RESEARCH ON ACQUIRING DESIGN KNOWLEDGE BASED ON ASSOCIATION RULE

Jingmin Li[1], Yongmou Liu[1,2], Jiawei Yang[1], Jin Yao[1]
[1]Department of Manufacturing Science and Engineering, Sichuan University, China; Email: hjljm@163.com. [2]IMAG Asia Ltd.

Abstract: This study presents a method of acquiring product design knowledge. By analyzing the process and mathematics expression of product design, design attributes are taken as item sets and associated relationship among them are taken as analyzing target. Based on the expression and application for domain knowledge, design data are pretreated. Then the correlations knowledge among design attributes is obtained by association rules. A case study is given to illustrate that the method proposed is valid and practicable.

Key words: Product Design, Data Mining, Association Rule, Domain Knowledge.

1. INTRODUCTION

As we all known, product design consists of two kinds of knowledge: explicit knowledge and tacit knowledge[1]. Explicit knowledge means the standard knowledge which is like books and manuals and easy to express, to capture, to encode; tacit knowledge means the knowledge which is just experience derived from experts long-term working, such as how to choose materials or parameters, how to optimize the relationships among attributes, and how to make decision according to input conditions. This kind of knowledge can not be summed up easily since it is private, underlying and bitty. Gaining them is a challenging problem that we have to solve.

Relevant researches focus on the development of expert system and knowledge system[2-3], in which more attentions are paid to the expression and reasoning for domain knowledge. But the essential problem of design knowledge acquisition hasn't been solved and becomes the bottleneck of system further development. Data mining, as one of the promising technologies since 1990s, is to some extent a non-traditional data-driven

This project is supported by Sichuan Science and Technology department of China under the grant No.04JY029-003-2.

Please use the following format when citing this chapter:
Li, Jingmin, Liu, Yongmou, Yang, Jiawei, Yao, Jin, 2006, in International Federation for Information Processing (IFIP), Volume 207, Knowledge Enterprise: Intelligent Strategies In Product Design, Manufacturing, and Management, eds. K. Wang, Kovacs G., Wozny M., Fang M., (Boston: Springer), pp. 103-108.

method to discover novel, useful, hidden knowledge from massive data sets. Association rule mining is one of mining techniques and has attracted more and more attention of academic and practitioners with applications such as customer relation management[4], market baskets[5], fault diagnosis system[6], decision support system[7], time-series analysis[8] and so on, but to a less extent in engineering design. This paper tries to apply association rules to the field of product design for acquiring associated knowledge among design attributes, and the key technology is investigated.

2. PROBLEM STATEMENT AND DEFINITIONS

2.1 The graphic description of product design

The process of product design is an iterative analyzing-synthesizing-decision-making process. This kind of design problem can be expressed in the tree as shown in figure 1. P is function requirements, which can be divided into some sub-problems $\{P_1, P_2, ..., P_n\}$, each of which also can be divided into $\{F_1, F_2, ..., F_j\}$, until into $\{SP_1, SP_2, ..., SP_k\}$ which can not be divided any more. These sub-problems that can't be separated are called fontal-problems. Each fontal- problem corresponds with many solutions.

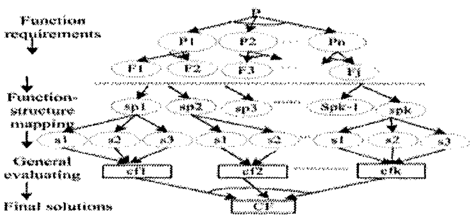

Figure 1. And & Or tree for product design problem

But since these fontal-problems are interrelated and interact on each other, corresponding solutions are various, similar, complicated, and sometimes contradictory on each other. Designer must make correct decision based on much knowledge to implement this mapping from problems to solutions.

2.2 The mathematic expression of product design

Product design problem can be described into the problem field and conclusion field in the form of attributes and attribute-values. The design case can be expressed in mathematics implication as follows:

$$P = < E, C, D, V, f > \quad (1)$$

The variable $E =< E_1, E_2, ..., E_n >$ is the set of design case; $C \cup D = A$ is the set of design attributes, in which C expresses condition attributes including user requirements, design environment states and so on, and D expresses conclusion attributes resulted from condition attributes set; $V = E_{a \in A} V_a$ is the set of attribute values; $f : E_x A \to V$ is the mapping rules from attributes to attribute-values. During the design process, designer must perform the mapping from condition attributes C to conclusion attributes D on the constraints of f. Discovering this kind of strong correlations or independence among all attributes is the problem that this paper will solve.

