ICPS 2024

The 7th IEEE International Conference on Industrial Cyber-Physical Systems
May 12-15, 2024, St. Louis, MO, USA.

Tutorial/Workshop on
Design and Synthesis of Certifiably Safe AI-enabled Cyber Physical Systems with Focus on Human-in-the-Loop Human-in-the Plant Systems

Dr. Ayan Banerjee and Dr. Sandeep K.S. Gupta
Arizona State University

Dr. Imane Lamrani
Nikola Motors

- Brief description:

Advent of Large Language Models (LLM) and generative AI has introduced uncertainty in operation of autonomous systems with significant implications on safe and secure operation. This has led to the US government directive on assurance and testing of trustworthiness of AI. This tutorial aims at introducing the audience to the arising safety issues of AI-enabled autonomous Cyber Physical systems (CPS) and how is affects dependable and safe design for real life deployments. With the advent of LLMs and deep AI methods, CPS are becoming vulnerable to uncertainties. It will introduce a new human in the loop human in the plant design philosophy that is geared towards assured certifiability in presence of human actions and AI uncertainties while reducing data sharing between the CPS manufacturer and certifier. We will provide a landscape of informal and formal approaches in ensuring AI-based CPS safety at every phase of the design lifecycle, defining the gaps, current research to fill those gaps, and tools for detection of commonly occurring software failures such as doping. This tutorial also aims at emphasizing the need for operational safety of AI-based CPS and highlight the importance of explainability at every stage for enhancing trustworthiness. There has been significant research in the domain of model-based engineering that are attempting to solve this design problem. Observations from the deployment of a CPS are used to: a) ascertain whether the CPS used in practice match the proposed safety assured design, b) explain reasons for a mismatch in CPS operation and the safety assured design, c) generate evidence to establish the trustworthiness of a CPS, d) generate novel practical scenarios where a CPS is likely to fail.

- Relevance, target audience, and interest for the ICPS community

AI has been widely adopted in different domains including autonomous vehicles and IoT medical device. In a competitive environment, engineers and researchers are focused on developing innovative applications while minimal attention is provided to safety engineering techniques that cope with the fast pace of technological advances. As a result, recent failures and operational accidents of AI-based system highlight a pressing need for the development of suitable stringent safety monitoring techniques. We advocate for a change in the linear CPS development lifecycle from design, validation, implementation, and verification by incorporating feedback from the field of operation. This will result in a circular CPS development lifecycle, where operational data can be used to identify novel states and can be used as feedback. This will enable an agile proactive redesign policy that can predict failures and propose techniques to circumvent any safety risks. The tools used in this circular lifecycle will provide interpretable reports to the appropriate stake holders such as certification agencies, developers and users at different stages. This tutorial directly relates to the Cyber physical systems, ML, topic of ICPS.

- Duration:
180 mins

- Outline:
This tutorial aims to familiarize the audience with the topic and introduce them to two software tools HyMN and FaultEx. HyMn automatically learns a verification model from operational data of a CPS. It helps regulatory agencies compare the operation of the control system with the specifications given by the manufacturer to ensure that the system's operation conforms with the safety assured design of a CPS, thus facilitating the detection of intentional/unintentional corruption scenarios. FaultEx is a framework that derives a hybrid system representation of the cyber-physical system operation in deployment from the observe input/output traces and matches it with a finite state machine-based simplification of the CPS code. In this tutorial, the following topics will be covered.

- Basic Definitions and Introduction.
- Application: Medical devices, Aviation, Autonomous cars, Gesture-Based Communication
- Modelling Dynamic Behaviours, Components interaction.
- Formal Analysis and Verification.
- Human-CPS interaction safety.
- Regulation of CPS: safety standards and certification.
- Large Language models to plan safe CPS operation
- Certification Game for the Safety Analysis of CPS
- Model Checking
- Evaluation platforms for CPS systems, demos.

- Specific Goals and Objectives
There are two specific objectives: a) to familiarize the audience with theoretical approaches towards validating CPS operation against its design, explanation interfaces for explaining failures, generation of evidence of correct operation to improve trust, and generating novel scenarios for CPS, and b) give a hands on experience on safe and dependable design of real life examples including artificial pancreas and a heavy vehicle braking system, through the usage of two software tools HyMN and FaultEx.

-Brief CV:
Sandeep K. S. Gupta is the Associate Dean for Research in Fulton School of Engineering and a Professor of Computer Science and Engineering in the School for Computing and Augmented Intelligence (SCAI), ASU Tempe, AZ. Dr. Gupta heads the IMPACT Lab (http://impact.lab.asu.edu). IMPACT Lab has significant experience in hosting tutorials. Previously we have hosted tutorials at the Body Sensor Network conference, and at the Food and Drug Administration (FDA) on safe mobile medical control systems.

Ayan Banerjee, is an Assistant Research Professor at ASU. His research interest lies in safe, secure and sustainable AI enabled AAS. His expertise include model based analysis and design of AAS, implementation of AAS with embedded computing, and applications of wearable sensor based control systems in domains such as medical control systems, or gesture recognition.

Imane Lamrani is a ADAS engineer at NMC. She received a PhD in Computer science from ASU. Her research goal is to develop rigorous safety verification approaches to evaluate the correct operation of AI-enabled AAS in the field, perform root-cause analysis, and verify the operational safety of AI-enabled AAS.

- Relevant publications:
Operational Data Driven Feedback for Safety Evaluation of Agent-based CPS
Certification Game for the Safety Analysis of AI-Based CPS