Provably Safe Control Design and Learning: Recent Advancements
Pre-Tutorial Workshop Proposal for ECC

Mayank S Jha [1] and Bayu Jayawardhana [2]

[1] CRAN, CNRS, Université de Lorraine, France.
[2] Faculty of Science and Engineering, University of Groningen
Groningen, The Netherlands.

Abstract— **Safety-critical autonomy increasingly relies on learning-enabled controllers that must operate under uncertainty, limited data, and changing environments. This full-day ECC pre-tutorial workshop surveys recent advancements in safe control design and safe learning, emphasizing methods with provable guarantees and demonstrated performance on real robotic and cyber-physical platforms. The workshop begins with safe reinforcement learning perspectives that embed Control Barrier Function (CBF) invariance constraints into reward shaping and enable safe exploration via input-to-state safety concepts. We then connect optimization-based control and learning, highlighting how data-driven model updates can be integrated within MPC-style safety filters without sacrificing real-time feasibility. A central theme is robustness to uncertainty and nonstationarity: we cover distributionally robust control approaches that hedge against partially known disturbance distributions, and robust conformal prediction techniques that maintain probabilistic safety under interaction-driven and more general distribution shifts beyond the i.i.d. regime. The workshop further addresses scalability and structure through constrained multi-task reinforcement learning using natural policy gradient and actor-critic methods in both centralized and decentralized settings. Complementing these approaches, we discuss indirect and direct data-driven safety certification, including Hamiltonian learning from trajectory data to construct conservative safe sets. Finally, we examine distributed safety guarantees for multirobot systems via distributed control barrier functions for safe formation control, supported by experimental demonstrations.**

Keywords: safe control, learning-enabled control, control barrier functions, safe reinforcement learning, distributional robustness, conformal prediction, data-driven safety, multi-agent systems

I. Motivation and Scope

Learning-enabled controllers are increasingly deployed in safety-critical domains such as robotics, autonomous vehicles, and networked cyber-physical systems. These systems must satisfy hard constraints and safety specifications while operating under model uncertainty, stochastic disturbances, and distribution shifts induced by nonstationary environments and by the controller itself. This workshop focuses on *recent* advances at the intersection of safe control design and learning, emphasizing methods that (i) provide formal safety guarantees, (ii) remain computationally tractable for real-time use, and (iii) demonstrate scalability to complex and multi-agent settings.

II. Format

The workshop is designed as a full-day pre-tutorial event, combining invited technical talks with Q&A, and concluding with a moderated discussion/panel on open problems and practical deployment barriers. Each talk slot is planned for up to 60 minutes (including Q&A), with two short breaks and a 90-minute lunch.

III. Confirmed Speakers/Organizers and Contributions

The following invited talks constitute the proposed technical program.

A. Mayank Shekhar Jha (Université de Lorraine / CRAN, CNRS)

Talk title: Safe Reinforcement Learning with Provable Guarantees

Talk abstract:

> This talk presents recent advances in Safe Reinforcement Learning (Safe RL) methodologies applicable to both discrete-time and continuous-time systems with provable guarantees. The talk will introduce reinforcement learning (Adaptive Dynamic Programming) and motivate Safe RL. Then, Safe RL approaches developed for discrete-time systems will be presented where safety is enforced by augmenting Control Barrier Functions (CBFs) into the reward structure, thereby ensuring the forward invariance of predefined safe sets. Further Safe Exploration problems will be elaborated, highlighting how Input-to-State Stability (ISS) properties can be exploited to maintain safety during the exploration phase. In this context, the concept of Input-to-State Safety (ISSf) is introduced, offering a novel framework that promotes rich, risk-aware exploration while preserving formal safety guarantees, thereby enabling more informative exploration and improved tracking performance. The talk further covers recent developments addressing input saturation constraints, strategies for boundary-focused exploration, and model-free methods for system dynamics learning, contributing to a more robust and generalizable Safe RL framework.

Bio: Dr. Mayank Shekhar Jha is an Associate Professor at École Polytechnique de l'Université de Lorraine (Polytech Nancy) and researcher at CRAN (CNRS) since 2018. He obtained his Ph.D. in 2015 at École Centrale de Lille (France) and previously held postdoctoral positions at INSA

Toulouse (France) and a Research Associate position at the Rolls-Royce Technology Centre at the University of Sheffield (UK) in 2017. He has authored around 30 publications, leads a work package in an ANR-funded project on safe heterogeneous robot fleets, and has served as PI/Co-PI on multiple industrial projects (CNES, Dassault Aviation). He is an external collaborator and visiting researcher at NASA Ames Research Center. His current interests include Safe RL, deep-learning-based prognostics, and adiabatic quantum computing.

