PURDUE UNIVERSITY | Elmore Family School of Electrical and Computer Engineering

***** CONFIDENTIAL*****

Please do not distribute outside of campus.

Faculty Candidate Seminar
Defense Innovation Search

James Ferlez

DARPA Innovation Fellow

Tuesday, February 25, 2025
Presentation: 9:30 A.M. – 10:30 A.M.
WANG 1004

Assured Autonomy for AI-enabled Systems

Abstract: The last decade has marked a fundamental paradigm shift in autonomous systems design: autonomous systems can now be designed by allowing machines to learn for themselves, most successfully using Neural Network (NN) models. However, autonomous systems with machine-learned NNs lack formal assurances, which makes them prone to sporadic, inexplicable failures. This lack of formal assurances is a critical impediment to the wider deployment of NNs in safety-critical autonomous systems, where they promise to enable next-generation capabilities in self-driving vehicles, aerospace applications, assistive robotics and smart cities.

In this talk, I will discuss my work on developing both the theory and tools necessary to automatically design assured NNs for autonomous systems. My work addresses this problem at multiple stages of the NN design pipeline: (i) designing assured NN architectures; (ii) verifying the formal properties of NNs after training; (iii) repairing trained NNs that fail assurances; and (iv) enhancing the performance of assured NNs without compromising those assurances. In particular, I will show that a unique semantic NN architecture, the Two-Level Lattice architecture, is particularly amenable to computationally efficient assured architecture design, formal verification and repair, especially for the control of autonomous systems. In addition, I will describe my work on provably-safe enhancement of NNs for an autonomous driving application, where I consider the problem of provably safe vehicle-to-edge NN offloading.

Bio: James Ferlez is a DARPA Innovation Fellow. Before that he was a Postdoctoral Scholar at the University of California, Irvine in the Resilient Cyber-Physical Systems Laboratory. His research interests include formal methods for control, neural networks, machine learning and stochastic control. He received his PhD from the University of Maryland, College Park under the supervision of Steve Marcus and Rance Cleaveland. As a PhD student, he was selected as an A.J. Clark School of Engineering Future Faculty Fellow. His recent tool, FastBATLLNN, was recognized by the International Verification of Neural Networks Competition 2022 (VNN-COMP'22), where it won the Category Winner award as the top performing tool on the tllverifybench benchmark.

Hosts
Shreyas Sundaram ~ sundara2@purdue.edu
Jason McKinney ~ mckinnjd@purdue.edu