

Abstract

An Ontology can be defined as a formal representation of a set of concepts within a domain and the relationships between those concepts. The development of the semantic web initiative is rapidly increasing the number of publicly available ontologies. In such a distributed environment, complex applications often need to handle multiple ontologies in order to provide adequate domain coverage.

Surprisingly, there is a lack of adequate frameworks for enabling the use of multiple ontologies transparently while abstracting the particular ontological structures used by that framework. Given that any ontology represents the views of its author or authors, using multiple ontologies requires us to deal with several significant challenges, some stemming from the nature of knowledge itself, such as cases of polysemy or homography, and some stemming from the structures that we choose to represent such knowledge with.

The focus of this thesis is to explore a set of techniques that will allow us to overcome some of the challenges found when using multiple ontologies, thus making progress in the creation of a functional information access platform for structured sources. In this thesis we try to address the question “How do we integrate and use information contained in a set of heterogeneous ontologies?” This question is becoming crucial as the world gets more connected. As the amount of information grows, so does the amount of resources that encode domain knowledge in different forms. At the same time, as we deal with problems that span a large number of domains, such as search, question answering and semantic analysis, the need for concurrent access to multiple ontologies becomes ever more self-evident.

We model our approach on the general framework proposed in Federated Search and transpose the set of problems to the ontological domain. We start by illustrating the variety of available ontologies and focusing on automatic ontology creation. We proceed to describe the use of pro-active ontology selection to guide the ontology selection process at the query level. We then propose a set of operators for ontological search centered on the user’s information needs and expand the basic operator set through composition of basic operators. We address the problem of graph merging within Ontology Search and propose a scoring metric for results that is based in the proposed operator set.

Finally, we present results of using the Federated Ontology Search (FOS) engine in a set of task based evaluations. Our task set is comprised of type checking for question answering, concept coverage, concept disambiguation and content recommendation. We show that in all of the evaluated tasks, the FOS system outperforms the use of individual ontologies.