Extracting Hypernym Pairs from the Web

Erik Tjong Kim Sang, ISLA, Informatics Institute, Univ. of Amsterdam

Goal

Comparing automatic methods for extending lexical resources like WordNet with additional information, in particular hypernymy information: *furniture* is a hypernym of *table*.

Approaches

- Extraction of lexical hypernym patterns from text (Snow et al., 2005)

- Combination of hypernym evidence extracted with different corpus patterns (Snow et al., 2005)

- Deducing hypernym candidates from morphological information (Sabou et al., 2005)

- Extraction of hypernym evidence from coordinations (Caraballo, 1999) collected from the web

Evaluation

- Extracted hypernym-hyponym pairs were compared with the Dutch part of EuroWordNet

- For each hyponym, only one candidate hypernym was considered

- Ancestor hypernyms were also accepted as correct hypernyms.

- The average distance between hyponyms and accepted candidate hypernyms was measured

Results: summary

- Hypernym extraction from the web outperformed extraction from corpora.

- Hypernym deduction from morphological information outperformed all other methods.

- A combination of the morphology approach and web extraction obtained the best recall score.

Results: scores

Method	Prec.	Recall	F	Dist.
corpus: *N zoals N*	0.22	0.0068	0.013	2.01
corpus: combined	0.36	0.020	0.038	2.86
corpus: coordinations	0.31	0.14	0.19	1.98
web: *N zoals N*	0.23	0.089	0.13	2.06
web: coordinations	0.39	0.31	0.35	2.04
morphology	0.54	0.33	0.41	1.19
web + morphology	0.48	0.45	0.46	1.64
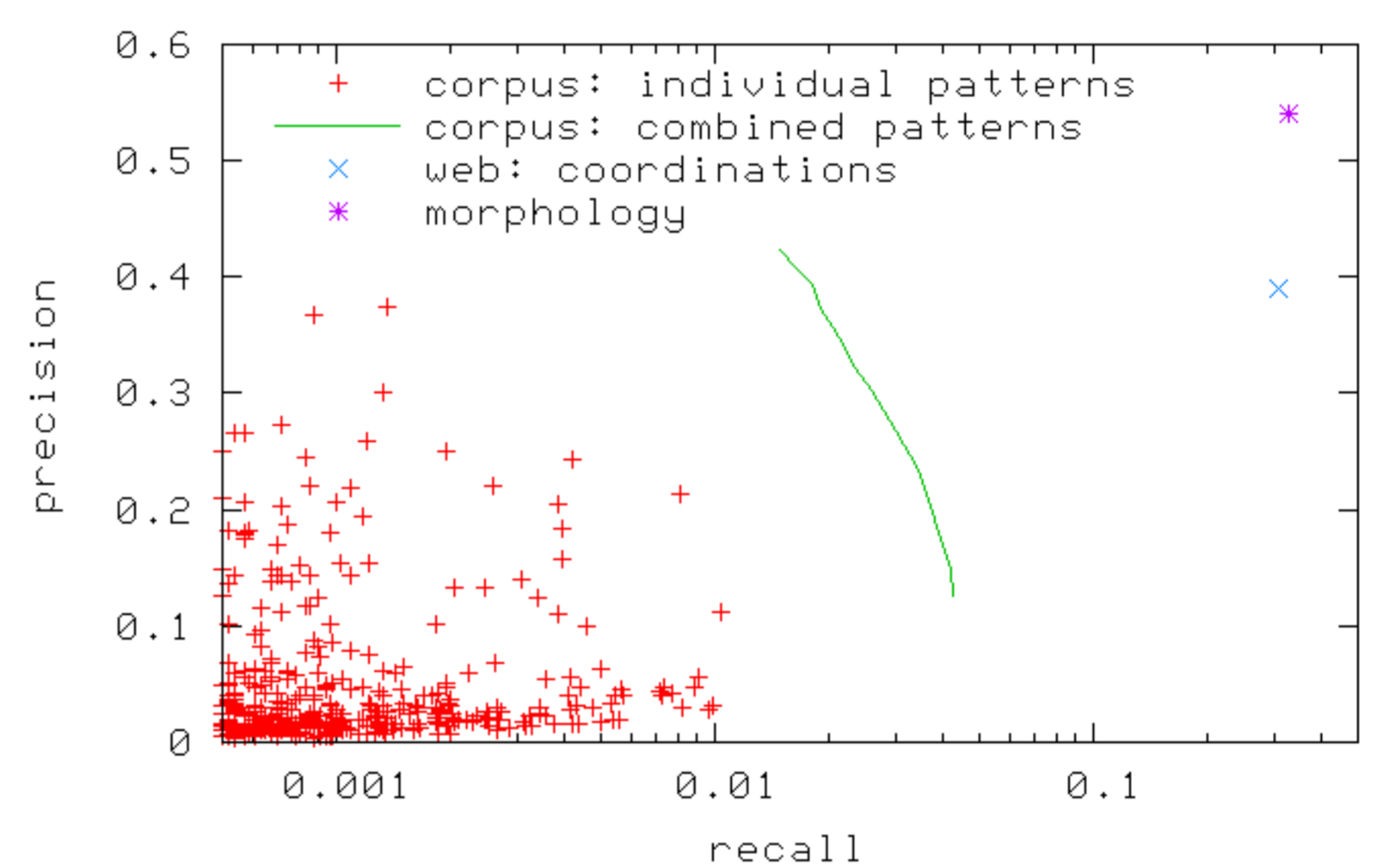
References

- Sharon A. Caraballo. 1999. Automatic construction of a hypernym-labeled noun hierarchy from text. In *Proceedings of ACL-99*. Maryland, USA.

- Marta Sabou, Chris Wroe, Carole Goble, and Gilad Mishne. 2005. Learning domain ontologies for web service descriptions: an experiment in bioinformatics. In *Proceedings of WWW2005)*. Chiba, Japan.

- Rion Snow, Daniel Jurafsky, and Andrew Y. Ng. 2005. Learning syntactic patterns for automatic hypernym discovery. In *Proceedings of NIPS 2005*. Vancouver, Canada.

- Rion Snow, Daniel Jurafsky, and Andrew Y. Ng. 2006. Semantic taxonomy induction from heterogenous evidence. In *Proceedings of COLING/ACL 2006*. Sydney, Australia.