Approaches to Attributes Reduction Based on Ant Colony Optimization

Honghua Chen Zheng Pei Li Zhang

School of Mathematics & Computer Engineering, Xihua University, Chengdu 610039, China

Abstract
As we known, ants are blind and single ant’s capacity is limited. But researchers have studied that ant colonies have egregious ability to search shorter path which loads between their nests and foods. Inspired from this and based on ant colony optimization(ACO), we propose a new approach about attributes reduction of rough sets. Using this approach, we can reduce attributes as maximum as possible, obtain many different results synchronously and reduce the scope of core attributes. Experiment shows that by this new approach, better results can be obtained.

Keywords: Rough sets, Attributes reduction, Ant colony optimization, Binary discernibility matrices

1. Introduction

The need for discovering knowledge from information increases with the forward rapid development of the recent civilization [1]-[3]. One of the most useful methods for knowledge discovering is the rough set theory(RST) which proposed by Z. Pawlak[4]. And attributes reduction is one of the most important subjects of RST. Due to attributes reduction is a NP-hard problem, most of approaches about attributes reduction are heuristic[6], [7]. As an approach of attributes reduction, binary discernibility matrices have many interesting properties[8]. In this paper, we propose a new approach about attributes reduction which based on binary discernibility matrices and ant colony optimization(ACO).

As we known, a wonderful self-organization behavior will usually be produced from the collective behavior of social animals. Take a colony of ants for example, biologists had studied the phenomenon carefully and found that ants cooperate to find shorter routing path by means of indirect communications using a kind of substance called “pheromone”[10]. Inspired from this, ant system(AS) was proposed by Italian researchers M. Dorigo, V. Maniezzo and A. Colorni[13]. This algorithm was proposed to solve the travelling salesman problem(TSP) firstly. These years, M. Dorigo etc. have expanded AS algorithm, and proposed a new generaless optimization technique: ant colony optimization(ACO)[10]. All the algorithms which accord with the frame of ACO can be called as ACO algorithm[14]. Many scholars are attracted to study ACO and in the past ten years the algorithm has been widely applied to the fields of combinatorial optimization, network routing, data mining etc., [15]-[18].

In the next Section, some preliminaries knowledge about attributes reduction and binary discernibility matrix is introduced. Section 3 is devoted to propose a new approach based on ACO and binary discernibility matrix. Experiment about the new approach is given in Section 4. Conclusion is Section 5.

2. Preliminaries knowledge

Let an information system be $S = \langle U, A, V, f \rangle$, where $U$ is a non-empty finite set of objects, $A$ is a non-empty finite set of attributes, $V = \bigcup_{a \in A} V_a$ and $V_a$ is the domain of $a$, $f : U \times A \rightarrow V$ is information function. Let $R$ be an equivalence relation on $U$. And a set of equivalence classes with respect to $R$ is as follows:

$$U/R = \{[x]_R | x \in U\},$$

where $[x]_R = \{x_j | (x, x_j) \in R\}$. Each non-empty subset $B \subseteq A$ determines an indiscernibility relation as follows:

$$R_B = \{(x, y) \in U^2 | f(x, a) = f(y, a), \forall a \in B\},$$

then $R_B$ is an equivalence relation on $U$ and let $[x]_B = \{x_j | (x, x_j) \in R_B\}$.

Definition 1 [5] $(U, A, V, f)$ is an information system, for $B \subseteq A$, if $R_A = R_B$, then $B$ is a partition consistent set. If $R_A = R_B$, and $B' \subseteq B$, s.t., $R_{B'} = R_A$, then $B$ is a partition reduct.
For an information system, attributes reduction is a NP-hard problem. But it can be solved by heuristic approaches. In heuristic approaches, weights of attributes are obtained as follows [11]:

\[ M = \begin{pmatrix}
(x_1, x_2) & 1 & 1 & 1 & 1 \\
(x_1, x_3) & 1 & 0 & 1 & 1 \\
(x_1, x_4) & 1 & 1 & 0 & 1 \\
(x_1, x_5) & 1 & 1 & 1 & 0 \\
(x_2, x_3) & 0 & 1 & 0 & 1 \\
(x_2, x_4) & 1 & 0 & 1 & 1 \\
(x_2, x_5) & 1 & 0 & 0 & 1 \\
(x_3, x_4) & 1 & 1 & 1 & 0 \\
(x_3, x_5) & 1 & 1 & 0 & 1 \\
(x_4, x_5) & 0 & 1 & 0 & 1 \\
\end{pmatrix}. \]

According to Equation (2), the weight of attribute \( a_1 \) can be calculated as follows:

\[ W(a_1) = \frac{1}{5} + \frac{1}{4} + \frac{1}{3} + \frac{1}{4} + \frac{1}{3} + \frac{1}{4} + \frac{1}{4} \]

\[ = \frac{132}{60}. \]

### 3. Approach based on ant colony optimization

Let \( b_i(t_i)(i = 0, 1, \ldots, |A|) \) denote the number of agents (artificial ants) in node \( i \) at the time \( t_i \), then

\[ m = \sum_{i=0}^{n} b_i(t_i) \]

deals with the total number of agents. Let \( \tau_i(t_i) \) denote the amount of pheromone on the edge: \((i, j)\), and \( tabu_i \) denote the nodes which agent \( k \) has gone.

**Definition 4** For a general information system \((U, A, V, f)\) and its binary discernibility matrix \((\overline{U}, A, \{0, 1\}, \psi)\), \( \forall a_i \in A, \) let \( a_i \) denote a node, and add a virtual node noted as \( \Theta \). \( \forall (a_i, a_j) \in (A \cup \{\Theta\})^2 \), there is an edge connects between node \( a_i \) and node \( a_j \), and \( \forall (x_i, x_j) \in U, \) let \( \psi((x_i, x_j), \Theta) = 0. \) Then, let all the agents move from \( \Theta \) till making circles. And the anticipant degree of moving from \( a_i \) to \( a_j \) is defined as follows:

\[ \eta_{a_i a_j} = \sum_{(x_i, x_j) \in U} \psi((x_i, x_j), a_j) - \nabla_{a \in \Theta \cup tabu_k} (\psi((x_i, x_j), a_j) \wedge \psi((x_i, x_j), a)). \]

### Table 1: A general information system

<table>
<thead>
<tr>
<th>( U \setminus A )</th>
<th>humidity ((a_1))</th>
<th>rainy ((a_2))</th>
<th>frosty ((a_3))</th>
<th>sunny ((a_4))</th>
<th>windy ((a_5))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monday(x_1)</td>
<td>Low</td>
<td>No</td>
<td>Yes</td>
<td>Shady</td>
<td>Low</td>
</tr>
<tr>
<td>Tuesday(x_2)</td>
<td>Average</td>
<td>Yes</td>
<td>No</td>
<td>Commonly</td>
<td>Average</td>
</tr>
<tr>
<td>Wednesday(x_3)</td>
<td>Average</td>
<td>No</td>
<td>No</td>
<td>Sunshine</td>
<td>Average</td>
</tr>
<tr>
<td>Thursday(x_4)</td>
<td>High</td>
<td>Yes</td>
<td>Yes</td>
<td>Shady</td>
<td>Average</td>
</tr>
<tr>
<td>Friday(x_5)</td>
<td>High</td>
<td>Yes</td>
<td>No</td>
<td>Shady</td>
<td>High</td>
</tr>
</tbody>
</table>

**Example 1** Considering a general information system (See Table 1.), it is about the weather of five weekdays. Where \( U = \{\text{Monday, Tuesday, Wednesday, Thursday, Friday}\}, A = \{\text{humidity, rainy, frosty, sunny, windy}\} \). The matrix \( M \) of \((\overline{U}, A, \{0, 1\}, \psi)\) can be calculated as follows according to Definition 2.
Where \( a_j \in J_k = (A - \text{tabu}_k) \cup \{\emptyset\} \), and \( J_k \) denotes the nodes which agent \( k \) can go. When \( \text{tabu}_k = \{\emptyset\} \), \( \eta_{a_j} = \sum_{(x_i, x_j) \in U} \psi((x_i, x_j), a_j) \).

