Reading Assignment 2
Hamza Rashid, 260971031
COMP 550, Fall 2024

Paper Summary

The paper, *"A Structural Probe for Finding Syntax in Word Representations"*, by John Hewitt and Christopher D. Manning, proposed a structural probe that tests a (neural-network based) language model's ability to embed syntactic knowledge in its word representation space. Concretely, the authors test for syntactic knowledge in the form of parse trees, and the probe is structural in the sense that it approximates structure-preserving (linear) maps between a language model's contextualized word representation space and the spaces given by two metrics on the parse trees: one that encodes the distance between words in the tree, and one that encodes depth. The (english) language models in question are ELMo (embeddings from language model, a Bi-LSTM) and BERT (bidirectional encoder representations from transformers), the latter utilizing an architecture not covered in class. The authors use the Wall Street Journal section of the Penn Treebank corpus to test each model's ability to capture the Stanford Dependencies formalism. They utilize supervised learning – gradient descent with the PTB WSJ dataset – to train the probes, namely the matrices that approximate the linear transformations. To aid analysis, the language models were compared against baselines that were expected to encode features useful for training a parser, but not be capable of parsing themselves. While the proposed method effectively determines that ELMo and BERT capture syntactic knowledge, it is limited by its strict formulation and reliance on supervised learning.

The proposed metrics consist of a squared L2 distance that encodes the distance between words in the parse tree, and one in which squared L2 norm encodes depth in the parse tree, where distance and depth are measured in edges. Distances between pairs of words are important for capturing hierarchical behavior, such as subject-verb agreement, and for recovering the predicted tree (via the Minimum Spanning Tree algorithm) after applying the linear transformation on the model's output vector sequence. On the other hand, the depth norm captures edge directions and imposes a total order on the words in a sentence. Naturally, the authors formulated the probes (for each metric) as matrices. If such transformations exist, they define inner products on the original space that capture the desired syntactic knowledge. The authors utilized mini-batch gradient descent to optmize the probes' parameters, namely the entries of the matrices, with the loss for sample normalized by its squared length.

Strengths and Limitations

The approach is innovative in its simplicity and direct measurement of syntax structure, supporting evidence of hierarchical information in pretrained embeddings. The probe's reliance on linear transformations makes it computationally efficient and interpretable. However, limitations include reliance on supervised data, which may bias results, and the probe's inability to capture syntactic nuances that require more complex transformations. Additionally, understanding the probe's practical applications beyond syntactic verification could be expanded upon.

Points of Uncertainty

Some concepts in the probe's design and its assumptions about the geometry of vector space remain unclear. For example, how the squared L2 distances perfectly encode parse tree structures or why a linear transformation is assumed optimal for all syntactic relations. Further clarification on alternative geometries for embedding syntactic knowledge might be beneficial.

Probing Approach: Supervised vs. Unsupervised

The paper employs a supervised probing approach, where parse distances are learned using labeled syntactic data. While supervision enables direct and interpretable syntactic measurements, it limits generalizability across languages and treebank styles. An unsupervised approach could mitigate reliance on labeled data, potentially revealing latent syntactic structures across models, but might struggle to match supervised accuracy.

Importance of Syntactic Probing

Probing models for syntactic knowledge is essential both practically, as it guides model improvements and applications in syntax-sensitive tasks, and theoretically, as it reveals linguistic patterns in language embeddings. Establishing syntax presence in embeddings provides insights into the representation power of pretrained models, crucial for understanding their success in tasks like machine translation and parsing.

References