PhD position, MLIA team, ISIR, Sorbonne Université,

Paris, France

“Leveraging language models and reinforcement learning for generating instructions and interacting with robots”

Contact :

- Laure Soulier, laure.soulier@sorbonne-universite.fr
- Nicolas Thome, nicolas.thome@sorbonne-universite.fr
- Olivier Sigaud, olivier.sigaud@sorbonne-universite.fr

Location: Sorbonne Université, Pierre et Marie Curie Campus, 4 Place Jussieu, Paris, France

Candidate profile:

- Master degree in computer science or applied mathematics, Engineering school.
- Background and experience in machine learning and deep learning. Background in Natural Language Processing and/or Reinforcement Learning is appreciated.
- Good technical skills in programming.

How to apply: please send a cv, motivation letter, grades obtained in master, recommendation letters when possible to laure.soulier@sorbonne-universite.fr, nicolas.thome@sorbonne-universite.fr and olivier.sigaud@sorbonne-universite.fr

Start date: October/November 2023

Keywords: text generation, reinforcement learning, robot instructions

Context

This PhD is founded by the PILLAR European project (2023-2027) which aims at developing a new generation of robots that can build on the experience acquired during the robots’ lifetime to fulfill the wishes of their human designers/users in real-life applications.

Autonomous agents require reasoning and planning strategies for performing tasks. The semantics captured by large language models can enhance the decision process at different levels [1]. Natural language can serve for building and clarifying the planning strategy, and therefore the actions done by a robot. Several works have addressed instruction identification as abstract representation [2, 4, 5] or natural language expression, but the limited data supervision is often a challenge [3, 5].

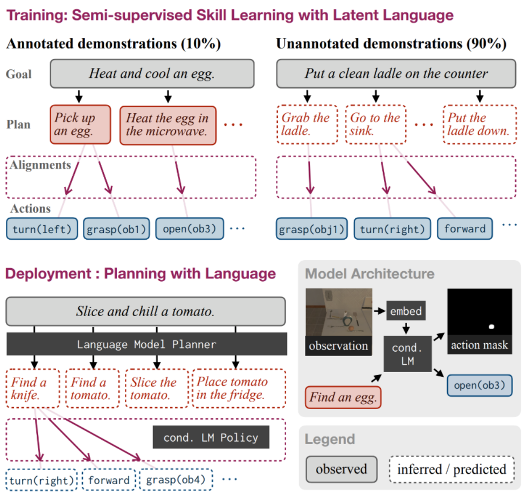

Figure 1. illustration from [5]

In this thesis, we envision working on the generation of natural language instructions and improving current models. Our objective is to enhance the semantics behind objects to identify the most relevant actions/sub-actions and design hybrid models combining reinforcement learning and language models to generate accurate instructions. An example of expected outputs is presented in Figure 1.

Research directions

Multi-modal representation learning for instructions

Learning from a variety of multi-modal inputs is crucial in robotics, where perceptive inputs of the environment are heterogeneous, *e.g.* images, instructions (text), times series, while output control commands can be diverse (continuous, discrete). Foundation models such as CLIP provide a promising solution for sustainable learning from text and images, since they can effectively perform zero-shot image classification. **The main goal of this axis is to propose solutions for designing new multi-modal robotic foundation models (RFM) that can handle perceptive inputs and output commands.** One key challenge is related to data collection; gathering representative and aligned samples with an increasing number of modalities quickly becomes prohibitive, and no such universal dataset is currently available. CLASP [7] is a very recent extension of CLIP for robotics, but the model is trained on a small and specific dataset**. To fully leverage the expressivity potential of LLM and LMM, the main hypothesis is to build on compositionality**. Firstly, we will explore instruction generation as a common intermediate representation between perception and control [10] such as proposing new ways of pre-training RFM with instructions on larger datasets, *e.g.* ALFRED. The assumption is that the compositionally power of LLMs can drive the compositionality of actions. We will also enforce disentangling properties in the embedding for learning semantic factors of variations, *e.g.* by adapting recent approaches [12] to the RFM context. Secondly, CLIP/CLASP losses are based on the alignment of modalities, which is arguably sub-optimal when the modalities convey complementary information. We will design losses able to learn more flexible ways of information sharing, *e.g.* based on masking. Finally, we want to explore the possibility to perform instruction fine-tuning with RFM, as done in NLP and text/vision. Especially, the possibility to use independently learned embedding for the different modalities and to use sustainable learning schemes, *e.g.* adapters, will be assessed.

