Proceedings of the 2012 16th International Conference on Computer Supported Cooperative Work in Design

Ontology-based Modeling of Manufacturing Information and its Semantic Retrieval

Shaoshuai Li and Lihong Qiao
Department of Industrial and Manufacturing Systems Engineering
Beihang University
Beijing, China
shaoshuailee@gmail.com and lhqiao@buaa.edu.cn

***Abstract*[1]---Due to the continuous growing complexity of manufacturing information (MI) and increasing need to exchange this information among various software applications, the unified fundamental manufacturing process data are necessary. This paper put forwards a MI model whose core contents are products, processes, resources and plants, meanwhile, their complex relations are also explained. Aiming to specify the concepts' unambiguous definition and their relationships both in syntax and semantics, this research proposes an ontology modeling approach and constructs the MI model in Ontology Web Language (OWL).Besides, the detailed characteristics of three typical semantic retrieval and visualization plug-ins in protégé are summarized, so as to realize a better understanding of the complex ontology model. In order to share and reuse this model in collaborative working environment, a semantic similarity algorithm is proposed and the procedures of realizing semantic retrieval is discussed, then a prototype system of semantic retrieval in web browser is developed. In the end, an example is given to demonstrate and verify the above work.**

Keywords- Manufacturing Information; Ontology Modeling; Semantics Retrieval

I. INTORDUCTION

With the deep usage of advanced manufacturing system and emergence of new manufacturing modes, the enterprise manufacturing process has been changed greatly, the accompanying MI are becoming more and more abundant and complex, they not only contain the physical resources that produce parts, such as the equipment, but also include all the information systems and datum.

Different information systems usually reflect different stages of product life cycle and have a close relationship in data transmission and exchange. For instance, computer aided design (CAD), computer aided process planning (CAPP) and manufacturing executive system (MES) respectively focus on product design, process and manufacturing phases, the design result in CAD software is the import of CAPP software. But different information systems often express product's attributes in different aspects, they may use different terms to mean the exact same concept or use the exact same term to mean very different concepts, so it's difficult to realize integration between different systems.

Meanwhile, the complexity of MI makes the convenient knowledge retrieval more and more necessary. The traditional keyword retrieval is usually only keyword matching, which stores the relevant data in database. It can't realize the semantic retrieval such as synonymous concepts (e.g. machine and equipment) and sub-super concepts (e.g. machining sequence and machining operation). In addition, it basically doesn't support the logical reasoning among the concepts and the retrieval results are mainly represented in the form of text hyperlinks. So the tradition method can't satisfy the retrieval and representation demands of complex domain model.

There have been some efforts on the two problems above. National Institute of Standards and Technology of the US put forward process specification language to ensure unambiguous exchange of process information within the manufacturing domain [1]. Lin, H.K from Taiwan investigated the manufacturing system engineering ontology model to facilitate communication and information exchange in inter-enterprise [2]. Lihong Qiao from Beihang University proposed an approach to modelling manufacturing process information based on PSL specifications [3]. Li Shanting from Zhejiang University implemented the network manufacturing resources intelligent retrieval based on semantic [4]. But the above research works mainly focus on the solid resources or other domains, and the format of search results is monotonous and non-visualized.

This paper proposes an ontology-based MI model with the core of products, processes, resources and plants, which describes and specifies all the information related to manufacturing process. In addition, the semantic retrieval and visualization of MI in Protégé is investigated for the sake of a better understanding of complex relations among the ontology. In order to realize the sharing and reuse of this model in a wider scope, a semantic similarity algorithm is proposed. At last, the semantic retrieval of MI in web browser is realized.

II. MANUFACTURING INFORMATION DOMAIN ONTOLOGY

A. Manufacturing information model

This work is supported by the National Natural Science Foundation of China (Grant 51075022), Significant Science and Technology Special Program of Advanced Machine tools and Elementary Manufacturing Equipment (2010ZX04016-011, 2011ZX04016-021).

