Towards Antonymy-Aware Natural Language Applications

Saif Mohammad (University of Maryland, saif@umiacs.umd.edu)
Bonnie Dorr (University of Maryland and HLT Center of Excellence, bonnie@umiacs.umd.edu)
Graeme Hirst (University of Toronto, gh@cs.toronto.edu)

Manually-created lexicons of antonym pairs have limited coverage and do not include most semantically contrasting word pairs. Although some approaches have been proposed before, it is only now that automatic measures of word-pair antonymy are approaching a maturity in precision and coverage (Turney 2008, Mohammad et al. 2008). However, the different computational approaches capture different facets of this intuitive, but complex, phenomenon. We describe several issues that must be negotiated before arriving at a successful antonymy-aware natural language system (outlined briefly in the next two paragraphs). Additionally, we detail various tasks that can benefit by knowing which terms are antonymous (outlined briefly in the last paragraph).

Over the years, many definitions of antonymy have been proposed. From an applications perspective, one has to determine whether a narrow or a broader definition of antonymy is more appropriate—will it be useful to know that a verb (say, *disband*) conveys a contrasting meaning to a noun (say, *conglomeration*) or should antonymy apply only to gradable adjectives (*hot–cold*)? For many applications it is useful to define antonymy to encompass the complete semantic range—a combined measure of the contrast in meaning conveyed by two words and the tendency of native speakers to call them opposites. So a useful measure of antonymy is one that is accurate for this wide range of input pairs.

Strictly speaking, antonymy (like any other lexical-semantic relation) applies to a pair of lexical units—combinations of surface form and word sense. This implies that two words may be antonymous in certain contexts, but not antonymous in others. For example, *freezing* in sentence (1), is not antonymous to *hot* in sentence (2):

(1) *Desert nights can be freezing.* (2) *The cops were in hot pursuit.*

However, given that two words are found in close proximity in a cohesive piece of text, one can assume that the intended senses of the two words are those that are close in meaning (have some lexical semantic relation). This follows from the assumption that in cohesive text, many words are used to describe related concepts. Further, antonymous words tend to be used in close proximity to express contrast. Thus, a natural language system may assume that in the sentence below *hot* and *freezing* express contrasting meaning, without requiring costly manual sense-annotation.

(3) *Desert weather is extreme; hot days are followed by freezing nights.*

A more sophisticated measure will use context to determine the degree of antonymy.

Such measures are useful for a number of applications, including identifying and generating paraphrases (*The dementors* ***caught*** *Sirius Black / Black could* ***not escape*** *the dementors*) and detecting contradictions (*Kyoto has a predominantly* ***wet*** *climate / It is mostly* ***dry*** *in Kyoto*). Of course, such "contradictions" may be a result of differing sentiment, new information, non-coreferent men-

tions, or genuinely contradictory statements. Antonyms often indicate the discourse relation of contrast. They are also useful for detecting humor, as satire and jokes tend to have contradictions and oxymorons. Antonyms can play a crucial role in multidocument summarization, especially that of opinions. Finally, it is useful to know which words are semantically contrasting to a target word, even if to filter them out, for example, in the automatic creation of a thesaurus.

Acknowledgments

This work is supported, in part, by the National Science Foundation under Grant No. IIS-0705832, in part, by the Human Language Technology Center of Excellence, and in part, by the Natural Sciences and Engineering Research Council of Canada. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the sponsor.

References

Saif Mohammad, Bonnie Dorr, and Graeme Hirst. October 2008. Computing Word-Pair Antonymy. *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP-2008)*, Waikiki, Hawaii.

Peter Turney. August 2008. A uniform approach to analogies, synonyms, antonyms, and associations. *Proceedings of the International Conference on Computational Linguistics (COLING-08)*, pages 905–912, Manchester, UK.