

# ADLR WS21 - Project Proposal Team 04

Marc Hauck  
*Robotics, Cognition, Intelligence*  
*Technical University of Munich*  
Munich, Germany  
ge65qoy@mytum.de

Baris Tura  
*Robotics, Cognition, Intelligence*  
*Technical University of Munich*  
Munich, Germany  
baris.tura@tum.de

## I. OBJECTIVE

Motion planning lies at the core of robotics. However, traditional methods are not only demanding in terms of computation, but also require prior knowledge of the environment. The real conditions of the environment in which a robot arm operates may be well outside of what the environment model entails, making swift responsiveness to unexpected obstacles a desirable trait for robots. Learning methods were found to be successful in unseen environments and in this work, we aim to harness reinforcement learning algorithms to guide a robotic arm through a dynamic environment where the instantaneous locations and trajectories of obstacles are unknown to the agent. Hereby, we aim to supplement local guidance of reinforcement learning with a global guidance algorithm, so as to keep the balance of exploitation and exploration. Whereas deterministic real-world environments are not in need of new technological advances, there is much work to be done in unknown and dynamic environments, and we consider this an important matter as environments with dynamic obstacles and other agents span a wide range of applications with increasing market potential, such as Human Robot Collaboration (HRC).

## II. RELATED WORK

Previous works consist of motion planning in static environments for joint robots, and motion planning in dynamic environments for mobile robots. Jurgenson et al. [1] propose a DDPG based algorithm with expert knowledge for a planar RRRR robot in a static yet unknown environment. They claim that supervised learning is heterogeneous in terms of data distribution and lacks sufficient data close to object boundaries; and that imitation learning leaves no room for exploration, especially in unseen environments as the expert never collides with the objects. Their results obtained with the expert knowledge supported algorithm, named DDPG-MP, show superior results to those of classic DDPG and HER algorithms.

Wang et al. [2] guide an agent through a 2D environment comprising of static and dynamic obstacles using ConvNets for scene understanding and an LSTM+MLP network for action selection. Here, they make use of A\* search as global guidance to enforce minimal traversal through the environment while still avoiding the dynamic obstacles. All the obstacles are unknown a priori, and are perceived within a certain field of view by the robot. For action selection, classic DDQN

algorithm is used in conjunction with an LSTM network to have temporal reasoning. They name the approach Globally Guided Reinforcement Learning and obtain state-of-the-art results.

## III. TECHNICAL OUTLINE

To begin, we will aim to guide a simple point robot through a 2D grid world with static obstacles using different algorithms such as DDPG and HER. By doing so, we hope to observe the effects different algorithms have on this particular application. Thereafter, we are going to test both expert knowledge (similar to [1]) and global guidance (similar to [2]) as an addition to the frame and see their impact. Both approaches will boost the performance of our algorithm by guiding the robot with a reachable reward, specially in larger environments where the reward can become sparse. By doing so, we hope to create the backbone of our model, deciding on a tentative algorithm and network structure. On top of that, we are going to compare the results we obtain with [1], as they state that a combination of DDPG-MP and classic motion planning methods may improve their results.

From this point, we will constantly extend the environment as well as our model. In addition to [1], we will experiment with different, more realistic obstacle shapes and investigate the effects of only partial knowledge of the environment by limiting the robots field of view (FOV). Building on our model with a point robot, we then plan to transition to a robot with three joints, while keeping the environment static and planar. This implies changes to the motion planning algorithm, as it is not only important for the end effector of the robot to avoid collisions, but also for the rest of the robot arm.

Having successfully guided the joint robot in the static environment, we aim to introduce motion to the obstacles. To make an agent capable of dealing with dynamic obstacles is important in order to create more realistic environments inspired by possible future applications like HRC. At this point, we foresee large changes to our previous model, as dynamic obstacles add a temporal dimension. While the planning method of [1] may already be able to cope with moving obstacles, it could be beneficial to extend the model with an LSTM network similar to [2] to extract the temporal information of the obstacle movements.

## REFERENCES

- [1] T. Jurgenson and A. Tamar, "Harnessing reinforcement learning for neural motion planning", *arXiv preprint arXiv:1906.00214*, 2019.
- [2] B. Wang, Z. Liu, Q. Li, and A. Prorok, "Mobile robot path planning in dynamic environments through globally guided reinforcement learning", *IEEE Robotics and Automation Letters*, 5(4), pp. 6932-6939, 2020.