The MI model concerns on resources and information that relates to manufacturing process. It mainly includes the product itself, the plant where the product is created, processes information that describe the procedures and sequences of products, physical resources, and their relationships. From the product lifecycle aspect, MI tries to describe the information related to the product design result, process planning, materials quota and manufacturing. The ultimate goal of this work is to allow support for the unified definition, dynamic evolution and sharing of the datum in the manufacturing process, also integration with other information systems, such as CAPP and MES.

Figure 1 shows the conceptual model of MI using UML, which mainly describes the entities and their relationships. Part, Configurable Attribute, Feature and Part Instance are the main four objects in product model. Part contains a union set of attributes from the design, process and manufacture to assembly process. Configurable Attribute expresses the different attributes in different parts, which consist of a part family. Part Instance is a specific part with the explicit attribute value. Part is made up of many Features.

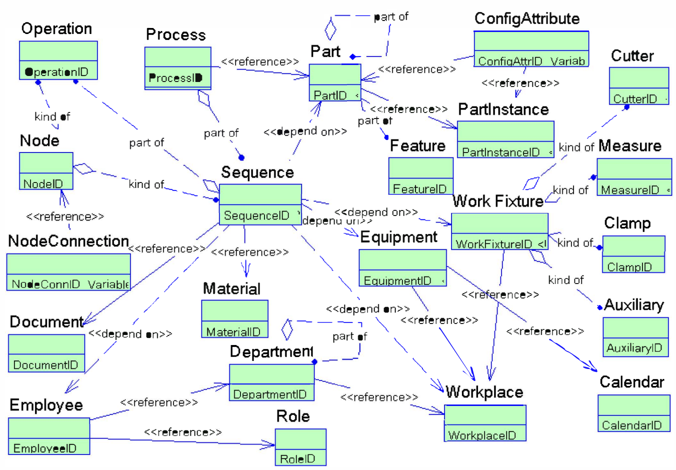

Figure 1. The main entities and their relations in manufacturing information conceptual model

The prime entities in resource model are Equipment and Work Fixture. Work Fixture consists of Cutter, Measure, Clamp and Auxiliary. The plant model mainly contains Employee, Department, Role and Workplace. Employee takes a certain role in the department and finishes the manufacturing processes of a part at certain workplace.

The process model primarily includes the following entities: Process, Sequence, Operation, Material, Document, Node and Node Connection. Process means the manufacturing method and procedure that makes a part evolving from the raw material to the finished part, it consist of many sequences while sequence is make up of many operations. From the point of view of the graph theory, Node and Node Connection constitute the whole process. Node maybe a sequence or operation or virtual node (e.g. and/or). Node Connection connects the previous node and the next node. Material and Document are the input or output of a process.

B. Ontology definition and its supporting technologies

Ontology is an explicit specification of a sharing conceptualization, which defines a set of representational primitives with which to model a domain of knowledge or discourse [5]. Inference support is another key capability being added to ontology applications. An inference tool is essential to make further assertions and discover new knowledge. Ontologies also often provide powerful tools for querying and information retrieval [6]. Thus, ontology can be used to share and reuse knowledge, and could be a solution to the interoperability between different information systems [7].

The ontology technology made great process in the last ten years. Several ontology definition languages have been created including Resource Description Framework (RDF), RDF schema and OWL, now OWL is the World Wide Web Consortium (W3C) standard semantic markup language for publishing, sharing and reuse of semantic data with great machine interpretability [8]. Another two important tools are Protégé and Jena. Protégé is an ontology editor and knowledge acquisition system that supports ontology building using RDF, RDFS, and OWL [9]. Jena is a JAVA framework for building semantic web application, it provides a programmatic environment for RDF, RDFS, and OWL, and also includes a rule-based inference engine [10].

Five meta-elements $Ontology = \{C, P, I, R, Cs\}$ are proposed to build the domain ontology model, which respectively stand for Class, Property, Instance, Relationship and Constraint.

