IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

REVIEW OF COMPUTER ENGINEERING STUDIES
Vol.2, No.3, 2015, pp.17-20
http://dx.doi.org/10.18280/rces.020304

THE ONTOLOGY EXPRESSING AND KNOWLEDGE BASE BUILDING FOR TCM ASTHMA BASIS

Lin Fanlei, Gong Can and Yan Hongcan*

College of Science, North China University of Science and Technology, 46 Xinhua Road, Tangshan, Hebei, China.

E-mail: yanhongcan@ncst.edu.cn

ABSTRACT

Traditional Chinese medicine case is the essence of the academic thoughts of the old Chinese doctor and the records of the experience of clinical medication, it is the valuable resources we must inherit. However, Chinese medicine case is of huge wealth, and it is mostly unstructured knowledge, so it is urgent need a method to express the knowledge effectively and an efficient reuse case base to support. This paper introduces the knowledge of information engineering into the construction of traditional Chinese medicine information, and it applies domain ontology theory to the formalized expression of the knowledge of traditional Chinese medicine, and then builds out the information case base of Chinese medicine. Through the clinical application of case-based reasoning of TCM Asthma. This method offers a technique approach for process and standardization of traditional Chinese medicine in TCM, and provides a feasible ideas and methods for information research of TCM and the inheritance of clinical experience.

Keywords: Traditional Chinese medicine case, Domain ontology, Asthma, Knowledge base.

1. INTRODUCTION

TCM medicine through thousands of years of development, has a vast literature of Traditional Chinese Medicine. In the process of modernization of Traditional Chinese Medicine also produced a large number of literature data and scientific data, but these documents and data are exist in scattered, sporadic, one-sided situation, and usually use natural language expression [1].Due to reasons such as history, geography, TCM's word meaning is abundant and a meaning have more than one word, the word meanings changes of The Times, so the words in ancient TCM books is different with modern Traditional Chinese Medicine word. This is the important obstacle to realize information resources sharing.

Semantic retrieval can analysis from the perspective of semantic understanding information semantic retrieval object and retrieval request, mainly from the natural language processing, the method based on the concept and main ideas of based on ontology, and so on three aspects to realize the semantic integration and application in information retrieval, the purpose is to improve the automation and intelligence of Internet [2].Knowledge Base is the base and key of reasoning and knowledge accumulation in semantic search engine, and Ontology is the foundation of the Knowledge Base. Domain

Ontology [3-4] provides a set of terms and concepts to describe a domain, the Knowledge Base is to use these terms to express the fact, to finish knowledge representation and semantic reasoning. Shared ontology layer provides semantic level, is the key to realize semantic search. Architecture in this paper, the introduction of information engineering knowledge base and construct knowledge base based on domain ontology theory, will be out of the knowledge of traditional Chinese medicine as a form of the background, using the Protégé tool and OWL language constructs based on the traditional Chinese medicine out of the ontology of the domain knowledge base as an example, can solve the problem of case knowledge representation and organization, for the next step of traditional Chinese medicine for the case-based reasoning research lay the foundation.

2. KNOWLEDGE BASE

Knowledge Base is the traditional DB (DataBase) technology and the combination of the AI (Artificial Intelligence), is a reasonable organization of a particular field of declarative knowledge and procedural knowledge collection of [5-6].

Domain knowledge is the concepts in a certain field, the relationship between concepts and relevant concepts of the set of constraints. Domain Knowledge Base is a term in the field of AI. In the field of artificial intelligence, domain knowledge is mainly used in the expert system of knowledge and natural language understanding system.

2.1 Knowledge

At present, there are three kinds of representative definitions for the concept of knowledge:

(1) Feigenbaum: Knowledge is by selecting cutting, shaping, interpretation, and transformation of messages.

(2) Bernstein: Knowledge is a process by a description of the specific areas, relationships, and composition.

