Multi-agent Continual Coordination via Progressive Task Contextualization

Lei Yuan[1,2], Lihe Li[1], Ziqian Zhang[1], Fuxiang Zhang[1,2], Cong Guan[1], Yang Yu[1,2,*]
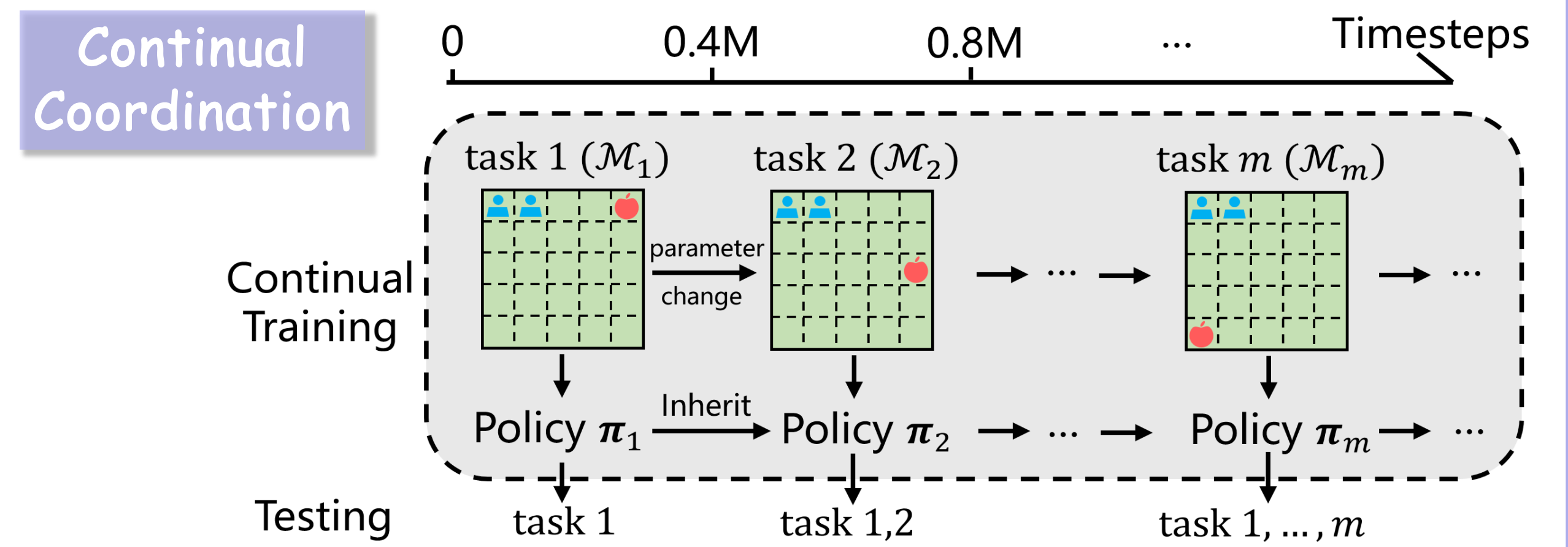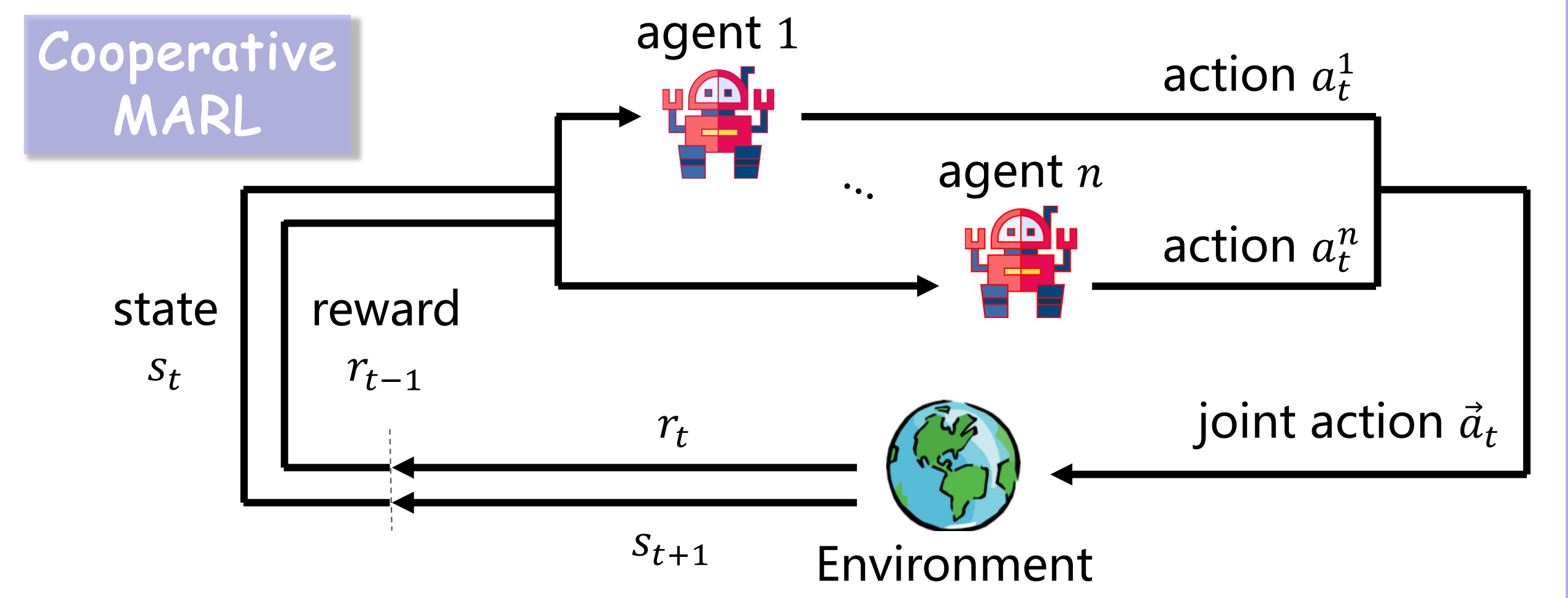[1]National Key Laboratory for Novel Software Technology, Nanjing University, [2]Polixir.ai

