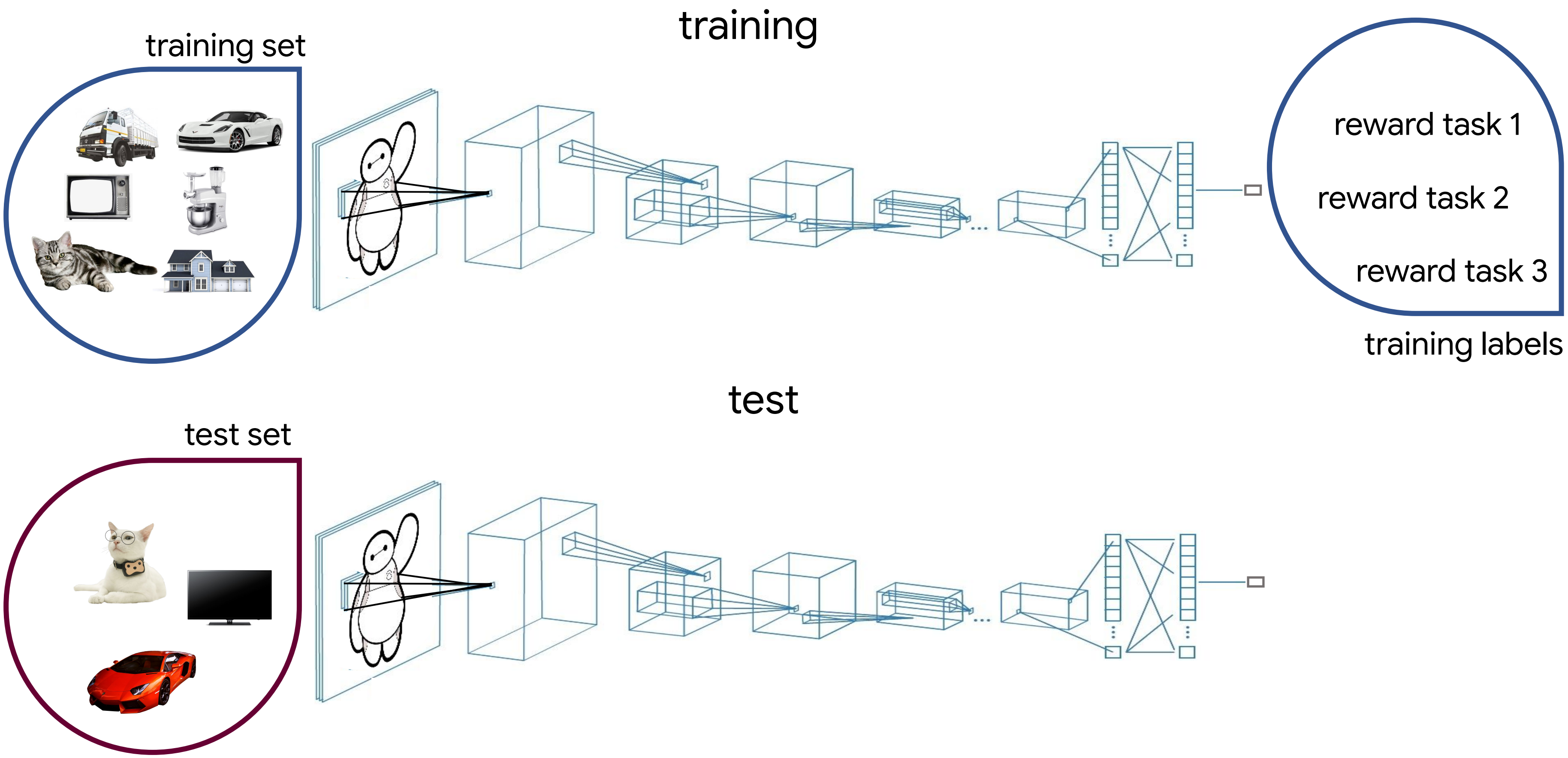
Supervised learning: generalization



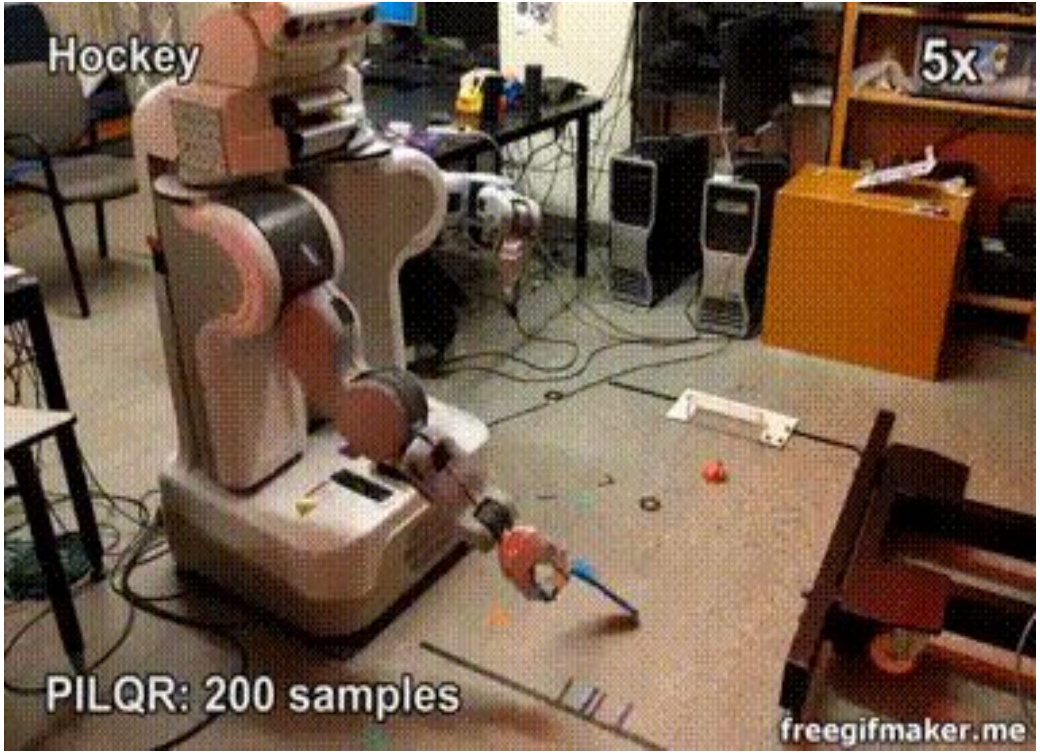
Single-task deep RL: generalization



Multi-task deep RL: generalization



Single-task deep RL: resets



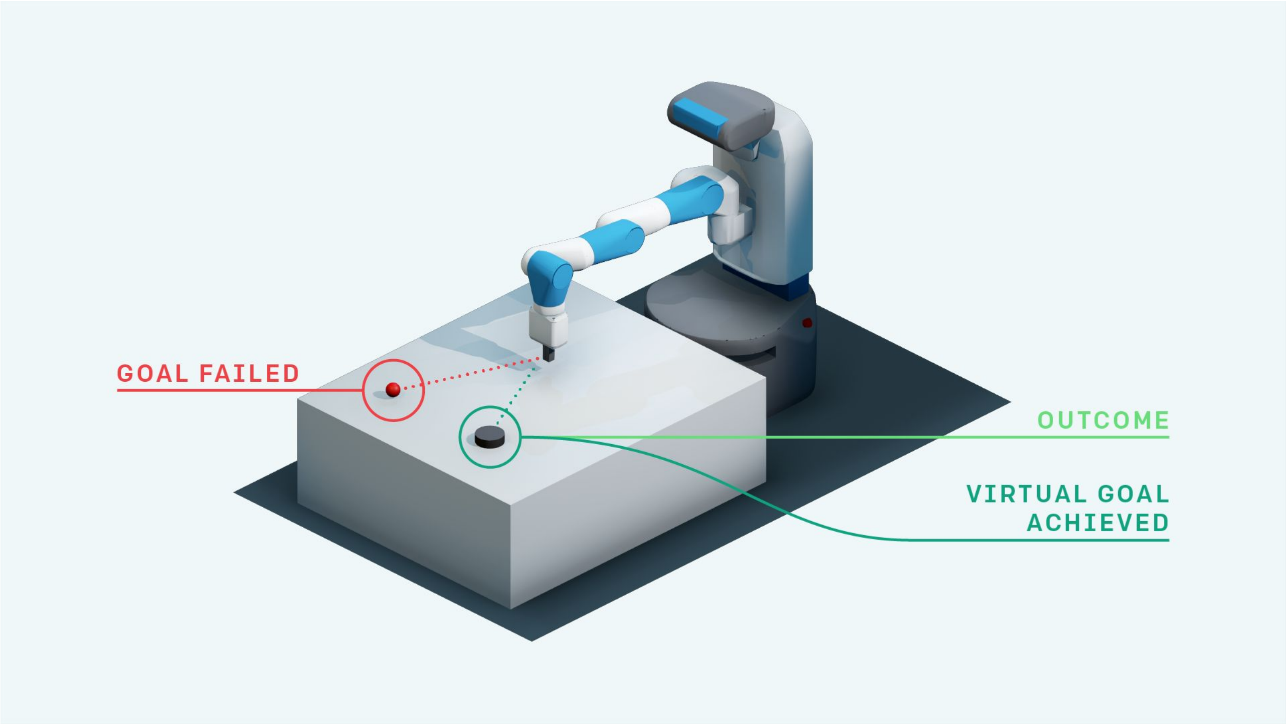
[Combining Model-Based and Model-Free Updates for Trajectory-Centric Reinforcement Learning, Chebotar*, Hausman*, Zhang*, et al., 2017]

Multi-task deep RL: resets

[Supervision via Competition: Robot Adversaries for Learning Tasks,
Pinto, et al., 2017]

Single-task deep RL: rewards

Multi-task deep RL: rewards



[Hindsight Experience Replay, Andrychowicz, et al., 2017]

Why **not** multi-task reinforcement learning?

Multi-task deep RL

Challenges

- task specifications, what constitutes a task, how to represent a skill?
- reuse of already-learned skills
- optimization of multiple tasks (conflicting gradients, gradient magnitudes)
- data imbalance issues (harder easier tasks, good exploration in all of them)
- multiple skills - multiple pains: rewards, setups, etc.
- efficient sequencing of skills at test time

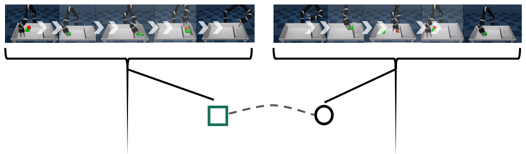
