International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2013): 4.438

Semi-Automatic Ontology Merging of Domain Specific Ontologies

C. P. Abinaya[1], V. P. Sumathi[2]

[1]PG Student, Department of Computer Science and Engineering, Kumaraguru College of Technology, Coimbatore

[2]Assistant Professor, Department of Computer Science and Engineering, Kumaraguru College of Technology, Coimbatore

Abstract: ***Ontologies aim at providing a formal, explicit and shared conceptualization and understanding of common domains between different communities. However, the distributed nature of ontology development resulted in the semantic heterogeneity between ontologies. Semantic heterogeneity constitutes the major obstacle against achieving interoperability between ontologies. To overcome this obstacle a novel method is designed for semi-automatic ontology merging. It combines syntactic and semantic measures for identifying similar concepts that will be merged in a signal one in the resulting merged ontology. The structural relationships between ontologies are solved by similarity measures. After combining the results, similar concepts are merged in a single one by normalizing the weights of all the similarity measures.***

Keywords: Domain ontology, interoperability, semantic, structural relationship, merging.

1. Introduction

Ontology has been developed to offer a commonly agreed understanding of a domain that is required for knowledge representation, knowledge exchange and reuse across domains. Therefore, ontology organizes information into taxonomies of terms (i.e., concepts, attributes) and shows the relationships between them. In fact, it is considered to be helpful in reducing conceptual confusion for users who need to share applications of different kinds, so it is widely used to capture and organize knowledge in a given domain.

Semantic is the process of adding information and description to the resources that help us to understand the meaning of these resources carried out in semantic web. Many researches carried out in semantic web among that ontology merging is the key issues in this era. The semantic web uses RDF to describe web resources with background in logic and artificial intelligences. Its utility depends on three issues such as Availability (existences of data), Accessibility (users can retrieve the data they want), Quality (user can judge the quality of the retrieved data).

Ontology is the platform for sharing the knowledge of domain that helps the machine to make intelligent decision. According to T.Gruber [13], Ontology is the explicit specification of a conceptualization. Conceptualization is a description of concepts and relationship that exist. It corresponds to an abstract of a domain which indentifies the relevant concepts and relationship.

Formal specification defines the machine readable with computational semantics. Ontology is developed by different people in different format which causes heterogeneity problem [1] that leads to an inaccurate search results in semantic web. Semantic heterogeneity is not resolved efficiently. The semantic heterogeneity is caused by different meaning or interpretation of data.

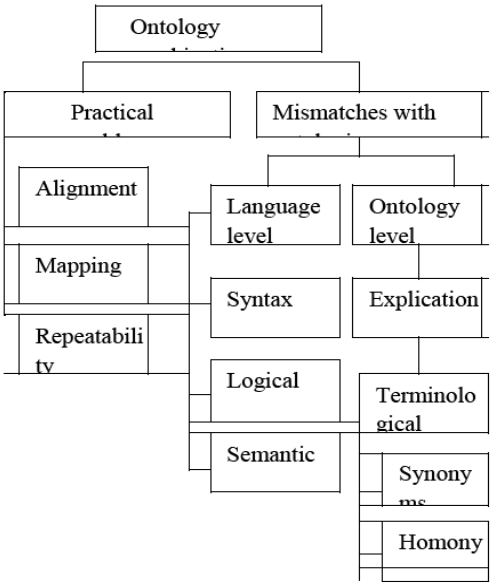

Figure 1: Problem on ontology combination

Different types of mismatches may occur between different ontologies. The identification of thesetypes of mismatches is essential in order to solve them during mapping, alignment, merging process. Ontology Merging as shown in Figure 1is the process of generating a single coherent ontology [8] from two or more existing and different ontologies related to the same subject. A merged single coherent ontology includes information from all sources ontologies but is more or less unchanged. Contemporary ontologies share many structural similarities. It describes instances, classes, attributes and relations. OWL ontology is interpreted as a set of axioms that provide semantics by allowing system to infer additional information based on the data explicitly provided.OWL is both syntax for describing and exchanging on ontologies, and has a formally defined semantics that gives the meaning. Protégé tool is open source tool used for creating, mapping and merging ontologies of same domain .

The contribution presented in this paper minimizes human involvement during ontology merging. Ontology merging

approach is suggested that semantic heterogeneity can be resolved with the help of ontology[5]. The matching strategies are proposed for indentifying the similarities and dissimilarities of source ontologies are merged as a global ontology that resolve the synonyms and relationship conflict among the domain specific ontologies[7]. To improve the accuracy of data conflict resolution very efficiently, novel architecture is being proposed.

