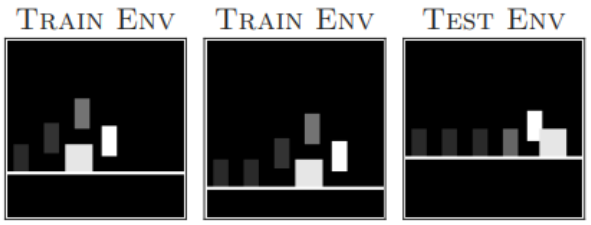
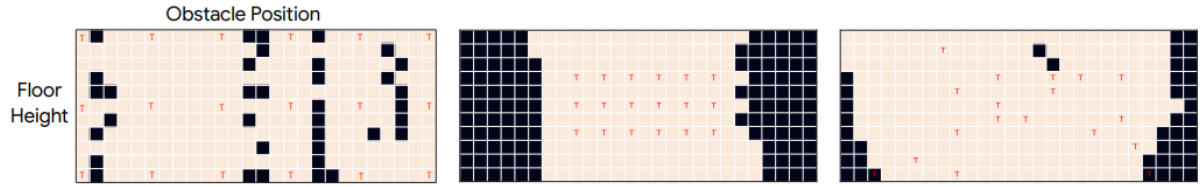
Project “Generalization in RL”

Aim: Understanding generalization of agents across a distribution of environments.

Tachet des Combes et al. (2018) introduced the Jumping Task, a configurable environment for testing the generalization of agents trained on a subset of the available environments. We will focus on two options to configure the environment, its floor height and the obstacle position. This environment has observations in pixel space, making it necessary to use deep learning methods to learn a policy.



A good experimental setup based on this environment is given by Agarwal et al. (2020). They propose three different sets of training environments, namely a wide, narrow and random grid. Testing environments are all other configurations.



(a) “Wide” grid (b) “Narrow” grid (c) Random grid

This project focuses on Behaviour Cloning, i.e. supervised learning where the optimal action given an input image is learned. Therefore the first task will be implementing a codebase that can generate the experimental setups shown in the figure above, suitable for Behavior Cloning. Thus, methods to generate optimal episodes, data-loading, training and evaluation procedures need to be implemented.
After setting up the codebase, there are two possible options for experiments. First, there is a selection of augmentations that can be applied to this setting to improve generalization. Second, one can use Deep Q-Learning instead of Behaviour Cloning to learn a policy and investigate the differences in generalization capabilities between those two.

The focus is a **clean codebase and reproducibility**, you will benefit your whole (study) career from those. You can use packages for the RL algorithms, but I recommend to program your own implementations as you learn more.

Environment Github: https://github.com/google-research/jumping-task
Environment Paper: https://arxiv.org/abs/1809.02591
Experimental Setup Paper: https://arxiv.org/abs/2101.05265 (Figures taken from here)
Augmentations in RL Paper: https://arxiv.org/abs/2004.14990