Human Inspired Word Segmentation for Language Models

Shinjini Ghosh (grad), Raul Alcantara (undergrad), Sophia Zhi (undergrad)

(from internal TA list under Tiwalayo Eisape)

Motivation

We are interested in projects at the intersection of cognitive science and linguistics. We believe that current state-of-the-art language models, which fail on natural language understanding and inference tasks, could benefit with human-inspired augmentations. Specifically, we want to answer the question "Can we apply intuitions about human cognition to improve word segmentation beyond the performance of current quantitative models?" in this project. We believe that an improved word segmentation algorithm would further the current NLP frontiers in the capabilities of neural and non-neural language models, especially because most models currently in place crucially reply on segmenting words correctly.

Project Outline

We wish to explore multiple avenues for word segmentation, including designing a useful metric and either collecting human judgments on various segmentations of words or looking at existing datasets, and eventually trying to build an algorithm inspired by human cognitive processes. We also want to look at, reimplement and possibly tweak current state-of-the-art algorithms and investigate into how ours differs from them. This also necessitates the development of a framework for comparing word segmentation algorithms: how do we know if a segmentation is "correct" or more "accurate" or considered more "natural" by humans? Eventually, we aim to have a comparison across existing and new word segmentation algorithms, which can help highlight the contributions induced by human judgements.

References

1. Wu, Y., Schuster, M., Chen, Z., Le, Q. V., Norouzi, M., Macherey, W., ... & Klingner, J. (2016). Google's neural machine translation system: Bridging the gap between human and machine translation. arXiv preprint arXiv:1609.08144.

2. Sennrich, R., Haddow, B., & Birch, A. (2015). Neural machine translation of rare words with subword units. arXiv preprint arXiv:1508.07909.

3. Dyer, C., Kuncoro, A., Ballesteros, M., & Smith, N. A. (2016). Recurrent neural network grammars. arXiv preprint arXiv:1602.07776.

4. Goldwater, S., Griffiths, T. L., & Johnson, M. (2009). A Bayesian framework for word segmentation: Exploring the effects of context. Cognition, 112(1), 21-54.