ISSN: 2321-8169
6660 - 6664

Ontological Framework for Semantic Web

Prof. Rashmi R. Janbandhu
Asst. Professor, IT
Rajiv Gandhi College of Engineering and Research, Nagpur, India
rashmi.janbandhu@gmail.com

Mr. Viplove Karhade
Scholar, Computer Science
North Eastern University
Boston, United States of America
viplovekarhade@gmail.com

Mr. Ritesh Selote
Director
Soft Skill Technology
Nagpur,India
Ritesh.Selote@gmail.com

Abstract— The web in today's era is an remarkable triumph due to the proliferating number of users and sources of information. However, the increasing complexity of the web is not reflected in the present-day state of Web technology. The load of gaining access to, extracting relevant information, interpretation and upholding is left to the human (user end).The Semantic Web should structure the content of web pages so that information is also structured in well-defined manner. Ontology structuring techniques is good choice to structure and model problem. It provides proper conceptualization of a domain that is communal to group of people. Ontology is basically representation of domain based theory or information in view of their semantics. Thus ontology is basically a structure or framework which underlies and enables machine understanding. The paper discusses ontology in context of semantic web.

Keywords- *Ontology; Semantic Web; Import; Extraction; Pruning; refinement; concept; Web Ontology Language; (OWL); Relational Description Framework (RDF); Extensible Markup Language (XML)*

*****

I. INTRODUCTION

"Ontology" is a deep-thinking discipline, a branch of philosophy that compacts with the nature and the organization of being. The term "Ontology" has been coined by Aristotle in Metaphysics, IV, 1. In the context of research on "Ontology", philosophers try to answer questions "what being is?" and "what are the features common to all beings?". Parmenides was the first to propose Ontology.

The novel idea dates back to Greek Philosophy: it is learning of what exists in the world. The next issue it handles is how the existence of a thing is influenced or is related to the existence of other things around it in the world .i.e. the relations between things. Ontology studies:

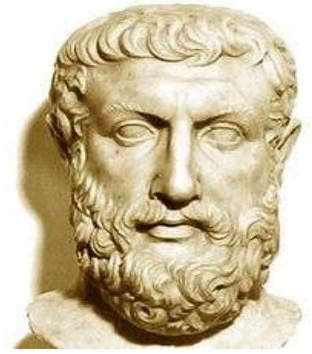

Figure.1. Parmenides

- How does a thing persist over time?
- Which category does the thing in existence belong or fit into?
- And lastly, What characteristics help us to classify a thing in a particular group or category? It also studies which characteristics change and the effect of it on classification?

The above questions are answered in allied branches of science such as in Metaphysics like Epistemology and Semantics.

Semantic Web basically or wholly deals with meaning. The word Semantics refers to synonyms, meaning etc. It can be collection of URL with similar meanings etc. Ontology can help to improve this collection by adding a parameter of concept to the meaning. It is useful when a single word has multiple meaning. The precision or accuracy lies in considering the correct meaning according to problem statement.

Ontology refers to classification of everything. Some things while being in existence, undergo change, these changes are also reflected in the ontology. It is framing an hierarchical architecture which is easy to store and access the information being framed. Over decades much work and research has been already done on ontology but implementing it in actuality is a tedious task which is being done at slow pace.

II. MEANING TRIANGLE

The previous work done in the field unnoticed the inescapable margins of communicating meaning via language and the ambiguities created by understood exchange of different "senses" of meaning. Understanding such ambiguities was a topic considered by Frege. Frege introduced a peculiarity of two types of meaning: the notion and the referent. The graphical elucidation of this peculiarity is referred to as the meaning triangle (Figure 2). The meaning triangle outlines the dealings between symbols or words, thoughts or ideas or notion or concept and things of the real world.

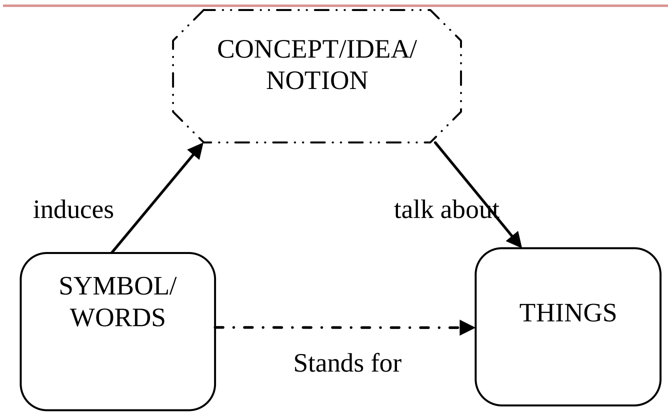

Figure 2. Meaning triangle

The diagram aims to illustrate that the word alone cannot give the exact meaning of the word. The word needs a concept or idea and thing to put forward the correct meaning. Thus, Concept, Symbol (word) and thing are in association. The association is indirect. The interpreter i.e the user who is willing to understand the word, processes the word in accordance of the concept and thing in being and then deduces the correct meaning of the word.

