INDUSTRIAL & SYSTEMS ENGINEERING

UNIVERSITY of WASHINGTON

Towards Data-Driven Safety Concept Synthesis for Human-Robot Interactions

Karen Leung

Assistant Professor

Abstract

Advances in the fields of artificial intelligence and machine learning have unlocked a new generation of "learning-enabled" robots that are designed to operate in unstructured, uncertain, and unforgiving environments, especially settings where robots are required to interact in close proximity with humans. However, as learning-enabled methods, especially deep learning, continue to become more pervasive throughout the autonomy stack, it becomes increasingly difficult to ascertain the performance and safety of these robotic systems, a necessary prerequisite for their deployment in safety-critical settings. In this talk, I will first discuss how Hamilton-Jacobi (HJ) reachability, a robust control technique, can complement a high-level, possibly learning-enabled, robot planner to produce minimally interventional safe control strategies for a robot whenever the robot is faced with an unexpected situation. The approach is validated through human-in-the-loop simulation as well as on an experimental vehicle platform, demonstrating clear connections between theory and practice. In the second part of the talk, we switch gears and take on a more philosophical stance and consider "what defines a safe or unsafe state?" Specifically, in the autonomous driving context, a number of safety concepts for trusted AV deployment have been recently proposed throughout industry and academia. Yet, agreeing upon an "appropriate" safety concept is still an elusive task. I show that the HJ reachability framework can serve as an inductive bias to effectively reason, in a data-driven fashion, what is considered a safe or unsafe state.

Bio

Karen Leung is an Assistant Professor and the Vagners & Christianson Endowed Faculty Fellow in Aeronautics & Astronautics at the University of Washington. She directs the Control and Trustworthy Robotics Lab (CTRL) which focuses on developing algorithms for safe, intelligent, and trustworthy robotics systems, especially those that interact with humans and operate in safety-critical settings. Her research combines techniques from control theory, machine learning, and formal methods to devise frameworks and machinery to achieve human trust in learning-enabled autonomy, starting from the algorithmic foundations of safe robot decision making and control, and incorporating further refinement through learnings from practical deployment. Prior to joining UW, Karen was a research scientist at NVIDIA, working in the Autonomous Vehicle Research Group, where she now currently holds a partial appointment. Karen received her M.S. and Ph.D. in Aeronautics and Astronautics from Stanford University, and a combined B.S./B.E. in Mathematics and Aerospace Engineering from the University of Sydney in Australia. She is a recipient of the UW + Amazon Science Hub Faculty Research Award, the William F. Ballhaus Prize for best Ph.D. Thesis Award, and the Qualcomm Innovation Fellowship.

Date/Time:

Jan. 10th, 2023

MEB 235

1:30 - 2:20 pm