Experimental Support for a Categorical Compositional Distributional Model of Meaning (Grefenstette and Sadrzadeh, [2011])

Overview.

The authors of this paper describe an implementation of Coecke et al. [1]. They use TF-IDF vectors over a subset of basis words to represent atomic words, and they describe a simple algorithm to learn the matrices of words with compound types (verbs) based on the co-occurrences of these words with atomic words. For example, for some transitive verb P, they compute the i, j element of P's matrix based on the number of times that both the i^{th} basis word has co-occurred with a subject of P and the j^{th} basis word has co-occurred with an object of P.

The authors also show how we can efficiently compute sentence meanings by representing the actions of adjectives, adverbs and verbs on atomic words and subsentences with pointwise multiplication. They then evaluate their trained models with both the word disambiguation task from (Mitchell and Lapata [3]) and a custom task that they design.

Comments.

One thing to note is that the authors' algorithm for learning compound type word matrices depends only on the weights of the atomic word projection vectors. The basis vectors themselves are not involved in the computation. This suggests that it would be quite simple to adapt this strategy to work with word vectors that we learn with low-dimensional word embedding techniques like word2vec. To do this, we would modify the algorithm to build the compound type word matrices from the elements of the learned embeddings of the atomic words. That is, for some transitive verb P, we would compute the i, j element of P's matrix based on the average magnitude of P's subject words' i^{th} vector elements and object words' j^{th} vector elements.

REFERENCES

[1] Coecke, B., Sadrzadeh, M., and Clark, S. (2010). Mathematical foundations for a compositional distributional model of meaning. *arXiv preprint arXiv:1003.4394*.

[2] Grefenstette, E. and Sadrzadeh, M. (2011). Experimental support for a categorical compositional distributional model of meaning. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 1394–1404. Association for Computational Linguistics.

[3] Mitchell, J. and Lapata, M. (2008). Vector-based models of semantic composition. *proceedings of ACL-08: HLT*, pages 236–244.