On Formal Representation of Procedure Knowledge and its Application to Interactive Electronic Technology Manual

Zhu Jun-wu
Department of Computer Science, Yangzhou University, Yangzhou, China
College of Information Science and Technology, Beihang University, China
Email: jwzhu@yzu.edu.cn

Jiang Yi[1], Li Bin[1,2]
[1]Department of Computer Science, Yangzhou University, Yangzhou, China
[2]State Key Lab of Novel Software Tech, Nanjing University, Nanjing, China
Email: {jiangyi, lb}@yzu.edu.cn

***Abstract*—The Interactive Electronic Technology Manual is a core element of the equipment support, and using informal means to express automatic reasoning and consistency checking of knowledge systems that can not be carried through by process knowledge. This paper, in cognizance of the high frequency of procedure description in IETM, taking troubleshooting process as an example, puts forward a procedural representation based on labeling transition system, and realizes the XML file storage in procedure description. As a result, the machine can read and operate these data. Meanwhile, this paper designs a quantitative calculating method of process semantic proximity to provide a candidate set of semantic proximity sorting as the result of procedural knowledge retrieval. Compared with existing methods, this method uses status to express the program node information, characterizes change of state by action, and comprehensively considers human-computer interactions. The XML-based procedural representation and storage facilitate interoperability between multiple systems. The formal representation of process knowledge well characterizes the fault diagnosis process in the project of Interactive Electronic Technology Manual.**

***Index Terms*—Interactive Electronic technology Manual, Procedure, Knowledge Representation**

I. Introduction

Technical information is an important factor in equipment support. Owing to the complication of technology material systems, and expansion of the technical information number and diversity of users' demand, the creation and use of IETM (Interactive Electronic Technical Manual) requires intelligence, sharability, mutual operability and versatility. Effective organization of domain knowledge and provision of uniform shareable knowledge management platform will be the premise and foundation of IETM and other knowledge-based systems.

IETM appeared in the 1990s.Now it has become an important component of the strategy of CALS (Continuous Acquisition and Life-cycle Support), which implemented by the United States and many other developed countries. It is also a hot issue of research and application of equipment supporting information technology [1, 2]. In the areas of IETM research, expert systems, fault diagnosis and CAI courseware have achieved some functions of IETM [3, 4]. However, the biggest features of existing IETM are its functional unity and lacking wide-area resource sharing capabilities. Using data as its base unit, existing IETM has no extraction and representation of domain knowledge, can not achieve functions of machine learning and automatic reasoning, etc. Integrated and multi-functional IETM studies are still on the tentative research stage abroad. With the development of network technology, Inmedius-publisher, Northrop Grumman and other companies have tried the network publication of electronic technical data [5, 6]. They win the greatest benefit of a single data source, content filtering, functional integration, and data live updating.

It is clear that, due to the lack of a shared knowledge management platform, IETM development can not achieve knowledge reuse and sharing. Having nothing to do with specific application, creation of IETM-oriented domain knowledge base aims at obtaining the knowledge base of domain-dependent formalized representation, realizing knowledge consistency check and knowledge learning based on the reasoning machine. Under the premise of a unified knowledge, application program developing should be on the basis of the specific needs of users. In order to ensure machine readability and maneuverability of the knowledge base, the primary issue is to solve the issue of knowledge formalized representation. We should distinguish between the static domain terms and dynamic procedure knowledge, give knowledge model or domain ontology to establish domain ontology as a basic "skeleton" of the knowledge base, and then use the basic concepts and relationships of the ontology, and relationship between individuals, to build IETM-oriented domain knowledge base.

doi:10.4304/jsw.5.12.1402-1408

This paper aims at creating the knowledge base of interactive electronic technical manual. Section II presents the electronic technological machine's readability and maneuverability and current situation of creating the domain ontology. Section III presents a IETM system architecture based on knowledge base, and gives the acquisition and representation of the static knowledge. Section IV presents the representation strategy of the procedure knowledge in IETM, XML file storage and semantic retrieval problem based on process title in detailed. Finally, this paper gives a summary and a future prospect.

