



Improving the efficiency of OWL based traversal using revisiting avoidance algorithm

R.S.Ramya, V.Praveena and A.Dinesh Kumar

Department of Computer Science and Engineering, N.G.P. Institute of Technology, Coimbatore-48.

ARTICLE INFO

Article history:

Received: 11 June 2013;

Received in revised form:

20 January 2014;

Accepted: 24 January 2014;

Keywords

Ontology,

Semantic web,

Owl,

Knowledge database.

ABSTRACT

Ontologies have become the best modeling tool, in many applications majorly semantic web. Creation of ontology remains as a difficult task. Ontology bootstrapping is one of the technique, depends on a set of pre-defined textual sources like web services. The advantage of web services usually consists of both WSDL and free text descriptors. And these are evaluated by three methods namely TF/IDF, web context generation and the bootstrapping technique. The bootstrapping method using large repository of real-world web services has a drawback of revisiting of the concepts again and again. Our proposed method for ontology creation with an revisiting avoidance algorithm, this algorithm helps us to overcome from revisiting strategies in the bootstrapping ontologies. The introduced algorithm will produce the efficient result as described this is paper.

© 2014 Elixir All rights reserved.

Introduction

The semantic web is the second generation of web used to share and reuse of data across the application, enterprise and community boundaries. The major component of the semantic web is ontology that is study of existence. Ontology legally represents the knowledge as a set of concepts within a domain and the relationship among those concepts. Ontologies are structural context for organizing information. It can be used to reason about the entities within that domain and may be used to describe a domain. Ontologies have been used commonly in many fields such as knowledge representation, information retrieval, natural language understanding and web services [1]. Semantic web contains many distributed ontologies with overlapping domains [2].

To allow for inter operation between applications on the web, these ontologies need to be related to each other through ontology mapping. Mapping of ontologies refers to naming of identical concepts or relations between different ontologies [3]. Ontology mapping is a challenging topic. It can be done either manually or using semi-automated or automated tools. Manual mapping becomes impossible as the size and complexity of ontologies increases [1].

The W3C standards for the semantic web include the web ontology language (OWL) [5]. Owl is an ontology specification language and the RDF (Resource Description Framework) is a language for describing resources that exist on the web. The disadvantage of natural language in a way to represent formal, unambiguous contents is overcome by using the semantic web technology [4]. The OWL is proposed to provide a language that can be used to describe the classes and the relationships between them that are essential in web documents and applications [5].

Knowledge base is the special kind of database for knowledge management. It is an information origin that provides a means for information to be collected, organized, search, share and utilized. It is of two types. 1. Machine readable stores the knowledge in computer readable form. They contain set of data often in form of rules that describes the knowledge in

logical consistent manner and 2. Human readable are designed to enable people to retrieve and use the knowledge they contain.

Related work

Aviv Segev [9] the author describes a model called bootstrapping model. This model is based on the continuous analysis of WSDL documents and employs an ontology model based on concepts and relationships [10].

Ondrej Svab [12] describes a mapping pattern as a graph structure. This paper going to examine about three simple patterns. And it is interest in two types of ontology design patterns: naming conventions which are related to naming classes, properties and/or instances and structural patterns concern the modeling options in using certain ontology entities and connecting them together. Yves et.al., [13] proposed a new method by accompanying the ambiguity modeling for ontology mapping using Naïve Bayes theorem. This methodology works without taking the degree of ambiguity into consideration with manual validation. Ryutaro Ichise [15] States that the different similarity measures such as concept similarities including the string based similarity, graph based similarity, instance classification similarity and knowledge resource similarity are proposed for the mapping process [15].

Owl

Semantic web is the vision for future of the web in which information is given explicit meaning, making it easier for machines to automatically process and integrate information available on the web. The first level above the RDF requirement for the semantic web is an ontology language that can formally describe the meaning of terminology used in web document. OWL facilitates greater machine interoperability of web content can be supported by XML, RDF and RDF schema by providing additional vocabulary along with a formal semantics.

Grigories Antoniou [16] has mentioned OWL could be an extension of RDF schema; OWL could use RDF meaning of classes and properties that support the richer effectiveness. RDF Schema has some very powerful modeling primitives, such as the `rdf:type` and `rdf:property`.

Owl has three different sub languages,

➤ **OWL Full:** The complete language is called as owl full, and uses all owl language primitives. It is also used to combine all the primitives with RDF and RDF schema. The fully upward is compatible with RDF, both syntactically and semantically: any legal RDF document is also a legal OWL Full document, and any valid RDF/RDF Schema conclusion is also a valid OWL Full conclusion. These are the some advantages of OWL Full.

