ABSTRACT

Lakshmi Narayan Vishnunandan Venkatesh, Purdue University, May 2025. Intuitive Learning in Multi-Robot Systems. Major Professor: Byung-Cheol Min.

In the dynamic field of robotics, the development of systems to adapt and reprogram robots for new tasks is a vital frontier, not only in unpredictable everyday environments but also in industrial settings where robots must repurpose themselves for varied functions without costly and expert-driven reprogramming. Traditional robotic systems, with their fixed programming, are often hamstrung by their rigidity, lack adaptability, and limit their deployment to repetitive and static environments. This limitation is particularly pronounced in multi-robot systems (MRS), where research lags behind single-robot applications, highlighting a significant gap in the field. Innovations are required that leverage natural, human-centric modes of communication, encompassing nonverbal cues through demonstrations and verbal instructions via natural language, to facilitate more intuitive and flexible interactions within these complex systems.

This dissertation introduces novel frameworks that bridge the gap between human instructors and robotic learners. By integrating Learning from Demonstration (LfD) and Large Language Models (LLMs), it enhances MRS with the ability to learn tasks from minimal human input. Through demonstrations and natural language instructions, robots can not only execute tasks but also adapt and evolve within their environment. By leveraging natural human communication, this work improves human-robot interaction, making robotics more accessible to non-experts and expanding its potential for widespread adoption.

This work is presented across four pivotal chapters, each addressing a unique aspect of MRS learning and adaptability. The first introduces a few-shot LfD framework that leverages visual demonstrations to model robot-object interactions. A key innovation is the use of Interaction Keypoints (IK) to distill demonstrations into actionable subtasks for efficient skill acquisition. Robots execute tasks using sensorimotor actions and adaptive reinforcement learning (RL) strategies. A distinctive feature is the ability to handle previously unseen contact-based skills, where Soft Actor-Critic (SAC) learning with a classifier-based reward function removes the need for manually crafted rewards, enhancing adaptability to environmental variations. The framework is rigorously tested across behavior-based and contact-based tasks, demonstrating its effectiveness in teaching complex multi-robot behaviors from minimal demonstrations.

Building on this foundation, the research introduces Demonstration-Driven Task Coordination and Trajectory Execution (DDACE), extending LfD beyond discrete task execution to task sequencing and trajectory-based learning in MRS. DDACE tackles coordinated multi-robot motion planning and heterogeneous teams by modeling temporal action sequences and spatial trajectory execution from demonstrations. It leverages graph-based learning such as Temporal Graph Networks (TGN) to capture inter-robot dependencies and Gaussian Processes (GP) to generate adaptable spatial trajectories. This enables MRS to generalize across different robot configurations and execute precise, flexible task policies in collaborative settings. Extensive validation demonstrates its effectiveness in multi-robot planning, trajectory execution, and large-scale coordination.

After acquiring new skills via demonstrations, the research shifts to optimizing task allocation when robots already possess the necessary skills. This introduces SMART-LLM, a framework that leverages LLMs for multi-robot task planning by translating high-level natural language instructions into structured task plans through task decomposition, coalition formation, and skill-based allocation. Evaluated on a specially designed benchmark dataset, SMART-LLM demonstrates its ability to interpret complex instructions and generate robust task plans for varying task complexities. This research highlights its role in streamlining MRS task planning, improving operational efficiency, and enhancing adaptability.

Building on SMART-LLM, this research explores language models in robotics through Zero-Shot Multi-Robot Context-Aware Pattern Formations (ZeroCAP). ZeroCAP enables context-aware pattern formation from natural language input without prior training or extensive datasets, reducing barriers to robotic coordination in dynamic environments. A key innovation is mitigating the spatial inaccuracies of Vision Language Models (VLM) by integrating advanced segmentation and shape descriptors, enhancing spatial awareness for complex formations. Extensive validation across diverse scenarios highlights ZeroCAP's efficacy over traditional methods, demonstrating its potential for real-world applications requiring precise and adaptive multi-robot coordination.

The dissertation concludes by synthesizing its findings and outlines future research directions, highlighting the impact of its methodologies on MRS. It envisions a future where robots evolve from mere tools to intelligent collaborators, capable of learning from and adapting to human instructions. This work lays the foundation for next-generation robotic systems that seamlessly integrate into daily life, transforming industries and personal environments.