II. RELATED WORK

As the basis of implementation of the product life-cycle management, the Interactive electronic technology manual supports the using, training, maintenance, modification, testing and other product life-cycle phases. IETM is for the use of different levels of users at different stages, which determines that needs diversity must be taken into account during the IETM developing process.

IETM Interoperability is an important issue that must be addressed by developers. Literature [7] gives the definition of IETM interoperability of the United States Armed Forces in the following three areas: 1) the ability of application system connected together as a single entity for the system operation; 2) ability of communication between hardware and software; 3) ability of effective operation between applications, units and users. The U.S. military now realize interoperability between systems or users through a series of data standards. As the application of artificial intelligence technology in the IETM, particularly the fifth-level IETM multi-functional needs of integration, a simple data and information can no longer meet the system needs. Therefore, reasoning in fault diagnosis and expert systems must be knowledge-based. It is necessary to build a knowledge base for specific domains in order to achieve multi-integration of the fifth level IETM and meet the needs of training, diagnostics, maintenance and complex systems repairing.

The analysis of the definition of the United States Armed Forces IETM interoperability is: the solutions of 1) and 2) rely on development of computer hardware and network technology the solution of 3) involves a different system, units and users. Clearly, the application systems are man-made product; units are a certain set of users, developers, and computer systems. Therefore, as interoperability needs Semantic Sharing, the premise of solution to the problem 3) is to establish a unified platform for knowledge representation among different groups.

All this shows that it is critical to provide a consistent data among different applications and users to achieve the fifth level IETM tasks. Intuitively, it is very easy to achieve; in fact, ambiguity is inevitable during information exchange and translation between different users or applications because the lack of sufficient, accurate semantic interpretation and definition of terms in our Technical Manual. For example, different people may use the same concepts and terminology to describe different things. Also, different people may use different terminology or concepts to describe the same things. So, the key of sharing is the connotation rather than simply symbolic concepts.

Ontology can be shared. Standard is the center of the current IETM interoperability and sharing issues, which plays a normative role in a certain area to effectively achieve the data sharing and interoperability. Ontology provides a shared mechanism treating users and applications as the center, which does not force them to follow a large number of specifications. Therefore, we introduced ontology to the fifth-level IETM to address the knowledge sharing and interoperability issues.

In the Knowledge Engineering research area, non-standardization organizations of knowledge have gradually become another big "stumbling block" after "knowledge acquisition". Heterogeneous form of organization makes the knowledge can not be shared, re-used and integrated [8, 9]. Ontology is a terminology set which is organized by inheritance and describes a certain domain. It has been widely used in domain knowledge modeling. Formal semantic description of terminology enables knowledge sharing and reuses [10,11,12]. The creation of ontology is subjective and knowledge engineering proposed needs of ontology reuse and sharing. Therefore, ontology requires constant integration and evolution to achieve ontology amendment and knowledge base expansion [13,14]. In view of large-scale of knowledge engineering [15], to be more practical, significance ontology integration and self-evolution must be automatically. On the basis of IETM, this paper achieves intelligent IETM and other knowledge-based systems on a unified knowledge management platform.

III. KNOWLEDGE-BASED IETM SYSTEM ARCHITECTURE

A. System Architecture

Technical manual in the text knowledge is implicit, and inconvenient for our application. We humans can read and understand the knowledge in the text, but computers can not understand the text symbols. Only when implicit knowledge in the original text files are structured that can allow computers to "recognize" each piece of knowledge. Structured activities include:

Concept extraction Concept is a term of a certain field, and it is abstract, which is a generalization of some individual with common attributes. The concept in technical manual contains two classes: vocabulary of terms and process activity. For example, some vocabulary of terms, such as fuel injector, fire box and nut, and some process activity, such as inspection, maintenance and disassembly.

Concept description can be defined by category. Concept description can be realized through connotation depiction, while the attribute and relation is given. Generally attribute corresponds to nouns, and relation corresponds to verbs, which exists in different concepts.

