

Construct Primary Education Semantic Ontology Library Based Mind Mapping

Dong-Hong Hu¹ and Xue-Jun YU²

¹ the Information Center of Civil Aviation Administration of China, Beijing 100710, China

² School of Software Engineering, Beijing University of Technology, Beijing 100124, China

Abstract. Researches conducted for Mind mapping application in primary education semantic ontology, while considering unique characteristics of primary education, found there were rare widely used ontology libraries and few connections between ontology libraries for information sharing and reuse. In addition, primary semantic ontology library lack precise definitions of the semantics. This paper proposed a solution based on cluster structure derived from mind mapping by providing logical description of the ontologies to precisely define semantics; Meanwhile, tags were adapted to associate different ontologies to form ontology library.

1 Introduction

Mind mapping is a thinking tool that presents a thought process, its application can be roughly summarized in the three aspects which are teaching, business management and personal knowledge management. Tony Buzan who is Britain's memories of father defined mind mapping as this: mind mapping is an expression of radioactive thinking, and therefore it is the natural functions of the human mind. It is a very useful graphics technology, is the master key to open the brain potential. Mind mapping can be used in various aspects of life, learning ability and clear thinking improved by mind mapping will improve people's behavior.[1]

The thinking process presents by mind map is a key of information organization and sharing.[2] The introduction of the mind map aided design in teaching, not only to give full play to the role of the students left and right brain, improve students' learning efficiency, but also can develop students' thinking ability, and form their own way of thinking.

Different from the traditional mind map creation mode, automatically generate mind map through semantic links is a new creation, this approach needs to be realized by means of semantic ontology library. Meanwhile, the mind map by showing a clear relationship can be used as a reference and basis of semantic ontology structure.

The ontology of educational resources can construct a clear knowledge architecture of learning content to conveniently organize and manage learning resources, can provide a guarantee for learning resources sharing and collaborative, provide the basis for learning activities and

optimize the process, greatly improve the reuse rate of educational resources. Although the application of education resource ontology has become a research hot spot, receives the widespread attention of scholars, and some achievements have been obtained, but its development especially in the primary school education still not mature enough, mainly in the following aspects: first, there were rare widely used ontology libraries; second, sharing and reuse of ontology model is not enough, and ontologies that constructed by many developers are according to theirs' subjective experience; third, the related technology of ontology that automatically acquired and expanded is still immature, domain ontology acquisition, organization, processing mainly in manual way; fourth, there is no unified standardized ontology modeling method, and lack corresponding standards of ontology evaluation.

Thus, the research goal is to define primary education domain ontology, including the analysis and design of ontology, to explore the method of creating primary education ontology and expanding primary education ontology library.

The rest of the paper is organized as follows. In Section 2, paper presents requirement analysis of ontology. Section 3 describes the construction ideas and creation method of primary education ontology. Section 4 contains the conclusions and future work.

2 Requirement Analysis of Primary Education Semantic Ontology

From the basic pattern of mind map view, its hierarchical structure can be regarded as a tree structure that divergent from the center to the periphery. There is a central node, surrounded by the divergence of the child node. The central node and child node is the inheritance relation, inclusion relation, whole and part relation or attribute relation. The relationship between sub nodes is paratactic. This tree structure basically satisfy the ontology definition of original hierarchical structure. This paper defined for primary education ontology is based on the hierarchical structure of mind mapping. Meanwhile, tags were adapted to associate different ontologies to form ontology library.

2.1 Requirement Analysis of Ontology

Primary education followed the gradual rules, old knowledge set the stage for learning new knowledge. There is a hyponymy or cross relation between knowledge points. A knowledge point usually includes the concept, the nature of concept, the relationship of concepts, the general rule of concepts, and so on. Therefore, the research of primary education ontology needs to explicitly establish links among concepts which according to generic relationship of concepts, and also needs to clearly define the attribute of the concept, value restriction of the attribute, process, relationship etc. Through the study to confirm the essential relation and subjection rapport between the concepts, to establish the complete structure of primary education field, which is conducive to obtain, analyze and use the knowledge.

