

Research of Component Integration Knowledge Acquirement Method Based on Rough Set

Xue Junfang^a, Wan Yuntao^b and Zheng Lei^c

Inner Mongolia university of technology, mechanical school, Hohhot, China

^axuejunfang@tsinghua.org.cn, ^b1328423730@qq.com, ^c405773856@qq.com

Keywords: Component integration; Rough set; Knowledge acquirement.

Abstract. The concept of component integration (CI) has been employed in several product design methods to reduce the total number of components and improve manufacturing quality and efficiency. However, none of existing methods provide the effective information needed to identify the appropriate components for integration. A methodology based on rough set is developed in this research to analyze expert experiences and acquire expert knowledge which wrapped in experiences data. It can build the decision rules for components integration. Regarding the significance of attributes as heuristic information applying to classical rough set, a new arithmetic for core and reduction of attribute is proposed. An example demonstrates that the arithmetic is simple and easy to achieve. An instance indicates that proposed method is effective and practical.

Introduction

The quantity of the parts in products directly affects the cost of each link in the life cycle, such as the product design, manufacture, use, and even recovery processing after the retirement. Component integration is to combine two or more parts into one part when meeting the certain conditions, and a designing optimization method satisfying the original functions of the product, which can effectively reduce parts quantity and cost, thus component integration is of important significance in design process^[1].

At present, component integration method are performed by the experience of designers or experts in such certain fields, there is no method which can effectively determine which parts could be combined together. Boothory and Dewhurt presented a method to determine the minimum number of parts in a product theoretically by answering three questions^[2]. Suh proposed observing function sharing principle to reduce the number of parts by designing one piece with integrating the designing parameters of other more pieces^[3]. Gu Tingquan indicated that arguing the analysis of component integration necessary conditions and feasibility conditions as the criterion whether parts could be combined^[4]. However, component integration is also conditioned by product function, processing technology, manufacturing difficulty, cost and other factors; this existing method may show excess subjectivity for engineering design, which largely depends on the experience and knowledge of the designers.

Due to the design process imply large number of expert experience and knowledge, and this experience can be transformed into knowledge by combination result, data analysis and reasoning, which is called combination rules. And the way to get access to the knowledge of experts in such fields is the key in solving the problem. Therefore, this paper presents a new method applying rough set method reasoning rules in order to get the knowledge of those experts.

The basic concept of rough set theory

Rough set theory being put forward by Pawlak, a Poland scholars, is a kind of mathematical tools to depict the imperfect and uncertainty, and suitable for finding those underlying, potentially useful regularity in data, and those relationship and features of internal data[5]. And there are many factors affecting component integration, and these data are always redundant and fuzzy, and it is difficult to

process and calculate using traditional methods. The main target based on the rough set data analysis is reduction, and find out the implicit rules in mass data.

According to the rough set theory, component integration information model data is stored in a information table, also known as information system, the definition as: $S = (U, A, V, f)$, $U = \{x_1, x_2, \dots, x_n\}$ is object's non empty set; $A = \{a_1, a_2, \dots, a_n\}$ is attribute's non empty set; $V = \bigcup_{a \in A} V_a$ is attribute's range set; f is information fun; $f: U \times A \rightarrow V$, $f(x_i, a) \in V_a$. If $A = C \cup D$, $C \cap D = \emptyset$, C is the condition attribute set, D is the decision attribute set, then this kind of information system is also known as decision-making system, expressed as $T = (U, C \cup D, V, f)$.

For any $B \subseteq A$, then the indistinguishable relation for B in U is defined as:

$$IND(B) = \{(x_i, x_j) \in U \mid b \in B, b(x_i) = b(x_j)\} \quad (1)$$

The indistinguishable relation is a kind of equivalence relation. Easy to see, in the information system, one attribute corresponds to one equivalence relation. For anyone object subset $X \subseteq U$ and equivalence class $IND(B)$, X and B lower approximation and upper approximation were defined as:

$$B_*(X) = \bigcup \{Y \in U / IND(B) : Y \subseteq X\} \quad (2)$$

$$B^*(X) = \bigcup \{Y \in U / IND(B) : Y \cap X \neq \emptyset\} \quad (3)$$

for decision-making system $T = (U, C \cup D, V, f)$, C relative to D is domain for $POS_C(D) = \bigcup_{X \in U/D} B_*(X)$, $POS_C(D)$, In fact, is the set among which classification elements are divided into such set by attribute set D depended on attribute set C directly.

Calculation method analyses

Attribute reduction usually is essential steps in data mining process. This paper, utilizing attribute importance as heuristic information, can efficiently get access to the core of system. This paper starts from computation reduction set, and then gradually finds a minimum attribute reduction set of system, further, gets least classification rule of system.

