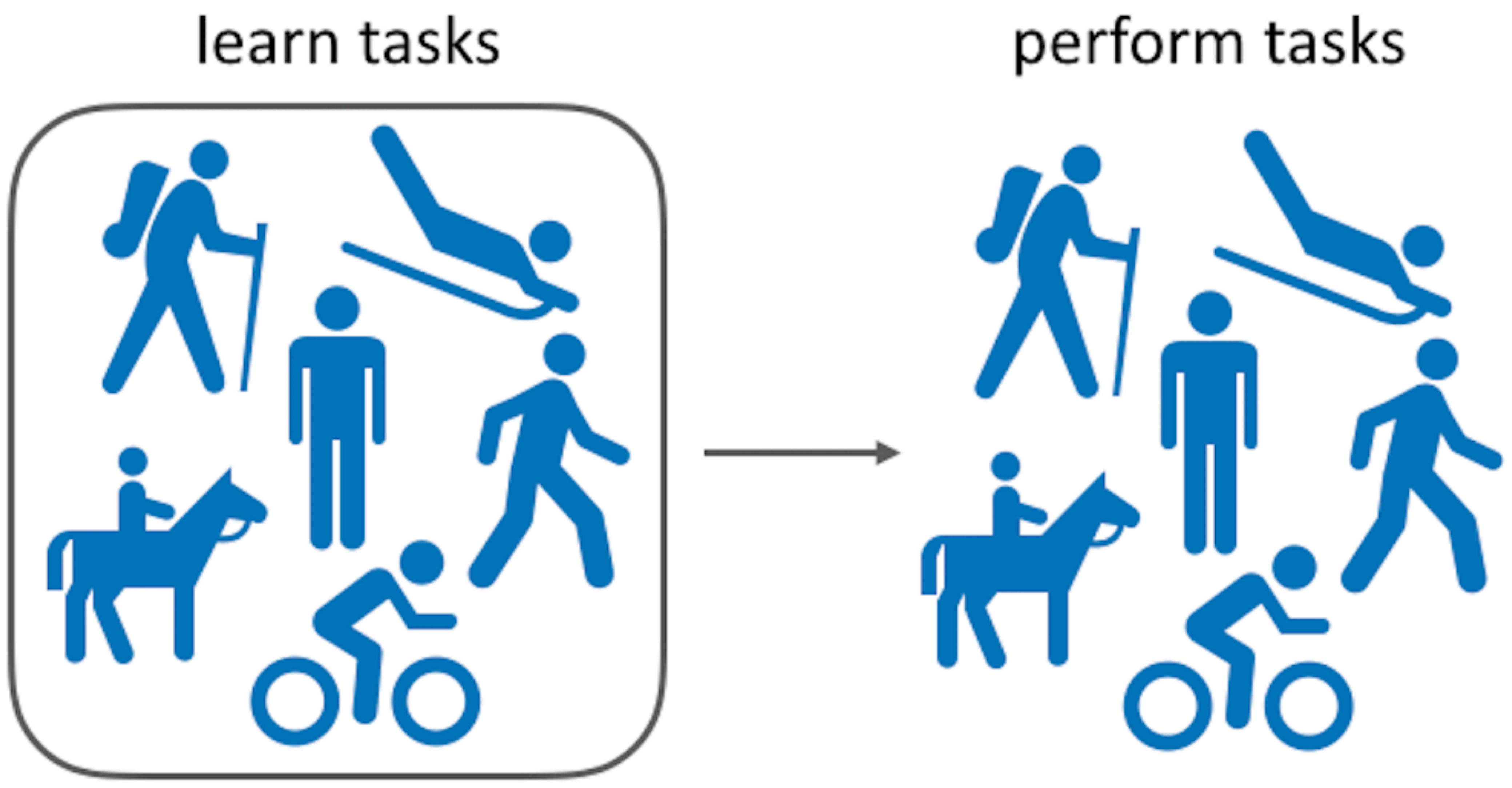
Meta-World: A Benchmark and Evaluation for Multi-Task and Meta Reinforcement Learning

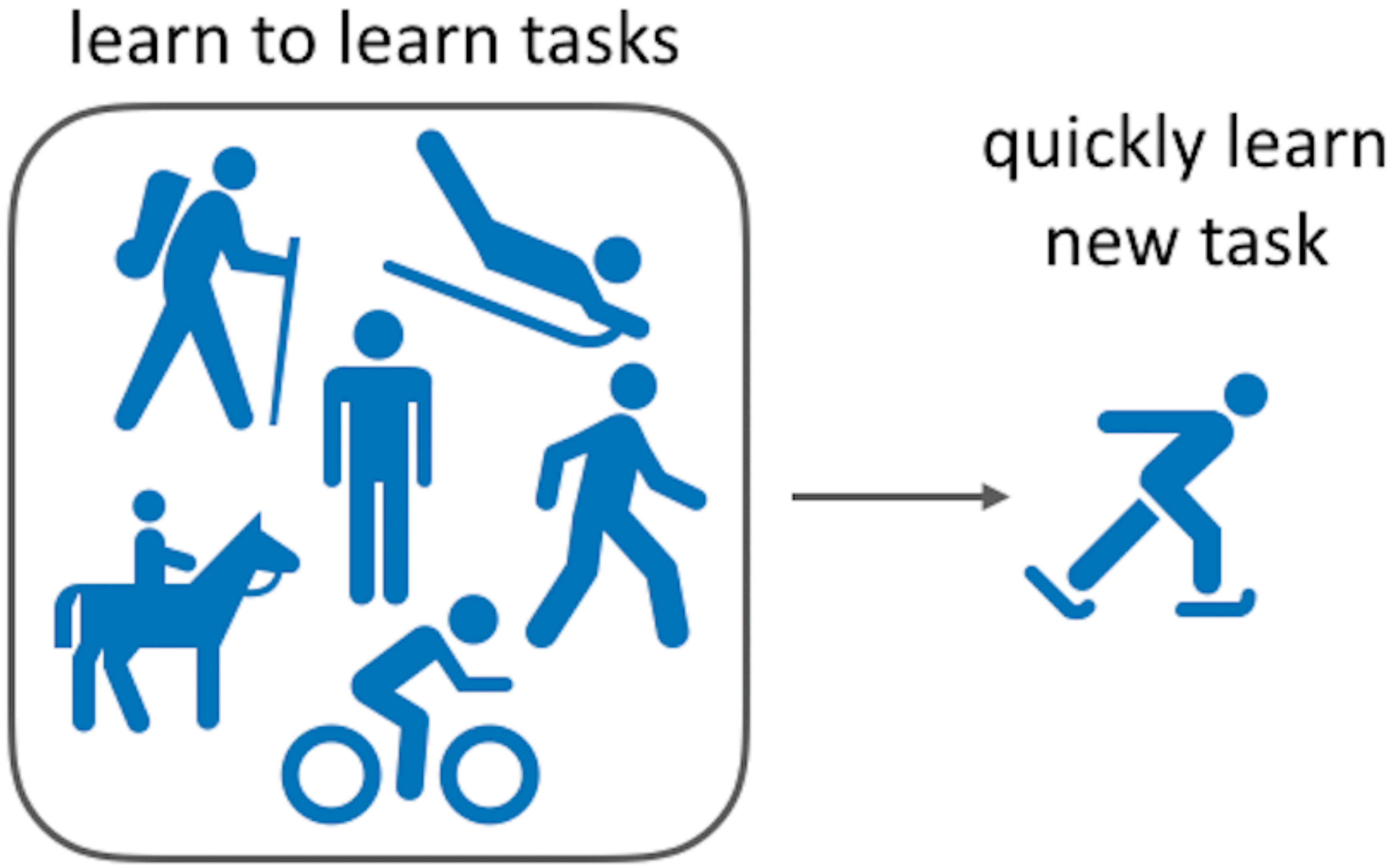
Tianhe Yu*, Deirdre Quillen*, Zhanpeng He*, Ryan Julian,
Karol Hausman, Chelsea Finn, Sergey Levine

CoRL 2019

Multi-Task Reinforcement Learning



Meta Reinforcement Learning



Multi-Task Reinforcement Learning

Promise: learn a single policy that can solve multiple tasks more efficiently than learning the tasks individually

Multi-Task Reinforcement Learning

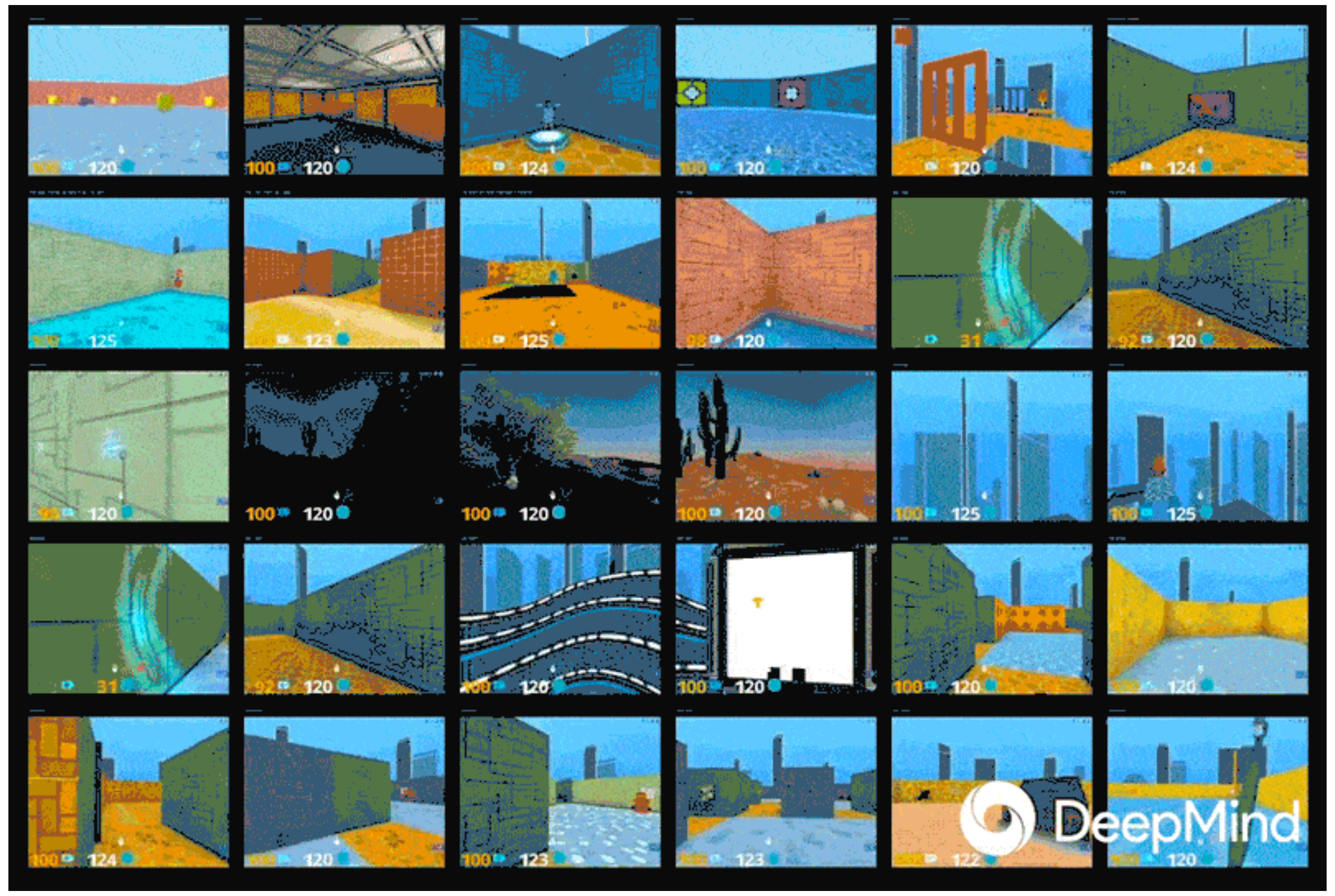
Promise: learn a single policy that can solve multiple tasks more efficiently than learning the tasks individually