B. Bayu Jayawardhana (University of Groningen, The Netherlands)

Talk title: Distributed Control with Safety Guarantee for Multirobot Systems

Talk abstract:

> Over the past two decades, the development of multirobot systems has grown rapidly, and such systems have been deployed across many application domains, including agriculture, manufacturing, mobility, and high-tech industries. As the number of deployed robots operating in dynamic and constrained environments increases, distributed control frameworks must be able to guarantee safe operation and enforce constraints. In this talk, we review recent progress in the field, with a particular focus on the design of distributed control barrier functions for achieving safe formation control. Experimental results are presented on the formation control of drones and mobile robots.

Bio: Bayu Jayawardhana received the B.Eng. degree from Institut Teknologi Bandung (2000), the M.Eng. degree from Nanyang Technological University (2003), and the Ph.D. degree from Imperial College London (2006). He is the Scientific Director of the Dutch Institute of Systems and Control, the Scientific Director of the Engineering and Technology Institute Groningen, and a Full Professor at the University of Groningen. His research interests include nonlinear systems, systems with hysteresis, optomechatronics, multirobot systems, and systems biology. He is Vice-Chair of Publications in the IFAC Technical Committee on Nonlinear Control Systems and a fellow of the Netherlands Academy of Engineering.

C. Melanie Zeilinger (ETH Zürich, Switzerland)

Talk title: Safe Learning in Optimization-Based Control

Talk abstract (proposed):

> Advancing autonomous systems requires not only improving the control of complex dynamical systems, but also achieving complex tasks in challenging environments. This leads to a range of uncertainties on all levels. Learning has emerged as a promising means to practically address these challenges; however, the recovery of guarantees, particularly concerning safety, is often still lacking. This talk will highlight our results towards addressing this problem by building on a constrained optimal control paradigm and integrating robustness with learning. I will begin by defining our notion of safety and how it can be effectively formulated as a planning problem. The talk will then address concepts for achieving robust constraint satisfaction, as well as safely learning dynamics, objective, and constraint functions. The results in this presentation will be illustrated with applications from autonomous racing and robotics.

Bio: Melanie Zeilinger is an Associate Professor at the Department of Mechanical and Process Engineering at ETH Zurich, where she is leading the Intelligent Control Systems. She received the diploma in Engineering Cybernetics from the University of Stuttgart in Germany in 2006 and the Ph.D. degree in Electrical Engineering from ETH Zurich in 2011. From 2011 to 2012 she was a postdoctoral fellow at the École Polytechnique Fédérale de Lausanne (EPFL), Switzerland. From 2012 to 2015 she was a Postdoctoral Researcher and Marie Curie fellow in a joint program with the University of California at Berkeley, USA, and the Max Planck Institute for Intelligent Systems in Tuebingen, Germany. From 2018 to 2019 she was a professor at the University of Freiburg, Germany. Her awards include the ETH medal for her PhD thesis, an SNF Professorship, the Golden Owl for exceptional teaching at ETH Zurich 2022 and the European Control Award 2023. Her research interests include learning-based control with applications to robotics and biomedical systems.

D. Lars Lindemann (ETH Zürich, Switzerland)

Talk title: Safe Control under Distribution Shifts with Robust Conformal Prediction

Talk abstract:

> Accelerated by rapid advances in machine learning and AI, there has been tremendous success in the design of learning-enabled autonomous systems in areas such as autonomous driving and robotics. These exciting developments are accompanied by new fundamental challenges that arise regarding the safety and reliability of these increasingly complex systems due to imperfect learning, system unknowns, and uncertain environments. Conformal prediction (CP) — a statistical tool for uncertainty quantification — has gained popularity due to its ability to deal with these challenges. However, CP-based safety guarantees assume i.i.d. data, an assumption that is violated when system changes induce shifts in the data distribution.
>
> In this talk, I will provide new insight to design safe controllers under distribution shifts using robust CP. I will begin by advocating for the use of CP due to its simplicity, generality, and efficiency as opposed to existing optimization-based verification techniques. I will then provide an introduction to CP and summarize existing work that uses CP to design probabilistically safe controllers in dynamic environments. Subsequently, we will look into interactive settings where the system's behavior may change the environment's behavior, and vice versa. This circular dependency creates an interaction-driven distribution shift that invalidates existing safety guarantees. To deal with this chicken-and-egg problem, we propose an iterative framework that episodically updates the controller while robustly maintaining safety guarantees by quantifying the potential impact of a controller update on the environment's behavior. We realize this via adversarially robust CP where we perform a regular CP step in each episode using observed data under the current controller, but then transfer safety guarantees across controller updates by analytically adjusting the CP result to account for distribution shifts. Lastly, we will discuss ways to deal with more general distribution shifts that go beyond this interactive setting using adaptive and distributionally robust CP.

