tasks. Experimental results on the Stanford Sentiment Treebank show that AdaMC significantly improves the baselines with fewer parameters. We further compare the distribution similarities of composition functions for phrase pairs, and the results verify the effectiveness of AdaMC on modeling and leveraging the semantic categories of words and phrases in the process of composition. There are several interesting directions for further research studies. For instance, we can evaluate our method in other NLP tasks. Moreover, external information (such as part-of-speech tags) can be used as features to select the composition functions. In addition, we can mix different types of composition functions (such as the linear combination approach in RNN and the tensor based approach in RNTN) to achieve more flexible choices in the adaptive composition methods.

Acknowledgments

We gratefully acknowledge helpful discussions with Richard Socher. This research was partly supported by the National 863 Program of China (No. 2012AA011005), the fund of SKLSDE (Grant No. SKLSDE-2013ZX-06), and Research Fund for the Doctoral Program of Higher Education of China (Grant No. 20111102110019).

References

Aerts, D., and Czachor, M. 2004. Quantum aspects of semantic analysis and symbolic artificial intelligence. *Journal of Physics A: Mathematical and General* 37(12):L123.

Baroni, M., and Zamparelli, R. 2010. Nouns are vectors, adjectives are matrices: Representing adjective-noun constructions in semantic space. In *EMNLP*, EMNLP '10, 1183–1193.

Clark, S.; Coecke, B.; and Sadrzadeh, M. 2008. A compositional distributional model of meaning. In *Proceedings of the Second Quantum Interaction Symposium*, 133–140.

Duchi, J.; Hazan, E.; and Singer, Y. 2011. Adaptive subgradient methods for online learning and stochastic optimization. *JMLR* 12:2121–2159.

Georgii, H. 2011. *Gibbs Measures and Phase Transitions*. De Gruyter studies in mathematics. De Gruyter.

Grefenstette, E., and Sadrzadeh, M. 2011. Experimental support for a categorical compositional distributional model of meaning. In *EMNLP*, 1394–1404.

Grefenstette, E.; Dinu, G.; Zhang, Y.-Z.; Sadrzadeh, M.; and Baroni, M. 2013. Multi-step regression learning for compositional distributional semantics. In *Proceedings of the 10th International Conference on Computational Semantics*.

Hermann, K. M., and Blunsom, P. 2013. The role of syntax in vector space models of compositional semantics. In *ACL*, 894–904.

Jones, M. N., and Mewhort, D. J. 2007. Representing word meaning and order information in a composite holographic lexicon. *Psychological review* 114(1):1.

Klein, D., and Manning, C. D. 2003. Accurate unlexicalized parsing. In *ACL*, 423–430.

Landauer, T. K., and Dutnais, S. T. 1997. A solution to plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological review* 211–240.

Mitchell, J., and Lapata, M. 2008. Vector-based models of semantic composition. In *ACL*, 236–244.

Mitchell, J., and Lapata, M. 2010. Composition in distributional models of semantics. *Cognitive Science* 34(8):1388–1439.

Pang, B., and Lee, L. 2005. Seeing stars: exploiting class relationships for sentiment categorization with respect to rating scales. In *ACL*, 115–124.

Plate, T. 1991. Holographic reduced representations: Convolution algebra for compositional distributed representations. In *IJCAI*, 30–35. Citeseer.

Polajnar, T.; Fagarasan, L.; and Clark, S. 2013. Learning type-driven tensor-based meaning representations. *CoRR* abs/1312.5985.

Rudolph, S., and Giesbrecht, E. 2010. Compositional matrix-space models of language. In *ACL*, 907–916.

Rumelhart, D.; Hinton, G.; and Williams, R. 1986. Learning representations by back-propagating errors. *Nature* 323(6088):533–536.

Smolensky, P. 1990. Tensor product variable binding and the representation of symbolic structures in connectionist systems. *Artificial Intelligence* 46(12):159 – 216.

Socher, R.; Lin, C. C.; Ng, A. Y.; and Manning, C. D. 2011. Parsing Natural Scenes and Natural Language with Recursive Neural Networks. In *ICML*.

Socher, R.; Huval, B.; Manning, C. D.; and Ng, A. Y. 2012. Semantic compositionality through recursive matrix-vector spaces. In *EMNLP-CoNLL*, 1201–1211.

Socher, R.; Bauer, J.; Manning, C. D.; and Ng, A. Y. 2013a. Parsing With Compositional Vector Grammars. In *ACL*.

Socher, R.; Perelygin, A.; Wu, J. Y.; Chuang, J.; Manning, C. D.; Ng, A. Y.; and Potts, C. 2013b. Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank. In *EMNLP*, 1631–1642.

Wang, S., and Manning, C. 2012. Baselines and bigrams: Simple, good sentiment and topic classification. In *ACL*, 90–94.

Widdows, D. 2008. Semantic vector products: Some initial investigations. In *Second AAAI Symposium on Quantum Interaction*, volume 26, 28th. Citeseer.

Yessenalina, A., and Cardie, C. 2011. Compositional matrix-space models for sentiment analysis. In *EMNLP*, 172–182.

Yu, D.; Deng, L.; and Seide, F. 2013. The deep tensor neural network with applications to large vocabulary speech recognition. *Audio, Speech, and Language Processing, IEEE Transactions on* 21(2):388–396.