ISSN : 0974 - 7435 Volume 8 Issue 3

BioTechnology
An Indian Journal

FULL PAPER

BTAIJ, 8(3), 2013 [298-305]

Research on modular ontology modeling based on domain vocabulary

Liu Ziyu[1]*, Zhang Rong[2], Qu Ying[1], Song Wei[1], Liu Bin[1]
[1]School of Economics and Management, Hebei University of Science and Technology, ShiJiaZhuangHeBei 050018, (CHINA)
[2]School of Foreign Language,Hebei University of Science and Technology, ShiJiaZhuangHeBei 050018, (CHINA)
E-mail: purpleyuliu@sohu.com

ABSTRACT

More and more scholars begin to pay attention to modular ontology, and it can solve the problems of the construction of large ontology, ontology reuse and the efficiency of reasoning. Based on the research of construction and connection method of modular ontology and the characteristics of domain vocabulary, this paper presents the structure and the formal definition of modular domain ontology, and puts forward a method of modular ontology modeling based on domain vocabulary. This paper details six parts of the method: determine the range of domain vocabulary, the division of ontology modules, ontology modules modeling, extraction the connection between concepts of ontology modulars, connect the ontology modules and generate the domain ontology and consistency test. Under the guidance of this method, an example of the modular ontology of high-speed railway domain is developed to verify the feasibility of the modeling steps, in which using owl as the ontology description language and using the semantic extension method of owl:import to connect ontology modulars.

KEYWORDS

Modular ontology;
Domain vocabulary;
Ontology modular;
High-speed railway.

INTRODUCTION

Since the importance of ontology in knowledge organization and knowledge management has been noticed in domains, and domain ontology can effectively organize knowledge, share knowledge and reuse it better in the domain. Domains such as the high-speed railway, logistics, agriculture and other Industries are generally composed of multiple sub-domains. For instance, the high-speed domain consists of sub-domains such as EMU, operations management, safety and rescue, traction power supply, and public works. In such cases, it is a complex and huge project to construct domain ontology once and for all.

More and more scholars begin to focus on modular ontology[1], and it can solve the problems of the construction of large ontology, ontology reuse and the efficiency of reasoning given the significant advantage of modularization and its effect in application. So if the domain is decomposed into a plurality of modules, which are then connected according to the relationship between them to generate the domain ontology model, we can solve the problems of complexity of ontology construction. In addition, in most cases different sub domain experts are only interested in their particular field of knowledge. Thus it is easier to construct the sub module domain ontology separately. In the use of domain ontology, sub module domain ontology can be used alone, while the connected domain ontology can be used for the sake of multiple domain knowledge. In this way, sub domain ontology can be maintained separately, and the huge ontology's bid problems can be solved such as difficult to maintain and reuse, and reasoning complexity.

Because most domains composed of multiple sub domains, have their own vocabulary such as dictionaries, thesauri, classification etc., the construction of ontology combined with ontology engineering on the basis of thesauri is one of the research hotspots of domain ontology modeling. At present, the modular ontology modeling involves little modeling process in detail, and the domain vocabulary has its own characteristics, we should choose the right connection method and connection language for ontology module according to the characteristics of domain vocabulary. Based on the research of construction and connection method of modular ontology and the characteristics of domain vocabulary, this paper presents the structure and the formal definition of modular domain ontology, and puts forward a method of modular ontology modeling based on domain vocabulary. Under the guidance of this method, an example of the modular ontology of high-speed railway domain is developed to verify the feasibility of the modeling steps. We determine the range of domain vocabulary and divide it into modules. Then we build ontology models for modules and connect these ontology modules using the semantic relationship between concepts in ontology modules.

REVIEW OF RESEARCH ON MODULAR ONTOLOGY

What's the so-called modular ontology is that the complex domain to be modeled is top-down decomposed into a plurality of interdependent and interconnected sub domains in accordance with the appropriate principles. In general, every sub domain only pays attention to concepts of one important aspect of the object domain; then, we construct the model separately for each sub domain and form the so-called ontology module; finally, each module will be assembled through a certain method and eventually form domain ontology. We make a following review of Modular ontology modeling and modular ontology language.

