Institute for Design and Control
of Mechatronical Systems

Master's Thesis:

Overtaking on Country-Roads based on Reinforcement Learning

Author: Stefan Angerbauer

Supervisors: Prof. Dr. Luigi del Re
DI. Daniel Adelberger

Finished: June 2021

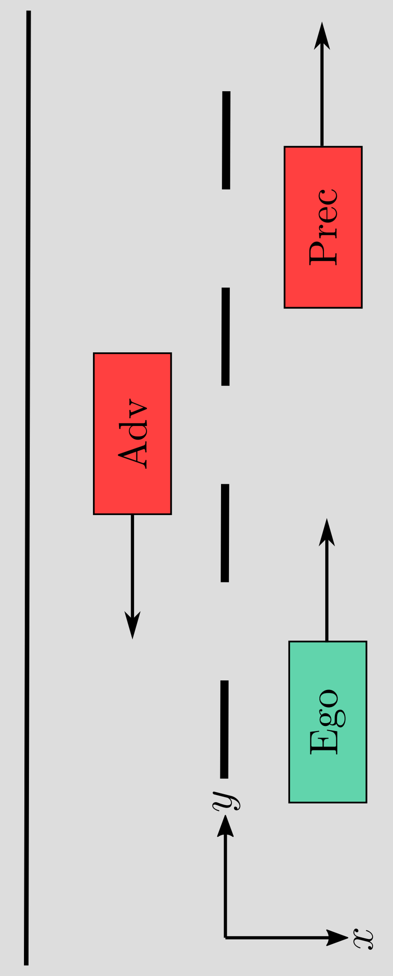
Abstract

Artificial Intelligence (AI) has made significant progress within the recent years, reaching from image-recognition over natural-language-processing to applications in robotics. One field of research of artificial intelligence, which can be used for model-free control of physical systems is so-called Reinforcement-learning (RL). Most algorithms of this class aim to learn desired behavior from a reward-distribution, designed by humans. The discipline promises high potential for solving autonomous driving tasks.

In this work a subproblem of autonomous driving is considered: planning and carrying out overtaking-maneuvers on County-Roads autonomously. To tackle this issue, so called Deep-Q-Networks (DQNs) are used. DQNs are artificial neural networks, trained by a specific algorithm. To enable this training, a training-environment is created, which is equipped with real-world-measurements to generate realistic traffic-situations.

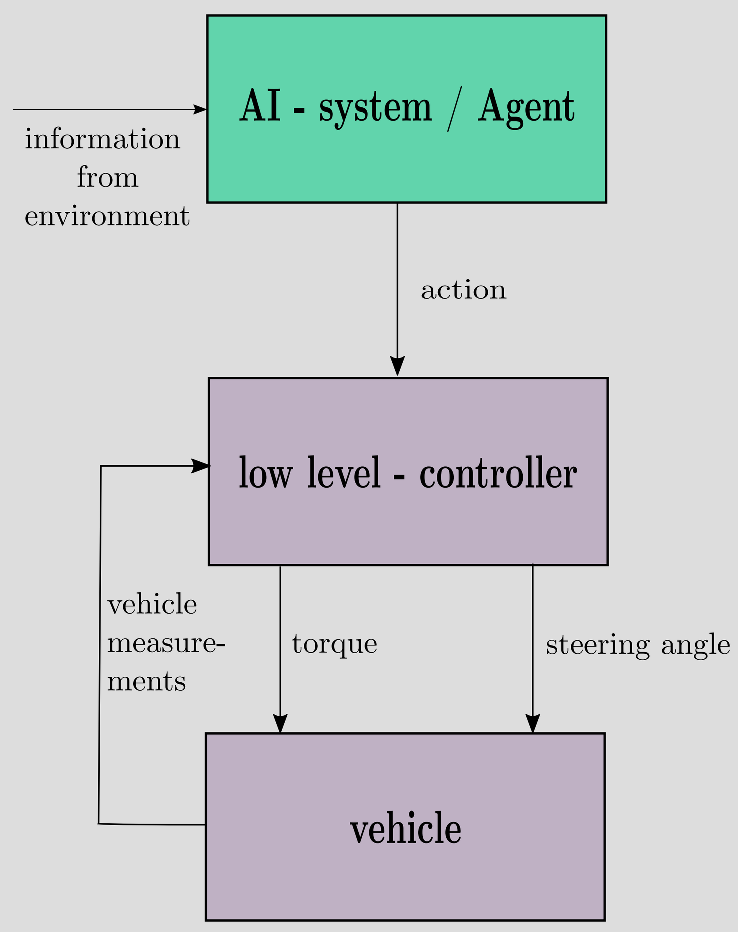
The topic of communicating with the algorithm is discussed. Concretely the questions, how to represent the environment in a manner, such that the neural networks can use it, and how to select the reward-distribution do achieve desired behavior, are answered. This knowledge is used, to train an AI-system, to solve the overtaking task as good as possible. Successes and failures of the algorithm are discussed. Finally for one selected case the decision boundaries of the DQN are visualized and the systems behavior is discussed in this light.

