HSFDONES: A Self-Leaning Ontology-Based Fault Diagnosis Expert System Framework

XiangBin Xu

School of Mechatronics Engineering, East China Jiaotong University,
330013 Nanchang, P.R. China
champagnewq@yahoo.com.cn

Abstract. HSFDONES is an expert system fault diagnosis which makes the fault diagnosis working more intelligently, HSFDONES uses the ontology-based self-leaning theory and technology to build fault diagnosis expert system. The fault diagnosis knowledge structure is defined and the relevant structure ontology and core fault ontology is researched in HSFDONES; the fault diagnosis data warehouse is built, the decision tree and Apriori algorithm are used to acquire fault knowledge to realize HSFDONES's ontology self-learning. HSFDONES offers system framework for building intelligent fault diagnosis system. Finally the agricultural machinery's hydraulic fault diagnosis expert system was developed on the basis of the framework.

Keywords: HSFDONES, Ontology, Fault Diagnosis, Self-learning, Data Mining.

1 Introduction

HSFDONES (Hydraulic System Fault Diagnosis Ontology Expert System, hereinafter referred to as HSFDONES) is an ontology-based hydraulic system fault diagnosis expert system. The system architecture of HSFDONES is shown in Fig.1.
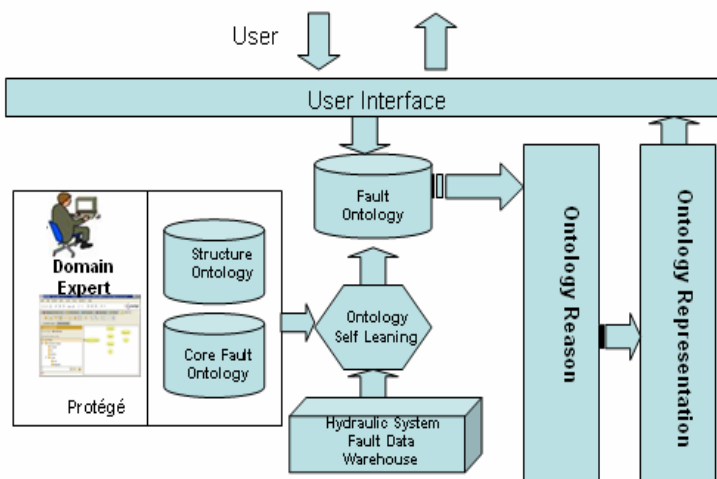

Fig. 1. The System Architecture of the HSFDONES

HSFDONES takes data warehouse technology to organize and manage heterogeneous system fault information, the ontology is used as HSFDONES' knowledge backbone. In order to realize ontology self-leaning, HSFDONES takes learning algorithm in the field of machine learning to realize diagnosis ontology self intelligent learning, and finally HSFDONES uses Jena and Topic Maps as its reference and presentation engine respectively.

2   HSFDONES Data Warehouse

2.1   Data Warehouse Structure

Hydraulic system fault information is stored in the media such as database, Excel, TXT and other file formats, the different format data can not offer knowledge efficiently. Data warehouse technology is an effective solution to solve this problem [1]; data warehouse technology can extract, handle and standardize large, varied data. The data in warehouse is organized in the form of the data cube, after data 's standardization and consistency, the data can be loaded and imported into the data warehouse's cubes, the hydraulic system fault data warehouse's structure is shown in Fig.2.
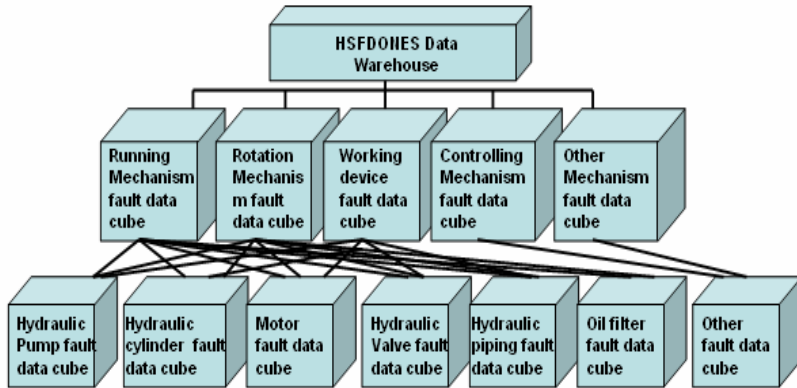

Fig. 2. HSFDONES Data Warehouse Structure

2.2   Data Warehouse Construction

HSFDONES take Mondrian as its OLAP[2] engine to build the hydraulic system fault diagnosis data warehouse,Mondrian is java-based, open source data warehouse engine, the hydraulic system fault diagnosis data warehouse and data cube of each module built by Mondrian are presented in the form of XML and data tables. HSFDONES use the JDBC technology to realize data extraction, cleaning and transformation, after that the data can be loaded into Mondrian's data warehouse. Once HSFDONES' fault data warehouse is constructed, the knowledge-learning algorithm can extract knowledge from it efficiently.