Modeling of modular ontology

In the modeling of Modular ontology, as a kind of knowledge organization and representation, modular ontology provides a new way for collaboration, integration, reuse of knowledge. B. Guenca Grau[2] thinks ontology modularization is similar to the module of software engineering, and a large-scale ontology can be decomposed into simpler ontology modules. H. Stuckenschmidt[3] thinks a modular ontology is the sets of ontology modules and these sets are connected by the external definitions, which are all contained in the modular system. Moreover, the modularization not only makes the construction and maintenance of the ontology easier, but also can promote the reuse of knowledge. A lot of languages and construction tools of the modular ontology have been proposed, such as P-DL[4],ε-connections[5], Swoop, Protégé, ProSé[6],etc.

There are also many domestic scholars make modular ontology research in progress. Jiang Cuiqing[7] puts forward the knowledge organization model of mechanical product manufacturing enterprises based on the Modular Ontology. They divide knowledge ontology into 4 modules and use assertions to connect the modules. Lin Songtao discusses the related theory problems of construction of modular ontology in his doctoral dissertation[8]. Zhang Weiyi, Lu Ruzhan[9] study the communication mode and structure of each module through the ontology instances of automobile driving training domain.

Language of modular ontology

The formal description language of ontology is the basis of ontology research. Researchers have proposed a variety of languages of modular ontology from the different application scenarios. The distributed description logic DDL[10] is proposed by Alexander Borgida in 2002. C-OWL[11] using the bridge rules is proposed by Paolo Bouquet in 2003. And ε-connections[12] is proposed by Oliver Kutz in 2003 which through extending the attribute to combinate the ontology module. In view of the problems of not supporting contain and transfer reuse between modules in DDL0ε-connections0C-OWL, Jie Bao puts forward P-DL[13] and extended owl: import[14]. They use the selective import mechanism to integrate the ontology module so that the ontology module can partially overlap, reduce the strictly disjoin demand and provide partial reuse of the ontology module. In addition, there are also many literatures discussing the method of the semantic description and integration[15] of the modular ontology, but they basically are based on P-DL[13] and extended owl: import[14]. At the same time, C-OWL[16] is a modular ontology language which increase DDL formal and extend OWL. Although it has strong ability of semantic expression, there is no reason to support it.

In summary, in the present study of modular ontology, the process of modeling of modular ontology is less involved and not aimed at the modeling methods of modular ontology based on domain vocabulary. In the choice of the modular ontology language, ε-connections and extended owl: import are used more. ε-connections express the dependence by defining and using the link property and the link property is a binary relation. Its first element is the instance of class in source ontology and the second element is the individual in the target ontology. Because the ontology module based on vocabulary is mainly connected by the synonymous, inclusion, ancestors, brotherhood, user-defined relation between the module concepts, that is, the binary relation connects the conceptions between two ontology modules, so we choose the connection of ontology module through the method of extended owl: import proposed by Jie Bao and Vasant Honavar.

MODELING METHOD OF MODULAR ONTOLOGY BASED ON DOMAIN VOCABULARY

Structure of modular domain ontology

The connection between the ontology modules based on the domain vocabulary is main associated by the relations between different ontology modules. In the structure of modular domain ontology, the ontology can be divided into three levels: domain ontology layer, concept relation ontology layer and ontology module layer, as shown in Figure 1.

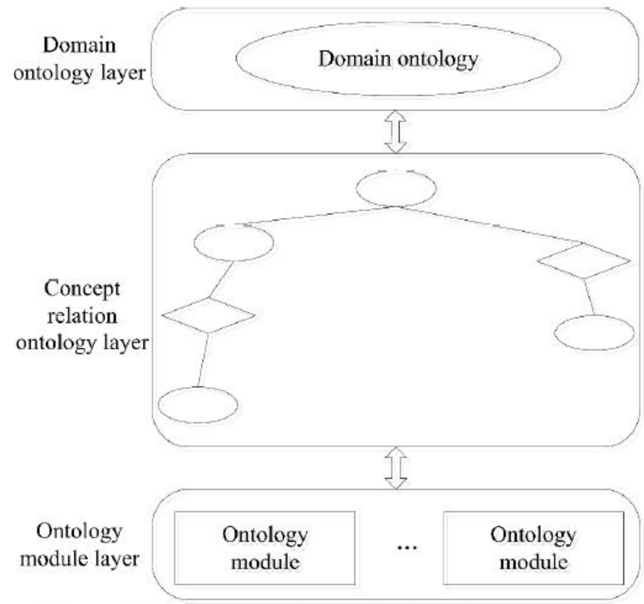

Figure 1 : The structure of modular domain ontology

Domain ontology layer, which defines a global ontology, is unified domain ontology that generated by the concept relation ontology layer. Although domain ontology layer is the integrated top domain ontology and can be used to be access to the ontology module layer, it can be regarded as virtual presence, dynamically generated by the concept relation ontology layer. And it can change with the change of the concept relation ontology layer.