DAI 2023

LAMDA Learning And Mining from DatA https://www.lamda.nju.edu.cn

POLIXIR 南栖仙策

Introduction

□ Many real-world problems consist of multiple interactive agents, and can be modeled as multi-agent systems (MAS).

□ We can apply reinforcement learning (RL), a machine learning paradigm that involves agents learning to make decisions, to train multi-agent systems to solve different tasks.

□ If all agents in the multi-agent system share the same reward, this learning paradigm can be formulated as cooperative multi-agent reinforcement learning (CMARL).
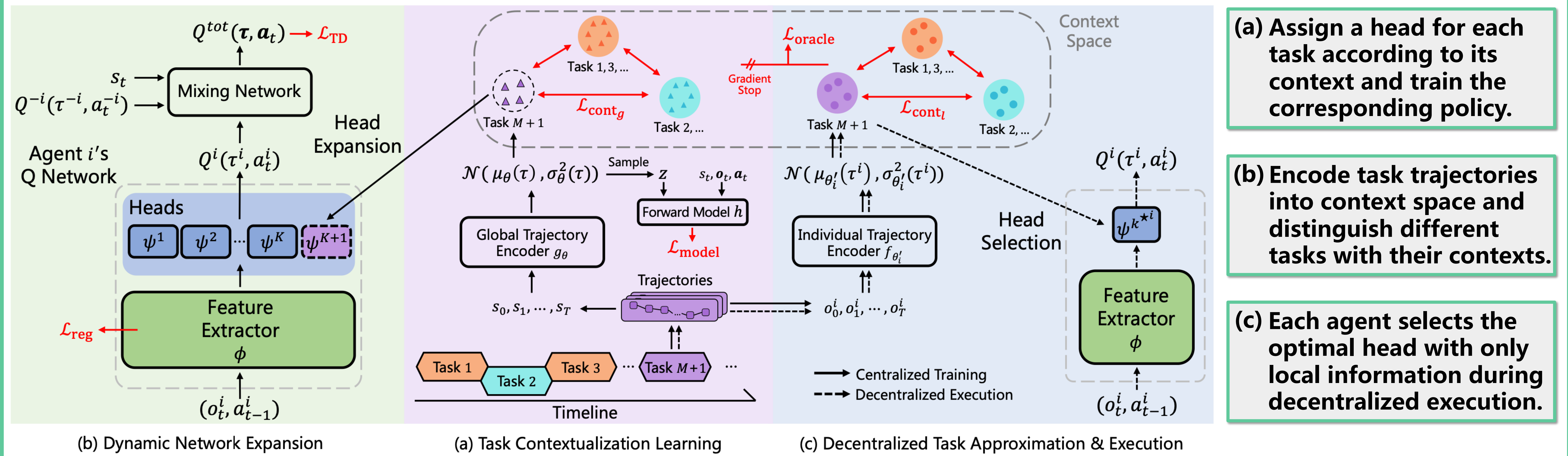
However,