Current Multi-Task RL benchmarks

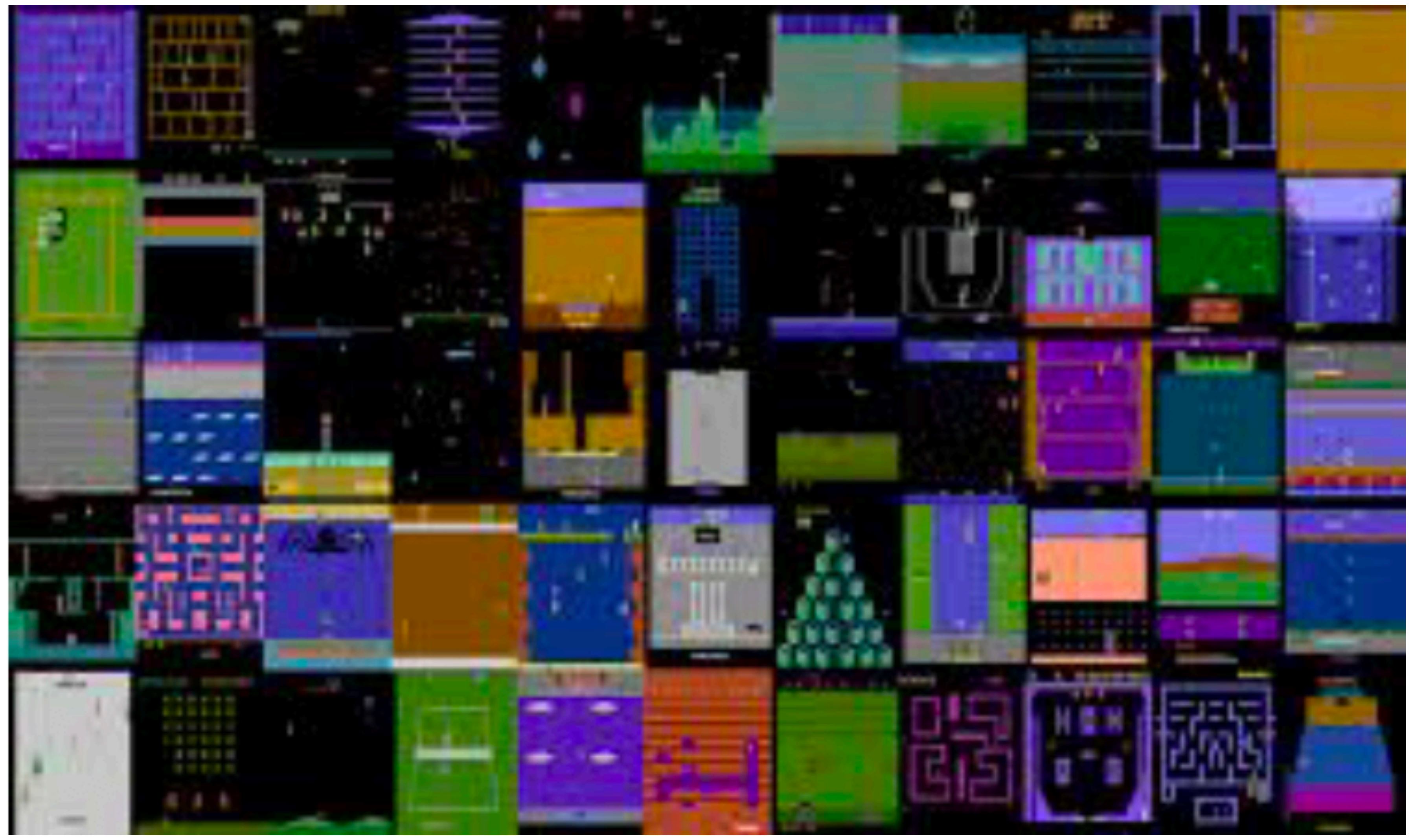
Multi-Task Reinforcement Learning

Promise: learn a single policy that can solve multiple tasks more efficiently than learning the tasks individually

Current Multi-Task RL benchmarks



DM Lab



Atari

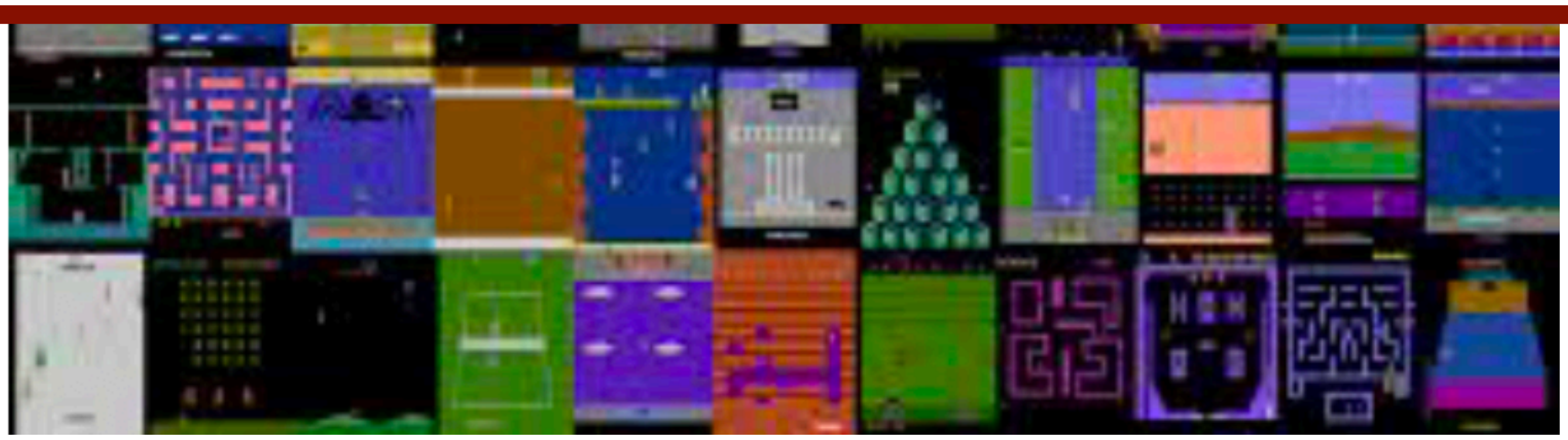
Multi-Task Reinforcement Learning

Promise: learn a single policy that can solve multiple tasks more efficiently than learning the tasks individually

Current Multi-Task RL benchmarks

- Limited to game settings and lack of realistic use cases
- Little efficiency to be gained on disjoint games

DM Lab


Atari

Meta Reinforcement Learning

Promise: efficiently acquire new tasks by leveraging experiences from past tasks

Meta Reinforcement Learning

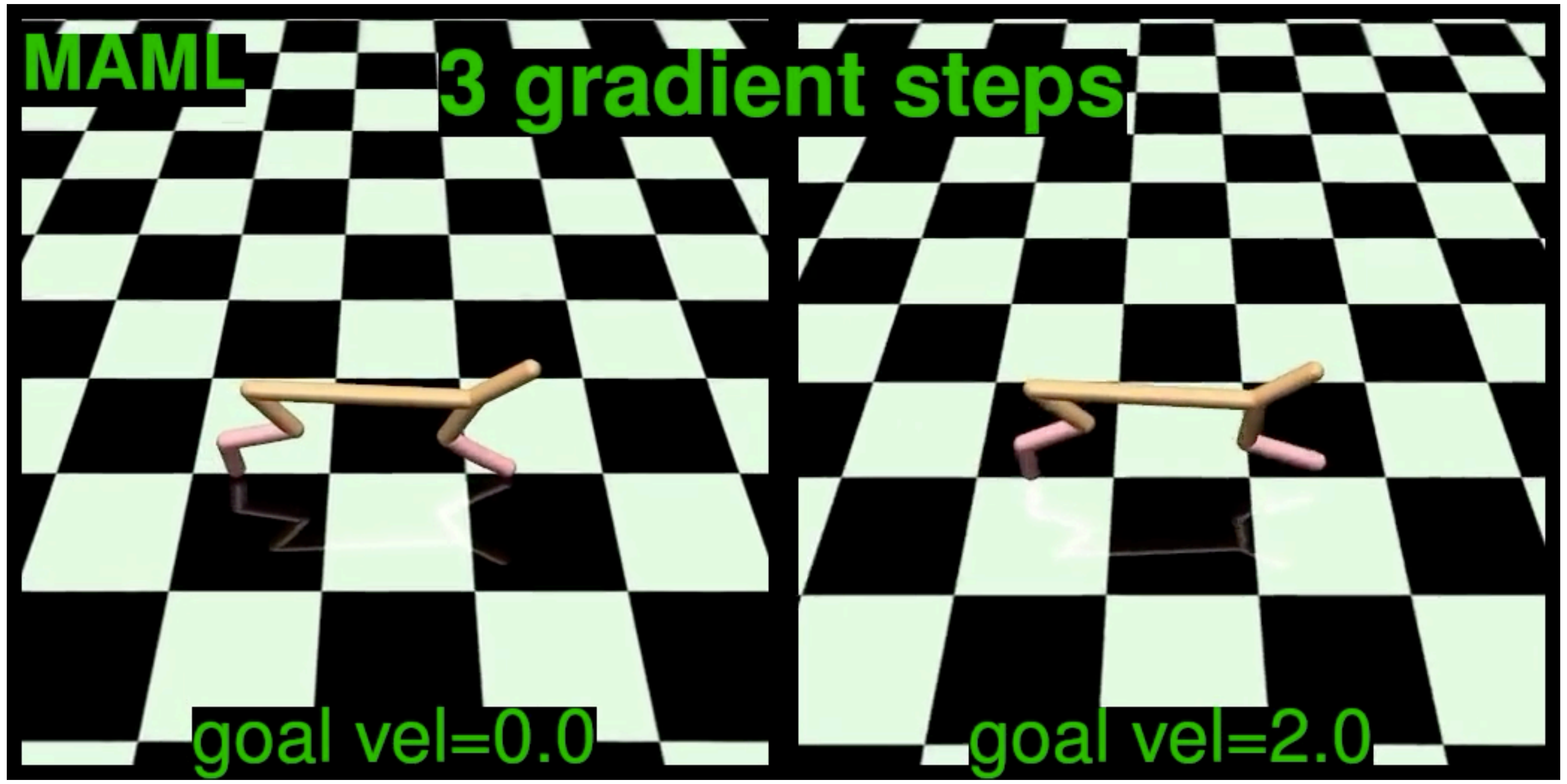
Promise: efficiently acquire new tasks by leveraging experiences from past tasks