.

Multi-task deep RL

Challenges

- task specifications, what constitutes a task, how to represent a skill?
- reuse of already-learned skills
- optimization of multiple tasks (conflicting gradients, gradient magnitudes)
- data imbalance issues (harder easier tasks, good exploration in all of them)
- multiple skills - multiple pains: rewards, setups, etc.
- efficient sequencing of skills at test time

Multi-task deep RL



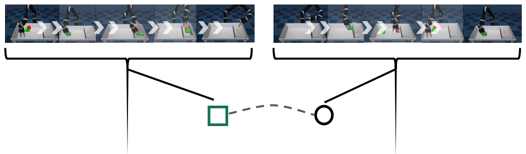
Skill Representation and Reusability

- task specifications, what constitutes a task, how to represent a skill?
- reuse of already-learned skills

Supervision and Efficiency

- multiple skills - multiple pains: rewards, setups, etc.
- efficient sequencing of skills at test time

Multi-task deep RL



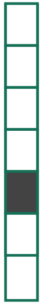
Skill Representation and Reusability

- task specifications, what constitutes a task, how to represent a skill?
- reuse of already-learned skills

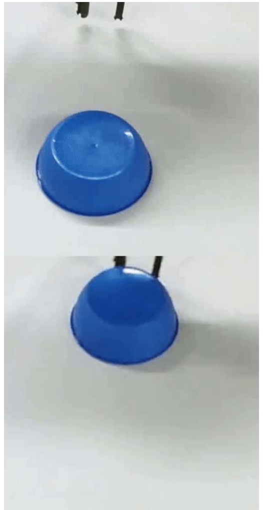
Supervision and Efficiency

- multiple skills - multiple pains: rewards, setups, etc.
- efficient sequencing of skills at test time