III. Ontology Learning Architecture

The components of the Ontology learning framework are designed as per requirement. Four main components have been devised:

A. *Ontology Engineering and Management Environment:*

Engineering ontology means that developing a architecture will be capable of storing the ontology. It also deals with framing a environment to store and access information stored in ontological conception. Management environment deals with designing proper user interface which complements the Ontology Engineering at back end.

B. *Data Importing and processing Unit:*

This unit contains techniques for discovering the data, importing the data, retrieving the data, transmuting the data to produce more relevant output. The unit contains various numbers of techniques for each task done in this unit for better results.

The important sub module of this unit is natural language processing system. The other sub module is ontology wrapper and importer; this unit offers a framework which inputs the data in the same ontological structure the system is working in. Next module is ontology merging; many a times there is situation when two ontologies present in the system should be merged into single ontology. In this scenario ontology merging module plays a vital role.

The next module is Document Wrapper this component offers particular processing techniques for wrapping the data into an algorithm-specific representation.

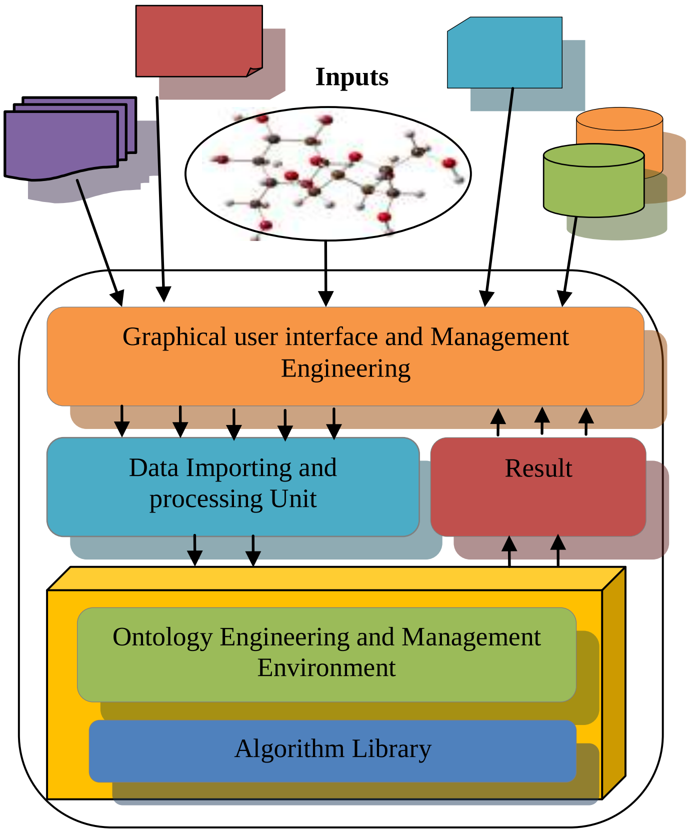

Figure 3. Components of Ontology Learning Architecture

C. *Algorithm Library:*

This unit contains many ontological based algorithms. This unit is as good as the back bone of the architecture. As many algorithms are implemented in this unit the problem that arises is standardization of input and output. A standard way has to be devised to take the input and give the output.

All the algorithms are implemented in such a way that the user can use it with minimum knowledge of how algorithm works. The algorithms which are implemented in the Ontology Learning architecture can be classified in two categories:

1) *Ontology Extraction*: The algorithms which work on extracting the associations and relation between the object, subjects and words in the input and thus create ontology come under this category of algorithms.
2) *Ontology Maintenance*: This algorithms aims for reducing ontology and prune it. The refinement and betterment of the existing algorithm.

D. *Graphical user interface and Management Engineering:*

This unit deals with designing a user friendly environment which lets the user enter the information in familiar pattern and retrieve the same information as and when required.

The graphical user interface and Management Engineering unit accomplish many tasks:

- It helps the ontology engineer to select the relevant data.
- It helps to input the correct and needed format of the data for underlying on ontology algorithms.

- The results in the user expected format can be displayed using this module.

