ICML International Conference On Machine Learning

Controlling Behavioral Diversity in Multi-Agent Reinforcement Learning

UNIVERSITY OF CAMBRIDGE

Matteo Bettini, Ryan Kortvelesy, Amanda Prorok

Department of Computer Science and Technology, University of Cambridge

PROROKLAB

Website

Multi-agent and multi-robot cooperative tasks often need or benefit from behavioral heterogeneity.

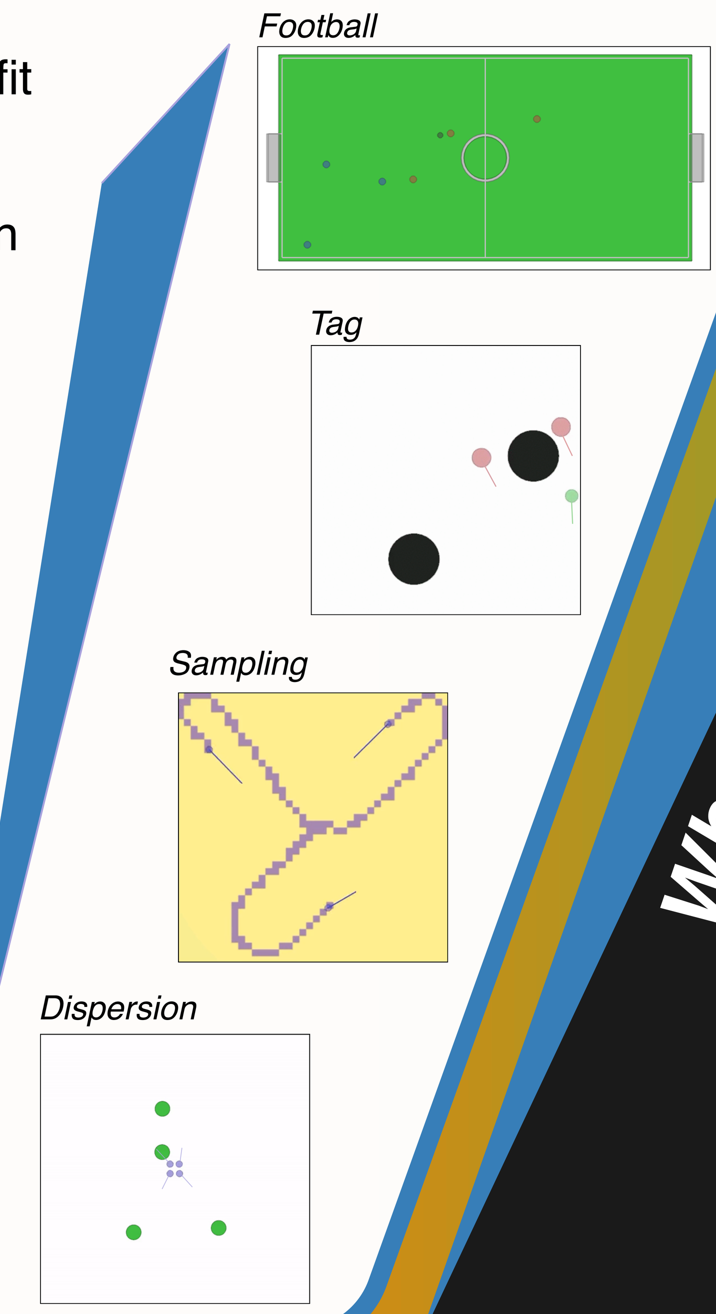
Diversity has been shown to be **key to collective intelligence** in both natural and artificial systems.

Despite this, **existing approaches either do not consider diversity or blindly boost it by using**:

- Intrinsic rewards
- Additional loss functions

Thus effecively modifying the learning objective and having no principled measure for it.

We introduce *Diversity Control (DiCo),* ***the first method able to control diversity to an exact value*** *or range of a given metric* ***without changing the learning objective.***



Why care?

