Departamento de Computación

UNIVERSIDADE DA CORUÑA

Seminar on Dependency Parsing

Friday, October 18th, 2013
Aula de Graos Víctor Gulías
Facultade de Informática, Campus de Elviña

[11:30] David Vilares (Universidade da Coruña):
"Dependency parsing for mining opinions on Twitter"

Abstract: In this talk we will present a system that classifies the polarity of Spanish tweets adopting a hybrid approach, which combines linguistic knowledge acquired by means of Natural Language Processing with machine learning techniques. As an initial step, a preprocessing of the tweets is performed to address some characteristics of the language used in Twitter. Next, the syntactic structure of tweets is obtained by means of a dependency parser. This structure is then used to address three of the most significant linguistic constructions for the purpose in question: intensification, subordinate adversative clauses and negation. All the information extracted is used as features for a supervised classifier. A system implementing this approach has successfully competed in the recent campaign TASS 2013.

[12:00] Joakim Nivre (Uppsala Universitet, Sweden):
"Training Deterministic Parsers Using Non-Deterministic Oracles"
(joint work with Dr. Yoav Goldberg, Bar-Ilan University)

Abstract: Greedy transition-based dependency parsers are among the fastest available systems for natural language parsing but tend to suffer from search errors and subsequent error propagation. This problem is aggravated by the fact that they are normally trained using oracles that are deterministic and incomplete in the sense that they assume a unique canonical derivation and are only valid as long as the parser does not stray from this path. In this talk, I will introduce oracles that are non-deterministic and complete in the sense that they allow different derivations of the same parse tree and make optimal predictions for all parser states, including states from which the gold parse is no longer reachable. Experimental evaluation on a wide range of data sets shows that using these oracles to train greedy parsers gives substantial improvements in accuracy because of reduced error propagation. Moreover, this improvement comes at no cost in terms of efficiency, unlike other techniques like beam search.

Organised by:

Lengua y Sociedad de la Información (www.grupolys.org)
Carlos Gómez Rodríguez (cgomezr@udc.es)

Supported by:

Rede Galega de Procesamento da Linguaxe e Recuperación de Información (redplir.grupocole.org)

XUNTA DE GALICIA