Knowledge Grounding in Language Models: An Empirical Study

Martin Fixman[♠], **Tillman Weyde**[♠], **Chenxi Whitehouse**[♠], **Pranava Madhyastha**[♠]
[♠] City St. George's, University of London

Abstract

Large language models, which exploded in quality and prevalence, have the propensity to produce hallucinations which presents a critical challenge where precision and correctness are crucial. Retrieval-Augmented Generation (RAG) has been proposed as as solution to this problem, but still presents problems as it's unclear when a large language model chooses to generate answers using the context provided by RAG over the knowledge on its parametric memory. We explore the *knowledge grounding* of large language models to understand the source chosen by the model when presented with a prompt that contains a context with information that contradicts its parametric knowledge.

Our findings show that smaller models and Seq2Seq are biased towards choosing knowledge from the context over larger and Decoder-only models. Retrieved information about the *perplexity* of an answer is used to create a predictor of the source of an answer.

1 Introduction

2 Related Work

3 Experimental Setup

4 Experimental Results

5 Discussion

6 Conclusions

This is an example of an ACL2025 reference (Izacard et al., 2022).

References

Gautier Izacard, Patrick Lewis, Maria Lomeli, Lucas Hosseini, Fabio Petroni, Timo Schick, Jane Dwivedi-Yu, Armand Joulin, Sebastian Riedel, and Edouard Grave. 2022. Atlas: Few-shot Learning with Retrieval Augmented Language Models.