3   HSFDONES Knowledge Structure

3.1   Oontology-Based Knowledge

HSFDONES core ontology includes two parts: the hydraulic structure ontology and the hydraulic fault ontology. Hydraulic structure ontology describes the structure of hydraulic system and assembly relations; Hydraulic fault diagnosis ontology describes the hydraulic system's critical fault information, which means hydraulic system failure and their relationship.

In order to design and build the HSFDONES core ontology, first we should define the domain area[3] of hydraulic system's fault diagnosis, which can be done by consulting domain experts, filed researching and statistics of historical data. After that, we can make abstraction to identify and define the ontology classes, class hierarchy, class property, restrictions, characteristics (such as symmetric, transitive, etc), finally build individual instances of ontology classes to get HSFDONES diagnosis ontology model.

3.2   HSFDONES Fault Knowledge Definition

The hydraulic system's fault is described in three parts, which are fault phenomenon, fault causes and failure sources (the parts or component which causes the failure). Fault phenomenon is the presenting symptoms, fault source is parts or components where the fault raises, fault causes is information chain which can track from fault phenomena to fault source. HSFDONES fault knowledge can be described in fault triples:

<center><Fault phenomena, Fault causes, Fault source></center>

HSFDONES fault knowledge should figure out the connection which links the fault phenomenon to the fault source through some logical analysis.

3.3   HSFDONES Fault Ontology Construction

HSFDONES hydraulic system fault ontology is built in the seven-step ontology way. After consulting domain experts, we get the following fault diagnosis knowledge:

1) IF "the fuel tank can not change direction" AND "the valve can not return to the center position" THEN "the electromagnetic iron is damaged";

2) IF "the valve can not return to the center position" AND "the electromagnetic iron is good" THEN "the gap is too small" OR "the oil is polluted";

3) IF the electromagnetic iron is good" THEN "the coil is aging" OR "the switch is circuited".

Let's take the first diagnosis knowledge as an example to show the process of ontology constructing:

The fault knowledge triples are:

< Fuel tank can not change direction, Valve can not return to the center position, Electromagnetic iron is damaged>

According to the fault knowledge, we can define related ontology knowledge in the ontology knowledge base, namely:

Fuel tank's changing direction: Can not change -> fault phenomenon
Valve resetting position: Can not reset -> fault cause
Electromagnetic iron Status: Is damaged -> fault source

The fault phenomena, fault causes and fault source described in HSFDONES fault knowledge triple can modeled as ontology's class, their status (such as: "can not change direction ", "can not reset position", etc.) are modeled as class properties associated the corresponding ontology class.

Different from the hydraulic system structure ontology, we just build the fault diagnosis core ontology, and the fault diagnosis core ontology can self-learn new ontology once the hydraulic system new failure occurs, which means fault diagnosis core ontology can grow dynamically.

The fault diagnosis should be formalized so as to be understood by machine, we build the fault diagnosis ontology model in Protégé IDE[4], after that we should transfer and formalize the fault diagnosis ontology model, HSFDONES use OWL as the ontology coding language to formalize and describe fault diagnosis ontology model[5], OWL-based ontology can be shared between systems.

4   HSFDONES' Self-learning Algorithms

4.1   Self-learning Algorithms

HSFDONES ontology self-learning algorithm can excavate fault knowledge from the hydraulic fault data warehouse, which can be transferred into ontology by referring to the hydraulic system structural ontology and core fault ontology, and the new-learned fault ontology are stored in the fault ontology library.

Decision tree[6] algorithm is efficient classification algorithm in data mining, it can extract decision-making knowledge from the training data, decision tree's knowledge is described in the tree-style, which includes internal nodes, fork and leaf nodes, the internal nodes in decision tree represent a test property, the fork in decision tree represent classification, the leaf in decision tree nodes represent a diagnosis result, the decision tree can display knowledge in a graphical way, the decision tree uses ID3 and the C4.5 as learning algorithm, ID3 is mainly applied to discrete attributes, C4.5 classification can leaning from both discrete and continuous attributes, C4.5 builds decision tree based on information entropy theory, which calculate the information gain for each attribute recursively, select the highest information gain attributes as the split attribute, the information gain is calculated as following:

1) Compute the entropy of dataset:

$$Info(S) = -\sum_{i=1}^{k}((freq(C_i,S)/|S|)*\log_2(freq(C_i,S)/|S|))$$

2) Compute the Weighted entropy sum of property:

$$Info_x(T) = -\sum ((|T_i|/|T|) * Info(T_i))$$

3) Compute the Info Gain of property:

$$Gain(X) = Info(T) - Info_x(T)$$

The freq (Ci, S) represents the sample size belongs to class Ci in training set S, Gain (X) represents information gain of the property X. C4.5 learning algorithm finds max information gain property as the root of decision tree, and then call the algorithm recursively in the split data set, until a whole decision tree is constructed; once the decision tree is built, we should prune it to get a refined, non-redundant decision tree.

