An Improving for Ranking Ontologies Based on the Structure and Semantics

S.Anusuya, K.Muthukumaran

K.S.R College of Engineering

Abstract

Ontology specifies the concepts of a domain and their semantic relationships. It is widely used for solving the information heterogeneity problems on the web because of their capability to provide explicit meaning to the information. Different ontologies are created by developer for a same domain. To solve the heterogeneity problem among the specific domain several matching strategies are designed such as linguistic matcher, semantic similarity, structural comparison. Similarities among ontologies are obtained based on the term, concept relationship and terminological relationship as synonyms, hyponyms and homonyms. SWRL (Semantic Web Rule Language) rule is developed for common repository knowledge base to detect the homonyms. A merging framework combines the matching strategies for indentifying the similarities and dissimilarities of source ontologies then the similar concept are automatically merged and dissimilarities are directly merged in to the global ontology that resolve the synonyms and homonyms conflict among the domain specific ontologies.

Keywords: Ontology, Linguistic, Semantic similarities, Homonyms Conflict, SWRL, Merging.

1. Introduction

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 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.

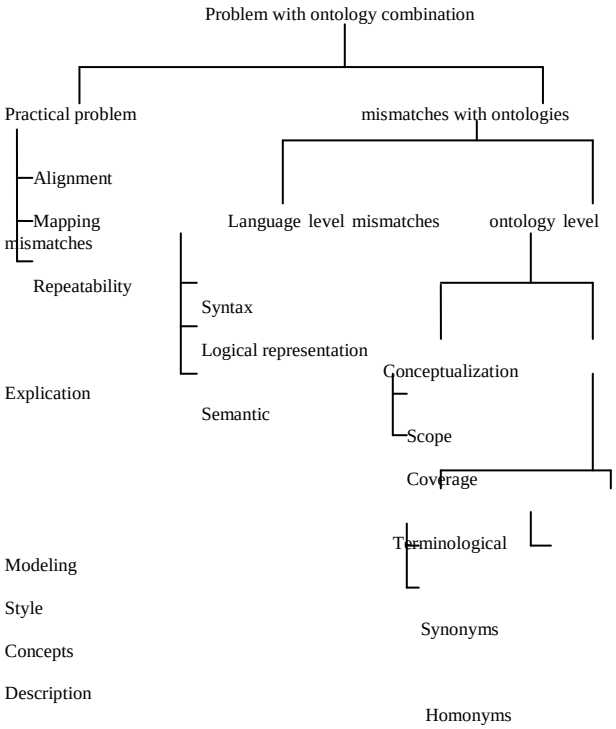

Figure1. Problem on ontology combination

Different types of mismatches may occur between different ontologies. The identification of these types of mismatches is essential in order to solve them during mapping, alignment, merging process. Ontology Merging is the process of generating a single coherent ontology 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 them meaning. SWRL is a standard OWL language to detect the similar rules and then cluster based on their similarity. SWRL atoms that govern the interaction between SWRL and OWL [11].SWRL rule based on by analysis the domain and range of object property and also analysis the same OWL classes and object properties.

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. 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 homonyms conflict among the domain specific ontologies. 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. The result of proposed framework is discussed and 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.

Siham Amrouch and Sihem Mostefai [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 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

Prasenjit Mitra and Gio Wiederhold [10] proposed linguistic similarities to match the terms. The source ontology object is designed in a different format. The articulation rules establish the semantic relationship among the ontology structure. The matcher uses the word similar table based on thesaurus and corpus methodologies to resolve the terminological heterogeneity.

Kamel Hussein Shafa'amri and Jalal Orner Atoum [4] proposed a multi matching framework to reduce the complexity of space and time. Three stages are involved in the framework to obtain the relationship type among the given ontologies of matched entities. System reads all sub and super classes of object then it matches the RDF statements and class hierarchies. Later assign matching relationship to the properties of object and data. The drawback of this framework is failed to find all possible entities of ontologies.

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. Heuristic similarity checking of properties are called to check the similar properties of classes.The process is repeated for every class of owlfile.

