

The goal of N-agent ad hoc teamwork (NAHT) is...

*To create a set of autonomous agents that are able to efficiently and robustly collaborate with previously unknown teammates on tasks to which they are all individually capable of contributing as team members.*



# N-Agent Ad Hoc Teamwork

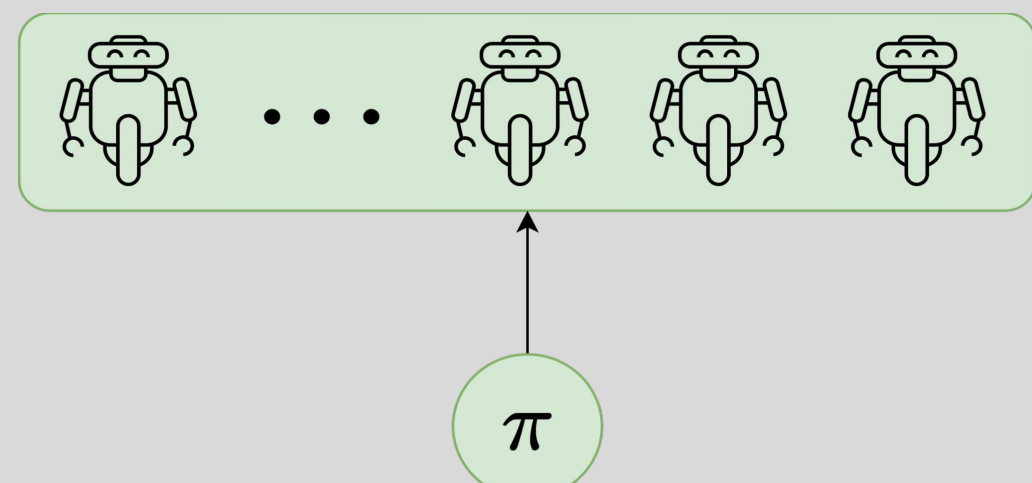
Caroline Wang<sup>1</sup>, Arrasy Rahman<sup>1</sup>, Ishan Durugkar<sup>2</sup>, Elad Liebman<sup>3</sup>, Peter Stone<sup>1,2</sup>

Corresponding author: caroline.l.wang@utexas.edu

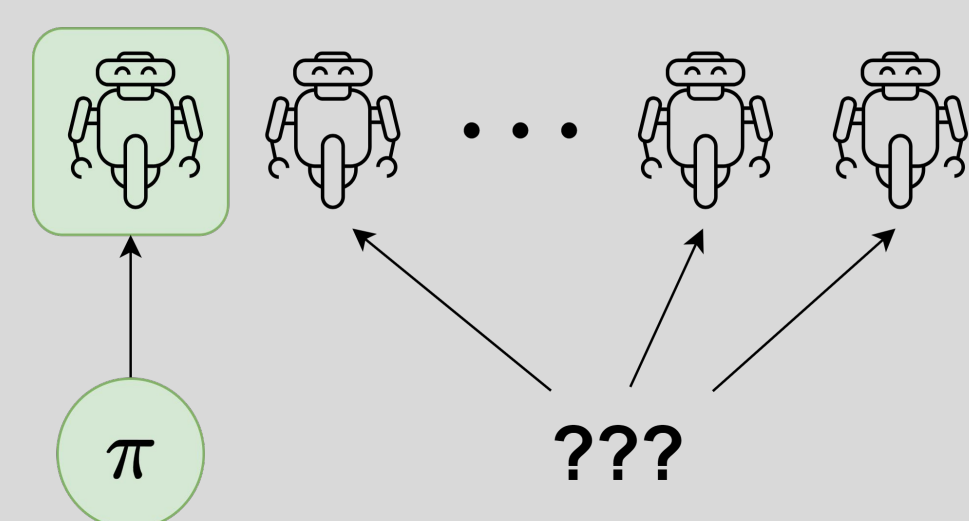
<sup>1</sup>The University of Texas at Austin, <sup>2</sup>Sony AI, <sup>3</sup>Amazon Science

