Methodology Research of Ontology Building in Semantic Web

Gang Wang

Computer Software and Engineering, Jiangsu University of Science and Technology

Zhenjiang, Jiangsu, 212003, China

E-mail: wanggang3275@163.com

Abstract

This paper describes two concepts on the Semantic Web and Ontology and points out the core role of Ontology in Semantic Web. Moreover, the writer gives a Double-Channel Helix Methodology and expounds its advantages at the structure and the task-flow.

Keywords: the Semantic Web, Ontology, Double-Channel Helix Methodology

1. The Semantic Web

The Web is a system interlinking lots of documents via the Internet. In the Web, people can read lots of books, watch the online TV, publish contents in blogs and so on.

The Semantic Web is a web of data and is about two things (Ivan Herman et al., 2010). One is common formats for integration and combination of data drawn from diverse sources. Another is about language for recording how the data relates to real world objects. In simple terms it provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries. Fig. 1 shows the Semantic Web layer cake. The philosophy was thought up by Berners-Lee.

Where, the Table 1 describes the specifications each layer.

2. Ontology

Ontology as a branch of philosophy is the science of what is, of the kinds and structures of the objects, properties and relations in every area of reality(Smith Barry et al., 2003, 155-166). Each scientific field will of course have its own preferred ontology, defined by the field's vocabulary and by the canonical formulations of its theories. Ontologies play an important role in fulfilling semantic interoperability as described in the seminal article on the Semantic Web (Berners-Lee et al., 2001, pp.35-43).

Why Ontology is so essential in the Semantic Web? The first reason is that the search for XML, RDF and RDFS have being rather ripper, which provided the language basis. Another reason is that the upper projects, researched on rules, logic framework, proof and trust, could not get improving smoothly and quickly without the support of Building Ontology. However, it is a hard work and needs a lot of research by scholars and engineers. In all the difficulties, what is the most important?

The answer is how to build a available Ontology. Meantime, it must to meet five rules, veracity and objectivity, integrity, consistency, extensibility and the least restriction (Peng Bo, 2009, pp.2610-2611).

3. Double-Channel Helix Methodology

Nowadays, there are many methodologies of Ontologies building, e.g., methodology by Ushold and King for Enterprise Ontology, METHONTOLOGY, On-To-Knowledge Methodology and so on (AIAI et al., 2003; Fermandez-Lopez et al., 1999, pp.37-46; Staab et al., 2001, pp.26-34). But un-commonality is the important limitation. Fig.2 shows a Double-Channel Helix Methodology, a better methodology. Its process of Ontology building is step-and-step. It has two "channel", each of which can make the process be shifted backward or forward. Considering from the engineering theory of ontologies building, the methodology is satisfied to people's acknowledge and logic thinking.

3.1 Semantic Web knowledge

In this layer, the Semantic Web, next generation of the World Wide Web, will consist of data defined and linked in such a way that it can be used for effective discovery, automation, integration, and reuse across various applications (S. Bechhofer et al., 1999, pp.33-36). Moreover, the distributed and dynamic character of the Semantic will cause that many versions and variants of Ontology knowledge will arise. To build a real Semantic Web, it is essential that the knowledge that is represented in the different versions of ontologies is *interoperable*.

3.2 Conception Extraction

The extraction of conception in the Semantic Web is not an easy task. It requires skills and is still an art rather

technology. As below, it shows a method.

Firstly, we have to gather all conception in brain storming, by clustering collected and by investigating refined. Though this step, decide the scope defined for conceptions in the Semantic Web. Secondly, make sure the attribution of every conception. Select a natural word which one meaning for each concept. If there is no accessible word for representing a concept, then create a new one. Thirdly, sum up the relation between the different conceptions. Organize these concepts in an *is-a* hierarchy.

In practice, these three steps are done not in a waterfall manner. Users can go back and forth during the process.

