On-policy Safe Reinforcement Learning under Input Saturation and State Constraints for Nonlinear Discrete Time Systems

Satya Vinay Chavan Marthi*[1], Mayank Shekhar Jha , and Didier Theilliol

[1]Centre de Recherche en Automatique de Nancy – Centre de Recherche en Automatique de Nancy (CRAN) – France

Résumé

This work proposes an on-policy Policy Iteration (PI) safe reinforcement learning (RL) framework for control of nonlinear discrete-time control-affine systems with input saturation constraints, ensuring safety throughout exploration and learning phases. This work emphasizes the importance of exploration noise while also addressing the challenges resulting from the conflict between control input saturation and exploration. Against the typically adopted off-policy PI based approaches for safe RL, our approach proposes on-policy PI algorithm under input saturation to iteratively learn the optimal control policy. This work also develops novel theoretical guarantees on the convergence to optimality under safety restrictions and evaluates the effectiveness through simulation studies, comparing its performance against the traditional Quadratic Programming based Control Lyapunov and Control Barrier function (QP-CLF-CBF) approach.

Mots-Clés: Safe Reinforcement Learning, Input Saturation, Control Lyapunov function, Control Barrier function, Optimal control, Neural Networks

*Intervenant