(a) Classes represent domain concepts and can be arranged in inheritance hierarchies. They should give the specific definition of concepts both in syntax and semantics level.

(b) Properties are used to describe and differentiate a class from the others. A primary property with the unique value is usually necessary to define a class.

(c) Instances are individuals of the classes with the properties having specific values.

(d) Relationship is an important semantic expression among different classes. There exits three main relations among classes: inheritance, composition and reference. Inheritance or kind-of reflects the relation that one class is a subclass of another. Composition or part-of relation defines the ownership. Reference means associating with or depending on other classes. Inheritance and composition relations make up the ontology tree. The three relations together create the network structure. In OWL, it can also state that two classes are the same (owl:sameClassAs), equivalent (owl: equivalentClass) [11].

(e) Constraints represent restrictions on the valid values of a certain property, such as the minimum and maximum of the value. It also includes the knowledge inference rules.

C. Manufacturing information ontology construction

When constructing the MI domain ontology, both the Bottom-Up and Top-Down modeling strategy are used. Bottom-Up strategy means collecting MI and data from the low level, Top-Down strategy means dividing MI into several modules first with the help of experts. The knowledge sources of MI include international manufacturing data standards such as ISA-95, relevant handbooks (e.g. machining process manual), enterprises' production experiences, and the previous study of the unified manufacturing data model [12] proposed by Beihang University and so on.

Corresponding to the five meta-elements mentioned above, protégé provides the "Class-Slot-Forms-Instances-Queries" plug-ins to build the ontology stored in owl or pprj file. Apart from the default relations in ontology, protégé provides the "Instance" and "Class" data type to respectively express the reference and super relation with other classes. Besides, the ontology file can be transferred to XML for further use, so here choose protégé to build the MI ontology. The construction of turner operation ontology using protégé is as shown in Figure 2.

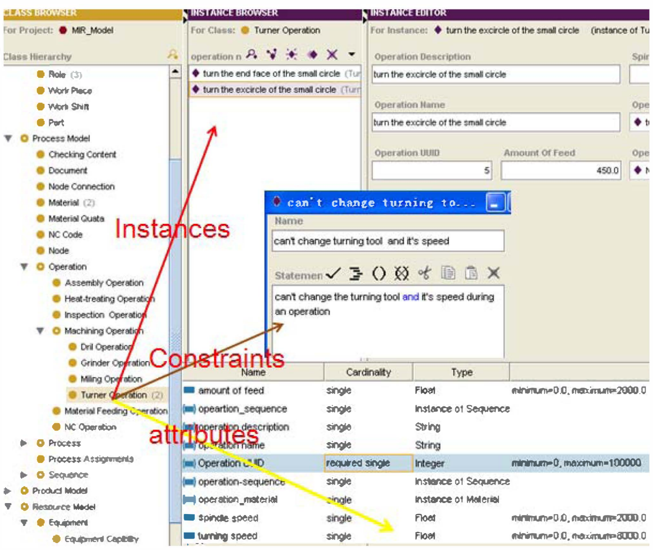

Figure 2. The construction of turner operation ontology using protégé

III. SEMANTIC RETRIEVAL AND VISUALIZATION OF MANUFACTURING INFORMATION

Semantics is defined as the "meaning and the use of data", it is all pervasive and covers many aspects such as the interpretation and the user perspective of the data [13]. Semantics is sometimes viewed primarily in terms of capturing the similarities between objects while others consider the identification of the relationships between objects as the key. Semantic retrieval pays more attention to information's actual meaning in applications, instead of just keyword matching.

A. *Semantic retrieval in protégé*

Protégé has provided the following 2D semantic retrieval and visualization plug-ins: TGViz, OntoViz and Jambalaya. In addition, it contains the full text retrieval plug-ins: Lucene Query and String Search, and the latter's searching range is wider and deeper than the former's.

