384.195 Robot Learning

Course Project: From Single-Task to Multi-Task Reinforcement Learning in Meta-World

> **Project Duration:** Nov. 28th, 2025 – Jan. 31st, 2026

---

1 Project Overview

Key Administrative Information

- **TA Consultation:** Fridays, 10:00–12:00 (Dec. 05, 2025 – Jan. 16, 2026), no appointment required.

- **Report Submission:** Jan. 23, 2026, 23:59 (IEEE format, 3–6 pages).

- **Intermediate Presentation Submission:** Dec. 17, 2025, 23:59.

- **Final Presentation Submission:** Jan. 28, 2026, 23:59.

- Include **names, photos and student numbers** of all group members.

- Submit **one report per group** via TUWEL.

2 Project Background & Motivation

Meta-World Benchmark

Meta-World is an open-source benchmark for meta-reinforcement learning and multi-task RL, consisting of **multiple simple robotic manipulation tasks**. It includes three standard task suites:

- MT1: single-task learning

- MT10: 10-task multi-task learning

- MT50: full 50-task multi-task learning

Documentation: https://metaworld.farama.org/

Reinforcement learning has achieved remarkable progress in single-task environments; however, scaling RL methods to dozens of tasks simultaneously introduces significant challenges, including:
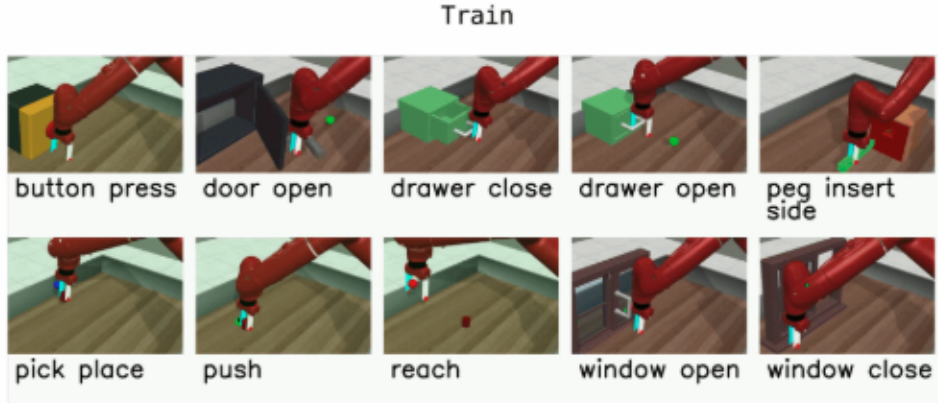
Figure 1: Visualization of the MT10 taskset of Meta-World.

- Interference and instability in shared representations.

- Catastrophic forgetting when tasks are learned jointly.

- Generalization difficulty and task-conditioning complexity.

This project provides hands-on experience in addressing these challenges through:

- Developing a single-task policy with RL baselines.

- Exploring the generalized policy scaling behavior in multi-task settings.

- Extending algorithms for multi-task learning of MT10 taskset.

3 Project Objectives

By completing this project, you will:

- Implement and train an RL agent on Meta-World single tasks (i.e. reach, push, pick-place).

- Adapt and improve a suitable RL algorithm for multi-task training towards above listed three tasks (https://metaworld.farama.org/introduction/basic_usage/#custom-benchmarks).

- Compare single-task and multi-task performance in terms of sample efficiency (training steps), stability (mean reward & standard deviation), and generalization (tasks average performance).

- Evaluate the influence of architectural choices when scaling to larger task sets.

- Optionally explore advanced strategies such as attention-based task conditioning, curriculum learning, or transfer learning.

4 Project Requirements & Deliverables

4.1 Project Deliverables

Deliverables

- An intermediate short presentation describing:

  - Environment setup & algorithm baseline selection.
  - Initial algorithmic idea.

- Well-documented and reproducible implementation code.

- Experimental results including:

  - Three single-task policy experiments (reach, push, pick-place).
  - One multi-task policy experiment on the above listed three tasks.
  - Exploring scaling experiments toward MT10 taskset.

- A comparative analysis of convergence behavior, reward trends, stability, and inter-task interference.

- A final report and presentation summarizing methods, results, and key insights.

4.2 Technical Requirements & Tools

- Use the Meta-World environment as described in the official documentation.

- Select and implement algorithms capable of multi-task learning (e.g., SAC, PPO) using an RL framework you are familiar with.

- Ensure reproducibility through fixed seeds, logging, and clear visualization of learning curves.

- Optionally investigate enhancements to the multi-task setup, such as incorporating temporal-aware networks, shared backbones, multi-head architectures, task conditioning, or other architectural improvements.

4.3 Evaluation Criteria

The project will be assessed based on:

- **Performance**: performance is evaluated in terms of sample efficiency (training steps and convergence speed), stability (mean reward and the standard deviation across evaluation tasks), and generalization (average success rate across tasks) on both the MT1, MT3, and MT10 task sets. The grading is based on the following performance criteria:

- Three single-task experiments: reach (success rate $> 90\%$), push (success rate $> 30\%$), pick-place (success rate $> 30\%$).
- Multi-task policy for the three tasks: average success rate $> 40\%$.
- MT10 task-set policy: average success rate $> 30\%$.

- **Depth of analysis**: insightful interpretation of experimental results.

- **Clarity**: professional presentation of methodology, findings, and conclusions.

- **Innovation(bonus)**: use of advanced, improved methods to address the potential limitations.

5 Possible Extensions & Stretch Opportunities

- Balance the weighting of losses across tasks.

- Explore shared versus task-specific network architectures.

- Incorporate task embeddings mechanisms.

- Apply curriculum-learning strategies to enhance multi-task performance.

- Investigate continual-learning dynamics and mitigate catastrophic forgetting.

- Leverage the expert policy as guidance for policy learning.

6 Summary

This project offers an opportunity to engage with modern reinforcement learning challenges by scaling from single-task to multi-task learning on a large and diverse benchmark. Through implementation, experimentation, and analysis, you will gain valuable insights into generalization, representational sharing, learning dynamics, and system scalability in multi-task RL.

References

1. Sutton, R. S., & Barto, A. G. Reinforcement Learning: An Introduction, 2nd Edition, 2020.

2. Yu, Tianhe, et al. "Meta-world: A benchmark and evaluation for multi-task and meta reinforcement learning." Conference on robot learning. PMLR, 2020.

3. https://metaworld.farama.org/

4. https://gymnasium.farama.org/