Keynote Talk 1 (DSML)

From Neural Network Verification to Formal Verification for Neuro-Symbolic Artificial Intelligence (AI)

Taylor T. Johnson

Abstract

The ongoing renaissance in artificial intelligence (AI) has led to the advent of data-driven machine learning (ML) methods deployed within components for sensing, perception, actuation, and control in safety-critical cyber-physical systems (CPS). While such learning-enabled components (LECs) are enabling autonomy in systems such as autonomous vehicles and robots, ensuring such components operate reliably in all scenarios is extraordinarily challenging, as demonstrated in part through recent accidents in semi-autonomous/autonomous CPS and by adversarial ML attacks. We will discuss formal methods for assuring specifications---mostly robustness and safety---in autonomous CPS and subcomponents thereof using our software tools NNV and Veritex, developed partly in DARPA Assured Autonomy and Assured Neuro Symbolic Learning and Reasoning (ANSR) projects. These methods have been evaluated in CPS development with multiple industry partners in automotive, aerospace, and robotics domains, and allow for formally analyzing neural networks and their usage in closed-loop systems. We will then discuss how these methods are enabling verification for the third wave of AI systems, namely neuro-symbolic systems. In particular, we will discuss a class of neuro-symbolic systems we have been developing called neuro-symbolic behavior trees (NSBTs), which are behavior trees (a form of hierarchical state machine) that may call neural networks and are becoming more widely used in robotics, and for which we have been developing verification methods implemented in a tool called BehaVerify. We will also discuss relevant ongoing community activities we help organize, such as the Verification of Neural Networks Competition (VNN-COMP) held with the International Conference on Computer-Aided Verification (CAV) the past few years and the Symposium on AI Verification (SAIV) this year, as well as the AI and Neural Network Control Systems (AINNCS) category of the hybrid systems verification competition (ARCH-COMP) also held the past few years. We will conclude with a discussion of future directions in the broader safe and trustworthy AI domain, such as in new projects verifying neural networks used in medical imaging analysis and malware classifiers.

Short Biography

Dr. Taylor T. Johnson, PE, is A. James and Alice B. Clark Foundation Chancellor Faculty Fellow and an Associate Professor of Computer Science (CS) in the School of Engineering (VUSE) at Vanderbilt University, where he directs the Verification and Validation for Intelligent and Trustworthy Autonomy Laboratory (VeriVITAL) and is a Senior Research Scientist in the Institute for Software Integrated Systems (ISIS). Dr. Johnson's research has been published in venues such as AAAI, CAV, EMSOFT, FM, FORMATS, HSCC, ICSE, ICDM, ICCPS, IJCAI, NFM, RTSS, SEFM, STTT, TNNLS, UAI, among others. Dr. Johnson earned a PhD in Electrical and Computer Engineering (ECE) from the University of Illinois at Urbana-Champaign in 2013, where he worked in the Coordinated Science Laboratory with Prof. Sayan Mitra, and a BSEE from Rice University in 2008. Dr. Johnson is a 2023 recipient of an Outstanding Reviewer award at EMSOFT, 2022 recipient of the Best Artifact Evaluation Award at FORMATS, a 2018 and 2016 recipient of the Air Force Office of Scientific Research (AFOSR) Young Investigator Program (YIP) award, a 2016 recipient of the ACM Best Software Repeatability Award at HSCC, a 2015 recipient of the National Science Foundation (NSF) Computer and Information Science and Engineering (CISE) Research Initiation Initiative (CRII), and his group's research is or has recently been supported by AFOSR, ARO, AFRL, DARPA, Mathworks, NSA, NSF, NVIDIA, ONR, Toyota, and USDOT.

Keynote Talk 2 (DSML)

Machine Learning for Privacy Compliance in Software Ecosystems

Guangdong Bai

Abstract

In recent years, many countries have implemented legislation to regulate the collection, use and sharing of personal data. These regulations have imposed stringent obligations on data controllers and data processors, with significant penalties for any infringement of user privacy. In this talk, we will highlight our recent progress in assessing privacy compliance in modern applications, focusing on our studies conducted on the Internet of Things (IoT), Android, and Virtual Personal Assistant (VPA) apps. We explore the role of machine learning techniques in these endeavors. We will introduce the analysis of privacy-related documents using natural language processing and machine learning, and the automatic understanding of data handling practices using software analytics. Furthermore, we will discuss the research opportunities facilitated by advanced large language models in understanding user opinions and nudging developers in addressing evolving privacy challenges. Through our work, we aim to foster a machine learning-enhanced privacy-fair environment for both application users and developers.

Short Biography

Guangdong Bai is an Associate Professor in the School of Electrical Engineering and Computer Science at the University of Queensland, Australia. He obtained his PhD degree from the National University of Singapore in 2015. His research spans responsible machine learning, security, and privacy. His work has appeared in top security and software engineering venues such as IEEE S&P, NDSS, USENIX Security, ICSE, and FSE. He has served as program/general (co-)chair of international conferences such as NSS, ICECCS, and ICFEM. He is an Associate Editor of IEEE Transactions on Dependable and Secure Computing.