Internship/Forschungspraxis (possible to be extended as Master Thesis)

Large Language Model based Task and Motion Planning

Background

The integration of large language models in robotic manipulation represents a remarkable advancement in the field of robotics and artificial intelligence. Owing to the emergent abilities of large language models (LLMs) [1], embodied agents can be empowered by the rich knowledge and commonsense about the world and the strong reasoning capabilities from LLMs. Therefore, the robotics research community has seen some recent attempts to leverage LLMs to solve Task and Motion Planning (TAMP) problems [2,3,4]. However, most of them are limited to very simple action sets. It is underexplored how to ground the high-level planning into executable low level robot motions. In this study, we plan to establish a real-world experimental setup involving multiple manipulation tasks, and compare the performance based on different action sets (End-to-End learned policy and our skill library [5][6]).

Your Tasks

1. Build a real-world setup with several manipulation tasks.
2. Develop a software interface to covert task plans generated from LLMs to executable motion plans or Behavior Tree based on our existing skill library.
3. Experimental validation on a set of manipulation tasks.

Requirement

- Highly self-motivated;
- Experiences or knowledge from related Robotics courses;
- Python programming experience.

Supervisor: Prof. Sami Haddadin

Advisor: Yansong Wu, Dr. Fan Wu

f.wu@tum.de yansong.wu@tum.de

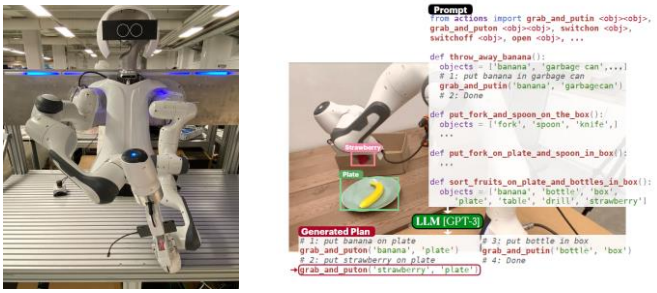

Figure 1 Overview of the experiment setup.

Lehrstuhl für Robotik und Systemintelligenz
TUM School of Computation, Information and Technology
Technische Universität München

[1] J. Wei, Y. Tay, R. Bommasani, C. Raffel, B. Zoph, S. Borgeaud D. Yogatama, M. Bosma, D. Zhou, D. Metzler et al., "Emergent abilities of large language models," Transactions on Machine Learning Research (TMLR), 2022
[2] Liang J, Huang W, Xia F, et al. Code as policies: Language model programs for embodied control[C]//2023 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2023: 9493-9500.
[3] Ahn M, Brohan A, Brown N, et al. Do as i can, not as i say: Grounding language in robotic affordances[J]. arXiv preprint arXiv:2204.01691, 2022.
[4] Singh, Ishika, et al. "Progprompt: Generating situated robot task plans using large language models." *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023.
[5] Johannsmeier L, Haddadin S. Can we reach human expert programming performance? A tactile manipulation case study in learning time and task performance[C]//2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2022: 12081-12088.
[6] Yansong, Fan, Lingyun, et al. 1 kHz Behavior Tree for Self-adaptable Tactile Insertion. (under review)