Knowledge Grounding in Retrieval-Augmented LMs
Quick Update

Martin Fixman et. al

City, University of London
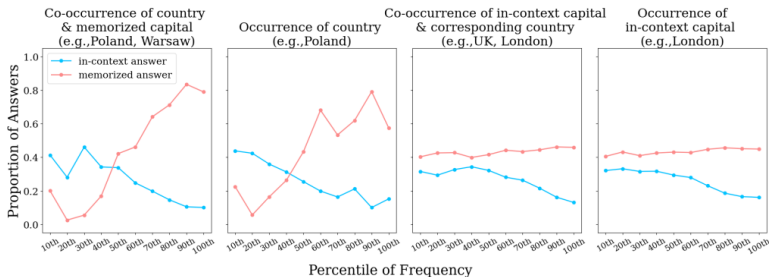
6 September 2024

Reading List

"Characterizing Mechanisms for Factual Recall in Language Models"
Yu et. al

Thanks Chenxi

Reading List

"Language Models are Few-Show Learners"
Brown et. al

"Large Language Models Struggle to Learn Long-Tail Knowledge"
Kandpal et. al

"What makes in-context learning work?"
Min et. al

New models

Currently experimenting with new models.

Flan-**T5** An encoder-decoder Seq2Seq model.

Atlas Flan T5 fine-tuned with index data.

Technical issues

Some technical problem on calculating cross-entropy.

Description of these challenges will be a good part of the written report!
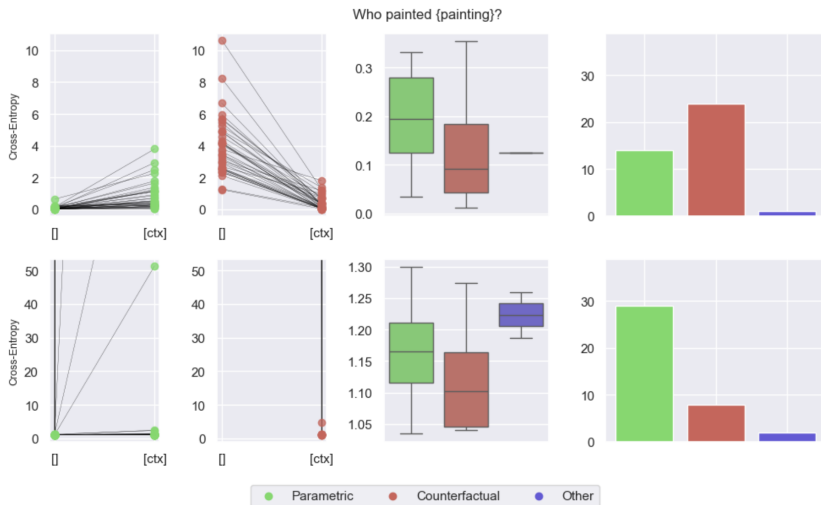
Fine-tuning out hypothesis

Can we predict whether the result from a query with added context came from parametric knowledge, or from the original context?

Does this change depending on the model? Is there a "best" model for preferring contextual knowledge?
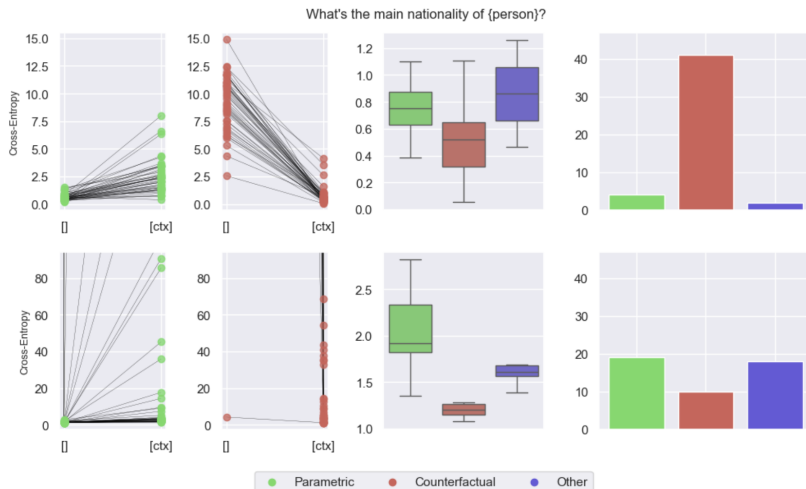
Does this depend on which question is being answered?

(Work in progress)

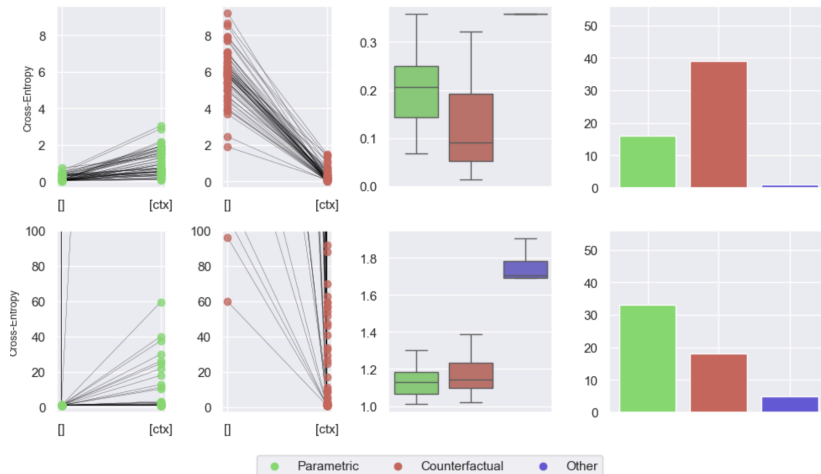# More results (now a bit more understandable)
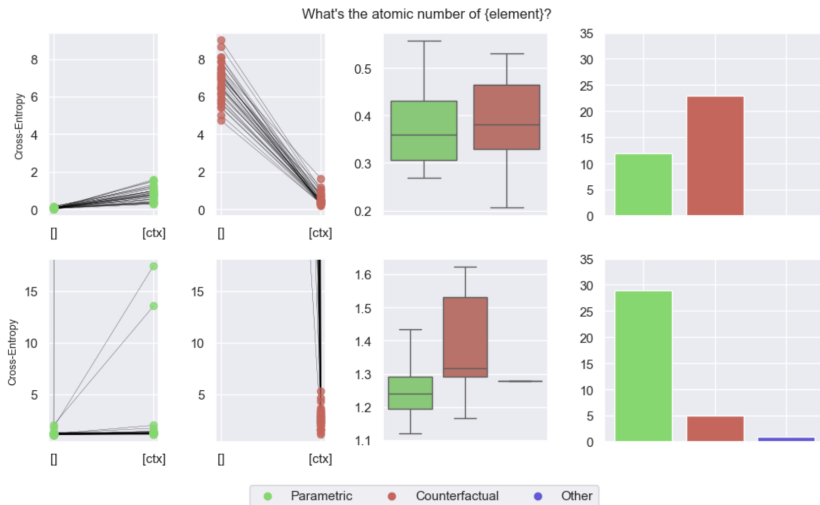
More results (now a bit more understandable)

More results (now a bit more understandable)

What year did {historical_event} happen?

More results (now a bit more understandable)

Any questions?

Any Questions?