Safe Reinforcement Learning in Shadow Mode

Technical University of Munich

Department of Informatics

Chair of Robotics, Artificial Intelligence and Real-time Systems

Background

With its successes in recent years, reinforcement learning (RL) has become very popular for training agents to complete a wide variety of tasks. However, it is not yet competitive for many cyber-physical systems, such as robotics, process automation, and power systems. One of the main reasons is that training on a system with physical components cannot be accelerated because simulation models do not exist or the simulation-to-reality gap is too big. During the long training time required, expensive equipment cannot be used and might even be damaged due to inappropriate actions of the reinforcement learning agent.

We want to address this problem by training a reinforcement agent in a so-called shadow mode while the system is operated by a conventional controller, which instantaneously performs well as it does not have to be trained. While learning the actual control task in shadow mode, the agent simultaneously learns in which situations it performs better than the conventional controller. and takes over, once confident enough and can keep learning, thus increasing the number of situations where it outperforms the baseline. We therefore always ensure that the performance is superior compared to only using conventional controllers or reinforcement learning and minimizes regret during learning.

One relevant domain for applying this principle is Autonomous Driving.

Description

The goal of this thesis is to expand the use of training in shadow mode by incorporating safety guarantees into the existing framework. The idea of training in shadow mode is useful for systems which cannot be fully and accurately simulated and can be applied to all systems for which some (sub-optimal) controller already exists. Initially, the system is controlled by the existing baseline controller. The RL agent is trained by simulating a subsystem, for which the dynamics are known and under the control of the agent. The observations of the RL agent are taken from the actual environment by adapting them according to the simulation. For example, a human might steer a vehicle, while the RL agent trains in the background. It simulates the ego-vehicle (the vehicle it is steering) and uses the real observations (such as distances to other vehicles) to infer the observations for its next state. A decision algorithm decides which action is executed on the real vehicle for each time step, i.e. it decides when the RL agent takes over. The topic of this work is extending existing implementations to incorporate safe reinforcement learning, such that unsafe actions will be avoided, while still providing useful feedback for training. This can be done by only allowing for safe actions and masking all unsafe actions through a safety layer, for example.[1] You will implement the existing verification methods of our chair for the shadow mode framework and evaluate the performance of the training algorithm compared to non-safe versions.

All programming will be done in our python-based CommonRoad-RL library.

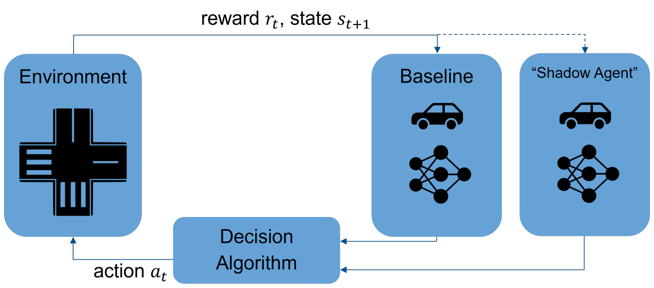

Figure 1: Framework for RL in shadow mode

Supervisor:
Prof. Dr.-Ing. Matthias Althoff

Advisor:
Philipp Gassert

Research project:
Reinforcement Learning in Shadow Mode

Type:
BA/MA

Research area:
Reinforcement Learning and Safety

Programming language:
Python

Required skills:
some familiarity with RL

Language:
English, German

Date of submission:
5. August 2023

For more information please contact us:

Phone: +49 (89) 289 18100

E-Mail: philipp.gassert@tum.de

Website:
www.ce.cit.tum.de/air/home/

Tasks

- Literature review on safe RL and background of shadow mode
- Setup of safety layer for shadow mode
- Training and tuning of models in shadow mode
- Evaluation of performance and identification of exemplary scenarios

References

[1] Hanna Krasowski, Xiao Wang, and Matthias Althoff. Safe reinforcement learning for autonomous lane changing using set-based prediction. In *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*, pages 1–7. IEEE, 2020.

Technical University of Munich

Department of Informatics

Chair of Robotics, Artificial Intelligence and Real-time Systems