Current Meta-RL benchmarks

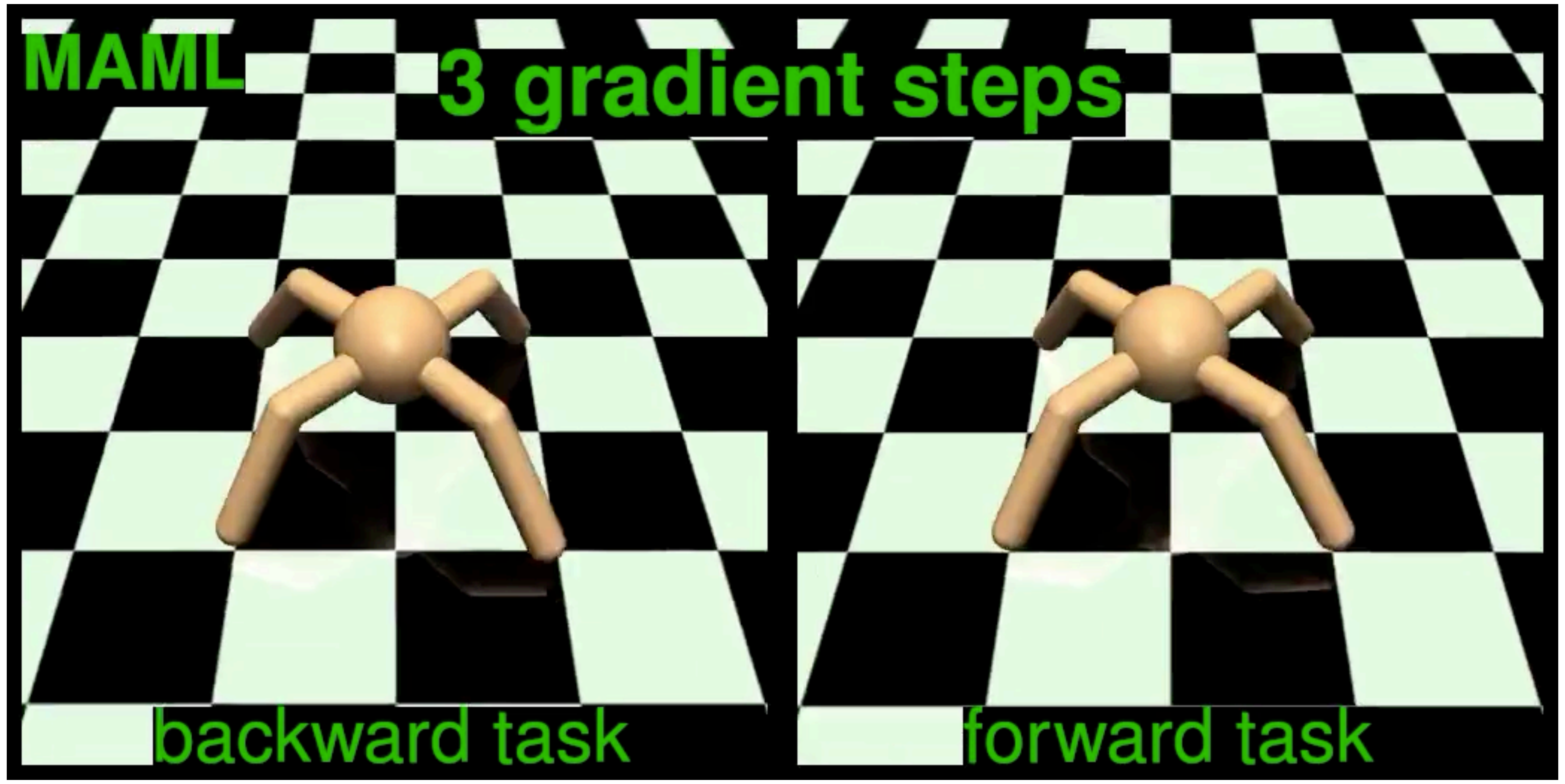
Meta Reinforcement Learning

Promise: efficiently acquire new tasks by leveraging experiences from past tasks

Current Meta-RL benchmarks



Half-Cheetah Velocity



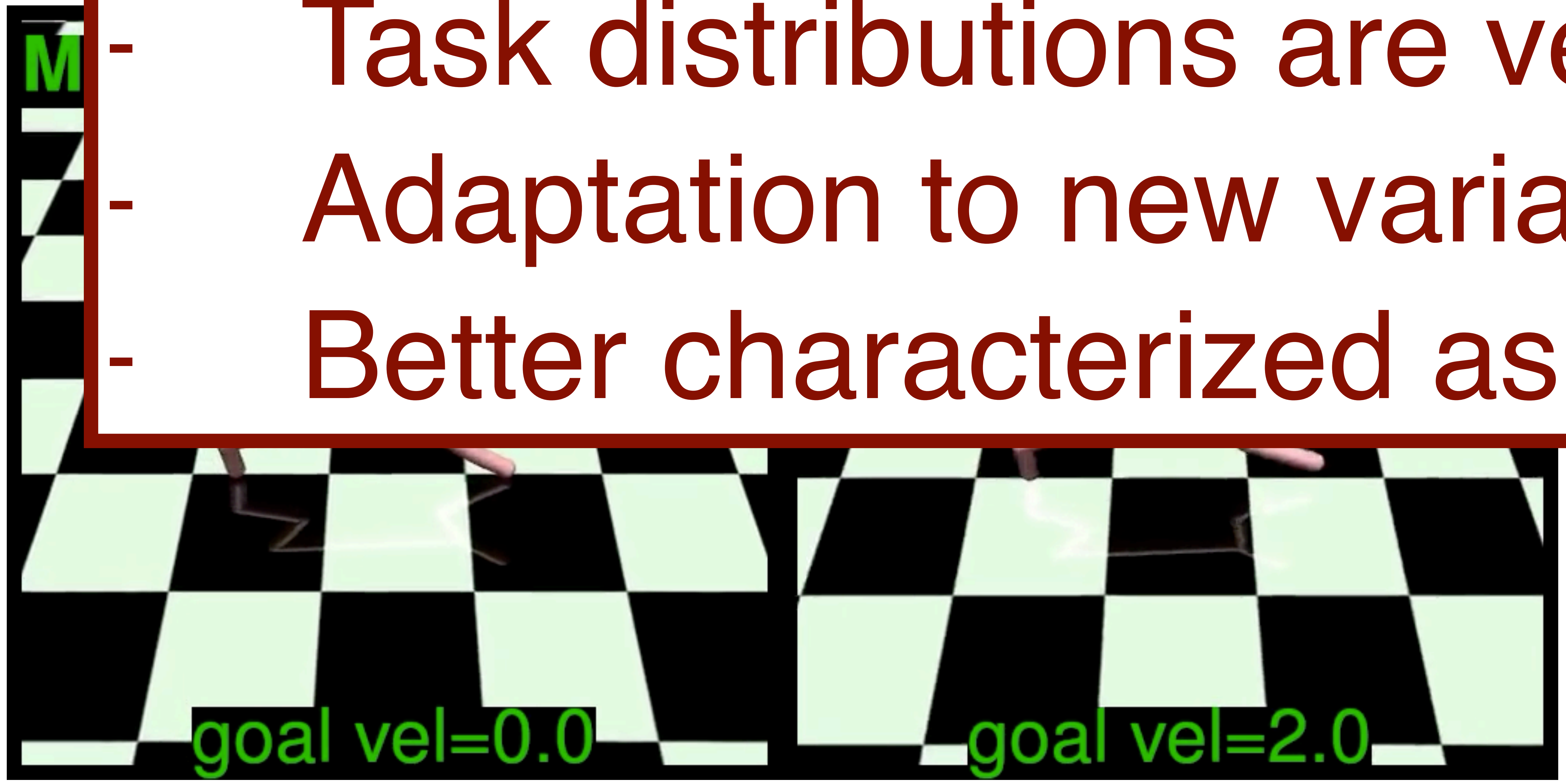
Ant Direction

Meta Reinforcement Learning

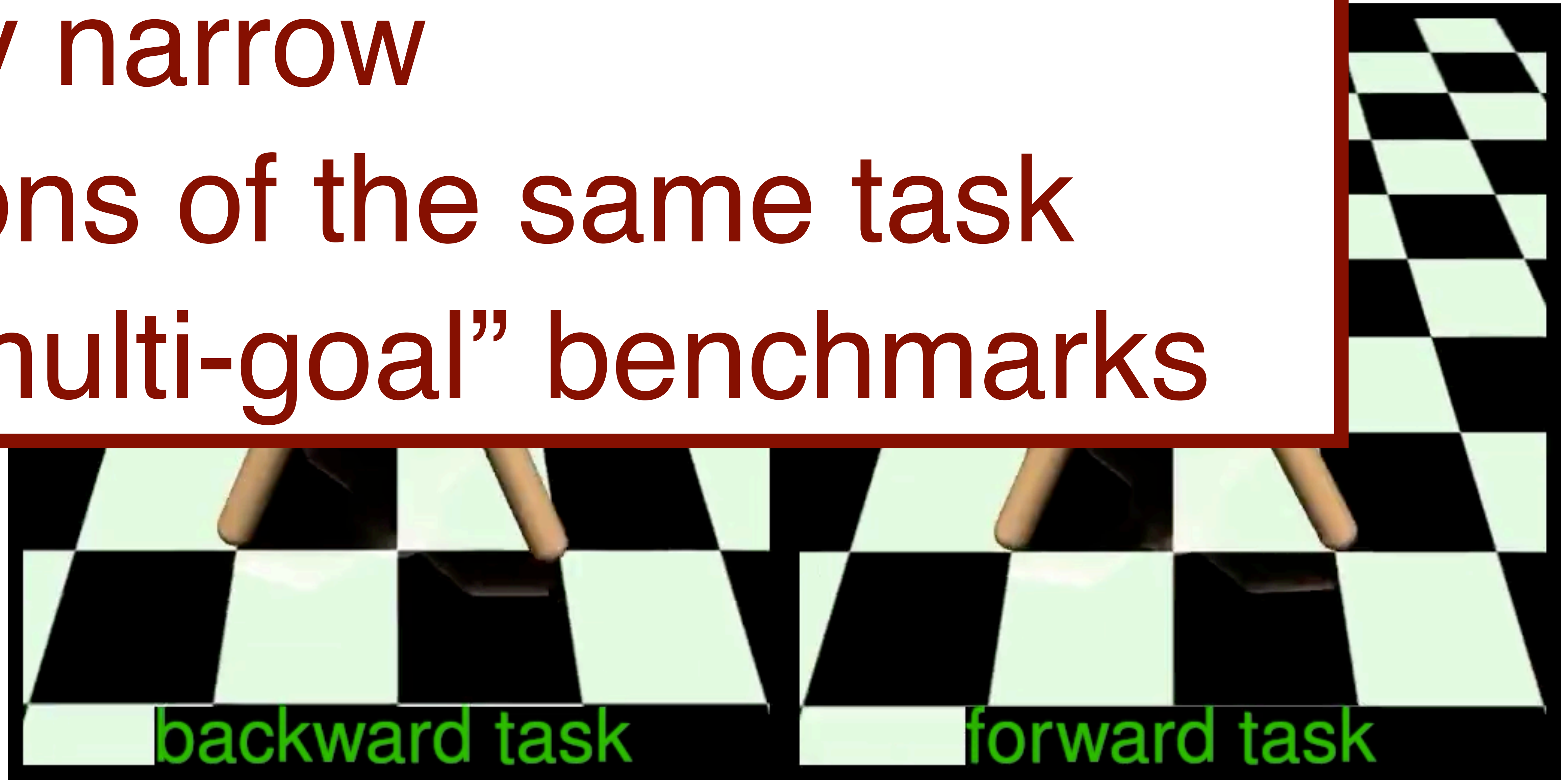
Promise: efficiently acquire new tasks by leveraging experiences from past tasks

Current Meta-RL benchmarks

- Task distributions are very narrow
- Adaptation to new variations of the same task
- Better characterized as “multi-goal” benchmarks



Half-Cheetah Velocity



Ant Direction

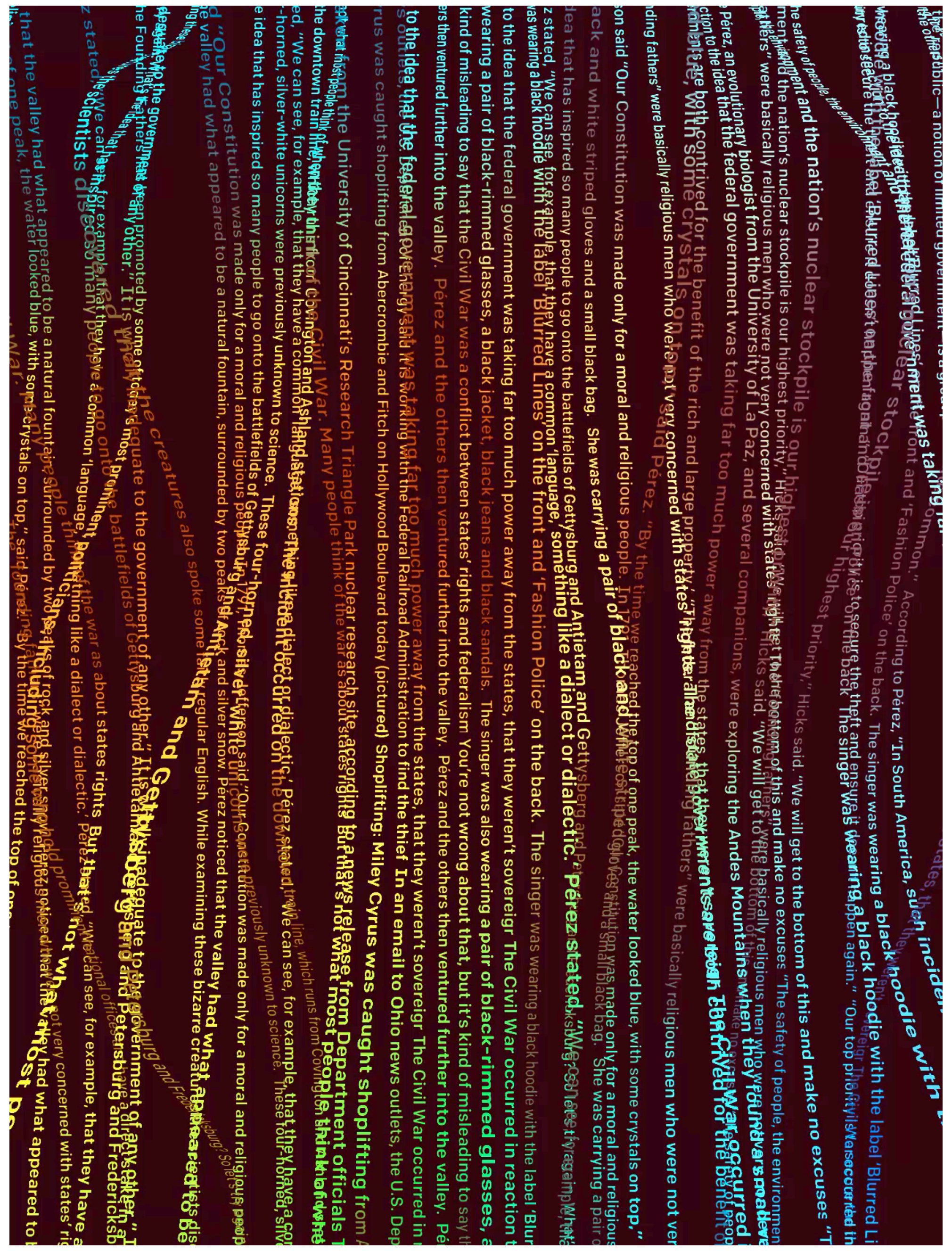
Goal: enable meta-RL to generalize to new distinct skills and evaluate the generalization performance

Goal: enable meta-RL to generalize to new distinct skills and evaluate the generalization performance

Imagenet, Russakovsky et al. '14

Goal: enable meta-RL to generalize to new distinct skills and evaluate the generalization performance

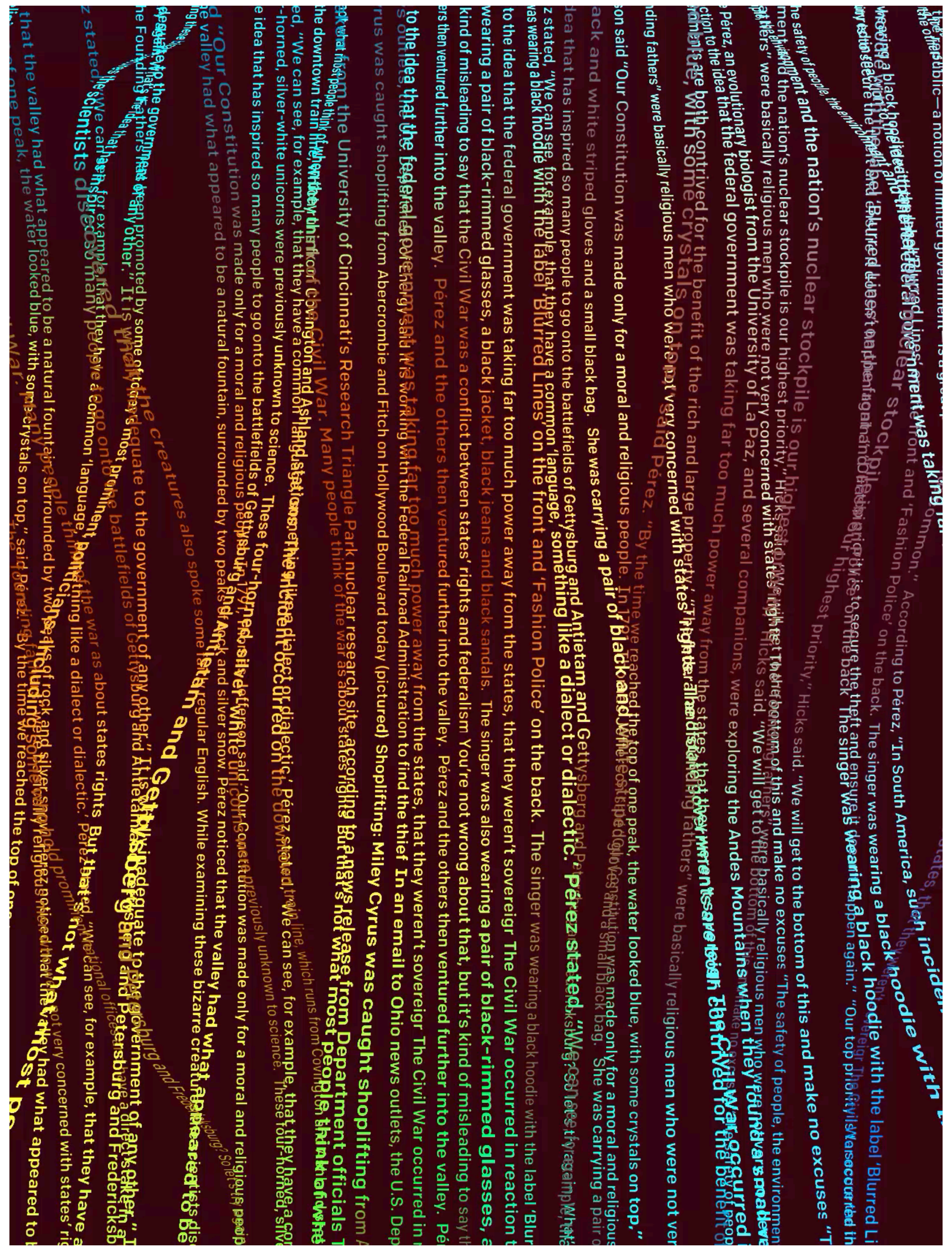
Imagenet, Russakovsky et al. ‘14



GPT-2, Radford et al. ‘19

Goal: enable meta-RL to generalize to new distinct skills and evaluate the generalization performance