3.3 Tools Chosen

Sincerely, many of tools has been developed for enterprises, which include mainly Ontolingua, Ontosaurus, WebOnto, WebODE, OntoEdit, OILEd, Protégé and so on. All tools are assistant well for Ontology to edit, modify, explore and maintain. In Table 2 (Wang Chang-xia et al., 2009, pp.26-28, 31) , it shows the scheme of key identification among seven tool. Users should select an available tool for Ontology building and even that choose the different tools for different stage to finish task better.

3.4 Ontology Building

Edit the coding about the conceptions, attributes and relations using the basic elements, e.g. *Class, Property, subClassOf, subPropertyOf* and so on. Finally, provide main interface for Ontology merging, Ontology mapping, Ontology translation resolution in the next stage.

3.5 Ontology Evaluation

Same to software engineering, Ontology needs also to evaluate. However, there is no authoritative software to detect new Ontology building. Recently, it is evaluated by extensibility, visibility and inferenceability and so on. Finally, if new Ontology building does not meet standards, the process will be shifted backward continuing loop. Otherwise, the process of Ontology building ends.

4. Conclusion

After analyzing the limitations of traditional methodology on Ontology building, the author give a better Double-Channel Helix model. Of course, there are something worthy to refine in the further task.

References

AIAI, IBM, Lloyd's Register, Logica UK Limited & Unilever. (2003). The Enterprise Ontology. [Online] Available: http://www.aiai.ed.ac.uk/project/enterprise/enterprise/ontology.html (Mar. 15, 2010).

Berners-Lee T., Hendler J. & Lassila O. (2001). The Semantic Web. *Scientific American 2001*. 284(5):35-43.

Fermandez-Lopez, M., A.Gomez-Perez & J. Pazos Sierr. (1999). Building a Chemical Ontology Using Methodology and the Ontology Design Environment. *IEEE Intelligent Systems*, 14(1), 37-46.

Ivan Herman, Harry Halpin, Sandro Hawke, Eric Prud'hommeaux, Dave Raggett & Ralph Swick (2010). W3C Semantic Web Activity. [Online] Available: http://www.w3.org/2001/sw/ (Mar. 15, 2010).

Peng Bo. (2009). Research about Ontology Building Method (2009). *Compute Knowledge and Technology*. 5(10), 2610-2611.

S. Bechhofer, I. Horrocks, P.F. Patel-Schneider & S. Tessaris. (1999). A proposal for a description logic *Proceedings of the International workshop on Description Logics (DL'99)*, pp.33-36, Linkoping, Sweden, July 30 –Aug. 1,1999.

Smith Barry & L. Floridi. (2003). How to Build an Ontology. *Blackwell Guide to the Philosophy of Computing and Information.* pp.155-166. [Online] Available: http://ontology.buffalo.edu/smith/ articles/ontologies.htm. (Mar. 15, 2010).

Staab. S. H., P. Schunurr, R. Studer & Y. Sure. (2001). Knowledge processes and ontologies. IEEE Intelligent Systems. *Special Issue on Knowledge Management*, 16(1), (2001)26-34.

Wang Chang-xia, Li Guan-yu & Chen Bu-wei. (2009). Study on Present Situation and Development Trend of Ontology Construction Tools. *Computer and Modernization*, 2009, 167(7):26-28, 31.

Table 1. The description of the Semantic Web Layer Cake

Layer	Description
URI Uniform Resource Identifiers	This layer makes sure of providing the uniform identifiers of lots of resource. It likes the strings starting with "http:" or "ftp:" that often find on the World Wide Web.
Unicode	This layer provides the uniform standards for all kinds of languages in the world to coding characters.
XML Extensible Markup Language	XML replaces HTML for its individual advantages. In addition, it gives a definition on the methods describing data.
Namespace	This layer provides many ways to differentiate names, so the resources which have same names and different means are still used.
RDF M&S Resource Description Framework Model and syntax	Here reserve specifications refer to model and syntax of RDF.
RDF Schema Resource Description Framework Schema	It is a kind of languages for describing RDF vocabularies and has some basic elements, e.g. *Resource, Class, Property, subClassOf, subPropertyOf, range, domain.*
Ontology	It is a formal, explicit specification of a shared conceptualization. It need to make sure all knowledge can be known by people and computer.
Rules	Many principles or regulations between the upper and the lower in the level structure;
Logic Framework	It provides a logic inferenceability on the knowledge of Ontology describing.
Proof	On being logic inferenceability, give a proof on whether a statement is right or wrong.
Trust	Detect whether the web information is trusted or not.
Signature	A method to make sure the environment more safety
Encryption	A method to make the environment more safety