In the iterative process: \( t_t \rightarrow t_{t+1} \), the probability of agent \( k \) to select node \( a_j \) when it is in node \( a_i \) is defined as follows:

\[
p^k_{a_i, a_j}(t_t) = \begin{cases} \frac{[\tau_{a_i, a_j}(t_t)]^\alpha \eta_{a_j}^\beta}{\sum_{a_k \in J_k} [\tau_{a_i, a_k}(t_t)]^\alpha \eta_{a_k}^\beta}, & \text{if } a_j \in J_k; \\ 0, & \text{otherwise.} \end{cases}
\]

(4)

Where, \( \eta_{a_i, a_j} \) is a heuristic parameter which denotes the anticipant degree of moving from node \( a_i \) to node \( a_j \), and \( \alpha, \beta \) are both parameters which denote the significance of accumulated information and heuristic factor respectively. In many cases, \( \eta_{a_i, a_j} = \frac{1}{d_{a_i, a_j}} \), where \( d_{a_i, a_j} \) denotes the distance between node \( a_i \) and \( a_j \). But in this approach, \( \eta_{a_i, a_j} \) is obtained according to Equation (3). After agents making circles, update the amount of pheromone on these paths according to the following equation:

\[
\tau_{a_i, a_j}(t_{t+1}) = (1 - \rho) \cdot \tau_{a_i, a_j}(t_t) + \Delta \tau_{a_i, a_j}(t_t),
\]

(5)

where \( \rho(0 < \rho < 1) \) denotes the vaporizing parameter of pheromone, \( 1 - \rho \) denotes the supportable parameter of pheromone, and \( \Delta \tau_{a_i, a_j}(t_t) \) denotes the increase amount of pheromone on the path: \( a_i \rightarrow a_j \) in the iterative process: \( t_t \rightarrow t_{t+1} \). And let

\[
\Delta \tau_{a_i, a_j}(t_t) = \sum_{k=1}^{m} \Delta^k \tau_{a_i, a_j}(t_t).
\]

(6)

Where, \( \Delta^k \tau_{a_i, a_j}(t_t) \) denotes the amount of pheromone which agent \( k \) leaves on the path: \( a_i \rightarrow a_j \) in the iterative process: \( t_t \rightarrow t_{t+1} \), and \( m \) denotes the number of agents. Let

\[
\Delta^k \tau_{a_i, a_j}(t_t) = \begin{cases} Q \frac{\psi((x_i, x_j), a_j \in \text{tabu}_k)}{N_b(t_t)}, & \text{if agent } k \text{ moves through } \text{path: } a_i \rightarrow a_j; \\ 0, & \text{otherwise.} \end{cases}
\]

(7)

Where \( Q \) is a constant and \( N_b(t_t) \) denotes the number of nodes that agent \( k \) has passed(except the node \( \emptyset \) ) in the iterative process: \( t_t \rightarrow t_{t+1} \).

**Remark** 1 Based on ACO, we can construct an ant system with respect to an information system. Then let agents simulate the real ants to search shorter paths. As we known, ants select paths according to the pheromone left on the paths. Then we define some parameters in order to guide these agents to make circles. And these circles are all the partition consistent sets of the information system.

**Theorem 1** For a general information system \((U, A, V, f)\) and an agent in node \( a_i \), \( \forall a_j \in (A - \text{tabu}_k), \eta_{a_i, a_j} = 0 \) if and only if \( \text{tabu}_k - \{\emptyset\} \) is a partition consistent set.

