PhD position 2024–2027

Safe reinforcement learning control for Markov jump linear systems

Location

University of Sao Paulo (Brazil) with long visit at CRAN Université de Lorraine, CNRS (France)

Advisors

In Brazil: Oswaldo Costa, André Marcorin
In France: Jamal Daafouz, Romain Postoyan

Contact

oswaldo.costa@usp.br, jamal.daafouz@univ-lorraine.fr

Funding

This PhD is funded by a CNRS–USP grant, in the context of the CNRS International Research Center (IRC) at USP, and will begin in Fall 2024, lasting up to 3 years.

Objective

The features of reinforcement learning make it a very relevant candidate to control general dynamical systems in an efficient manner (near-optimally). Our objective is to ensure that systems controlled by reinforcement learning algorithms exhibit safety guarantees, in terms of stability and robustness, which are crucially lacking in the literature [3].

We will concentrate our efforts on stochastic dynamical systems described by Markov jump linear systems (MJLS) [4], using reinforcement learning techniques that are rooted in dynamic programming [2, 9]. A first major milestone will be to identify under which conditions robust stability guarantees can be established for dynamical systems whose inputs are generated by dynamic programming algorithms without any learning techniques. This step is not only necessary to address stability for systems controlled by reinforcement learning but is also of utmost importance in control engineering. Existing results are limited to deterministic dynamical systems, see, e.g. [1, 5–8], and not stochastic systems, which are very natural to model systems interacting with humans. In this context, MJLS are relevant in a range of fields including power systems, satellite control, cell growth etc. We will establish stability guarantees for MJLS, first, with optimal inputs for discounted costs, second, with inputs generated by dynamic programming. Afterwards, we will finally remove the assumption of the knowledge of the model and consider reinforcement learning techniques, with the goal of endowing the closed-loop system with robust stability guarantees.

Scientific environment

The PhD will take place at Escola Politécnica of the University of São Paulo. The student will be co-supervised by Oswaldo L.V.Costa (USP) and Andre Marcorin de Oliveira (UNIFESP) on the Brazilian side and J. Daafouz (CRAN) and R. Postoyan (CRAN) on the French side. A yearly mobility budget is secured, allowing the student to visit the French partners. Such mobility periods are expected and will be strongly encouraged. The student will therefore benefit from a scientifically rich environment, with opportunities for exchange and interaction with members of two world-class research institutions having strong expertise in control systems.

Application

We look for motivated candidates with a background in either control or electrical engineering with a strong interest in the theoretical side of engineering. Interested candidates should send an email to oswaldo.costa@usp.br including the following items:

1. an up-to-date CV in English;
2. official transcripts from each institution that you have attended (in French, Portuguese, or English);
3. a motivation letter.

References

[1] F. Berkenkamp, M. Turchetta, A. Schoellig, and A. Krause. Safe model-based reinforcement learning with stability guarantees. *arXiv:1705.08551*, 2017.

[2] D. Bertsekas. *Reinforcement learning and optimal control*. Athena Scientific, 2019.

[3] L. Buşoniu, T. de Bruin, D. Tolić, J. Kober, and I. Palunko. Reinforcement learning for control: Performance, stability, and deep approximators. *Annual Reviews in Control*, 46:8–28, 2018.

[4] O. L. V. Costa, M. Fragoso, and R. Marques. *Discrete-time Markov Jump Linear Systems*. Springer Science & Business Media, 2005.

[5] M. Granzotto, O. L. de Silva, R. Postoyan, D. Nesic, and Z.-P. Jiang. Policy iteration: for want of recursive feasibility, all is not lost. *arXiv preprint arXiv:2210.14459*, 2022.

[6] M. Granzotto, R. Postoyan, L. Buşoniu, D. Nešić, and J. Daafouz. Finite-horizon discounted optimal control: stability and performance. *IEEE Transactions on Automatic Control*, 66(2):550–565, 2020.

[7] M. Granzotto, R. Postoyan, L. Buşoniu, D. Nešić, and J. Daafouz. When to stop value iteration: stability and near-optimality versus computation. In *Learning for Dynamics and Control*, pages 412–424. Proceedings of Machine Learning Research, 2021.

[8] Z.-P. Jiang, T. Bian, and W. Gao. Learning-based control: A tutorial and some recent results. *Foundations and Trends in Systems and Control*, 8(3), 2020.

[9] R. Sutton and A. Barto. *Reinforcement Learning: An Introduction.* MIT Press, Cambridge, U.S.A., 1998.