The concept relation ontology layer consists of the concept in several ontology modules and their relations which mainly include synonymy, inclusion, ancestors, brothers and user-defined relation.

The ontology module layer is composed of a plurality of ontology module and an ontology module corresponds to a module. The module division is performed according to the characteristics of the domain vocabulary: firstly the domain vocabulary is divided into several sub domain ontology modules and then the ontology module is generated by ontology modeling of the sub domain.

Related definitions of modular domain ontology

Ontology module as the representation of the conceptualized knowledge, describes concepts, relations between concepts and the rules establishing concepts and their relations. We define it as follows:

Definition 1 sub domain ontology module. Sub domain ontology module is a formal specification of a shared conceptualization. We will represent sub domain ontology for a eight tuple:

$SDO=\{C, P, H^c, R^s, R^{ud}, I, F, A\}$, thereinto:

SDO says domain ontology module, C says the concept (or classification) setÿP for type attribute set of Data type in domain ontology, H^c for subclass-of binary relations collection between types, R^s for synonymy relations collection between types, R^{ud} for user-defined relations collection between types, that is Object Property of types (including part-of relation which also use user-defined relations to describe collection), I represents a collection of domain concept instances, F represents a special relation between concepts, which can be expressed in the following form: $c_1 \times c_2 \times ... \times c_{n-1} \rightarrow c_n$, A is some tautology representing for axioms of concepts and their relations.

The sub domain ontology modules constructed on the basis of vocabulary, connect through various relations between concepts and form the whole domain ontology model. The connection between the sub domain ontology modules is called the layer of concept relationship ontology and we can define it as follows:

Definition 2 relation model between two sub domain ontology modules. Given two sub domain ontology modules SDO_i and SDO_j, relation model of the connecting SDO_i and SDO_j can be expressed as $R_{SDOi,SDOj}=\{SDO_i, SDO_j, SDOH^c, SDOR^s, SDOR^{ud}\}$.

thereinto, SDO_i and SDO_j represent the names of two sub domain ontologyÿ$SDOH^c$ represents subclass-of binary relations collection between two sub domain ontology conceptsÿ$SDOR^s$ represents synonymy relations collection between two sub domain ontology concepts, $SDOR^{ud}$ represents user-defined relations collection between two sub domain ontology concepts.

Definition 3 domain ontology model. Domain ontology model consists of definition 1 and 2, and can be expressed as:

$$DO = \{\bigcup_{i=1}^{n} SDO_i, \bigcup_{(i,j=1)\&(i<j)}^{n} R_{SDO_i,SDO_j}\}$$

Modeling method of modular ontology based on domain vocabulary

We use the domain vocabulary such as category and subject thesaurus as the base of the ontology construction. In the division of modular ontology, we need determine the partition principle. In the connection method of ontology modules, extract the ontology module concept and connect it with other ontology module concepts. The steps of modular ontology modeling as show in figure 2.