## Background & Motivation

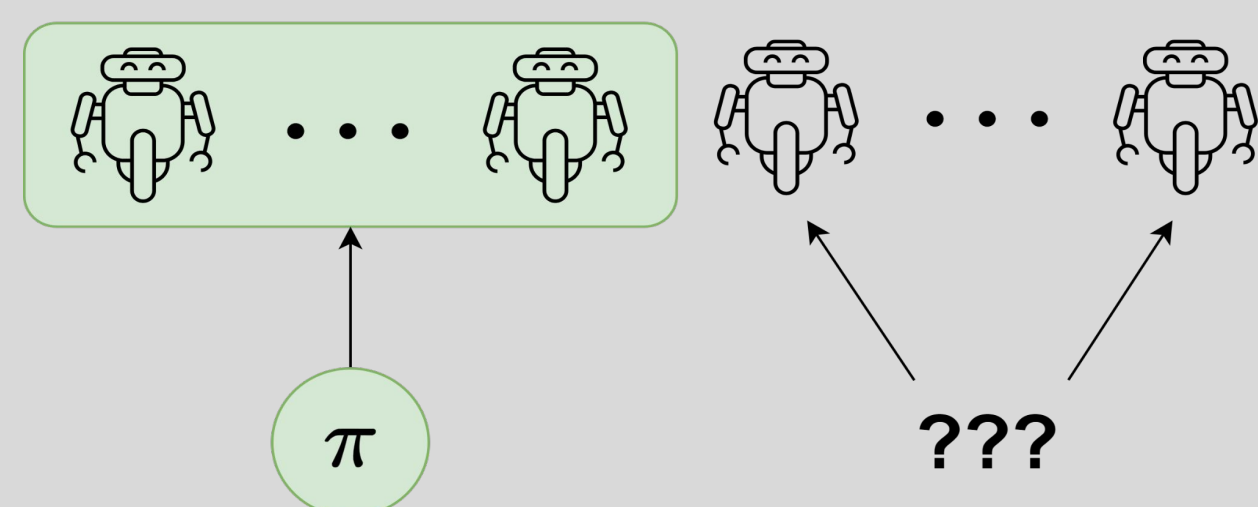
- **Cooperative MARL** (Foerster et al., 2018; Rashid et al., 2019): assumes all agents are under control of learning algorithm



