

Towards Antonymy-Aware Natural Language Applications

Saif Mohammad¹, Bonnie Dorr^{1,2}, and Graeme Hirst³
 University of Maryland¹, HLT Center of Excellence², University of Toronto³

Antonyms

Kinds

- Clear opposites:
 - create-destroy*
 - hard-soft*
 - promoted-demoted*
 - wet-dry*
- Contrasting word pairs:
 - fired-employed*
 - promoted-censured*
 - hard-fluid*
 - large-small-scale*
 - flinch-advance*
 - cogent-unconvincing*

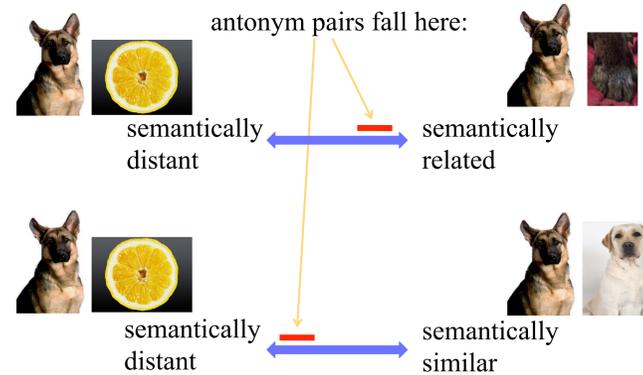


largely unrecorded

Domain

- In the strictest sense:
 - pairs of gradable adjectives
 - *hot-cold, tall-short*
- In a broader sense:
 - pairs of nouns, pairs of verbs, pairs of adjectives
 - *life-death, ascend-descend*
- In the broadest sense:
 - any two words that express a contrast in meaning
 - *city-farm, lifeless-life*

Relation with semantic distance



Antonym pairs simultaneously convey a sense of both distance and closeness:

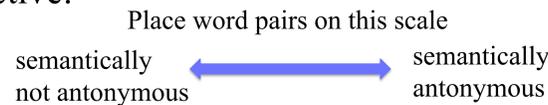
- semantically related;
- but not semantically similar.

Why be Antonymy-Aware

- Detecting incompatibles:
 - contradictions
 - Mad-Eye Moody finds the dementors charming.*
 - Mad-Eye Moody detests the dementors.*
 - differing sentiment/opinion
 - Cornelius Fudge is an incompetent minister of magic.*
 - Fudge is one of the finest ministers of magic ever.*
 - non-coreferent entities
 - Viktor is short and shy.*
 - Viktor is an imposing quidditch player from Romania.*
- Detecting paraphrases
 - Sirius Black could not evade the dementors.*
 - The dementors caught Sirius Black.*
- Detecting humor
 - I don't suffer from insanity; I enjoy every minute of it.*
 - Procrastinate now!*
- Separating antonymous words from those that are semantically similar, as in a distributional thesaurus (Lin, 1998).

Computing Word-Pair Antonymy

Objective:



Hypotheses:

- **Co-occurrence hypothesis of Antonyms**
Antonym pairs co-occur more often than random.
- **Distributional Hypothesis of Antonyms**
Antonym pairs occur in similar contexts.

Central idea (Mohammad et al., 2008):

1. Identify whether two words have a contrast relation.
 - a. generate seed antonym pairs:
 - (i) using antonym generating affix rules
 - (ii) from WordNet
 mark corresponding thesaurus categories as contrasting.
 - b. consider adjacent thesaurus categories to be contrasting.
2. Determine degree of antonymy.
 - a. The degree of antonymy between two contrasting categories is proportional to their semantic closeness: **distributional hypothesis for antonyms.**
 - b. The degree of antonymy between two words across a contrasting category pair is proportional to their tendency to co-occur: **co-occurrence hypothesis for antonyms.**

Example:

All word pairs across categories **HIDING** and **REVEALING** are marked to have a contrast relation because of seed antonym pair *cover* and *uncover*.

cover and *uncover*: strong tendency to co-occur suggests high degree of antonymy.

unnoticed and *uncover*: moderate tendency to co-occur suggests medium degree of antonymy.

curtain and *spill*: weak tendency to co-occur suggests low degree of antonymy.