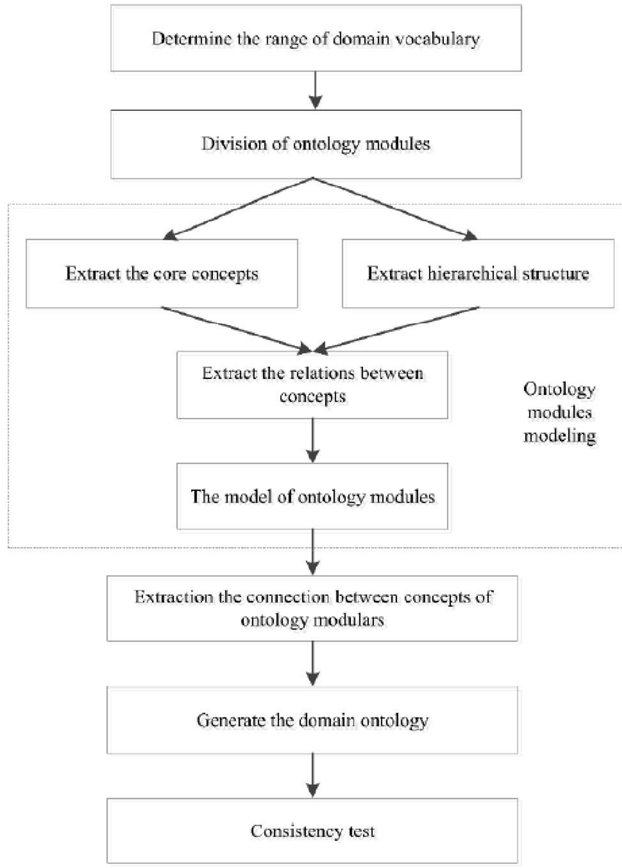

Figure 2 : The modeling method of modular ontology

(1) Determine the range of domain vocabulary

First, we should determine the range of the domain vocabulary. For example, when construct the domain ontology, we choose thesauruses, category list or subject vocabulary? In the choice of vocabulary, try to choose a recognized vocabulary in this domain.

(2) Ontology modules division

After defining the important concepts and their relations in the domain, divide the target domain into several sub domains according to certain principles, which is the division process of ontology module. We should follow the following principles when dividing ontology module based on the domain vocabulary:

(1)The concepts of the same subject category form the same ontology module and the concepts of the different subject category are divided into the different ontology module. The concepts of the same subject are put into the same ontology module, and which can be developed independently by the experts in one field. The relation of the concepts between the different subjects is looser, so they should be put into the different modules. That can control the query and inference of one concept into the effective range.

(2)The division of the module should be understood easily. The goal of the research on ontology is making the computer understand the knowledge, but the most important thing is making people understand. That is conducive for share and reuse of the ontology and convenient to maintain and expand.

(3)The quantity of the modules which are related to a module should be small and the complicacy between the modules should be reduced. That can decrease the dependency and the coupling between modules. The smaller coupling between modules can make each module easy to reuse and share.

So the division of ontology module based on the domain vocabulary mainly depend the professional disciplines. That is conducive for the experts in a field to develop independently and easy to understand and the cohesion between the modules is good.

(3) Ontology modules modeling

For each ontology module, the experts in the field can cooperate to the construction of ontology. Each ontology module is respectively defined in different documents and their vocabularies are respectively specified by a series of namespaces and defined a series of the entities. Regard the module vocabulary as the base, we can extract the core concepts, hierarchical structure and the other relations between concepts. Then define the attributes and the instances of the concept. We use OWL to describe the model of ontology module.

(4) Extraction the connection between concepts of ontology modulars

Extracting the relation between the concepts of ontology module is the core of modeling of the dependency between ontology modules. The dependency also is the base on the next connection of modules. The relation between ontology modules includes synonymy, inclusion, ancestors, brothers and user-defined etc. Defining the relation needs the cooperation of the experts in the field.

(5) Connect the ontology modules and generate the domain ontology

Each module in the modular ontology can be connected by semantic relation, so the connection of ontology modules is based on the semantic connection. We will get the domain ontology after connecting all ontology modules. At present, the integration method between modules mainly includes import and link. The import method import the content of one ontology into the present ontology, and this method is used in owl:import system. The link method integrate the corresponding module into the system of knowledge when a module dependencies the knowledge of other modules. There are kinds of link methods, such as bridge rules, extended the attributes of ontology module, etc.

In this paper, we use the extended owl:import that is proposed by Jie Ba and Vasant Honavar. The model of domain ontology will be generated after connecting the concepts between.the different ontology modules.

(6) Consistency test

Consistency test is needed in the modeling and integration of ontology modules. Checking the consistency mainly includes checking the consistencies of module hierarchy declare and ontology modular logic reasoning consistency. Testing the compatibility of the implicit semantic relation in the module and between the modules is the necessary requirement of the context in each ontology module.

THE EXAMPLE ANALYSIS

An example of the modular ontology of high-speed railway domain is developed to verify the feasibility of the modeling steps. Because the original railway category list is not suitable for the high-speed railway, Ministry of railways organizes experts in different professional domains to reconstruct which is also suitable for literature classifications of high-speed railway and “High-speed railway basic datasheets”. “High-speed railway category list and Thesaurus” has 10 first level categories, 64 secondary categories, 208 third level categories, and 2488 corresponding keywords.

(1) Determine the range of domain vocabulary

We select the "High-speed railway category list and Thesaurus" as the basis of the division of ontology modules. And among "High-speed railway category list and Thesaurus", we select vocabulary lists of five core subjects including EMU, operations management, safety and rescue, traction power supply, and public works.