(3) Heyes-Roth: Epistemic=fact +belief + heuristic. Knowledge is often fuzzy, uncertain or incomplete. And knowledge is constantly in the process of dynamic change.

For any knowledge, usually adopt Heyes- Roth's knowledge of three-dimensional space to describe. Is the purpose of the scope of knowledge, knowledge, and the validity of knowledge? Range from specific to general, the purpose from description to specify, by certain to uncertain effectiveness. Knowledge of three-dimensional space description as shown in figure 1.

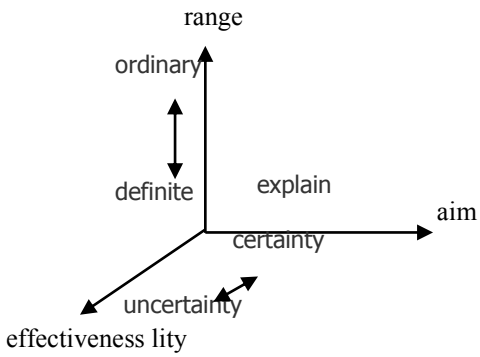

Figure 1. The three-dimensional description of knowledge

Date get value after a large amount of data processed, after analysis and processing of data become information, the role of information after the time limits and scope. In order to make information effectively in a long time, must carry out a series of internal processing, this process is called the comprehensive, integrated the information of the knowledge.

From the perspective of computer science, and knowledge is the result of information integrated processing. In the process of comprehensive, the information transmission and compare, combined into meaningful links. Data, information and knowledge with hierarchical relationships, their hierarchy is shown in figure 2.

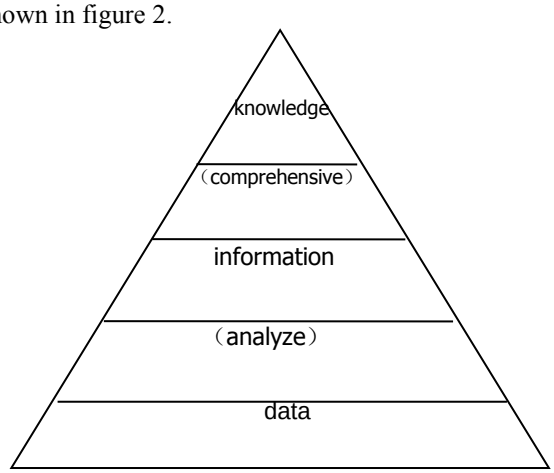

Figure 2. Hierarchy of Data information and knowledge

2.2 Domain knowledge

Domain knowledge is mainly used in the expert system based on knowledge and the concept of natural language understanding and the set of constraints. Knowledge engineering domain knowledge for the description of three aspects:

(1) Domain knowledge is a conceptual model, the conceptual model includes concepts and relationships between concepts.

(2) The domain knowledge is a constraint between concept and concept.

(3) Domain knowledge is stated how to calculation is derived the relationship between the new concepts and new concept of rules.

Two basic concepts of domain knowledge:

(1) Field characteristics of the concept: Is the conceptual domain knowledge, is the important concept in the field of all kinds of relevant semantic description.

(2) Field attributes: Refers to the concept of a field has the characteristics of the field feature concept can be a word, also can according to need to extend into a phrase or word list.

3. ONTOLOGY

3.1 The concept of ontology

Originally belong to the category of philosophy of ontology concept, the introduction of information science, its definition is constantly improve, Studer in 1998 put forward "ontology is clear of Shared conceptual model formal specification", is now recognized as the definition of ontology. This definition includes four meanings:

- Conceptual model: the abstract model of the phenomenon of the objective world.

2) Clear: concepts and links between them are precisely defined and defined.

3) Formal: computer readable precise formal description.

4) Sharing: reflect in the ontology knowledge is generally regarded as the concept of set in the related fields.

