

Prevalence of Negative Transfer in Continual Reinforcement Learning: *Analysis and a Simple Baseline*

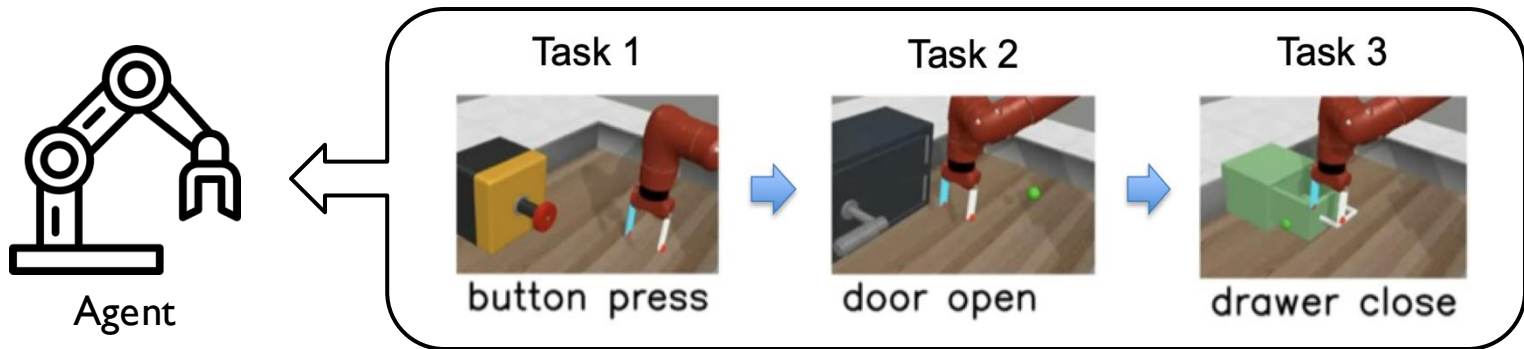
Hongjoon Ahn¹, Jinu Hyeon¹, Youngmin Oh², Bosun Hwang², and Taesup Moon¹

¹Seoul National University

² Samsung Advanced Institute of Technology

Continual Reinforcement Learning

- Continual Reinforcement Learning (CRL)

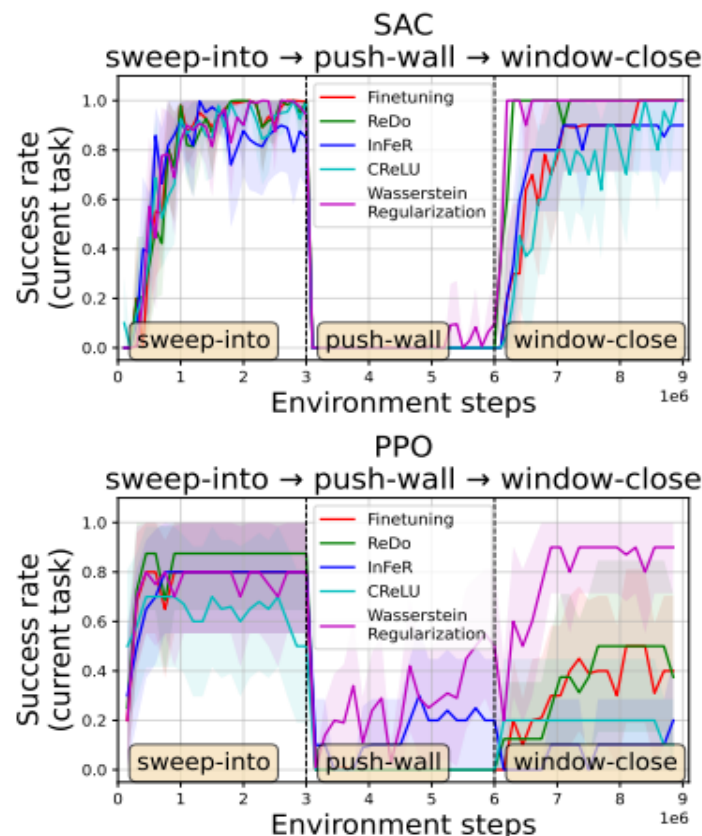
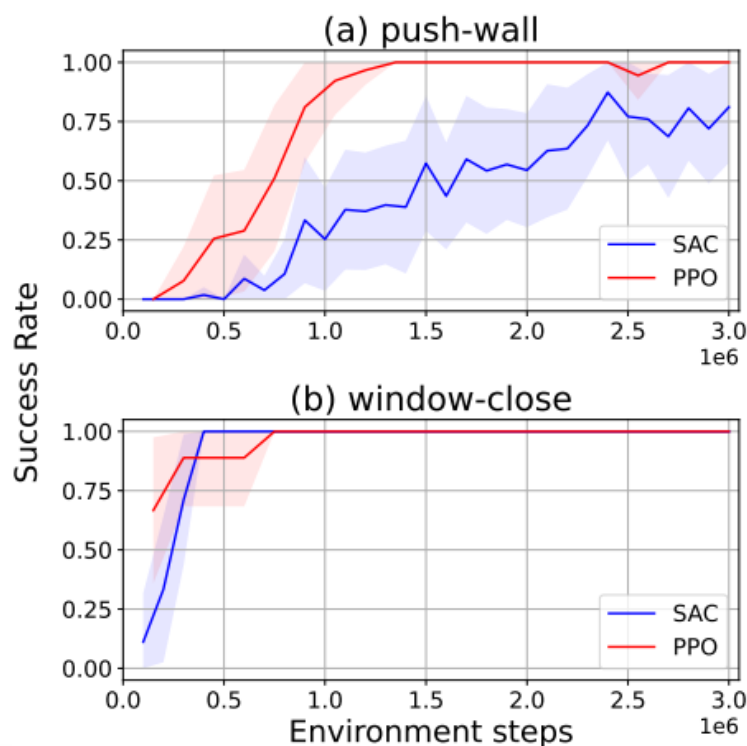