(2) Ontology modules division

According to the principle of the division method of ontology modules, discipline is a major consideration. And which is conducive for the experts in a field to develop independently and easy to understand and the cohesion between the modules is good. So, we divide the vocabulary lists of high-speed railway into five ontology modules including EMU, operations management, safety and rescue, traction power supply, and public works.

(3) Ontology modules modeling

We adopt Protégé that is developed by Stanford University to build our domain ontologies. At present, we construct five ontology modules including EMU, operations management, safety and rescue, traction power supply, and public works. The model of ontology mainly includes concepts, properties, relations and instances. We use OWL describing language for ontology module.

The first class concept main comes from the words of "thesaurus and thematic words of high-speed railway". The second class concept is the classification of the first class concept, and if the first class concept has two or more division methods, we often choice the most general division method as the second class concept. If the second class concept can still be classified, then divide it down, and until no classification. For example, EMU is the first class concept of the ontology of EMU professional field. EMU has three division methods: according to the dynamic configuration, according to the fashion of supply power and according to speed grade. Among these the most commonly used is according to the fashion of supply power in which EMU is divided into "EMU of power distributed" and "EMU of central power". So the second class concept of EMU is "EMU of power distributed" and "EMU of central power". The above method determines the core concepts and the hierarchical structure of ontology modules.

User-defined relations can be described by adding the special property of concept to connect two concepts. The names of user-defined relations need to be determined by domain experts. For example, add the special property of "supply power" to the concept of "pantograph", which can describe user-defined relation between the concepts of "pantograph" and "EMU". Therefore, we can get the relation: pantograph <supply power> EMU.

The property of the concept is the "datatype" property. For example, we can add "type", "manufacture", "country" and "starting acceleration" properties to EMU concept.

The instance needs domain experts to add. If the concept has sub-concept, the instance is added to the bottom concept. For example, we can add "CRH1", "CRH2", "CRH3" etc. instances to "EMU of power distributed".

(4) Extraction the connection between concepts of ontology modulars

For example, the concept of "contact network system" in ontology module of traction power supply associate with the concept of "pantograph" in ontology module of EMU through the relation of <supply power>. In addition, the relation between ontology modules also includes synonymy, inclusion, ancestors, brothers and defining the relation needs the support of the experts in the domain.

(5) Connect the ontology modules and generate the domain ontology

When we want to use two or more ontology modules, we can link them through the connections between the concepts in different ontology modules. The model of domain ontology will be generated after linking the all ontology modules.

We use owl:import to connect ontology modules, and the following is the key codes of the concept of "pantograph" in ontology module of EMU associate with the concept of "contact network system" in ontology module of traction power supply through the relation of <supply power>:

<rdf:RDF

xmlns="http://www.owl-ontologies.com/Ontology1215582893.owl#"

```
xmlns:protege=”http://protege.stanford.edu/plugins/owl/protege#”
xmlns:rdf=”http://www.w3.org/1999/02/22-rdf-syntax-ns#”
xmlns:xsd=”http://www.w3.org/2001/XMLSchema#”
xmlns:rdfs=”http://www.w3.org/2000/01/rdf-schema#”
xmlns:owl=”http://www.w3.org/2002/07/owl#”
xml:base=”http://www.owl-ontologies.com/Ontology1215582893.owl”>
<owl:Ontology rdf:about=””>
<owl:imports rdf:resource=”http://www.owl-ontologies.com/Ontology1219369205.owl”/>
</owl:Ontology>
<owl:ObjectProperty rdf:ID=” supply power “>
<rdfs:domain>
<owl:Class>
<owl:unionOf rdf:parseType=”Collection”>
<owl:Class rdf:about=”# pantograph “/>
<rdf:Description rdf:about=”http://www.owl-ontologies.com/Ontology1219369205.owl# contact network system “/>
</owl:unionOf>
</owl:Class>
</rdfs:domain>
</owl:ObjectProperty>
```

(6) Consistency test

We use Protégé to install plug-in of Run ontology tests for OWL ontology structure and logic consistency check.