Obviously, impact of each attributes is different for problem, the importance of this kind of influence can express through the auxiliary knowledge predefined or with a certain weight. This paper calculates importance degrees of attribute influence resulted from decision by decision-making system itself. In the decision-making system, utilizing mutual dependence of two attributes set $R \subseteq C$ and $P \subseteq D$ to define the importance of an attribute. Dependence degree of Attribute set between P and R is shown as $g_R(P)$. Definition as:

$$g_R(P) = \frac{card(POS_R(P))}{card(U)} \quad (4)$$

Among them, $card(U)$ indicates the number of elements contained in the collection U . Different attributes play different roles to dependencies between decision conditions attributes and decision attributes.

The attributes subset R' join attributes set R , the relative importance for the classification $U / IND(P)$ is defined as^[5]:

$$d_{R-R'}(P) = (g_R(P) - g_{R-R'}(P)) / g_R(P) \quad (5)$$

Reduction is a kind of attribute set that using one classification to express another classification from the perspective of the classification. All reduction intersections are defined as cores. And the degree of attribute importance shows that of the influence posed by classification decision, so the importance of attributes is used as heuristic information to calculate attribute core.

Parts combination relevant factors analysis

Based on rough set theory, parts combination information model data stored in an information table. Three questions proposed by Boothory are given as the necessary conditions to judge whether

two parts contacting with each other can be combined together. Those pair of parts that matches the conditions constitute a set of research object(line of the information system); factors that affect combination constitute the condition attribute set of the study object(row of the information system); results whether the parts combine together is defined as decision attribute, so the information system is decision-making system. The affecting factors of integration can also be considered as the related costs change caused by parts combination.

The cost of the assembly and disassembly operations. Since two parts are integrated to become one part, this part of the assembly and maintenance, remove operation of recovery processing stage will no longer be necessary during the assembly process, obviously parts integration will result in a reduction of the operating cost of the assembly and disassembly. As the detailed information of the parts at this time is unclear, it is difficult to give the exact numerical values, the degree of reduction of the detachably fees were respectively used 1 - low, 2 - high to representing.

Material use costs. The parts to be combination may be the same or may not be identical to their constituent material, thus use cost of the material invested in the merger may increase, decrease, or be unchanged, with 1 - decreased, 2 - the same, 3 - increased to respectively representing.

Manufacturing cost. Parts integration will usually cause manufacturing cost reduction. Original assembling features of parts are disappeared, processing surface must be reduced after merge, from this perspective, the manufacturing cost is reduced; However part merged later, the changes of component materials may lead to increase in manufacturing cost or may also result in lowering it. Generally manufacturing costs may increase may, decrease or be unchanged, respectively with 1 - decreased, 2 - the same, 3 - increased to respectively representing.

Design development and post-processing influence. Parts combination is a redesign optimization process in the original design that the design requirements of the optimization are met, so it will spend extra time and money. Meanwhile, after combination, because those structures are changed, which will cause related change for those pending combined related parts' structure and connection manner, which will increase the cost of the product design. The degree of increase of the design cost using 1 - low, 2 - high to represent.

Quality influence. The parts combination is a redesign optimization process based on the function of products, and the quality of those products should meet the requirements. But, as we all know, any parts of the processing errors are inevitable, so the parts' accumulated error will seriously affect the quality of the parts. Parts combination effectively reduces the number of parts of the product, so the cumulative error indirectly will also be reduced to different extent. Because of the difference that the importance, accuracy, etc, this kind of influence degree is also different, respectively for 1 - low, 2 -- high to represent. There are many factors affect parts merge, designers can added them depended on different requirement.

Case study

In table 1, it shows a part merger decision table of a controller. We take it for example to verify the proposed algorithm. We set the research object $U = \{x_1, x_2, \mathbf{L}, x_{12}\}$, $C = \{a_1, a_2, \mathbf{L}, a_5\}$ as condition attributes, $D = \{d\}$ as decision attribute, table 1 expressed the expert's design experience of whether parts can be integrated together. We should follow expert's experience to analyze merge mechanism, and then simplify this process thus obtaining merger rules of parts to guide design. It can be seen that there's no contradictory rules in the table, for it is a consistent decision table.