□ Mainstream CMARL methods are still restricted to being trained in one single task or multiple tasks simultaneously and thus cannot learn a stream of tasks in a continual manner.

□ To solve this issue, we formalize continual coordination and learn a MAS in a stream of tasks via progressive task contextualization.

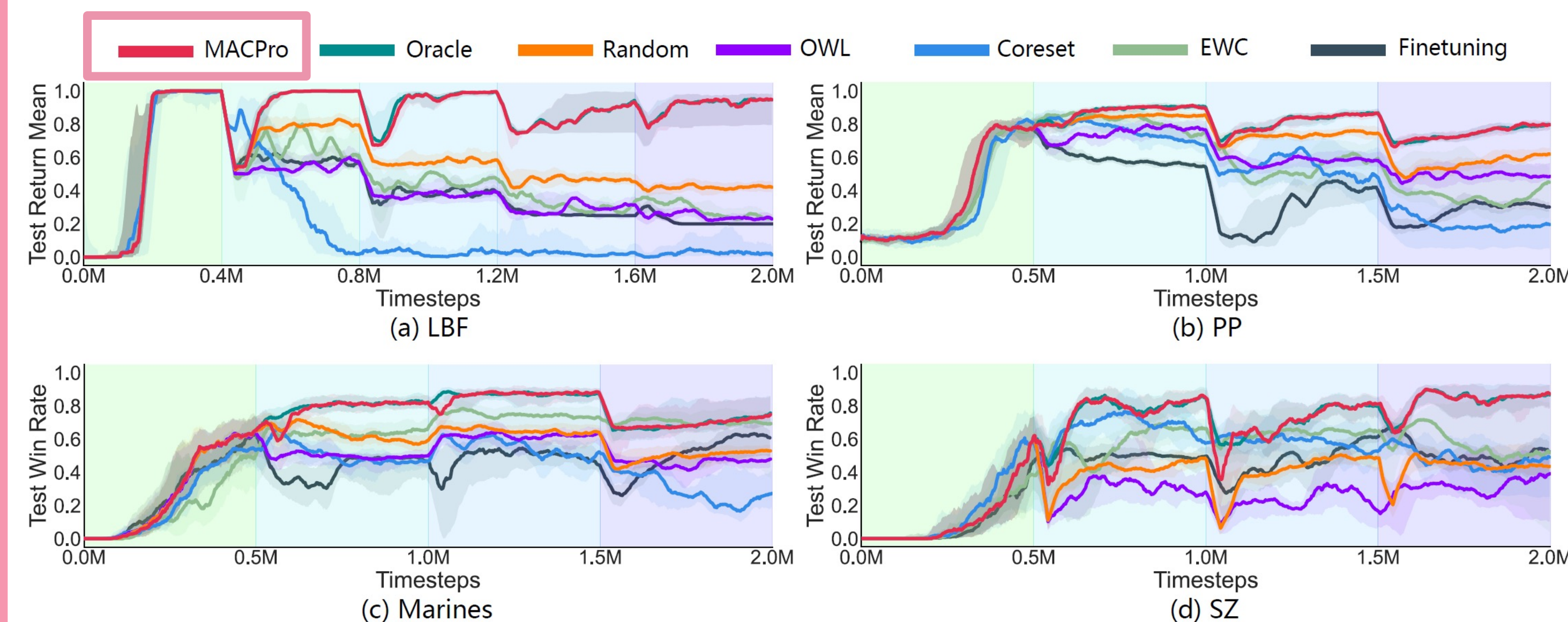Cooperative MARL

Continual Coordination

Method

□ The overall workflow of our method MACPro

(b) Dynamic Network Expansion (a) Task Contextualization Learning (c) Decentralized Task Approximation & Execution

(a) Assign a head for each task according to its context and train the corresponding policy.