3. THE METHOD AND IMPLEMENTATION

3.1 Association rule for design data

The basic principle of associations is shown as follows: let $I = (i_1, i_2, ..., i_m)$ be a set of items. Let D be a set of transactions where each transaction T is a set of items such that $T \subseteq I$. A transaction T is said to contain A if and only if $A \subseteq T$. An association rule is an implication of the form $A \Rightarrow B$, where $A \subset I$, $B \subset I$, and $A \cap B = \phi$. The rule $A \Rightarrow B$ holds in the transaction set D with support s and confidence c, where s is the percentage of transactions in D that contain both A and B, and where c is the percentage of transactions in A that also contain B. The mining problem is to find all rules that satisfy both a minimum support threshold and a minimum confidence threshold. The process are two steps: (1) finding all frequent itemsets: each of these itemsets will occur at least as frequently as a pre-determined minimum support count; (2) generating strong association rules from frequent itemsets. It requires that data are expressed in the form of attributes and attribute-values, which is consistent with the expressing form of design case.

3.2 Domain knowledge for pretreatment

The essence of association rules is to search the correlations among "1(T)" values in a relational table in which all attributes are Boolean variable. Many discretization techniques, which can be used to reduce the number of values for a given continuous attribute by dividing the range of attribute into intervals, are required and investigated widely. Since design data has some domain meaning instead of purely numerical data, the determination of the interval number and the break point of a continuous attribute must depend on expert's domain knowledge which can come from domain expert, or can be refined from the result of data mining. The application process of domain knowledge are shown in figure 2, in which more attentions are paid to expressing the relations among design attributes or records and to constraining values of some attributes or records instead of being used for reasoning like in expert system

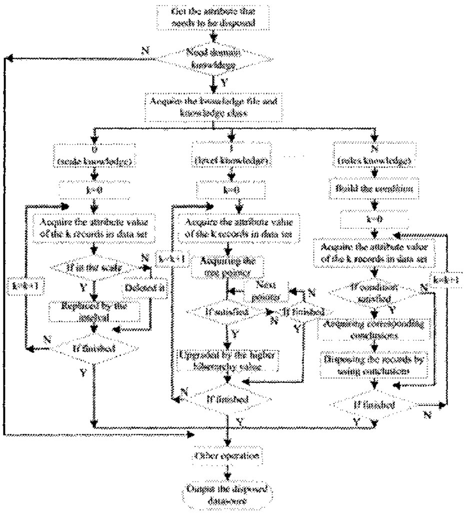

Figure 2. Domain knowledge application

4. A CASE APPLICATION IN DESIGN

Taking form-selection of gate valve as an example, each design case can be seen as a transaction T which includes many items such as working medium(WM), nominal pressure (PN), working temperature (WT), nominal diameter (DN), structure form and so on. These items have strong constraint relations. Acquiring these relations is the aim now. Firstly, the simple table D is set up shown in table 1. Since valve model has many different values that can't be conceptual, delete it; WT, PN and DN are all needed to be pretreated and corresponding domain knowledge is shown in figure 3.

Table 1. The simple table D

ID	Model	WM	PN	WT	DN	SS	WD	...
1	Z41T	water	1.0	100	50-450	B	W+S	...
2	Z65H	oil	2.5	425	15-50	A	W+S	...
3	Z40Y	acid	6.4	200	25-300	M	W+E	...
4	Z40H	oil	4.0	350	50-200	A	W+E	...
5	Z42W	gas	0.1	100	300-800	M	W+T	...
6	Z763F	petrol	6.4	30-80	700	M	P+S	...
7	Z43H	slop	0.6	120	50-350	M	P+S	...
8	Z44W	oil	1.0	100	50-400	M	P+T	...
...	...	...	...	...	...	...	...	...

The process of applying domain knowledge is as follows: (1) by searching for the relevant records according to "mining algorithm = association" and "Data sets = S-F", find relevant knowledge file and knowledge class; (2) in .\asso1, according to the method of acquisition and application for rule knowledge, pretreat DN value into four values: small_dia, middle_dia, big_dia and super_dia. In.\asso2, process WT value into five values: super_low_T, low_T, normal_T, middle_T and High_T. In .\asso3, upgrade PN value according to its conceptual hiberarchy tree. Then, Let each $< x, l_k, u_k > \in split(x)$ map into logical attribute A, and each attribute value will be transformed into many Boolean variable. All these attributes compose the itemsets required by mining. The original simple table D can be changed into Boolean transaction base, shown as in table 2.

