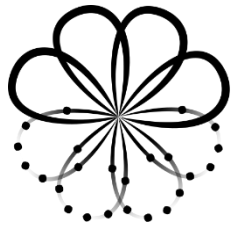


Institute of
Artificial Intelligence

Leibniz
Universität
Hannover

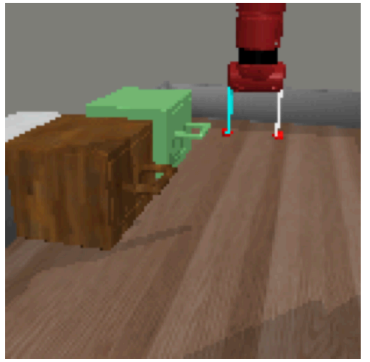
Online Meta-Exploration from Offline Data in Realistic Robotic Tasks [Rafailov et. al. 2021]

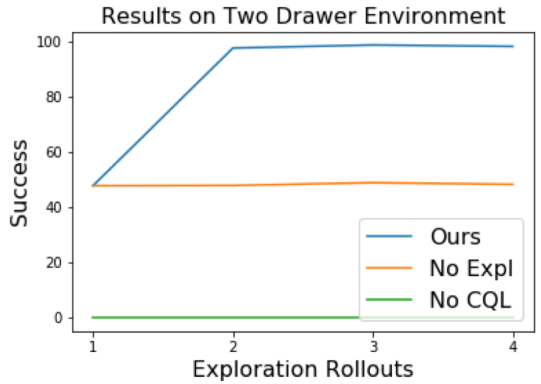
Artur Ganzha 10019651

Problem and Approach [Rafailov et. al. 2021]

- **Challenges from Real-World RL:**
  - High sample complexity
  - Distribution shifts
  - High-dimensional inputs (e.g., images)
  - Spurious memorization instead of learning
- **Current Solutions and Limitations:**
  - Meta-learning aids agents for quick adaptation to new tasks
  - requires a lot of data and is hard to optimize in high-dimensional spaces
- **Proposed Approach:**
  - Operates from images with sparse rewards
  - Meta-exploration from offline-data
  - Using variational filtering for representation learning (=)
  - Latent conservative model-free policy optimization (=)

Methodology and Results [Rafailov et. al. 2021]





- **Synthetic Exploration**
  - Creating extended exploration rollouts by mixing trajectories from different tasks
  - Assigning target tasks & zeroing out rewards for non-target tasks (to avoid memorization/correlation)
- **Variational Model Training**
  - Training a high-dimensional variational model to infer environment and tasks states from image sequences
- **Latent Conservative Model-Free policy Optimization**
  - Using CQL (Conservative Q-Learning) to slow over-estimation bias and distribution shift issues (primarily from offline RL)
  - **CQL** = Penalizes Q-Values of actions which are not well-supported by the offline data, causes reliable policy learning
- **Results**
  - Successfully solved a meta-learning task from a realistic problem
  - Outperformed naive combinations of previous approaches which already failed on the same task

Sources

[Rafailov et. al. 2021] https://offline-rl-neurips.github.io/2021/pdf/54.pdf