# Zachary Serlin, PhD

*Research Statement*

## Overview

My research focuses broadly on applying principles from formal methods and control systems theory to multi-agent autonomous systems to enforce guarantees on high-level system behaviors, particularly when these systems have learning-based components. My past work has developed approaches to learn high-level multi-agent strategies for adversarial games, to coordinate large teams of heterogeneous agents and sensors from high-level specification languages, and to ensure high-dimensional learning-based autonomous systems remain within performance envelopes and safe operation regions. My current and future research directions look to broaden safety guarantees for embodied AI systems and neuro-symbolic visual-language navigation problems.

## Research Experience

My past research has focused on multi-agent system coordination and control at the intersection of autonomous systems, formal methods, and machine learning. Details of my previous research work is available in my CV. As a researcher, I have developed both theoretical contributions and large scale hardware experiments demonstrating the practical utility of that theory. I have developed formal guarantees on multi-robot systems using concepts from temporal logics, verifiable control, and optimization theory. This work enabled a large team (10s to 100s) of robots to coordinate in sensor coverage and complex sequential tasks using linear temporal logic. This work was expanded to consider heterogeneous agents and more complex missions using a novel temporal logic framework called Capability Temporal Logic. This framework was also used to deploy a team of autonomous marine vessels with varying sensor capabilities on the Charles River with help of collaborators at MIT. Alongside this coordination work, I developed algorithms for distributed feature matching for computer vision applications that allowed the coordinated team to fuse image data and perform more efficient loop closures for the simultaneous localization and mapping problem. This work was extended to use features generated by the latent space of an autoencoder neural network. My recent work has also included developing learning-based safety guarantees for complex dynamical systems using neural control barrier functions. This work has also been done in collaboration with a professor at MIT and applied to tactical aircraft systems. My other current work also explores zero-shot Boolean task composition for Q-learning policies as a method for generating real-time strategies for adversarial games (specifically using an options-like framework in capture the flag). I have extensive experience with building and deploying robotics hardware (drones, marine vessels, subsea systems, and ground vehicles), using robotic operating system (ROS), and developing applications in MOOS-IvP. I also lead research programs focused on learning aircraft dynamics, learning-accelerated combinatorial optimization, and simulation to real transfer of learning-based perception and control systems. I currently maintain active collaborations with multiple professors at MIT, Boston University, Lehigh University, WPI, and BYU through internal funding, as well as grants from NASA, DARPA, and OSD(R&E).

## Research Goals

There is a pressing need for safety guarantees on AI systems as they become more prevalent in our society. My research has focused on safety guarantees for control systems with learned components, and I would like to extend these concepts to learning-based perception systems, neuro-symbolic visual-language navigation, and embodied AI systems. Building guarantees for these systems, especially as they begin to interact with the physical world, will be imperative to enable trust and wide adoption. I would specifically like to explore neuro-symbolic open-vocabulary systems for visual language mapping and navigation and develop safety guarantees on system operation and stability guarantees on representation embeddings. I would also like to eventually extend this field to the multi-agent setting and explore common concept representations that would allow a team of agents to explore an area and build a common open-vocabulary representation.