This paper is organized as follows. Some related research works are briefly reviewed in Section 2. Proposed approach is explained in Section 3 and in Section 4 conclusion is drawn and some future directions are pointed out.

2. Related Work

This section deals with the issues of ontology merging. The merging is the bottleneck in the research of semantic field. Recently, some interesting techniques and methodologies are focus on the interoperability among the domain specific data sources.

SihamAmrouch and SihemMostefai [12] proposed a syntactic and semantic similarity methods are important in the process of merging. The syntactic is computed based on Jaro Winkler distance which measures the similarity between the concepts of strings. Semantic technique uses word net[2] dictionary as an external resources to obtain the equivalent correspondence and then merged as single ontology.

Mohammed Maree and Mohammed Belkhatir [6] Heterogeneous problem is a main issue of merging the domain specific. Many approaches fail to produce the semantic among the ontologies. Proposed a name based approach finding the equivalent classes, properties of object. The statistical based technique using Normalized Retrieval Distance (NRD) function to define the missing concepts of knowledge base. It uses website for the information content. It requires more human intervention.

C.R Rene Robin and G.V.Uma [3] proposed a hybrid algorithm for automatic merging of ontologies. The approach consists of four strategies such as heuristic function, lexical, semantic matching and similarity checking. The two domain specific owl files are given as a input .The lexical and semantic compare the class names. The process proceed with the top- down strategy to avoid conflict among merging[9]. Heuristic similarities checking of properties are called to check the similar properties of classes. The process is repeated for every class of owl file.

Many approaches that were proposed are lack in handling the heterogeneity in an efficient way and failed to handle the homonyms conflict as a great issue, thus the resolution results of those are often inaccurate. Thus a system using knowledge base with the help of Wordnet for handling conflicts during ontology merging process. The semantic heterogeneity is handled with the help of ontology as it provides richer semantics such that conflicts are removed and precision is increased.

C.R Rene Robin and G.V.Uma [3] proposed a hybrid algorithm for automatic merging of ontologies. The approach consists of four strategies such as heuristic function, lexical, semantic matching and similarity checking. The two domain specific owl files are given as a input .The lexical and semantic compare the class names. The process proceed with the top- down strategy to avoid conflict among merging [9]. Heuristic similarity checking of properties are called to check the similar properties of classes. The process is repeated for every class of owl file.

Many approaches that were proposed are lack in handling the heterogeneity in an efficient way and failed to handle the relationship conflicts as a great issue, thus the resolution results of those are often inaccurate. Thus a system using knowledge base with the help of Wordnet for handling conflicts during ontology merging process. The semantic heterogeneity is handled with the help of ontology as it provides richer semantics such that conflicts are removed and precision is increased.

3. Proposed Work

This paper proposes a novel architecture for ontology merging as shown in figure 2.

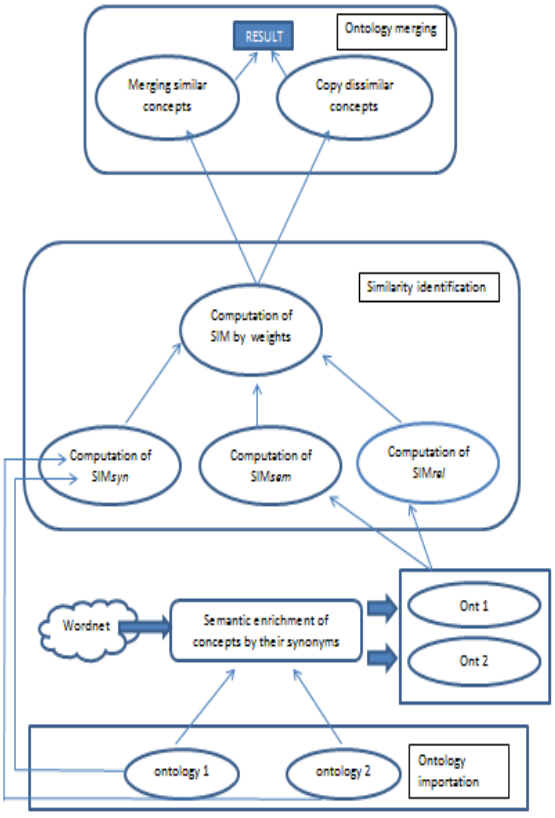

Figure 2: Architecture of ontology merging

3.1 Syntax similarity

Syntactic similarity will be calculated by using the following metrics :Levenshtein distance and Jaro–Winkler distance.
Levenshtein distance

In information theory and computer science, the Levenshtein distance is a string metric for measuring the difference

between two sequences. Informally, the Levenshteindistance between two words is the minimum number of single-character edits (i.e. insertions, deletions or substitutions) required to change one word into the other.