Concept system can clearly express the relation between concepts, which is also an effective measure. The structure of concept system is a approximately treelike structure and there are some common relations that form concept relation, such as ISA、PART_OF and MEMBER_OF etc.

Knowledge carrier in technical data contains text and image, which are respectively processed by text knowledge editor and image semantic tagging. Text knowledge editor mainly realizes the concept of text and relation extraction, carrying out description, and finally term architecture is given on the basis of concept attribute description. In image semantic tagging, it expresses the content semantics of image that the concept of term architecture is taken as key words. IETM system architecture based on knowledge is shown by the below figure.

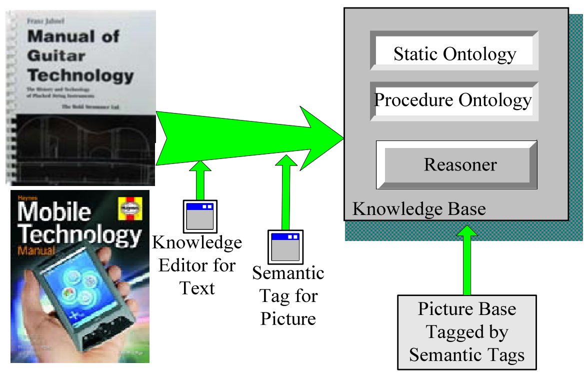


Figure1 The Architecture of Knowledge-based IETM

Taking some technical data as example, the text knowledge edit process is given as follows. According to attribute depiction of special tool in technical data, the below frame can be found, from text description to structure description.

Spare parts location function: (gasket (16), is pressed in the middle of bushing and nut to ensure tightness). Machine cannot directly understand natural language; however, natural language can be abstracted through mathematics and logical formula so that machine can process the formula. The abstract process is called formalization process.

Axiom in ontologies has the functions of limit and inference. Firstly, axiom with limit function possesses the constraint function for concept and attribute so that depiction and real world is closer. Secondly, in knowledge system, some knowledge is implicit, which can be found by inference and logical operators. Some axioms about tools with limit function are given as follows:

The transformation process from text knowledge to formalization knowledge is given above, and on the basis of it, automated reasoning is realized by constructing the formalization system. Process knowledge is a common operation knowledge description in technical manual, and its representation and storage will be detailed introduced in the next section.

IV. PROCESS COMPOSITION AND REPRESENTATION STRATEGY

To characterize the operation and maintenance, we need some basic terms (concept). For any concept interpretation and definition, we need a number of atomic concepts, which is the basis for other concepts. These concepts call for sufficient description to the domain concepts. To the fifth-level IETM, We have summed up the basic concepts of diagnosis and maintenance process, and have done a classification characterization. Characterization of the process ontology can refer to PSL (Process Specification Language) and PIF (Process Interchange Format). Although they cater to the manufacturing industry, they are also applied to the atomic concept in the process characterization.

A. *composition of the diagnosis program*

Solution consists of a series of boot states, guiding movements and user responses and other components. Boot state is a snapshot of guiding process when applies solutions, and is the embodiment of assembled information obtained when in the precursor diagnosis process. Guiding movement is the decision when system making a further diagnostic guide based on current obtained information. User response is the response that users respond to the system guiding steps, which includes selection (radio and check), confirmation (yes or no), prompt message entering of systems diagnosed, etc.

The system uses label transfer system to organize and describe the solution to the problem. That is, a solution is defined as the triples (S, T, L), in which the S is the set that guide the state; L is defined as a collection of user input; T: $S \times L$ is the mapping to S to define the diagnosis state changing after obtaining the users' information by guiding.

In order to easily make guiding decision based on current information, this system further maps the guiding decision-making to the state. On the other hand, since diagnosis guiding is a continuous process of trial and error, diagnosis process may involve backtracking. In order to determine whether it is necessary to facilitate backtracking, we add the attribute tag to the place node as possibly backtracking node.

Unified modeling of solution steps:

(1) State set: contains standards, the program initial state, intermediate state and the terminal state. The program original point transforms from the information which inherited from interaction into the program initial state through the loading process;

(2) Action set: contains input action set and output action set;

Input action, including machinery questions and automatic detection, is the trigger condition of state transition;

Output action provides information outside the system, such as prompt messages or tutorial messages.