Table 2. The comparison on the Ontology building tools

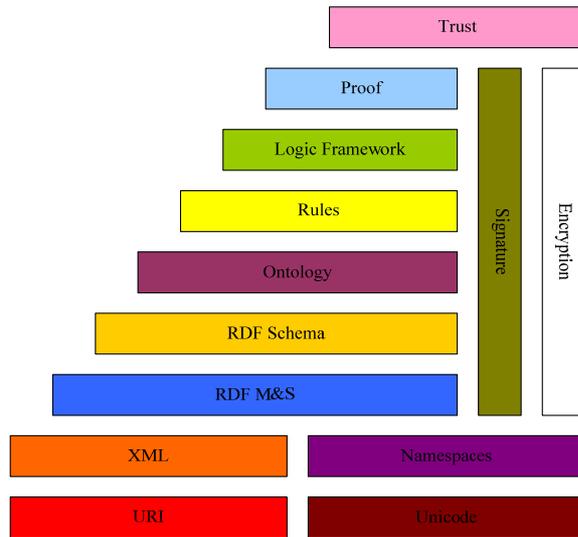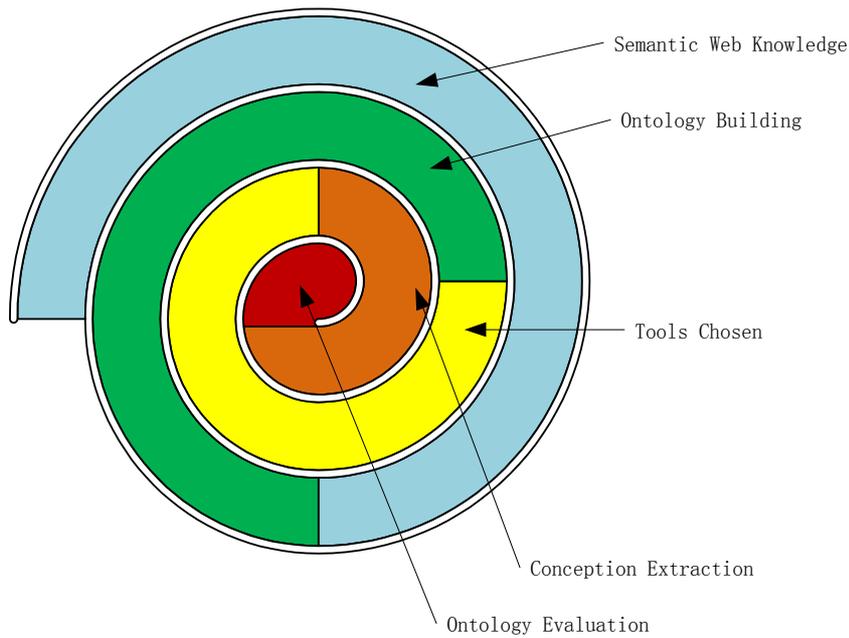| Tools Name	Ontoligus	Ontosaurus	WebOnto	WebODE	OntoEdit	OILED	Protégé
OWL Language	×	×	×	√	×	√	√
Chinese Supportion	×	×	×	×	×	×	√
Visiblity	×	×	√	√	×	×	√
Network Technique	√	√	√	√	√	×	×
Cooperation Development	√	√	√	√	√	×	×
Ontology Merging	√	×	×	√	√	×	√
Fuzzy Ontology	×	×	×	×	×	×	×

Figure 1. The Semantic Web Layer Cake

Figure 2. Double-Channel Helix Methodology