What can Control Theory do for Robust and Safe Learning?

Open Track for the IFAC 2023 World Congress organized by TCs 1.2 and 2.5. Track code: 8258t

Mario Sznaier * **Milad Siami** * **Eduardo Sontag** *
Tiago Roux Oliveira **

* *Northeastern University, Boston, MA 02115, USA.*
** *University of Rio de Janeiro, Rio de Janeiro, Brazil*

TRACK DESCRIPTION

The combination of the exponential explosion in sensing capabilities and unprecedented advances in Machine Learning (ML) have opened up the possibility of having truly autonomous systems capable of safely learning from and interacting with the environment. Indeed, recent results on reinforcement learning offer a tantalizing view of the potential of a rapprochement between control and learning. However, most of the research at the confluence of control and ML seeks to address the problem of "what can ML do for Control?" and concentrates on using Machine Learning techniques to synthesize controllers from data, often avoiding Systems Identification and model-based control design steps. This open track focuses on the less explored dual problem of "what can Systems Theory do for Learning?" We believe that this less explored area can be a rich source of problems at the confluence of Machine Learning and Control that do not necessarily involve learning a controller. Examples range from the use of control-theoretic tools to analyze and certify the learning properties of Neural Networks to the exploration of several aspects unique to learning dynamical systems that make these problems considerably harder than more traditional Machine Learning problems. For example the training of feedforward neural networks can be viewed as an optimal control problem, in which the weights of the network at successive stages are seen as inputs. This suggests the analysis of convergence properties of training as an optimal control problem. Related to this is the issue of robustness to adversarial perturbations, which can be analyzed in principle using the concept of "input to state stability" (ISS) from control theory (ISS). Similarly, the use of concepts from Realization Theory can help in determining the sample complexity of learning systems and analyzing the generalization properties of dynamical models learned from partial data. Finally, barrier function motivated ideas, coupled with computational tools rooted in semi-algebraic optimization can certify safety during learning. Examples of topics of interest to this open track include, but are not limited to:

- Control theoretic analysis of NN training, including convergence and robustness.
- Control theoretic analysis of Reinforcement Learning.
- Use of control theoretic tools to analyze robustness of NN against adversarial attacks.
- Use of control theoretic tools to prune NN during training or to derive frugal equivalents of a trained network.
- Sample complexity of robust learning of dynamical systems.
- Generalization properties of dynamical models learned from local data.
- Systems Theoretic motivated architectures for learning dynamical models
- Certified safe learning.

PROPOSING TECHNICAL COMMITTEES

(1) TC 1.2 Adaptive and Learning Systems
(2) TC 2.5 Robust Control

EVALUATING TECHNICAL COMMITTEE

(1) TC 2.5 Robust Control