Unsupervised Hierarchical Story Infilling

Citation

Daphne Ippolito, David Grangier, Chris Callison-Burch, and Douglas Eck. 2019. Unsupervised Hierarchical Story Infilling. In *Proceedings of the First Workshop on Narrative Understanding*, pages 37–43, Minneapolis, Minnesota. Association for Computational Linguistics.

Big Picture

Natural Language Processing ⊃ Automatic Story Generation ⊃ Story Infilling

Story infilling is the process of "filling in" a missing sequence of text in an intermediate location within a story. It is a subcategory of *text infilling* which considers any general text and is a type of cloze test task. Challenges include: text fluency, story line coherence, and non-generic or obvious text generation. Notice the first two challenges and last challenge are contradictions in extreme cases. The authors state their motivation: "to generate diverse text without compromising on coherence and fluency."

Their proposed process could be useful during revision and editing phases of writing.

Summary

The authors propose a hierarchical method for story-infilling: first a set of "rare" words is selected, and then the rest of the text is generated by conditioning on that set of words.

Baseline non-hierarchical approach with different decoding strategies:

1. Beam search
2. Sampling
3. Top-10 next words sampling

Proposed hierarchical approach:

Model 1: select set of non-generic words via random sampling
Model 2: generate sequences conditioned on Model 1 output using beam search

Their experiments use the self-attentive Transformer model with the Tensor2Tensor framework, trained on the union of books contained in both the Toronto Book Corpus and Project Gutenberg Corpus. For each training example, the target sequence (infilled text) was 5-50 tokens long, and the left and right contexts were 50 tokens each. The divisions between each of these segments were not constrained by linguistic boundaries (i.e. end-of-sentence punctuation). During evaluation, 10,000 spans were evaluated automatically, and paragraphs from three children's books were evaluated by human raters. The results conclude the baseline non-hierarchical approach with beam search generates non-diverse text, while sampling (top-10 next words) generates diverse but potential non-fluent and non-coherent text. The proposed hierarchical approach achieves both diverse and fluent generations, thereby achieving the authors' motivational objective.

© Anneliese Brei January 9, 2023