PAPER

Visual Knowledge Structure Reasoning with Intelligent Topic Map

Huimin LU[†a)], ***Member***, **Boqin FENG**[†], ***and*** **Xi CHEN**[††], ***Nonmembers***

SUMMARY This paper presents a visual knowledge structure reasoning method using Intelligent Topic Map which extends the conventional Topic Map in structure and enhances its reasoning functions. Visual knowledge structure reasoning method integrates two types of knowledge reasoning: the knowledge logical relation reasoning and the knowledge structure reasoning. The knowledge logical relation reasoning implements knowledge consistency checking and the implicit associations reasoning between knowledge points. We propose a Knowledge Unit Circle Search strategy for the knowledge structure reasoning. It implements the semantic implication extension, the semantic relevant extension and the semantic class belonging confirmation. Moreover, the knowledge structure reasoning results are visualized using ITM Toolkit. A prototype system of visual knowledge structure reasoning has been implemented and applied to the massive knowledge organization, management and service for education.

key words: *topic map, intelligent topic map, knowledge reasoning, knowledge visualization*

1. Introduction

Knowledge doesn't exist by itself, since knowledge is always related to other knowledge. According to the Constructivism Theory and Cognitive Load Theory perspective, the inherent relationships which are mutual complementation, explanation, and reinforcement of knowledge can contribute to achieving consistent with navigation features of human cognition learning, thereby enhancing the cognitive efficiency of knowledge [1]. The knowledge structure reasoning constructs knowledge based on some data structure (e.g., vector space, tree, graph, etc.). It mainly selects the appropriate search strategy and matches the pattern. The knowledge structure reasoning bodes well for knowledge and the relationships between them, but the lack of structure constraint. Knowledge reasoning cannot be guaranteed to be as effective as logical representation. The knowledge logical relation reasoning is often used to describe knowledge representation and reasoning based on the logical relation. It is rigorous, flexible and with a strict formal definition. A knowledge representation model should be built to integrate these two types of knowledge reasoning, in order to obtain the satisfactory knowledge reasoning results [2].

Intelligent Topic Map (ITM) [3], [4] extends the conventional Topic Map (TM) [5], [6] in structure and enhances its reasoning functions. We define a clustering level above the topic level. Furthermore, a knowledge element level is inserted above the resource level. Each cluster contains some closely related topics. A knowledge element is the smallest unit of anatomic, explicit, formally defined knowledge content, a record of some form of externalization viewed as a single organized unit both from a conceptual and from a technical perspective. It is composed of a grouping of formatted information objects which can not be separated without substantial loss of meaning together with meta-data describing the element [7] (e.g., definition, theorem, algorithm, etc.). There are some interdependent relationships (e.g., preorder, postorder, example, reference, cause, etc.) with implicit manner in the internal resource file or between the resource files. Modern Cognitive Science insists that the storage structure of the human brain is a network of knowledge structure. When human analyzing and solving the problems, they search the corresponding knowledge on this network according to the inherent relationship of knowledge, instead of searching all the knowledge orderly [8]. Knowledge elements can help users to access to more detailed knowledge information and provide knowledge elements navigation for e-learning. ITM establishes a novel knowledge logical organization which organizes knowledge from four levels: cluster level, topic level, knowledge element level and resource level. It constructs multi-granularity knowledge representation architecture which includes clusters, topics, knowledge elements, associations, and occurrences.

We extend the syntax and semantics of XTM (XML Topic Maps) [9] so that it can describe clusters, knowledge elements, the association between knowledge elements, the relation between topic and knowledge element. It not only supports the topics navigation by semantic relation, but also supports the knowledge elements navigation based on the inherent relationship according to the Cognitive Science. The Extended XTM (EXTM) provides a model and grammar for representing the structure of ITM and defining the reasoning rules. EXTM makes XML extend to the semantic field further, which defines an abstract, graphics-based knowledge association model and allows logical reasoning to discover new knowledge.

According to ITM, we propose a novel method for knowledge reasoning, which can efficiently implement both the knowledge structure reasoning and the knowledge logi-

Manuscript received January 8, 2010.
Manuscript revised May 25, 2010.

[†]The authors are with the School of Electronic and Information Engineering, Xi'an Jiaotong University, No.28 Xianning West Road. Shaanxi 710049, China.

[††]The author is with the Department of Machine Learning, School of Computer Science, Carnegie Mellon University, Pittsburgh 15213, USA.

a) E-mail: luhm.cc@gmail.com

DOI: 10.1587/transinf.E93.D.2805

cal relation reasoning. The knowledge logical relation reasoning offers a set of typical inference services, such as the satisfiability of knowledge points, the implicit associations reasoning between knowledge points, knowledge consistency checking, etc. The meaning of "knowledge point" in this paper is topic or knowledge element. The knowledge structure reasoning can implement the semantic implication extension (e.g., when querying "Operating System", it also can obtain the subclass belonging to this topic such as "UNIX"), the semantic relevant extension (e.g., when querying "UNIX", it also can obtain the relevant topic such as "Windows", "Linux") and the semantic class belonging confirmation (e.g., when querying "Microsoft Word", it also can obtain the class belonging topic such as "Office Software"). The knowledge structure reasoning results are visualized using ITM. It provides a visual knowledge map, which is available for users to acquire the knowledge and associations among them. Visual navigation tools capable of exploiting the created knowledge structures are based on hyperbolic geometry concepts and provide users with intuitive access mechanisms to the required knowledge [10].