3.2 Ontology building

According to the principle of 'Clarity', 'Coherence', 'Extendibility', 'Minimal Encoding Bias', 'Minimal Ontological Commitment' of ontology construction, domestic and foreign scholars put forward a method of building ontology actual operation ,summarized as two kinds of patterns:

1) Use of the existing literature and domain experts use manual way of creating concept.

2) To transform existing table into ontology, or USES the learning mechanism, automatic or automatic ontology construction.

The internationally recognized body manual build methods[7-8] mainly include the skeleton method, enterprise modeling method, Methontology, KACTUS, circulating acquisition method, IDEF-5, seven steps ontology construction method, etc. Stanford university school of medicine, the development of ontology construction seven steps described as:

(1)Determine the professional field and scope of ontology;

(2)Investigate the possibility of reusing existing ontologies;

(3)List the ontology of important terms;

(4)Define the class and the class hierarchy;

(5)Define the class attribute;

(6)Custom properties of facets;

(7)Create an instance;

In this paper, we use the seven steps to construct ontology knowledge base for asthma syndrome of traditional Chinese medicine.

3.3 Ontology construction tools

At present there are a lot of ontology development tools, Different ontology development tool applied in different fields. Which developed by Stanford university's main advantages are Protégé (Using Java as operation platform, and open resources):

1) A graphical user interface;
2) Open network resources;
3) Various storage formats;
4) Open modular style.

We choose Protégé as ontology modeling tool，OWL (Ontology Web Language) as the ontology description Language. Protégé provides a predominantly knowledge model the behavior of the structure system to support a variety of expression form (such as OWL and RDF, Dublin core element, etc.) of the ontology construction, in Protégé 4.1.0 editor, Displayed in a tree structure of hierarchical directory structure ontology structure, For ontology maintenance operations (such as increased classes, subclasses, and properties, instance) is very intuitive, don't have to consider specific ontology description language, only in the level of the concept of design domain ontology model.

4. ONTOLOGY AND KNOWLEDGE BASE BUILDING

Collect and organize "Wu Jutong Medical cases" "Ding Ganren Medical cases. Typhoid fever cases", According to "The Chinese Medical" included 75 basis work involved in basis for the research object [9-10], With "the main Dyspnea" "breathing" "out of inverse Dyspnea" and "Dyspnea voice", "cough and asthma" as keywords as the core concept, Eventually determine the relevant property field more than ten, Established a relatively perfect card for basis of traditional Chinese medicine database. As shown in figure 3, as the research object in the form of background.

4.1 Define the class

From top-down of the ontology of top-down abstract base class, and hierarchical relationships, using Protégé gasping syndrome of traditional Chinese medicine domain ontology class relations preliminary model is set up. Among them, the class in the corresponding ontology system Class, The Instance of object corresponding ontology system or Individual. For example, exogenous type breath is a subclass of dyspnea of excess type, dyspnea of excess type is a subclass of dyspnea, using OWL describe the relationship between the specific as follows:

```
<owl: Class rdf: ID="Dyspnea type of exogenous">
<rdfs: subClassOf>
<owl: Class rdf: ID="dyspnea ofexcess type"/>
</rdfs: subClassOf>
</owl: Class>
<owl: Class rdf: about="dyspnea ofexcess type">
<rdfs: subClassOf>
<owl: Class rdf: about="Dyspnea"/>
</rdfs: subClassOf >
</owl: Class >
```

4.2 Define the hierarchical relationships

Each relationship is equivalent to a dual or multiple group function, the domain and range of each function is asthma disease ontology model within the ontology classes and the subclass object, We can use the ObjectProperty properties to set up domain and range of the function of the relationship, in order to facilitate the relation among the class of the ontology model through correlation, the following shows is the ObjectProperty properties of "phlegm type" and its special properties, Among them the domain of "phlegm type" is case, and the range of "phlegm type" consists of "white phlegm" "yellow phlegm" "lucency phlegm", the case is composed of three parts of the range, at the same time it also has a transmission properties.

```
<owl: TransitiveProperty rdf: about="#phlegm type"
<rdfs: domain rdf: resource="#case"/>
<owl: inverseOf rdf: resource="#compose"/>
<rdfs: range>
<owl: Class>
<owl: unionOf rdf: parseType="Collection">
<owl: Class rdf: about="#white phlegm"/>
<owl: Class rdf: about="#yellow phlegm"/>
<owl: Class rdf: about="#lucency phlegm"/>
</owl: unionOf>
</owl: Class>
</rdfs: range>
<rdf: type rdf: resource=http: //www.w3.org/2002/07/owl#ObjectProperty/>
</owl: TransitiveProperty>
```