4.2   Knowledge Extraction

The knowledge leaned by decision tree is showed in a "IF ... ... Then ... ..." style, we can extract the knowledge rule from decision trees as follows:

1) The fault knowledge rule is paths which link the root to each leaf respectively;
2) The internal test nodes along the path represent the premise of the rule (IF) section and each leaf node at the end of the path represent conclusions (THEN).

5   HSFDONES Knowledge Inference and Presentation

5.1   Knowledge Inference

HSFDONES take Jena as it's of inference engine, Jena is Java-based open source semantic web toolkit developed by HP, it offers a programming environment for parsing RDF, RDFS and OWL ontology, and Jena provides a rule-based inference engine. Jena's inference engine can be used in OWL-based ontology, its inference is based on direct graph, if can offer forward reasoning, backward reasoning and reasoning mixed. The rule sets Jena get are bound to the data model, which can form a new reasoning input, so the query results returned by the Jena inference engine include both original data triples and the new knowledge rules inferred from the ontology library.

In HSFDONES system,once the hydraulic fault diagnosis ontology OWL ontology model is imported into Jena's ModelFactory, the Jena inference engine can transform the OWL ontology into a an ontology inference graph, then the Jena inference engine can take the ModelFactory API to interact with the hydraulic fault diagnosis ontology, which means the end-user can query and get hydraulic fault diagnosis knowledge.

In the HSFDONES system, the Jena API is embedded into the JSP code, the diagnosis knowledge inferred by the Jena is transformed in a XTM-formatted style, which finally is showed in the end-user's web browser.

5.2   Knowledge Presentation

HSFDONES take the Topic Maps to present the fault diagnosis knowledge, Topic Maps can organize and describe the ontology knowledge topics based on meta-data,

and the ontology knowledge organized and described on the topic-based can be navigated and explored easily. The HSFDONES knowledge topic map includes three basic elements: topic, association and the occurrence. In the topic map, HSFDONES fault source organized can be located easily,and the fault correlation can be displayed directly, which means a easy way for the user to navigate, explore and trace the hydraulic fault, this is mostly helpful for the fault detection and prevention.

HSFDONES take the TM4J as it's topic map engine, TM4J offer a user-friendly interface and a powerful a secondary development environment, which make it flexible and extensible, TM4J take a XML-based specification language-XTM to describe the ontology knowledge and its relation, so HSFDONES needs a data adapter, which bridge the gap between the OWL and XTM, the working processes of the adapter is showed in Fig.3.
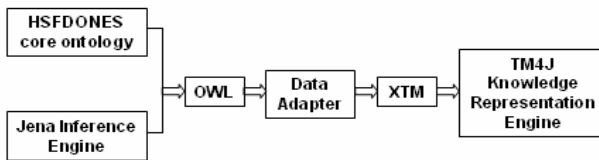

Fig. 3. HSFDONES Knowledge Representation

6 Conclusions

Taking the fault diagnosis expert system as an example, this paper details in the theory and implementation of fault information organization, ontology knowledge definition, ontology self-leaning algorithm, ontology inference and representation. The open source solution is used here and the HSFDONES was successfully developed based the researching result. The successful research and application of HSFDONES offer useful guidance for the development of intelligent diagnosis expert system.

Acknowledgements

References

1. Yuan, H.: OLAP and Data Warehouse Modeling Techniques. Computer Application Research 12, 61 (1999)
2. Liu, Y., Chang, G.: OLAP-Based Relational Database Research. Computer Engineering and Applications 38, 127 (2001)

3. Liu, J., Hu, L.: Skeleton of Boiler-Based Fault Detection System of the Domain Ontology Construction. Henan University of Science and Technology: Natural Science Edition 3, 36 (2008)
4. The Protégé project, `http://protege.Stanford.Edu`, 2009-04
5. Du, N., Huang, D.: Ontology-Based Fault Diagnosis for Chemical Process. Computer Integrated Manufacturing System 4, 587 (2004)
6. Wei, H., Kamber, M.: Data Mining Concepts and Techniques. Mechanical Industry Publishing, China (2002)
7. Lan, Y.: Excavator Hydraulic System Fault Diagnosis and Rule Out. Water Resources and Technology 2, 71 (2006)