CONCLUSIONS

The method of modular ontology is an effective method that it can solve the problems of the construction of large ontology, ontology reuse and the efficiency of reasoning. Not only the ontology module can be used alone, but also two or more modules can be used in combination by the relationship between the concepts of modules. This paper puts forward a method of modular ontology modeling based on domain vocabulary. First, the domain vocabulary to be modeled is divided into several interrelated ontology modules according to the polymerizable between modules. Then, we construct the ontology model for each module independently. Finally, we connect all ontology modules through the relations between concepts of ontology modulars, and generate the model of domain ontology. Compared with the traditional method of ontology modeling, this method divides the large scale ontology into a series of ontology modules. Each module only has the relationship between concepts in this module, and it can greatly reduce the difficulty of modeling and maintenance. Also, each module has its own function and is independent relatively, which reduce the complexity of the system and help to improve the ontology module’s ability of share and reuse.

ACKNOWLEDGEMENTS

This work was financially supported by the Natural Science Foundation of Hebei Province (F2012208018), the project of Science and Technology Department of Hebei Province of China (134576113D, 12457203D-30), the project of Natural Science Foundation of China (71301044, 71271076) and the project of Hebei University of Science & Technology (QD201017) under which the present work was possible.

REFERENCES

[1] C.G.Bernardo, B.Parsia, E.Sirin; Working with multiple ontologies on the semantic web[J]. Lecture Notes in Computer Science, **(3298)**, 620-634 **(2004)**.

[2] B.C.Grau, I.Horrocks, Y.Kazakov et al.; A logical framework for modular integration of ontologies. IJCAI-07, 298-303 **(2007)**.

[3] H.Stuckenschmidt, M.Klein; Integrity and change in modular ontologies. IJCAI-03, 900-908 **(2003)**.

[4] J.Bao, D.Caragea, V.G.Honavar; Package-based Description Logics-Preliminary Results[C]. Proceedings of the 5th International Semantic Web Conference-ISWC, 967-969 **(2006)**.

[5] O.Kutz, C.Lutz, F.Wolter et al.; ε-connections of Description Logics[OL]. [2009-12-16]. http://www.dcs.bbk.ac.uk/~michael/dl03final2.ps, **(2009)**.

[6] E.Jimenez-Ruiz, R.Berlanga, B.C.Grau; ProSé: A Protégé Plugin for Reusing Ontologies, Safe and

Economique User Manual[OL]. [2009-12-10]. http://krono.act.ujies/people/Emesto/safety-ontology-re-use/proSE-current-version/ProSE-Manual.pdf, **(2009)**.

[7] Jiang Cui-qing, Jia Hao, Ding Yong; Constructing knowledge organization model of mechanical products manufacturing enterprises based on modularization ontology. Machinery Design & Manufacture, **(10)**, 253-255 **(2012)**.

[8] Lin Songtao; Research on the construction of modular ontology.[D]. Beijing: Beijing University of Posts and Telecommunications, **(2006)**.

[9] Zhang Weiyi, Lu Ruzhan; Research and implementation of modular ontology[J]. Application Research of Computers, **24(11)**, 206-209 **(2007)**.

[10] A.Borgida, L.Serafini; Distributed Description Logics: Assimilating Information from Peer Sources[J]. Journal of Data Semantics, **(1)**, 153-184 **(2003)**.

[11] P.Bouquet, F.Giunchiglia, Frank van Harmelen et al.; C-owl: Contextualizing ontologies[J].Web Semantics, **1(4)**, 325-343 **(2004)**.

[12] O.Kutz, C.Lutz, F.Wolter et al.; E-connections of Description Logics[EB]. http://ftp.informatik.rwth-aachen. de /Publications/CEUR-WS/Vol-81/wolter-1.pdf, 2003-09-05, **(2003)**.

[13] J.ao, D.Caragea, V.Honavar; Package-based Description Logics Preliminary Results[A]. In the International Semantic Web Conference (ISWC)[C], 967-969 **(2006)**.

[14] J.Bao, V.Honavar; Adapt OWL as a Modular Ontology Language. In OWL: Experiences and Directions (OWLED 2006)[EB], http://ftp.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol 216/ submission-28.pdf, 2006-12-10, **(2006)**.

[15] J.Pan, L.Serafini; Semantic Import: An Approach for Partial Ontology Reuse[EB].http://ftp.informatik.rwth-aachen. de /Publications/CEUR-WS/Vol-232/,2007-09-06, **(2007)**.

[16] T.R.Gruber; Toward Principles for the Design of Ontologies Used for Knowledge Sharing[J]. International Journal of Human Computer Studies, **43(5-6)**, 907 Nov-Dec **(1995)**