

GRADUATE SEMINAR

Explaining Errors in Autonomous Driving: A Diagnosis Tool and Testing Framework for Robust Decision Making

Leilani Gilpin – Assistant Professor, Department of Computer Science and Engineering, University of California, Santa Cruz

Autonomous systems are prone to errors and failures without knowing why. In critical domains like driving, these autonomous counterparts must be able to recount their actions for safety, liability, and trust. An explanation: a model-dependent reason or justification for the decision of the autonomous agent being assessed, is a key component for post-mortem failure analysis, but also for pre-deployment verification. I will show a monitoring framework that uses a model and commonsense knowledge to detect and explain unreasonable vehicle scenarios, even if it has not seen that error before.

In the second part of the talk, I will motivate the explanations as a testing framework for autonomous systems. While it is important to develop realistic tests in simulation, simulation is not always representative of the corner cases in the real world. I will show how to use explanations in a feedback loop. The explanation ensures that the machine has done the right thing or it exploits a stressor to be modified and tested moving forward. I will conclude by discussing new challenges at the intersection of XAI and autonomy towards autonomous systems that are explainable by design.



Leilani H. Gilpin is an Assistant Professor in Computer Science and Engineering at University of California, Santa Cruz. Her research focuses on enabling opaque autonomous systems to explain themselves for robust decision-making, system debugging, and accountability. Before joining the faculty at UC Santa Cruz, she was a research scientist at Sony AI. She holds a PhD in Computer Science from MIT, an M.S. in Computational and Mathematical Engineering from Stanford University, and a B.S. in Mathematics (with honors), B.S. in Computer Science (with highest honors), and a music minor from UC San Diego.

November 11, 2021

12 – 1 PM

<https://wse.zoom.us/j/93432014600>