ICPS 2021

The 4th IEEE International Conference on Industrial Cyber-Physical Systems
May 10-13, 2021, Victoria, BC, Canada.

Tutorial Proposal

Title of the Proposal: Ensuring Safety and Establishing Trust for AI enabled Cyber-Physical Systems

- Presenter(s):

Sandeep K.S. Gupta, Ayan Banerjee, Imane Lamrani
Affiliation: Arizona State University
Institutional Email: (sandeep.gupta,abanerj3,ilamrani)@asu.edu

- Brief description:

Artificial intelligence (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.

This tutorial aims at introducing the audience to the arising safety issues of AI-enabled cyber-physical systems (CPSs). We will provide a landscape of informal and formal approaches in ensuring AI-based CPS safety at every phase of the system's development and defining the gaps. This tutorial also aims at emphasizing the need for operational safety of AI-based CPS. There has been significant research in the domain of model-based engineering that are attempting to solve this design problem. However, in this tutorial we are looking at this problem from a different perspective of a third part observer. Observations from the deployment of a CPS are used to: a) ascertain whether the CPS used in practice actually 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. The challenge is the uncertainty of the human-in-the-loop behavior and incompleteness of the environment's model used in the developed and safety verification of AI-enabled CPS. In this tutorial, we will talk about two aspects: a) 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) hands on usage of two software tools HyMN and Learn2Sign using data from real life examples including closed loop blood glucose control systems (artificial pancreas) and American Sign Language Learning system (Learn2Sign).

- Duration:

Presentation duration - 180 mins

- **Outline:**

Complex interactions between operational components of AI enabled CPS under different environmental conditions and inclusion of human-in-the-loop result in myriad of safety concerns all of which may not only be comprehensively tested before deployment but also may not even be detected during design and testing phase. Incident rates have increased in multiple fields (e.g. aviation and autonomous car industries) that have been applying traditional safety verification techniques. This suggest that existing safety verification tools are underperforming due to the high complexity of the systems and the increasing pace of technology. In addition, safety critical AI enabled CPS such as IoT medical devices should meet government regulatory requirements before marketing. However, it is difficult for current safety verification methods to keep up with the increasing pace of technological change. For example, current safety verification methods did not initially detect the Volkswagen's defeat device that allowed vehicles to improperly meet US standards during regulatory testing.

This tutorial aims to familiarize the audience with the topic and introduce them to two software tools HyMN and Learn2Sign. HyMn automatically learns a formal verification model from operational data of a CPS. HyMn algorithm aims to help regulatory or legal agencies compare the operation of the control system with the specifications given by the manufacturer to ensure that the system's operation in the real world conforms with the safety assured design of a CPS, thus facilitating the detection of intentional/unintentional corruption scenarios. Learn2Sign is a gesture learning application that can not only recognize gestures performed by a user but can also compare the gestures with an

expert and provide feedback to the user so that they can improve their gesture execution. Lear2Sign can explain failures of a user in replicating gestures.

In this tutorial, demos of the HyMn and Learn2Sign tool applications will be performed. The following topics will be covered.

- Basic Definitions and Introduction - AI enabled CPS, CPS with control loop feedback, Self-Adaptive control systems, Explainable AI.
- Application: Medical devices, Aviation, Autonomous cars.
- Modeling Dynamic Behaviors, Components interaction, Formal Verification Modeling.
- Analysis and Verification.
    - Traditional safety engineering and existing gaps.
    - Effect of self-adaptation on system safety.
    - Runtime monitoring to enhance safety.
- Human-CPS-AI interaction safety.
- Regulation of CPS-AI systems: safety standards and certification.
- Model Checking and Reachability Analysis.
- Evaluation platforms for CPS-AI systems,
- Demos for using HyMn and Lear2Sign.

-**Brief CV:**
Sandeep K. S. Gupta is the Director of the School of Computing, Informatics, and Decision Systems Engineering (SCIDSE) and a Professor of Computer Science and Engineering, Arizona State University, Tempe, AZ. Dr. Gupta heads the IMPACT Lab (http://impact.asu.edu) at Arizona State University. IMPACT Lab has significant experience in hosting tutorials. Previously we have hosted tutorials at the Body Sensor Network conference (two times), and at the Food and Drug Administration (FDA) on safety security and sustainability of mobile medical control systems.

Ayan Banerjee, is an Assistant Research Professor at Arizona State University. His research interest lies in safe, secure and sustainable AI enabled Cyber-Physical systems. His expertise include model based analysis and design of CPS, implementation of CPS with embedded computing, and applications of wearable sensor based control systems in domains such as medical control systems, or gesture recognition. He has continuing collaboration with Food and Drug's Administration and Mayo Clinic.

Imane Lamrani is a Post Doctoral Fellow in the School of Computing, Informatics, and Decision Systems Engineering (SCIDSE) at Arizona State University. She received a M.S. in Computer Systems and Software Design from Jacksonville State University. Her research goal is to develop rigorous safety verification approaches to evaluate the correct operation of AI-enabled Cyber Physical Systems (CPS) in the field, perform root-cause analysis, and verify the operational safety of AI-enabled CPS.

- **Relevant publications:**

Operational Data Driven Feedback for Safety Evaluation of Agent-based CPS
I Lamrani, A Banerjee, SKS Gupta
IEEE Transactions on Industrial Informatics

Non-linear Analysis for Operational Safety Verification of Cyber Physical Systems
A Banerjee, I Lamrani, SKS Gupta
2020 IEEE Conference on Industrial Cyberphysical Systems (ICPS) 1, 529-534

On evaluating the effects of feedback for Sign language learning using Explainable AI
P Paudyal, A Banerjee, S Gupta
Proceedings of the 25th International Conference on Intelligent User Interface

An Analytical Framework for Security-Tuning of Artificial Intelligence Applications Under Attack
K Sadeghi, A Banerjee, SKS Gupta
2019 IEEE International Conference On Artificial Intelligence Testing