Language-conditioned Meta-Reinforcement Learning

Robotics and Embedded Systems

Background

Have you ever imagined a scene where robots can execute multi-manipulation tasks according to your language instructions, such as ”open the door”, ”close the drawer”, and other tasks in daily life, as illustrated in Figure 1. This scenario can be realized through the utilization of imitation learning (IL) or reinforcement learning (RL) techniques. However, the successful execution of multi-manipulation tasks presents a significant challenge in the domains of IL and RL, due to the requirement of coordination among multiple actions. Furthermore, the reliance on extensive expert demonstrations or the design of carefully crafted reward functions as a means of grounding language instructions with manipulation tasks remains a limiting factor in the implementation of such systems.

Figure 1. A demonstration of language-conditioned method for multi-manipulation tasks [1].

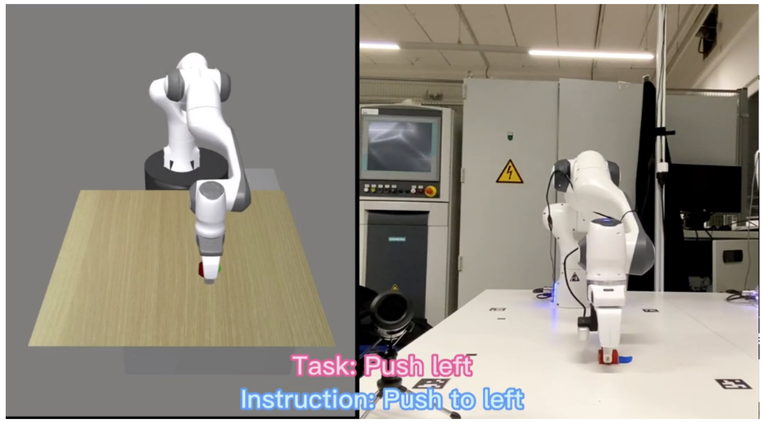
To alleviate the above limitations, meta-reinforcement learning (meta-RL) [3] is a good choice. As a subfield of reinforcement learning, meta-RL is concerned with the acquisition of the ability to learn to learn, and its goal is to develop a general policy that can be adapted to various tasks with limited fine-tuning. Language-conditioned meta-RL represents a specialized approach within meta-RL that leverages natural language instructions to specify the task at hand, making it easier for the agent to understand and execute the task. This utilization of language conditioning enhances the generalization capacity of the meta-RL algorithm, as the agent can learn from a wider range of tasks and instructions. Through language-conditioned meta-RL, the agent is capable of more effectively learning and executing multi-manipulation tasks, even those that may not have been present in the training process.

learn to learn

Your Tasks

In this thesis, you will develop language-conditioned meta-RL algorithms for the purpose of addressing multi-manipulation tasks, e.g., Kitchen Environment[2], Meta-World benchmark[4]. Specifically, your focus will be on:

- learning about language-conditioned meta-RL for manipulation tasks;
- learning how to transfer a trained policy to a real Franka Panda robot.



Requirement

- **Master/Bachelor Thesis**.
- High self-motivation and passion on research.
- Six month working time.
- Existing knowledge about RL or meta-RL will be a bonu.

Supervisor: Prof. Alois Knoll

Advisor: Xiangtong Yao: xiangtong.yao@tum.de

Lehrstuhl für Echtzeitsysteme und Robotik,

Fakultät für Informatik, Technische Universität München

[1] Michael Ahn, Anthony Brohan, Noah Brown, Yevgen Chebotar, Omar Cortes, Byron David, Chelsea Finn, Keerthana Gopalakrishnan, Karol Hausman, Alex Herzog, et al. Do as i can, not as i say: Grounding language in robotic affordances. *arXiv preprint arXiv:2204.01691*, 2022.

[2] Abhishek Gupta, Vikash Kumar, Corey Lynch, Sergey Levine, and Karol Hausman. Relay policy learning: Solving long-horizon tasks via imitation and reinforcement learning. *arXiv preprint arXiv:1910.11956*, 2019.

[3] Jane X Wang, Zeb Kurth-Nelson, Dhruva Tirumala, Hubert Soyer, Joel Z Leibo, Remi Munos, Charles Blundell, Dharshan Kumaran, and Matt Botvinick. Learning to reinforcement learn, 2016.

[4] Tianhe Yu, Deirdre Quillen, Zhanpeng He, Ryan Julian, Karol Hausman, Chelsea Finn, and Sergey Levine. Meta-world: A benchmark and evaluation for multi-task and meta reinforcement learning. In *Conference on robot learning*, pages 1094–1100. PMLR, 2020.