Embodied, Intelligent Communication for Multi-Agent Cooperation

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.

Introduction

While multi-robot systems are capable of executing time-sensitive, complex, and large-scale problems, it is challenging to efficiently coordinate such systems and to optimize the collaborative behavior among robots. Communication is a key necessity to achieve an effective coordinated policy among agents. This process, emulates high-performing human teams where communication is leveraged to build team cognition and maintain shared mental models to improve team effectiveness (Konan, Seraj, and Gombolay 2022).

In tackling the problem of designing or learning multi-agent coordination strategies through end-to-end models, I specifically focus my studies around three major problems that are less addressed in prior work: (1) heterogeneous teaming, (2) complex multi-faceted objective(s), and (3) *restless* and dynamic environments. In the following, I describe each of these challenges (Seraj 2022).

Heterogeneity in robots' characteristics and roles are introduced to leverage the relative merits of different agents' capabilities (Seraj, Chen, and Gombolay 2021). A group of heterogeneous robots that are collaborating on a number of co-dependent tasks to accomplish an overarching mission form a *composite robot team* (Seraj, Chen, and Gombolay 2021; Seraj, Wu, and Gombolay 2020). Due to inherently different state- and action-spaces of agents in a composite team, communication, coordination, and consequently collaboration is not straightforward and requires proper considerations to model the interactions among heterogeneous agents. Additionally, such heterogeneity in a composite team implicitly entails a multi-faceted team objective such that several disjoint and sometimes competing objectives need to be successively and actively carried out to successfully accomplish a mission. For instance, such multi-faceted team objectives occur when a team of surveillance robots is required to simultaneously monitor a wildfire while providing their human collaborators on the ground with online information regarding their proximity to the fire (Seraj and Gombolay 2020). Such multi-agent coordination problem becomes even more challenging when robots have to collaborate in a dynamic environment (e.g., a propagating wildfire) since dynamic environments are *restless*; meaning regardless of robots' collective actions, the states of the environment continually change (Seraj, Chen, and Gombolay 2021; Seraj and Gombolay 2020).

Copyright © 2023, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

Model-Based Multi-agent Planning and Control Methods

To enable a highly efficient and intelligent team behavior, in Seraj and Gombolay (2020), 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. An issue that we faced in (Seraj and Gombolay 2020) 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. (2021), 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.

I studied the problem of coordinated planning of a heterogeneous multi-UAV team at the high-level decision-making and the low-level control for cooperative surveillance and tracking of a restless environment in (Seraj, Chen, and Gombolay 2021; Seraj, Silva, and Gombolay 2022). In particular, in Seraj, Chen, and Gombolay (2021), 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 (Seraj, Chen, and Gombolay 2021; Seraj et al. 2022; Seraj, Wu, and Gombolay 2020). 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 (Seraj et al. 2022).

Learning End-to-End Multi-agent Coordination Policies

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 (Seraj et al. 2022; Pimentel et al. 2022), 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 . Our binarized communication model in Seraj et al. (2022) achieves 200× 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 and making strategic decisions by recursively reasoning about the actions (strategies) of other human members. Such hierarchical rationalization alongside with communication facilitate meaningful and strategic cooperation in human teams. Inspired by this behavior in strategic human teams, in (Konan, Seraj, and Gombolay 2022), 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, 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.

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 Multi-Agent Learning from Demonstration (MA-LfD) methods. Through LfD, we can enables 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?

By conducting rigorous human-subject studies (i.e., expert providing demonstration for robot team) 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

Konan, S.; Seraj, E.; and Gombolay, M. 2022. Iterated Reasoning with Mutual Information in Cooperative and Byzantine Decentralized Teaming. *arXiv preprint arXiv:2201.08484*.

Pimentel, L.; Paleja, R.; Wang, Z.; Seraj, E.; Pagan, J. E.; and Gombolay, M. 2022. Scaling Multi-Agent Reinforcement Learning via State Upsampling. *RSS 2022 Workshop on Scaling Robot Learning*.

Seraj, E. 2022. Embodied Team Intelligence in Multi-Robot Systems. In *AAMAS*, 1869–1871.

Seraj, E.; Azimi, V.; Abdallah, C.; Hutchinson, S.; and Gombolay, M. 2021. Adaptive Leader-Follower Control for Multi-Robot Teams with Uncertain Network Structure. In *(ACC 2021)*, 1088–1094. IEEE.

Seraj, E.; Chen, L.; and Gombolay, M. C. 2021. A hierarchical coordination framework for joint perception-action tasks in composite robot teams. *IEEE Transactions on Robotics*.

Seraj, E.; and Gombolay, M. 2020. Coordinated control of uavs for human-centered active sensing of wildfires. In *(ACC 2020)*, 1845–1852. IEEE.

Seraj, E.; Silva, A.; and Gombolay, M. 2022. Multi-UAV planning for cooperative wildfire coverage and tracking with quality-of-service guarantees. *JAAMAS*, 36(2): 1–39.

Seraj, E.; Wang, Z.; Paleja, R.; Martin, D.; Sklar, M.; Patel, A.; and Gombolay, M. 2022. Learning efficient diverse communication for cooperative heterogeneous teaming. In *AAMAS*, 1173–1182.

Seraj, E.; Wu, X.; and Gombolay, M. 2020. FireCommander: An Interactive, Probabilistic Multi-agent Environment for Heterogeneous Robot Teams. *arXiv preprint arXiv:2011.00165*.