The role of diversity in collective intelligence

TL;DR

Diversity is key to collective intelligence

↓

Existing approaches promote it blindly via:

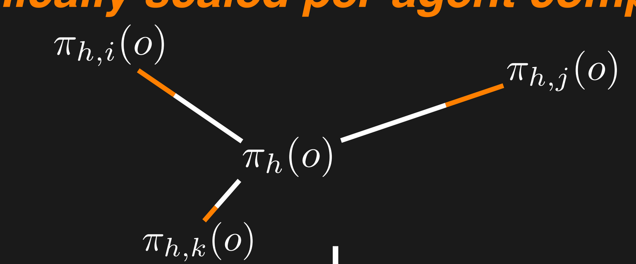
- Intrinsic rewards
- Additional loss functions

↓

We introduce ***Diversity Control (DiCo)***, the first method method ***to control diversity with no additional learning objective***

↓

We do this by representing policies as the sum of a ***parameter-shared component*** and ***dynamically scaled per-agent components***



↓

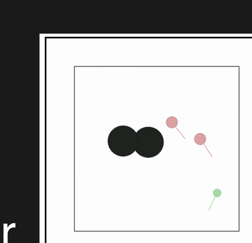
We provide theoretical proofs and demonstrate it empirically in a didactic ***case study***

↓

Our experiments show how DiCo can betemployed to ***boost performance and sample efficiency in MARL***

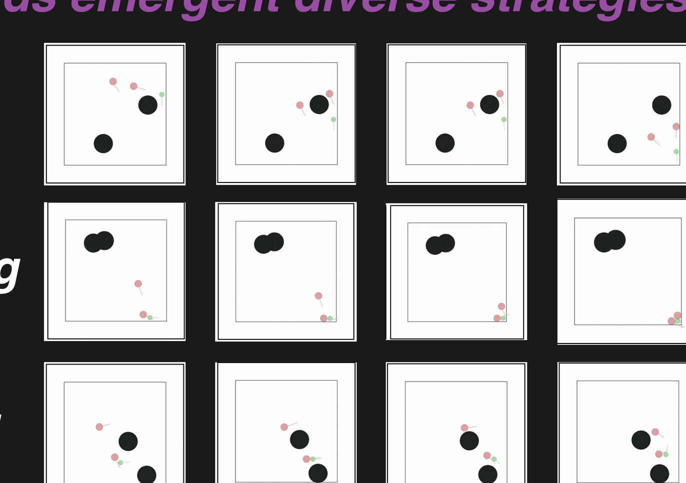
↓

In predator-prey tasks, where homogeneous and unconstrained heterogeneous predators all blindly chase after the prey

