ST/MT: Benchmarking robotic multi agent reinforcement learning

Motivation:
In many industrial settings it is desirable to maximise throughput in order to speed up production. One way of doing this is utilizing multiple robots. When dealing with multiple robots, in addition to solving the motion planning problem, we also need to decide which robot does what and at what time.

Formulating this as a reinforcement learning problem is relatively straightforward, but depending on the formulation, we need to make sure that the robots stay safe (i.e. do not avoid collisions or actions that would damage the objects we are interacting with).

An example for such a multi robot setting is typing with a multi fingered robot hand, or robotic pick and place settings. In settings like these, where we can achieve task success with only a single robot, reinforcement learning approaches tend to find conservative suboptimal policies, where e.g., only a single robot is utilized for solving the task in order to avoid penalties for doing the wrong thing with a robot.

Goal:
In this work, we want to set up a benchmark for multi robot settings, where coordination would help the overall completion time of the task, and train baseline agents (e.g. using PPO/SAC/TD3) to compare algorithms, and how they scale with data, or how they work at all.

Compared to existing multi-agent reinforcement learning benchmarks, we are interested in settings where a single (or a subset of all agents) is enough, but suboptimal to achieve the task.

In addition, we want to explore what can be changed in order to achieve better utilization of many robots, without having to specifically shape rewards separately in all environments.

Interested?
Ideally, you already have strong experience with Python and both theory and practice of robotic reinforcement learning.

For further information or to apply for this project, please contact Valentin Noah Hartmann (*valentin.hartmann@inf.ethz.ch*) with your CV and transcripts.