IV. PHASES OF ONTOLOGY LEARNING

The ontology learning has four phases. The ontology Learning framework is process oriented. Each phase of ontology learning uses processed relevant data. The ontology learning process the data using modules or units described in the Ontology Learning Architecture above. The phases of Ontology are:

i. *Import*: This is the first phase of Ontology Learning. Importing input to process it into ontology is the first phase. It also provides mechanism to import a ontology itself.
ii. *Extraction:* This is the second phase of Ontology Learning. The data coming as input is processed to extract ontology from it in this second phase.
iii. *Pruning:* This is the third phase of Ontology Learning. Pruning is cutting short ontology to keep the ontology in size for the application.
iv. *Refinement:* This is the fourth phase of Ontology Learning. It deals with refining the fine granulites in the ontology.

The above process can be followed to produce any new ontology as and when needed or required. It should be noted that each phase is individual of each other. Therefore, any phase can be skipped if not needed, for example, if ontology is imported it can skip the second phase directly be transferred to pruning.

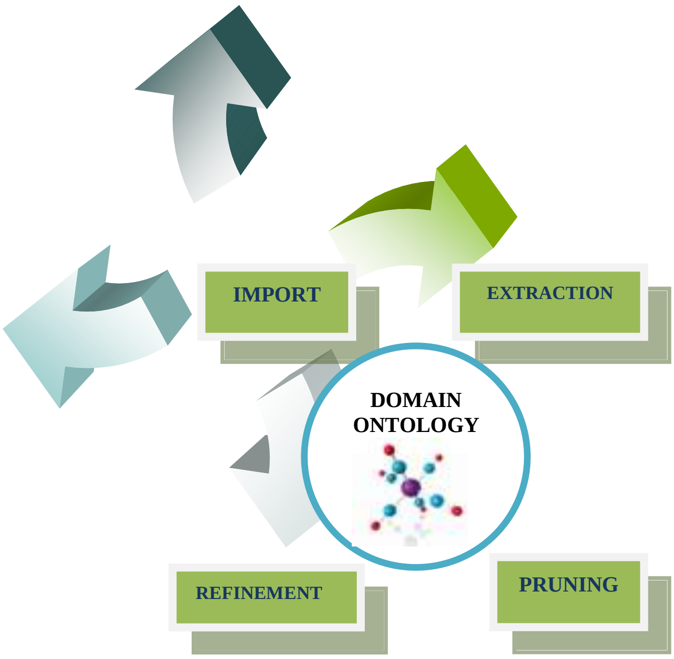

Figure 4. Phases of Ontology Learning

and in this time many domain developed their ontology. The architecture of ontology should have a mechanism to import these ontology; import and reuse is the mechanism developed to do so. Import and refuse are roughly separated in two different divisions:

1) The first parts deals with transforming one ontology framework to other. This transformation is also addressed as ontology wrapper. One has to convert the particular ontology into a representation that was inculcated in the ontology learning framework.

2) The second part of import and reuse is merging of ontology. It is sometimes required that two already existing ontologies should be merged to create a new ontology according to need. This need is facilitated using the second part.

Researchers are developing algorithms for ontology wrapping techniques as well as merging of ontology.

A. Extract

In the extract phase the major work done is to make the most of input data to bring out relevant data for framing up the ontology from it. The process of extraction is iteratively growing model. The process stars with seed information and iterates over the information. It updates the ontology in its iteration. Various Algorithms for ontology Extraction are

1) Lexical Entry Extraction: Lexical Entries are done with accordance of words and terms and their relationships.

2) Taxonomy Extraction: Various taxonomical methods of extraction are used under this arena. Some of them are hierarchical clustering , pattern based clustering mechanisms.

3) Non-Taxonomy Extraction: Non taxonomical techniques such as core learning algorithm, Relation Extraction including Background Knowledge etc.

B. Pruning:

Ontology pruning becomes necessary if large ontologies are imported into the system. There are two strategies of pruning

a. *Baseline Pruning*: Baseline pruning adopts the idea of term extraction in an inverse sense.

b. *Relative Pruning:* The ideas of the relative pruning algorithm are that domain relevance is considered as the relative frequency of given ontological entities with respect to frequencies obtained from a general corpus.

C. Ontology Refinement:

Refining is as good as extracting. The difference between these two cannot be shown distinctly other than the difference that ontology refinement works more on already existing ontology, associations and relation while as the extraction works more on scratch raw information in order to develop new ontology.

Ontology Refinement is a incremental process. It keeps increamenting as and when changes occur in relations and associations with time or circumstances

V. Web Data Semantics

A. *Semantic Web:*

The Semantic web is aiming to build a web which is processed by both, humans and machines. The data in semantic web is in machine processable format. The need for semantic web has increased because the data has increased tremendously in size, and it has become unmanageable by human so systems should be developed which themselves are capable of understanding the data being uploaded in it. Such intelligent systems are thus coded with modules such as back end knowledge base, importing, analyzing, filtering, classifying etc.