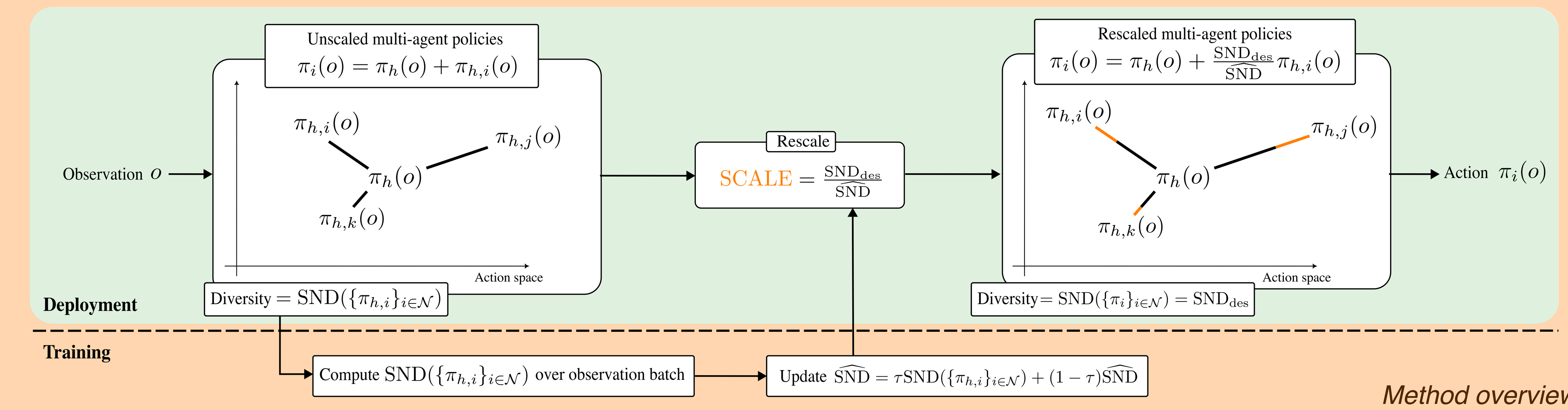


↓

DiCO finds emergent diverse strategies



Why care? · Method · Case study · Experiments



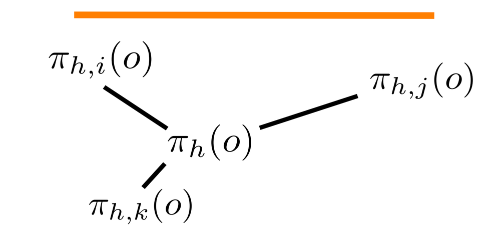
Method overview

We use System Neural Diversity $\text{SND}(\{\pi_i\}_{i\in\mathcal{N}})$ to measure behavioral diversity

Divesrity Control (DiCo)

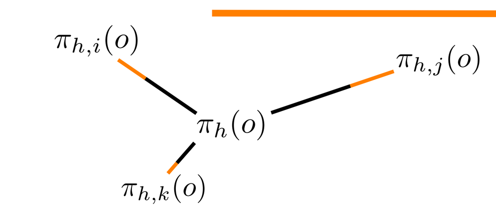
① We represent multi-agent policies as **heterogeneous deviations** $\pi_{h,i}(o)$ from a **homogeneous reference** $\pi_h(o)$

$$\pi_i(o) = \pi_h(o) + \pi_{h,i}(o)$$



② Policies are then **rescaled dynamically** w.r.t. the homogeneous reference, given the **desired diversity** SND_{des} and their **current diversity** $\widehat{\text{SND}}$

$$\pi_i(o) = \pi_h(o) + \frac{\text{SND}_{\text{des}}}{\widehat{\text{SND}}}\pi_{h,i}(o)$$



