Distinguishing Recalled versus Imagined Events in Humans

Background and Methods

The narrative flow of a story, described as "sequentiality", measures the coherence in a story. This coherence can serve as a tool to differentiate between autobiographical memories and imagined events. The foundational paper for this project introduces sequentiality to compare imagined stories and autobiographical tales. Drawing inspiration from the research paper that introduced sequentiality as a metric to measure narrative flow, this project aimed to replicate and validate the major findings depicted in Figure 2 of the paper using HippoCorpus Dataset.

To compute sequentiality, each story's sentences were evaluated for their likelihood given the prior context (history length upto 5). By comparing this contextual likelihood to a topical likelihood (provided by a story summary), the sequentiality of the story was computed. Our implementation deviated slightly from the paper by using the GPT-2 model in lieu of GPT-3, history length 5 in lieu of history length 9 due to computation limitation, anticipating potential variations in the results. Each story's sentences were evaluated for their likelihood given the prior context (sentences before them).

$$c(si,h) = -\frac{1}{|si|}\left[\log p_{LM}(si|T) - \log p_{LM}(si|T, si_{i-h:i-1})\right] \quad (1)$$

In the formula, $c(s_i,h)$ measures the sequentiality of a given segment s_i with respect to its historical context of length h. Here, $|s_i|$ is the length of the segment s_i. The term $p_{LM}(s_i|T)$ is the probability of the segment s_i as predicted by a language model given a topic T , while $p_{LM}(s_i|T, s_{i-h:i-1})$ is the probability of the segment s_i predicted by the same language model but now conditioned not just on the topic T, but also on the h preceding segments $s_{i-h:i-1}$.

Hypotheses

Building upon the preliminary observation from the raw data and considering the inherent structured nature of autobiographical memories:

1. Imagined stories will exhibit higher sequentiality scores than autobiographical stories, hinting at the fact is it is well structured narrative because it is imagined.

2. The sequentiality of autobiographical stories will increase when the memories are recounted several months later hinting the building up of narrative over time.

Analysis

.The dataset "Hippocorpus," containing 6,854 English short stories about autobiographical and imagined events, was employed. The autobiographical events are subjected to distinguishable events, retold and recalled . The inclusion of both recalled and imagined narratives in a controlled environment allows for a direct comparison between real-life events and fictional tales using sequentiality. The data for the study was obtained from a dataset comprised of 6,854 English diary-like short stories categorized into imagined, recalled and retold. Package NLTK's Punkt model is used to quantify the lexicons to understand the realis of the events.

Without delving deep into specific figures, a cursory overview of the computed sequentiality scores indicates a trend in alignment with our hypotheses. The imagined narratives generally showcased a higher sequentiality compared to the autobiographical accounts. Also, as anticipated, the sequentiality of recounted autobiographical stories surged compared to the initially narrated versions. These observations, though derived from GPT-2, appear to echo the essence of the results obtained with GPT-3 in the referenced research.

Results

The results successfully captured the trends highlighted in the research paper, albeit with some inconsistencies, which might be attributed to the model difference(in lieu, GPT2 vs GPT3-Large):

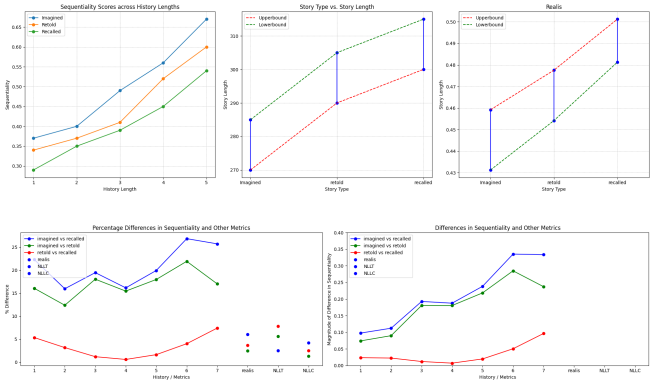

Figure 1: Seqeuntiality across story lengths, story length vs count, realis across story length & type and Effect size

The derived sequentiality showed imagined stories had a slightly higher sequentiality compared to autobiographical ones. Retold narratives also exhibited a greater sequentiality than recalled ones, aligning with the paper's findings. However, the percentage changes and its Magnitude slightly off relatively, this might be due to the LLM used, reduced history sizes and sample space about 3000 stories.

Discussion and Future Plans

In our approach, we made a minor adjustment from the original paper by opting for the GPT-2 model instead of GPT-3, so we expect some differences in the outcomes. The imagined events tend to have more sequentiality, which may be due to the fact it is structered. In Phase 2, we're focusing on: 1. Extracting additional linguistic aspects from the narratives. 2. Creating a model that can differentiate between Imagined and recalled events. We'll gauge the model's effectiveness using a previously defined validation set. By employing advanced NLP using the HippoCorpus Dataset, we can extract linguistic patterns and characteristics such as sentiment, complexity,

and nuances. Along with this data, the data collected through Phase 1 can be used to analyze and develop a computational model that can classify imagined and recalled events. This can be verified using the pre-determined validation set in the corpus dataset.

References

Sap, M., Jafarpour, A., Yejin, C., Smith, N., Pennebaker, J., & Horvitz, E. (2022, 11). Quantifying the narrative flow of imagined versus autobiographical stories. *Proceedings of the National Academy of Sciences of the United States of America*, *119*, e2211715119. doi: 10.1073/pnas.2211715119