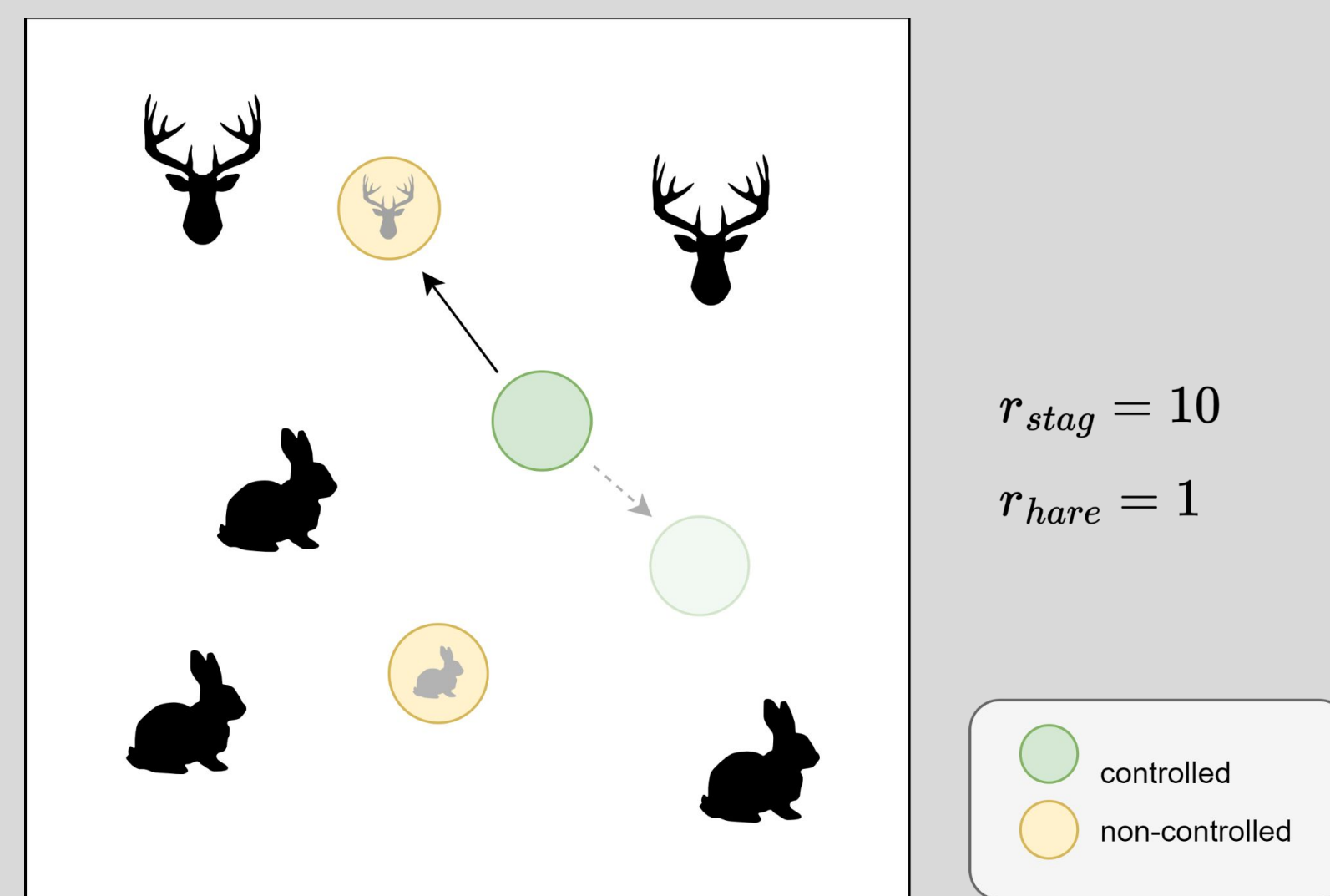
- **Zero Shot Coordination (ZSC) & Ad Hoc Teamwork (AHT)** (Hu et al., 2020; Mirsky et al., 2022): assumes one agent under control of learning algorithm



