



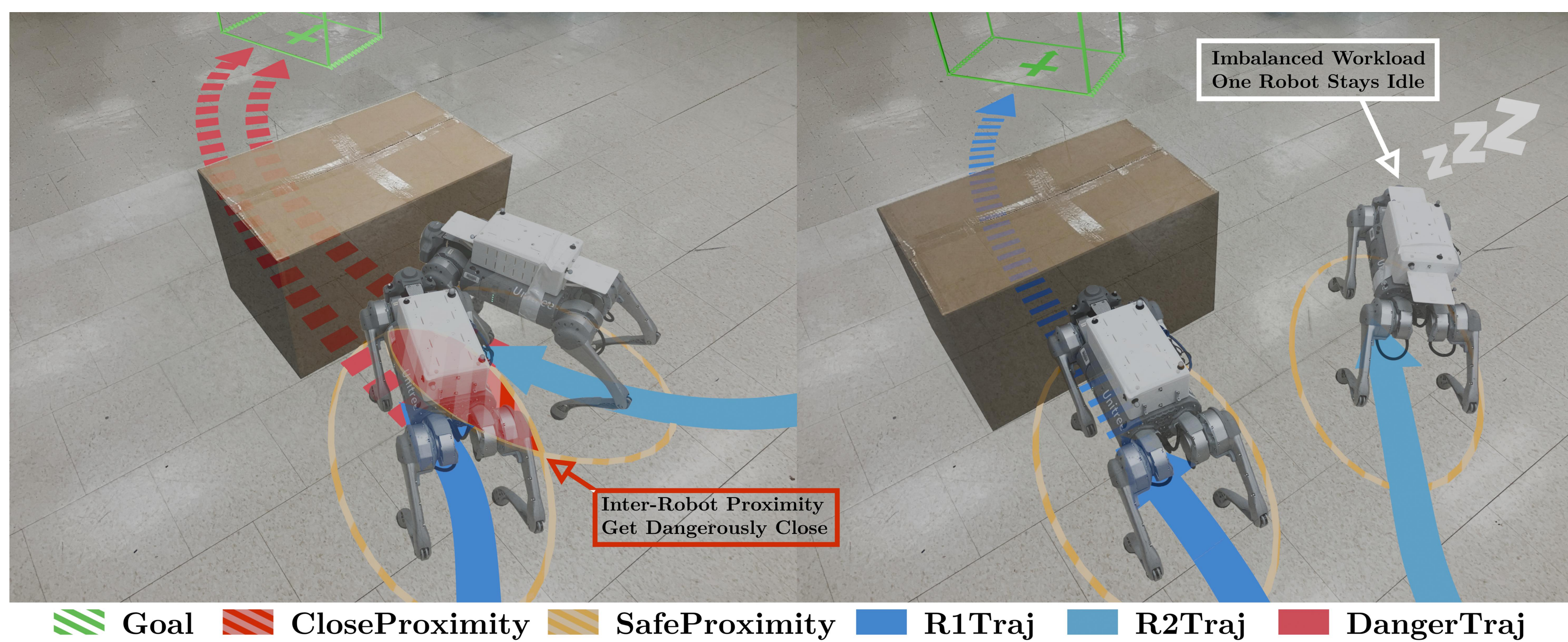
Safe and Fair Multi-Quadruped Collaborative Box Pushing

Yuhao Li, Williard Joshua Jose and Hao Zhang

Human-Centered Robotics Lab, University of Massachusetts Amherst, USA

Motivation

- Multi-quadruped collaborative box pushing must satisfy three needs at once: task progress, safe inter-robot proximity, and balanced effort.
- Reward designs that emphasize task-only progress collapse into one of two characteristic failure modes – role imbalance on easy tasks, unsafe crowding on hard tasks.