Table. 1 parts merger decision table

For U merger of parts	Loading and operational cost a_1	Material use cost a_2	Manufacturing cost a_3	Design and development and post-processing cost a_4	Quality influence a_5	Whether merger D
x_1	2	3	2	1	1	Yes
x_2	2	1	1	2	2	Yes
x_3	2	2	1	2	2	No
x_4	1	2	2	2	1	No
x_5	1	1	2	1	1	No
x_6	1	2	3	1	2	No
x_7	2	1	2	2	2	Yes
x_8	1	1	2	2	2	Yes
x_9	2	3	2	2	1	No
x_{10}	1	3	3	2	1	No
x_{11}	2	1	2	1	2	Yes
x_{12}	1	2	3	2	2	No

Calculation attributes core

Step 1: make initial attribute core $CORE_D(C)=f$

Step 2: using the formula (5) to calculate the related importance that condition attribute is to decision attribute.

$$U / IND(D) = \{ \{x_1, x_2, x_7, x_8, x_{11}\}, \{x_3, x_4, x_5, x_6, x_9, x_{10}, x_{12}\} \}$$

$$U / IND(C) = \{ \{x_1\}, \{x_2\}, \{x_3\}, \{x_4\}, \{x_5\}, \{x_6\}, \{x_7\}, \{x_8\}, \{x_9\}, \{x_{10}\}, \{x_{11}\}, \{x_{12}\} \}$$

$$POS_C(D) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}\}$$

$$g_C(D) = \frac{card(POS_C(D))}{card(U)} = \frac{12}{12} = 1$$

To calculate the related importance that a_1 is to D , a_1 from the decision table must be removed, and after that:

$$U / IND(C - \{a_1\}) = \{ \{x_1\}, \{x_2\}, \{x_3\}, \{x_4\}, \{x_5\}, \{x_6\}, \{x_7, x_8\}, \{x_9\}, \{x_{10}\}, \{x_{11}\}, \{x_{12}\} \}$$

$$POS_{C-\{a_1\}}(D) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}\}$$

$$g_{C-\{a_1\}}(D) = \frac{card(POS_{C-\{a_1\}}(D))}{card(U)} = \frac{12}{12} = 1$$

$$d_{C-\{a_1\}}(D) = (g_C(D) - g_{C-\{a_1\}}(D)) / g_C(D) = 1 - \frac{12/12}{12/12} = 0$$

For a_1 , because of $d_{C-\{a_1\}}(D)=0$, according to the algorithm 1, it can be concluded that $a_1 \notin CORE_D(C)$. By the same token can reach the conclusion that attribute a_2 and a_3 , a_4 , a_5 to attribute of the relative importance of D respectively:

$$d_{C-\{a_2\}}(D) = (g_C(D) - g_{C-\{a_2\}}(D)) / g_C(D) = 1/6 \neq 0$$

$$d_{C-\{a_3\}}(D) = (g_C(D) - g_{C-\{a_3\}}(D)) / g_C(D) = 0$$

$$d_{C-\{a_4\}}(D) = (g_C(D) - g_{C-\{a_4\}}(D)) / g_C(D) = 1/6 \neq 0$$

$$d_{C-\{a_5\}}(D) = (g_C(D) - g_{C-\{a_5\}}(D)) / g_C(D) = 0$$

Step 3: according to the algorithm 1, it can be concluded that: $CORE_D(C) = \{a_2, a_4\}$

Calculations attribute reduction set

Step 1: let $B = CORE_D(C) = \{a_2, a_4\}$, $A = C - CORE_D(C) = \{a_1, a_3, a_5\}$

Step 2: according to the algorithm 2, for $B = \{a_2, a_4\} \neq f$

$$U / IND(B) = \{ \{x_1\}, \{x_2, x_7, x_8\}, \{x_3, x_4, x_{12}\}, \{x_5, x_{11}\}, \{x_6\}, \{x_9, x_{10}\} \}$$

$$POS_B(D) = \{x_1, x_2, x_3, x_4, x_6, x_7, x_8, x_9, x_{10}, x_{12}\}$$

$$d_B(D) = (g_c(D) - g_B(D)) / g_c(D) = 1 - \frac{10/12}{12/12} = 1/6 \neq 0$$

$$d_{BU\{a_1\}}(D) = (g_c(D) - g_{BU\{a_1\}}(D)) / g_c(D) = 0$$

$$d_{BU\{a_3\}}(D) = (g_c(D) - g_{BU\{a_3\}}(D)) / g_c(D) = 1 - \frac{10/12}{12/12} = 1/6 \neq 0$$

$$d_{BU\{a_5\}}(D) = (g_c(D) - g_{BU\{a_5\}}(D)) / g_c(D) = 0$$

Step 3: according to the algorithm 2, it can be concluded that the minimum attribute reduction such as: $RED_D(C) = BU\{a_1\} = \{a_1, a_2, a_4\}$ and $RED_D(C) = BU\{a_5\} = \{a_2, a_4, a_5\}$. take $RED_D(C) = \{a_1, a_2, a_4\}$ for instance, removing the same decision-rule, after reduction, table is shown in chart 2.

