Natural Language Engineering **21** (5): 661–664. © Cambridge University Press 2015
doi:10.1017/S1351324915000078

Word from the editors

ZORNITSA KOZAREVA[1], VIVI NASTASE[2] and RADA MIHALCEA[3]

[1]*Yahoo! Labs, Sunnyvale, USA*
[2]*Human Language Technologies, Natural Language processing, Fondazione Bruno Kessler, Trento, Italy*
[3]*Computer Science and Engineering, University of Michigan, Ann Arbor, MI, USA*

(*Received 18 February 2015; revised 7 April 2015; first published online 20 April 2015*)

1 Graphs and natural language processing

Graph structures naturally model connections. In natural language processing (NLP) connections are ubiquitous, on anything between small and web scale. We find them between words – as grammatical, collocation or semantic relations – contributing to the overall meaning, and maintaining the cohesive structure of the text and the discourse unity. We find them between concepts in ontologies or other knowledge repositories – since the early ages of artificial intelligence, associative or semantic networks have been proposed and used as knowledge stores, because they naturally capture the language units and relations between them, and allow for a variety of inference and reasoning processes, simulating some of the functionalities of the human mind. We find them between complete texts or web pages, and between entities in a social network, where they model relations at the web scale. Beyond the more often encountered ‘regular’ graphs, hypergraphs have also appeared in our field to model relations between more than two units.

Graphs have been rigorously studied, both mathematically and computationally, providing a well-developed theoretical and practical base to the many natural language processing problems that map into this framework.

In syntax, part-of-speech tagging was tackled using graph clustering (Biemann 2006) and dependency parsing using minimum spanning trees (McDonald *et al.* 2005). Related to parsing is the task of prepositional phrase attachment, which found interesting solutions in semi-supervised based learning (Toutanova, Manning and Ng 2004). Min-cut algorithms have also been used in text processing applications, for instance for the problem of coreference (Nicolae and Nicolae 2006).

In semantics, graphs have been used to construct semantic classes (Widdows and Dorow 2002) through networks of words built from very large corpora. On similar word networks, work has also been done on understanding lexical network properties (Ferrer i Cancho and Solé 2001), or extracting words that follow certain semantic relations such as synonymy (Weale, Brew and Fosler-Lussier 2009). A significant amount of effort has also been put into the measurement of semantic distance using path-based algorithms on semantic networks (Lin 1998) or random-walks (Ramage, Rafferty and Manning 2009). These random-walk algorithms have been successfully applied to other problems in semantics, such as word sense disambiguation (Sinha and Mihalcea 2007; Agirre and Soroa 2009), name disambiguation (Minkov, Cohen and Ng 2006) among others. The problems of sentiment and subjectivity analysis were also tackled using graph-based methods, such as clustering over graphs to identify the polarity of adjectives (Hatzivassiloglou and McKeown 1997), or min-cut algorithms for sentence-level subjectivity analysis (Pang and Lee 2004).

There are also a number of natural language processing applications that found successful solutions in the use of graph-based methods. These include text summarization (Erkan and Radev 2004; Mihalcea and Tarau 2004), semi-supervised passage retrieval (Otterbacher, Erkan and Radev 2005), keyword extraction (Mihalcea and Tarau 2004), text mining (Kozareva and Hovy 2011), deriving semantic classes (Kozareva, Riloff and Hovy 2008), topic identification (Syed, Finin and Joshi 2008; Coursey, Mihalcea and Moen 2009), topic segmentation (Malioutov and Barzilay 2006), machine translation (Zens and Ney 2005), cross-language information retrieval (Monz and Dorr 2005), and question answering (Molla 2006).

2 Overview of the issue

The four papers in the current special issue each showcase and exploit a different aspect/facet of graphs for different tasks in natural language processing.

Kotlerman *et al.* present a novel twist on the graphs as a framework for the organization of knowledge. *Textual Entailment Graphs* expand the notion of the ontology as a network of concepts to larger text units that convey complex information. The *entailment* relation replaces the *is-a* relation from ontologies, and text fragments replace concepts, to obtain a structure that organizes text units by subsumption of the conveyed information.

Jadidinejad *et al.* exploit the graph structure of knowledge repositories for the computation of semantic relatedness between texts. Previously, ontologies were used to provide bag-of-concept representations for given texts, and these unstructured collections were used for similarity/relatedness computations. The approach presented in this paper shows that the structure itself is useful: using the graph structure of the ontology through a clique-based semantic kernel can lead to improvements in semantic relatedness estimations.

Fernández *et al.* and Mitra *et al.* reveal and exploit sub-structures – communities of nodes representing words with related meanings – in word co-occurrence graphs. Fernández *et al.* apply frequency-based filtering and ranking and clustering algorithms to form and expose communities in word co-occurrence graphs. These are taken to approximate word senses, that together with bilingual dictionaries help perform sense-level translations.

Mitra *et al.* start with word-specific co-occurrence graphs, which are clustered using the Chinese Whispers algorithm to form sense-specific clusters. This method is applied on text collections representing disjoint time frames. Changes in the clusters obtained from data from different time spans are analysed and interpreted as effects of diachronic sense changes.