**Proof** \( \forall a_j \in (A - \text{tabu}_k), \eta_{a_i, a_j} = \sum_{(x_i, x_j) \in U} \psi((x_i, x_j), a_j \in (A - \text{tabu}_k)) \). Due to \( \forall (x_i, x_j) \in U, \psi((x_i, x_j), a_j) \geq \psi((x_i, x_j), a_j \in (A - \text{tabu}_k)), \psi((x_i, x_j), a_j) \Rightarrow \forall (x_i, x_j), (x_i, x_j, a_j) \in (A - \text{tabu}_k), \psi((x_i, x_j), a_j) \Rightarrow \forall (x_i, x_j) \in U, \psi((x_i, x_j), a_j) \in \emptyset \Rightarrow \forall (x_i, x_j) \in U, \psi((x_i, x_j), a_j) = 0 \)

Therefore, \( \forall a_j \in (A - \text{tabu}_k), \eta_{a_i, a_j} = 0 \) if and only if \( \text{tabu}_k - \{\emptyset\} \) is a partition consistent set.

**Note** 1 For agent \( k \) is in node \( a_i \), then \( \forall a_j \in J_k, \eta_{a_i, a_j} = 0 \), for any \( a_j \in J_k \), let

\[
p^k_{a_i, a_j}(t_t) = \frac{[\tau_{a_i, a_j}(t_t)]^\alpha \eta_{a_j}^\beta}{\sum_{a_k \in J_k} [\tau_{a_i, a_k}(t_t)]^\alpha \eta_{a_k}^\beta}.
\]