TGViz allows users to navigate gradually making visible parts of the graph. A variable radius of visibility is used to limit the size of the graph in smaller, more manageable sizes. The user may also expand or retract nodes, hide them, and change the node on focus by double clicking on it. In OntoViz, it is possible for the user to choose which ontology features will be displayed, as well as to prune parts of the ontology from the Config Panel on the left. Right-clicking on the graph allows the user to zoom-in or zoom-out. Jambalaya provides a set of tools including several node presentation styles, configuration of display properties and different overview styles. Table 1 shows the above plug-ins' detailed characteristic of semantic retrieval and visualization.

Figure 3 and Figure 4 demonstrate the semantic retrieval result of operation respectively using TGViz and Jambalaya. It's easy to find out that the result not only contains the relevant entities, but also their instances and relations with other entities. In addition, the dynamic reticular is much more comprehensive than the traditional text hyperlinks.

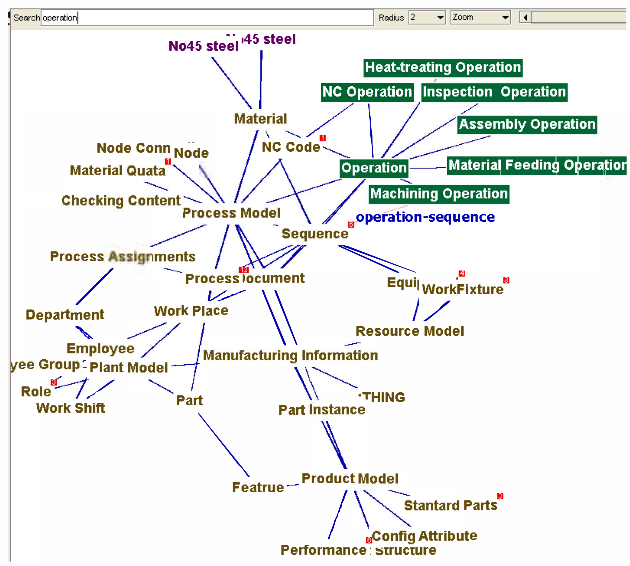

Figure 3. The semantic retrieval result of operation using TGViz

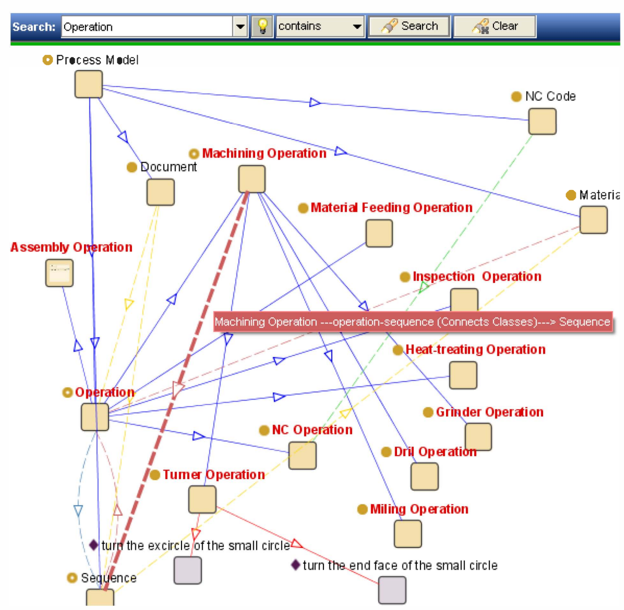

Figure 4. The semantic retrieval result of operation using Jambalaya

Table 1. The function comparison of semantic retrieval and visualization plug-ins in Protégé

