Embodied Team Intelligence in Multi-Robot Systems

Esmaeil Seraj
Georgia Institute of Technology
Atlanta, GA, United States of America
eseraj3@gatech.edu

ABSTRACT
High-performing human teams leverage intelligent and efficient communication and coordination strategies to collaboratively maximize their joint utility. Inspired by teaming behaviors among humans, I seek to develop computational methods for synthesizing intelligent communication and coordination strategies for collaborative multi-robot systems. I leverage both classical model-based control and planning approaches as well as data-driven methods such as Multi-Agent Reinforcement Learning (MARL) to provide several contributions towards enabling emergent cooperative teaming behavior across both homogeneous and heterogeneous (including agents with different capabilities) robot teams. In future work, I aim to investigate efficient ways to incorporate humans’ teaming strategies for robot teams and directly learn team coordination policies from human experts.

KEYWORDS
Multi-Agent Systems; Distributed Control and Planning; Collaborative Decision Making; MARL; Heterogeneous Teaming

2 MODEL-BASED MULTI-AGENT PLANNING AND CONTROL METHODS
To enable a highly efficient and intelligent team behavior, in Seraj and Gombolay [12], I design a low-level control strategy toward a human-centered robot coordination. Such a system is desired in a variety of applications; as an example, deploying a fleet of Unmanned Aerial Vehicles (UAVs) to actively monitor a propagating wildfire in support of human firefighters on the ground. In [12], I propose a decentralized control framework that leverages a model-predictive mechanism to coordinate a UAV team for tracking the moving firefronts while simultaneously enabling the firefighters to receive online information regarding their time-varying proximity to fire. To facilitate our multi-faceted objective, we develop a dual-criterion objective function based on Kalman uncertainty residual and weighted multi-agent consensus protocol. Our simulated evaluations as well as physical robot experiments in a multi-robot testbed demonstrated efficacy of our framework and a significant team coordination boost over prior model-based and RL-based methods.

An issue that we faced in [12] was that a connected communication graph was required at all times for the control architecture to work properly. The team’s communication network, however, may be disconnected at times and its links may have varying strengths due to environment constraints. In Seraj et al. [10], we tackle this problem by designing a model-reference multi-agent adaptive controller that achieves team convergence even for a network of robots with a disconnected communication graph. We derive an adaptive control law for a leader-follower networked system that provably converges to mimic any desired network structure even though the real communication topology remains unknown to the robots.
My works in [10] and [12] enable adaptive, decentralized, and multi-objective controllers to coordinate a team of robots at the low-level control input to collectively perform active target tracking and field monitoring tasks. A natural next step then is to study the team coordination and collaboration strategies at the high-level planning/scheduling and decision-making stages. I studied the problem of coordinated planning of a multi-UAV team for cooperative surveillance and tracking of a restless environment in [13] (currently under peer-review in JAAMAS). In Seraj et al. [13], I utilized a similar model-predictive mechanism as in [12] to enable UAVs with the ability to reason about their cooperation plan for collaborative surveillance through actively estimating the changing states of their environment. Particularly, we consider safety-critical and time-sensitive scenarios where only a limited number of UAVs are available for allocation. A central contribution of our work in [13] is a set of analytical temporal and tracking-error bounds that allow UAVs to enable probabilistically-guaranteed coordination in tracking dynamic targets. Our quantitative evaluations validate the performance of our method accumulating 7.5x and 9.0x smaller tracking-error than two model-based and RL benchmarks.