- A well-known dilemma in CRL



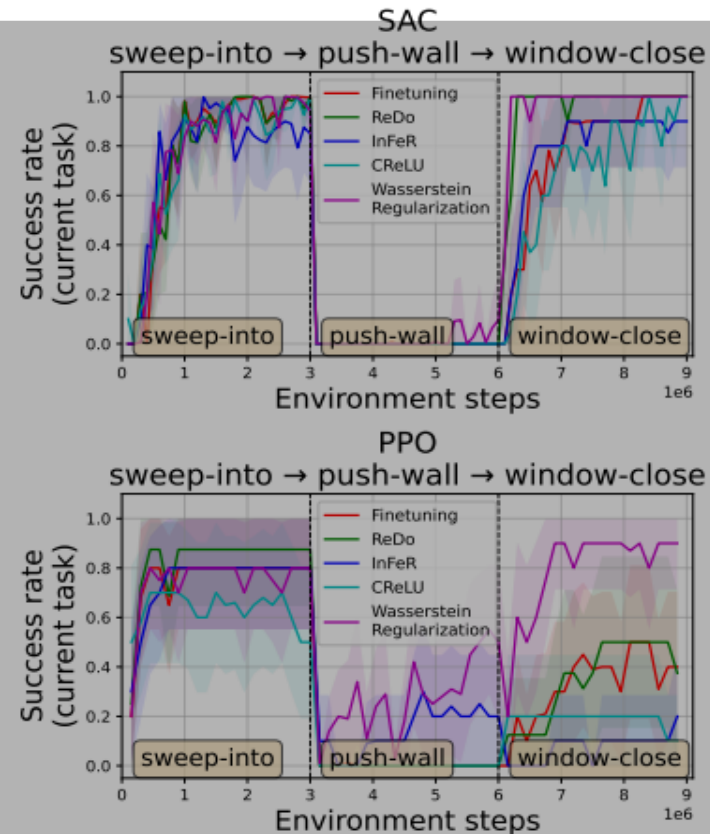
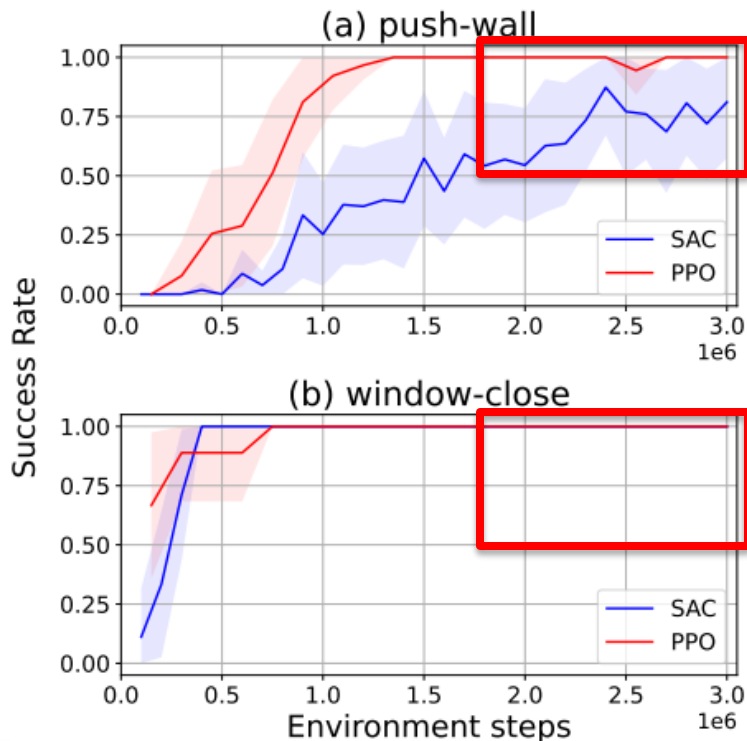
Negative Transfer in CRL