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 SWRL [11] rules 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. This architecture consists of four matching strategies a) Linguistic comparisons, b) Structure comparisons, c) Semantic comparisons, d) Homonyms detection to resolve the conflict among the ontologies.

The source ontology are of same domain are merged to generate the global ontology. Semantic inconsistency is resolved with the help of word net and knowledge base is used to solve the homonyms conflict between ontologies. The algorithm combines the matching strategies for identifying the similarities and dissimilarities of source ontologies then the similar concept are automatically merged.

3.1 Linguistic Comparison

The linguistic matcher finds the possible pairs of term from two ontologies. The similarity is computed based on the Jaro Winkler distance(1). Similarity score are assigned to each pair if it matches. If the similarity score is greater than the threshold then the similarity of each pair is determined.

$$DW=Djaro+ (l*0.1(1-Djaro))$$

Djaro is the Jaro distance for string s1, s2
L is the length of common prefix of the string up to a maximum of 4 characters.
p is a constant scaling factor 0.1

3.2 Semantic Comparison

The semantic comparison determines the similarity between concepts based on their terminological relationships such as synonyms and hyponyms. This approach requires the use of auxiliary sources, such as documents or annotations. The word net is a lexical database. The relation among words in the word net is synonymous called synset. Synonyms have a unique index and share its properties such as gloss definition. The semantic relationship is captured. Two ontologies are taken as input. Concept of ontology1 is similar with concept of ontology2 then both concepts are similar.

$$SIM\ SEM\ (c1, c2) = \frac{2* (synset\ (c1) \cap synset\ (c2))}{Synset\ (c1)) + synset\ (c2)}$$

3.3 Structural Comparison

The structural approach exploits relationships between concepts that appear together in a structure. Concepts and their relations are represented in a graph so that different kinds of structural related elements are identified for matching. Estimate the similarities between two concepts need to compare different kinds of their neighbor elements such as the parents, children or the leaves subsumed by them. It checks the relationships between the concepts of the level in the two ontologies and merges the similar concept.