➤ **OWL DL:** It was designed to provide the maximum expressiveness possible while retaining computational completeness. It is a sub language of OWL Full which restricts the way in which the constructors of OWL and RDF can be used. It permits the efficient reasoning support which is a major advantage of OWL DL. The disadvantage of OWL DL is, it will loose full compatibility with RDF.

➤ **OWL Lite:** It was originally intended to support the users primarily needing a classification hierarchy and simple constraints. The advantage of this is a language that is both easier to grasp and easier to implement. The disadvantage is restricted expressivity.

RDF

The W3C has defined a number of description languages for creation of ontology. The first published language was the RDF. Marvin P Palson [6] describes that RDF uses the XML (Extensible Markup Language) as a syntax model. RDF-Grap is one of the abstract models. Both syntaxes describe the RDF-statements or RDF-triples and every triple has a resource (Subject), a property (Predicate) and a property value (object). These three elements are the major essentials of a RDF-statement and also a relationship between things. A RDF-statement is an explanation of a knowledge fact [6]. The RDF enables the creation and exchange of resource metadata as normal web data. To interpret these metadata within or across user communities, RDF allows the definition of suitable schema vocabularies (RDFS) [7]. The main aim in using ontologies is [8],

➤ To share common understanding of the structure of information among people or software agents.

➤ To analyze domain knowledge.

➤ To make domain assumptions explicit.

➤ To enable reuse of domain knowledge.

➤ To separate domain knowledge from the operational knowledge.

Bootstrapping model

The bootstrapping model is used to create ontologies which centers on 1). The combination of the use of two different extraction methods, TF/IDF and web based concept generation, and 2). The verification of the results using a free text description verification method is done by analyzing the external service descriptor [9]. The Fig. 1 shows the overall bootstrapping process. There are four main steps in the process. The token extraction step is used to extract the tokens representing relevant information from the WSDL document. This step extracts all the name labels, parses the tokens, and performs initial filtering.

The Second step analyzes in parallel the extracted WSDL tokens using two methods. In particular, TF/IDF analyzes the most common terms appearing in each web service document and appearing less frequently in other documents. The Web context extraction uses a set of tokens as a query to a search engine, clusters the results according to textual descriptors, and classifies which set of descriptors identifies the context of the web service.

The Concept evocation step identifies the descriptors which appear both in the TF/IDF method and the web context method.

These descriptors identify possible concepts names that would be utilized by the ontology evolution. The context descriptors also assist in the merging process of the relations between concepts. The final step is ontology evolution that expands the ontology as required according to the newly identified concepts and modifies the relation between them [9].

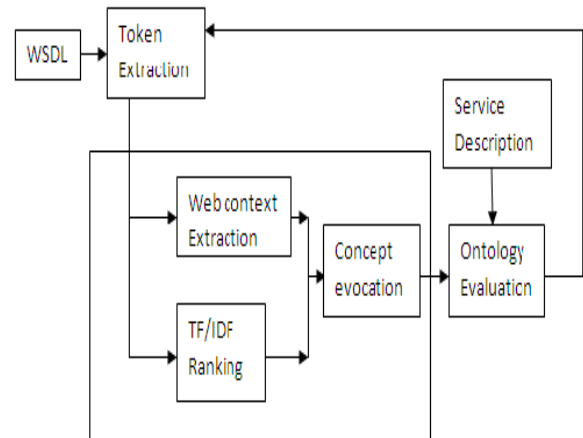


Fig 1. Web Service Ontology Bootstrapping Process

This model is used to create ontologies in an effective and efficient manner. It used the concept of WSDL language and the OWL and RDF Language. It is one of existing method for better creation of ontology. The drawback with this method is revisiting of concept again and again decreases the overall performance of the search engine. We propose an algorithm called revisiting algorithm.

Revisiting Algorithm

The RDF language split the searching key into three parts and sends to the crawler. The crawler will make different combinations of the parts and search for the exact data and put the data into the knowledge database or annotated database. The ontology is used to group the similar meaning words and form a cluster for easier identification of the searched data. Searching of different keys with same meaning will be done easily with the help of existing bootstrapping model and the revisiting algorithm to overcome from the drawbacks of this model. Revisiting algorithm is used to overcome from the above drawback and improving the performance and searching accuracy in minimum time.

Algorithm

Let t be the average searching time of the node in knowledge/Annotated database.

Let a_n and a_m be the searching nodes in the ontology.