2. ITM Description

ITM establishes a novel multi-source knowledge organization which depicts the hierarchical relationship of "cluster - topic - knowledge element - occurrence". The structure of ITM is shown in Fig. 1 [11]. ITM provides strong paradigm and concept for the semantic structuring of linked networks. It can establish the relationship among unstructured information resources, thereby allowing to link heterogeneous, unmodified resources of information semantically by creating a semantic web and implement concrete objects to be joined with abstract concepts. Knowledge logical organization based on ITM adapts to human's own cognitive pattern, it lays a foundation for high-quality knowledge structure reasoning and points a new direction for knowledge reasoning. ITM is a technology for encoding knowledge and connecting this encoded knowledge to relevant information resources, it is used as a formal syntax for representing and implementing ontologies. We define an ITM Model (ITMM) as following eleven tuples:

$ITMM = (C, T, KE, O, AT, AKE, \alpha, \beta, \theta, \varphi, \gamma)$

$C = \{c_1, c_2, \cdots, c_k\}(k > 0)$, C denotes a non-empty finite set of clusters. $T = \{t_1, t_2, \cdots, t_n\}(n > 0)$, T denotes a non-empty finite set of topics. $KE = \{ke_1, ke_2, \cdots, ke_m\}(m > 0)$, KE denotes a non-empty finite set of knowledge elements. $O = \{o_1, o_2, \cdots, o_z\}(z > 0)$, O denotes a set of information resources. $AT = \{at_1, at_2, \cdots, at_x\}(x > 0)$, AT denotes a set of topic association types. $AKE = \{ake_1, ake_2, \cdots, ake_y\}(y > 0)$, AKE denotes a set of knowledge element association types. $\alpha \subseteq (T \times T \rightarrow AT)$, α denotes the associations between topics. $\beta \subseteq (KE \times KE \rightarrow AKE)$, β denotes the associations between knowledge elements. $\theta \subseteq (C \times T \rightarrow \{0, 1\})$, θ denotes the relations between cluster and topic. $\theta(c_i, t_j) = 1(c_i \in C, t_j \in T)$ denotes c_i including t_j, $\theta(c_i, t_j) = 0$ denotes c_i not including t_j. $\varphi \subseteq (T \times KE \rightarrow \{0, 1\})$, φ denotes the relations between topic and knowledge element. $\varphi(t_i, ke_j) = 1(t_i \in T, ke_j \in KE)$ denotes t_i involving ke_j, $\varphi(t_i, ke_j) = 0$ denotes t_i not involving ke_j. $\gamma \subseteq (KE \times O \rightarrow 0, 1)$, γ denotes the relations between knowledge element and resource. $\gamma(ke_i, o_j) = 1(ke_i \in KE, o_j \in O)$ denotes o_j including ke_i, $\gamma(ke_i, o_j) = 0$ denotes o_j not including ke_i.

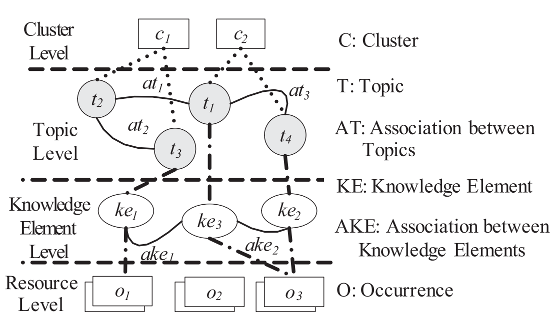


Fig. 1 The structure of intelligent topic map.

3. Basic Principle

The knowledge structure reasoning includes four parts: knowledge consistency checking, implicit associations reasoning between knowledge points, knowledge structure reasoning and visualization of reasoning results. First, knowledge consistency checking can eliminate knowledge redundancies, contradictions and mistakes. It can help us obtain the optimal description of ITM. Second, the implicit associations reasoning between knowledge points based on rules can help us obtain new knowledge. Third, the knowledge structure reasoning is the main stage, which returns all the knowledge elements, topics, cluster, and occurrence which are associated with the knowledge point within a certain knowledge radius. Finally, the visual knowledge map based on ITM is constructed.

3.1 Knowledge Consistency Checking

In the process of ITM constructing, conflicts can be caused by many reasons, like the differences of people's understanding, the annotation of knowledge resources, and the constructing of knowledge organization. These conflicts cause information redundancies, contradictions and mistakes, and lead to inconsistencies in knowledge reasoning. Knowledge consistency checking is a key part of knowledge reasoning. It includes the reflexivity checking, loop transitivity checking, knowledge redundancy checking and knowledge contradiction checking.