(3) User interaction, automatic detection and state changing mechanism: the system automatically detects to achieve transforming of problem solving state through

interaction with customers in a certain state. We adopt "?Q" and "!Q" to distinguish between "interactive question" and "automatic detection". In Figure A, "$ M" is optional, which provides users an optional multi-media information when question to the user, and then asks another question.

（1）Title of state：the name of state；

（2）Content of state：it is consisted of a set of facts, and every fact is a 3-tuple, such as (object, property, value of property). For example, a inter-state can be descript as：

Title of state: “Operation System”

Content of state: (Operation System， Language， Chinese) & (Operation System, State, Out of Work)

Taking the diagnostic process of a device as an example, the following diagram illustrates this representation.

B. *XML file storage of Fault Diagnosis program*

In order to store the above-mentioned label transfer system in the system knowledge base, this system uses XML language to describe the label transfer system, and stores in the knowledge base in text form.

Here is a XML language description of a label transfer system for a certain equipment's diagnostic idea.

```
<?xml version="1.0" encoding="utf-8"?>
<statementchart>
  <state1 name="S1">
    <property>not</property>
      <tips>
        <tip id="1">
              <widget>checkbox</widget>
<content>Is a software installed before blue screen
    appears </content>
      </tip>
        <tip id="2">
      <widget>checkbox</widget>
<content> Is a hardware installed before blue screen
appears </content>
  </tip>
    </tips>
      <onChoose1>
      <nextStatement>S2</nextStatement>
    </onChoose1>
      <onChoose2>
<nextStatement>S3</nextStatement>
    </onChoose2>
      <onChoose1and2>
<nextStatement>S4</nextStatement>
    </onChoose1and2>
              <onChooseNothing>
      <nextStatement>S5</nextStatement>
    </onChooseNothing>
  </state1>
      <state2 name="S2">
              <property>not</property>
<tips>
      <tip>
      <content>uninstall the software</content>
    </tip>
  </tips>
              <onSuccess>
      <nextStatement>finish</nextStatement>
  </onSuccess>
              <onUnsuccess>
              <nextStatement>S5</nextStatement>
  </onUnsuccess>
</state2>
      ……
</state10>
      <state11 name="S11">
              <property>not</property>
              <tips>
                    <tip>
              <content>kill the virus </content>
    </tip>
  </tips>
              <onSuccess>
      <nextStatement>finish</nextStatement>
  </onSuccess>
              <onUnsuccess>
      <nextStatement>failed</nextStatement>
  </onUnsuccess>
</state11>
</statementchart>
```

Corresponding to the above label transfer system, the XML document illustrates the label on the document as below. Statementchart: is the XML document's root element. State1-state11: stands for 11 states. Each element has a name attribute, namely S1-S11. Property elements: correspond to the properties in the state diagram. Tips and tip: tips are a number of prompt messages corresponding to states, and tip is tips' sub-elements, which is each specific prompt message. Tip may have two sub-elements which are widget and content. Widget indicates the controls which information needs to show, and content means the content of the prompt messages. On ｘ ｘ ｘ: All the events are given in this form. For example: onChoose1 means choosing the first prompt message, onSuccess means successfully solving the failure. Nextstatement: is the next statement. There are two final statements. If it successfully diagnoses the problem, then the final statement is finish, otherwise failed.

C. *selection of solutions*

To build the shared semantic space, the distance of concepts come from different ontologies should be computed. Owing to the realizing the selection from data base, the content of their title is in a form like <verb>+<noun>. So, the similarity of title of segment of knowledge stored in system should be computed firstly, and it is the fundament to construct semantic connector. To given title just like S=<verbi>+<nouni>, we should design a mechanism to value the similarity of user requirement and title of knowledge segments. In above structure, there are two components and they are <verb> and <noun>. By computing the distance, a list of knowledge segment titles will be given after ranking according similarity of clause structure.

