

Investigating graded meaning representations

Many words are polysemous. Different meanings can be very distinct, like the financial and the embankment senses of "bank", or they can be closely related, as in "show a picture" and "show a theorem". The task of determining the meaning of a word in a given context is usually phrased as one of word sense disambiguation (WSD), choosing the right sense out of a list of dictionary senses. However, there are issues with this enterprise both in terms of cognitive validity and adequacy for computational linguistics applications.

Research on the psychology of concepts shows that categories in the human mind are not simply sets with clear-cut boundaries: Some items are perceived as more typical than others, and there are borderline cases on which people disagree more often, and on whose categorization they are more likely to change their minds. Word meanings are certainly related to mental concepts. This raises the question of whether there is any such thing as the one appropriate sense for a given occurrence.

This talk examines the case for a graded notion of word meaning through two annotation experiments. The first task used WordNet senses, but asked for graded annotator judgment. The second task asked annotators to judge the similarity of pairs of usages of a target word, without recourse to dictionary senses. We find that the graded responses correlate with annotations from previous datasets, but sense assignments are used in a way that weakens the case for clear cut sense boundaries. The responses from both experiments also correlate with the overlap of paraphrases given by annotators in the English lexical substitution task, which bodes well for the use of substitutes as a proxy for word sense.

Focusing on the graded sense judgments dataset, we then consider the question of computational modeling. We study two models for the task of graded word sense assignment. The first model interprets the confidence level returned by a traditional word sense disambiguation model as a graded sense judgment. The second models positions senses as "landmarks" in a vector space representation of word meaning. We also propose evaluation measures for graded sense assignment systems, among them an extension of the notions of precision and recall to the graded case.