- Simple demonstration in 3-task experiment



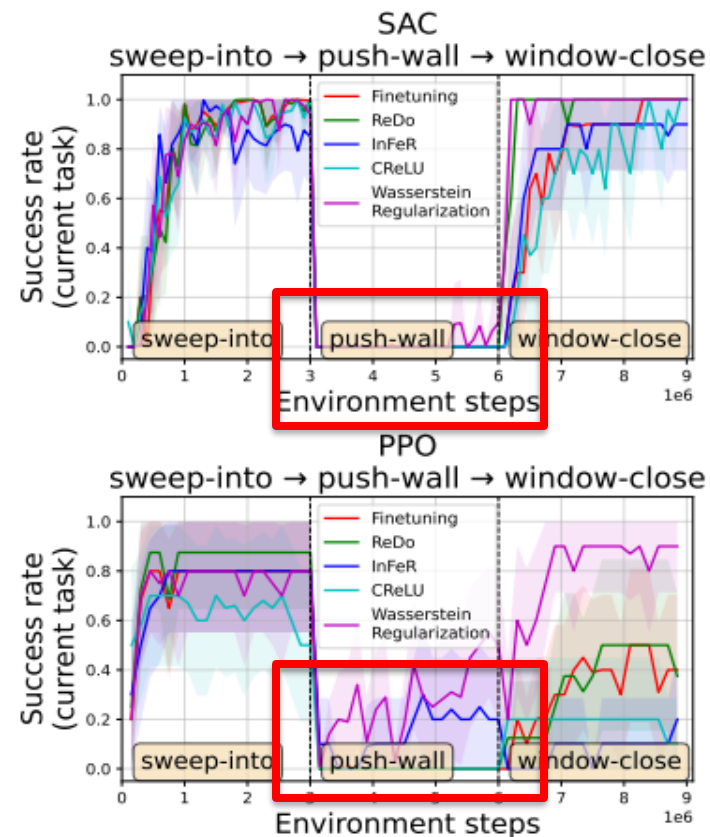
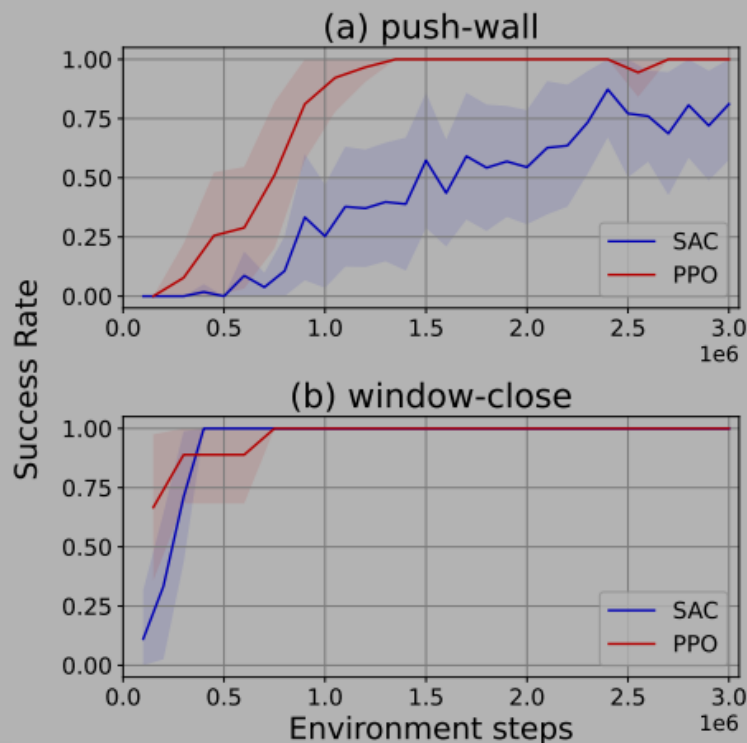
Negative Transfer in CRL

- Simple demonstration in 3-task experiment
 - Each of the tasks is easy to learn by RL algorithms **from scratch**