Skill representation



task ID



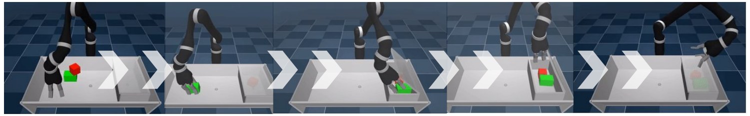
[Visual Reinforcement Learning with Imagined Goals, Pong et al. 2018]

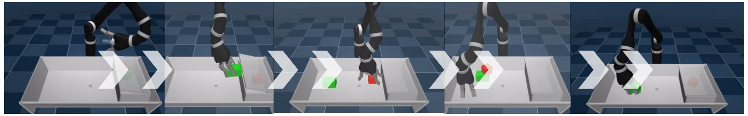
[Progressive Growing of GANs for Improved Quality, Stability, and Variation, Karras et al. 2018]

Latent space in images and policies

Images

Policies

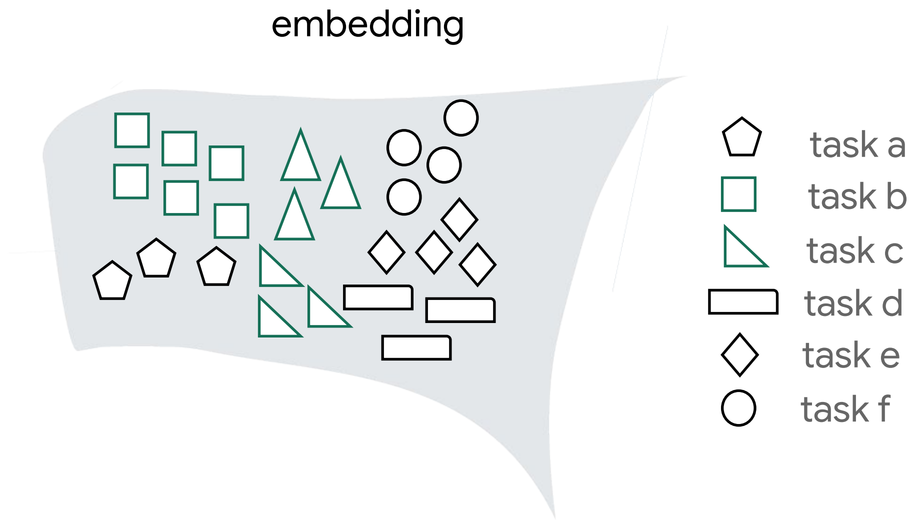




Robot skill embeddings

main idea: learn multiple re-usable skills and their skill embedding

embedding can represent different solutions for every task:

