



Peter D. Turney
National Research Council
Canada

Friday, August 30
2:30PM, Doherty Hall 2315

Experiments with Three Approaches to Recognizing Lexical Entailment

Inference in natural language often involves recognizing lexical entailment (RLE); that is, identifying whether one word entails another. For example, buy entails own. Two general strategies for RLE have been proposed: One strategy is to manually construct an asymmetric similarity measure for context vectors (directional similarity) and another is to treat RLE as a problem of learning to recognize semantic relations using supervised machine learning techniques (relation classification). In this talk, we experiment with two recent state-of-the-art representatives of the two general strategies. The first approach is an asymmetric similarity measure (an instance of the directional similarity strategy), designed to capture the degree to which the contexts of a word, a , form a subset of the contexts of another word, b . The second approach (an instance of the relation classification strategy) represents a word pair, $a:b$, with a feature vector that is the concatenation of the context vectors of a and b , and then applies supervised learning to a training set of labeled feature vectors. Additionally, we introduce a third approach that is a new instance of the relation classification strategy. The third approach represents a word pair, $a:b$, with a feature vector in which the features are the differences in the similarities of a and b to a set of reference words. All three approaches use vector space models (VSMs) of semantics, based on word-context matrices. We perform an extensive evaluation of the three approaches using three different datasets. The proposed new approach (similarity differences) performs significantly better than the other two approaches on some datasets and there is no dataset for which it is significantly worse. Along the way, we address some of the concerns raised in past research, regarding the treatment of RLE as a problem of semantic relation classification, and we suggest it is beneficial to make connections between the research in lexical entailment and the research in semantic relation classification.

Bio: Peter Turney is a Principal Research Officer at the National Research Council of Canada and an Adjunct Professor at the University of Ottawa. He obtained his PhD in 1988 from the University of Toronto and joined the NRC in 1989. His recent work focuses on machine learning applied to natural language. He is the author or co-author of more than eighty publications. In the past, he has been an Editor of Canadian Artificial Intelligence magazine, an Editorial Board Member, Associate Editor, and Advisory Board Member of the Journal of Artificial Intelligence Research, and an Editorial Board Member of the journal Computational Linguistics. He was involved in initiating the Wiki of the Association for Computational Linguistics in 2006 and continues to play an active role in its maintenance. His paper Mining the Web for Synonyms won the ECML PKDD 10 Years Award in 2011.