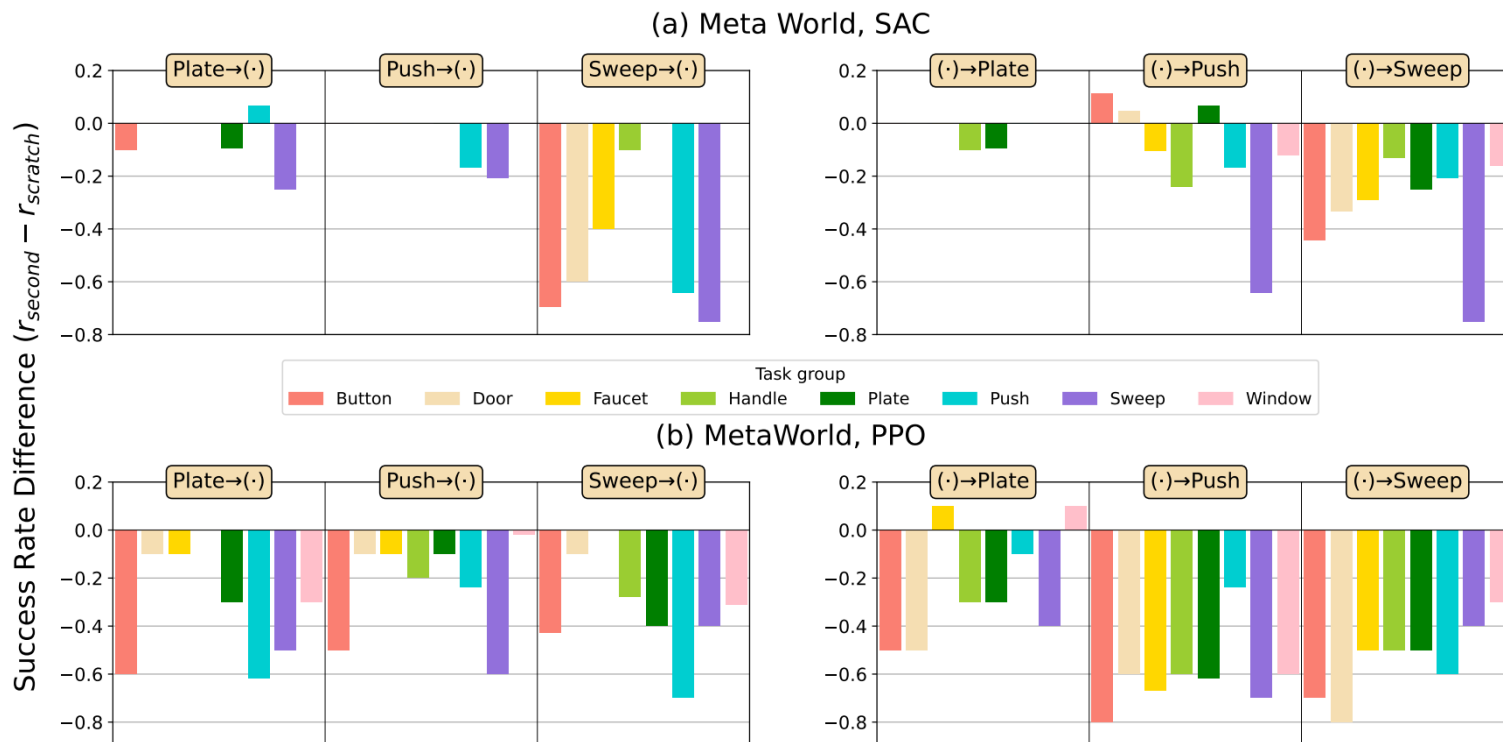
Negative Transfer in CRL

- Simple demonstration in 3-task experiment
 - But the performance collapses when it's learned **after other tasks**