If $(a_n > a_m)$

a_n = visit all the nodes related to a_n

else

a_m = visit all the nodes related to a_m

if $(a_n || a_m > 5)$

don't visit either of nodes calculate the average time for visiting all related nodes in a_n and a_m using,

$t(a_n + a_m)$

The a average time for visiting the related nodes in a_n and a_m is less when compared with ordinary search engine.

Result And Discussion

The analysis graph for the existing search engine and the proposed search engine is shown in fig 2 and this graph shows the improvement in the performance of the OWL based traversal in semantic web technology. As shown in the fig 2 the performance of the proposed search engine is comparatively high when compared to the existing search engine. It also used to identify total number of records retrieved per minute in

existing search engine as well is proposed search engine. The number of records retrieved by the proposed search engine is comparatively high. The searching time of the existing search engine is 0.75 milliseconds; our proposed search engine will produce the result in 0.01 milliseconds. Thus the searching time is reduced in proposed search engine. Thus the proposed revisiting avoidance algorithm helps us to improve the performance and decrease the searching time.

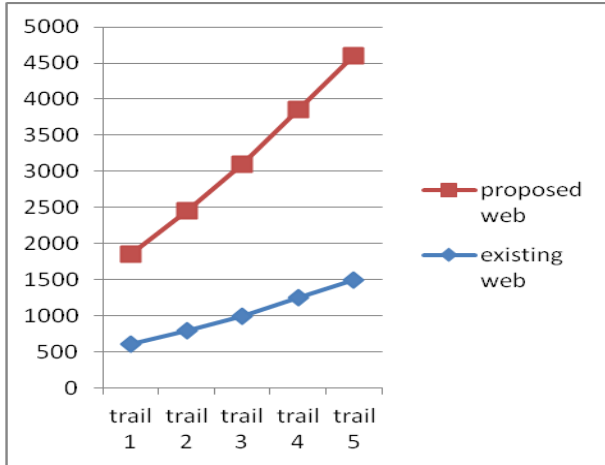


Fig 2 analysis of no.of records retrieved per minute
Conclusion

The Semantic web is the emerging search engine for performing the meta data based search. It is used to overcome from the drawbacks such as content or title based search and no relationship between any data and pages. The ordinary search engine uses the goggle bot to get the effective result as like semantic web. The semantic search engine produces the feasible result than the ordinary search engine. The existing technique bootstrapping ontology is used to create an effective ontology for the semantic web. But the drawback with this existing method is revisiting of same page again and again. The proposed work is used to overcome from the revisiting of pages. So the proposed technique resulted in improvement in the efficiency of the OWL based traversal and produces the efficient search result by decreasing the searching time.

References

- [1] Chitra S and Aghila G, "SOMPER: Mapped Mets Information (MMI) for Semantic Web Using Ontology Mapping Patterns", *International Journal of Computer Applications*, vol 14, no-5, Jan 2011.
- [2] Hendrik Thomas, Declan O' Sullivan, Rob Brennan, "Evaluation of ontology Mapping Representations", *Knowledge and Data Engineering group*, 2009.
- [3] Doan, A; Madhavan, J; Domingos, P; Halvey, "A Learning to map Between Ontologies on the Semantic Web", *Proceedings of the Eleventh International WWW Conference*, 2002.
- [4] Yajing Zhao and Jing Dong, "Ontology Classification for Semantic-Web-Based Software Engineering", *IEEE Transactions on Services Computing*, vol 4, no 4, Oct-Dec-2009.
- [5] Web Ontology Language OWL, <http://www.w3.org/2004/owl>, 2009.
- [6] Marvin P Polson and Sumathy Eswaran, "Ontology Creation and Semantic Web for Paddy", *International Journal of Computer applications*, vol 41, no.9, Mar 2012.
- [7] Grigoris Antoniou, "A Semantic Web Primer", 2nd Edition and 2008 *Frank Van Harmelen, MIT Press*.
- [8] D.Brickley, R.V.Guha, "Resource Description Framework (RDF/S) Specification".
- [9] Aviv Segav and Quan Z. Sheng, "Bootstrapping Ontologies for Web Services", *IEEE Transactions on Services Computing*, vol.5, no.1, Jan-Mar-2012.
- [10] Siham Amrouch and Sihem Mostefai, "Ontology Interoperability Technique, the state of the Art", *Journal of Information Organization*, vol.2, no.1, Mar 2012.
- [11] Ondrej Svab, "Exploiting Patterns in Ontology Mapping", *6th International Conference on Semantic Web Conference*, 2008.
- [12] Yves R. Jean-Marya E, Patrick Shironoshitaa, Mansur R. Kabuk, "Ontology Matching with Semantic Verification", *Web Semantics: Science, Services and Agents on the WWW*, vol.21, 2009.