Acknowledgements

We would like to thank all the authors who submitted their papers to our special issue. We are also grateful to the reviewers who spent time over several rounds of revisions to provide feedback and suggestions to the authors. The editorial staff of the Journal of Natural Language Engineering has been extremely helpful along the way, to assist us with information regarding entire process of putting together a special issue, and to remind us of deadlines. Rada Mihalcea's contribution has been partially supported by National Science Foundation CAREER award #1361274.

References

Agirre, E., and Soroa, A. 2009. Personalizing pagerank for word sense disambiguation. In *Proceedings of the Conference of the European Chapter of the Association for Computational Linguistics*, Athens, Greece, pp. 33–41.

Biemann, C. 2006. Unsupervised part-of-speech tagging employing efficient graph clustering. In *Proceedings of the COLING/ACL 2006 Student Research Workshop*, Sydney, Australia, Association for Computational Linguistics, pp. 7–12.

Coursey, K., Mihalcea, R., and Moen, W. 2009. Using encyclopedic knowledge for automatic topic identification. In *Proceedings of the Conference on Natural Language Learning*, Boulder, CO.

Erkan, G., and Radev, D. 2004. Lexrank: graph-based centrality as salience in text summarization. *Journal of Artificial Intelligence Research* **22**: 457–479.

Ferrer i Cancho, R., and Solé, R. V. 2001. The small world of human language. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **268**(1482): 2261–2265.

Hatzivassiloglou, V., and McKeown, K. 1997. Predicting the semantic orientation of adjectives. In *Proceedings of the Conference of the European Chapter of the Association for Computational Linguistics*, Madrid, Spain, pp. 174–181.

Kozareva, Z., and Hovy, E. 2011. Insights from network structure for text mining. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, Portland, Oregon, pp. 1616–1625.

Kozareva, Z., Riloff, E., and Hovy, E. 2008. Semantic class learning from the web with hyponym pattern linkage graphs. In *Proceedings of the Association for Computational Linguistics*, Columbus, Ohio, pp. 1048–1056.

Lin, D. 1998. An information-theoretic definition of similarity. In *Proceedings of the 15th International Conference on Machine Learning*, Madison, Wisconsin.

Malioutov, I., and Barzilay, R. 2006. Minimum cut model for spoken lecture segmentation. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics (COLING-ACL 2006)*, Sydney, Australia, pp. 9–16.

McDonald, R., Pereira, F., Ribarov, K., and Hajic, J. 2005. Non-projective dependency parsing using spanning tree algorithms. In *Proceedings of the Empirical Methods in Natural Language Processing*, Vancouver, Canada.

Mihalcea, R., and Tarau, P. 2004. TextRank – bringing order into texts. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2004)*, Barcelona, Spain.

Minkov, E., Cohen, W., and Ng, A. 2006. Contextual search and name disambiguation in email using graphs. In *Proceedings of the International ACM SIGIR conference on Research and Development in Information Retrieval*, Seattle, Washington, pp. 27–34.

Molla, D. 2006. Learning of graph-based question answering rules. In *Proceedings of TextGraphs: the Second Workshop on Graph Based Methods for Natural Language Processing*, New York City, Association for Computational Linguistics, pp. 37–44.

Monz, C., and Dorr, B. J. 2005. Iterative translation disambiguation for cross-language information retrieval. In *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, Salvador, Brazil.

Nicolae, C., and Nicolae, G. 2006. Bestcut: a graph algorithm for coreference resolution. In *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing*, Sydney, Australia.

Otterbacher, J., Erkan, G., and Radev, D. 2005. Using random walks for question-focused sentence retrieval. In *Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing*, Vancouver, British Columbia, Canada, Association for Computational Linguistics, pp. 915–922.

Pang, B., and Lee, L. 2004. A sentimental education: sentiment analysis using subjectivity summarization based on minimum cuts. In *Proceedings of the 42nd Meeting of the Association for Computational Linguistics*, Barcelona, Spain, July.

Ramage, D., Rafferty, A., and Manning, C. 2009. Random walks for text semantic similarity. In *ACL-IJCNLP TextGraphs-4 Workshop*, Singapore.

Sinha, R., and Mihalcea, R. 2007. Unsupervised graph-basedword sense disambiguation using measures of word semantic similarity. In *Proceedings of the International Conference on Semantic Computing*, Irvine, California, pp. 363–369.

Syed, Z., Finin, T., and Joshi, A. 2008. Wikipedia as an ontology for describing documents. In *Proceedings of the Second International Conference on Weblogs and Social Media*, Seattle, Washington.

Toutanova, K., Manning, C. D., and Ng, A. Y. 2004. Learning random walk models for inducing word dependency distributions. In *ICML '04: Proceedings of the 21st International Conference on Machine Learning*, New York, NY, USA, p. 103.

Weale, T., Brew, C., and Fosler-Lussier, E. 2009. Using the wiktionary graph structure for synonym detection. In *Proceedings of the Workshop on The People's Web Meets NLP: Collaboratively Constructed Semantic Resources*, Suntec, Singapore, pp. 28–31.

Widdows, D., and Dorow, B. 2002. A graph model for unsupervised lexical acquisition. In *Proceedings of the 19th International Conference on Computational Linguistics*, Taipei.

Zens, R., and Ney, H. 2005. Word graphs for statistical machine translation. In *Proceedings of the ACL Workshop on Building and Using Parallel Texts: Data-Driven Machine Translation and Beyond*, Ann Arbor, Michigan, pp. 191–198.