4.3 Build conception tree

Including all the name of the attribute, value, type and other constraints, the corresponding ontology system of Datatype Propety. Attributes are binary relation between individuals, That is, properties make two individuals together. Described the relationship between the concept of ontology mainly has four kinds:

(1) Part and the overall relations between concepts

(2) Fathers and sons inheritance relationships between concepts

(3) The relationship between the concepts and the instance of the concept of

(4) A concept is another attribute of the concept of relationship

The intrinsic relationship between concepts constitutes the rich semantic relationships between concepts, by the rules of logic reasoning can obtain the implication of relationship between concepts.

In order to realize the semantic extension of keyword queries, ontology model must support the original meaning, synonymy, nearly righteousness and up and down a relational query. For example, "micro-asthma" can have "asthma" and other synonyms and synonyms，any of these keywords, the input, can not only examine the information of keywords original meaning, but can show its synonyms for information. Was the key to improve the case retrieval recall in the future?

At last, take shape of a concept of tree. Ontology mainly include two ways of file storage and database storage, the file is stored is stored in OWL and RDF file format. Asthma disease ontology model stored in OWL file format, the case knowledge base data is stored in relational database. Asthma

syndrome of traditional Chinese medicine ontology based

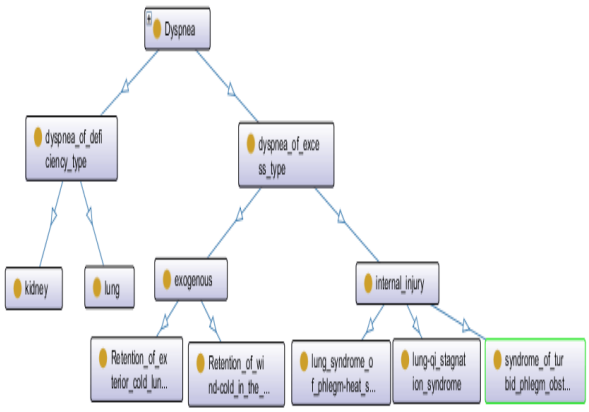

Figure 3. Ontology fragment of TCM

tool Protégé fragment as shown in figure 3.

In order to verify the ontology model is reasonable, we can use Jean reasoning machine to validate instance defined in the Class, and a definition, extension, examples of attributes and attribute values. As same as validate the rationality of the ontology model and the concept of satisfiability, and makes the corresponding adjustment to the unreasonable place.

4.4 Export the owl ontology file

At present, there are four files in Protégé type, this article choose OWL RDF files, can support the OWL ontology file browsing and editing tools are stored, read and modified, for knowledge reasoning [11-12] and retrieval provides a resource description.

Asthma syndrome of traditional Chinese medicine basis of ontology concept semantic tree as shown in figure 4.