Imagenet, Russakovsky et al. ‘14



GPT-2, Radford et al. ‘19

Large, diverse task set → Generalization to new tasks

Meta-World Task Suite

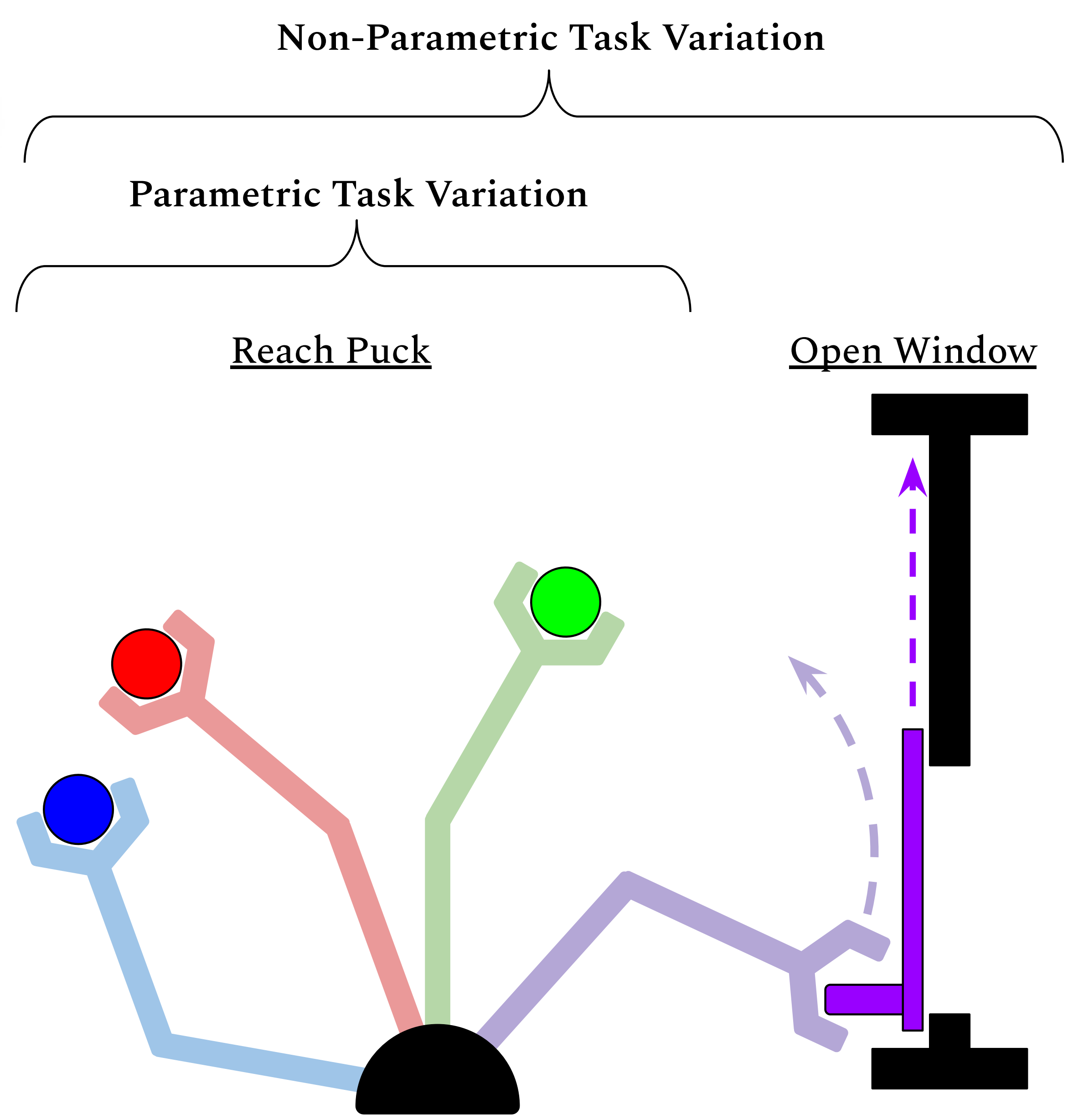
Meta-World Task Suite

What's different: A new multi-task and meta-RL benchmark with a wide range of tasks to study how meta-RL accelerates acquisition of new tasks

- 50 qualitatively different manipulation tasks using a simulated Sawyer robot

Meta-World Task Suite

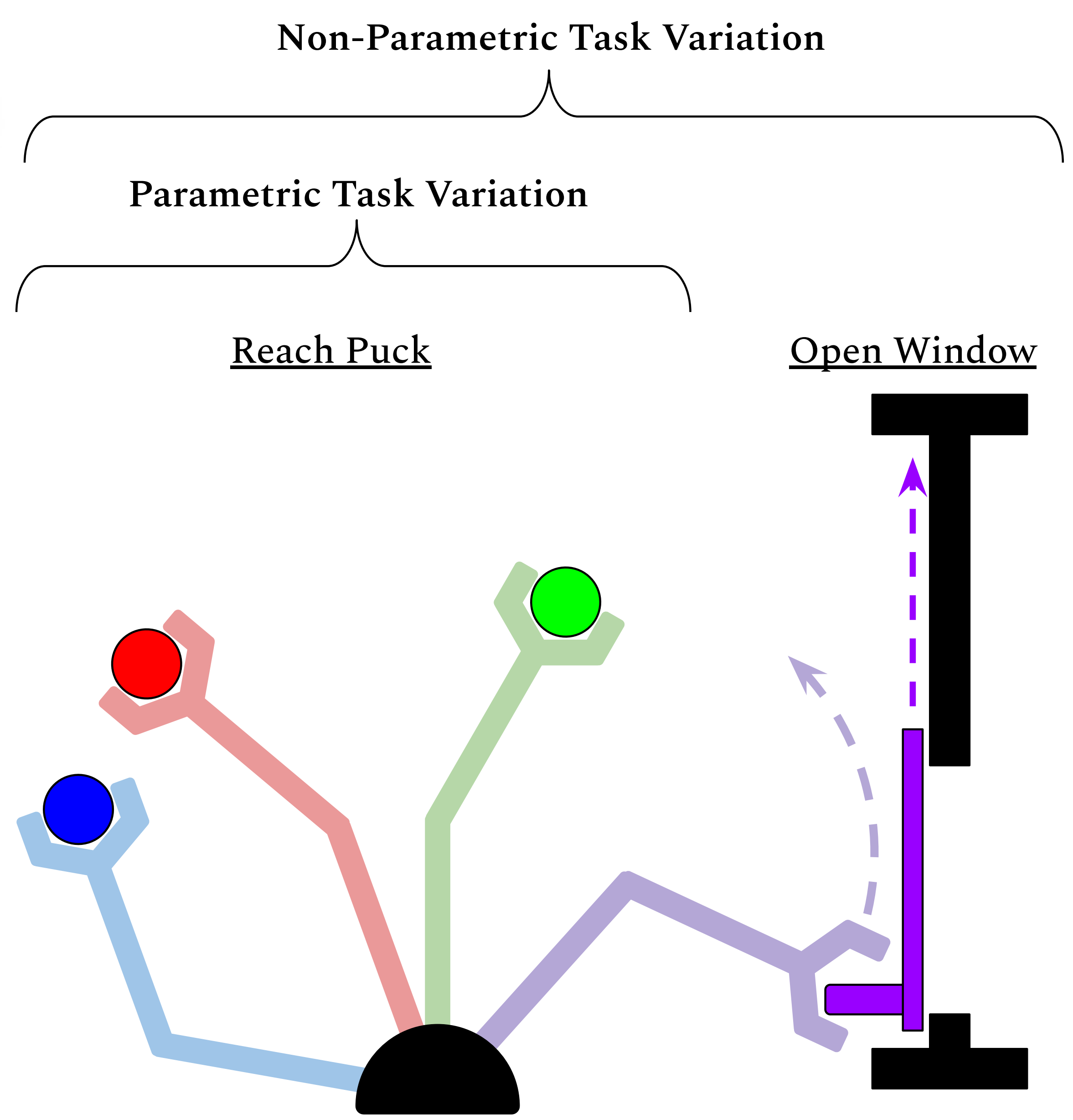
What's different: A new multi-task and meta-RL benchmark with a wide range of tasks to study how meta-RL accelerates acquisition of new tasks



- 50 qualitatively different manipulation tasks using a simulated Sawyer robot
- Parametric variability with tasks, e.g. goal positions

Meta-World Task Suite

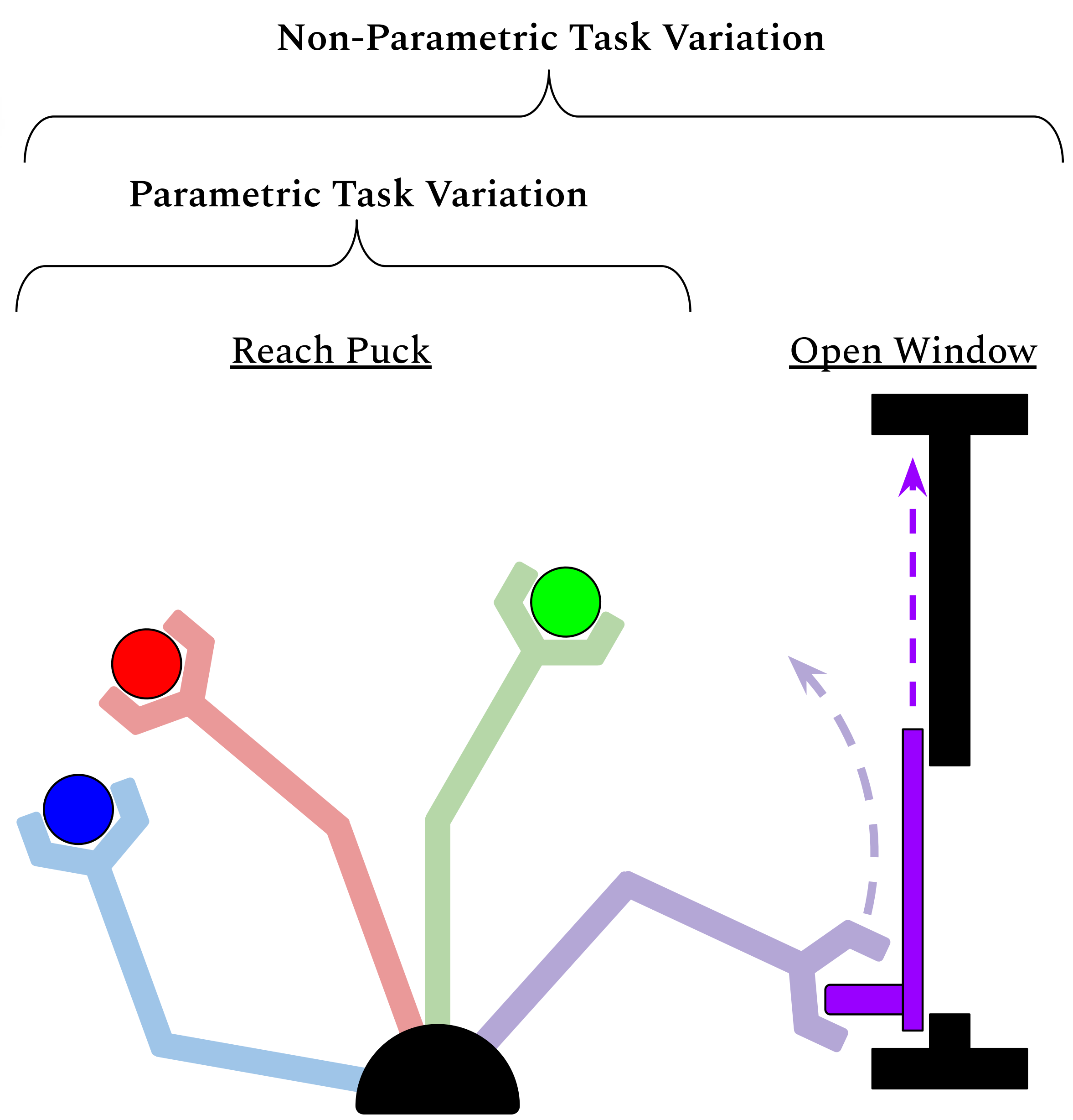
What's different: A new multi-task and meta-RL benchmark with a wide range of tasks to study how meta-RL accelerates acquisition of new tasks



- 50 qualitatively different manipulation tasks using a simulated Sawyer robot
- Parametric variability with tasks, e.g. goal positions
- Unified observation and action space

