Supplementary Materials: Exploring Lexical Relations in BERT using Semantic Priming

Kanishka Misra[1], Allyson Ettinger[2], Julia Taylor Rayz[1]

[1]Purdue University, West Lafayette, IN, USA
[2]University of Chicago, Chicago, IL, USA
{kmisra, jtaylor1}@purdue.edu, aettinger@uchicago.edu

Measures used in this study

To help guide how we measure priming in BERT, consider the following example, where the word originally in place of ___ is ***agreement***:

(1) a. ***negotiation.*** We have an ___ .
b. ***fancy.*** We have an ___ .

In this example, (1a) is referred to as the related context, (R, C) – since the prime word in this case is a lexical associate of the target (T), ***agreement***, and (1b) is the unrelated context, (U, C). We assess whether BERT is getting (lexically) primed by measuring the surprisal of the target word (***agreement***) in place of ___ in contexts (R, C) and (U, C). The surprisal measure quantifies the level of surprise in encountering particular events given context. It is defined as:

$$\textit{Surp} = -\log P(w \mid C),$$

where *Surp* is the amount of surprise in encountering word, w in context, C. If BERT tracks lexical associations between words, it should be more surprised in encountering T in context (U, C), than in context (R, C), i.e., the difference of surprisals, *Surp*(T | U, C) − *Surp*(T | R, C) should be positive. We refer to this difference of surprisals as the Facilitation (F).

Contextual Constraints

Contexts tend to vary considerably in terms of how they contribute towards anticipation of particular words. Consider the following example where ___ is the word *key*:

(2) a. He lost the ___ yesterday.
b. She unlocked the door using her ___ .

Here, (2a) is less constraining than (2b) since the set of the objects that can satisfy the semantic role of THEME of the event LOSE is considerably larger than the set of objects that can satisfy the semantic role of INSTRUMENT of the event UNLOCK-DOOR. To take this into account, we measure priming under various contextual constraints. We measure the amount of constraint on the blank position as the probability of the most expected word according to BERT. In our final dataset (refer to the video on how we compiled the dataset), we group our sampled contexts into 10 equal bins with width of 0.1 constraint score each, obtaining contexts with constraint scores ranging from 0.1 to 10. Following previous work in assessing priming by sentence contexts in humans (Schwanenflugel and LaCount, 1988), we include a neutral context: ”*the last word of this sentence is* ___ . This context intuitively poses the least amount of constraint on the blank position since any word can complete it.

Results

We plot two quantities to infer our conclusions about the priming behavior of BERT: (1) the average facilitation values, F, averaged across all stimuli for each constraint bin, and (2) the proportion of primed instances, where $F > 0$. We plot our results across different lexical relations that are annotated in the Semantic Priming Project (Hutchison et al., 2013), from which we extract our priming triples (T, R, U). These results convey the degree to which BERT tracks specific lexical relations in predicting word probabilities in context.

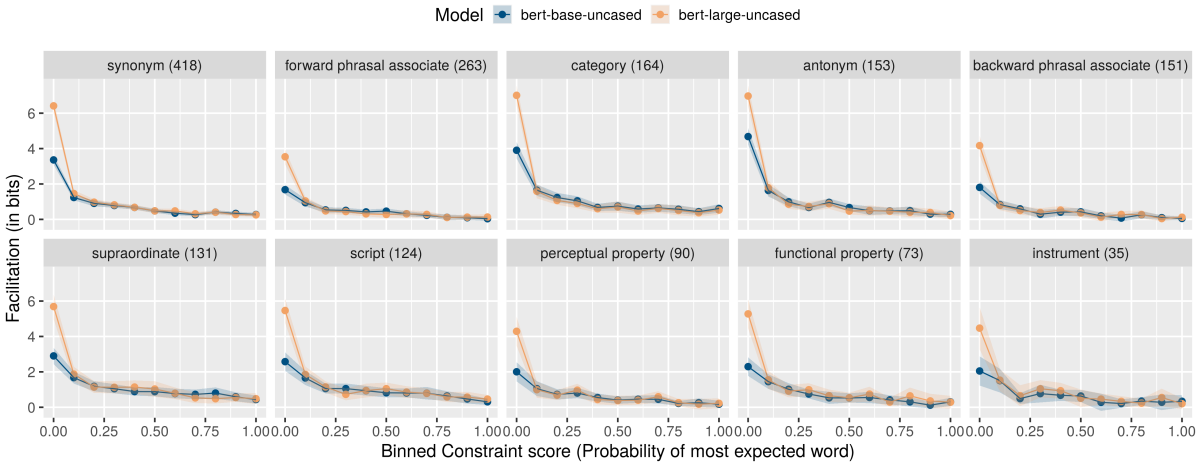

Figure 1: Average facilitation values vs. binned constraint score across annotated lexical relations in the Semantic Priming Project. Neutral contexts are shown at constraint score = 0.0.

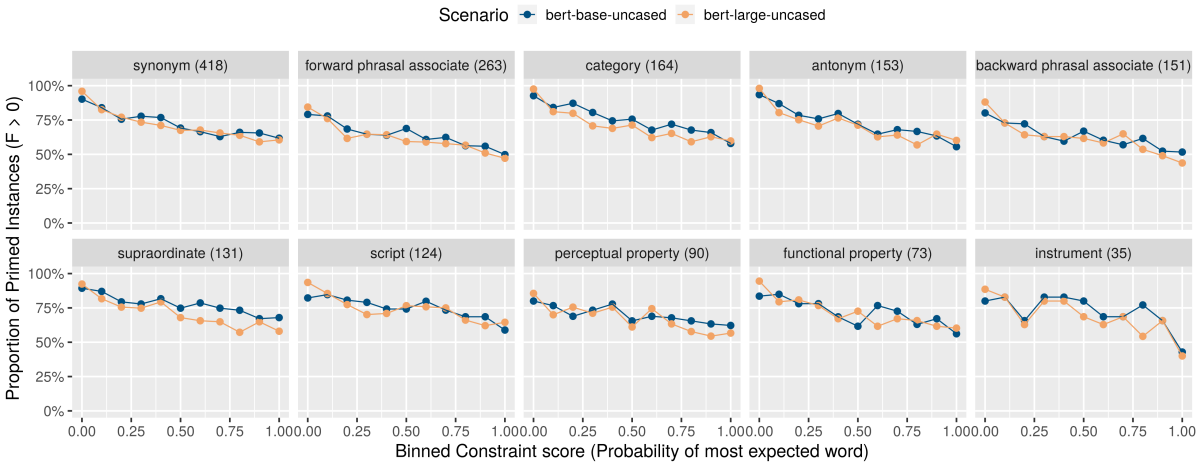

Figure 2: Proportion of primed instances ($F > 0$) vs. binned constraint score across annotated lexical relations in the Semantic Priming Project. Neutral contexts are shown at constraint score = 0.0.

References

Keith A Hutchison, David A Balota, James H Neely, Michael J Cortese, Emily R Cohen-Shikora, Chi-Shing Tse, Melvin J Yap, Jesse J Bengson, Dale Niemeyer, and Erin Buchanan. 2013. The semantic priming project. *Behavior research methods*, 45(4):1099–1114.

Paula J Schwanenflugel and Kathy L LaCount. 1988. Semantic relatedness and the scope of facilitation

for upcoming words in sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(2):344.