Heterogeneous Multi-Robot Reinforcement Learning

Matteo Bettini, Ajay Shankar, Amanda Prorok

Department of Computer Science and Technology, University of Cambridge

Paper

Video

Taxonomy

How is a system heterogeneous?

We introduce a **taxonomy to classify heterogeneous systems**

Robots can have *differences*:

- **Physical** differences $\mathcal{P}$
- **Behavioral** differences $\mathcal{B}$
  - Due to different objectives $\mathcal{B}_d$
  - Even with same objective $\mathcal{B}_s$

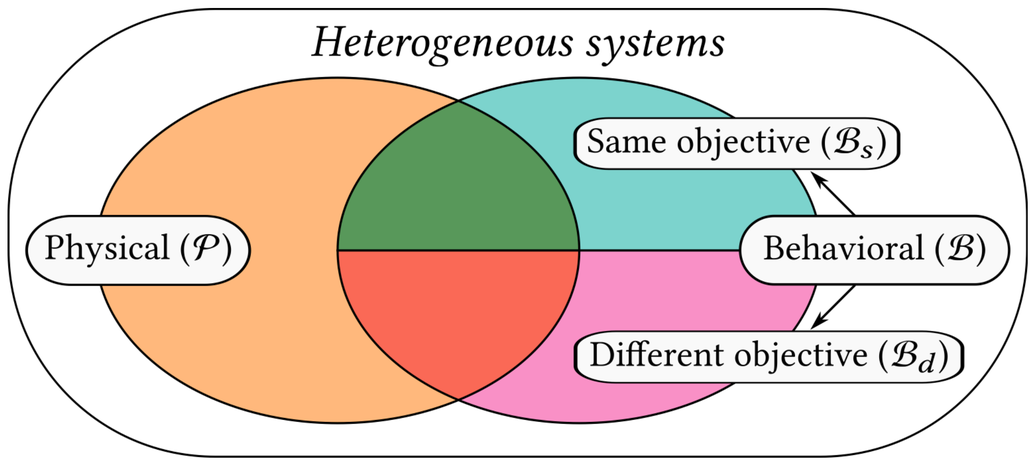
Agents are physically different but share the same behavioral model — $\mathcal{P} \setminus \mathcal{B}$

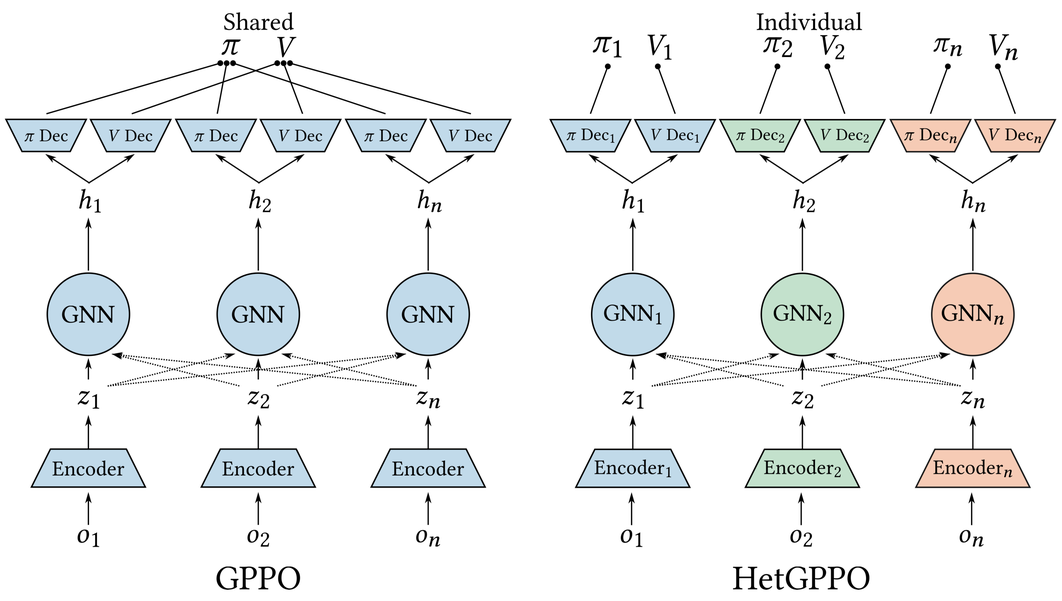
Agents are physically different and differ in behavioral models and objectives — $\mathcal{P} \cap \mathcal{B}_d$

Agents are physically different and differ in behavioral models, but share the same objective — $\mathcal{P} \cap \mathcal{B}_s$

Agents are physically identical and share the same objective, but differ in behavioral models — $\mathcal{B}_s \setminus \mathcal{P}$

Agents are physically identical but differ in behavioral models and objectives — $\mathcal{B}_d \setminus \mathcal{P}$





Each agent has:

- a stochastic **policy** $\pi_i(a_i|o_{\mathcal{N}_i})$
- a **value** function $V_i(o_{\mathcal{N}_i})$

For a given observation o, for all robots i, j in the system

$\pi_i(o) = \pi_j(o)$ GPPO

$\pi_i(o) \neq \pi_j(o)$ HetGPPO

Model

How we learn heterogenous policies

We introduce **GPPO** and **HetGPPO**, two actor-critic models for **Multi-Agent Reinforcement Learning**