Approach

- We add two team-level reward terms to the per-agent task-progress reward and optimize under shared-parameter MAPPO.
- Both shaping terms are normalized so task progress, proximity safety, and effort fairness contribute comparable magnitudes.

$$r_i(t) = r_{\text{push},i}(t) - \lambda_{\text{prox}} \phi_{\text{prox}}(t) - \lambda_{\text{fair}} \phi_{\text{fair}}(t)$$

Task progress

Cosine alignment of box velocity with the goal direction, gated when the box is near-stationary.

Proximity safety

Smooth inverse-distance penalty over all robot pairs. Penalises tight formations before contact.

Effort fairness

Pairwise gap between robot-to-box distances. Discourages single-robot pushing regimes.

- Training: centralized critic + decentralized actor (CTDE), shared-parameter MAPPO, 2048 parallel envs.

Policy interface

Each robot observes box, goal, and teammate features in its local frame, then outputs planar twist commands (v_x , v_y , ω).

Team coupling

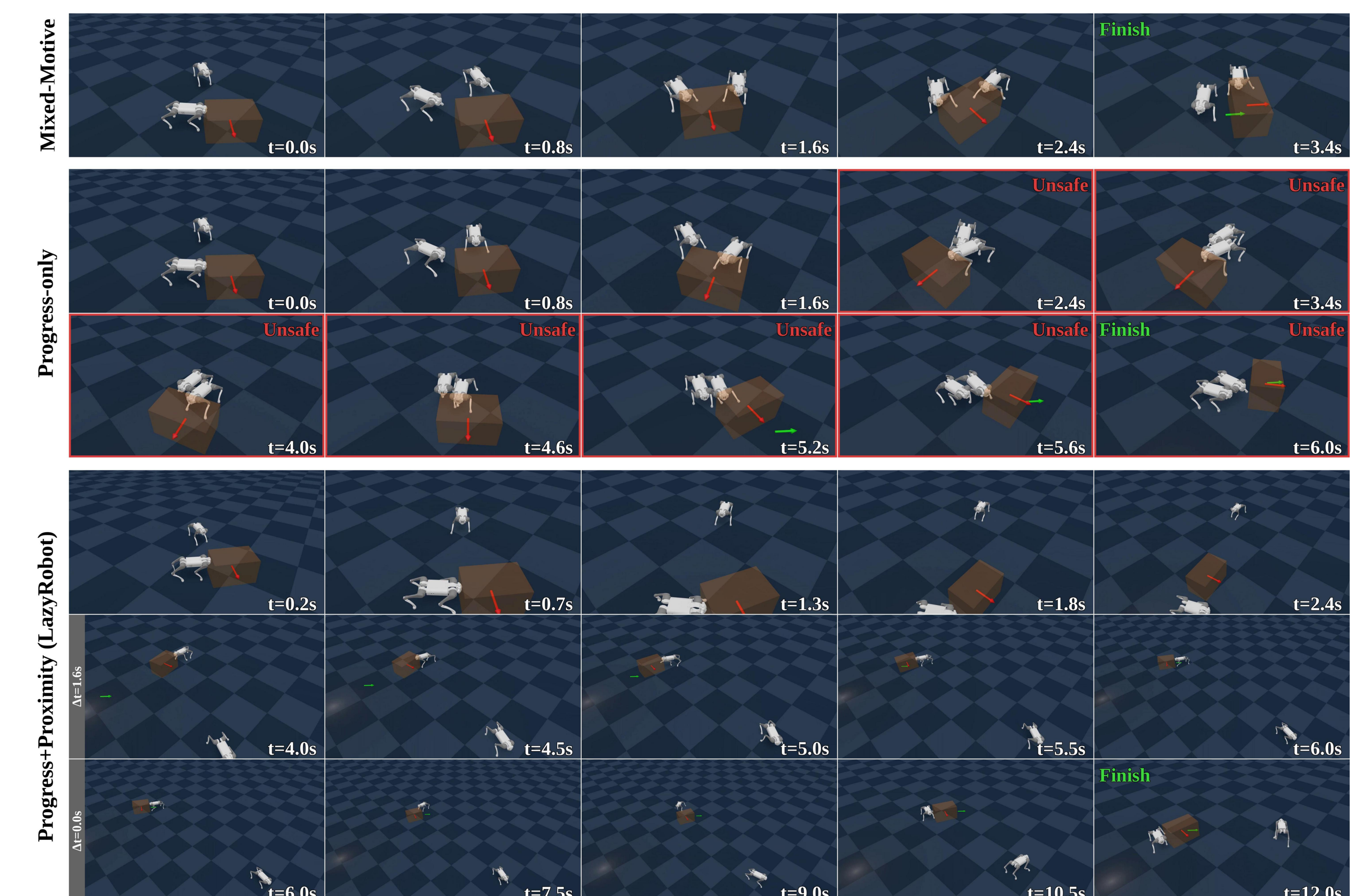
Proximity and fairness penalties are shared by all agents, aligning decentralized actions with team safety and workload balance.

