Reachability under Bounded Delayed Inputs and Constrained Reinforcement Learning for Control Loop Safety and Efficiency

Josué Nkaya-Nkaya, Christophe Combastel (dir.), Jérôme Cieslak (co-dir.)

IMS Laboratory – Automatic group, ARIA team.
Univ. Bordeaux, CNRS UMR 5218, Bordeaux INP, F-33400 Talence, France

The following abstract refers to a contribution in the framework of the PhD thesis entitled: ***"Toward robust estimation and control algorithms hybridizing uncertain dynamic models and reinforcement learning"***

"Reinforcement learning is learning what to do — how to map observed situations to actions — so as to maximize a numerical reward signal. The learner is not told which actions to take, but instead has to discover which actions yield the highest reward by trying them" *(R.Sutton and A.Barto).* Using these learning techniques, high average efficiency can often be obtained to perform certain complex tasks which remain hard to model. However, the deployment of these solutions remains limited for safety critical applications. In this thesis work, reachable set computations are used to formally verify that the actions taken by a learning agent satisfy required safety properties, while enhancing the adaption to uncertain environments compared to pure model-based schemes.

This paper deals with safety verification of control loops hybridizing a baseline model-based controller and a constrained reinforcement learning (RL) agent aiming at exploiting available robustness margins to improve given performance metrics, without sacrifying proven safety properties. In order to preserve control-loop safety, the agent actions are constrained consistently with the underlying system physics. In this work, this is obtained by jointly bounding not only RL action values at some time, but also their variations between consecutive time steps. Disturbances modelling some lack of knowledge about the (physical) environment are characterized similarly. Then, safety verification is achieved through an original zonotopic reachability algorithm addressing the case of bounded inputs with bounded variations, reformulated as bounded delayed inputs. The resulting algorithm is shown to provide significant improvements compared to the classical bounded input case. A numerical example based on an aircraft roll dynamic illustrates the joint safety verification and performance enhancement achieved by the proposed hybrid control-loop scheme.

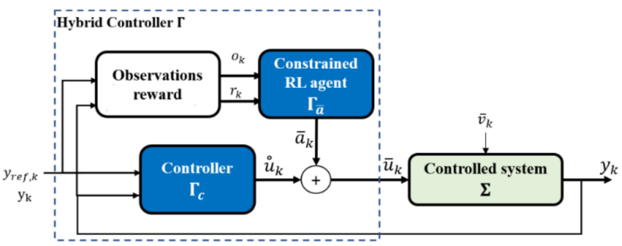

Fig. 1: Considered hybrid control setup

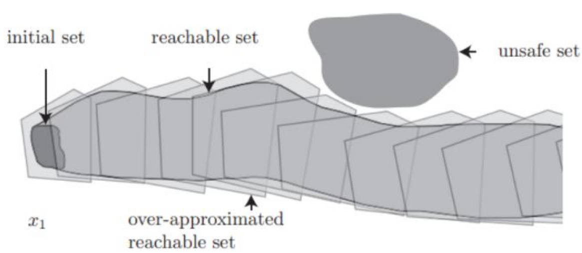

Fig. 2: Safety Verification & Set-based Reachability

Keywords
Safety, Robustness, Verification, Uncertain dynamic systems, Reachable sets, Zonotopes, Control, Reinforcement Learning.