

Curriculum knowledge organization and representation based on ontology

Zhang Yuanfeng and Zhang Liang
School of Information Science and Engineering
Hangzhou normal University
Hangzhou, China
yuanfeng_zhang@hotmail.com

Abstract—Currently most present web-based course still organized knowledge in accordance with sections. In view of the existing problems of the above knowledge organization form, this paper takes the curriculum of *computer organization* as an example to discuss ontology-based knowledge representation. In the extraction of concepts and their relationships, this paper not only borrows ideas from WorldNet, a quite common ontology that has considerable influence in the international computing linguistics, but also depends on the actual need of the course to expand and redefine them, in order to improve the sharability and reusability of the established curricular ontology. Moreover, methods of storage and exchange of ontology-based database are introduced in detail for providing better knowledge display and organization in actual applications.

Keywords—*computer organization curriculum; knowledge organization; knowledge representation; ontology*

I. INTRODUCTION

With the advance of computer and network technique, more and more adaptive and intelligent web-based teaching system has been successfully developed. Online education form which combines with the contemporary advanced technology has gradually been accepted and approved as a new way for people to gain knowledge and improve their own. In the meantime, knowledge organization and knowledge representation of web-based education becomes the focus of attention in related fields. At present, lots of online course still not get rid of traditional textbook knowledge organization form which is no good to play the advantages of online teaching environment. Hence, this paper takes the curriculum of *computer organization* as an example to illustrate the ontology-based course content organization and representation.

So far, there are many systems of ontology being created. According to the different application domain, the processes of constructing ontology with the applicable methods are also different. In the field of education, the development of ontology undoubtedly requires the participation of domain experts. Meanwhile, construction of ontology also needs the help from computer science ontology engineer, which increases the difficulty for the development of ontology to popularize. The diversity of domains leads to the knowledge (including concepts and Furthermore, in the same subject

areas knowledge may be represented by different particle size with difficulty levels. Currently, there is scarcely recognized evaluation and quality assurance standards of ontology design. Hence, how to organize and represent course ontology is discussed in following sections based on WorldNet and other related theory.

II. RELATED WORK

The knowledge representation and acquisition techniques are the basis of knowledge engineering technology, and the knowledge representation is also a core issue in the field of knowledge engineering. The current semantic-based knowledge representation is a hot research area, leading to the study about ontology and Semantic Web increasing fast. In the definition of the ontology, Gruber [1] put forward "ontology is an explicit specification of a conceptualization" in 1993, widely accepted and showed high citation rate. In short, ontology makes use of concepts, terminologies, and their interrelationships standardization to explain and describe a particular field or even wider range of domain knowledge concepts and their relations, which can be shared for well recognized, clearly defined and uniquely identified. For the convenience of communication among peoples and machines, Semantic Web, intelligent information retrieval and information integration have been widely used in the applications of ontology [2].

Formal definition of ontology is a five-tuple [3] $O = \{ C, R, H^c, Rel, A^o \}$, which C represents the set of concepts, R the set of relationships, H^c the set of classification relations between concepts, Rel the set of non-classified relations between the concept, A^o the set of axioms. Thus, the process of the extraction and definition of the above five elements should be included in building an ontology. In the construction of domain ontology, researchers have put forward many guidelines, methods and principles. Because of the relative immaturity of ontology engineering, the methods used by the researchers in different application fields are also different, which have their special features, advantages and defects. At present, massive ontology have been developed in many fields, such as medical ontology UMLS[4], the method for software requirements acquisition based on ontology[5] and Lu Ru-qian's common sense knowledge systems[6]. Concerning literatures about specific

domain ontology construction, [7] studied the construction of geometry knowledge ontology.

Throughout above literatures, there is less case focusing on course knowledge ontology construction about specific subject. Here, curriculum of *computer organization* is taken as an example to discuss knowledge organization and representation of curricular ontology. In the course of concepts and their relations defining, a general ontology database, WorldNet, is introduced to construct curricular ontology to meet online Web-education requirements, with reusable, re-definable and extendible capability. Methods and technologies of this case provide valuable reference for other courses' knowledge organization.

III. USING ONTOLOGY TO ORGANIZE AND REPRESENT COURSE KNOWLEDGE

This paper uses the Protégé [8] to build the curricular ontology of *computer organization*. The process of ontology-based knowledge representation firstly needs object-oriented abstraction of information and knowledge resources, and then the extraction of content and relationship of objects or concepts, according with the syntax of ontology knowledge representation to establish abstract classes (including concepts, attributes and relations, etc.).

In the process of building curricular ontology the definition of the main concepts and relationships here mainly consult WorldNet [9], which is an English dictionary based on cognitive linguistics designed by psychologists, linguists and computer engineers of Princeton University. It also can be seen as an English lexical semantic network system. All English vocabulary in WorldNet is divided into five types: nouns, verbs, adjectives, adverbs and function word. Each word has a series of meaning (Sense), which has close mean with a set of synonym words (Synset) respectively. Word and word will be associated by relations which are hyponyms, hypostasis, membership, similarity, etc. These relations will associate one word with another forming a network structure. Thus, WorldNet can be regarded as a kind of general ontology. At present, WorldNet is used for computational linguistics, text analysis and other related field. Now it has considerably influence in the international computing linguistics.