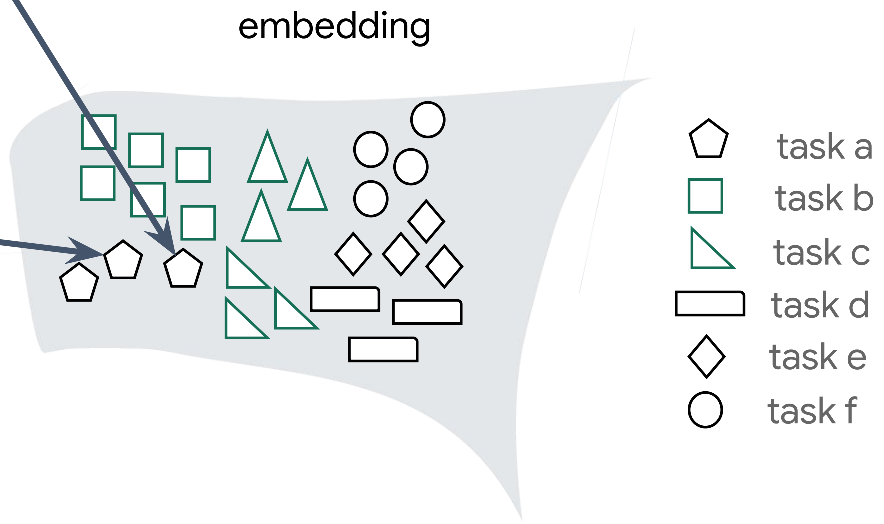


Robot skill embeddings

main idea: learn multiple re-usable skills and their skill embedding

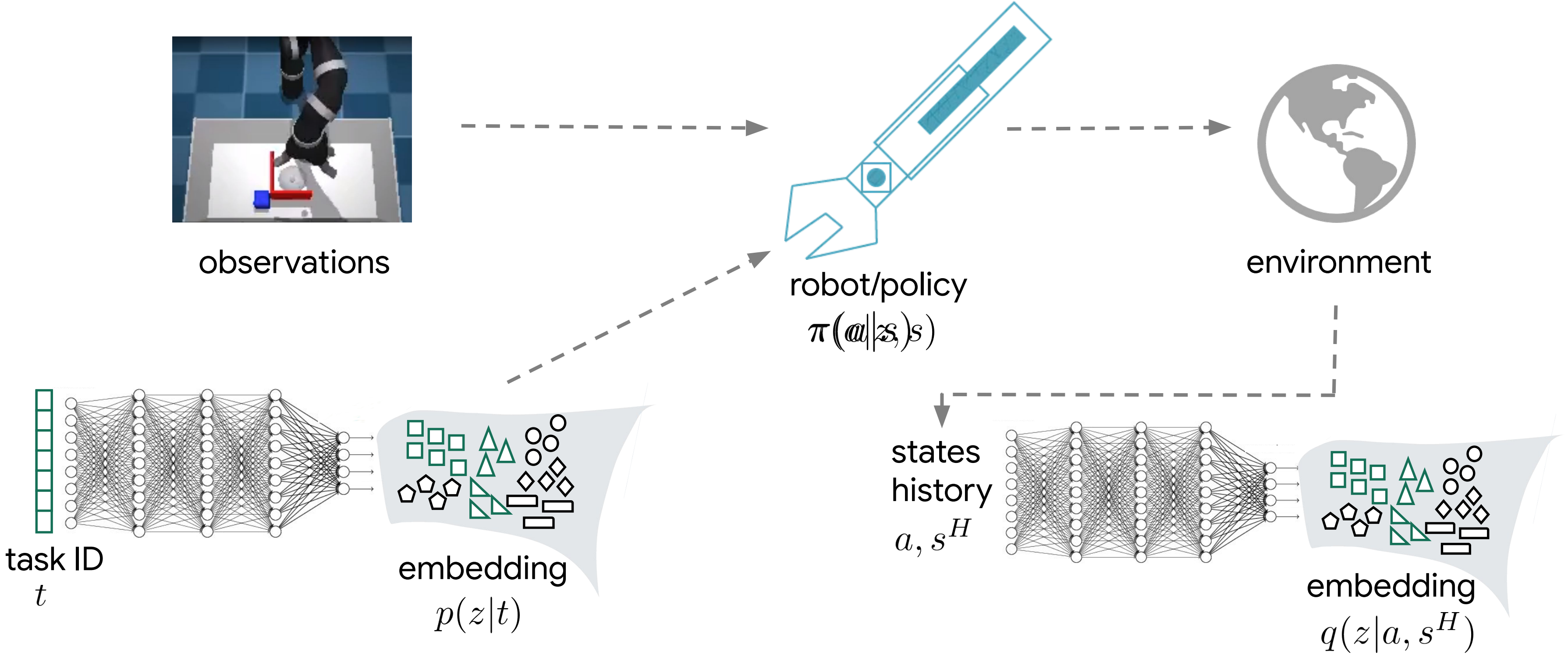
embedding can represen[redacted]for every task:

[redacted]



Robot skill embeddings

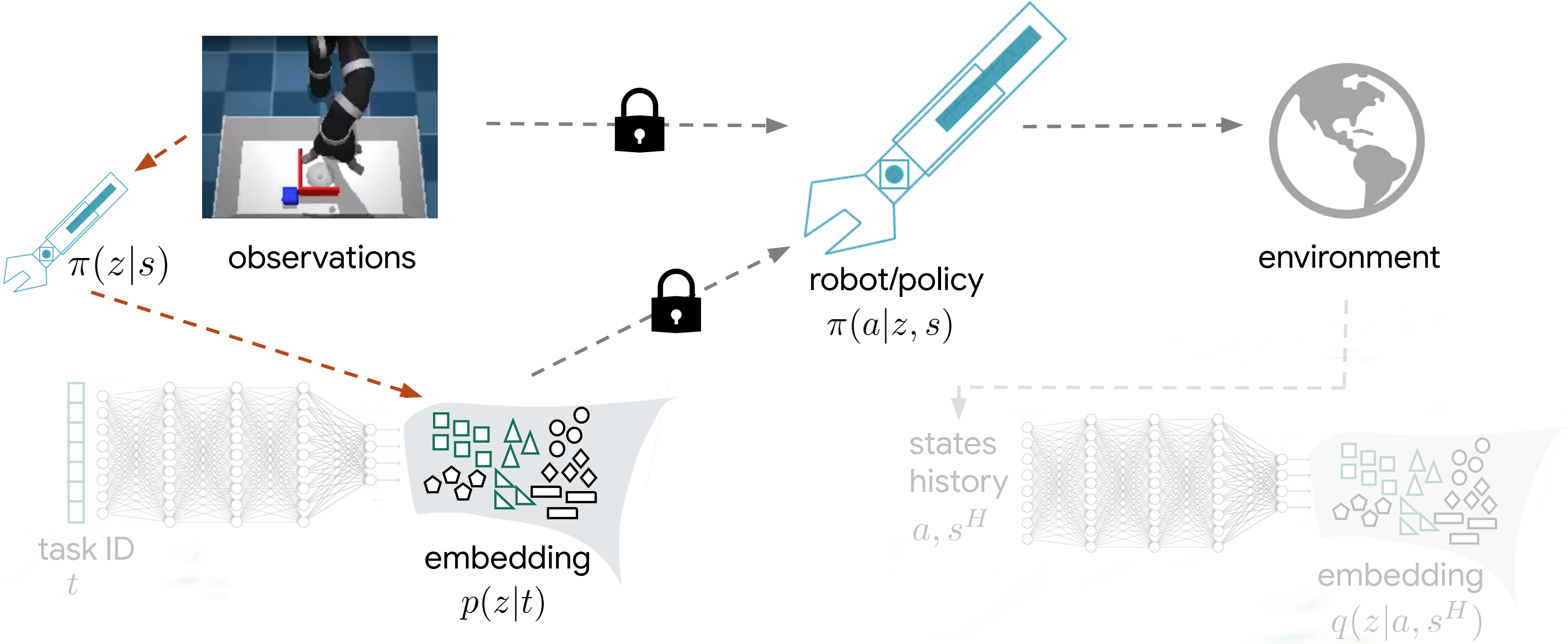
training:



[Learning an embedding space for reusable robotic skills, Hausman et al.]

Robot skill embeddings

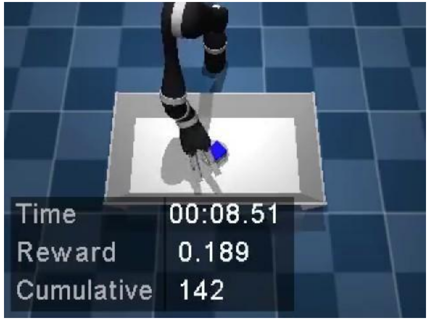
test:



[Learning an embedding space for reusable robotic skills, Hausman et al.]