Plug-in	Classes and Instances	Class Hierarchy	Multiple Inheritance	Role Relations	Properties	Semantic Retrieval
TGViz	Classes and instances are represented as labels of different colors	Lower level nodes are displayed around their parent and connected to it with a "sub" link	There is a link from the node to both its parents	Links with labels on mouse over are used	Properties are displayed in a separate window	Yes, but only for the part of the ontology that is already visible
OntoViz	Rectangle nodes with different color for classes and instances	Child nodes are placed under their parents and linked with an "is-a" link	The child node is linked with all the parents	They are represented with labeled links	Properties are displayed on the node	No
Jambalaya	Represented as rectangles inside their parent node	Lower levels are presented by smaller size rectangles and placed inside their parent nodes	Child nodes are placed under both parents.	Supported through the properties, as directed links with their label visible as a tooltip	Properties are displayed as an embedded form if the selected node is zoomed-in	Yes, offer several types of the searched item and can search between the results

B. Semantic retrieval in web browser

Although realizing the semantic retrieval of ontology in protégé is feasible, it is not enough for a wider knowledge sharing and collaborative work, after all, installing and using protégé is challenging for ordinary users. A better solution is to realize the semantic retrieval of MI in web browser.

As mentioned above, the three main relations among classes produce the ontology's dynamic network structure. In order to make these relations more determined and objective, graph theory is used to describe the ontology and to quantify the semantic relation between two classes [14]. First, two definitions are introduced as follow.

Definition 1: Ontologies can be considered as an undirected weighted graph $G=(V,E)$. V stands for the set of vertexes; a vertex is a class in ontology. E means the set of edges; an edge is the relation between two classes.

Definition 2: Semantic similarity. Supposing A and B are two vertexes in E, the semantic similarity between A and B is called $S(A,B), S\in[0,1]$. $S(A,B)\in(0,1]$ means vertex A and B have certain similarity, the value of $S(A,B)$ is proportional to their similarity criterion. If A and B are the same or equivalent class, their similarity is 1. If A and B totally have no relation, their similarity is 0.

In the author's opinion, the similarity degree of the three relations (kind-of, part-of and inference) are different, the closeness of the first two relations is tighter than the last relation's. Suppose that parameters α,β,γ respectively standard for their similarity coefficient and parameters x,y,z respectively standard for their quantity between two vertexes, α,β,γ should meet the constraint $1<=\alpha<=\beta<\gamma$.

In the graph, there exists at most one direct path and probably many indirect paths between two vertexes. The similarity of two vertexes connecting by a direct or indirect path is $\frac{1}{1+\alpha x+\beta y+\gamma z}$, typical suggested values are $\alpha=\beta=1,\gamma=2$, and the true semantic similarity between two vertexes is the maximal similarity among all feasible paths.

The semantic similarities among vertexes A, B, C and D in figure 5 are givens as follow.

$$S(A,B)=\max\{\frac{1}{1+\alpha x+2\gamma z},\frac{1}{1+\alpha x+\beta y}\} \quad (1)$$

$$S(A,C)=\frac{1}{1+\gamma z}, S(A,D)=\frac{1}{1+\alpha x+\gamma z} \quad (2)$$

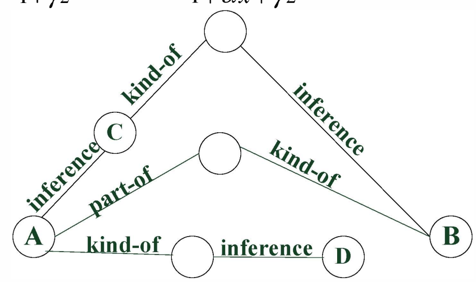

Figure 5. A schematic diagram about the relations between classes in ontology

The procedures of realizing the semantic retrieval in web browser can be described by Figure 6, which includes the data model level, retrieval logic level and client input level. The data flow is viewed from the user's aspect, while the control flow is viewed from the backend to the front end. These procedures mainly contain the following steps.