In this paper, we defined primary education ontology as a five tuple $\{N,F,A,R,T\}$, where N is name of education ontology, F is parent ontology name of N, A is attributes set of ontology, R is ontology network, composed by basic relationship between ontologies, T is a set of tag that attached to the ontology.

2.2 Requirement Elicitation of Ontology

Based on the construction of primary education ontology, this paper use the classification system of education theme for reference, design ontology structure from the primary education ontology itself needs. The goal is to be able to describe knowledge structure of primary education field by providing primary education ontology, provide knowledge expression and reasoning support for primary education informatization.

Based on the above understanding, at the same time, considering the knowledge characteristics of primary education field, this paper divided primary education ontology into three major categories which are education activities, education subject, education architecture, then, divided the class into two level. For example, education activities can be divided into the education management

activities, teaching activities; education subject can be divided into some branches, such as Chinese, Mathematics, English; education architecture is divided into various types of education and all levels of education. The rest can be done in the same manner, concept can be subdivided step by step, thus forming the concept hierarchy of primary education ontology. Concept hierarchy of primary education ontology is shown in Figure 1.

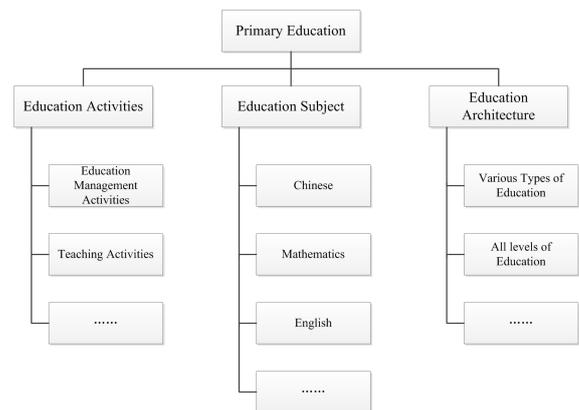


Figure.1 Structure of Primary Education Ontology

2.2.1 Name of Ontology and Father Ontology

Through the above analysis and the test of practical application, concept that appeared in the hierarchy of primary education ontology become the name of ontology and father ontology.

2.2.2 Attribute Sets

In this paper, attribute is relative to the relation of ontology. As shown in Figure 1, sub ontology can be seen as one of parent ontology properties. The attribute definition of primary education ontology not only fully express the relationship and characteristics of the education concept, but also should not be too complicated. Although the complex attribute definition has strong expressiveness, it also bring difficulty to knowledge sharing, and knowledge sharing is the original intention of the construction of primary education ontology. Considering the meaning of the concept and the simple relationship, more specific concept needs to choose as the attribute.

2.2.3 Ontology Network

At the beginning of this section mentioned that there are four relations between central nodes and sub nodes of mind mapping. They are inheritance, inclusion, part and whole, attribute. To a certain extent, there are some similarity between the relationship of ontology and mind mapping. Ontology has four basic relationship, they are kindOf, partOf, instanceOf and attributeOf. KindOf express the relation of category and predictability between ontologies,

the equivalent of inheritance in object-oriented, as defined in ontology: if X is a subclass of Y, each instance of X at the same time is the instance of Y, such as the education management activities is the lower class of the education activities. KindOf relation is transitive, if X is a subclass of Y, Y is a subclass of Z, then X is a subclass of Z. PartOf express the relation of part and whole between ontologies, part and whole relation is also transitive. InstanceOf express the relation between instance and ontology, similar to relationship with an instance of a class in object-oriented[3], such as the specific subject is the instance of curriculum. AttributeOf express the attribute of ontology, it can express attribute relationship between ontologies, such as age and name are the attributes of student. Based on the four basic relations, we can carry on the expansion and deletion. There will be a cross between these relations, and then forming ontology network.

2.2.4 Tag Sets

Usually, there are some phrases or sentences between the nodes of mind mapping, their role is to describe the relationship between two nodes or is to make sign for sub node. Tag is the key word that marked product target and classification or content. Tags can classify the ontology, usually mark multiple category to different ontologies, point out affiliation between different ontologies. In addition, tag describes the key attribute of ontology, more importantly, social attribute of tag consistent with original intention that constructed the primary education semantic ontology library to realize knowledge sharing. In this paper, the tag is based on attribute significance of ontology.

