Exploring Spatial Reasoning Abilities of Large Language Models

Shengqiang Zhang

January 2024

- Supervisor: Shengqiang Zhang

- Examiner: Prof. Hinrich Schütze

- Open to: MSc.

- Prerequisites: Good programming background (Python, PyTorch/TensorFlow, etc), basic knowledge and experience of natural language processing (NLP) and deep learning.

- Introduction:
  Though large language models (LLMs) show impressive performance on various tasks, the spatial reasoning abilities of LLMs are less studied by prior works. However, the spatial reasoning abilities are crucial for some tasks such as robotic path planning and understanding spatial relationships of objects in human's instructions.

  This project aims at evaluating and improving the spatial reasoning abilities of LLMs. Compared to prior work, we are more focused on LLMs' long-horizon spatial relationship understanding and sequential spatial planning capabilities. We will create several tasks and corresponding datasets to evaluate current popular LLMs. Moreover, we will try to propose methods to improve the LLMs' performance on these tasks.

Recommended readings

1. Yamada, Yutaro, et al. "Evaluating Spatial Understanding of Large Language Models." arXiv preprint arXiv:2310.14540 (2023).

2. Sharma, Manasi. "Exploring and Improving the Spatial Reasoning Abilities of Large Language Models." arXiv preprint arXiv:2312.01054 (2023).

3. Aghzal, Mohamed, Erion Plaku, and Ziyu Yao. "Can large language models be good path planners? a benchmark and investigation on spatial-temporal reasoning." arXiv preprint arXiv:2310.03249 (2023).

4. Li, Fangjun, David C. Hogg, and Anthony G. Cohn. "Advancing Spatial Reasoning in Large Language Models: An In-Depth Evaluation and Enhancement Using the StepGame Benchmark." arXiv preprint arXiv:2401.03991 (2024).