Identifying Patterns of Negative Transfer

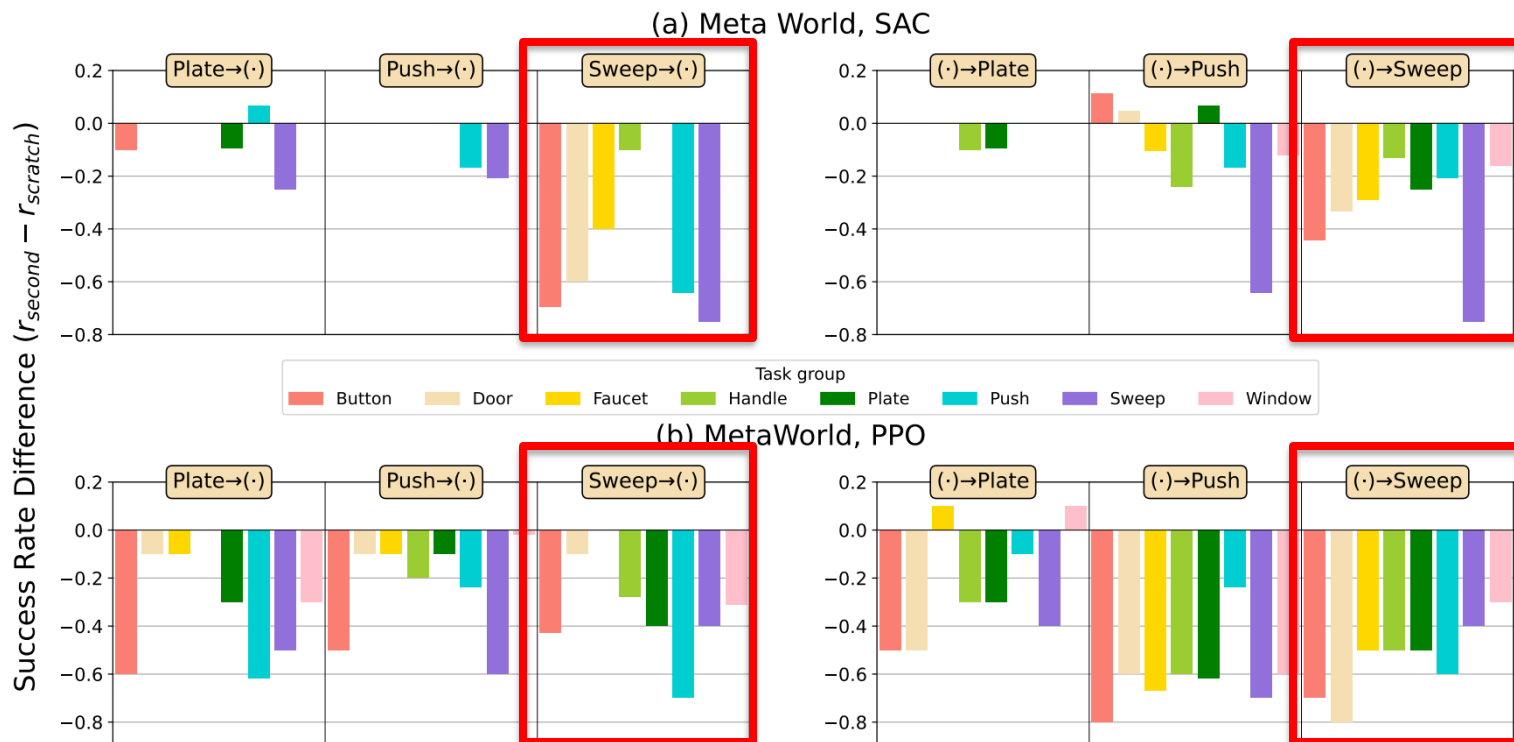
- 2-task experiments on Meta World



* Results from Atari and DeepMind Control can be found in our paper
ICLR 2025

Identifying Patterns of Negative Transfer

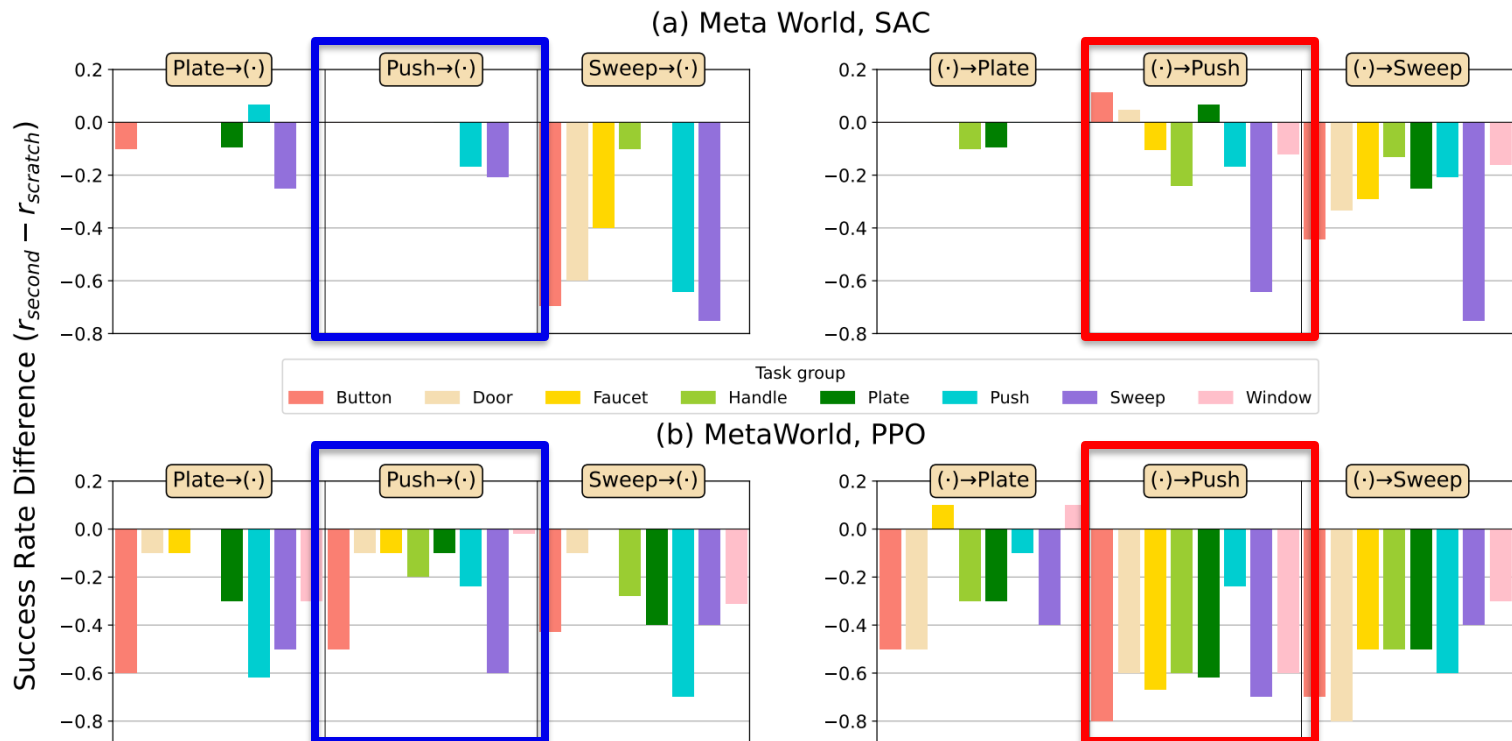
- 2-task experiments on Meta World
 - ‘Sweep’ tends to suffer from negative transfer