- Reflexivity checking. We define the rule as follows: Rule 1: If a topic (or knowledge element) is associated with itself, there exists reflexivity conflict. It is described as follows:
  Topic reflexivity:

$$\exists t \in T, t\ AT\ t \tag{1}$$

Knowledge element reflexivity:

$$\exists ke \in KE,\ ke\ AKE\ ke \tag{2}$$

For example, "$subClassOf$" denotes that a topic is a subclass of another topic. If $subClassOf(t_1, t_1)$, there is a semantic error for topic t_1 is the subclass of t_1. When the reflexivity conflict is detected, the association between the same topics (or knowledge elements) would be deleted.

- Loop transitivity checking. We define the rule as follows:
  Rule 2: If there is an association loop between the two directly related topics (or knowledge elements), there exists a loop transitivity conflict.
  Topic transitivity:

$$\exists t_1 \in T,\ \exists t_2 \in T,\ t_1\ AT\ t_2\ \wedge\ t_2\ AT\ t_1 \tag{3}$$

Knowledge element transitivity:

$$\begin{aligned}&\exists ke_1 \in KE,\ \exists ke_2 \in KE,\\ &ke_1\ AKE\ ke_2\ \wedge\ ke_2\ AKE\ ke_1\end{aligned} \tag{4}$$

As the same way, there is a semantic error if $subClassOf(t1, t2)$ and $subClassOf(t2, t1)$. When the transitivity conflict is detected, one of the associations between the topics (or knowledge elements) would be deleted.

- Knowledge redundancy checking. There exists redundancy if have the same topics (or knowledge elements) in an ITM.
  Topic redundancy:

$$\exists t_1 \in T,\ \exists t_2 \in T,\ t_1 = t_2 \tag{5}$$

Knowledge element redundancy:

$$\exists ke_1 \in KE,\ \exists ke_2 \in KE,\ ke_1 = ke_2 \tag{6}$$

Though knowledge redundancy is not a mistake on semantics, it would be solved when it is detected for ensuring certainty and uniqueness. Given an ITM, topics t_1 and t_2 (or knowledge elements ke_1 and ke_2), $\forall t_1, t_2 \in T,\ \forall ke_1, ke_2 \in KE$:
Step 1: Finding the same topics (or knowledge elements). We adopt a similarity measure algorithm for topics (or knowledge elements) which called Subject Identity Measure (SIM) [12]. This algorithm describes how similar the related topics (or knowledge elements) are. It is used to calculate the syntactic similarity by analyzing the character composition of topics (or knowledge elements). For a topic pair (t_1, t_2), we calculate the similarity as follows:

$$SIM(t_1, t_2) = \frac{2c}{|t_1| + |t_2|} \tag{7}$$

The c denotes the number of characters of the largest common substring contained in two topics. We can calculate how many words are matched between t_1 and t_2. If the value of $SIM(t_1, t_2)$ is higher than a threshold, it would be considered that t_1 is same as t_2.
Step 2: Merging the same topics (or knowledge elements). If t_1 has high similarity with t_2 in ITM, they would be merged into a single one. When two topics are merged, the association merging would be considered, e.g., an association $A_a(t_1, t_3)$ exists between t_1 and t_3, an association $A_b(t_2, t_4)$ exists between t_2 and t_4. If t_1 has high similarity with t_2, the merged topic t_1 would have two associations, i.e., $A_a(t_1, t_3)$ and $A_b(t_1, t_4)$. The same is true for knowledge elements.

- Knowledge contradiction checking. Knowledge contradiction is a logical error, e.g., $preorderOf(ke_1, ke_2)$ and $postorderOf(ke_1, ke_2)$ are contradictive each other. When knowledge contradiction conflicts are detected, the contradictive associations would be deleted.

Through knowledge consistency checking, we can obtain an ideal ITM description. It lays a foundation for the knowledge structure reasoning.

3.2 The Implicit Associations Reasoning

The implicit associations reasoning can discover the new associations between knowledge points. ITM contains an abundance of association types. In this paper, we mainly discuss the association of topics, e.g., $subClassOf$, $instanceOf$, $memberOf$, and the association of knowledge elements, e.g., $preorderOf$ and $postorderOf$.

- $subClassOf$. It is a typical binary association between two topics where one topic is a subclass of another. For example, $subClassOf(t_1, t_2)$ indicates topic t_1 is a subclass of t_2, t_1 is called sub-topic and t_2 is called parent-topic. Knowledge reasoning rules based on $subClassOf$ is as follows:
  Rule 3: Transitivity rule. If t_a is a subclass of t_b, t_b is a subclass of t_c, then t_a is a subclass of t_c.

$$\begin{aligned}&subClassOf(t_a, t_b) \wedge subClassOf(t_b, t_d)\\ &\rightarrow subClassOf(t_a, t_d)\end{aligned} \tag{8}$$

Rule 4: Attribute inheritance rule. "Attribute" means the inherent characteristic of a certain class of topics owned. If t_a is a subclass of t_b, t_b has an attribute A, then t_a has an attribute A.

$$\begin{aligned}&subClassOf(t_a, t_b) \wedge HasAttribute(t_b, A)\\ &\rightarrow HasAttribute(t_a, A)\end{aligned} \tag{9}$$

Rule 5: Property inheritance rule. "Property" means the characteristic of a certain class of topics owned, which is different with other class of topics. If t_a is a subclass of t_b, t_b has a property P, then t_a has a property P.

$$\begin{aligned}&subClassOf(t_a, t_b) \wedge HasProperty(t_b, P)\\ &\rightarrow HasProperty(t_a, P)\end{aligned} \tag{10}$$

Rule 6: Instance inheritance rule. If t_a is a subclass of t_b, e is the instance of t_a, then e is the instance of t_b.

$$subClassOf(t_a, t_b) \wedge instanceOf(e, t_a) \rightarrow instanceOf(e, t_b) \tag{11}$$

We can judge whether there is a *subClassOf* association between two topics and check the topic's instance, and so on.

