

Distributional Analysis of Related Synsets in WordNet for a WSD Task

Kostas Fragos¹, Yannis Maistros²

1 Department of Computer Engineering, National technical University of Athens, Iroon
Polytexneiou 9 15780 Zografou Athens Greece
http://nts.ece.ntua.gr/nlp_lab

kfragos@ece.ntua.gr

2 Department of Computer Engineering, National technical University of Athens, Iroon
Polytexneiou 9 15780 Zografou Athens Greece

maistros@ece.ntua.gr

Abstract.

This work presents a new method for an unsupervised word sense disambiguation task using WordNet semantic relations. In this method we expand the context of a word being disambiguated with related synsets from the available WordNet relations and study within this set the distribution of the related synset that correspond to each sense of the target word. A single sample Pearson-Chi-Square goodness-of-fit hypothesis test is used to determine whether the null hypothesis of a composite normality PDF is a reasonable assumption for a set of related synsets corresponding to a sense. The calculated p-value from this test is a critical value for deciding the correct sense. The target word is assigned the sense, the related synsets of which are distributed more "abnormally" relative to the other sets of the other senses. Our algorithm is evaluated on English lexical sample data from the Senseval-2 word sense disambiguation competition. Three WordNet relations, antonymy, hyponymy and hypernymy give a distributional set of related synsets for the context that was proved quite a good word sense discriminator, achieving comparable results with the system obtained the better results among the other competing participants.