Hybridizing RL and LLMs for richer instruction-following agents

Reinforcement learning is the natural framework to endow our robot with task learning capabilities. However, in the context of the PILLAR project we do not want the reward function corresponding to each possible task to be engineered in advance. Rather, we want the robot to leverage natural language interactions with its users to shape its own reward function on demand. This raises the issue of grounding these natural language interactions into the sensorimotor interactions between the robot and its environment so as to turn instructions, descriptions or feedback into an actionable reward function. So far, existing instruction-following agents have provided an unsatisfactory solution to the corresponding issues, as they generally rely on a restricted vocabulary and language structure instead of leveraging the full richness of natural language interactions [6]. A line of research in this PhD will consist in moving towards a more powerful framework to guide a reinforcement learning agent with natural language interactions.

Proactive interaction for solving uncertainty in instruction generation

The objective of this third part is to address uncertainty during text generation. Indeed, while generating instructions, some actions might be ambiguous and give rise to different interpretations, and therefore actions. For instance, the instruction “cutting a lemon” might imply to cut the lemon into slices or into wedges. Also, in the case of a complex environment with sensors, “activate the lever” might imply different actions depending on the lever texture, location, or size. These ambiguous or complex situations engage the text generation model in a proactive setting in which complementary information might be provided by humans (i.e., if the goal is wrongly worded), external documents/resources (i.e., technical notices) or the environment itself (i.e., from sensors). Our aim here is to design instruction generation models able to 1) **identify the uncertainty** while generating an instruction, 2) **formulate the different possible options** to interact with a human, external resources, or the environment, and 3) **revise the instruction generation** according to the external feedback. This setting opens a new perspective of designing proactive agents to perform their actions while anticipating possible errors and generating interactions with humans. From the technical point of view, the literature review around this task is centered around two main concepts:

- Uncertainty quantification in multi-modal context and language models [8,9]
- Disambiguation and clarification with language models [10, 11]
- Explaining and planning in robotics [13]

References

[1] Michael Ahn et al. “Do As I Can and Not As I Say: Grounding Language in Robotic Affordances”. In: arXiv preprint arXiv:2204.01691. 2022.

[2] Jacob Andreas et al. “Learning with Latent Language”. In: NaaCL 2018. url: https://aclanthology.org/N18-1197

[3] Haonan Chen et al. “Enabling Robots to Understand Incomplete Natural Language Instructions Using Commonsense Reasoning”. In:ICRA 2020. url: https://doi.org/10.1109/ICRA40945.2020.9197315

[4] Athul Paul Jacob et al. “Multitasking Inhibits Semantic Drift”. In: Naacl-HLT 2021. url: https://aclanthology.org/2021.naacl-main.421

[5] Pratyusha Sharma et al. “Skill Induction and Planning with Latent Language”. In: ACL 2022. url: https://doi.org/10.18653/v1/2022.acl-long.120

[6] Luketina, Jelena, et al. "A survey of reinforcement learning informed by natural language." *arXiv preprint arXiv:1906.03926* (2019).

[7] Rana, Krishan, Andrew Melnik, and Niko Sünderhauf. "Contrastive Language, Action, and State Pre-training for Robot Learning." *arXiv preprint arXiv:2304.10782* (2023).

[8] Baan et al.,”Uncertainty in Natural Language Generation: From Theory to Applications”. 2023. url: https://arxiv.org/pdf/2307.15703.pdf

[9] Zhanf et al. “Multimodal Chain-of-Thought Reasoning in Language Models”. 2023. url: https://arxiv.org/pdf/2302.00923.pdf

[10] Kuhn et al. “CLAM: Selective Clarification for Ambiguous Questions with Generative Language Models”. 2023. url: https://arxiv.org/pdf/2212.07769.pdf

[11] Deng et al. “Prompting and Evaluating Large Language Models for Proactive Dialogues: Clarification, Target-guided, and Non-collaboration” 2023. url: https://arxiv.org/pdf/2305.13626.pdf

[12] S. Hao, K. Han, K. Wong. Learning Attention as Disentangler for Compositional Zero-shot Learning. CVPR 2023.

[13] Wang et al. Describe, Explain, Plan and Select: Interactive Planning with Large Language Models Enables Open-World Multi-Task Agents. 2023