Table. 2 parts integration decision table after reduction

U	a_1	a_2	a_4	D
x_1	2	3	1	Yes
x_2, x_7	2	1	2	Yes
x_3	2	2	2	No
x_4, x_{12}	1	2	2	No
x_5	1	1	1	No
x_6	1	2	1	No
x_8	1	1	2	Yes
x_9	2	3	2	No
x_{10}	1	3	2	No
x_{11}	2	1	1	Yes

Attribute value reduction. Be similar with the attribute reduction, but also excess conditions attribute value can be cut from decision table. Since the size of the decision table has been reduced greatly, so the data analysis can be used, it has the advantages of simple calculation^[6].

Take the first decision rule in table 2 $d_1: (a_1, 2) \wedge (a_2, 3) \wedge (a_4, 1) \rightarrow (D, \text{yes})$ for instance, $(a_2, 3) \wedge (a_4, 1) \rightarrow (D, \text{yes})$ is the consistent decision rule, $(a_1, 2) \wedge (a_4, 1) \rightarrow (D, \text{yes})$ and x_{11} keep consistently, still $(a_1, 2) \wedge (a_4, 3) \rightarrow (D, \text{yes})$ and x_9 are contradictorily, so $(a_1, 2), (a_2, 3)$ can be about to divide out, but $(a_4, 1)$ cannot. According to this method, it can get all core value after the condition attribute reduction, which was shown in table 3.

Table. 3 the decision table which only contains attributes core value

U	a_1	a_2	a_4	D
x_1	-	-	1	Yes
x_2, x_7	-	1	-	Yes
x_3	-	2	-	No
x_4, x_{12}	-	2	-	No
x_5	1	-	1	No
x_6	-	-	-	No
x_8	-	1	2	Yes
x_9	-	3	2	No
x_{10}	-	3	-	No
x_{11}	2	-	-	Yes

Minimum decision classification rule. According to the core value table above, it can further obtain reduction of every decision rule^[7]. Merging the same rules, it can get the minimum decision rule, which is about the equivalent of merger parts with the table 1.

Rules 1: parts pair of the decreasing cost of loading with the medium degree and of redesigned cost increase with the low degree can be integrated together.

Rules 2: parts pair of the increasing material use cost and of redesigned cost increase with the low degree can be integrated together.

Rules 3: parts pair of the decreasing cost of loading with the medium degree and of decreasing material use cost can be integrated together.

Rules 4: parts pair of the decreasing material use cost and of redesigned cost increase with the medium degree can be integrated together.

It can be seen, in the redesigning section, mainly considering that it can be effectively decided whether parts pair can be integrated together depending on the redesign increasing cost and material use changing cost caused by parts integration.

Summary

This paper puts forward a kind of application with the rough set theory mining parts integration concerning expert decision knowledge method and puts to use the attribute importance as heuristic information to calculate the attribute core and attribute reduction, its characteristic is simple, quick, and easy to implement, and thus shows great practical value for decision analysis. Example is exhibited here testing its validity. However, the product structures in this real world differs in thousands ways, this paper only provides a method of digging implicit knowledge. More knowledge of integration decision can be got through the analysis of the typical product structure, and further, the whole knowledge will be expanded and perfected.

Acknowledgements

This work was financially supported by the Inner Mongolia Natural Science Foundation (2013MS0713).

References

- [1] Uno Motoo, Nakajima Taku. Recycling technologies in Hitachi [A]. Proceedings of First International Symposium on Environmentally Conscious Design and Inverse Manufacturing. Tokyo, Japan, 1999: 273~276.
- [2] YangJianGuo, LiBeiZhi, shao shi-huang, an approach to. But assembling design again [J]. J micro computer applications, 1997 (4): 37 to 39.
- [3] Chou Fu-Mao. A Methodology for Component Integration in Product Design [M]. New York, 2002, 5.
- [4] GuTingQuan, GaoGuoAn, XuXiangYang. Based on the analysis of the function of quantitative design for assembly method [J]. Mechanical engineering, 1998, (6): 3-5.
- [5] Pawlak Z. Rough set [J]. International Journal of Computer and Information Sciences, 1982, 11(5):341~356.
- [6] LiuQing. Rough sets and Rough reasoning [M]. Beijing: science press. 2001.
- [7] HuangLin. Rough set theory and its application [M]. Chongqing: chongqing press. 1999.