

# Proposal Draft for ADLR Project

Group 06: Katharina Hermann and Ferenc Török

The aim of our project is to train a Reinforcement Learning agent, which is able to carry out movements from a starting point to a goal point through challenging environment. Under challenging environment we understand a workspace with obstacles in the manipulation space with which the robot arm must not collide during movement.

The focus for training this agent is, that we use a modification of Hindsight Experience Replay [1]. When the robot's chosen trajectory fails to reach the goal in a specific workspace, instead of simply relabeling the goal to match the trajectory, as it is done for HER [1], we sample a new goal and a new workspace, for which the trajectory would have been a good solution.

In [2] the authors argue that with HER the robot does not learn to navigate through narrow passages, since the high probability of collisions with the obstacles lead to a high negative reward in those regions and prevent the robot from exploring those regions. To remedy this issue, they use expert demonstrations (generated with traditional MP algorithms) to show the robot a way to navigate through these challenging workspaces with narrow passages. This method led to very good performance results, regarding efficiency and accuracy.

However, we believe that generating challenging environments to relabel the robot's goal is another way to show the robot how a perfect trajectory for a very narrow passage would have looked like, without having to model the perfect trajectories through challenging environments via traditional motion planners first.

The guidelines along which we will present our final project proposition are as follows:

- We are going to model the robot as a point robot in 2D
- We are going to apply a grid on the workspace to be able to distinguish between free space and obstacles.
- The sampled workspace which matches the failure trajectory should be challenging for the robot. Therefore the obstacles are placed such that they form a narrow passage along the trajectory.
- We are going to incorporate the ideas of Neural Motion Planning [2], Hindsight Experience Replay [1] and Generalized Hindsight Experience Replay [3].
- The final goal is to evaluate the performance of the trained agent in previously unseen environments and compare it to the performances of the DDPG-MP agent [2]

## Planned milestones

1. Initial setup
  - Setup of the environment - a grid with pixels that indicates whether there is an obstacle or not.
  - Implementation of the point robot's logic
2. Training Process with Hindsight Experience Replay and relabeling of the environment

- Implementing the DDPG algorithm or getting familiar with an off-the-shelf implementation.
  - Implementation of a strategy to generate challenging workspaces for relabeling
  - Training of the agent with the extended HER strategy of workspace relabeling
3. Evaluation and Comparison to other methods such as DDPG-MP [2]
- Evaluation of the trained agent in previously unseen workspaces possibly with different configurations (robot’s complexity, complexity of the environment, variation in the generation strategies of the new workspace)
  - Comparison against the DDPG-MP [2] algorithm with the metrics used in the paper.

## Bibliography

- [1] M. Andrychowicz, F. Wolski, A. Ray, J. Schneider, R. Fong, P. Welinder, B. McGrew, J. Tobin, O. P. Abbeel, and W. Zaremba, “Hindsight experience replay,” in *Advances in neural information processing systems*, 2017, pp. 5048–5058.
- [2] T. Jurgenson and A. Tamar, “Harnessing reinforcement learning for neural motion planning,” *arXiv preprint arXiv:1906.00214*, 2019.
- [3] A. C. Li, L. Pinto, and P. Abbeel, “Generalized hindsight for reinforcement learning,” *arXiv preprint arXiv:2002.11708*, 2020.