3 Construction of Primary Education Ontology Library

Ontology learning is the way that obtain domain knowledge from existing information sources, and automatically construct or extend ontology.[4] Now, there is no mature scheme for ontology learning, manual method of ontology construction is time-consuming, therefore, the way of construction by automated means, using semi manual mode to realize primary education ontology library is an effective compromise.

Machine learning and statistical techniques are the main automation means of ontology learning, for the processing of structured, semi-structured or unstructured data through these methods, domain knowledge is acquired from the data. Structured data mainly come from the database, and the XML file; the main source of semi-structured and unstructured data is HTML Webpage, plain text files, electronic documents, etc. The main method of processing structured data is mapping model to the ontology structure, for example, mapping the relation model of relational database to ontology structure[5]; for semi-structured or

unstructured data, using machine learning and natural language analysis to obtain the implicit domain knowledge.

For the construction of primary education semantic ontology, a more convenient way is to make use of education thesaurus, education dictionary and corpus set of primary education field for machine learning, to get the knowledge for constructing ontology, and then manually construct primary education ontology. The specific constructing process is shown in Figure 2.

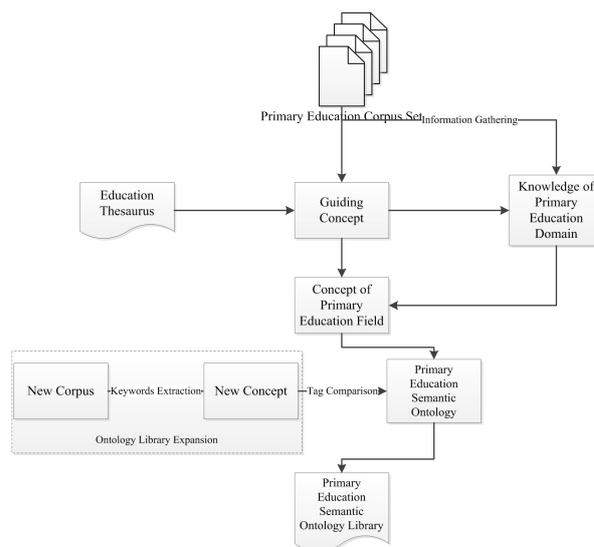


Figure.2 Flow Chart of Constructing Primary Education Semantic Ontology

3.1 Collection and Acquisition of Primary Education Domain Knowledge

Primary education domain knowledge is easy to understand, but it is rich and wide in content. It is important to collect the relevant knowledge of primary education domain before construct primary education ontology. Part concept of education thesaurus and faceted classification relationship of word are some basis for this research; basic knowledge of primary education is an important supplement; tacit knowledge of authoritative expert in primary education domain is key to select concepts of primary education ontology and to delineate the relationship. In addition, the relevant documents for constructing domain knowledge otology, the existing rules, the primary education major literature in network and the books of primary school are object of primary education domain knowledge for collection and acquisition.

3.2 Concept Selection

The main task of this step is to choose the basic vocabulary of terms for primary education domain, to accurately define concept of vocabulary. The core concept that extended from top to bottom form reasonable and complete concept

hierarchy. The choice of primary education vocabulary and the definition of hierarchy usually needs to draw on existing thesaurus, professional dictionary, dictionary, primary education professional books and terminology. In addition, original concept can be got through the Chinese word segmentation. For example, a definition of the parallelogram which means is "The four sides of the diagonal are parallel to each other." could get the following string by word segmentation: "diagonal / mutual divide /the/ quadrilateral // parallelogram ". "the " and "is " is stop word that need to be give up. "diagonal", "mutual divide ", "quadrilateral ", "parallelogram " these four phrases, namely original concept, need to be retain.

3.3 Establishing Attribute Relationship

The Attributes and relationships are the foundation of constructing primary education ontology.[6] In this step, the selected concept vocabulary need to be standardly defined according to the basic relationship that mentioned above section. These logical relationships arranged in a crisscross pattern, so as to form a stereo and intuitive semantic net. For example, the relationship among "diagonal line ", "Divide equally between each other ", "Quadrilateral", "Parallel quadrilateral " these four concepts is shown in table 1.