Meta-World Task Suite

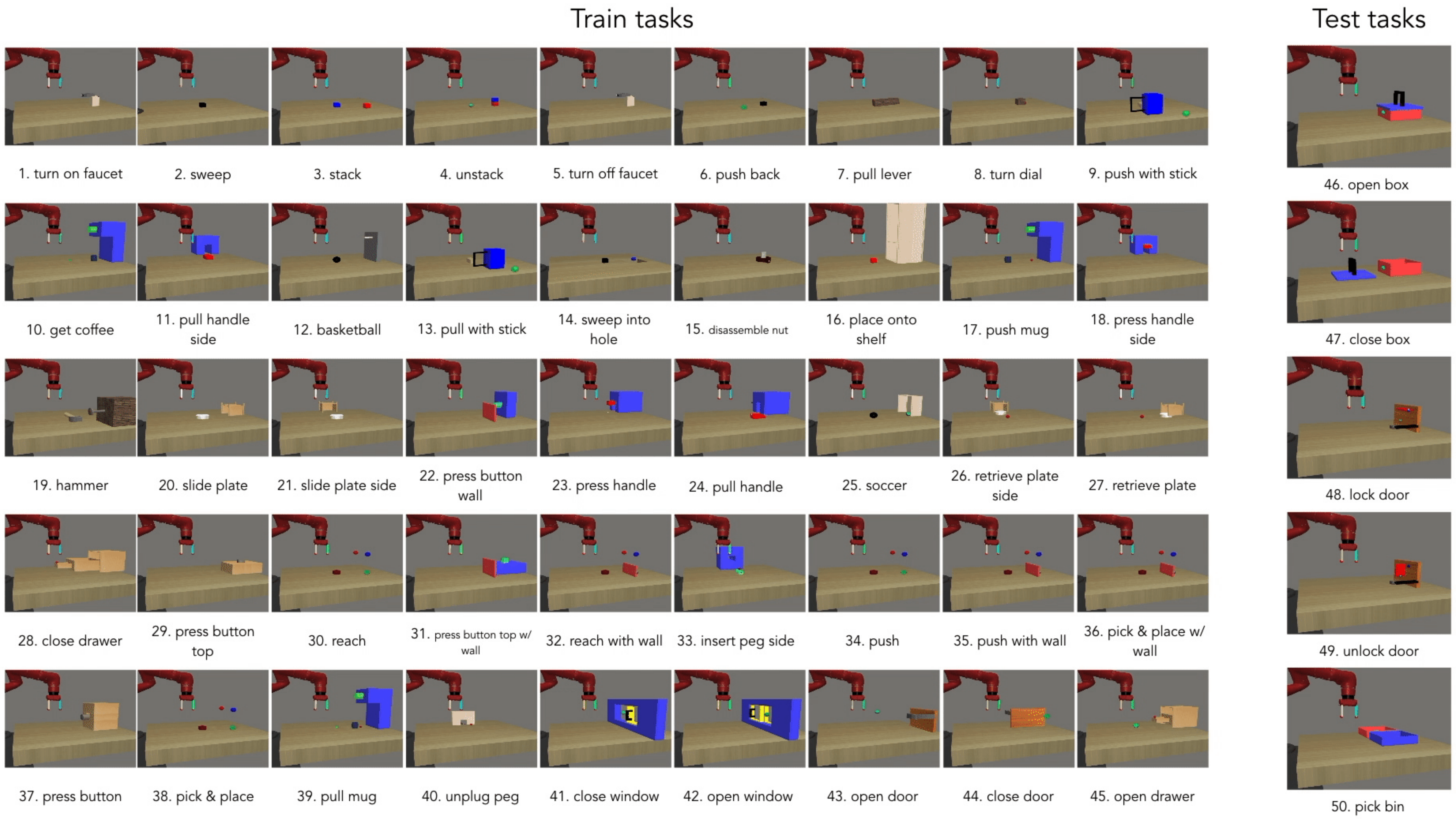
What's different: A new multi-task and meta-RL benchmark with a wide range of tasks to study how meta-RL accelerates acquisition of new tasks

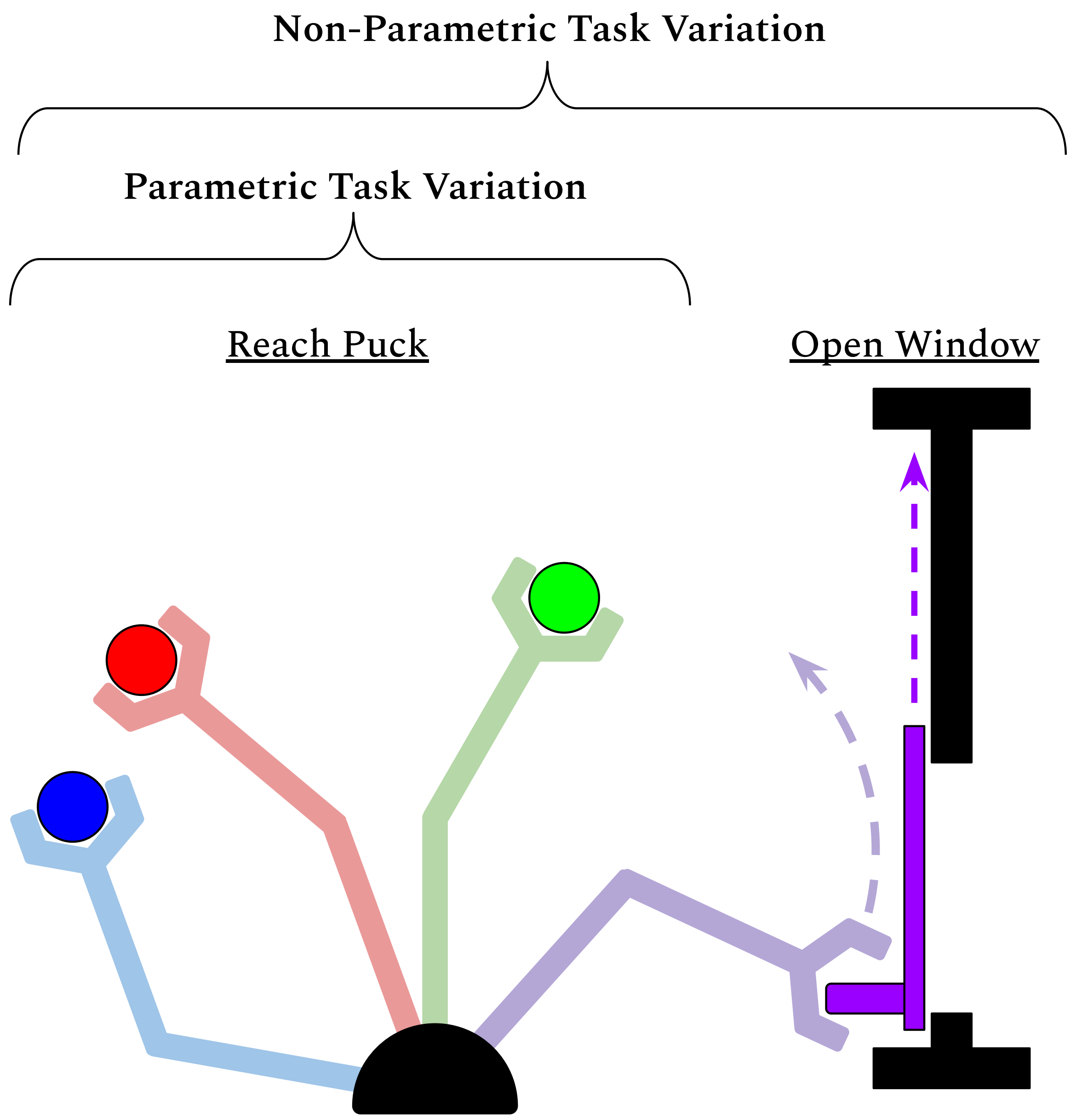


- 50 qualitatively different manipulation tasks using a simulated Sawyer robot
- Parametric variability with tasks, e.g. goal positions
- Unified observation and action space
- Structured, multi-component rewards for all tasks

Meta-World Task Suite

What's different: A new multi-task and meta-RL benchmark with a wide range of tasks to study how meta-RL accelerates acquisition of new tasks





50 qualitatively different manipulation tasks using a simulated Sawyer robot

Parametric variability with tasks, e.g. goal positions

Unified observation and action space

Structured, multi-component rewards for all tasks

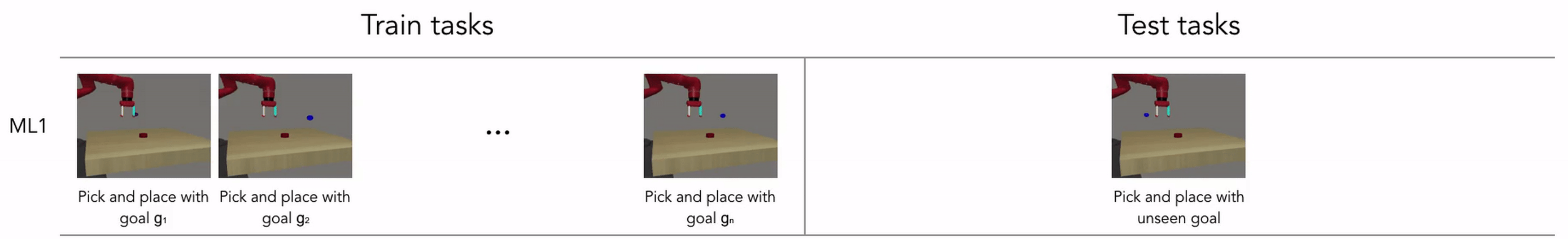
Evaluation Modes

Evaluation Modes

ML1
Train: Goals
Test: Unseen goals

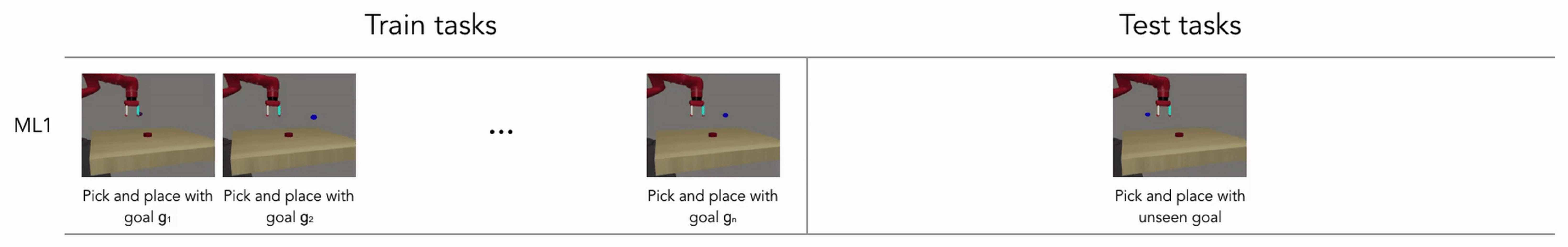
Evaluation Modes

ML1
Train: Goals
Test: Unseen goals



Evaluation Modes

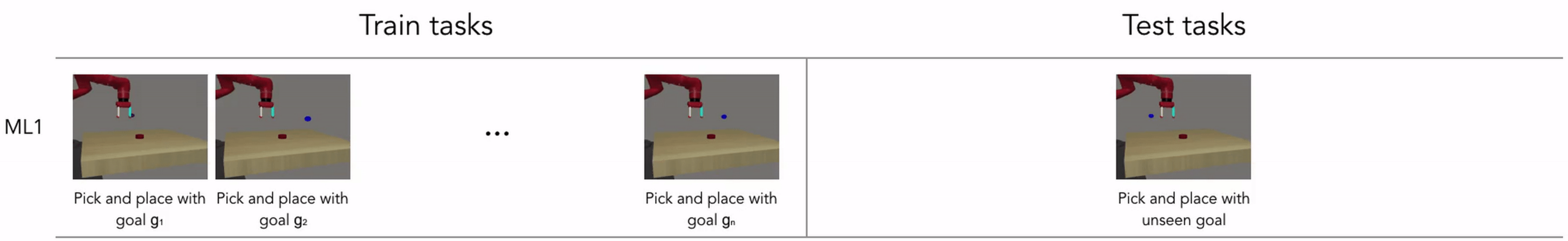
ML1
Train: Goals
Test: Unseen goals



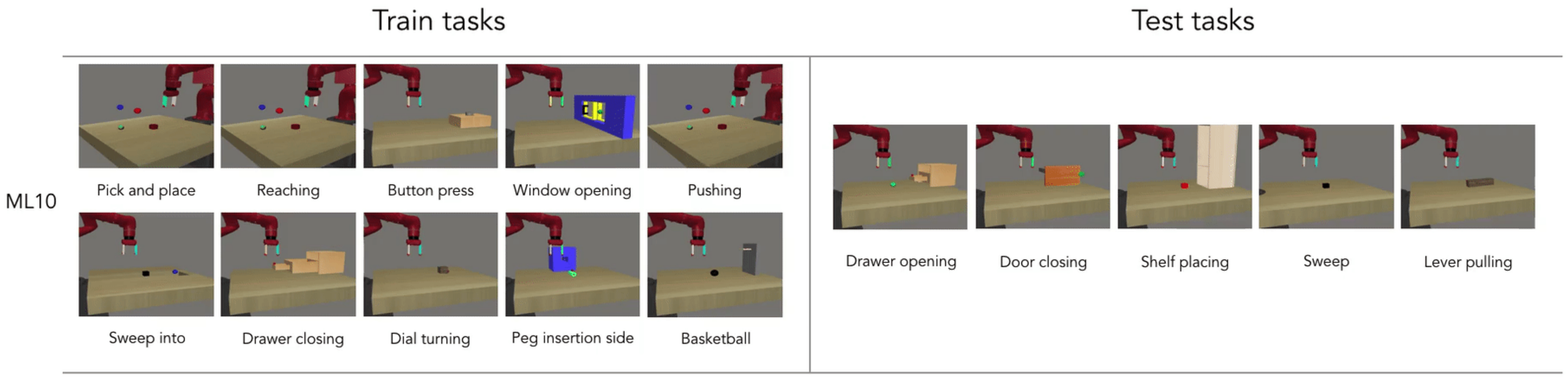
ML10/MT10
Train: 10 /10 tasks
Test: 5 unseen tasks