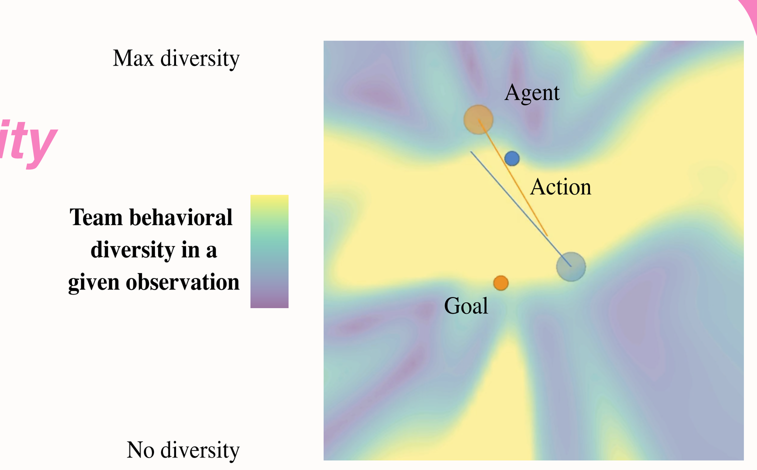
Method

Dynamically rescaling policies

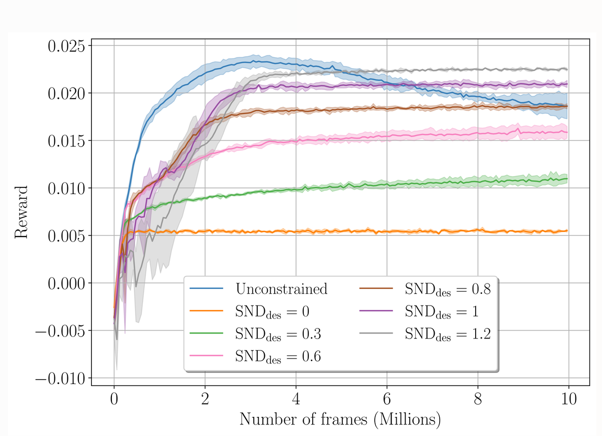
Case study

A simple example of controlling diversity

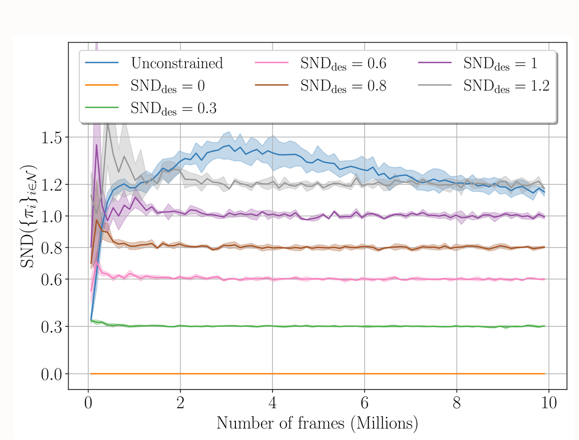
We demonstare our method in a **Multi-Agent Navigation** task, where each **agent observes all goals and has to navigate to a specific one.** Thus, agents **need to be diverse** to go to different goals.

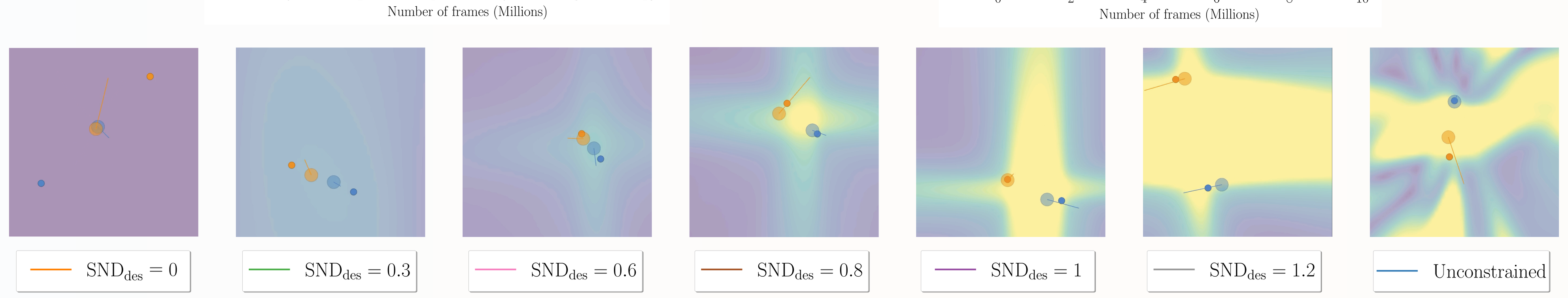


Reward



Diversity





Experiments

Improving MARL sample efficiency and performance

We run experiments, in **cooperative and competitive** tasks, that show how DiCo can be employed as a novel paradigm to **increase performance** and sample efficiency.

By controlling diversity, we can search the beahvioral space at different heterogeneity levels, leading to the **emergence of novel strategies** not found by unconstrained agents.

Agents constrained with high diversity find better policies

Unconstrained diverse agents settle to lower diversity