Next, in an attempt to simultaneously tackle both the high-level planning and the low-level control stages of the coordination hierarchy, in Seraj et al. [11], I develop an efficient hierarchical coordination framework for a composite robot team composed of perception-only and action-only agents. In a perception-action composite team, perception agents are first tasked to explore an unknown environment to find an initial set of dynamic targets. Estimated target-states are then communicated to action agents (unable to sense) to perform a specific manipulation on those targets [1, 11, 14, 15]. Accordingly, agents in such composite teams can inherently have different state, action, and observation spaces and yet, must still coordinate efficiently to cooperatively accomplish their mission [2, 14]. My proposed framework in [11] consists of two modules: (1) a Multi-Agent State-Action-Reward-Time-State-Action (MA-SARTSA) algorithm under partially observable Semi-MDPs as the high-level decision-making module to enable perception agents to learn to surveil in a restless environment with unknown number of dynamic targets and (2) a low-level coordinated control module that ensures probabilistically-guaranteed support for action agents. Our empirical and physical experiments show that our method enables effective collaboration in a perception-action composite team to accomplish complex missions, such as aerial wildfire fighting.

3 LEARNING END-TO-END MULTI-AGENT COORDINATION POLICIES

As a next step to my previous studies, I made several contributions towards leveraging data-driven approaches to develop end-to-end differentiable models for learning coordination strategies.

High-performing human teams benefit from communication to build and maintain shared mental models to improve team effectiveness. However, typical communication patterns across human teams widely differ based on the responsibility or role the human assumes. Inspired by heterogeneous communication patterns across human teams, in [14], we propose Heterogeneous Policy Networks (HetNet) to learn efficient and diverse communication models for coordinating cooperative heterogeneous robot teams. The key to our approach is the design of heterogeneous graph attention networks for an end-to-end communication learning model with a differentiable, binarized encoder-decoder channel to account for the heterogeneity of inter-class messages. HetNet enables "translating" the encoded messages into a shared, intermediate language among agents of a composite team, such as the perception-action teams described in Section 2. Our binarized communication model in Seraj et al. [14] achieves 200x reduction in the communication bandwidth over the best performing baseline while also setting a new SOTA in team performance, achieving an 8.1% to 434.7% performance improvement over baselines and across domains.

In addition to communication, individuals in high-performing human teams also benefit from the theory of mind [3] and making strategic decisions by recursively reasoning about the actions (strategies) of other human members [4]. Such hierarchical rationalization alongside with communication facilitate meaningful and strategic cooperation in human teams. Inspired by this behavior in strategic human teams, in [6], we propose a novel information-theoretic, fully-decentralized cooperative MARL framework, called InfoPG, where agents iteratively rationalize their action-decisions based on their teammates’ actions. We study cooperative MARL under the assumption of bounded rational agents and leverage action-conditional policies into policy gradient objective to accommodate our assumption. By leveraging the k-level reasoning paradigm from cognitive hierarchy theory [5], we propose a cooperative MARL framework in which naive, nonstrategic agents are improved to sophisticated agents that iteratively reason about the rationality of their teammates for decision-making. Our quantitative results show that InfoPG sets the SOTA performance in learning emergent cooperative behaviors by converging faster and accumulating higher team rewards over several recent prior work.

4 FUTURE WORK: FROM HUMAN EXPERTS TO ROBOT TEAMS

In future work, I intend to investigate efficient ways to incorporate humans’ teaming strategies for robot teams and directly learn team coordination policies from human experts through developing methods for Multi-Agent Learning from Demonstration (MA-LfD). Through LfD, we can enable robot teams to learn humans’ preferred way to communicate/collaborate, which may not be optimal but may be more natural. Moreover, LfD methods resolve the reward specification challenges in MARL. Particularly, I intend to tackle:

1. How can we teach a team of robots by showing them the preferred way to do a task?
2. How can we leverage heterogeneous demonstrations from a group of human experts for teaching a composite robot team to accomplish a shared task (i.e., teaching a robot soccer team, or a perception-action robot team)?
3. What are the dynamics/logistics required for an effective collaboration between a team of humans and a team of robots in cooperative and mixed-cooperative tasks?

Through conducting rigorous human-subject studies where a group of humans provide collaborative strategies to satisfy a shared objective, we can collect required data for training an MA-LfD architecture generalized for multiple teachers and students, particularly, to learn heterogeneous coordination policies for composite teams.
REFERENCES