Table.1 Example of the relationship between the concepts

Concept	Relationship
diagonal line	attribute Of
Divide equally between each other	attribute Of
Quadrilateral	—
Parallel quadrilateral	kind Of

These relationships can simply construct part ontology of parallelogram.

3.4 Adjunction of Semantic Tags

In this paper, tag that has rich semantic information is used to represent the similarity between ontologies. These information form a tag similarity calculation method that based on hierarchical structure of ontology. In the following, we collectively named the labels that generated by upper ontology as tag in the absence of ambiguous cases. For example, in Figure 3, f is the upper ontology of a, a is the lower ontology of f, so f make a mark to a.

Figure 3 take solid shape for instance to show hierarchical structure of ontology and tag.

Because of hierarchy of tag and multi-tag of ontology, in using the tag hierarchy to measure the similarity between two ontologies should also satisfy some properties, in this paper, we use sim to represent the similarity.

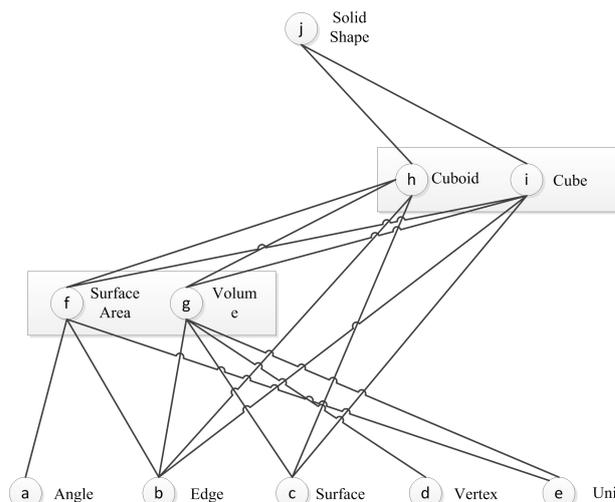


Figure.3 Hierarchical Structure of Tags

Character 1: expansion. If the two ontologies only share a tag, then the tag marked more ontologies, the similarity of these two ontologies is lower. For example, in Figure 3, a and b only share tag f, b and e only share tag g, of which three ontologies have been annotated with f, but there are four ontologies have been marked the tag g, so you can have this conclusion: $sim(a, b) > sim(b, e)$.

Character 2: distance. If common tag of two ontologies close to them, then the similarity of the two ontologies is more than the similarity of the two ontologies which public tag far from them. As shown in Figure 3, the common tag f of ontology a and b is the sub node of tag h which is public tag of ontology b and c, so $sim(a, b) > sim(b, c)$.

Character 3: binding. If the two ontologies have a plurality of public tags, then the similarity of these two ontologies is more than the similarity of these two ontologies which have a single common tag, is smaller than the summation of each tag similarity. As shown in Figure 3, b and c have a common ontology tag g and h, so $max(sim(b,c|T) \leq sim(b,c) \leq \sum T sim(b,c|T)$, where T refers to the common tag set {g, h}. Tag can be regarded as brief description of some attribute of ontology, two ontologies have public tag means they have same attributes. More public tags they have, more similarity of attributes they have. However, the attributes of each tag may be repeated, the actual similarity is smaller than calculation of each tag.

3.5 Extension Methods of Ontology Library

During or after the construction of ontology library, there will be a larger space to supply and improve. Quickly increasing vocabulary, the development of semantic technology and subdivision of primary education domain ontology type ask us to construct semantic ontology with considering to extend the ontology library.

3.5.1 Keywords Extraction

There are many words to constitute new corpus. Vector dimension of text is very large if see all words as keywords without distinguishing. This will not only have an impact on the operational efficiency, but also make the key features are submerged in the non-key words. Therefore, keywords of categories can be accessed by machine learning, and to reach the purpose of increasing the discrimination and reducing dimensions.

