«Utilizing Ontologies and Using Semantic Web Technologies for Knowledge Representation, Management and Discovery over the Internet»

M.Sc. Thesis

Dimitrios Koutsomitropoulos

ABSTRACT

Current research on the evolution of the World Wide Web is moving towards more intelligent means for information discovery and extraction. The *Semantic Web* is an extension of the current Web where information is given well defined meaning, thus enabling machines to process data more efficiently and to "understand" the information that currently they could only render. For the Semantic Web to function, computers should have access to structured information collections known as *ontologies*. Ontologies provide for knowledge representation on the Semantic Web and consequently they can be leveraged by computing systems in order to conduct *automated reasoning*.

Several languages have been proposed and developed for the description and representation of Semantic Web ontologies: The most prominent include *DAML+OIL* (DARPA) and the state-of-the-art *OWL* (W3C). These languages, among others, are compared and evaluated in a systematic manner, using a thorough evaluation framework. Moreover ontology-based models and approaches are provided, for managing knowledge through its whole lifecycle.

Description Logics provide the theoretical background where the representational and syllogistic power of ontologies is based upon. Description Logics are a well-defined, decidable subset of First Order Logic and their formal properties are shown to be suitable for knowledge representation and discovery in the Semantic Web. In addition, some very up-to-date results for reducing OWL to expressive Description Logics are provided, along with concrete reasoning complexity results.

Therefore, it is recommended to utilize Description Logics systems for discovering implied information. Such systems are proven to be advantageous, for the time being, in contrast to alternative approaches, like *Theorem Provers* and *Rule-Based* systems. Three Description Logic inference engines are thus considered and evaluated both empirically and experimentally.

RACER was chosen as the most suitable to provide its reasoning services for the prototype *knowledge discovery interface* that has been implemented in order to demonstrate

inferring capabilities over web-distributed OWL documents. The knowledge discovery interface implements an expressive inferring methodology and supports intuitive and declarative construction and submission of intelligent queries by the user.

Using this interface, a set of experimental inferences is conducted, based on cultural heritage knowledge, as expressed by the *CIDOC-CRM* ontology. CIDOC-CRM was ported into OWL syntax and extended, in a distributed manner, with concrete instances and more expressive structures. This semantic enrichment was possible after expressing the ontology in OWL.