Experiments

- Simulation eval uses fixed randomized test seeds and the same environment / CTDE-MAPPO setup for all variants. Bins: Easy 30-60° (1x diag), Normal 90-120° (2x), Hard 150-180° (4x).
- 4-way reward ablation (same backbone, env, seeds): P · P+Proximity · P+Fairness · Full-Shaping (ours).

Bin	Method	Success ↑	Comp. Time (s) ↓	Unsafe Prox. % ↓	Single-Robot % ↓
Easy	Progress-only	54	1.6	10.1	51.8
	P+Proximity	54	5.0	0.0	47.1
	P+Fairness	72	1.4	7.5	47.6
	Full-Shaping	86	1.8	3.4	44.5
Normal	Progress-only	64	4.0	53.7	49.4
	P+Proximity	32	9.6	0.1	51.3
	P+Fairness	22	3.6	26.7	29.3
	Full-Shaping	68	3.4	1.1	25.6
Hard	Progress-only	18	8.2	54.4	30.1
	P+Proximity	4	13.2	4.3	55.1
	P+Fairness	4	12.5	45.8	24.4
	Full-Shaping	26	11.3	18.7	26.5

Metrics: Success over all episodes; other columns over successful episodes. Unsafe Prox. = any robot pair too close; Single-Robot = exactly one robot near the box.

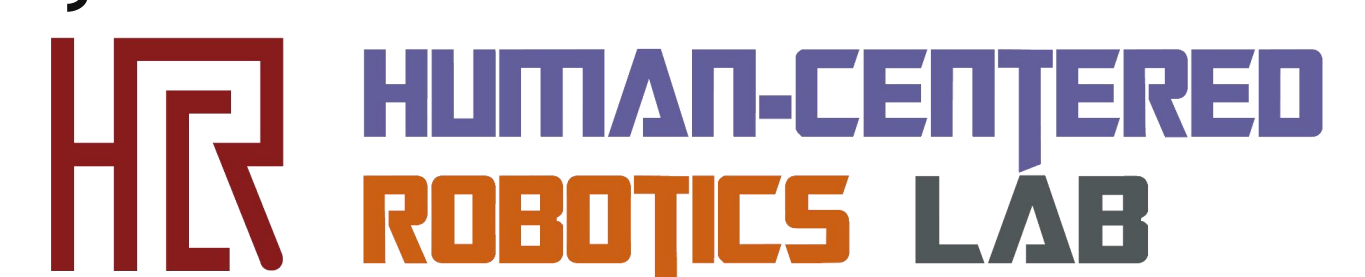


Summary

- Proximity safety and effort fairness deserve primary-objective status.
- Single-term shaping backfires: proximity-only induces lazy disengagement; fairness-only induces crowded interference.
- Full-Shaping reaches the strongest success across all difficulty bins.

Contact

Yuhao Li
Human-Centered Robotics Lab
CICS @ UMass Amherst
yuhao.li@umass.edu



ICRA 2026 Workshop