UNIVERSITÄT BIELEFELD
Faculty of Technology

Master Thesis

Embodied Commonsense Reasoning with Fine-Tuned Models

Semantic Computing Group
Jan-Philipp Töberg
jtoeberg@techfak.uni-bielefeld.de

Intelligent agents are challenged by unknown situations in open worlds. They cannot perform everyday tasks like cutting food or pouring drinks without encountering unknown motions, objects or environments. To mitigate this problem, providing these robots with commonsense knowledge is a possible way to increase their world understanding and support their planning capabilities [1]. However, this household-specific commonsense knowledge is, despite its potential, not often benchmarked or compared between different approaches.

In our current work, we are creating a *benchmark* focused on different, household-specific commonsense aspects like tool usage or prototypical object locations. We employed this benchmark to evaluate different LLMs with regard to their capabilities in reasoning about this *embodied commonsense*. In related work [2], the authors reported that fine-tuning the employed models on any commonsense reasoning benchmark increases the performance on downstream reasoning-adjacent tasks.

In this thesis, you will investigate this claim and try it on our created benchmark. We are interested to see whether non-embodied commonsense reasoning datasets like PIQA [3] or SocialIQA [4] can increase the performance and how the different tasks in our benchmark can influence each others performance. Important research questions are the following:

- What model can be used? How can the fine-tuning be setup?
- How do non-embodied commonsense reasoning benchmarks influence the performance?
- How does fine-tuning the model on one of the other tasks in our benchmark influence the performance?

No prior knowledge regarding Robotics or NLP is required. You can use the programming language of your choice, but Python is recommended. The thesis can be taken in English or German.

Related literature

[1] J.-P. Töberg, A.-C. N. Ngomo, M. Beetz, and P. Cimiano, ‘Commonsense knowledge in cognitive robotics: a systematic literature review’, Front. Robot. AI, vol. 11, 2024, doi: 10.3389/frobt.2024.1328934.

[2] S. Adak, D. Agrawal, A. Mukherjee, and S. Aditya, ‘Text2Afford: Probing Object Affordance Prediction abilities of Language Models solely from Text’, in Proceedings of the 28th Conference on Computational Natural Language Learning, L. Barak and M. Alikhani, Eds., Miami, FL, USA: Association for Computational Linguistics, Nov. 2024, pp. 342–364. Accessed: Nov. 11, 2024. [Online]. Available: https://aclanthology.org/2024.conll-1.27

[3] Y. Bisk, R. Zellers, R. Le bras, J. Gao, and Y. Choi, ‘PIQA: Reasoning about Physical Commonsense in Natural Language’, in Proceedings of the AAAI Conference on Artificial Intelligence, Apr. 2020, pp. 7432–7439. doi: 10.1609/aaai.v34i05.6239.

[4] M. Sap, H. Rashkin, D. Chen, R. Le Bras, and Y. Choi, ‘Social IQA: Commonsense Reasoning about Social Interactions’, in Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), Hong Kong, China: Association for Computational Linguistics, 2019, pp. 4462–4472. doi: 10.18653/v1/D19-1454.

The Semantic Computing Group researches and develops methods that enable machines to acquire relevant knowledge as well as linguistic capabilities. Using methods from *natural language understanding* and *machine learning*, we are aiming at machines that are capable of knowledge acquisition by reading unstructured textual data. In particular, the group focuses on methods for information extraction, semantic parsing, ontology learning, sentiment analysis, entity linking, as well as question answering.

More information is available at: `http://www.sc.cit-ec.uni-bielefeld.de/`

Interested? @mail to `jtoeberg@techfak.uni-bielefeld.de`