TECHNISCHE UNIVERSITÄT MÜNCHEN
LEHRSTUHL FÜR INFORMATIONSTECHNISCHE REGELUNG
ORDINARIA: UNIV.-PROF. DR.-ING. SANDRA HIRCHE

May 24, 2024

M A S T E R ' S T H E S I S

Fusion of Model Predictive Control and Reinforcement Learning for the Safe Exploration of Unknown Systems with Latent States

Problem description:

Reinforcement learning (RL) has been used very successfully in recent years to learn optimal actions in complex environments. Unlike other learning-based methods, RL algorithms do not just passively rely on the availability of informative data and instead provide a principled approach to balancing exploration and exploitation. However, a severe concern with standard RL algorithms is that they generally do not provide safety guarantees for the learning phase, in which the agent actively explores its potentially safety-critical environment. In order to overcome this drawback, there is a growing interest in combinations of model-based RL and model predictive control (MPC) as MPC techniques offer a systematic way to incorporate safety considerations in the form of constraints [1].
An example of such a combination is presented in [2], where a posterior sampling RL approach is combined with MPC and formal safety and performance guarantees are derived. While the analysis assumes that all states of the system can be measured, this is not the case in many real-world applications. Recently, however, it has been shown that specific sampling techniques can be used to formulate learning-based MPC problems with constraint satisfaction guarantees even in the case that the states of the system are not directly measurable [3].
The aim of this work is thus to combine these two approaches into a secure RL algorithm for systems with latent states. Furthermore, formal safety and performance guarantees shall be derived, and the resulting algorithm shall be evaluated using simulations.

Tasks:

- Literature research on learning-based MPC and safe RL
- Implementation of a posterior sampling RL algorithm for systems with latent states
- Formal regret analysis
- Numerical evaluation of the proposed algorithm

Bibliography:

[1] L. Brunke, M. Greeff, A. W. Hall, Z. Yuan, S. Zhou, J. Panerati, and A. P. Schoellig, "Safe learning in robotics: From learning-based control to safe reinforcement learning," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 5, pp. 411–444, 2022.
[2] K. P. Wabersich and M. N. Zeilinger, "Cautious Bayesian MPC: Regret analysis and bounds on the number of unsafe learning episodes," *IEEE Transactions on Automatic Control*, 2022.
[3] R. Lefringhausen, S. Srithasan, A. Lederer, and S. Hirche, "Learning-based optimal control with performance guarantees for unknown systems with latent states," *arXiv preprint*, 2023.

Supervisor:	M.Sc. Robert Lefringhausen
Start:	2024
Delivery:	tba

(S. Hirche)
Univ.-Professor