Project Title: Safe decision-making in interactive environments

Recipient/Grant (Contract) Number: Carnegie Mellon University, Grant #: 69A3552344811

Center Name: Safety21 National University Transportation Center for Promoting Safety

Research Priority: Promoting Safety

Principal Investigator(s): Yorie Nakahira Carnegie Mellon University

PI Contact: yorie@cmu.edu

Project Partners:

- Jitsik LLC
- City of Pittsburgh, Department of Mobility & Infrastructure

Research Project Funding: $196,000.00

Project Start and End Date: 07-01-2023 to 06-30-2024

Project Description:

Autonomous ground vehicles must safely operate in highly interactive environments with human uncertainties. Safe actions depend on context, interactions, and absolute (mathematical measures for safety) may differ from how humans perceive as safe and reliable behaviors. However, producing context-dependent and interaction-aware safe actions is non-trivial due to the following technical challenges. Challenge 1: Hardness in the identification of interaction mechanisms. When interaction mechanisms are known or can be learned, many safe learning and control techniques can be employed. However, in many interactive environments, it may be fundamentally difficult to fully identify the mechanisms of the opponent's interaction due to unobserved confounders. Challenge 2: Latent risk and latent variables. There exist latent risks (such as occlusions) and unobserved variables (e.g., awareness and intention), and safe actions depend on such factors. For example, pedestrians can decide to enter a crosswalk, but such intent may not be directly observable. Decisions that ignore such factors may experience unexpected risks. Challenge 3: Tensions between long-term safety vs. computation. Accounting for long-term outcomes in risk quantification and control is challenging due to stringent computation vs. time-horizon tradeoffs, particularly for rare events. Myopic safety can be efficiently certified, but ensuring long-term safety may require prohibitive computation. Although distribution shifts are often approached by finding actions that are robust to changes or avoid changes, much less work explores how to proactively induce desirable opponent behavior changes in interactive environments. For example, whereas humans can slowly squeeze their way through crowded environments, autonomous systems programmed to maintain worst-case distance may not find feasible solutions (e.g., freezing robot problems). Proximity can be safe or risky depending on opponents' interaction mechanisms, but this is not captured in a simple risk measure of distance. The proposed research aims at realizing such capabilities by accounting for interaction models and anticipating unobserved unknowns in risk quantification and decision-making. Specifically, we propose the following research. Thrust 1: We will establish an efficient risk quantification method with theoretical guarantees. We will leverage an integrative view of stochastic systems, MC, PDEs, and Physics-informed neural networks (PINNs) to estimate intervention risk from heterogeneous data and exploit low-dimensional structures. Our prior work has derived four types of long-term safe probabilities as unique solutions to deterministic linear PDEs, which characterize the relation between risk probabilities of different time horizons and initial states. PINNs with these PDE constraints have been shown to be able to infer risk probabilities beyond available data with provable generalization. Here, we will explore such characterization and enable long-term risk to be quantified using shorter-term interaction data. Thrust 2: We will develop efficient long-term safety certificates for interactive environments. While treating statistical models as mechanistic models may neglect important latent risk factors, little effort has been made to rigorously differentiate the underlying mechanistic models vs. observed statistical models in the design of safety certificates. Here, we will build upon our prior work on probabilistic invariance to differentiate the two models, control the risk probability using observed statistics, and handle information constraints arising from delayed and rate-limited communication.

Outputs:

We anticipate the following outputs:

(1) Theory, methodology, and algorithms for risk quantification, control, and learning techniques that jointly realize the following merits. Merit 1: Produce safe and interaction-aware actions Merit 2: Anticipating risks from unobserved factors Merit 3. Physics-Informed Risk Quantification with provable generalization Merit 4: Interaction-Aware Safety Certificate without Complete Interaction Mechanisms
(2) Integrated software package, data sets from experiments, and testbed for future research: The code for the proposed techniques, experimental data, and testbed will be made publicly available through publications. We will include detailed descriptions so that they can be useful for research, engineering, and education. These tools will also be disseminated in PI's classes for students to better understand available techniques for safe interactions in intelligent transportation systems.
(3) A set of demonstrations that illustrate the capability to handle complex interactions and latent risks. The proposed system is expected to exhibit flexible behaviors such as slowly squeezing forward a way through a congested area, maintaining sufficient distance from agents who do not actively avoid obstacles, following the flow of a crowd (known as a flow-following strategy), reducing speed in anticipation of latent risks (occlusions).

Outcomes/Impacts:

Safety: The proposed techniques are expected to navigate crowded, interactive environments with better safety and feasibility. The performance and robustness of the proposed methods will be evaluated and compared to the existing techniques for autonomous ground vehicles operating in diverse environments. Reliability: Be better account for human driving models, the proposed techniques are expected to provide more comfort for humans sitting in automated vehicles. The reliability will be tested in human experiments sitting in real-sized vehicles virtually experiencing automated driving. We will use the experiments to understand Durability and cost: The proposed research studies how to produce safe actions using reduced computation and unreliable communication. These efforts will contribute to the durability and cost-reduction of intelligent vehicles and connected autonomy. Through deployment and case studies, we will also study additional design considerations such as optimal allocation of onboard resources, offloading computation offline, and deployment in embedded devices.