* Results from Atari and DeepMind Control can be found in our paper
ICLR 2025

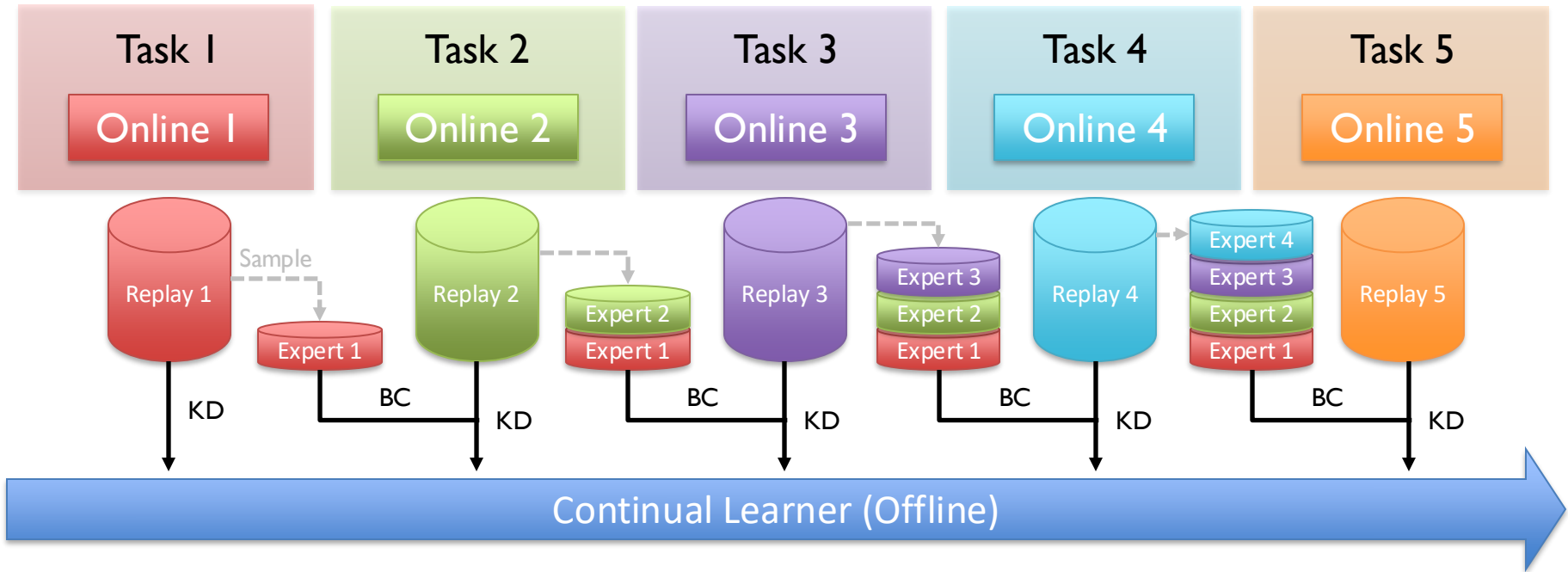
Identifying Patterns of Negative Transfer

- 2-task experiments on Meta World
 - ‘Push’ is learned fine when they come first, but struggle when they come second



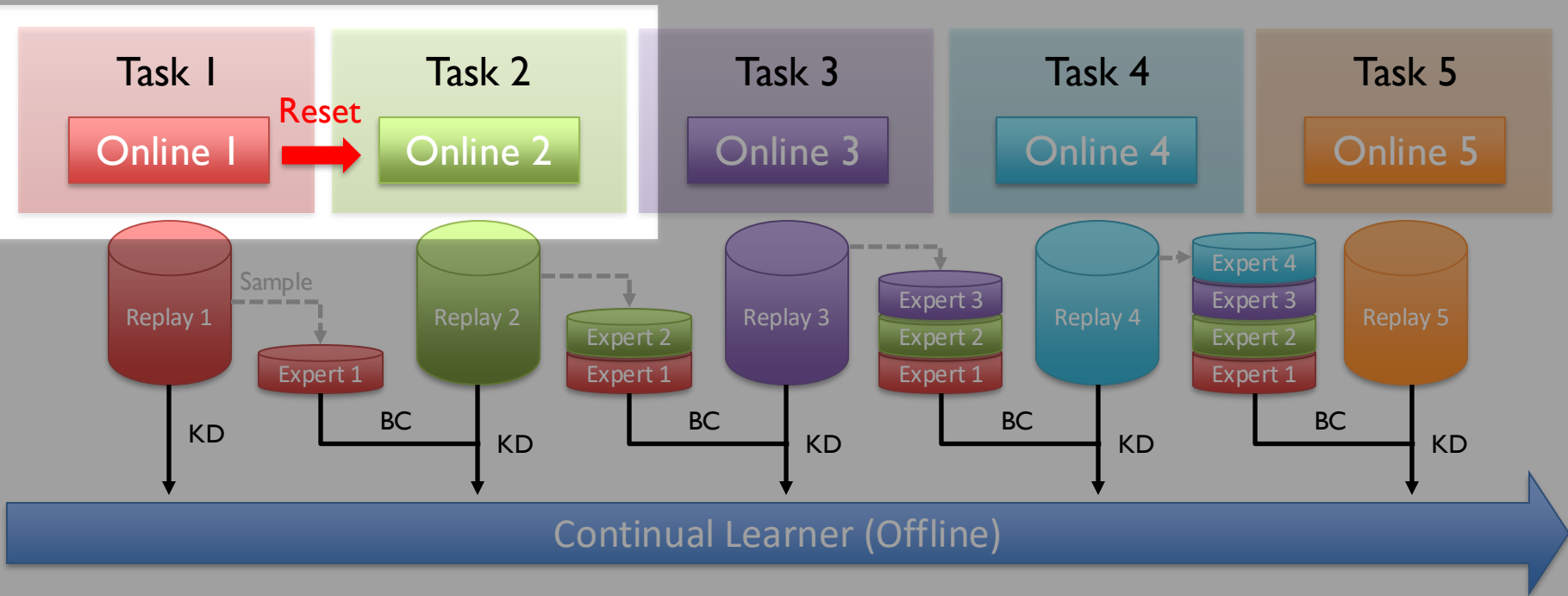
* Results from Atari and DeepMind Control can be found in our paper
ICLR 2025