Manually create list of affixes that tend to generate antonyms:

<i>x-abx</i>	<i>normal-abnormal</i>
<i>x-disx</i>	<i>trust-distrust</i>
<i>x-unx</i>	<i>classified-unclassified</i>
<i>x-inx</i>	<i>consistent-inconsistent</i>
<i>x-antix</i>	<i>clockwise-anticlockwise</i>
<i>x-nonx</i>	<i>aligned-nonaligned</i>
<i>x-imx</i>	<i>mobile-immobile</i>
<i>x-malx</i>	<i>practice-malpractice</i>
<i>x-misx</i>	<i>fortune-misfortune</i>
<i>lx-illx</i>	<i>legitimate-illegitimate</i>
<i>rx-irx</i>	<i>regular-irregular</i>
<i>imx-exx</i>	<i>implicit-explicit</i>
<i>inx-exx</i>	<i>introvert-extrovert</i>
<i>upx-downx</i>	<i>upstream-downstream</i>
<i>overx-underx</i>	<i>overdone-underdone</i>
<i>xless-xful</i>	<i>harmless-harmful</i>

False positives such as *part-depart* and *tone-intone* did not effect results much.

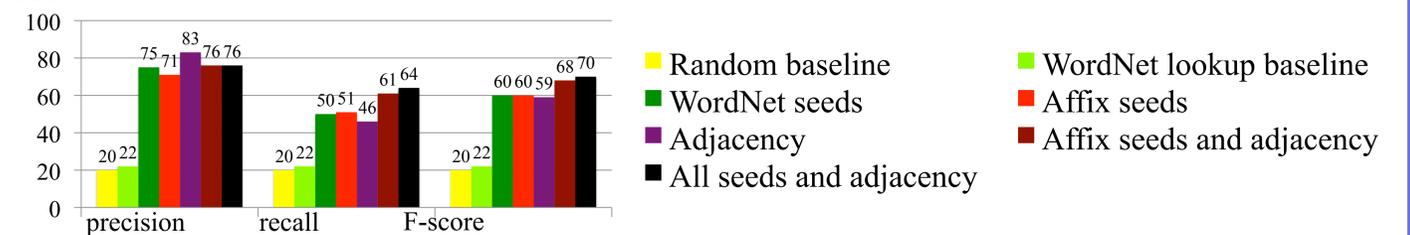


Evaluation: Solve 950 GRE closest-opposite questions

Examples

- Astute*
- shrewd* ← near-synonym
 - foolish* ← answer
 - callow* ← another close opposite
 - winning*
 - debating*
- Obdurate* (hardened in feelings)
- meager*
 - unsusceptible*
 - right*
 - tender* ← answer
 - intelligent*
- Obdurate* (resistant to persuasion)
- yielding* ← answer
 - motivated*
 - moribund*
 - azure*
 - hard*

Results



Conclusions

- Proposed a computational measure of antonymy. Geared towards natural language applications. Captures semantic contrast.
- Used the structure of a thesaurus and distributional hypothesis. Small set of affix rules found to be potent. WordNet helped, but can be done without.

Future Work

- Compute word-pair antonymy in a resource-poor language by combining its text and an English thesaurus.
- Using affix-rule information from different languages to improve performance in a target language.
- Creating a wide coverage polarity lexicon.
- Using word-pair antonymy for text summarization.

References

Dekang Lin. 1998. Automatic retrieval and clustering of similar words. In *Proceedings of the 17th International Conference on Computational Linguistics*, pages 768–773, Montreal, Canada.

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

Acknowledgments

This work was 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. We thank Smaranda Muresan and Siddharth Patwardhan for their feedback.