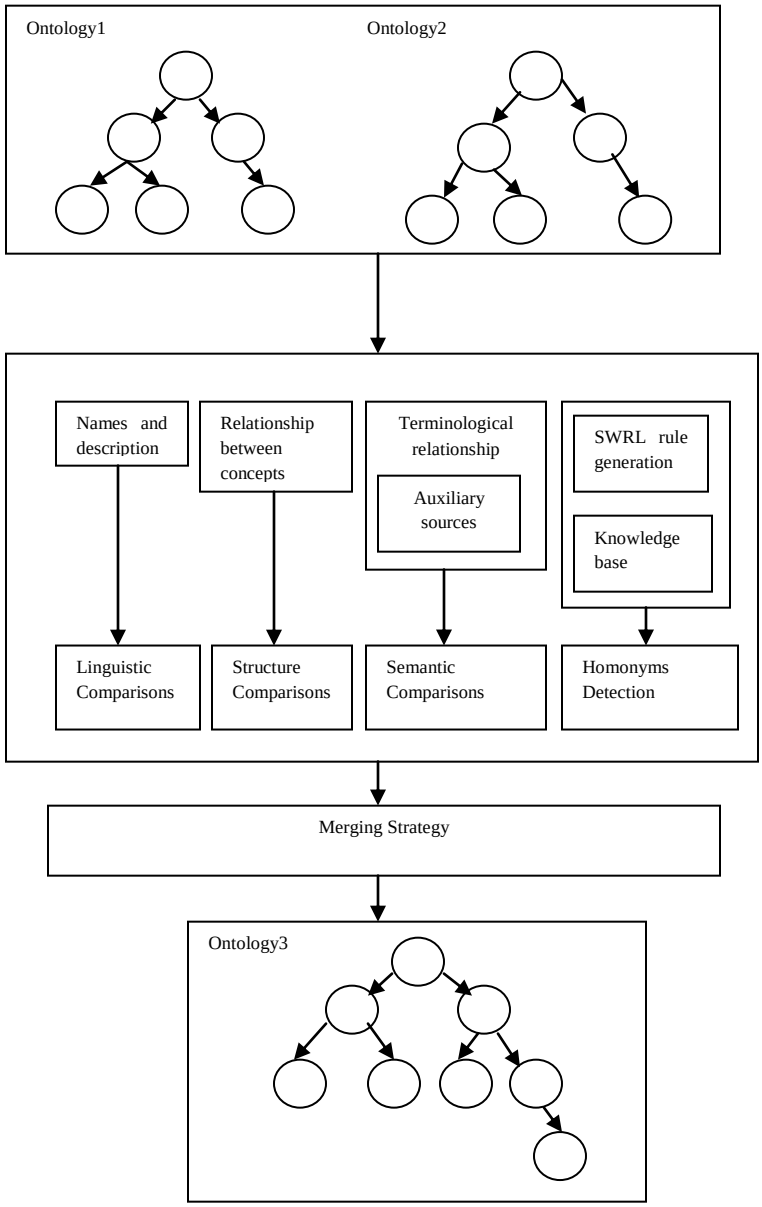

Figure2. Proposed Architecture

3.4 Homonyms Detection

Semantic Web Rule Language (SWRL) based on a combination of the OWL DL and OWL Lite. It includes a high level abstract syntax for horn- like rules. The abstract syntax contains sequence of axiom such as antecedent and consequent. Knowledge base is constructed using SWRL. Finding the similarity of class of given ontologies and then check the similarity of attributes and relationship among the attributes. By checking the similarity can detect the homonyms using SWRL.

Name (? x) ^ has Label (? X, Author) →Author (Name)

3.5 Merging Strategy

To create a common repository knowledge base using SWRL to avoid overlapping between existing ontologies for that using ontology merging and to detect the homonym. A merging framework combines the matching strategies for identifying the similarities and dissimilarities of source ontologies. When class are found similar through lexical and semantic matching are merging in to global ontology and dissimilar classes are added directly in to the global ontology. For merging two concepts need to specify a threshold. Similar concepts and properties with a similarity value higher than the threshold are merged recursively. Merging source ontologies initially helps to resolves the homonyms and synonym issues. A framework increases the accuracy of the search result.

4. Result

Domain specific ontologies are created by protégé tools.

Protégé is a free, open source ontology editor and knowledge-base framework. The Protégé platform support two main ways of modeling ontologies via the Protégé-Frames and Protégé-OWL editors. Protégé ontologies are exported into a variety of formats includes RDF(S), OWL, and XML Schema.

The BOOK domain ontology is considered for merging. The owl file1 contain classes such as Author has Name as subclass, Book, Publisher, Article. The owl file2 contain classes such as Author, Article, Publisher, Book has Name as subclass.

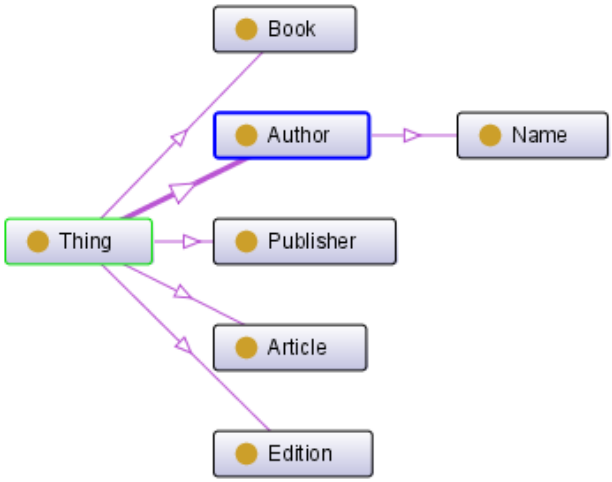

Figure3. Owl File 1

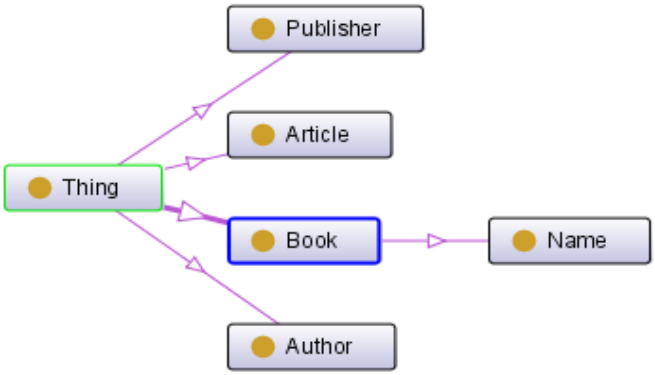

Figure4. Owl File 2

Similarities among the classes of ontologies are calculated by Jaro Winkler distance. The similarities are determined by threshold (0.6).some input of book domain specific ontologies are