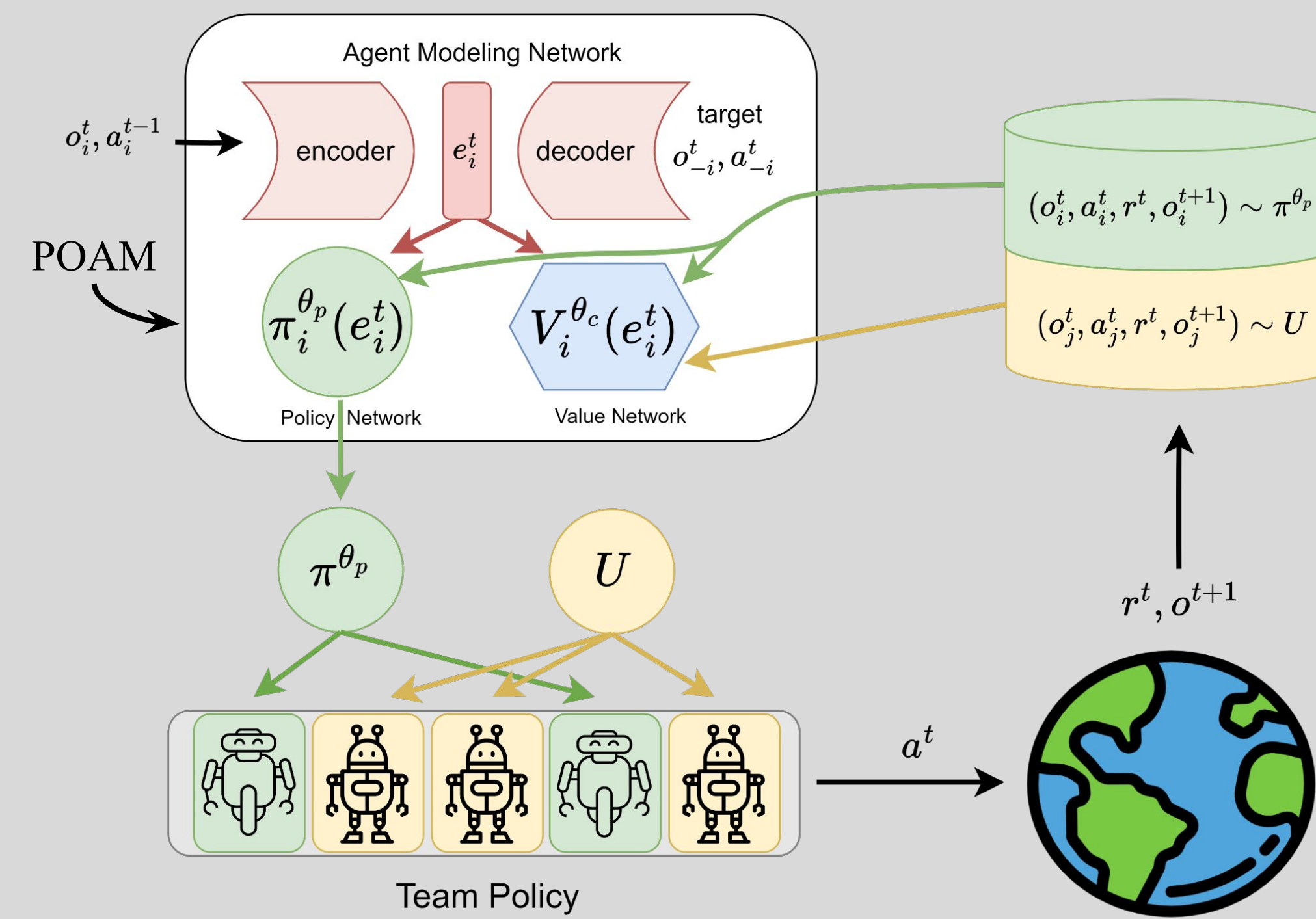
- **N-Agent Ad Hoc Teamwork (NAHT)**: how can sets of agents coordinate with each other?



## AHT vs NAHT



## Policy Optimization via Agent Modeling



## Experiments

- **Domains:** StarCraft II, Multi-agent Particle Environment (MPE) - predator prey
- **Baselines:** IPPO-NAHT, naive MARL, POAM-AHT

Figure 1: Test returns of IPPO-NAHT vs POAM against unseen teammates on predator prey.

