論文の内容の要旨

An Approach for Semantic Search in the Web including Databases

(データベースを含む Web のセマンティック検索に向けたアプローチ)

マムドゥーヘ　ファルーケ　モハメド　ファルガリ

Recently, huge amount of data is published on the Web and it continues to increase with an explosive speed. Finding information on the Web became harder. Semantic web is an extension to the current web in which web data is organized in a manner that facilitates getting information. The key point in the new web is data representation. The data should be represented into a semantic format that enables web agent to understand the data. Many semantic languages have been specified for representing data in semantic web. Resource Description Framework (RDF) is a standard language for semantic web that is supported by W3C.

The main problem of semantic web is depending on domain ontology for data representation. However, in this research we propose representing data into Concept Description Language (CDL) which does not depend on domain ontology. An approach for converting web databases to CDL semantic format is proposed.

Searching the semantic web data is challenge to exploit the semantic representation of data and satisfy users' needs. A semantic search engine based on CDL is proposed for query CDL data. the proposed search technique based on semantic relation between query concepts and semantic

expansion for query keywords. The proposed technique is based on semantic graph matching. Moreover, two different semantic hierarchies are used for semantic expansion for query keywords. The first hierarchy is Universal Word (UW) semantic hierarchy and the second one is WordNet thesaurus.

A prototype for the proposed search engine is implemented and tested to show the effectiveness of the proposed technique. A micro web is a dataset represented into CDL. It includes web DBs converted to CDL. Micro-web dataset is used to test the proposed search engine. The results of the proposed approach is compared by the results Google search and UNL semantic search engine. Our results are very promising comparing to others.

Finally, an injection technique is proposed to inject RDF ontologies with CDL comments. The injected CDL comments are used by the proposed search engine to make ontology mapping for properties belongs to different ontologies. The generated mapping is used for SPARQL query expansion. SPARQL is a standard query language for RDF data.