

ABSTRACT

Title: Semantic Similarity Measures Using Ontology and Information Content based approaches for Multiple Subsumed Concepts

Semantic Web, the vision of Tim Berner's Lee, is an extension of conventional web and not a separate web. The Semantic Web gives clearly defined meaning to the information on the web, which enables people and computers to work in coordination. Ontology is considered to be a significant component of Semantic Web which describes concepts, their relationships, and properties within a domain. To develop any Knowledge System the most important step is the construction of domain ontology. Ontology development is mostly done either from scratch or fully automatic with very less user intervention. This thesis proposes a "Semi-Automatic Reusability Approach to Ontology Building" which utilizes the online ontologies and proves to be better than the earlier methods. To support interoperability, mapping and merging of ontologies on the Semantic Web, it is necessary to find the semantic similarity of any two concepts in ontology.

The main objective of this thesis is to handle the problem raised by existing measures during semantic similarity computation, especially for those concepts which exhibit multiple inheritance. This thesis proposes Ontology-Based and Information Content (IC) Based measures for computing semantic similarity. The proposed measures are applied on concept pairs of two datasets MeSH.owl and Human.owl. To evaluate the efficiency of the proposed measures in similarity calculation, the similarity values are correlated with Human Judgement (HJ) values. In the experimental results, the existing Edge-Based measure shows 60% correlation with HJ values which is less than the proposed measure whose correlation is 73%. In IC based measures, the proposed similarity measure shows 85% and 81% correlation in two cases with HJ values and the existing measures show 43% and 39% correlation in respective cases.