- *instanceOf*: for the topic t and its instance set St, the association between c ($c \in St$) and t is called *instanceOf* association. $instanceOf(c, t)$ denotes c is an instance of t. Knowledge reasoning rule based on *instanceOf* is as follows: Property Inheritance:

$$instanceOf(c, t) \wedge HasProperty(t, P) \rightarrow HasProperty(c, P) \tag{12}$$

The *instanceOf*-based knowledge reasoning is implemented by the inheritance.

- *memberOf*: it represents the association between the member M and the object set W. $memberOf(M, W)$ denotes M is a member of W. *memberOf* and *instanceOf* are two kinds of completely different associations. The *memberOf* does not has transitivity, attribute inheritance and property inheritance, it emphasizes on the association between topics.
- *preorderOf*, *postorderOf*: they reflect some inherent relationships of knowledge from the cognitive perspective for e-learning. For example, users should learn the definition of "Angle" first, and then learn the definition of "Triangle". The *preorderOf* represents that a knowledge element A is produced before another knowledge element B, denoted as $preorderOf(A, B)$. The *postorderOf* represents that A is produced after B, denoted as $postorderOf(A, B)$. Knowledge reasoning rules based on the *preorderOf* and *postorderOf* associations are as follows:
Transitivity:

$$preorderOf(A, B) \wedge preorderOF(B, C) \rightarrow preorderOF(A, C) \tag{13}$$

$$postorderOf(A, B) \wedge postorderOF(B, C) \rightarrow postorderOF(A, C) \tag{14}$$

Inverse relation between *preorderOf* and *postorderOf*:

$$preorderOf(A, B) \rightarrow postorderOF(B, A) \tag{15}$$

$$postorderOf(A, B) \rightarrow preorderOF(B, A) \tag{16}$$

Besides the above association types, there are some other inherent relationship between knowledge elements, e.g., *causeOf*, *referenceOf*, *exampleOf*, etc. We implement the implicit associations reasoning based on EXTM.

3.3 The Knowledge Structure Reasoning

Since knowledge is highly correlated with each other, we implement the semantic implication extension, the semantic relevant extension and the semantic class belonging confirmation in order to acquire the complete knowledge structure. Moreover, we try to better reflect the relations of level and class structure from the results. The appropriate search strategy such as breadth-first search, depth-first search, depth-first heuristic strategy is selected, and the corresponding pattern is matched in knowledge structure reasoning. According to the characteristics of ITM, we propose an extended algorithm based on knowledge unit circle, named Knowledge Unit Circle Search (KUCS) strategy. Before discussing what can be reasoned based on knowledge structure in ITM, we would like to define three concepts, i.e., knowledge path, knowledge radius and knowledge unit circle.

- ***Definition 1:*** Knowledge path. In ITM, if there is a sequence $C_p, C_1, C_2, \cdots, C_m, C_q$, and there are association between $(C_p, C_1), (C_1, C_2), \cdots, (C_m, C_q)$ respectively in ITM, then we said that there exists a knowledge path between concept C_p and C_q. C represents topic or knowledge element. Association between (C_i, C_j) denotes that C_i is directly related to C_j.
- ***Definition 2:*** Knowledge radius. A knowledge path is a sequence of consecutive topics (or knowledge elements) in ITM, and the knowledge radius is the number of topics and knowledge elements traversed in a knowledge path, i.e., the length of the path.
- ***Definition 3:*** Knowledge unit circle. Knowledge unit circle is all the topics and knowledge elements which are associated with the knowledge point, when the knowledge radius is equal to 1.

KUCS is described as follows:
Input: ITM, knowledge point t (assuming that t is a topic) and knowledge radius R.
Output: All the topics, knowledge elements, cluster, and occurrence which are associated with t.
Step 1: Searching knowledge point t in ITM, knowledge radius variable r is equal to 1.
Step 2: Searching all the topics which are associated with t, the knowledge radius is equal to 1. The results are stored in $setT_1$, which is a $HashSet$.
Step 3: Searching all the topics which are associated with topics in $setT_1$, and then using them as the center of knowledge unit circle respectively, the knowledge radius is equal to 1. The results are stored in $setT_2$. Increasing r by 1.
Step 4: Continue searching new topics which are associated with the topics in $setT_2$, and then using them as the center of knowledge unit circle respectively, increase r by 1, until r is equal to R.
Step 5: Searching all the knowledge elements which are associated with the topics acquired, and stored in $setKE$.
Step 6: reasoning the associations between topics and the associations between knowledge elements.
Step 7: Searching the cluster which is associated with the topics, and the occurrences which are associated with the knowledge elements, and then constructing the visual

knowledge map based on ITM finally.

