D3G: Learning Multi-robot Coordination from Demonstrations

Yizhi Zhou, Wanxin Jin and Xuan Wang

Motivation. The control and coordination of large-scale multi-robot systems have long been challenging, as traditional optimization approaches [1] are not directly applicable due to large joint states and real-world communication constraints. To address this issue, game theoretic formulations allow agents to achieve distributed coordination, where decision making is based solely on agents' local observations. Building on this, determining what coordination strategy agents should employ can be accomplished through differentiable programming for inverse learning [2]. However, existing approaches for inverse games [3] heavily rely on centralized training, which is computationally inefficient. Motivated by this, we develop a *Distributed Differentiable Dynamic Game (D3G)* framework equipped with new distributed Nash-seeking algorithms and inverse game gradient solvers. This framework enables robots to efficiently learn scalable coordination strategies from desired demonstrations and execute them, both in fully distributed settings.

Problem Statement. We represent the coordination of a multi-robot system as a parameterized dynamic game $\mathbf{P}(\Theta)$. This game consists of local optimal control problems $P_i(\theta_i)$, each possessed by a robot. Here, $P_i(\theta_i)$ describes the robot's unknown dynamics and its tunable objective function, the value of which also depends on other robots' behaviors. Building on this, we associate robot coordination with the Nash Equilibrium (N.E.) of $\mathbf{P}(\Theta)$, which can be adapted by tuning the objective and dynamics of each robot through θ_i. By developing the D3G framework, our objective is to allow each robot to automatically tune θ_i in a distributed manner by minimizing loss function $\mathcal{L}_i(\cdot)$, which represents the mismatching between the desired demonstrations and the robots' trajectories. In this way, robots will learn and reconstruct the embedded coordination strategies.

Main Results. As visualized in Fig 1, the learning scheme of D3G features a new design. It contains a forward-pass of computing the N.E. of a dynamic game for which we develop a distributed shooting-based method for numerical efficiency. In the backward-pass, we first associate the gradient of N.E. with a Differential Pontryagin's Maximum Principle condition and then propose a distributed solver to compute the gradient. All inter-robot communication happens locally among connected neighbors, eliminating the need for global information exchange or centralized coordination.

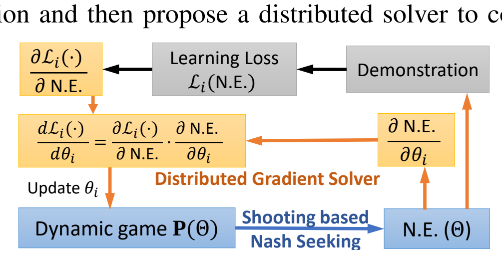

Fig. 1: D3G: learning multi-robot coordination from demos.

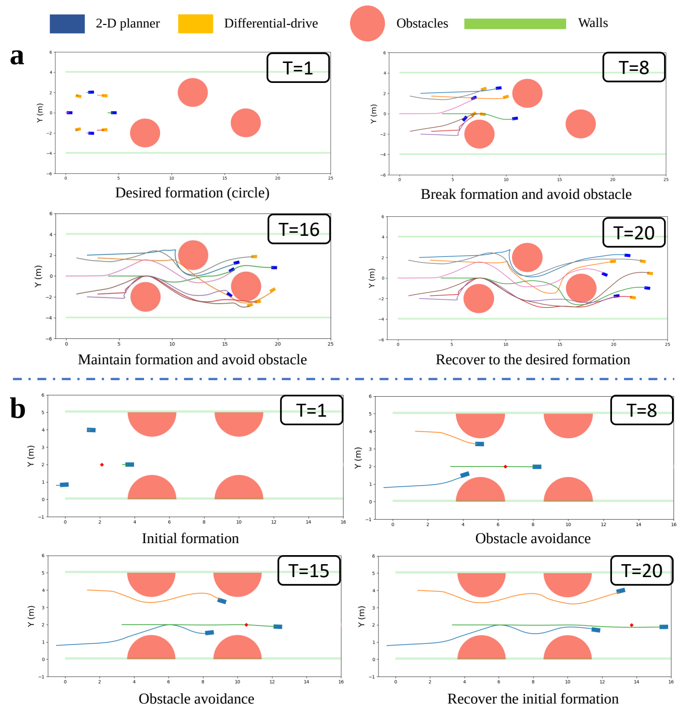

Fig. 2: Simulated experiments for two scenarios.

We validate the effectiveness of D3G through experiments:

Scenario (a). Eight heterogeneous robots including four differential-drive and four planners aim to maintain a desired (circle-like) formation through navigation. Based on D3G, each robot learns a local objective function that accounts for both collision avoidance and formation maintenance. In the provided demonstrations, the robots navigate through narrow spaces among the obstacles while, to the maximum extent possible, maintaining the desired formation.

Scenario (b). Three robots cooperatively transport a slung payload to a final destination. Throughout this process, the robots need to learn objective functions that allow for maintaining clearance of the goods from the ground, reducing sling tension, and handling obstacle avoidance. The robots successfully accomplished the task based on D3G framework.

References

[1] T. Halsted, O. Shorinwa, J. Yu, and M. Schwager, "A survey of distributed optimization methods for multi-robot systems," *arXiv preprint arXiv:2103.12840*, 2021.

[2] W. Jin, Z. Wang, Z. Yang, and S. Mou, "Pontryagin differentiable programming: An end-to-end learning and control framework," *Advances in Neural Information Processing Systems*, vol. 33, pp. 7979–7992, 2020.

[3] K. Cao and L. Xie, "Game-theoretic inverse reinforcement learning: A differential pontryagin's maximum principle approach," *IEEE Transactions on Neural Networks and Learning Systems*, 2022.