Looking for Hyponyms in Vector Space

Marek Rei
SwiftKey

Ted Briscoe
University of Cambridge

Hyponym Generation

The system needs to find all possible hyponyms for a given word.

- **vehicle** : car, ship, train, ...
- **fruit** : apple, pear, melon, ...
- **treatment** : ?
- **bird** : ?

Useful for query expansion, summarisation, entailment detection, smoothing language models, etc.

Approach

Using pairwise scoring:

- We take a very large set of candidate words.
- Score them with a vector similarity measure.
- Rank them according to the score.

Candidates:

- All words in the BNC that occur at least 10 times.

Evaluation dataset:

- Hypernyms from WordNet with all hyponyms
- Includes indirect hyponyms and synonyms
- Excludes infrequent hyponyms
- Training (1,230 hypernyms), development (922) and test (922) sets.

Vector Spaces

Comparison of different models

- **Window-based** baseline
- **Collobert & Weston** (2008) neural language model
- **Mnih & Hinton** (2007) neural language model
- **Word2vec** (Mikolov et al., 2013)
- **Dependency-based** features

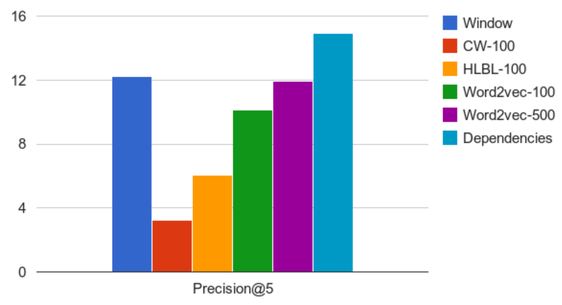

Figure 1: Precision with different vector spaces

Vector Offset

Evaluating the **vector offset** method (Mikolov et al., 2013b) for hyponym generation.

$$king - man + woman = queen$$

$$bird - fish + salmon = eagle?$$

The method did not give a consistent improvemnt on this task.

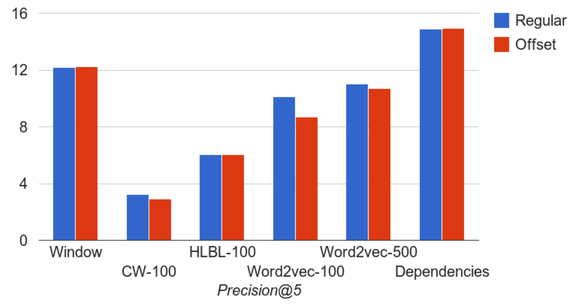

Figure 2: Precision using vector offset

Similarity Measures

We found two properties that improved accuracy:

- Assigning more importance to **features that are shared** between the two words
- Assigning more importance to **features of the broader word**

$$weight(f) = \begin{cases} (1 - \frac{rank_b(f)}{|F_b|+1}) \times (1 - C) + C & \text{if } f \in F_a \cap F_b \\ C & \text{otherwise} \end{cases}$$

	Pattern	Cosine	Lin	BalPrec	BalAPInc	WCosine	Combined
MAP	0.51	2.73	2.01	1.88	1.68	2.85	**3.51**
P@1	8.14	25.41	21.17	17.48	15.85	25.84	**27.69**
P@5	4.45	14.90	12.23	11.34	9.66	15.46	**18.02**

Table 1: Evaluation of different vector similarity measures on the test set of hyponym generation

Examples

scientist	sport	treatment
researcher	**football**	**therapy**
biologist	**golf**	**medication**
psychologist	club	patient
economist	**tennis**	procedure
observer	**athletics**	surgery
physicist	**rugby**	remedy
sociologist	**cricket**	regiment

Table 2: Example output from the best system

Conclusion

- Dependency-based feature vectors outperformed neural network models on this task, as they were able to better capture both the context and function of words.
- The vector offset method did not give an improvement, but performance could be improved by being more selective about the training examples.
- Symmetric similarity measures outperformed most existing directional measures on hyponym generation.
- We constructed a directional measure that achieved the best results on this task.

Download

Three vector sets are publically available:
www.marekrei.com/projects/vectorsets/