A. Determination of the basic concepts

The concept of ontology will be organized in the taxonomic point of view, formed a system of classification structure. There are many methods to define the hierarchy of concepts. The most commonly used methods are the following three kinds [10]:

1) Top-down method: This method first defines the most conventional concepts in the field, followed by the more specific concepts.

2) Bottom-up method: This method begins with defining the most detailed concepts, namely leaves part of the hierarchical structure, and then combining these concepts into a more general and upper class concepts.

3) Synthesis method: This method is a combining method of top-down and bottom-up method.

We mainly adopt the top-down method, from the top level concept "computer system" downward particulars. After enter the word, computer system, into WorldNet to retrieve, we will see its definition and relationship to other words as Fig. 1 showed.

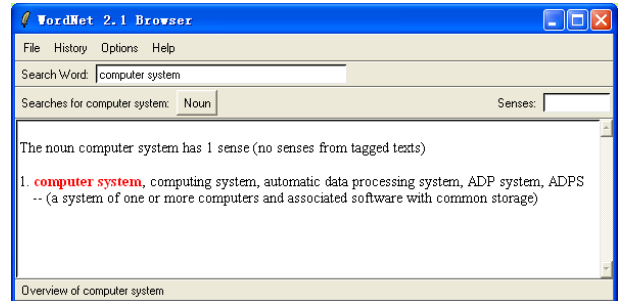


Figure 1. The defined word of WorldNet

From the query result, we can extract the definition and synonyms of the concept. According to the need for trade-offs, following this method downward to the definition of other concepts, we combine the retrieval result of WorldNet with the actual need of the curriculum. Concept of "computer system" is divided into two parts: computer hardware system and computer software system, then refines these two concepts respectively, forming the whole course concepts hierarchical graph as Fig. 2 showed.

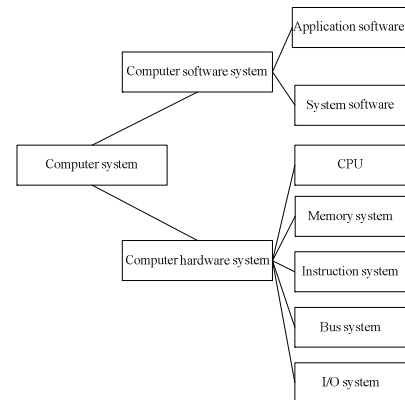


Figure 2. The concept hierarchical graph

B. Definition of properties of the concepts

Concepts can be defined in two categories of properties: static properties and dynamic properties.

1) Static properties

It is mainly used to describe the meaning of the concept and characteristics of the static properties, called Datatype Property, needing no behavior support. For the concept of this course, we mainly define the following static properties:

a) C_Id: Concept identifier.

b) C_Name: Concept name.

c) C_Type: Knowledge type of the concept, mainly includes: factual knowledge, conceptual knowledge,

procedural knowledge, metacognitive knowledge, separately denoted as F, T, P, M, where F = Fact, T = Concept, P = Process, M = Meta.

d) C_Objective: Knowledge goal of the concept, mainly includes: memory, understand, apply, analysis, separately denoted as A, B, C, D.

e) C_Difficulty: Knowledge difficulty of the concept, mainly includes: very simple, simple, medium, difficult, very difficult, separately denoted as 1,2,3,4,5.

f) C_Description: Description of the concept.

g) C_mediaLink: Media resources links, the resources related to this concept, such as: the web, video, audio, animation and so on.

h) C_Father: Father concept set.

i) C_Child: Child concept set.

j) C_Previous: Precursor concept set

k) C_Subsequent: Subsequent concept set.

l) C_Reference: The relevant concepts are contributed to understand of the concept, such as similar concepts, concepts in the same classification level.

2) Dynamic properties

They are dynamic properties, called Object Properties, referring to specific definition, computing, process, events of call, which describe binary relation between concepts. Clicking the Noun label, WorldNet will show us the defined relationship of a word as shown in Fig. 3.

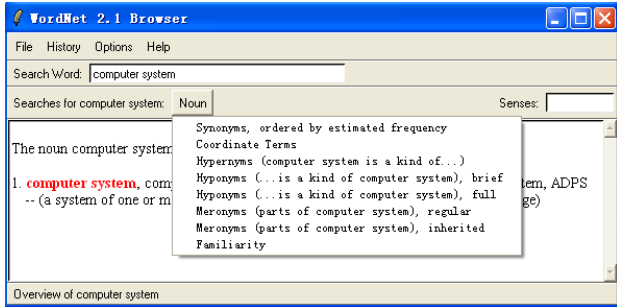


Figure 3. The defined relationships of WorldNet

The defined relations in WorldNet can be seen, including synonyms, coordinate, hyponyms, is-a, parts-of, etc. According to the curriculum needs and referencing above relationships defined, we mainly define and use the following relations in this course ontology.

a) isPartOf: It expresses the sub-concept is an integral part of the parent. For example, memory system is an integral part of computer hardware systems.

b) NarrowThan: It describes the sub-concept is in more detailed features than its father, which can represent a more narrow face. Such as memory systems can be divided into: main memory, auxiliary storage, cache memory.

c) hasInstance: It describes the sub-concept is an instance of its father in certain characteristics, such as output devices with instances monitor, printers, etc.

d) ParallelWith: It means concept A and concept B have the same father. Because input devices and output devices have same parent peripherals, they are brothers at the same level.