Keyword extraction can use the word frequency statistics to complete.[7] First, select the text that matches theme features as the corpus, randomly select a part of corpus as training sample, the rest is test sample; then use Chinese segmentation to divide training sample, and make frequency statistics of the training samples. Vast majority of text are likely to appear stop word, it will produce "noise " effect on the automatic text categorization. Therefore, stop word need to be filtered in the preprocessing stage. After making word frequency statistics and removing stop words, vocabulary weight can be calculated by using the TF.IDF method. Finally, word frequency is sorted according to the calculation results of TF.IDF. Part of vocabulary with high weight value are selected as keywords. The number of keywords need to be set initial value, then test it by using test samples, select the best value according to the repeated experimental results.

3.5.2 Tag Comparison

After knowledge collection, concept selection, establishing attribute relationship and adjunction of semantic tags, the corpus that formed by extracting keywords become a new ontology. Then, to add the ontology into ontology library needs to use the way of tag comparison.

The way that presented in reference[8] can almost meet the three properties which are proposed in section 3.4. This method is concept similarity calculating method based on ontology.[9] The formula of this method is as follows, assuming that the node p is the parent node of node c in ontology, the distance between node c and node p is calculated for:

$$wt(c,p) = (\beta + (1-\beta) \frac{E}{E(p)}) (\frac{d(p)+1}{d(p)})^\alpha [IC(c) - IC(p)] T(c,p) \quad (1)$$

d(p) means the hierarchy depth of node p, E(p) means the connection number of the child nodes, E represents the number of average child link nodes in the whole hierarchy, T(c,p) represents a connection type, α and β control the effects that made by depth and local density of the node on edge weight value. The information entropy of each node is calculated as $IC(c) = -\log P(c)$, where P(c) refers to the occurrence probability of node c and all its child nodes.

Reference[8] said that using shortest distance reciprocal of public tags to represent the similarity of ontology c1 and

c2, i.e. $sim(c1,c2) = 1/\min pdt(c1,p,c2)$. The range of similarity between [0,1].

In this paper, tag comparison is finished by similarity computation of tag. If similarity of the two tags is greater than 0.7, we believe that these two tags are similar, then the ontology of these two tags are similar. In reality, a ontology comprises a plurality of tags, so we need to calculate the similarity of all tags, and then determine the similarity by variance calculations of the value.

4 Conclusions and Future Work

To construct and improve the ontology is an iterative process, a method that easy to model and expand can guide the construction of ontology library. This paper propose a method to construct primary education semantic ontology library based on analysis of primary education domain knowledge, which take hierarchical structure of mind mapping for model, depend on and reference to the existing method of ontology construction with some concept relationship of education thesaurus. This method can show up knowledge concept and relationship of primary education domain in the form of ontology library. Of course, any one of ontology that successfully constructed will have a larger space to supplement and improve. In a word, the construction of primary education library is a huge and tedious task, effective construction method and extended rules need to be further research and improve.

References

- [1] Tony. Buzan. Mind Mapping[M]. LISi, translate. Beijing: Writers Publishing House, 1999, (3):56.
- [2] Anonymous. Maricl, Inc.; Cutting-Edge Social Mind Mapping Web Site Marici.com Launches[J]. China Weekly News, 2008(1):121-122.
- [3] LIU Chun-lei. Study on Methodology for Subject Knowledge Modeling Based on ontology[D]. Chongqing: Chongqing University, 2008.
- [4] LIU Bai-song. A Study on Web-Based Domain Independent Ontology Learning [D]. Zhejiang: Zhejiang Universtiy, 2007.
- [5] CHA Song-li, MA Zong-ming, JIAO Xiao-long. Survey on methodology for constructing ontology based on structured information source[J]. Application Research of Computers, 2012, 29(7):2407.
- [6] HAN Ren, HUANG Yong-zhong, LIU Zhen-lin, et al. Research on OWL ontology building method[J]. Computer Engineering and Desing, 2008, 29(6):1398.
- [7] YANG Yin, DAI Bin. Chinese Keyword Extraction Method Based on Multi-Features[J]. Computer Application and Software, 2014, 31(11):110.
- [8] Sun Yuan-shuai. Recommendation Algorithms in the Big Data Era[D]. Fujian: Xiamen University, 2014.
- [9] Jiang J, Conrath D. Semantic Similarity Based on Corpus Statistics and Lexical Taxonomy[C]. In Proceedings of ROCLING'1997. New York: ACM Press, 1997:19-33.