Robot skill embeddings - multi-task learning

skills: push



lift

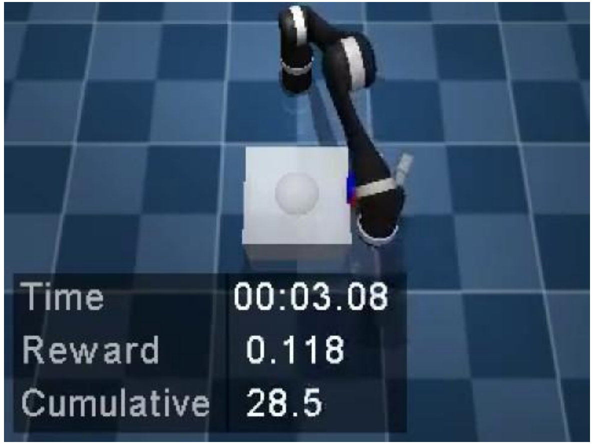
transfer: push around a wall

Robot skill embeddings - multi-task learning

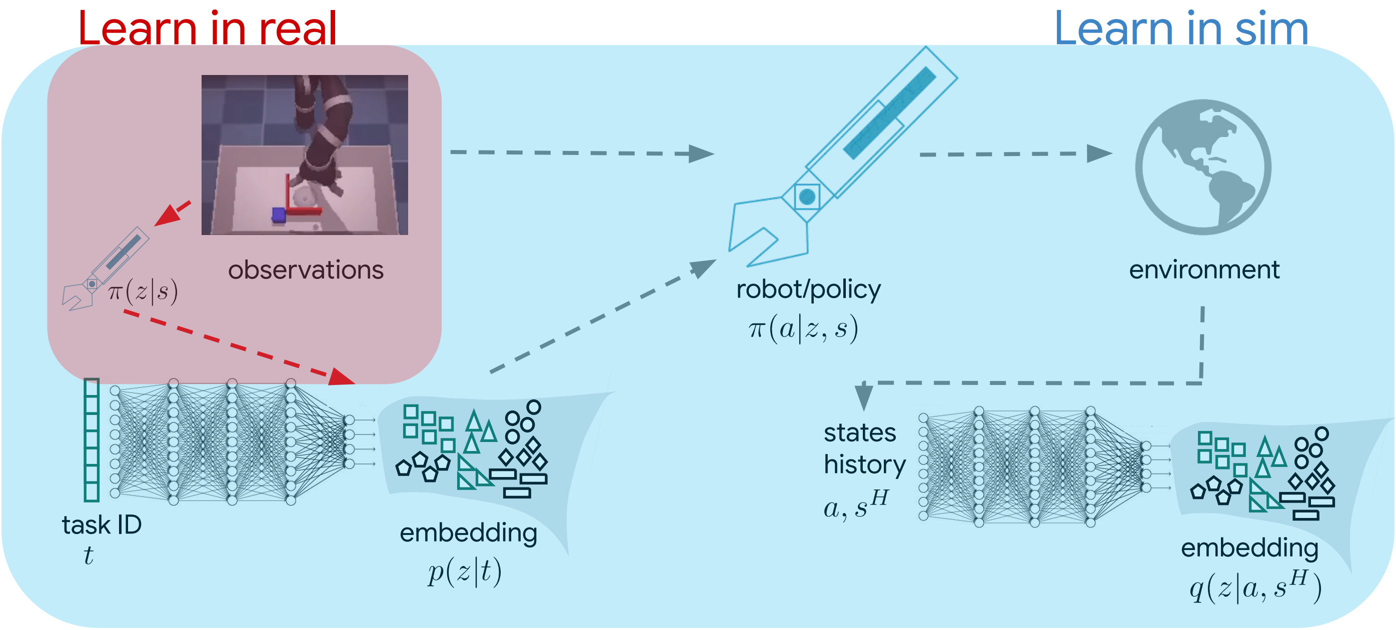
skills: lift on a rail

push on a table

transfer: lift and then push



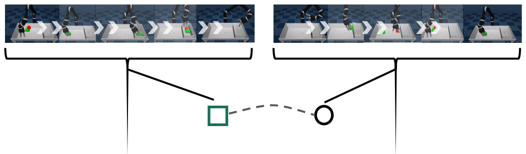
Robot skill embeddings - sim2real transfer



[Scaling simulation-to-real transfer by learning composable robot skills
Julian, et al., 2018]
[Zero-Shot Skill Composition and Simulation-to-Real Transfer by
Learning Task Representations, He, at al., 2018]

Robot skill embeddings - sim2real transfer

Multi-task deep RL



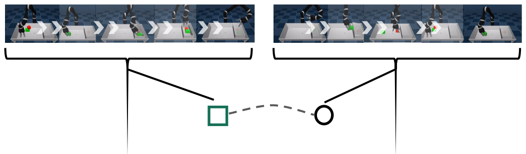
Skill Representation and Reusability

- task specifications, what constitutes a task, how to represent a skill?
- reuse of already-learned skills

Supervision and Efficiency

- multiple skills - multiple pains: rewards, setups, etc.
- efficient sequencing of skills at test time

Multi-task deep RL



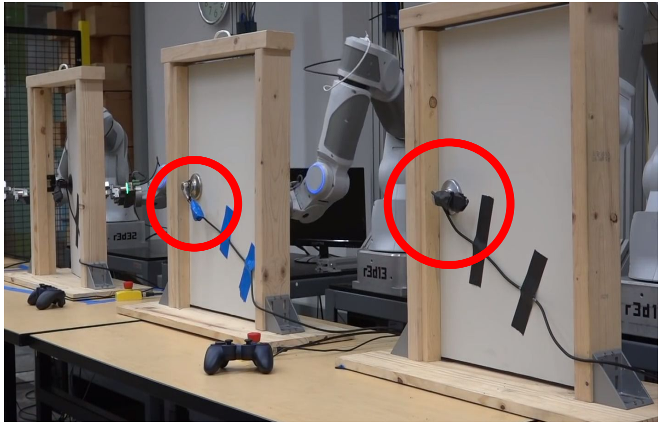
Skill Representation and Reusability

- task specifications, what constitutes a task, how to represent a skill?
- reuse of already-learned skills