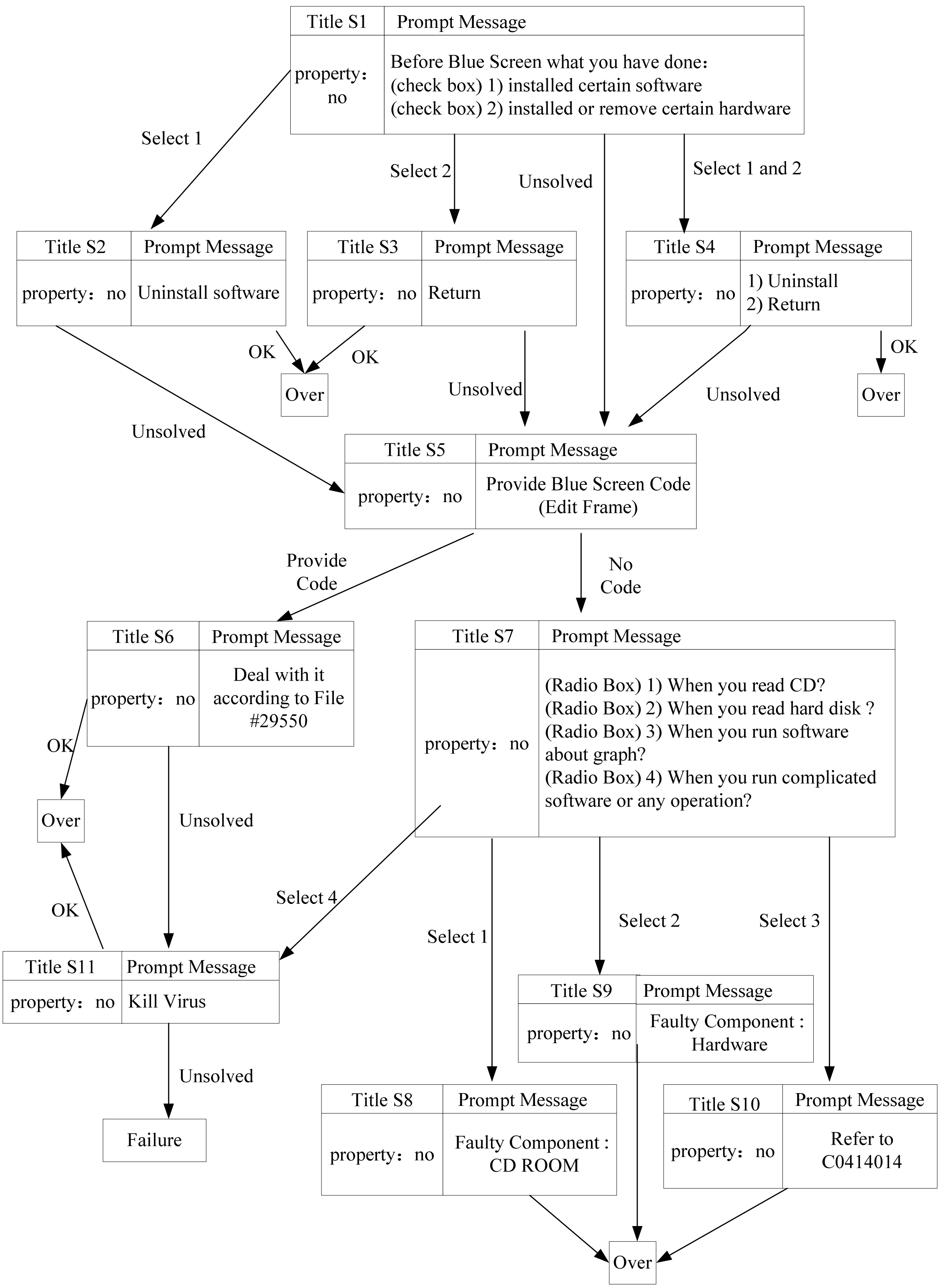


Figure 2 Representation of a computer diagnostic process

Take <noun> as an example, we can compute the distance between <noun> and <nouni> which stored in computer ontology. There are three class of relations between <noun> and <nouni>:

1) Equal relation: the similarity is 1，and denoted as *Sim(noun, nouni)=1;*

2) Super-sub relation: computing the absolute value of subtraction of the concept's depth, denoted as Depth(noun,nouni);

Sim(noun,nouni)=1/(1+ Depth(noun,nouni))

3) Other relation: computing the distance of the <noun> to the common ancestral, and then provide the absolute value of subtraction of the two distance, denoted as Depth(noun,comm-ancestor) and Depth(nouni,comm-ancestor). So,

Sim(noun,nouni)= abs(Depth(noun,comm-ancestor)-(nouni,comm-ancestor))/ (Depth(noun,comm-ancestor)+(nouni,comm-ancestor))

Calculating the semantic similarity of <noun> and <verb> respectively, and then the semantic similarity of clause structure can be get by given them different power.

Sim(sentence,sentence_i)

=(Sim(noun,nouni)+Sim(verb,verbi))/2;

Result=max(Sim_i(sentence,sentence_i)) 1<=i<=n;

As can be seen from above, the title of knowledge segment can be ranked by semantic similarity of sentence, after algorithm calculates the distance of <noun> and <verb> from U ontology and K ontology respectively.

Arithmetic for calculating semantic similarity.

Step1:descompose the sentence of users' question, and get the <verb> and <noun>. In knowledge we have done a pre-process by which we use an order pair to present the knowledge segment. A title set maps to the segments store in knowledge, denoted as

Title_set=

{(verb1,noun1),(verb2,noun2),…,(verbn,nounn)};

//using array a and b in practice;

Step2 For(i=1;i<=n;i++)

{if (noun==nouni) then Sim(noun,nouni)=1;

Else

{ //judging the relation between the <noun> and <nouni>, if their super-classes are same one, then we can draw a conclusion the relation between them are super-sub relation.

If (satisfy(relation(noun,nouni),super-sub))

Sim(noun,nouni)=1/(1+ Depth(noun,nouni)) Else

Sim(noun,nouni)= abs(Depth(noun,comm-ancestor)- (nouni,comm-ancestor))/

(Depth(noun,comm-ancestor)+ (nouni,comm-ancestor))

}}

a[i-1]=Sim(noun,nouni)； //a is a array used to store the semantic similarity of the <noun> and <nouni>

}

Step3：calculating the semantic similarity of Verb, and store it in array b.

Step4：For(j=1;j<=n;j++)

(Sim(sentence,sentence_i)=(Sim(noun,nouni)+Sim(verb,verbi))/2);

Return(Result=max(Sim_i(sentence,sentence_i)));

Focusing on text of fault diagnosis experience knowledge, we use a structured, formal mode to describe, and solve the following three problems. First, we design a process general representation of failure solutions based on the state transition diagram. On the basis of the state representation, we trigger the state transition by guiding movement and put in man-machine interactions. Secondly, we design the xml file representation and solutions to storage failures, and lay the foundation for knowledge sharing and interoperability. Finally, we provide a process-oriented selection scheme, which provides the alternates sorting order by calculating the semantic proximity.

V. Conclusion

Focusing on the intelligence, sharability, and interoperability requirements of the IETM, this paper proposes an IETM architecture of technology knowledge base. This architecture set domain ontology as formal semantic support of domain knowledge, emphatically solving the problems of formal representation transformation of text structuring and formalization, representation and storage of process knowledge, and process knowledge retrieval. The next step of the research work will focus on solution reuse of multi-fault diagnosis system, creation of solutions ontology, and to achieve automatic fault diagnosis design and implementation.

Acknowledgment

The authors wish to thank the chairs and the co-chairs of international conference CCCM2009, because they provide many useful comments for this paper so as to be published on journal. This research was supported in part by the Chinese National Natural Science Foundation under grant No. 60903130, Natural Science Foundation of Jiangsu Province under contract BK2009699 and BK2009698.