Through the knowledge structure reasoning, we can obtain all the knowledge elements, topics, cluster, and resource occurrence which are associated with the knowledge point within a certain knowledge radius.

3.4 The Visual Knowledge Map Constructing

The tools of TM editing and navigating are all based on TM model thus still cannot support clusters and knowledge elements editing and navigating. Based on the ITM logical representation of knowledge, ITM toolkit is designed. It is coded in Java which to assist users in sharing and navigating the domain knowledge for e-learning (Toolkit uses Tough Graph as the graphic engine). ITM toolkit includes three layers. The function layer directly interacts with EXTM documents. The logic layer is responsible for handling the logic of user's actions. The user interface layer is served as Graphic Interface which listens to user's actions and notifies the proper module in logic layer. There are three critical issues to be considered in designing the ITM Toolkit [13].

- ***Information overload mitigation.*** The ITM document is visually displayed as a double-layer network. As shown in Fig. 4, clusters, topics and topic associations are represented in the upper layer in which the light color nodes are regarded as topics. The deep color node is regarded as a cluster and each edge is regarded as an association of topics. When user clicking the edge, it will display the corresponding association type. Knowledge elements and their associations are in the lower layer in which the deep color node is regarded as an occurrence. When clicking the nodes in the knowledge element layer, it will display the topics which are associated with the knowledge element.
- ***Consistency checking.*** In order to maintain the consistency in the ITM, consistency checking is preformed at the end of each operation. There are the syntax checking and the semantics checking. Syntax checking mainly focuses on the correctness and completeness of input data. Semantics checking mainly focuses on the semantics in knowledge element associations.
- ***Concurrent update management.*** The ITM Editor accepts new versions of the EXTM and reproduces old versions on request. Concurrent update management of EXTM involves two main tasks, i.e., version recording and merging. Version recording records the generation tree of versions and merging integrates information from diverse sources into a coherent new topic map automatically.

4. Empirical Evaluation

To verify the efficacy of the knowledge structure reasoning, a set of experiments was conducted. We apply our method to a part of the knowledge domain of "Computer Network".

Table 1 Knowledge consistency checking results.

Checking item		Checking results	Times
Reflexivity checking		Interior gateway protocol, TCP protocol, etc.	4
Transitivity checking		Transport layer functions: network layer functions, Application layer protocol: HTTP, Internet interconnection protocol: IP, etc.	48
Redundancy checking	Topic(or knowledge element) redundancy	The main function of TCP protocol, TCP protocol function, protocol, Computer network, etc.	7
	Association redundancy	IP addresses composition: IP address property, Network: protocol, etc.	9
Contradiction checking		Network: service, Network: protocol, etc.	3

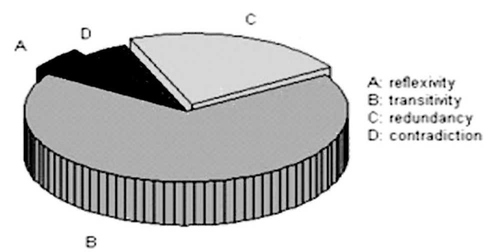


Fig. 2 The statistical analysis of knowledge consistency checking results.

The Experimental data includes 1394 topics, 3104 knowledge elements, 816 associations between topics, 906 associations between knowledge elements and 617 associations between topic and knowledge element.

4.1 Knowledge Consistency Checking Experiment

We implement the reflexivity checking, loop transitivity checking, knowledge redundancy checking and knowledge contradiction checking experiment respectively. Knowledge consistency checking results are shown in Table 1. The statistical analysis of knowledge consistency checking results is depicted in Fig. 2. It shows that the main conflict type is loop transitivity conflict, which makes up 60%-70% of total conflicts, knowledge redundancy conflict type makes up 20%-25% of total conflicts, and knowledge reflexivity conflict and knowledge contradiction conflict make up 10%-15% of total conflicts. Conflicts can be caused by many reasons. The ITM corpus construction is a process that needs many people's collaboration and many times of revision. Differences of the people's understanding for the relationships between topics (or knowledge elements) and the different input order of the data set are the basic cause of leading to the conflicts by analyzing the results of knowledge consistency checking experiment. In order for the local ITM to be reused, they first need to be fused or aligned to one another to produce a single integrated and reconciled global ITM that deals with a larger domain of interest. Knowl-

Table 2 The implicit association reasoning results.

Checking item	Checking results	Times
Super ordinate association between topics	Protocol: IPX, Transmission medium: Shielded Twisted Pair, Internet: IP, etc.	316
Preorder association between knowledge elements.	OSI network system structure: Network layer provides two types of services, The development of URL: URL syntax features, IP address definition: The difference between TCP protocol and IP protocol, etc.	421

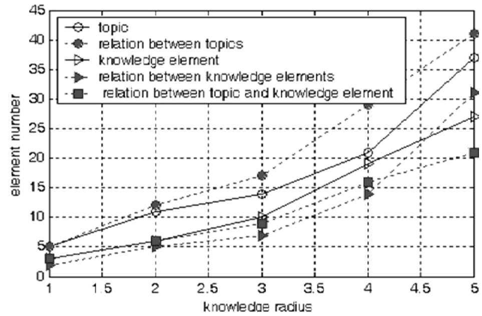


Fig. 3 The knowledge structure reasoning results.