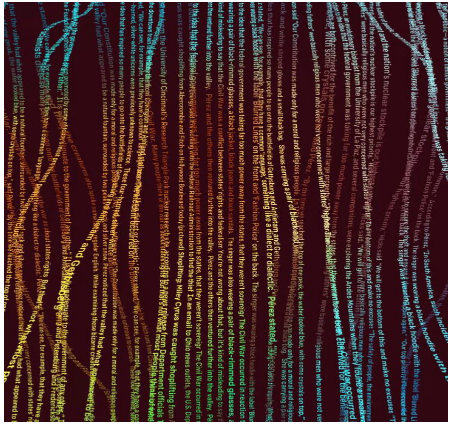
Supervision and Efficiency

- multiple skills - multiple pains: rewards, setups, etc.
- efficient sequencing of skills at test time

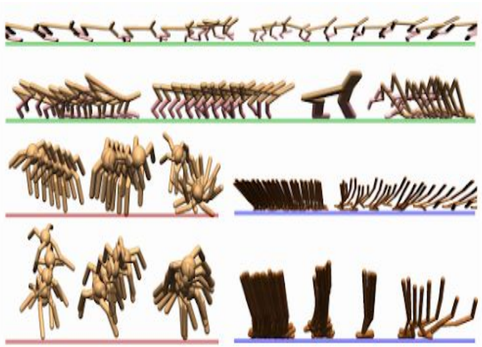
Supervision



[Collective robot reinforcement learning with distributed asynchronous guided policy search, Yahya et al. 2017]

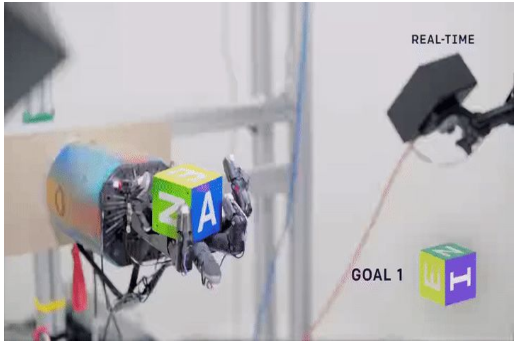


[Better Language Models and Their Implications, OpenAI Blog, 2019]



[Diversity is all you need, Learning Diverse Skills without a Reward Function, Eysenbach, 2018]

Efficiency



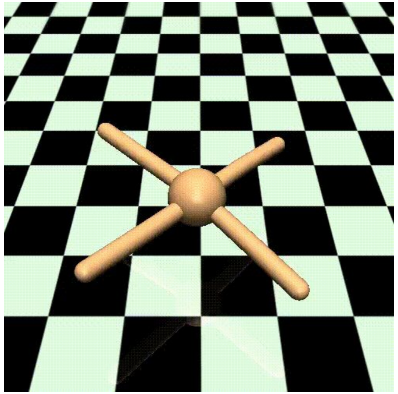
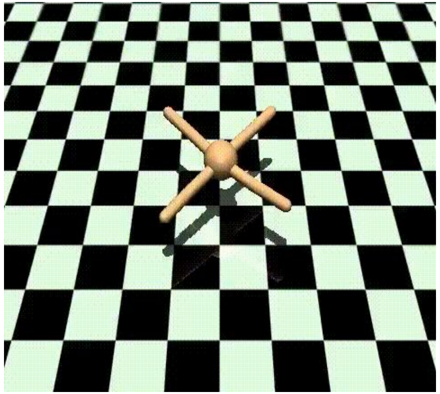
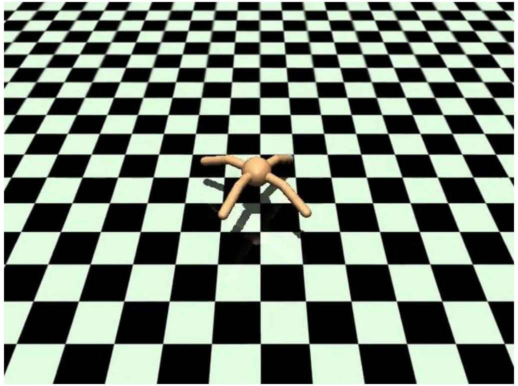
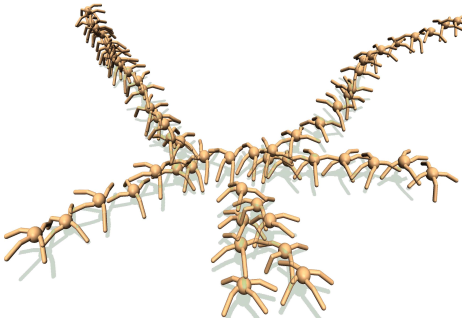
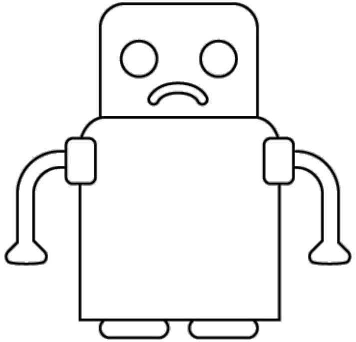
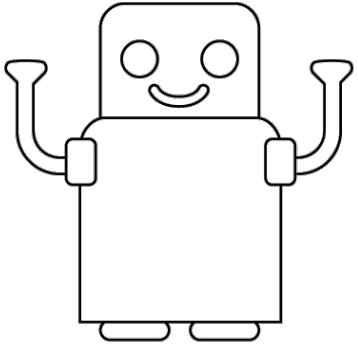
~100 years of experience

~1 hour of experience

[Learning Dexterous In-Hand Manipulation, OpenAI et al. 2018]

[SOLAR: Deep Structured Representations for Model-Based Reinforcement Learning, Zhang et al. 2019]

Global vs Behavior-Specific Dynamics Models

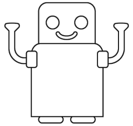
Dynamics-Aware Unsupervised Discovery of Skills (DADS)

main idea: use empowerment to simultaneously optimize for skills and their specific dynamics

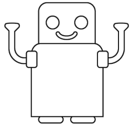
mutual information objective:

$$I(s'; z|s) \geq \mathbb{E}_s \mathbb{E}_z \mathbb{E}_{p(s'|s,z)}[\log \frac{q_\phi(s'|s,z)}{p(s'|s)}]$$
$$\approx \mathbb{E}_s \mathbb{E}_z \mathbb{E}_{p(s'|s,z)}[\log \frac{q_\phi(s'|s,z)}{\sum_{i=1}^{L} q_\phi(s'|s,z_i)} + \log L]$$