References

[1] Xie Hong-cheng, Liu Xu-dong On navy equipment without paper(Research Report)） Institute of East China Shipbuilding 2002.8

[2] DoD Tri-Service IETM Interoperability Task Force Preliminary Joint IETM Architecture (JIA) Handbook March 31, 1999 Prepared by: The DoD IETM Interoperability Core Architecture Team (CAT)

[3] Eric L.Jorgensen《DoD CLASSES OF ELECTRONIC TECHNICAL MANUALS》http://www.dt.navy.mil/tot-shi-sys/tec-inf-sys/etm/tec-std/pdf/CLASSES.PDF 2004

[4] Wang xue-qi, Xiao Ming-qing Study on Interactive Electronic Technology Manual and it Application 2007.05

[5] http://www.ietmsolutions.com 2008.04

[6] Browser-Based IETM Viewing Software Solutions http://www.inmedius.com/downloads/Publisher.pdf2008.1

[7] F Liu, H Zhang, J Zhu Ontology-Based Intelligent Interactive Electronic Technology Manual: Ontology-Based Intelligent Interactive Electronic Technology Manual: An Overview of the OIIETM Project In Pro of Convergence and Hybrid Information Technology, Nov. 2008,226-229

[8] Astrova I. Reverse engineering of relational database to ontologies. Proc. of the ESWC 2004. Heidelberg: Springer-Ver1ag，2004．327-341．

[9] Maedche A, Motik B. Stojanovic L. Stojanovic User-driven ontology Evolution Management. In Proc of the 1 3th European Conf. on Knowledge Engineering and Knowledge Management, Spain．Oct．2002

[10] Deng Zhi-hong, Tang Shi-wei, Zhang Ming. Ontology Research. Journal of Beijing University 2002, 38(5)：730～738

[11] Mello RdS, Heuser CA． A bottom-up approach for integration of XML sources. In Proc of the WWW．Brazil，2001．118-124．

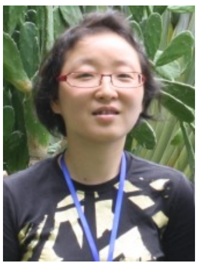
[12] Maedche A, Staab S. Ontology learning for the semantic Web. IEEE Intelligent System， Special Issue on the Semantic Web，2001，16(2): 72-79．

[13] M． Klein and N. Noy. A component-based framework for ontology evolution[C] In Proc. Of the WS on Ont. and Distr Sys. IJCAI'03, Acapulco，Mexico，Aug．9.2003.

[14] Ventrone V ，Heiler S． Semantic Heterogeneity as a Result of Domain Evaluation． SIGMOD Record Special Issue. Semantic Issues in Multi-database Systems. 1991.20(4)：46～54

[15] Heflin J, Hendler J. Dynamic Ontologies on the Web. In Proc. of the Seventeenth Nationa1 Conf. on Artificia1 Intelligence(AAAI 2000)． AAAI/MIT Press, Menlo Park，CA，2000．443～449

J.W. ZHU was born in Jiangdu, Jiangsu, P.R.China, in July 11, 1972. He received Master degree and Doctor Degree in computer science from Jiangsu University of Science and Technology and Nanjing University of Aeronautics and Astronautics, P.R. China in 2005 and 2008 respectively. He is currently associate professor of computer science, and the member of the Chinese Computer Society. He has published more than 30 papers in journals including the international journal of systems & cybernetics and Journal of southeast university or in IEEE CS proceedings. His research interest includes ontology, intelligent system, knowledge representation, semantic web, knowledge engineering and Interactive Electronic technology Manual. Dr Zhu recent research has been supported by the Chinese National Natural Science Foundation partially and Natural Science Foundation of Jiangsu Province, China. In additional, Dr Zhu has organized several national conferences and workshops and has also served as a program committee member for several major international conferences.



Y. Jiang was born in Shanghai, P.R.China, in Jan 15, 1974. She received Master degree in computer science from Jiangsu University of Science and Technology, P.R. China in 2005. She is currently lecture of computer science in Yangzhou University, and she has published more than 10 papers in recent years. Her research interest includes ontology, management information system, distributed database and system integrating.