Proposed Scenario:



The considered scenario is a straight road with two lanes. It is assumed that three vehicles are on the road at any time. The one, that is controlled by the system to be developed is named Ego. The two other vehicles are called Prec and Adv. Prec is the vehicle, that should be overtaken and is assumed to be initially located in front of the Ego. Adv is the vehicle, that drives on the opposite lane and in the opposite direction as Prec and Ego. The goal is, to overtake Prec as soon as possible while not crashing in any of the vehicles. The reason to choose this scenario is that it includes most of the relevant traffic scenarios that can lead to crashes

Proposed Approach:



To solve the problem, machine-learning, more precisely Reinforcement-Learning, is applied. To increase the algorithms performance, all low-level-control-related tasks, such as adjusting the steering angle and engine torque are outsourced to a control system. The AI-algorithms task should be only to dictate, which action should be carried out, where one action denotes an entire sequence of control inputs to the physical system, such as turning the steering wheel to the left, until the center of the road is reached, and then turning right until the center of the left lane. This way, the algorithm only has to learn those informations, that would be difficult to hard-code.

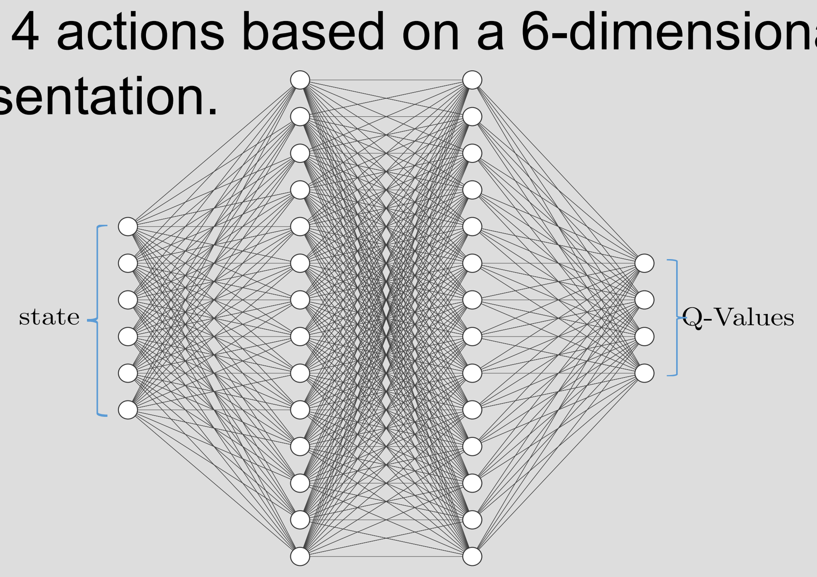
DQNs

Deep-Q-Networks (DQNs) are neural networks used to estimate the Q-value:

$$Q^{\pi}(s_t, a_t) = E_{\pi}\{R_t | s_t, a_t\}$$

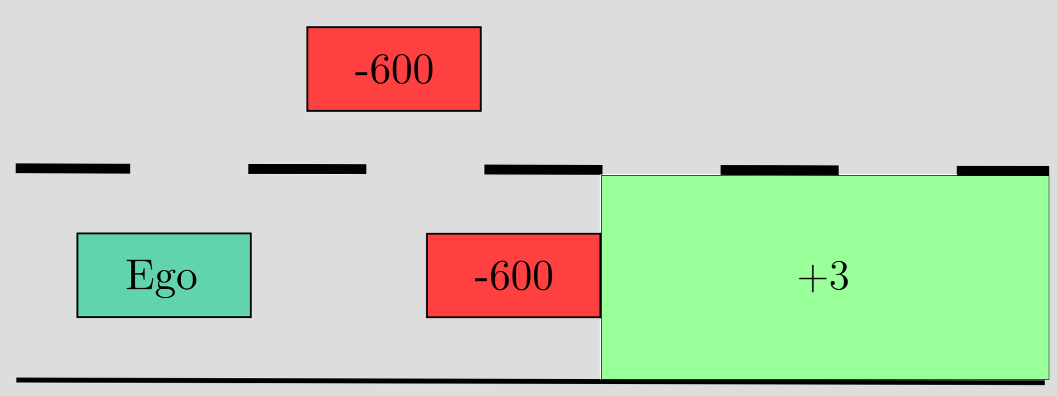
$$R_t = \sum_{k=0}^{\infty} \gamma^k r_{t+k+1}$$

Thereby a_t denotes an action, s_t a state and r_t the reward at time t. The neural networks task is to estimate the Q-value, given a state. It is a measure for how much reward can be expected performing an action in a given state under the current policy π. The following picture shows a DQN rating 4 actions based on a 6-dimensional state-representation.



Outcomes

The so-called hyperparameters for the algorithm had to be determined experimentally. The reward-function depicted below is an example of such a hyperparameter. It turned out to be a good choice to only punish crashes and unneccessary lane changes and only reward being in the desired state.



Using this reward-distribution a DQN was trained. Evaluation on a large number of attempts yielded:

- over 70 % of maneuvers were successfull
- under 0.4 % failure/crash
- neither success nor failure in all other cases

Conclusions and Outlook

In this thesis an AI-based overtaking-assistance system was designed. Thereby the scenario was restricted to be a straight road with the agent itself and two other vehicle, Prec and Adv, on it. A suitable behavior was found using the DQN-approach and carrying out many simulations to determine a good reward-distribution, state-representation and hyper parameter-choice. The obtained agent was able to succeed in approximately 70% of all cases, crashed in just under 0.4% and did not overtake, but not crash either in the rest of cases.
Future applications of the designed system are either as one component of a larger mechatronical overtaking-system or as an initialisation to a more complex AI-based overtaking system.