Reset and Distill (R&D)



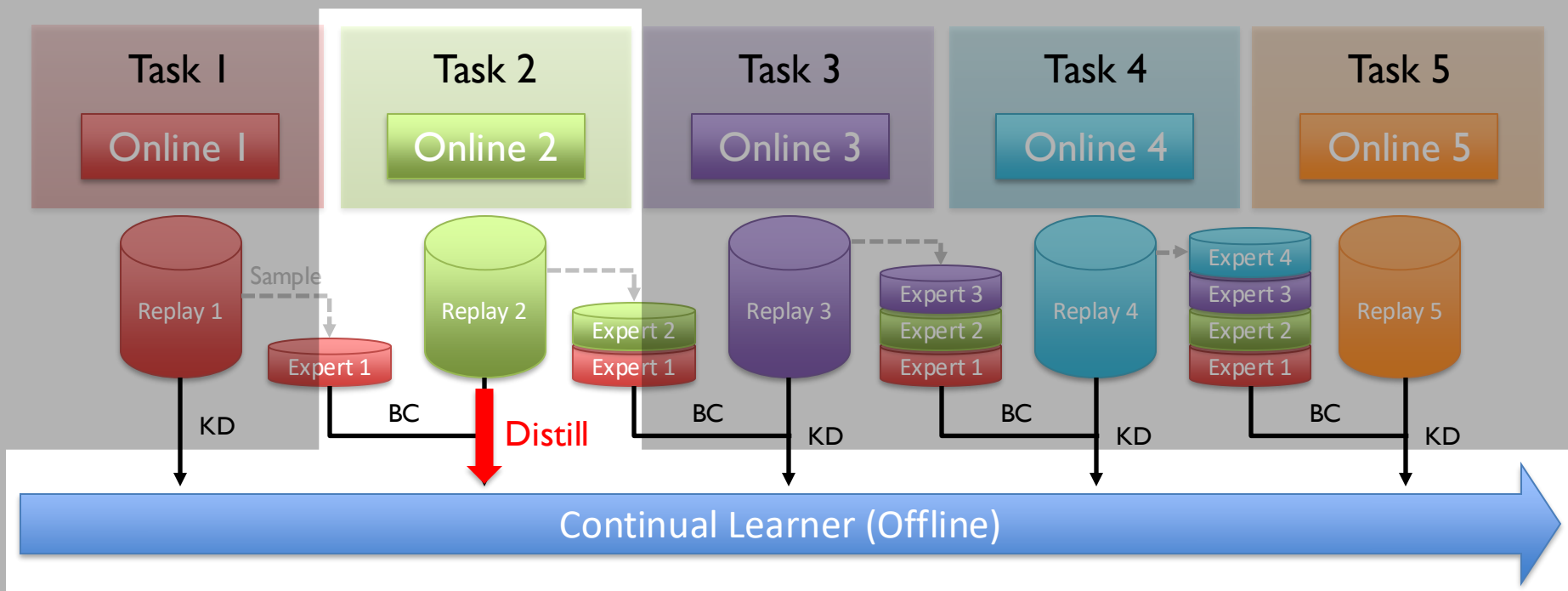
1. **Reset** the online learner after finishing each task
2. The online learner learns a novel task from scratch
3. Using KD, **distill** the knowledge of current task from the online learner to the offline(continual) learner
4. Store small sample of current task to prevent forgetting

Reset and Distill (R&D)



1. **Reset** the online learner after finishing each task
2. The online learner learns a novel task from scratch
3. Using KD, **distill** the knowledge of current task from the online learner to the offline(continual) learner
4. Store small sample of current task to prevent forgetting

Reset and Distill (R&D)

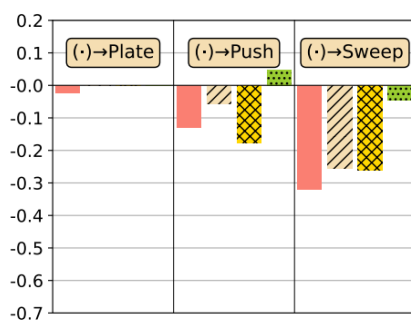
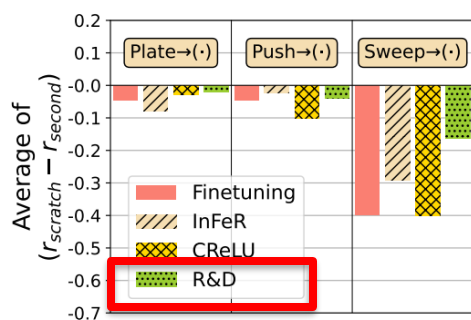


