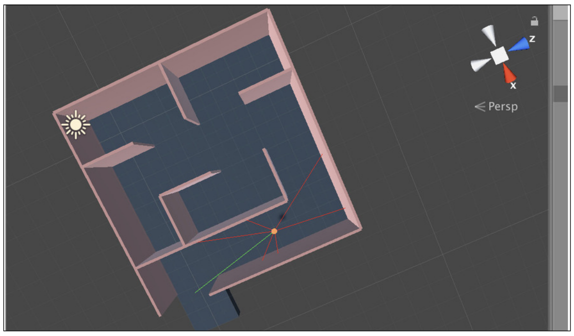
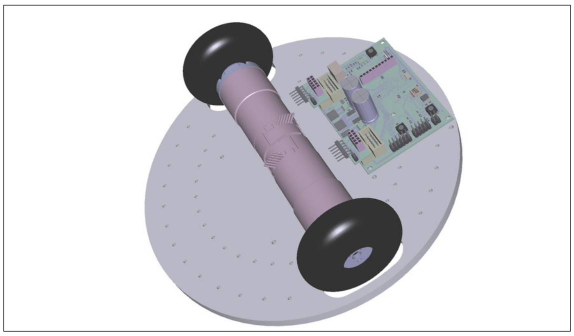
David
Wright

Student	David Wright
Examiner	Prof. Dr. Guido Schuster
Subject Area	Sensor, Actuator and Communication Systems

A Reinforcement Learning Based Rover Navigation



Inference of a trained policy in the simulation environment in Unity
Own presentment

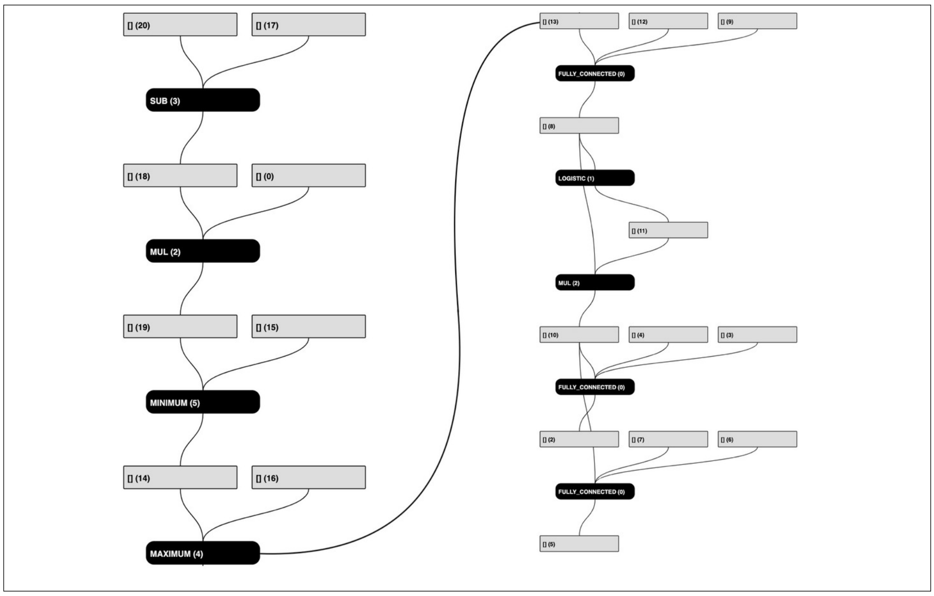


A 3D model of the rover
Erwin Brändle

Introduction: While reinforcement learning and the underlying theory have been well established since the 1950's, it had been largely unnoticed up until fairly recently. Although lately there has been a significant increase in research and development on this topic, only little thereof studied embedded system applications. The goal of this project was to overcome the obstacles presented by the combined tasks of simulation, training and implementation on the target hardware. Therefore, a rover navigation was implemented, which is modelled by a policy that was trained using reinforcement learning techniques.

Approach / Technology: The environment in which the rover was to learn to navigate was a maze. The observations which the navigation was based on were LIDAR measurements resulting from seven VL53L1X sensors by STMicroelectronics. For the target hardware, a Drive Control Unit based on a Tiva C Series Cortex-M4 processor by Texas Instruments was used. This thesis begins by establishing the theoretical fundamentals of reinforcement learning as well as the theory it is based on, after which the simulation, training and deployment procedures are described in detail. After the conclusion of the theoretical foundation, the simulation and training procedure, based on Unity and the ml-agents toolkit, are presented. Thereafter, the workflow for the policy inference on the target hardware is demonstrated, using the tensorflow lite port for microcontrollers.

Result: The results presented in this thesis consist of a simulation for the agent and environment, a setup for the training of the policy, and a bare metal implementation for the inference thereof. This thesis serves as a basis for future projects which aim at the goal of developing reinforcement learning applications for embedded systems by providing a complete toolchain and workflow for its implementation. Moreover, the results can also be utilized as a guideline for any tensorflow based machine learning project targeting bare metal microprocessor platforms.



The resulting model for the policy
Own presentment

HSR
HOCHSCHULE FÜR TECHNIK
RAPPERSWIL

FHO Fachhochschule Ostschweiz