

Contextual Ontology Learning

Dr. Lobna KAROUI^a

^a *Université Paris-Sud Orsay, France*

Abstract

Keywords:

1. Introduction

Context is an elusive concept that cannot be defined automatically by a machine. Many researches, related to text analyses, word sense disambiguation or ontology learning, have used the context as a sentence, a syntactic structure, a set of sentences to extract concepts but without tacking into account the structure of the document and the relation between the contexts.

2. Aim

In our work, we define the contexts based on the document structure to improve the conceptual quality of the ontological concepts.

3. Material and methods

For our purpose, we define a context that exploits the html structure and the location of words to select the semantically closer cooccurents for each word and to improve the word weighting. Guided by this context definition, we propose a contextual clustering algorithm that refines the context of each word cluster to obtain semantically extracted concepts. The unsupervised hierarchical clustering divides recursively each cluster based on an automatic context definition. We experiment our algorithm on French HTML document corpus related to the tourism domain. The results show that the appropriate context definition and the successive refinements of clusters improve the relevance of the extracted concepts in comparison with a kmeans algorithm.

4. Results

Defining automatically different contexts based on linguistic context, documentary context or structural context enhance the set of extracted concepts and obtain more specific ones. Our results show that our context-based algorithm improves the clusters' conceptual quality and the relevance of the extracted ontological concepts.

January 21, 2018

5. Conclusions

As known, knowledge acquisition is a difficult task. In this work, we have defined contextual aspects to extract semantic from the Web. These contexts are used to refine the clustering approach to improve the concept extraction and the ontology learning. We continue our approach by defining “the impact of contexts on the intelligent interpretation” and extend the contextual model to others types of documents such as XML, pdf, word, etc.

6. Keywords

Context, machine learning, ontology, concept, clustering, learning, semantic