

Research Scholar: Nadia Imdadi

Supervisor: Dr. S.A.M. Rizvi

Department of Computer Science

Title: *Global Framework for Automatic Semantic Integration Incorporating Semantic Repositories*

ABSTRACT

Ontologies are conceptual representation of domains in a formal language that make data machine processable over the web. They are key elements that allow knowledge to be represented in a structured way so that a higher degree of interoperability amongst the various heterogeneous resources on the web may be achieved. They are the hinges upon which Semantic Web is built upon. A key factor for the success of semantic web is availability of technologies for the efficient and effective reuse of ontological knowledge.

Ontology engineering, the process of building ontology, is a time consuming activity which also requires domain specific skills generally given by experts of the field. The approach to ontology development can be broadly categorized into two areas one where creation is done from scratch and another through reuse, which generally is in the form of merging, integration, alignment, mapping or translation. The former form of development is painstaking while the latter makes use of already developed formal domain representations and though it requires a diligent attention it definitely cuts down on the time of development.

With standardization and maturity of semantic web languages that support description logic ontologies, on the web have mushroomed and are on the rise. The availability of these semantic resources augur well as they help to achieve the idea of the semantic web and form the necessary infrastructure where software agents can make decisions by inferring knowledge from a variety of resources. As interoperability amongst heterogeneous resources on the web is the operative term in context of semantic web the methodologies for efficient and effective reuse of information from different

resources that have extensions from various ontologies have to evolve. Reuse of existing ontological knowledge on the web to build ontologies for the semantic web may be helped by the efforts in the field of ontology engineering.

This study forwards a global framework for automatic semantic integration incorporating semantic repositories where the focus was to evolve strategies for development or engineering of ontologies through reuse of existing online semantic repositories namely the ontologies. Word Sense Disambiguation technique is used for identification of ontologies to be used during the phase of identification of relevant ontologies from the web. Hash based bucket algorithm is used for retrieval and subsequently ranking of the ontologies is done using an aggregation function. Significant components of a class definition are identified and concept retrieval across the retrieved ontologies is done based on the semantic relations that already exist along with the new ones learned. A Relative Semantic Similarity Measure is formulated to find similar versus dissimilar feature between two classes based on the significant components of a class. A novel method of presentation of result in form of adjacency matrix is suggested to demonstrate relations amongst the retrieved classes to assist ontology developer. For evaluation of created ontology, human based assessment on set of parameters is suggested in absence of reference ontology. In case reference ontology for a domain is available then evaluation may be done based on golden standard methodology.

The framework is evaluated based on existing golden standard methodology of ontology evaluation with promising results. Concepts from food and academic domain are used for demonstrating and evaluation of the methods proposed by the framework to address the various issues at all stages of ontology development through reuse of semantic repositories available via the Semantic Web.