1. **Reset** the online learner after finishing each task
2. The online learner learns a novel task from scratch
3. Using KD, **distill** the knowledge of current task from the online learner to the offline(continual) learner
4. Store small sample of current task to prevent forgetting

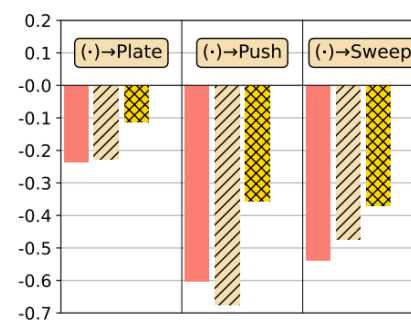
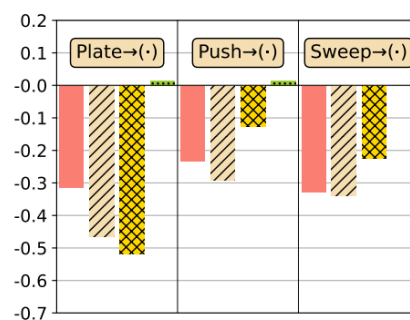
Experimental Results

- 2-task setups
 - **R&D** outperforms other baselines devised to improve plasticity

(a) SAC

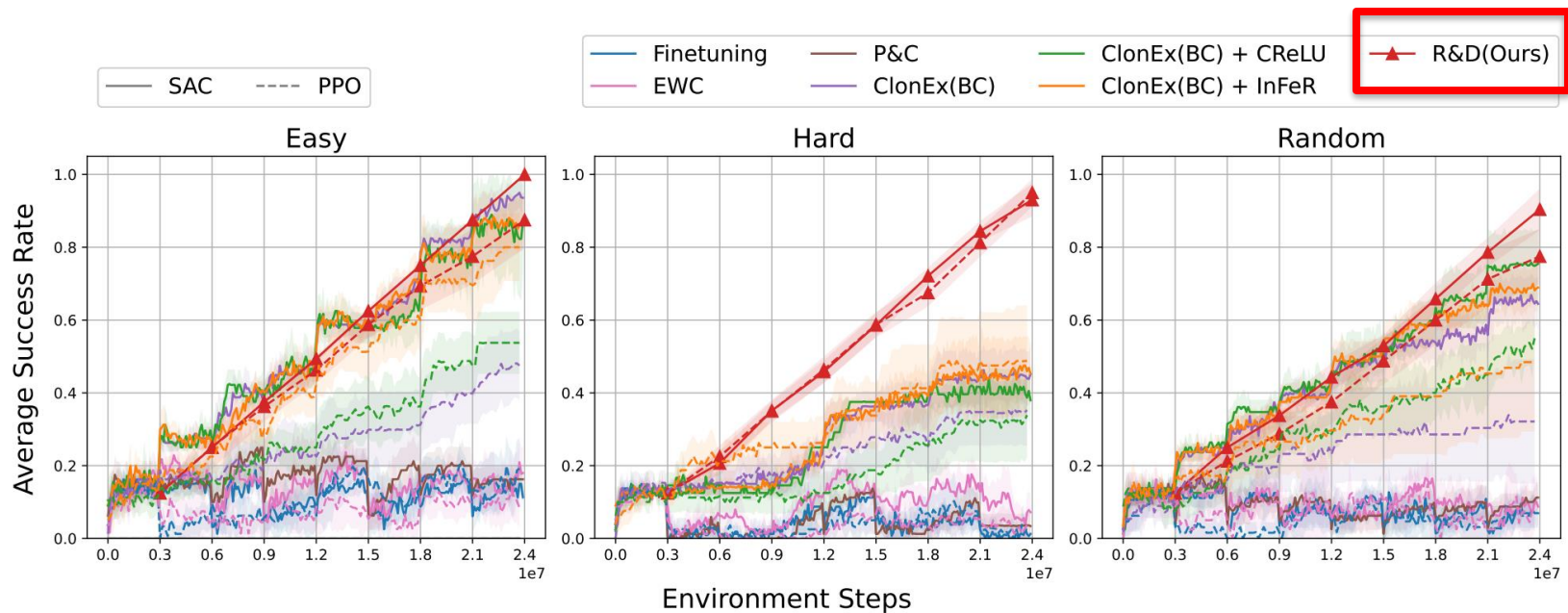


(b) PPO



Experimental Results

- Long sequence setups
 - **R&D** achieves higher success rate, especially in hard or randomly ordered sequences



Conclusion

- ***Negative transfer*** is pervasive in CRL
- We propose ***R&D***, a simple method to address negative transfer
- Experiments show that mitigating negative transfer is prerequisite for advancing CRL