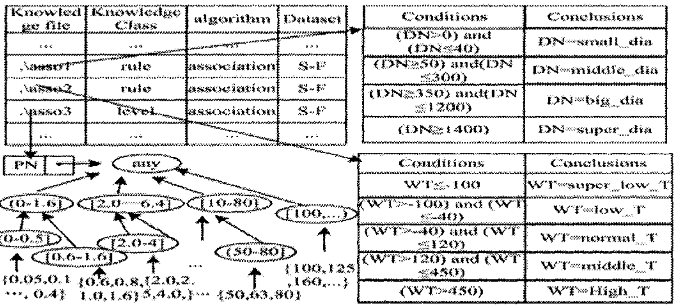

Figure 3. Store structure of domain knowledge

Last, based on the algorithm of association rules, Boolean design transaction D is processed and many rules can be obtained. Part of strong rules acquired in this example are shown as follows:

Table 2. Boolean design transactions table

ID	PN1	...	WM1	...	WT1	...	DN1	...	SS1	...	WD1	..
1	T	...	F	...	T	...	T	...	T	...	T	..
2	T	...	T	...	F	...	F	...	F	...	T	..
3	F	...	F	...	F	...	T	...	T	...	F	..
4	T	...	T	...	F	...	T	...	T	...	F	..
5	F	...	F	...	T	...	F	...	T	...	F	..
6	F	...	F	...	T	...	F	...	T	...	F	..
7	F	...	F	...	T	...	T	...	T	...	F	..
8	T	...	T	...	T	...	T	...	T	...	F	..
...	...	...	...	...	...	...	...	...	...	...	...	..

Rules1: $PN1 \wedge WT2 \Rightarrow WD3, \sup = 7.6\%, conf = 50.5\%$ .When PN is(0.5-6] and WT is high, the designs that the form of Wedge disc is parallel and single, are 7.6 percent in all design transactions. The designs that the form of wedge disc is parallel and single are 50.5 percent in all the design

transactions that PN is (0.5-6] and WT is high. This rule shows that parallel and single wedge disc fits the situation of high temperature and low pressure.

Rules2: $PN3 \wedge WM1 \Rightarrow WD1, sup = 8.3\%, conf = 65.2\%$. When PN is (10-80] and WM is water and steam, the designs that the form of wedge disc is wedge and single, are 8.3 percent in all design transactions. The designs that the form of wedge disc is wedge and single are 65.2 percent in all designs that PN is (10-80] and WM is water and steam. This rules shows wedge and single gate valve fits the situation of high pressure and low-viscidity medium.

These rules are the heuristic information with some supports and confidences, and through expert trimming, selecting, and evaluating, can be changed into useful and regular knowledge which are consistent with expert's thinking mode and provide important references for designer.

5. CONCLUSION

In this paper, based on domain knowledge discretization, association rule is introduced to the field of product design for acquiring the correlative knowledge among design attributes. This method integrates the computing ability of computer with expert domain knowledge. In the future, much more researches are needed to further improve its efficiency and accuracy in order to acquiring more complex design knowledge.

6. REFERENCES

1. Dienes Z,Perner J.A.,(1999), Theory of Implicit and Explicit Knowledge, *Behavioural and Brain Sciences*, vol. 22, pp. 735-808.
2. Tao Shanxin, Tang Wenxian, Li Limin,(2003), "Research and Implementation on Product Development System Based on Knowledge Driven", *Journal of Computer Engineering and application* ,vol.22, pp.129-131.
3. Liu Ying, Ni Zhonghua, (2005), The Expert System of Cooperative design Knowledge Based on machine tool products Full Life Cycle, *Group Technology & Production Modernization* ,Vol.22, pp.10-13.
4. Jiawei Han and Micheline Kamber, Data Ming: concepts and Techniques (China Machine Press, Beijing, 2001).
5. Haiming Chen, Wu qing, Chunhua Zhao, (2004), "Associated Rules Mining in DIAGNOSIS Based on Oil Monitoring", *Journal of WUHAN UNIVERSITY Of Technology (Transportation Science & Engineering)*, vol. 28, pp.729-731.
6. Show-Jane Yen, Arbee L.P. Chen Member, (2001), "A Graph-based Approach for Discovering Various Types, of Association Rules", *IEEE Trans.Knowledge and data Eng*, vol.13, pp.1354.
7. Hou Xue-bo, Tian Bin, (2005), "Application of association rules techniques in electric marketing analysis", *Proceedings of the CSU-EPSA*, vol. 17, pp.67-725.
8. Li Ming, (2005), "Application of Mining Association Rules with Multiple Minimum supports in Sales Data", *Computer Engineering*, vol.29, pp. 92-93.