Abstract

In this paper we present a method for automatically extracting definitional contexts from restricted domains in Spanish. Definitional contexts are textual fragments where there is an implicit definition that can be identified by taking into account verbal patterns linking a term and its corresponding definition. Our interest is in definitional contexts with analytical definitions.
Therefore, we focus on the extraction of textual fragments with a term and a hypernym. Then, hypernym is used for filtering non-relevant contexts by means of the occurrence frequency. It is assumed most frequent hypernyms have a higher probability of giving true definitional contexts than those less frequent hypernyms. We captured regularity in analytical definitions by means of a chunk grammar. This method achieves acceptable results in precision and recall compared with other works.

Reference

- Aguilar, C. 2009 Análisis Lingüístico de Definiciones en Contextos Definitorios. Tesis de doctorado, UNAM.

Index Terms
Computer Science  Computational Linguistics

Key words
Natural language processing
Information extraction
Conceptual extraction
Lexical relations