(b) Encode task trajectories into context space and distinguish different tasks with their contexts.

(c) Each agent selects the optimal head with only local information during decentralized execution.

Experiments

□ Performance on different benchmarks

Legend: MACPro, Oracle, Random, OWL, Coreset, EWC, Finetuning

(a) LBF (b) PP (c) Marines (d) SZ

□ MACPro achieves the **best** continual learning performance on all multi-agent benchmarks.

□ Swift knowledge transfer to new tasks (**plasticity**) while avoiding catastrophic forgetting (**stability**).

□ Visualization Analysis

Legend: MACPro (ours), # Heads of MACPro (ours), MACPro w/o adaptive expansion, # Heads of MACPro w/o adaptive expansion, Finetuning, # Heads of Finetuning

Merge task 3 to task 1 Expand a new head for task 6 Projection at the end
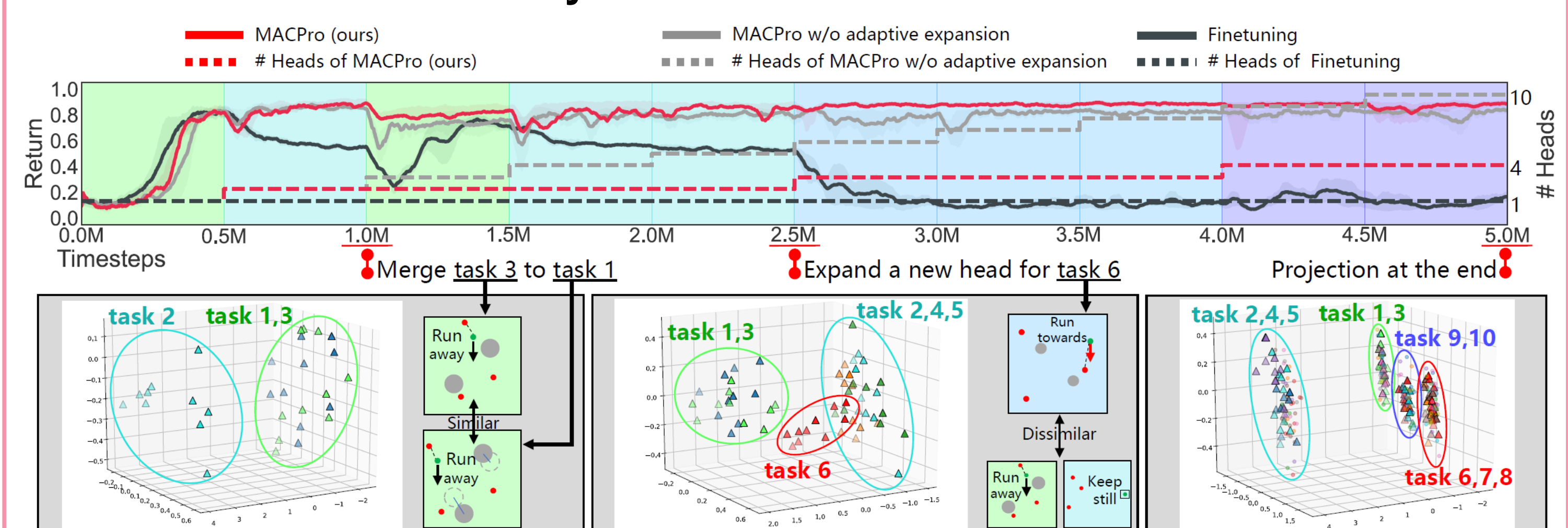
□ Expand a new head only for **dissimilar** tasks, without sacrificing performance.

□ Enjoy high **scalability** and learning **efficiency**.