IEEE ICPS 2020 - http://icps2020.fi

3rd IEEE International Conference on Industrial Cyber-Physical Systems (ICPS)

Tutorial on

"Ensuring Safety and Establishing Trust for AI enabled Cyber-Physical Systems"

Date: 10 June 2020

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

Keywords: AI-enabled CPS, Trustworthy CPS, Operational Safety, Autonomous Cars, Medical IoT devices, Explainable AI

Aim and Learning Objectives

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).

Content Summary

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

1. Show how HyMn can automatically learn a formal verification model (control logic and environment predictive model) of the artificial pancreas control system using data collected from the operation of the Medtronic Minimed 670G.
2. Show how Learn2Sign can be applied to increase trust in AI based gesture recognition systems by providing explanations of the system's decision and corrective feedback to enable the learner to replicate a gesture with similar qualities as that of an expert.

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.

Target Audience

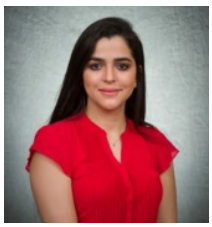
- Researchers in AI related applications, Cyber Physical Systems (CPS), Software verification and testing.
- Regulators such as personnel from US FDA or FAA
- Personnel from federal institutes such as NSF or NIH
- Industry partners in the domain of medical devices, aviation, autonomous cars, or power systems.

CV of the presenters:

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. He received the BTech degree in Computer Science and Engineering (CSE) from the Institute of Technology, Banaras Hindu University, Varanasi, India, the M.Tech. degree in CSE from the Indian Institute of Technology, Kanpur, and the MS and PhD degrees in Computer and Information Science from The Ohio State University, Columbus, OH. He has served at Duke University, Durham, NC, as a postdoctoral researcher; at Ohio University, Athens, OH, as a Visiting Assistant Professor; and at Colorado State University, Ft. Collins, CO, as an Assistant Professor. His current research is focused on safe, secure and sustainable cyber-physical systems with focus on AI-enabled systems such as Artificial Pancreas and Autonomous Transportation. His research has been

funded by the US National Science Foundation (NSF), The National Institute of Health (NIH), Science Foundation of Arizona (SFAz), the Consortium for Embedded Systems (CES), the Intel Corp., Raytheon, Northrop Grumman, and Mediserve Information Systems. Dr. Gupta has published over 150 peer reviewed conference and journal articles (Google Scholar h-index 53) and has advised over 15 PhD and over 25 MS students. He has co-authored two books: Fundamentals of Mobile and Pervasive Computing, McGraw Hill, and Body Sensor Networks: Safety, Security and Sustainability, Cambridge University Press. He currently is or has served on the editorial board of Elsevier Sustainable Computing, IEEE transactions on Parallel & Distributed System, IEEE Communications Letters and Wireless Networks. Dr. Gupta is a Senior Sustainability Scientist, in the Global Institute of Sustainability, ASU. His awards include a Best 2009 SCIDSE Senior Researcher, a Best Paper Award for Security for Pervasive Health Monitoring Application, and two best paper award nominations. His research has been highlighted on various research news, sites and blogs from various sources including NSF, ACM, ASU, and the Discovery channel. He was TPC Chair of BodyNets 2008 conference and TPC co-chair for Greencom 2013 conference and SI co-editor for IEEE Pervasive, IEEE Transactions on Computers, IEEE Transactions on Knowledge and Data Engineering, and IEEE Proceedings. Dr. Gupta heads the IMPACT Lab (http://impact.asu.edu) at Arizona State University.

Ayan Banerjee is an Assistant Research Professor at School of Computing Informatics and Decision Systems Engineering, Arizona State University. His research interests include pervasive computing in healthcare and analysis, and safety verification of embedded system software. He is also interested in hybrid system-based modeling and safety verification of closed loop control systems which interact with the physical environment, also known as Cyber-Physical Systems. In addition, he also works on explainable AI for failure analysis and feedback in human computer interface systems. Ayan has worked in close collaboration with FDA and Mayo clinic on practical AI driven systems.



Imane Lamrani is a current Computer Engineering Ph.D. Candidate in the School of Computing, Informatics, and Decision Systems Engineering (SCIDSE) at Arizona State University. She is also a member of Dr. Sandeep Gupta ‘s iMPACT Lab at SCIDSE. Before joining ASU, she received a M.S. in Computer Systems and Software Design from Jacksonville State University. She also obtained a M1 in Intelligent Systems from Faculté des Sciences de Kénitra and a Bachelor’s degree in Electronics, Telecommunications, and Computer Science from Faculté des Sciences et Techniques de Fès. Imane has been one of the 10 laureate U.S doctoral students chosen to participate in the French-American Doctoral Exchange Seminar (FADEx) 2016-CPS. 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.