Bio: Lars Lindemann is an Assistant Professor for Algorithmic Systems Theory in the Automatic Control Laboratory at ETH Zürich. From 2023 to 2025 he was an Assistant Professor in the Thomas Lord Department of Computer Science at the University of Southern California. From 2020 to 2022 he was a Postdoctoral Fellow in the Department of Electrical and Systems Engineering at the University of Pennsylvania. He received his Ph.D. degree in Electrical Engineering from KTH Royal Institute of Technology in 2020. His research interests include systems and control theory, formal methods, machine learning, and autonomous systems. He received the Outstanding Student Paper Award at the 58th IEEE CDC and the Student Best Paper Award (as an advisor) at the 60th IEEE CDC, and was finalist for several best paper awards at CPS/Hybrids venues.

E. Ryan K. Cosner (Tufts University, USA)

Talk title: Theory-Driven Safe Robot Autonomy

Talk abstract:

> Robots can only achieve safe, lifelong autonomy if they can navigate the complex, stochastic uncertainties of the real world. Moreover, these systems must be able to make risk-aware decisions with limited data, sensing, and computational resources. In this talk, I discuss how I use tools from control theory, machine learning, and robotics to achieve this goal. In particular, I consider Control Barrier Functions (CBFs) as a method for ensuring safety and propose techniques that extend their guarantees to real-world systems. I first discuss my contributions to the traditional deterministic safety paradigm, which relies on worst-case, adversarial assumptions on uncertainty. Next, I present my work on an alternative stochastic, risk-sensitive paradigm that allows control algorithms to intelligently manage the system's level of risk. To further improve a system's ability to navigate the real world and enable it to balance robustness and performance, I demonstrate how theory can guide machine learning-based performance improvements while maintaining safety. To verify the utility of these methods and the validity of their theoretical guarantees, I deploy them on a variety of bipedal, quadrupedal, wheeled, and flying robot platforms.

Bio: Ryan K. Cosner is the Glenn R. Stevens Assistant Professor of Mechanical Engineering at Tufts University. He received his Ph.D. in Mechanical Engineering from Caltech in 2025, the M.S. degree from Caltech in 2021, and the B.S. degree from UC Berkeley in 2019. In 2022, he interned with the Autonomous Vehicle Research Group at NVIDIA. His interests include nonlinear and stochastic control and machine learning, with applications to dynamic, risk-aware safety-critical robotics.

F. Thinh T. Doan (UT Austin, USA)

Talk title: Natural Policy Gradient and Actor-Critic Methods for Constrained Multi-Task Reinforcement Learning

Talk abstract:

> Constrained reinforcement learning has been extensively recognized as a promising approach for safe autonomy. In this talk, I'll present our recent work on constrained multi-task reinforcement learning, where the goal is to find a single safe policy that effectively solves multiple tasks at the same time. We consider solving this problem both in the centralized setting, where information for all tasks is accessible to a single server, and in the decentralized setting, where a network of agents, each given one task and observing local information, cooperate to find the solution of the globally constrained objective using local communication. We first propose a primal–dual algorithm that provably converges to the globally optimal solution of this constrained formulation under exact gradient evaluations. When the gradient is unknown, we further develop a sampled-based actor–critic algorithm that finds the optimal policy using online samples of state, action, and reward. Finally, we study the extension of the algorithm to the linear function approximation setting.

Bio: Thinh T. Doan is an Assistant Professor in the Department of Aerospace Engineering and Engineering Mechanics at UT Austin. He received his B.S. from Hanoi University of Science and Technology (2008), and his M.S. (University of Oklahoma) and Ph.D. (UIUC) in Electrical and Computer Engineering. He was a TRIAD postdoctoral fellow at Georgia Tech (2018–2020) and previously an Assistant Professor at Virginia Tech. He received the AFOSR YIP and NSF CAREER Awards in 2024 and the 2025 IEEE CSS Antonio Ruberti Young Researcher Award.

G. Astghik Hakobyan (National Polytechnic University of Armenia / CSIE)

Talk title: Safety-Critical Control Under Uncertainty Using Distributionally Robust Approaches

Talk abstract:

> Ensuring safety in autonomous systems operating under uncertainty remains a fundamental challenge, particularly in real-time and large-scale settings. This talk explores distributionally robust control frameworks for safety-critical decision-making when disturbance and model uncertainty distributions are unknown or only partially characterized. Leveraging risk metrics and tools from distributionally robust optimization, we present methods that promote safe operation while maintaining the computational efficiency required for real-world implementation. The discussion spans both sampling-based control for single-agent robotic systems and distributed control strategies for multi-agent systems, highlighting how safety constraints can be enforced without relying on heavy online optimization.