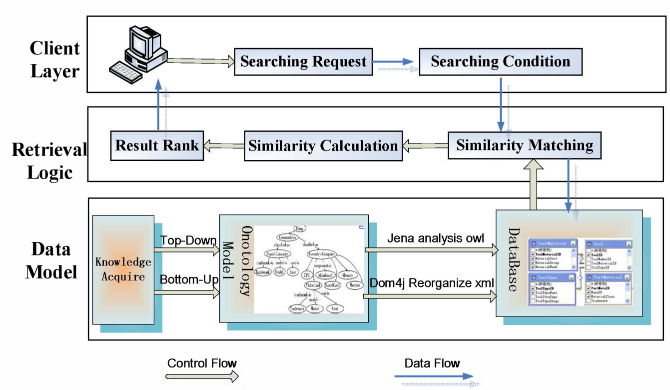

Figure 6. The procedures of semantic retrieval

Step 1: Acquire the domain concepts from all the relevant knowledge sources, and then use the Bottom-Up and Top-Down modeling strategy to construct the domain ontology.

Step 2: Use Jena to analyze the ontology file whose format is usually owl, or use dom4j to reorganize the exported ontology file with xml format, and then extract all the concepts and their relations into database.

Step 3: Transform all the set of relations to an undirected weighted graph.

Step 4: Compute the semantic similarities between the target concept inputted in the web browser and other concepts, then store them in a descending order.

Step 5: Show the retrieval results in the web browser whose similarities are top ten or bigger than the expected value. And the concept with biggest similarity will be listed first.

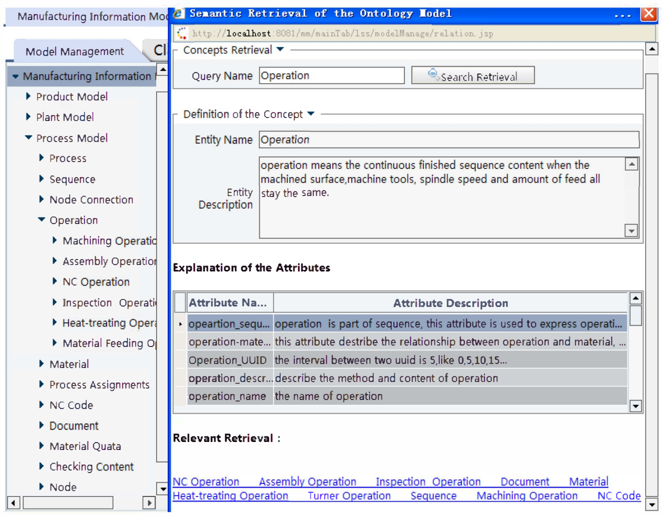

Figure 7. The semantic retrieval result of operation in web browser

According to the above method, a MI data management and semantic retrieval prototype system is developed, users can manage the concept's definition, attributes and relations with other concepts just like they do in Protégé. Figure 7 shows the semantic retrieval results of operation in this system, which include operation's definition, attributes, and relevant concepts. The relevant concepts are not just confined to those concepts that contain the characters 'operation', such as NC operation and assembly operation, while these results are the traditional searching engine like the plug-in Lucene Query in protégé can deal with. The relevant retrieval also includes concepts that have practical relations with operation, like material and NC code. When users click the relevant concept, they can understand its further information.

IV. DISSCUSSION

Ontology modelling theory has been a new solution to the isolation and chaos of information among different functional application systems, a set of normative terminology is helpful to the unification of the same data in multi-systems, also to the reuse of knowledge. In addition, the diversified visualization plug-ins in protégé make the domain model much easier to be understood, while the data model built through the relational database is indigestible, especially when the relations among entities are very intricate.

Although the ontology theory has advantage in domain modeling than the database, it relies on the modeling and management tools, such as Protégé, meanwhile, these tools are relatively independent closed systems. So an embarrassing phenomenon appears: there exist many ontology-based domain models but very few practical applications. One of the contributions in this paper is providing a feasible approach to transfer the ontology model into database and develop an ontology semantic retrieval prototype system. Compared with the traditional keyword matching method, the semantic retrieval based on ontology theory has better performance on the precision, due to the extension in relations among the concepts.