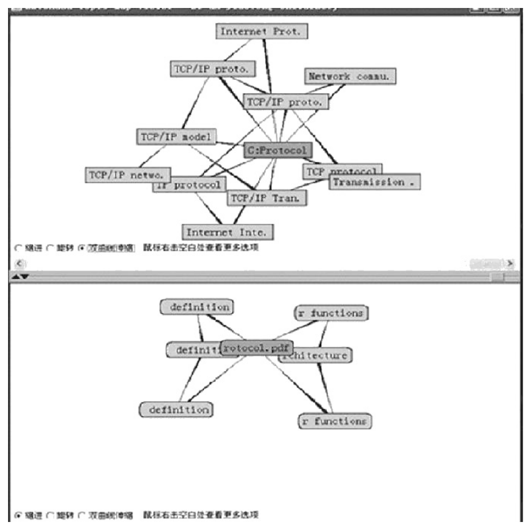


Fig. 4 The visual knowledge structure.

edge redundancy conflicts may exist in the process of ITM merging. Knowledge reflexivity conflicts are mainly caused by people's carelessness and knowledge contradiction conflicts are mainly caused by the differences of people's understanding. Conflicts detection and resolution is a key part of knowledge reasoning strategy.

We use performance measurement of information retrieval such as P (Precision) and R (Recall). We get true-positive set (TP) which includes correctly identified checking times, false-positive set (FP) includes false checking times, and false-negative set (FN) which includes missed checking times. We can evaluate the quality of automatic checking process by the following expression.

$$P = \frac{TP}{TP + FP}, \quad R = \frac{TP}{TP + FN} \tag{17}$$

As the results, P is equal to 71.2% and R is equal to 81.5%. In order to improve the precision and recall of consistency detection, the semantic matching between topics and between knowledge elements should be considered.

4.2 The Implicit Association Reasoning Experiment

The implicit Association reasoning can reason out the new associations between topics (or knowledge elements), and make knowledge structure with more detailed semantic associations. For example, we experimentalized the implicit associations reasoning functions based on the super ordinate association between topics and the *preorder* association between knowledge elements. The implicit associations reasoning results are shown in Table 2. It shows that the implicit associations reasoning can discover the implicit associations between topics (or knowledge elements). It provides inherent relevant characteristics of knowledge to constructing the complete knowledge structure, but we find that some reasoning associations between topics (or knowledge elements) are not tight enough. It is not conducive to the user's understanding. For example, we obtained the implicit association "Internet: IP", although there is a certain association between "Internet" and "IP", but the association is not tight enough. Next, we should consider the cross-correlation between topics (or knowledge elements) and knowledge radius.

4.3 The Knowledge Structure Reasoning Experiment

We select a topic "TCP/IP protocol" as knowledge point and different knowledge radius to carry out the knowledge structure reasoning experiment. It returns all the knowledge elements and topics which are associated with the knowledge point within a certain knowledge radius. The results are shown in Fig. 3. With the knowledge radius increasing, the number of topics, knowledge elements and relations increase continuously. When knowledge radius is equal to 2, the knowledge structure reasoning results include eleven topics (e.g., "IP protocol", "TCP/IP protocol", "TCP protocol", etc.) and five associations between the topics, six knowledge elements (i.e., "TCP protocol definition", "IP protocol definition", "TCP/IP protocol definition", etc.) and five associations between the knowledge elements, and six relations between the topic and knowledge element. The visual knowledge structure is depicted in Fig. 4. In order to improve the display speed, the user interface does not display the full name of topics (or knowledge elements). When user selects a topic (or knowledge element) by clicking the node on the topic (or knowledge element) layer, the corresponding full name will be displayed.

5. Relate Works

So far, the knowledge representation model which is able to integrate logical reasoning and structure reasoning includes XML, RDF, ontology, TM (Topic Map), etc. XML provides a flexible, general, rich structured information representation and convenient for the cooperative processing of heterogeneous knowledge, but it is powerless for heterogeneous semantic processing [14]. RDF uses resources as the center, and the relationship is one-way which must be point from subject to object. Moreover, the semantic expression is more complicated, the semantic level is primary, and the associated semantics expression is not rich enough. XML tags do not have any restrictions as well as the attribute set in RDF. XML and RDF can not handle the synonyms and the homographs. Ontology description language OWL has a strong ability of knowledge representation and reasoning. The grammar is independent, scalable and applicable to the knowledge representation of distributed systems [15], [16]. However, it is not in an intuitive and graphical way to display knowledge, and there is no relationship between the resources and the related concepts contained. OWL based on description logic. There are limitations in expressing general rules, and it can not properly express the user's preferences and constraints.