The semantic web has knowledge base as a very important module. It is the knowledge base which helps the machine to process the machine processable data; analyze it, process it, classify it, if required filter and produce user desired results. Thus semantic web is cable of performing decision making. The data in the back end should be ontologically stored which makes it more manageable for the system to perform intelligently.

B. *Ontology Language for the semantic Web:*

The Ontology Languages for Semantic web include Web Ontology Language (OWL), Relational Description Framework (RDF), Extensible Markup Language (XML), etc. Extensible markup language helps to format the data in the machine processable format. XML provides the syntactical formats and peculiarities but the semantically XML has no contributions. Thus to introduce the semantic contribution to the data RDF is constructive. It stores the data in subject predicate object triplet. The triplet id written as *P(S,O)* and the graphical rep[resentaion of the above triplet is shoen in fig5. below:

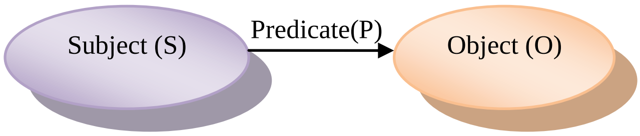

Figure 5. RDF Subject Predicate Object

Example of the above triplet can be as in fig 6.:

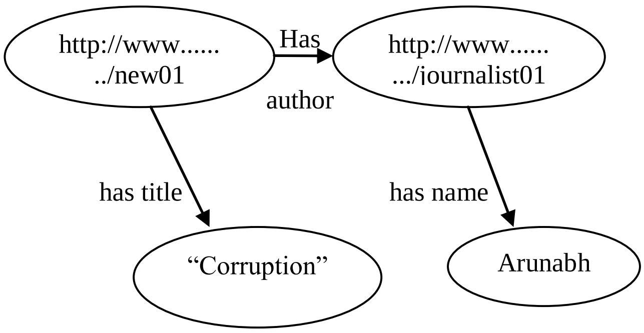

Figure 6. RDF Subject Predicate Object Example

The OWL language covers more semantic details as compared to RDF. OWL adds semantics such as vocabularies, properties classes etc. which are not supported by RDF.

Conclusion

The semantic web based on ontology can assist human to handle the overgrowing size of web. All it needs is to frame ontological data at the backend with the help of architecture and process of ontology learning described in the paper.

Acknowledgment

The paper contains all the knowledgeable content best known to the authors.

References

[1] Jim Blythe, Jihie Kim, Surya Ramachandran and Yolanda Gil." An Integrated Environment for Knowledge Acquisition". Proceedings of the International Conference on Intelligent User Interfaces 2001.

[2] Natalya F. Noy and Deborah L. Mc.Guinness, "Ontology Development: A guide to creating your first ontology", Stanford University, Stanford, CA, 94305.

[3] Bill Swartout, Ramesh Patil, Kevin Knight and Tom Russ. "Toward Distributed Use of Large-Scale Ontologies". *Proceedings of the Tenth Knowledge Acquisition for Knowledge-Based Systems Workshop*, November 9-14, 1996.

[4] Michael Bada, Robert Stevens, Carole Goble, Yolanda Gil, Michael Ashburner, Judith A. Blake, J. Michael Cherry, Midori Harris, Suzanna Lewis," A Short Study on the Success of the Gene Ontology", Vol 1, No 2 (2004): Special Issue: 2003 World Wide Web Conference

[5] Aditya Kalyanpur, Bijan Parsia, Evren Sirin, Bernardo Cuenca Grau, James Hendler, "Swoop: A Web Ontology Editing Browser", Vol 4, No 2 (2006): Special Issue: Semantic Grid - The Convergence of Technologies.

[6] Michael Compton, Payam Barnaghi, Luis Bermudez, Raul Garcıa-Castro, Oscar Corcho, Simon Cox, John Graybeal, Manfred Hauswirth, Cory Henson, Arthur Herzog, Vincent Huang, Krzysztof Janowicz, W. David Kelsey, Danh Le Phuoc, Laurent Lefort, Myriam Leggieri, Holger Neuhaus, Andriy Nikolov, Kevin Page, Alexandre Passant, Amit Sheth, Kerry Taylor, “The SSN Ontology of the W3C Semantic Sensor Network Incubator Group”, Vol 17 (2012).

[7] http://ontobroker.semanticweb.org/ontos/swrc.

[8] http://www.ontology.org

[9] http://www.w3.org/XML/1998/06/xmlspec-report

[10] Silvio Peroni, David Shotton, “FaBiO and CiTO: ontologies for describing bibliographic resources and citations”, Vol 17 (2012)