V. CONCLUSION

This paper focuses on the development of ontology of manufacturing process concepts and terms along with their respective formal and unambiguous defines, aimed at offering a standardized and consolidated domain model and supporting the information sharing and interoperation between the multiple application systems. For the purpose of a better understanding of the complex model, the semantic retrieval and visualization of MI with the format of dynamic reticular is realized in protégé. Because of the convenience to share and reuse knowledge with internet, this paper comes up with a semantic similarity algorithm and realizes the semantic retrieval of MI in web browser. The demonstrations of operation show that ontology has a better inner mechanism to ensure the explicitness and sharing of the domain model; and the completeness of semantic retrieval is higher compared with the effect of the traditional full-text searching engine. Our ongoing study involves improving the semantic similarity algorithm to enhance retrieval performances both on the precision and the recall.

ACKNOWLEDGMENT

This work is supported by the National Natural Science Foundation of China (Grant 51075022), Significant Science and Technology Special Program of Advanced Machine tools and Elementary Manufacturing Equipment (2010ZX04016-011,2011ZX04016-021). The authors would like to express their appreciation to the agencies.

REFERENCES

[1] Grüninger, Michael1, Menzel, Christopher, "The process specification language (PSL) theory and applications," AI Magazine, vol. 24, no. 3, pp. 63-74, 2003.

[2] Lin, H.K, Harding, J.A, "A manufacturing system engineering ontology model on the semantic web for inter-enterprise collaboration," Computers in Industry, vol. 58, no. 5, pp. 428-437, 2007.

[3] Lihong Qiao, Shuting Kao, Yizhu Zhang "Manufacturing process modelling using process specification language," International Journal of Advanced Manufacturing Technology, vol. 55, no. 5-8, pp. 549-563, 2011.

[4] Li Shanting, Research on intelligent retrieval of network manufacturing resources based on semantics, (MS., Zhejiang University, China, 2008), pp.48-60.

[5] What' ontology [EB/OL]. http://tomgruber.org/writing/ontology-definiti -on-2007.htm

[6] Zhan, Pei, Jayaram, Uma, Kim, OkJoon, Zhu, Lijuan, "Knowledge representation and ontology mapping methods for product data in engineering applications," Journal of Computing and Information Science in Engineering, vol. 10, no. 2, pp. 1-11, 2010.

[7] Staab, S. and Maedche, A., 2000, "Ontology Engineering beyond the Modeling of Concepts and Relations", 14th European Conference on Artificial Intelligence; Workshop on Applications of Ontologies and Problem-Solving Method, Berlin, Germany.

[8] W3C, 2004, OWL Web Ontology Language Overview [EB/OL]. http://www.w3.org/TR/ owl-features/

[9] The Protégé Ontology Editor and Knowledge Acquisition System [EB/OL].http://protege. stanford.edu/

[10] JENA—A Semantic Web Framework for JAVA [EB/OL]. http://jena. sourceforge.net/

[11] H.K. Lin, J.A. Harding, "A manufacturing system engineering ontology model on the semantic web for inter-enterprise collaboration" Computers in Industry, vol. 58, no. 5, pp. 428-437, 2007.

[12] Liu Wei, Qiao lihonghong, "A Unified Manufacturing Resource Modeling Architecture Base on meta-model," Computer Integrated Manufacturing System, vol. 13, no. 10, pp. 1903-1908, 2007.

[13] Sheth, A., 1996, "Data Semantics: What, Where, and How?" Data Semantics (IFIP Transactions), R. Meersman and L. Mark, eds., Chapman and Hall, London, pp. 601–610.

[14] Liuchen, Long Hongneng, Yin Guofu, "Describing Method of Semantic Ontology for Process Knowledge of Complex Product," Journal of Sichuan University, vol. 32 no. 2, pp. 243-247, 2007