The structure of Topic Map composed of Topics, Associations and Occurrences (TAO) [17], which describes the concepts and the semantic relationships between them and can locate the resource which are associated with the concept [18], [19]. TM is known as "a bridge between information management and knowledge management", and "the GPS of information world". It can provide intuitive navigation of information resources. TM technology has attracted many scholars to research and explore. The research mainly involves three fields, i.e., basic research and standards, TM technology and tools, and TM practical application. Basic research and standards are used and extended. For example, TMDM (Topic Map Data Model) [20] defines a formal data model for TM, which will be used to define the XTM syntax. TMCL (Topic Map Constraint Language) [21] is a language for defining schemas and constraints on TM models. Specifically, TMCL can be used to constrain instances of TMDM. Clients can access the unified TM either by using the TMQL (Topic Map Query Language) [22] query interface or programmatically through a TMAPI [23] compatible interface. A number of tools that support flexible extendable architecture, visualization for interactive exploration and editing of TM, as well as implementations of constraint languages, query and reasoning are available [24]. Several tools on editing and navigating have been proposed such as TM4J (Topic Map for Java) [25], TM4L (Topic Map for e-Learning) [26] and UNIVIT (Universal Interactive Visualization Tool) [27], etc. TM is widely used in knowledge management, Web applications, semantic merging and other fields [28]–[31]. TM can be applied to cross-system since the XTM syntax is based on XML and is an exchangeable data standard. XML are all extended in the semantic field by RDF/OWL and XTM. The abstract and graphics-based correlation models are all defined and the reasoning measures are all allowed. However, RDF/OWL is mainly resource-oriented and is designed for the machine, while the TM is mainly user-oriented and is designed for human. The biggest advantage of RDF/OWL is reasoning, while the greatest advantage of TM is the discovery and visualization of knowledge architecture [32], [33].

6. Conclusion

The proposed visual knowledge structure reasoning model provides us with a means to organize, discovery and display knowledge. Visual knowledge structure reasoning based on ITM not only achieves better knowledge reasoning results and provides visual knowledge navigation, but also lays the foundation for high-quality knowledge value-added services. The ongoing work includes two aspects: (1) How to implement the knowledge reasoning for multi-source effectively? Cloud computing with huge computing ability and storage capacity suitable to manage massive, isomeric and distributed knowledge and realize knowledge reasoning. We will study the visual knowledge structure reasoning based on cloud computing. (2) Semantic matching method for knowledge consistency checking needs further improvement. Combing the Comprehensive Information Theory with the structure and semantic information of ITM, the similarity measure algorithm will be used, in which the syntactic matching, semantic matching and pragmatic matching are considered comprehensively.

Acknowledgements

This research was supported by the National High-Tech Research and Development Plan of China under Grant No.2008AA01Z131; The National Natural Science Foundation of China under Grant No.60803162.

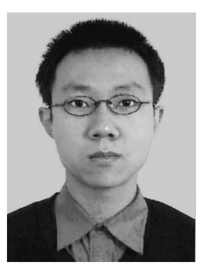
References

[1] J. van Merriënboer and P. Ayres, "Research on cognitive load theory and its design implications for e-learning," Educational Technology Research and Development, vol.53, no.3, pp.5–13, 2005.
[2] Q. Wang, L. Rong, and K. Yu, "Visual knowledge reasoning on typed categorical structure," Proc. 5th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD-08), pp.684–688, 2008.
[3] H. Lu and B. Feng, "An intelligent topic map-based approach to detecting and resolving conflicts for multi-resource knowledge fusion," Information Technology Journal, vol.8, no.8, pp.1242–1248, 2009.
[4] H. Lu and B. Feng, "Distributed knowledge integration based on intelligent topic map," Information Technology Journal, vol.9, no.1, pp.132–138, 2010.
[5] ISO/IEC 13250 Topic Maps Second Edition, Information Technology Document Description and Processing Languages, May 2002.
[6] ISO/IEC JTC 1/SC 34, ISO/IEC 13250-2: Information Technology-Topic Maps-Part 2: Data Model, http://www.isotopicmaps.org/sam/sam-model/data-model.pdf, 2008.