Evaluation Modes

ML1
Train: Goals
Test: Unseen goals

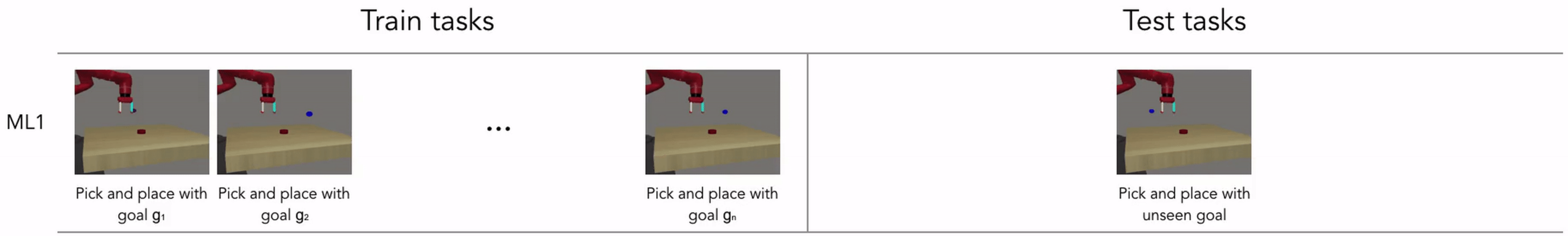


ML10/MT10
Train: 10 /10 tasks
Test: 5 unseen tasks

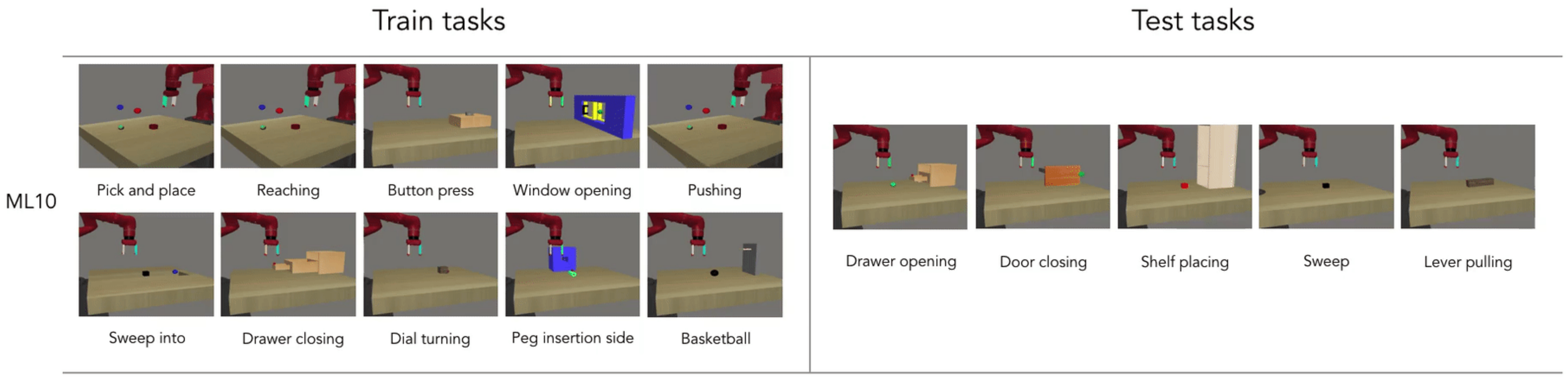


Evaluation Modes

ML1
Train: Goals
Test: Unseen goals



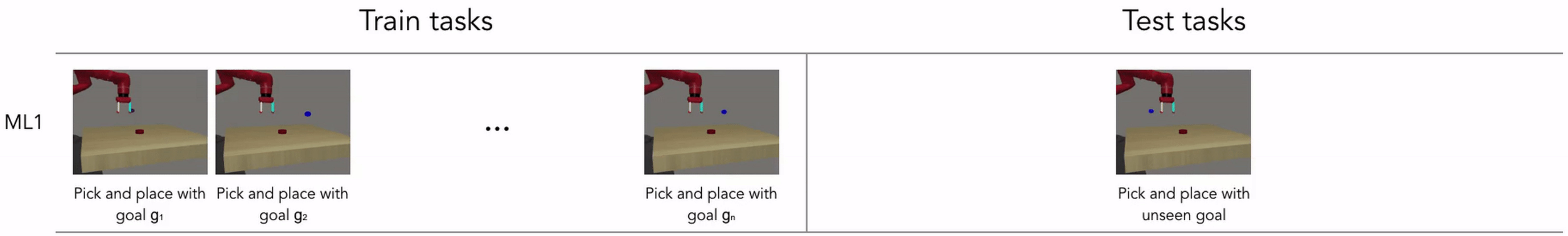
ML10/MT10
Train: 10 /10 tasks
Test: 5 unseen tasks



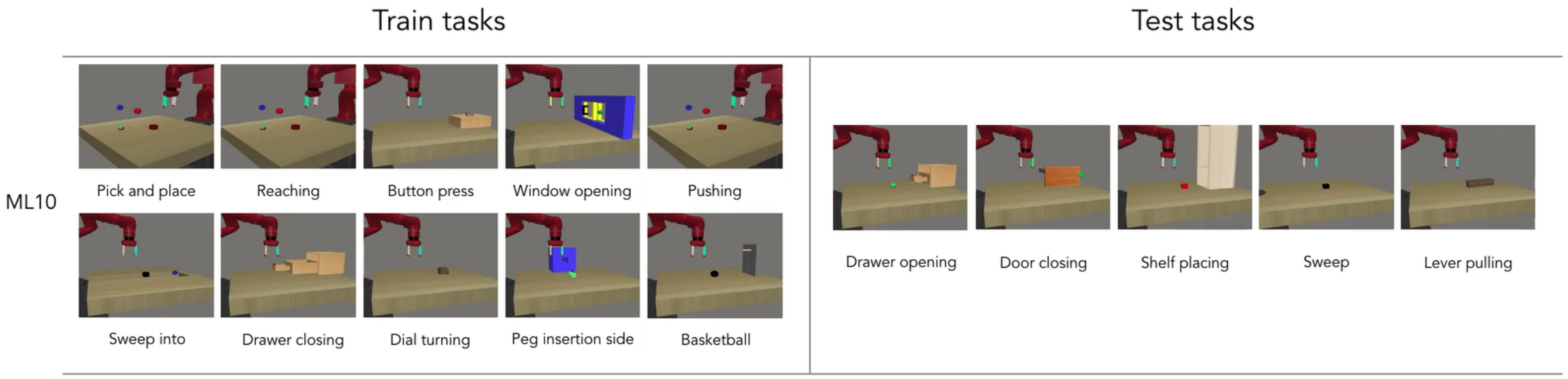
ML45/MT50
Train: 45 / 50 tasks
Test: 5 unseen tasks

Evaluation Modes

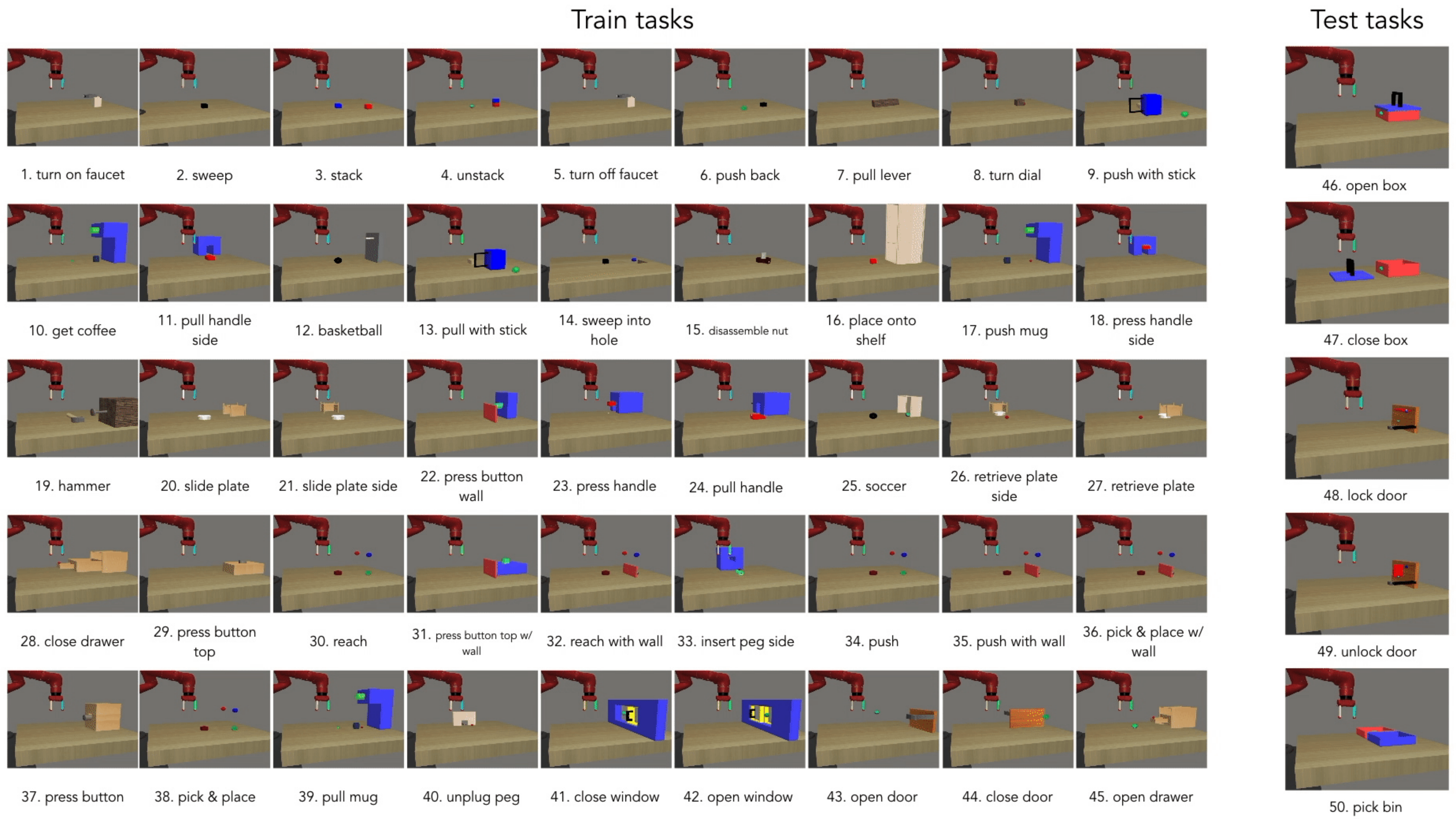
ML1
Train: Goals
Test: Unseen goals



ML10/MT10
Train: 10 /10 tasks
Test: 5 unseen tasks



ML45/MT50
Train: 45 / 50 tasks
Test: 5 unseen tasks



Evaluation Results

Evaluation Results

Multi-Task RL

Methods:

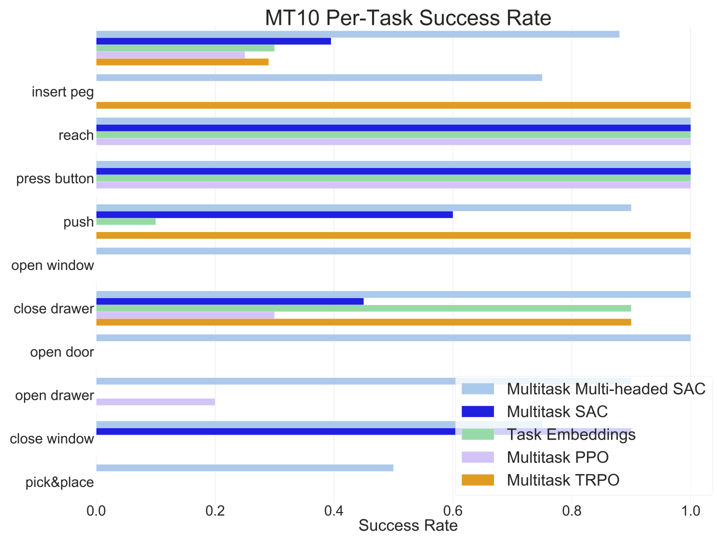
- Multi-Task PPO
- Multi-Task TRPO
- Multi-Task SAC
- Task Embeddings
- Multi-Task Multi-Headed SAC

Evaluation Results

Multi-Task RL

Methods:

- Multi-Task PPO
- Multi-Task TRPO
- Multi-Task SAC
- Task Embeddings
- Multi-Task Multi-Headed SAC