C. Using Protégé to construct curriculum ontology

In the process of constructing ontology through Protégé software, it generally includes four basic steps: creating documentation, defining of class and class hierarchy, defining attributes and allowed values and adding instance. Here, we use Protégé3.4.4 to construct course ontology. One snapshot of part of the concept hierarchical graph is showed as Fig. 4.



Figure 4. Part of the concept hierarchical graph

IV. THE STORAGE AND APPLICATION OF ONTOLOGY BASED ON RELATIONAL DATABASE

The development of semantic web applications typically includes the establishment of domain ontology and the browsing, querying and searching based on domain knowledge. Constructing ontology is the premise condition for the realization of domain knowledge sharing and information automated processing. The built ontology needs a durable and platform-independent way to store, and then it can serve as the medium for people and machine to cognize the domain knowledge.

In this paper, Jena [11-12] will be used to save the built ontology into relational database MySQL. Jena is a Java-based open source semantic web toolkit developed by Hewlett-Packard, which can be used for providing a programming environment and a rule-based inference engine to parse RDF, RDFS, and OWL ontology. Jena provides the interface to store the RDF data into a relational database. Model, Resource, Query and other interfaces can be used to access and maintain RDF data in the database. When processing the data, the applications do not have to directly manipulate the database (but through Jena's API), also needn't know the database schema. Preserving ontology into the database includes the following steps: loading the database JDBC driver, creating a database connection, creating a ModelMaker for the database and creating a model for the ontology. Then we can construct a query or use the inference engine to operate on the ontology to obtain the necessary search. For example, we can enter a keyword to extract the words which have hyponymy or synonymy to the input word to achieve semantic-based information service.

V. CONCLUSION

Along with the rising and developing of semantic Web, ontology knowledge representation and processing technology gradually becomes the focus of attention in related field, which embody thoughts of semantic Web ideas. This paper explores the ontology-based knowledge organization and expression, extracting and defining the key concepts and their relations of the curriculum with reference to WorldNet, and according to the actual teaching need to be re-defined or extended, so that the established course ontology can be established as robust, reusable and scalable.

This paper discusses the construction of curriculum knowledge ontology. Because the research achievement of online course ontology knowledge organization form is not comprehensive and complete, few available construction method and the corresponding reference standards can be got, and therefore the built curriculum ontology will inevitably

exist omissions. In the follow-up, we need to keep the curriculum ontology being revised and expanded in order to better serve as a basis for the online course.

REFERENCES

- [1] Gruber T R. A Translation Approach to Portable Ontology Specifications. *Knowledge Acquisition*, Vol.5, No.2, 1993, pp.199-220.
- [2] DU Xiao-Yong, LI Man, WANG Shan. A Survey on Ontology Learning Research. *Journal of Software*, Vol.17, No.9, Sep. 2006, pp.1837-1847.
- [3] Maedche A. *Ontology Learning for the Semantic Web*. Boston: Kluwer Academic Publishers, 2002.
- [4] Boden reider O. The unified medical language system(UMLS):integrating biomedical terminology. *Nucleic Acids Research*, 2004, 32, Database issue D: 267-270.
- [5] LU RQ, Zhang SM. PANGU—An Agent-Oriented Knowledge Base. *World Computer Congress*, 2000.
- [6] Jin Zhi.Ontology-Based Requirements Elicitation.Chinese Journal of Computers, Vol.23,No.5, May.2010, pp.486-492.
- [7] ZHONG Xiu-Qin, FU Hong-Guang, SHE Li, HUANG Bin. Geometry Knowledge Acquisition and Representation on Ontology. *Chinese Journal of Computers*, Vol.33, No.1, Jan.2010, pp.167-173.
- [8] John H. Gennari, Mark A. Musen et al. The evolution of Protégé: an environment for knowledge- based systems development. *International Journal of Human-Computer Studies*, Vol.58, Jan.200, pp.89-123.
- [9] Miller G, Beekwith R, Fellbau C.Introduction to WordNet:an on-line lexical database. In *International Journal of Lexicography*, Vol.3, No.4, 1990, pp.235-244.
- [10] Ashenhurst. *Ontological Aspects of Information Modeling*. *Minds and Machines*, Jun.1996, pp.287- 394.
- [11] Jena.<http://jena.sourceforge.net/>
- [12] Kevin Wilkinson.Jena Property Table Implementation.The Second International Workshop on Scalable Semantic Web Knowledge Base,2006.