Jaro–Winkler distance
In computer science and statistics, the Jaro–Winkler distance is a measure of similarity between two strings. It is a variant of the Jaro distance metric, a type of string edit distance, and was developed in the area of record linkage (duplicate detection) (Winkler, 1990). The higher the Jaro–Winkler distance for two strings is, the more similar the strings are.

3.2 Semantic similarity

Semantic similarity will be calculated by using WorldNet and Description of the concept for two ontologies. WordNet is a lexical database for the English language. It groups English words into sets of synonyms called synsets, provides short definitions and usage examples, and records a number of relations among these synonym sets or their members. WordNet can thus be seen as a combination of dictionary and thesaurus. While it is accessible to human users via a web browser, its primary use is in automatic text analysis and artificial intelligence applications. The database and software tools have been released under a BSD style license and are freely available for download from the WordNet website. Both the lexicographic data (lexicographer files) and the compiler (called grind) for producing the distributed database are available.

3.3. Relationship similarity

The Relationship similarity is computed by combining the similaritiesbetween the parents and children of the entity being compared. Asentities may contain multiple parents and children, the similaritycalculation is normalized in order to restrict the results between 0and 1. For example, each parent of the source entity is paired withthe closest parent of the target entity; all parent-pair similaritymeasures are summed and then normalized by dividing by thetotal number of parent entities for both the source and target entities.

3.4 Ontology Merging

The two domain ontologies are merged together by measuring the similarities of syntactic, Semantic and relationship between both ontologies and the weights are normalised to merge the two ontologies into single domain specific ontology.

4. Conclusion

The method has been proposed to reduce the heterogeneous problem by providing a semi- automated merged framework. In the proposed approach the domain specific global ontology is Created by measuring the syntactical, semantic and relationship conflicts wordnet.. The similar classes and instance are combined as a single ontology. In the future work, aim to enhance the ontology merging of different domain.

References

[1] B.Organ, M.Dras, A.NayakG.James, “Approaches for a semantic interoperability between domain ontologies”, International Conference,Vol.35.pp,(2011).
[2] Cho,Kim,Kim.P,”A New Method for Ontology Merging based on Concept using Word net”, in advanced communication technology, Vol 3,(2006).
[3] C.R. Rene Robin, G.V. Uma, “A novel algorithm for a fully automated ontology merging using hybrid strategy”, Journal of scientific research, vol.47 no.1, (2010).
[4] Kamel Hussein Shafa’amri, Jalal Omer Atoum, “A framework for improving the performance of ontology matching techniques in semantic web”, International journal of advanced computer science and application vol.3.no.1 (2012).
[5] Michel Klein, “Combining and relating ontologies: an analysis of problem And solution”, International conference,(2009).
[6] Mohammed Maree, Mohammed Belkhatir, ”A Coupled Statistical and Semantic Framework for Merging Heterogeneous Domain Specific Ontologies”,International Conference on tool with Artifical Intelligence, Vol 3 (2010).
[7] Muhammad Fahad, Muhammad AdbulQadir,”A Framework for Ontology Evaluation”, pp, (2010).
[8] Natalya F. Nov, Dedorah LMc Guinness, “Ontology development: A guide to creating first ontology”, International Conference,(2009) .
[9] Nora Maiz, Muhammad Fahad ,Omar Boussaid, FadliaBentayeb, “Automatic Ontology Merging by Hierarchical Clustering and Inference Mechanisms” , proceeding of I- KNOW (2010),pp.81-93.
[10] Prasenjit Mitra, GioWiederhold, “Resolving Heterogeneity in Ontologies”, IEEE TKDE, Vol.99 (2010).
[11] Saeed Hassanpour, Martin J.o’Connor and Amar K.Das,”Visualizing Logical Dependencies in SWRL rule Bases”, International Conference, Vol.35.pp, (2012).
[12] Siham Amrouch, Sihem Mostefai, ”Syntatico- Semantic algorithm for an automatic ontology merging”, International Conference information technology and e-service, vol.5.pp, (2012).
[13] T.Gruber,”A translation approach to portable ontology specifications”, Knowledge acquisition, Vol.5.pp.199-220. (1993).
[14] R. Jean-Mary, Mansur R. Kabuka, ”ASMOV: Ontology Alignment with Semantic Validation”, International Conference on tool with Artifical Intelligence, Vol 3,Page(s):326-332, August(2010).