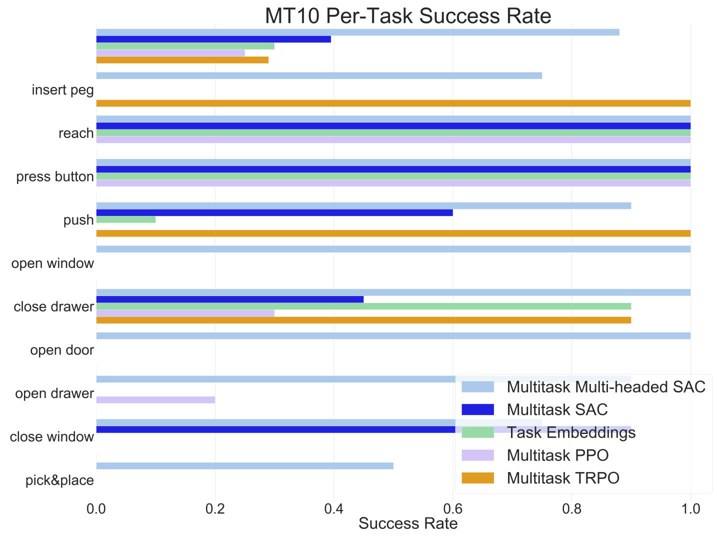


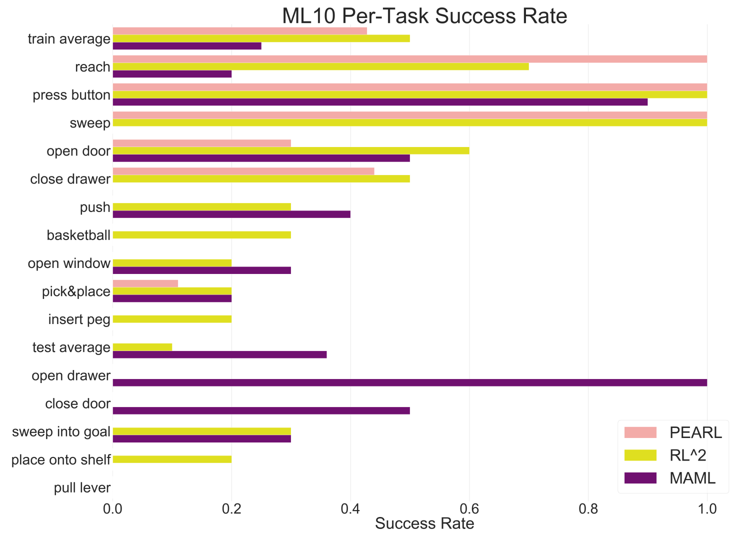
Evaluation Results

Multi-Task RL

Methods:

- Multi-Task PPO
- Multi-Task TRPO
- Multi-Task SAC
- Task Embeddings
- Multi-Task Multi-Headed SAC





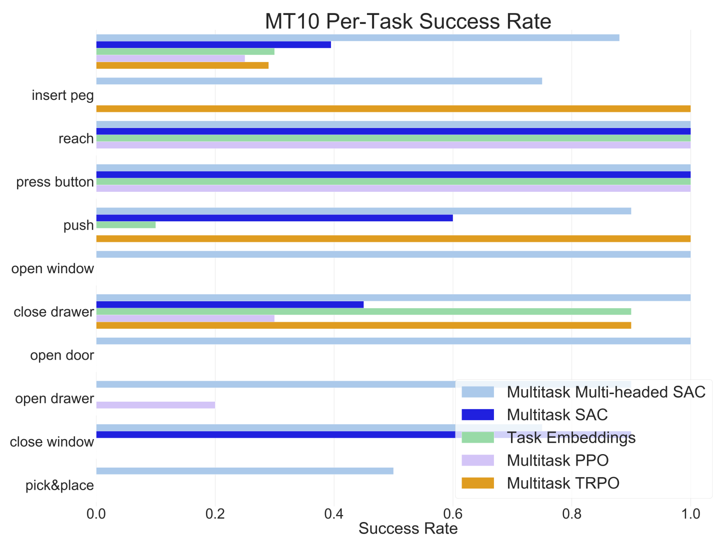
Evaluation Results

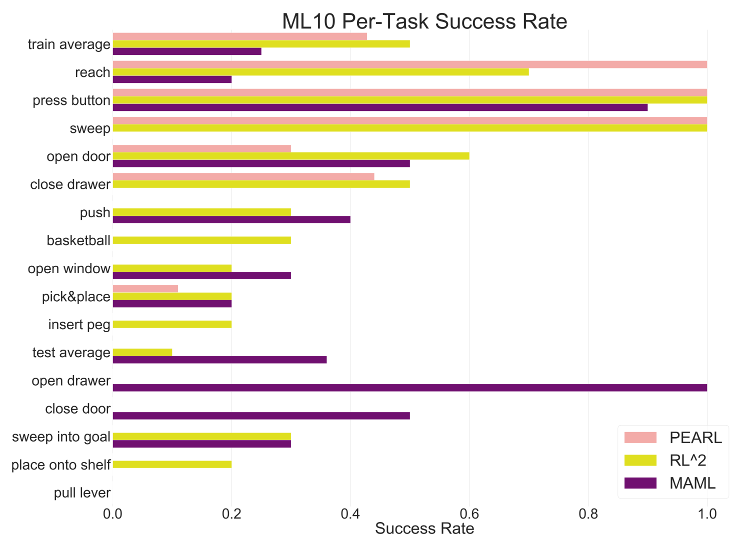
Multi-Task RL Methods:

- Multi-Task PPO
- Multi-Task TRPO
- Multi-Task SAC
- Task Embeddings
- Multi-Task Multi-Headed SAC

Meta-RL Methods:

- MAML
- RL²
- PEARL





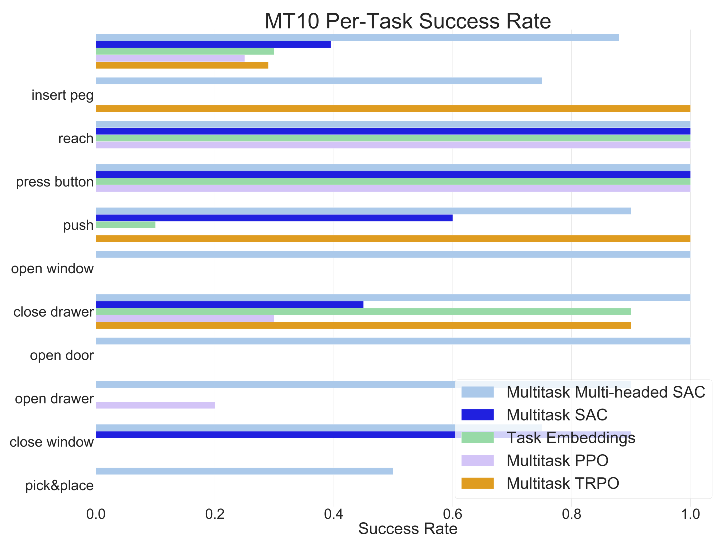
Evaluation Results

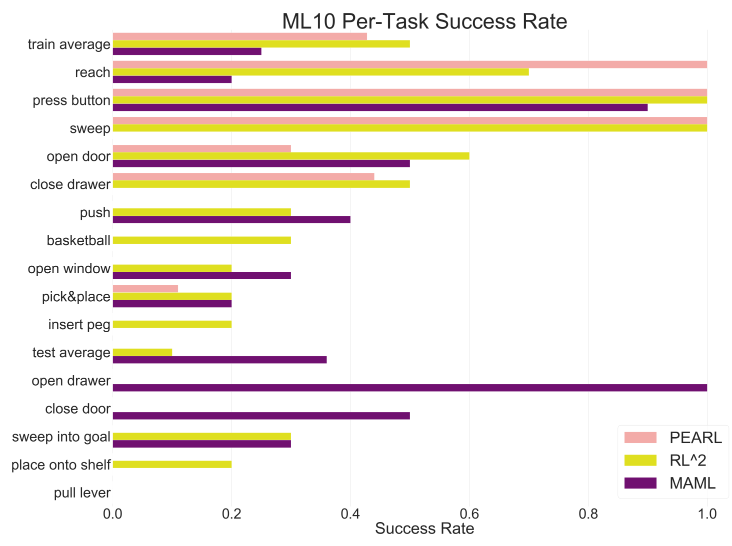
Multi-Task RL Methods:

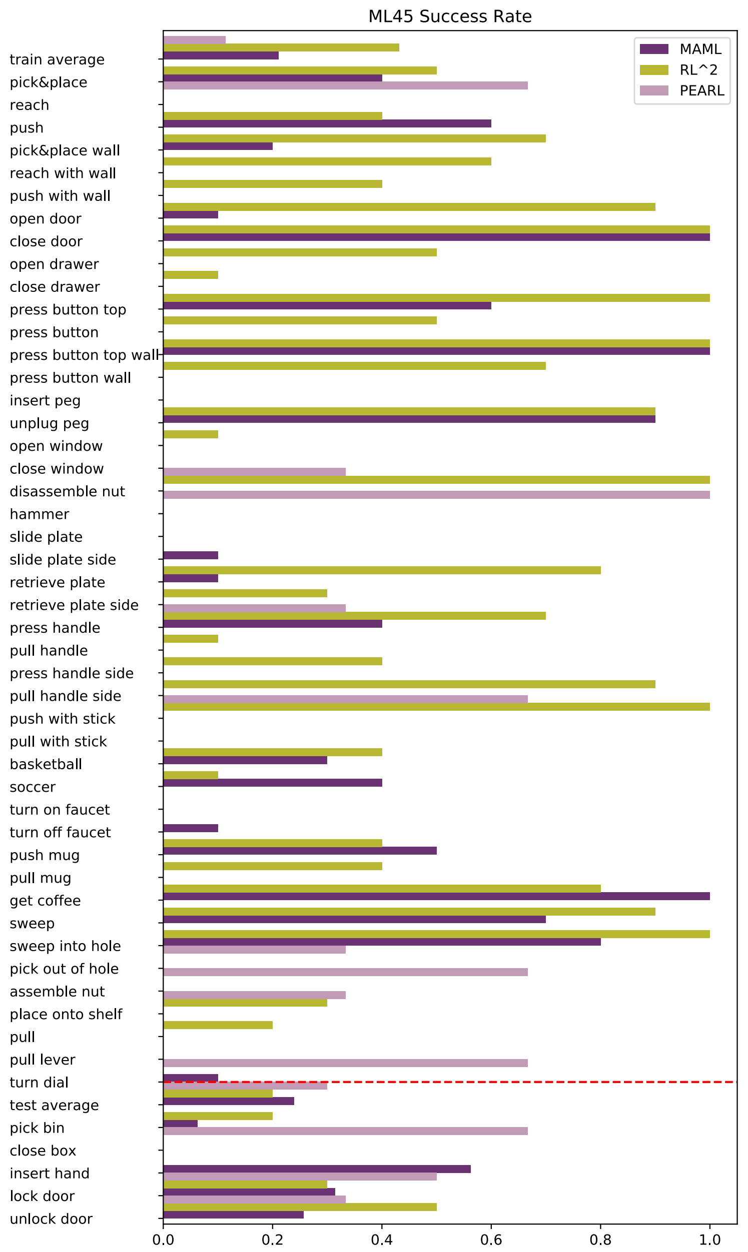
- Multi-Task PPO
- Multi-Task TRPO
- Multi-Task SAC
- Task Embeddings
- Multi-Task Multi-Headed SAC

Meta-RL Methods:

- MAML
- RL^2
- PEARL







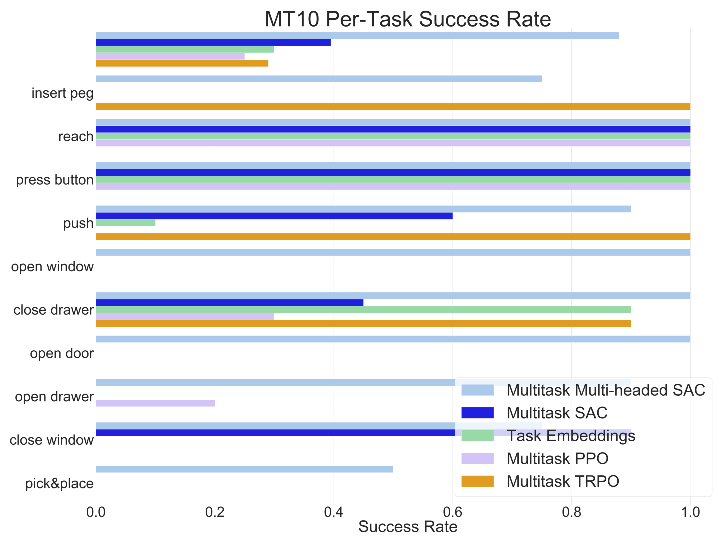
Evaluation Results

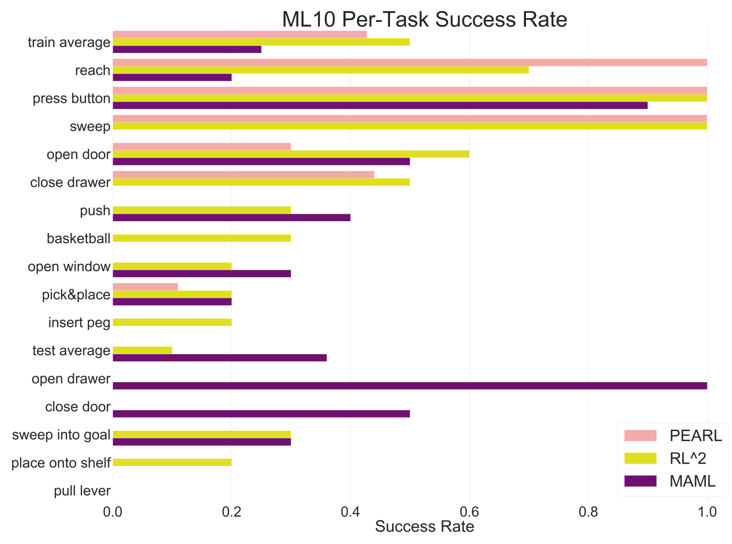
Multi-Task RL Methods:

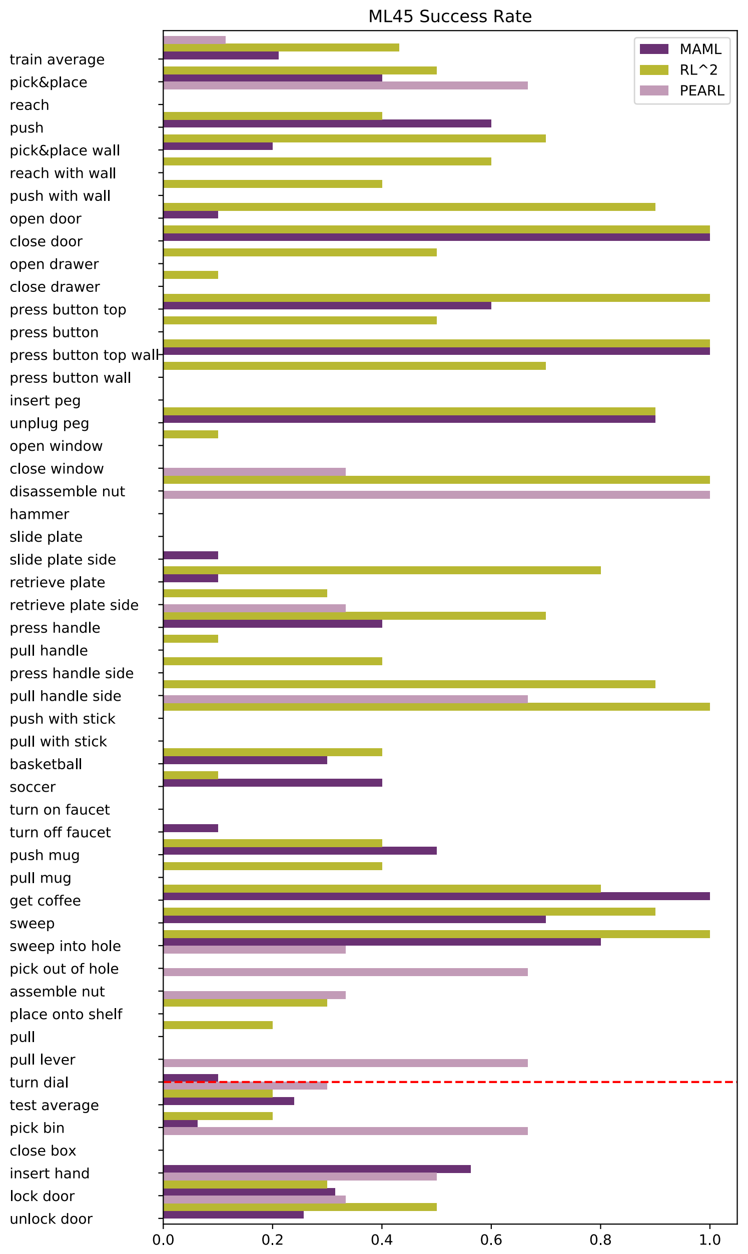
- Multi-Task PPO
- Multi-Task TRPO
- Multi-Task SAC
- Task Embeddings
- Multi-Task Multi-Headed SAC

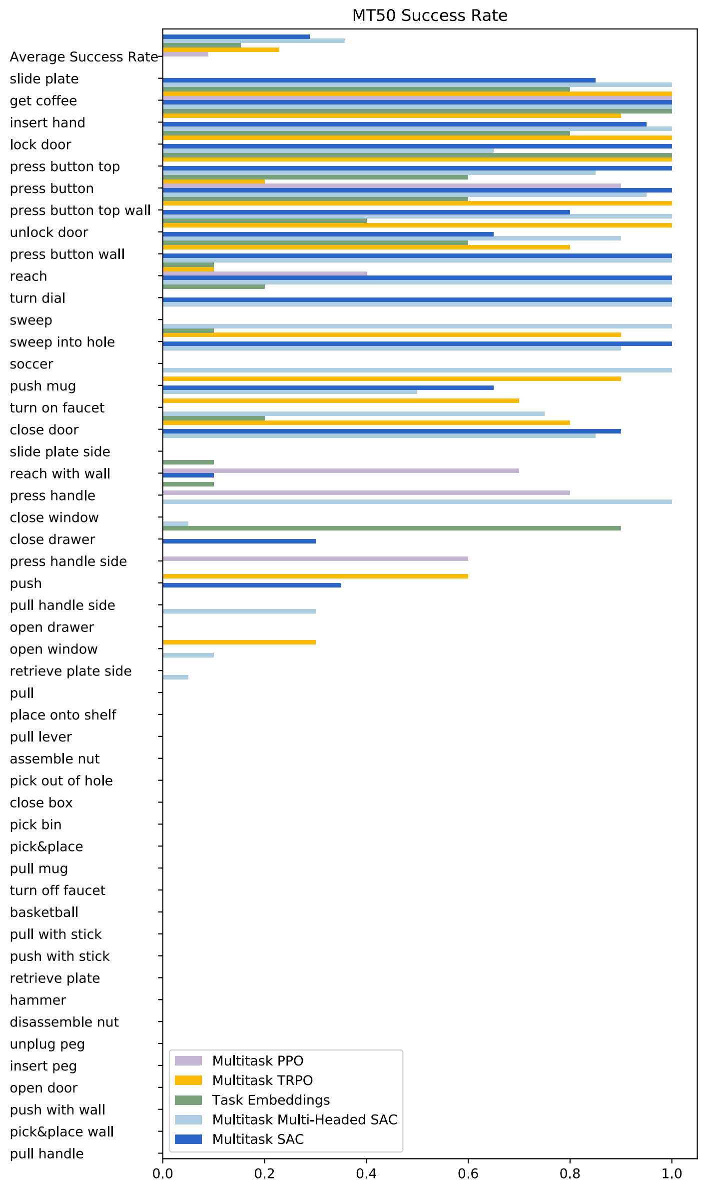
Meta-RL Methods:

- MAML
- RL^2
- PEARL









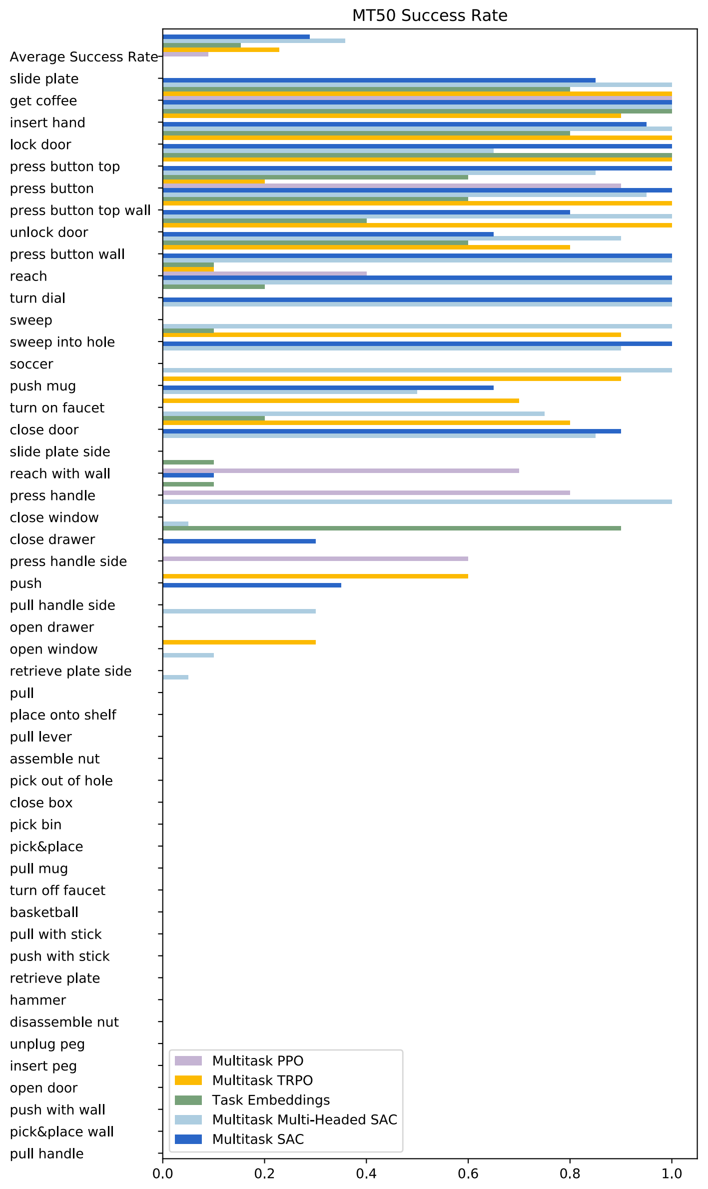
Evaluation Results

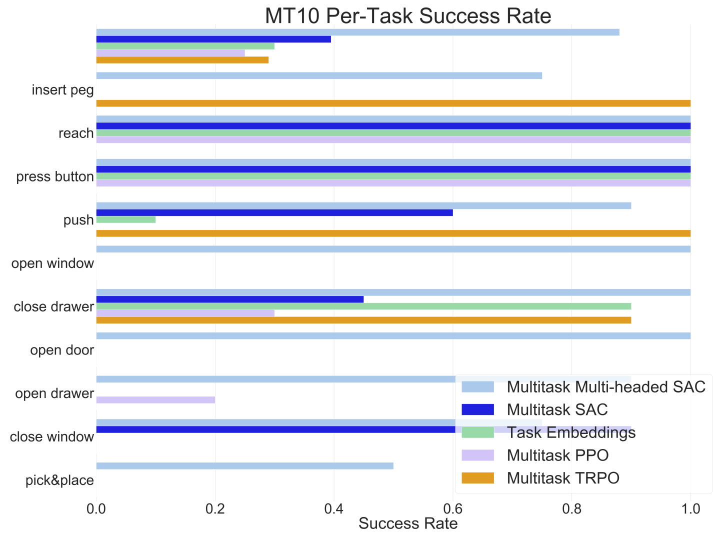
Multi-Task RL Methods:

- Multi-Task PPO
- Multi-Task TRPO
- Multi-Task SAC
- Task Embeddings
- Multi-Task Multi-Headed SAC

Meta-RL Methods:

- MAML
- RL^2
- PEARL



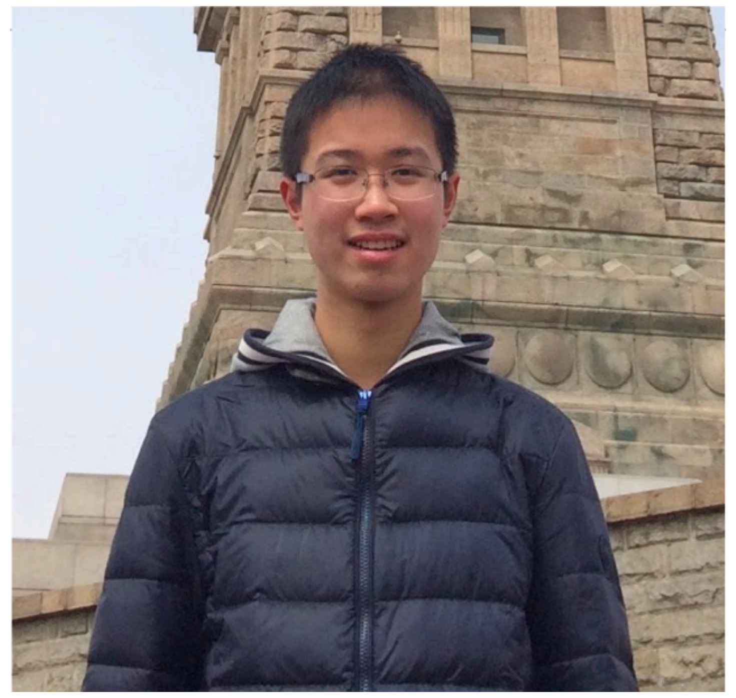


For full evaluation results of all methods, please come to our poster session!

Takeaways

- An open-sourced multi-task and meta-RL benchmark with 50 robotics manipulation tasks
- Thorough evaluations on current multi-task and meta-RL algorithms on five different modes

Collaborators


Tianhe Yu


Deirdre Quillen


Karol Hausman


Chelsea Finn


Sergey Levine

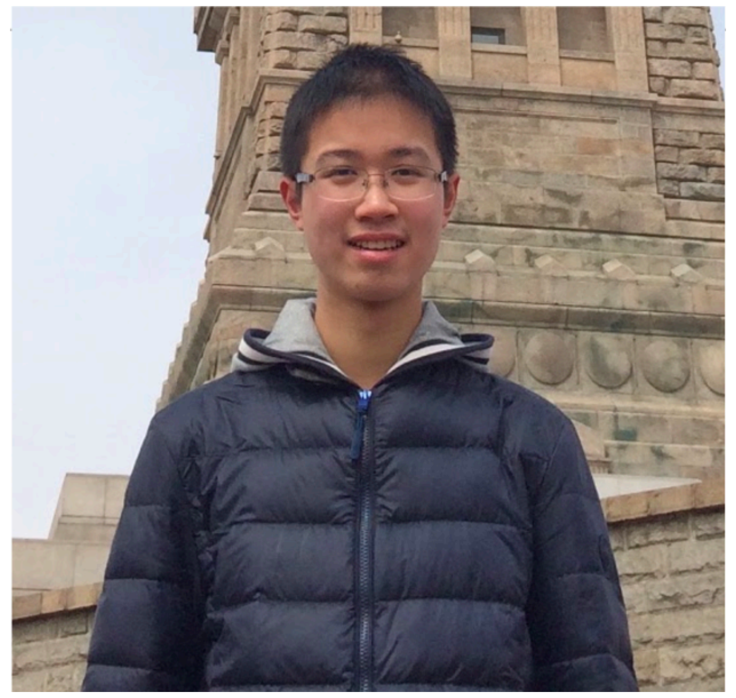
Video and Code

https://meta-world.github.io/

Takeaways

- An open-sourced multi-task and meta-RL benchmark with 50 robotics manipulation tasks
- Thorough evaluations on current multi-task and meta-RL algorithms on five different modes

Collaborators


Tianhe Yu


Deirdre Quillen


Karol Hausman


Chelsea Finn


Sergey Levine

Video and Code

https://meta-world.github.io/

Please come to our poster for a demo!