[7] R. Maier, Knowledge Management Systems, 3rd ed., Springer, Berline, 2007.
[8] J.R. Anderson, The Architecture of Cognition, Harvard University Press, Cambridge, Massachusetts, 1983.
[9] S. Pepper and G. Moore, XML Topic Maps (XTM) 1.0, http://www.topicmaps.org/xtm/1.0/index.html, 2002.
[10] S. Smolnik and I. Erdmann, "Visual navigation of distributed knowledge structures in groupware-based organizational memories," Proc. 6th International Conference on Information Visualization, pp.353–360, 2002.
[11] H. Lu, B. Feng, Y. Zhao, Q. Zheng, and J. Liu, "A new model for distributed knowledge organization management," Proc. 7th International Conference on Grid and Cooperative Computing (GCC-08), pp.261–265, 2008.
[12] L. Maicher and H.F. Witschel, "Merging of distributed topic maps based on the subject identity measure (SIM) approach," Proc. Berliner XML tags'04, pp.301–307, 2004.
[13] L. Jiang, J. Liu, Z. Wu, Q. Zheng, and Y. Qian, "ETM toolkit: A development tool based on extended topic map," Proc. 13th International Conference on Computer Supported Cooperative Work in Design (CSCWD-09), pp.528–533, 2009.
[14] C. Baru, A. Gupta, B. Ludäscher, R. Marciano, Y. Papakonstantinou, P. Velikhov, and V. Chu, "XML-based information mediation with MIX," ACM SIGMOD Record, vol.28, pp.597–599, 1999.
[15] P.A. Silva, C.M.F.A. Ribeiro, and U. Schiel, "Formalizing ontology reconciliation techniques as a basis for meaningful mediation in service related tasks," Proc. ACM first Ph.D. Workshop in CIKM'07, pp.147–154, 2007.
[16] J.L. Seng and I.L. Kong, "A schema and ontology-aided intelligent information integration," Expert Systems with Applications, vol.36, pp.10538–10550, 2009.
[17] S. Pepper, The TAO of topic maps-finding the way in the age of infoglut, http://www.gca.org/papers/xmleurope2000/papers/s11-01.html
[18] L.M. Garshol, What are topic maps, http://www.xml.com/pub/a/2002/09/11/topicmaps.html, Sept. 2002.
[19] S. Pepper and L.M. Garshol, The XML papers — Lessons on applying topic maps, http://www.ontopia.net/topicmaps/materials/xmlconf.html, ontopia.
[20] L.M. Garshol and G. Moore, JTC1/SC34, Topic maps-Data Model, http://www.isotopicmaps.org/sam/sam-model/, ISO 13250: Topic Maps, 2008.
[21] ISO, IEC, Topic Maps Constraint Language (ISO 19756), http://www.isotopicmaps.org/tmcl/, 2010.
[22] ISO/IEC 18048, Topic Maps Query Language (TMQL), http://www.isotopicmaps.org/tmql, 2008.
[23] L. Heuer and J. Schmidt, "TMAPI 2.0," Proc. Fourth International Conference on Topic Maps Research and Applications (TMRA-2008), pp.129–136, 2008.
[24] A. Papastergiou, A. Hatzigaidas, G. Grammatikopoulos, Z. Zaharis, P. Lazaridis, D. Kampitaki, and G. Tryfon, "Introducing an advanced Topic Map software tool towards the deployment of a TM-based system for managing melanoma cases images," WSEAS Trans. Information Science and Applications, vol.4, no.3, pp.452–459, 2007.
[25] TMNav 0.2.5, TM4J — Topic Maps 4 Java, http://tm4j.org/tmnav.html
[26] W. Dandan, D. Dicheva, C. Dichev, and J. Akouala, "Retrieving information in topic maps: The case of TM4L," Proc. ACM Southeast Regional Conference, pp.88–93, 2007.
[27] B.L. Grand and M. Soto, Information management — Topic Maps visualization, XML Europe 2000, Paris, France, June 2000.
[28] T. Schwotzer, "Modelling distributed knowledge management systems with topic maps," Proc. 4th International Conference on Knowledge Management (IKNOW-04), pp.1–9, 2004.
[29] R. Kannan, "Topic Map: An ontology framework for information retrieval," Proc. National Conference on Advances in Knowledge Management 2010, pp.195–198, 2010.
[30] I.X. Chen, C.Z. Yang, and T.L. Hsu, "Design and evaluation of a panoramic visualization environment on semantic web," Information and Software Technology, vol.48, no.6, pp.402–409, 2006.
[31] T. Neidhart, R. Pinchuk, and B. Valentin, "Semantic integration of relational data sources with topic maps," Proc. Fifth International Conference on Topic Maps Research and Applications (TMRA-09), pp.185–192, 2009.
[32] H. Lu, B. Feng, Y. Zhao, Q. Zheng, and J. Liu, "Distributed knowledge management based on extended topic maps," Proc. International Conference on Computer Science and Software Engineering (CSSE-08), pp.649–652, 2008.
[33] A. Korthaus, M. Aleksy, and S. Henke, "A distributed knowledge management infrastructure based on a topic map grid," Int. J. High Performance Computing and Networking, vol.6, pp.66–80, 2009.

Huimin Lu received the M.S. degree in Electronic and Informationl Engineering from Xi'an Jiaotong University, Xi'an, China, in 2005. She is a doctor candidate at the Department of Computer Science and Technology, School of Electronic and Information Engineering, Xi'an Jiaotong University, China. She is a student member of ACM, IEEE and CCF. Her current research interests include: knowledge management, knowledge fusion and knowledge service. She has published 12 papers in referred journals and International conferences.

Boqin Feng is a professor at the Department of Computer Science and Technology, School of Electronic and Information Engineering, Xi'an Jiaotong University, China. His current research interests include: knowledge management, knowledge service and data mining. He has published 52 papers in referred international conferences and journals.



Xi Chen received the M.S. degree from Carnegie Mellon University in 2008. He is currently a Ph.D. candidate at Machine Learning Department in the School of Computer Science, Carnegie Mellon University. His interests include knowledge reasoning, machine learning and texting mining.