Wittgenstein embedded: large language models, their successes and limitations.

In the 1950s, just at the birth of Artificial Intelligence, Ludwig Wittgenstein and John Firth stated famous sentences that had an enormous impact on modern natural language processing (NLP) -- one of the fastest-developing branches of AI. They both agreed that the meaning of a word comes from its use in the language [1, 2] so that we can “understand” words by analyzing their contexts. This idea gave rise to distributional semantics, which is widely used in NLP to model word meanings as vectors of numbers called word embeddings, which can be easily compared and processed by AI methods.

In recent years large language models (LLMs), which aim to predict a given word based on its context have become a massive milestone in the NLP field [3]. They achieve state-of-the-art results on multiple NLP tasks such as text classification focused on assigning labels to texts (e.g., sport, politics), automatic speech recognition, text generation, and question answering, to mention a few. However, popular LLMs learn word representations using textual inputs only, which is the cause of several limitations. For example, one of the recent papers showed that because commonsense knowledge is rarely expressed explicitly in internet-based documents used to prepare LLMs, it is beneficial to inject so-called affordances directly into those models to improve the performance on tasks requiring commonsense-reasoning abilities [4]. Moreover, the lack of visual inputs is also limiting as humans learn mostly by observing and interacting with the world. After the first hyper-optimism caused by the huge successes of LLMs, scientists started to discuss again whether LLMs can really understand words [5].

[1] J. R. Firth. A synopsis of linguistic theory 1930-55. In Studies in Linguistic Analysis (special volume of the Philological Society), volume 1952-59, pages 1–32, Oxford, 1957. The Philological Society.

[2] Wittgenstein, Ludwig. Philosophical investigations. John Wiley & Sons, 2010.

[3] Devlin, Jacob, et al. "Bert: Pre-training of deep bidirectional transformers for language understanding." arXiv preprint arXiv:1810.04805 (2018).

[4] Gretkowski, Andrzej, Dawid Wiśniewski, and Agnieszka Ławrynowicz. "Should We Afford Affordances? Injecting ConceptNet Knowledge into BERT-Based Models to Improve Commonsense Reasoning Ability." International Conference on Knowledge Engineering and Knowledge Management. Springer, Cham, 2022.

[5] Bender, Emily M., and Alexander Koller. "Climbing towards NLU: On meaning, form, and understanding in the age of data." Proceedings of the 58th annual meeting of the association for computational linguistics. 2020.