Bio: Astghik Hakobyan is an Assistant Professor at the National Polytechnic University of Armenia (NPUA) and a Leading Researcher at the Center of Scientific Innovations and Education and Aerial Robotics Education (CSIE). She earned her B.Sc. in Automation and Control from NPUA (2018) and her M.Sc. and Ph.D. in Electrical and Computer Engineering from Seoul National University (2020, 2023). She received the Distinguished ECE M.S. Dissertation Award and Distinguished ECE Ph.D. Dissertation Award from SNU. Her research focuses on control and optimization, motion planning, and safe autonomous systems.

H. Jason Choi (UCLA, USA)

Talk title: Data-driven Safety Frameworks—Indirect vs. Direct Approaches

Talk abstract:

> Ensuring safety in autonomous systems operating under uncertainty is a central challenge for realizing reliable, large-scale autonomy. Classical model-based safety frameworks—such as Hamilton–Jacobi (HJ) reachability and Control Barrier Functions (CBFs)—provide rigorous guarantees but rely heavily on accurate system models. To overcome this limitation, recent research has sought data-driven extensions that incorporate empirical information from real-world operation. This talk presents a unified perspective contrasting two broad categories of such frameworks: indirect and direct data-driven safety approaches.
>
> In indirect frameworks, data are used to model or bound the uncertainty in dynamics, which are then embedded into existing model-based frameworks like HJ reachability or CBFs. These approaches enable scalable learning of safety guarantees and have been successfully applied to flight envelope protection for emerging electric vertical takeoff and landing (eVTOL) vehicles. However, they remain reliant on intermediate model-learning steps and domain-specific assumptions.
>
> In contrast, direct data-driven frameworks, exemplified by the proposed Data-Driven Hamiltonian (DDH), bypass explicit model identification and instead infer safety certificates directly from trajectory data. By approximating the Hamiltonian with observed state–velocity pairs, the DDH method constructs safe sets and safety filters through a purely data-driven formulation without explicit dynamics model and guarantees conservative (inner) approximations of the true safe set.
>
> Through this comparison, the talk highlights a conceptual shift from learning models for safety analysis to learning safety itself, outlining how direct data-driven frameworks can generalize safety guarantees across diverse dynamical systems.

Bio: Jason Jangho Choi is an Assistant Professor in the Electrical and Computer Engineering Department at UCLA and the principal investigator of the Safety and Collective Intelligence (SCI) Autonomy Lab. He received his Ph.D. in Mechanical Engineering from UC Berkeley (2025) and his bachelor's degree from Seoul National University. His research focuses on safety assurance for learning-enabled autonomous systems and decentralized multi-agent intelligence. He was recognized as a Robotics: Science and Systems (RSS) Pioneer in 2024.

IV. Tentative Schedule

Table I shows a tentative schedule consistent with an 08:05 start and a 18:00 finish, including two 15-minute breaks and a 90-minute lunch. All talk slots are planned as 60 minutes *maximum* (including Q&A).

TABLE I
Tentative workshop program (local time; each talk slot is 60 minutes incl. Q&A).

Time	Program item
08:00–08:05	Opening remarks and logistics (Organizers)
08:05–09:05	**Bayu Jayawardhana** — Distributed Control with Safety Guarantee for Multirobot Systems
09:05–10:05	**Melanie Zeilinger** — Safe Learning in Optimization-Based Control
10:05–10:20	Coffee break
10:20–11:20	**Lars Lindemann** — Safe Control under Distribution Shifts with Robust Conformal Prediction
11:20–12:00	**Astghik Hakobyan** — Safety-Critical Control Under Uncertainty Using Distributionally Robust Approaches
12:00–13:30	Lunch break
13:30–14:30	**Thinh T. Doan** — Natural Policy Gradient and Actor-Critic Methods for Constrained Multi-Task Reinforcement Learning
14:30–15:30	**Ryan K. Cosner** — Theory-Driven Safe Robot Autonomy
15:30–15:45	Coffee break
15:45–16:45	**Jason J. Choi** — Data-driven Safety Frameworks—Indirect vs. Direct Approaches
16:45–17:45	**Mayank S. Jha** — Safe Reinforcement Learning with Provable Guarantees
17:45–18:00	Panel discussion, open problems, and wrap-up (all speakers)

V. Organizers

Mayank Shekhar Jha (Université de Lorraine / CNRS) works on safe reinforcement learning and learning-enabled control for safety-critical systems, with academic and industrial collaborations.

Bayu Jayawardhana (University of Groningen) works on nonlinear systems and distributed control with applications to multirobot systems, and serves in leadership roles within the Dutch Institute of Systems and Control.

VI. Expected Outcomes

The workshop will provide ECC participants with a consolidated view of state-of-the-art safe control and learning methods, highlighting common mathematical structures, practical implementation constraints, and open research directions (e.g., safety under distribution shift, scalable multi-agent safety guarantees, and data-efficient certification).

Note: Titles and abstracts are included as provided by speakers; the schedule is tentative and can be adjusted based on ECC logistics and speaker availability.