HARMONY: Multi-Humanoid Motion Coordination

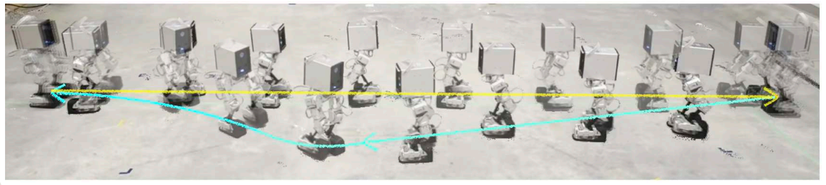
via Diffusion Planning & Learning-Based Control

Lufan Yang, Dun Yang, Yanshu Zhang, Jiaqi Tan, Ke Li, Hang Ma
School of Computing Science, Simon Fraser University

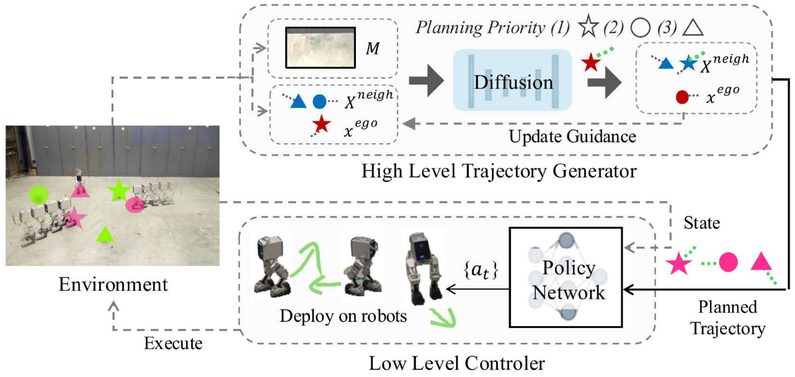
Motivation

Coordinating many legged humanoids in shared spaces is hard: high-level multi-robot planners ignore body dynamics, while low-level whole-body controllers cannot reason about neighbouring agents.

We propose HARMONY, a closed-loop framework that couples a guided diffusion planner with a learned whole-body RL controller — using the controller's value function as a continuous feasibility prior at test time — achieving collision-free coordination of up to 100 humanoids in simulation and successful hardware deployment on 2 / 3 / 4 real Mini-π bipeds.



Closed-Loop Framework



Guided Diffusion Planner

Each agent denoises a trajectory $\bar{\tau}$ conditioned on its own history, neighbors, and a local map. Test-time guidance shifts the predicted mean by a composite gradient — **no retraining needed**:

$$L_{total} = \lambda_{goal} \cdot L_{goal} + \lambda_{map} \cdot L_{map} + \lambda_{agent} \cdot L_{agent} + \lambda_{value} \cdot L_{value}$$

- **Goal:** softmin-weighted distance to target (or forward-progress when out of reach).
- **Map:** distance-based hinge penalizing penetration of static obstacles.
- **Agent:** distance-based hinge penalizing penetration of other agents (see next card for asymmetric responsibility).
- **Value:** L_value = (1/T_f) Σ exp(−V(ô_j)) — the critic's expected return at each waypoint.

Priority-Aware Collision Avoidance

Symmetric reciprocal avoidance (Trace, MMD) yields conflicting gradients → **oscillation and deadlock**. HARMONY adopts an asymmetric protocol inspired by Prioritized Planning:

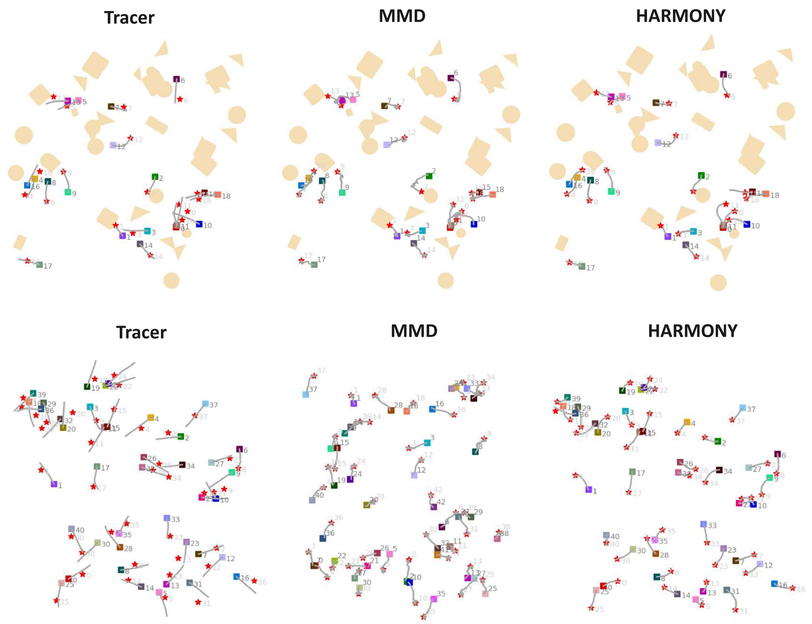
1. Sample priority ordering σ over N agents.
2. Higher-priority agent σ(j) plans first.
3. Its trajectory becomes a fixed dynamic obstacle for σ(j+1).
4. Repeat K times, keep the lowest-loss plan.

Each agent has **unambiguous** avoidance responsibility — gradients no longer cancel. K ∈ {5, 10} randomized orderings trades compute for quality (HARMONY-PP-k).

Open-Loop Planning Results

Method	Random Obstacles Random Obstacles		Empty Empty	
	Dev ↓	Path ↓	Dev ↓	Path ↓
Tracer	0.632	0.163	1.067	0.185
MMD	0.522	0.350	0.455	0.297
HARMONY	**0.209**	**0.137**	**0.112**	**0.141**
HARMONY-PP-5	**0.191**	**0.136**	**0.093**	**0.141**
HARMONY-PP-10	**0.189**	**0.136**	**0.087**	**0.141**

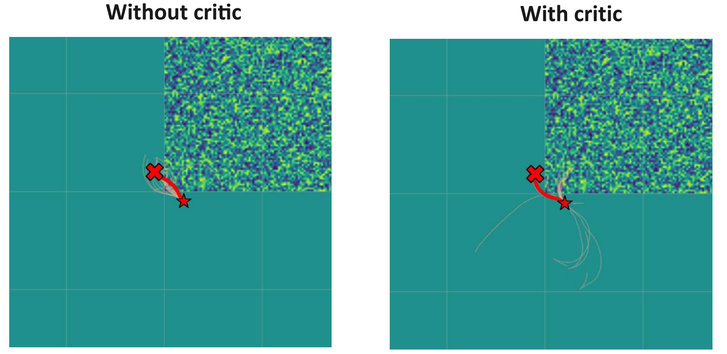
100 procedural scenes. Dev = mean distance to goal, Path = avg path length (lower is better).



Same scene, three planners. *Tracer collides at intersections; MMD detours unnaturally; HARMONY produces direct, collision-free arcs.*

Value Function as Dynamics Guidance

The RL critic gives a **continuous**, robot-specific feasibility signal.

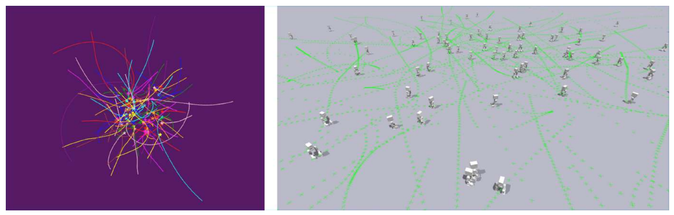


Bumpy terrain, start (★) → goal (✕). Without critic: cuts through rough ground. With critic: detours around.

Closed-Loop Results

Simulation (100 agents, open field)

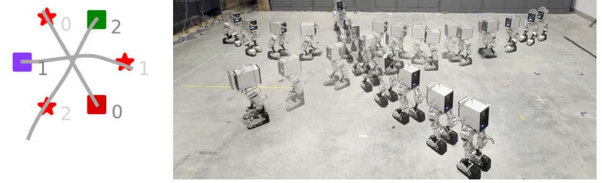
Left: planner trajectories at a later re-planning cycle. Right: execution on bipedals in simulation.



Real-World (Mini-π bipeds, 1 kHz PD torque)

Each row: planned trajectories (left) | motion-trail of executed paths (right).

3 agents



4 agents