Predictability

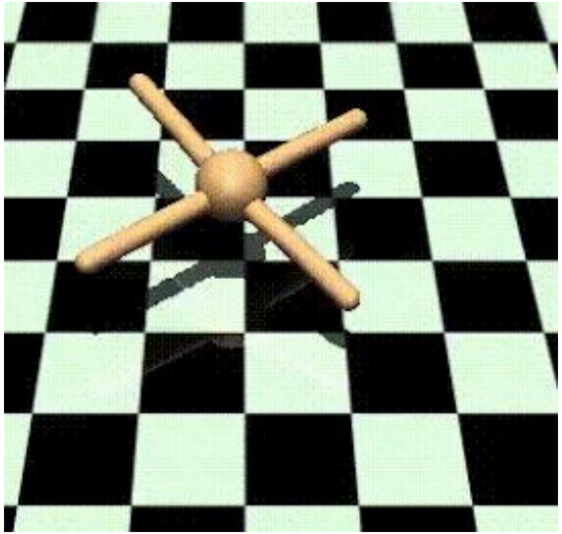


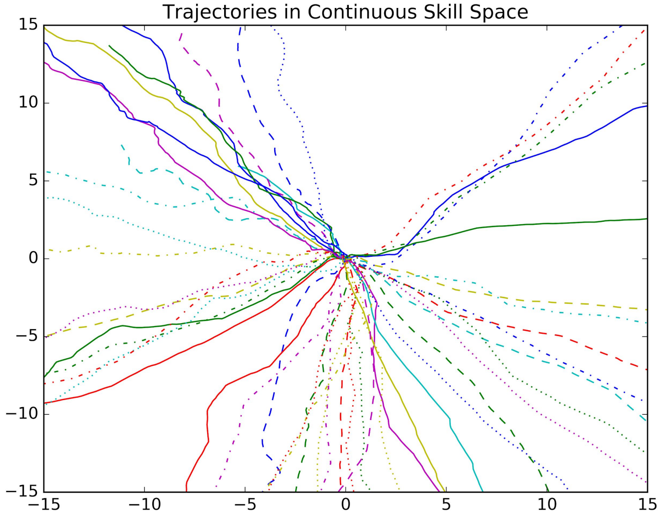
Diversity



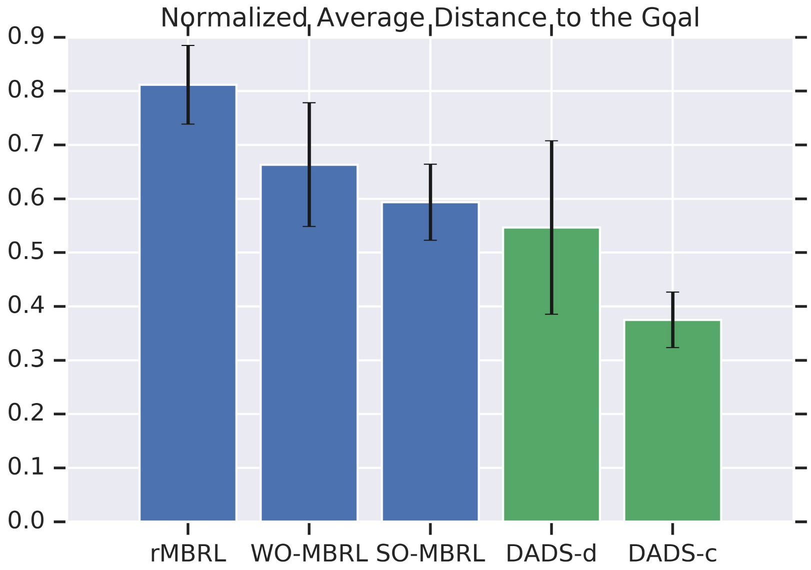
[Dynamics-Aware Unsupervised Discovery of Skills, Sharma, et al. 2018]

Dynamics-Aware Unsupervised Discovery of Skills (DADS)