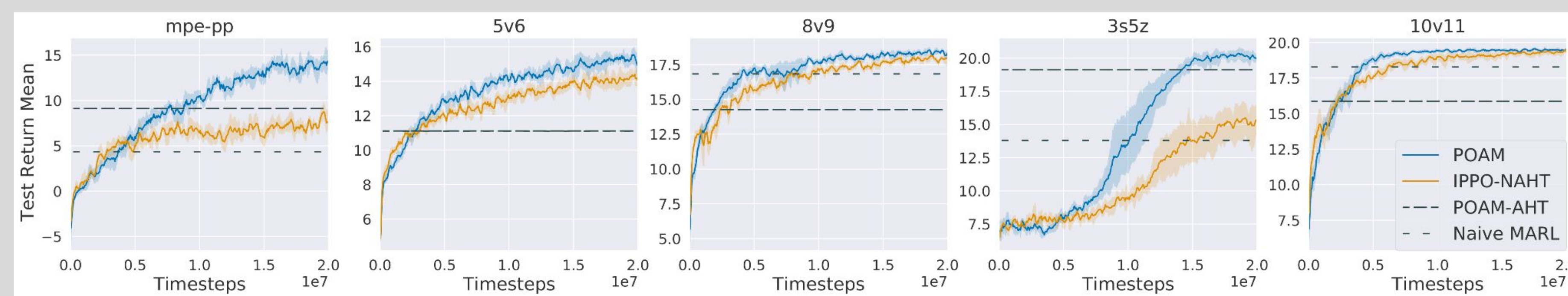
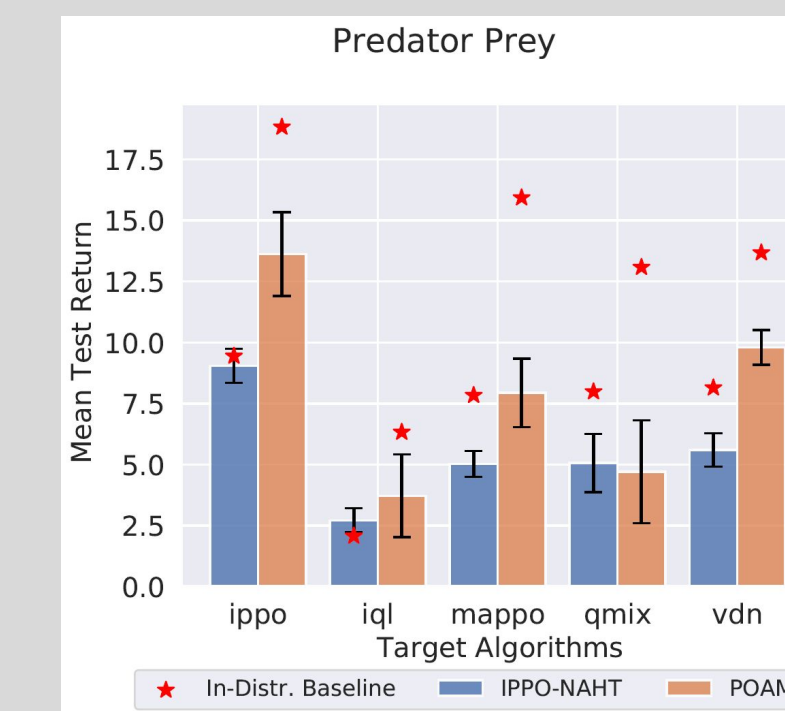


Figure 2. Sample efficiency of POAM vs IPPO-NAHT.

## NAHT Problem Formulation

### Dec-POMDP

$$(M, \mathcal{S}, \mathcal{A}, \mathcal{O}_i, \mathcal{T}, r, T) + (C, U, \mathcal{X})$$