HetGPPO learns *individual* agent policies

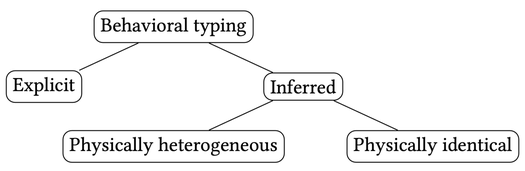
- Uses neighborhood communication to **overcome partial observability**
- Allows **decentralized training** of Graph Neural Networks (GNNs)

Behavioral typing

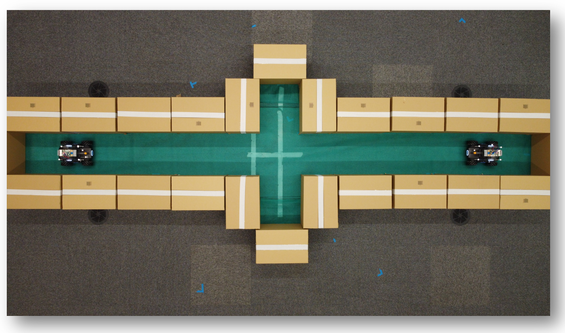
How homogeneous robots emulate heterogeneous behavior

We find that **homogeneous robots are able to infer behavioral roles** through observations, emulating heterogeneous behaviors

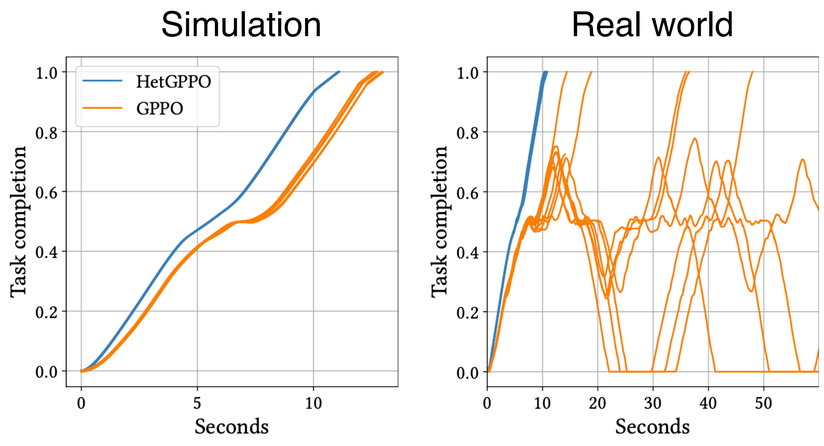
They **encode multiple roles in the same policy**, activating them based on the input observation



Behavioral typing is disrupted by sim-to-real transfer



Task: *Give Way*
Robots need to switch places in a narrow corridor



Results

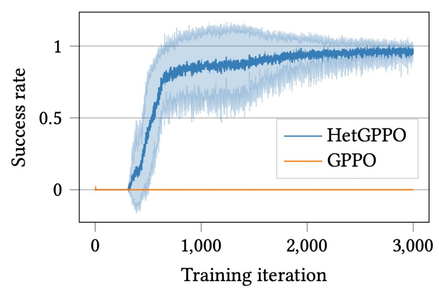
Tasks where heterogeneous robots achieve better performance and resilience

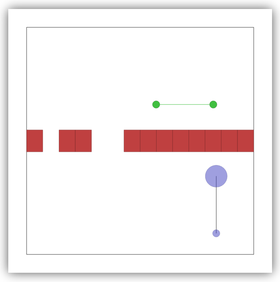
We perform evaluations of our heterogeneous (**HetGPPO**) and homogeneous (**GPPO**) models in **multi-robot cooperative scenarios**

We demonstrate the benefits of heterogeneous behavior in terms of **performance** and **resilience** in tasks with a shared global objective ($\mathcal{B}_s$)

Our results highlight the need for heterogeneity to achieve **collective intelligence**

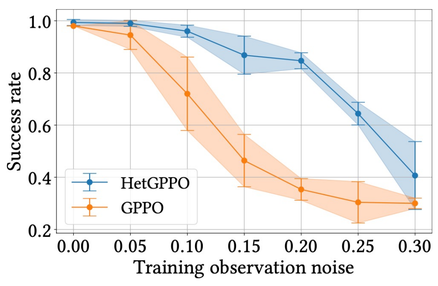
*Heterogeneous behavior is **vital for performance***

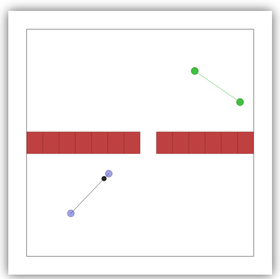




Task: *Different Size Joint Passage*
Different sized robots, connected by a linkage through revolute joints, need to cross a passage with different sized gaps while keeping the linkage parallel to it.

*Heterogeneous behavior **achieves better resilience***





Task: *Asym. Payload Joint Passage*
Physically identical robots, connected by a linkage through revolute joints, need to cross the passage while transporting a payload placed on one side of the linkage.

UNIVERSITY OF CAMBRIDGE

This work was supported by ARL DCIST CRA W911NF-17-2-0181, the European Research Council (ERC) Project 949940 (gAIa), and in part by a gift from Arm.

PROROKLAB