$$DW(\text{AUTHOR}, \text{ARTICLE}) = 0.6371$$

AUTHOR and ARTICLE of two local ontologies are dissimilarities classes found to be similar in the linguistic matcher to overcome these problems semantic using word net is developed. The AUTHOR and ARTICLE classes having different Synset in word net that resolve the similarities among the classes and homonyms are detected by SWRL rule such as NAME of one owl file class and NAME of other owl file class are similar in the linguistic and semantic matcher but NAME of first owl file determine the subclass of AUTHOR class and the NAME of second owl file determine the subclass of BOOK class.

NAME (? x) ^ has Label (? X, AUTHOR) →AUTHOR (NAME)

NAME (? x) ^ has Label (? X, BOOK) →BOOK(NAME)

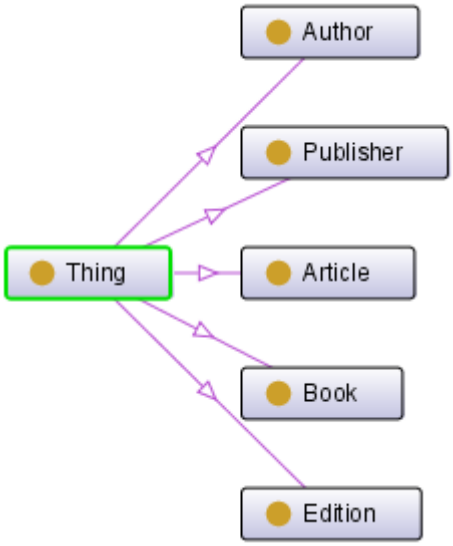

Figure5. Merged ontology

The global ontology is created by merging the similarities and dissimilarities are added directly in to the global ontology. The result increases the accuracy of the search result of the domain specific ontologies.

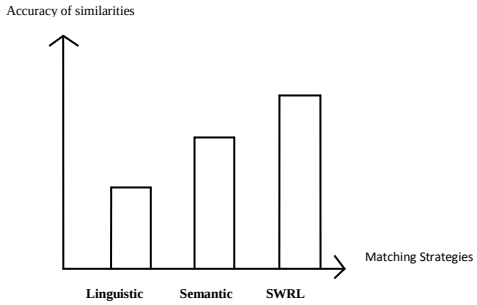

Fig 6 matching strategies accuracy result graph

Fig 6 shows the accuracy result of matching strategies Linguistic and semantic matching strategies resolve only the similarities among the ontologies with minimum accuracy of search result but SWRL resolve the heterogeneity problem completely such as Synonyms, homonyms which gives the accurate search result of domain specific ontologies among matching strategies

5. Conclusion

The method has been proposed to reduce the heterogeneous problem by providing a fully automated merged framework. In the proposed approach the domain specific global ontology is created by measuring the lexical, semantic and to detect homonyms conflict using set of SWRL rule in knowledge base. 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.Orgun, M.Dras, A.Nayak G.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 Adbul Qadir,"A Framework for Ontology Evaluation", pp, (2010).

[8]Natalya F.Noy, Dedorah L.McGuinness,"Ontology development: A guide to creating first ontology", International Conference, (2009).

[9]Nora Maiz, Muhammad Fahad ,Omar Boussaid, Fadlia Bentayeb, "Automatic Ontology Merging by Hierarchical Clustering and Inference Mechanisms" , proceeding of I- KNOW (2010),pp.81-93.

[10]Prasenjit Mitra, Gio Wiederhold, "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).