$C(\theta)$ - set of controlled agents, parameterized by  $\theta$

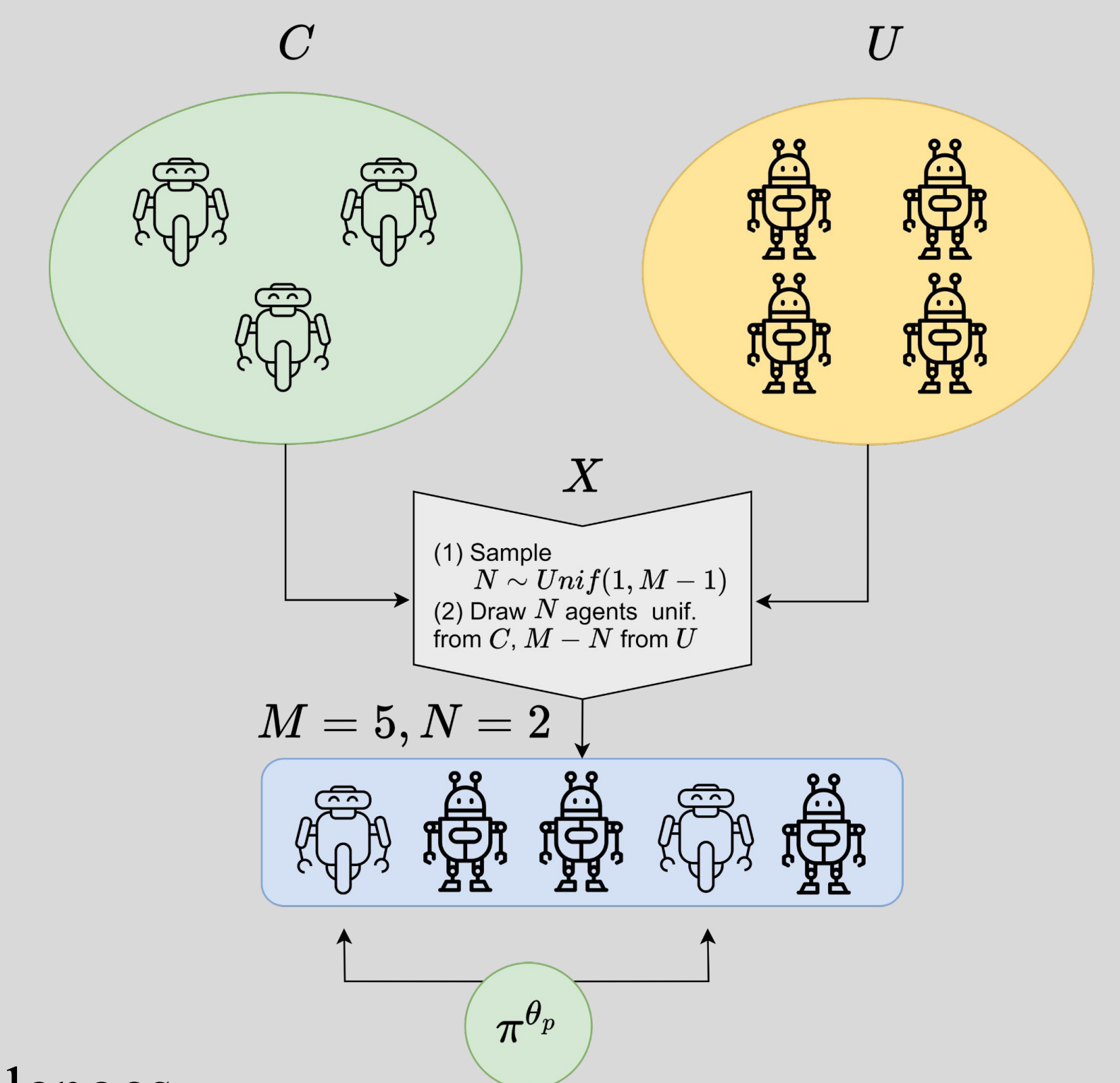
$U$  - set of uncontrolled agents

$\mathcal{X}$  - team sampling procedure

**Objective:**

$$\max_{\theta} \left( \mathbb{E}_{\pi^{(M)} \sim \mathcal{X}(U, C(\theta))} \left[ \sum_{t=0}^T \gamma^t r_t \right] \right)$$

### Team Sampling Procedure



### Challenges

- 1) **Generalization:** coordinating with non-controlled and potentially unknown teammates
- 2) **Openness:** coping with an unknown number of controlled teammates

## Future Directions

**Goal:** to enable multi-agent teams to efficiently cooperate with an unknown number of unseen agents.

- Explore teammate generation algorithms for NAHT
- Develop a benchmark for NAHT