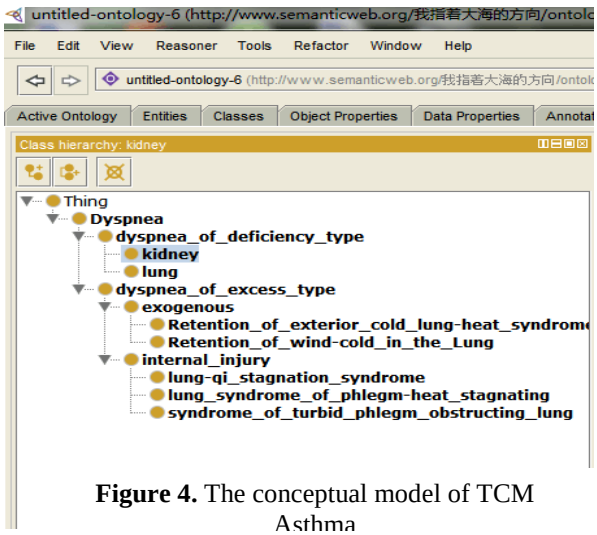

Figure 4. The conceptual model of TCM Asthma

5. CONCLUSIONS

As a new tool of organizing knowledge, ontology technology can be used to signify and organize the TCM (Traditional Chinese Medical) knowledge and relevant information resources multidimensionally visually and dynamically, ontology-based domain knowledge base can clearly signify the concept attribute handling method and internal relations of domain knowledge, those concept is benefit to knowledge expression and semantic reasoning. This paper, used ontology technology to the Traditional Chinese Medicine Asthma's information base constructing, is the preparation of the realization of Asthma for the knowledge retrieval and the preparation of case based reasoning, the paper aim to find a new way to Traditional Chinese Medicine informationalization and inherit Traditional Chinese Medicine.

ACKNOWLEDGMENT

This work was supported by college students' innovative projects(X2015064), the National Natural Science Foundation of China (61370168) and the National Natural Science Foundation of Hebei (F2014209238).

REFERENCES

1. Tian, T., Gao S. H., Application Status and Prospect of Ontology in TCM' [J], *China Journal of Traditional Chinese Medicine and Pharmacy*, pp.1832-1834, 2003.
2. Xu D. S., Zhang Z. X., Xin M. F., Ontology Knowledge Base of the Study of Visual Search [J], *Information Studies: Theory & Application*, pp.114-117, 2010.
3. Li J., Meng X. R., Su X. L., Development Method And Practice Research On Domain Ontology [M], Graduate School Of Chinese Academy Of Agricultural Sciences, 2009.
4. Zen. Z., Practice in Construction of Traditional Chinese Medicine Ontology [J], *Chinese Journal of Medical Library and Information Science*, pp.1-5, 2009.
5. Ma X. D., Den X. N., Peng W. J., Yan H. C., Based on Domain Ontology Knowledge Base Architecture and Implementation [J], *Journal of Hebei Polytechnic University: Social Science Edition*, pp.42-47, 2012.
6. Gao C. M., Bao H. F., Zhou Q., Ontology Construction Principle and Its Application in TCM [J], *Medical Information*, pp. 581-583, 2008.
7. Cao, Y, F., Cao, C, G., Knowledge Acquisition of Tongue Diagnosis an Ontology-Based Approach [J], *Application Research of Computers*, pp.31-35, 2006.
8. Zhang J. F., Building And Application Of Teaching Based On Ontology Knowledge Base [J], *Modern Computer*, 2011.
9. Yan, H, C., Li L. H., Ma H. X., Li, J., Bao J. T., Liu B. X., Analysis on Clinical Medicine Asthma Medical Records Association Rules Based on Rough Set Theory [J], *Liaoning Journal of Traditional Chinese Medicine*, pp.1218-1220, 2012.
10. Yu, N., Ontology-Based Construction of Hierarchical Categorization System of Domain Knowledge Base [M], Northeastern University, 1218-1220, 2005.
11. Rensel D., Harmelen F. V., Horrocks L., Etal. OIL: An Ontology Infrastructure for the Semantic Web [J], *IEEE Intelligent System*, 2001, 16(2): 38-45.
12. Baader F., Horrocks I., Sattler U., Description Logics as Ontology Languages for the Semantic Web, *Hutter D., Stephan W. Eds. Festschrift in Honor of Jorg Siekmann. Lecture Notes in Artificial Intelligence*, Springer, 2003.