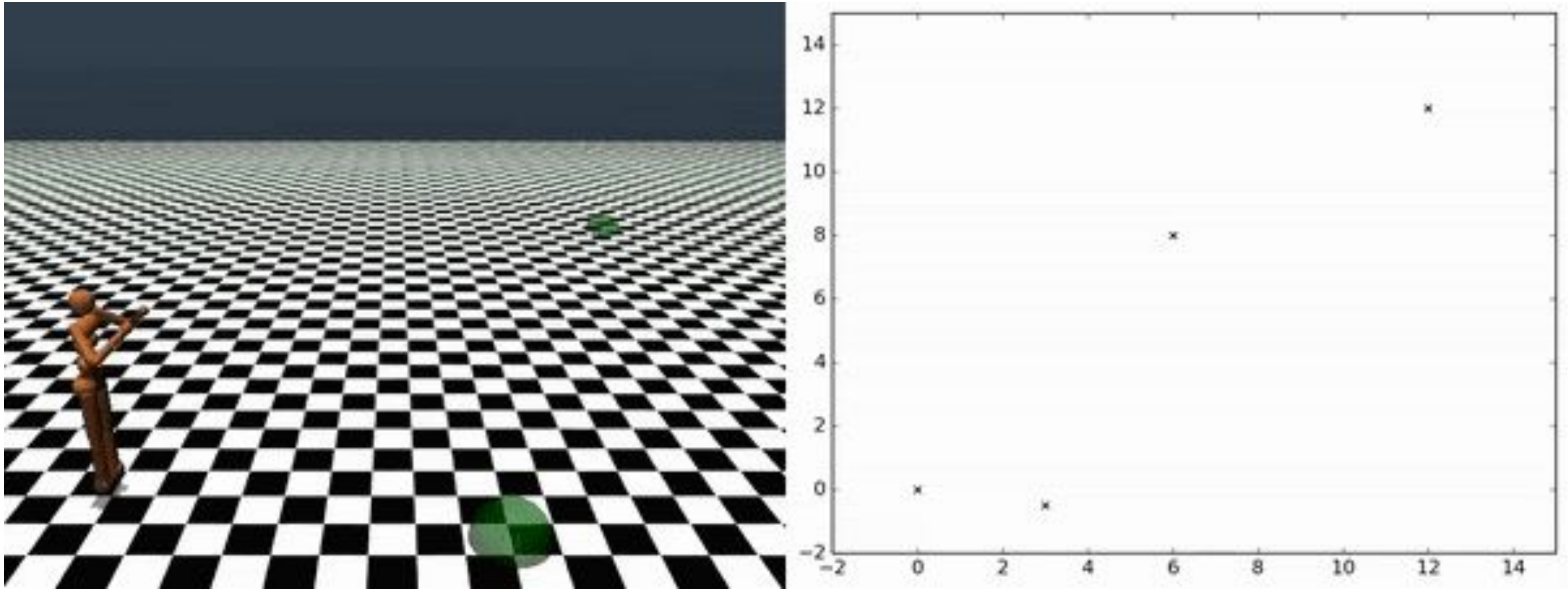




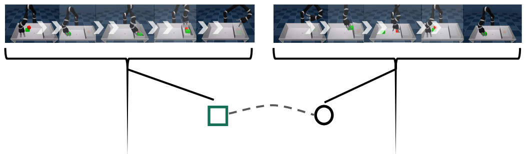
Dynamics-Aware Unsupervised Discovery of Skills (DADS)



Dynamics-Aware Unsupervised Discovery of Skills (DADS)



Multi-task deep RL



Skill Representation and Reusability

- task specifications, what constitutes a task, how to represent a skill?
- reuse of already-learned skills

Supervision and Efficiency

- multiple skills - multiple pains: rewards, setups, etc.
- efficient sequencing of skills at test time

Multi-task deep RL

Challenges

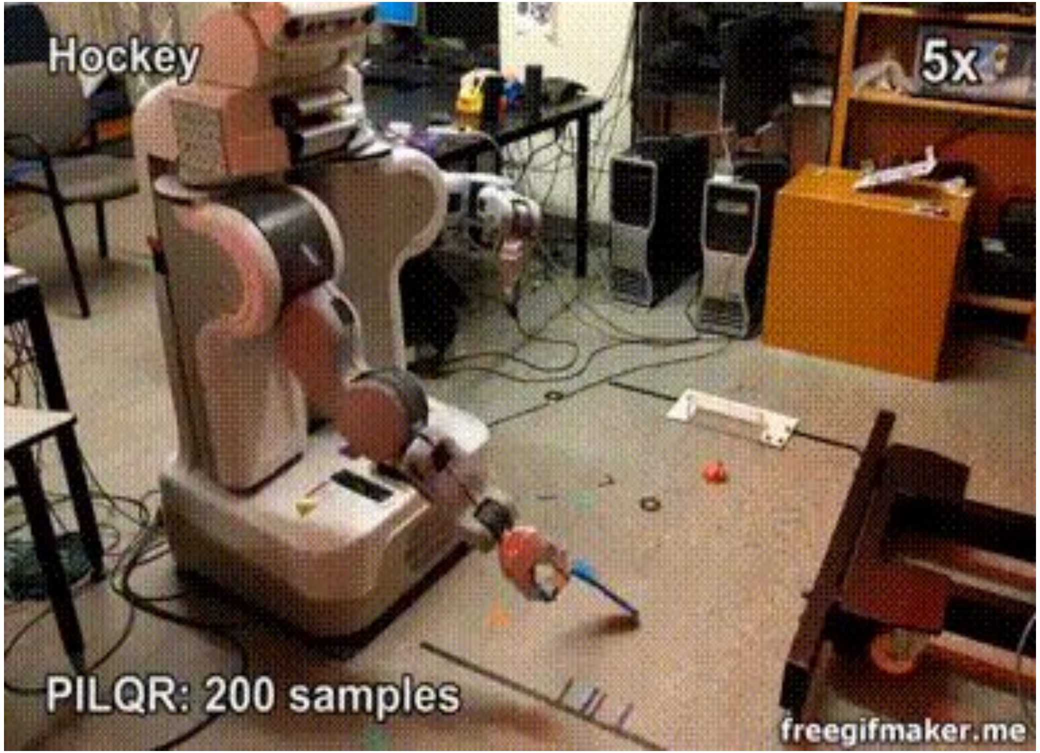
- task specifications, what constitutes a task, how to represent a skill?
- reuse of already-learned skills
- optimization of multiple tasks (conflicting gradients, gradient magnitudes)
- data imbalance issues (harder easier tasks, good exploration in all of them)
- multiple skills - multiple pains: rewards, setups, etc.
- efficient sequencing of skills at test time

Future Work

Multi-task deep RL

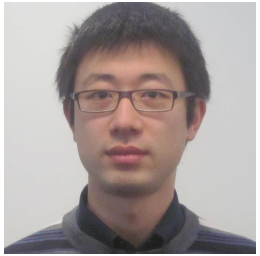
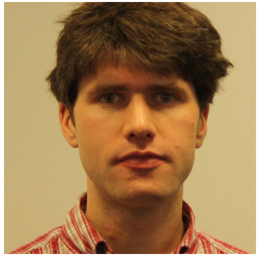
Challenges

- task specifications, what constitutes a task, how to represent a skill?
- reuse of already-learned skills
- optimization of multiple tasks (conflicting gradients, gradient magnitudes)
- data imbalance issues (harder easier tasks, good exploration in all of them)
- multiple skills - multiple pains: rewards, setups, etc.
- efficient sequencing of skills at test time
- and many more...


Hockey
5x
PILQR: 200 samples
freegifmaker.me

Learning an Embedding Space for Transferable Robot Skills, ICLR 2018

K. Hausman, T. Springenberg, Z. Wang, N. Heess, M. Riedmiller

Dynamics-Aware Unsupervised Discovery of Skills, NeurIPS 2019 Submission

A. Sharma, S. Gu, S. Levine, V. Kumar, K. Hausman