(8)
Table 2: Status of crops in China

<table>
<thead>
<tr>
<th>$U \setminus A$</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$a_3$</th>
<th>$a_4$</th>
<th>$a_5$</th>
<th>$a_6$</th>
<th>$a_7$</th>
<th>$a_8$</th>
<th>$a_9$</th>
<th>$a_{10}$</th>
<th>$a_{11}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1983</td>
<td>10,350</td>
<td>29,854</td>
<td>16,695</td>
<td>1,402</td>
<td>1,316</td>
<td>32</td>
<td>55</td>
<td>0</td>
<td>332</td>
<td>0</td>
<td>546</td>
</tr>
<tr>
<td>1984</td>
<td>10,839</td>
<td>30,679</td>
<td>15,840</td>
<td>1,690</td>
<td>1,445</td>
<td>37</td>
<td>59</td>
<td>149</td>
<td>432</td>
<td>0</td>
<td>619</td>
</tr>
<tr>
<td>1985</td>
<td>11,382</td>
<td>33,140</td>
<td>15,588</td>
<td>1,927</td>
<td>1,655</td>
<td>47</td>
<td>59</td>
<td>110</td>
<td>555</td>
<td>0</td>
<td>705</td>
</tr>
<tr>
<td>1986</td>
<td>11,896</td>
<td>33,719</td>
<td>16,623</td>
<td>2,112</td>
<td>1,796</td>
<td>59</td>
<td>62</td>
<td>188</td>
<td>555</td>
<td>333</td>
<td>824</td>
</tr>
<tr>
<td>1987</td>
<td>12,101</td>
<td>32,773</td>
<td>18,034</td>
<td>2,216</td>
<td>1,835</td>
<td>79</td>
<td>72</td>
<td>219</td>
<td>590</td>
<td>379</td>
<td>955</td>
</tr>
<tr>
<td>1988</td>
<td>12,538</td>
<td>34,222</td>
<td>20,153</td>
<td>2,480</td>
<td>2,018</td>
<td>96</td>
<td>80</td>
<td>274</td>
<td>696</td>
<td>419</td>
<td>1,061</td>
</tr>
<tr>
<td>1989</td>
<td>12,805</td>
<td>35,281</td>
<td>21,164</td>
<td>2,629</td>
<td>2,123</td>
<td>107</td>
<td>96</td>
<td>282</td>
<td>720</td>
<td>436</td>
<td>1,152</td>
</tr>
<tr>
<td>1990</td>
<td>13,021</td>
<td>36,241</td>
<td>21,002</td>
<td>2,857</td>
<td>2,281</td>
<td>126</td>
<td>107</td>
<td>323</td>
<td>795</td>
<td>475</td>
<td>1,237</td>
</tr>
<tr>
<td>1991</td>
<td>13,193</td>
<td>36,965</td>
<td>20,621</td>
<td>3,144</td>
<td>2,452</td>
<td>154</td>
<td>118</td>
<td>395</td>
<td>922</td>
<td>524</td>
<td>1,351</td>
</tr>
<tr>
<td>1992</td>
<td>13,485</td>
<td>38,421</td>
<td>20,733</td>
<td>3,431</td>
<td>2,635</td>
<td>180</td>
<td>125</td>
<td>454</td>
<td>1,020</td>
<td>564</td>
<td>1,557</td>
</tr>
<tr>
<td>1993</td>
<td>13,988</td>
<td>39,300</td>
<td>21,731</td>
<td>3,842</td>
<td>2,854</td>
<td>234</td>
<td>137</td>
<td>274</td>
<td>1,180</td>
<td>564</td>
<td>1,823</td>
</tr>
<tr>
<td>1994</td>
<td>14,919</td>
<td>41,462</td>
<td>24,053</td>
<td>4,499</td>
<td>3,205</td>
<td>327</td>
<td>101</td>
<td>755</td>
<td>1,479</td>
<td>609</td>
<td>2,143</td>
</tr>
<tr>
<td>1995</td>
<td>15,862</td>
<td>44,160</td>
<td>27,865</td>
<td>5,260</td>
<td>3,648</td>
<td>415</td>
<td>202</td>
<td>935</td>
<td>1,677</td>
<td>673</td>
<td>2,517</td>
</tr>
<tr>
<td>1996</td>
<td>13,361</td>
<td>36,284</td>
<td>23,728</td>
<td>4,584</td>
<td>3,158</td>
<td>356</td>
<td>181</td>
<td>833</td>
<td>1,897</td>
<td>736</td>
<td>2,813</td>
</tr>
<tr>
<td>1997</td>
<td>14,542</td>
<td>40,035</td>
<td>25,576</td>
<td>5,269</td>
<td>3,596</td>
<td>441</td>
<td>213</td>
<td>979</td>
<td>1,897</td>
<td>681</td>
<td>3,692</td>
</tr>
<tr>
<td>1998</td>
<td>14,803</td>
<td>42,256</td>
<td>26,904</td>
<td>5,724</td>
<td>3,884</td>
<td>480</td>
<td>235</td>
<td>1,056</td>
<td>2,921</td>
<td>745</td>
<td>3,907</td>
</tr>
<tr>
<td>1999</td>
<td>15,025</td>
<td>43,020</td>
<td>27,926</td>
<td>5,949</td>
<td>3,891</td>
<td>505</td>
<td>251</td>
<td>1,116</td>
<td>2,135</td>
<td>867</td>
<td>4,122</td>
</tr>
<tr>
<td>2000</td>
<td>15,152</td>
<td>44,682</td>
<td>29,032</td>
<td>6,125</td>
<td>4,031</td>
<td>533</td>
<td>274</td>
<td>1,208</td>
<td>2,243</td>
<td>919</td>
<td>4,279</td>
</tr>
<tr>
<td>2001</td>
<td>14,996</td>
<td>45,743</td>
<td>29,826</td>
<td>6,334</td>
<td>4,184</td>
<td>549</td>
<td>293</td>
<td>1,210</td>
<td>2,337</td>
<td>1,123</td>
<td>4,374</td>
</tr>
<tr>
<td>2002</td>
<td>15,189</td>
<td>46,292</td>
<td>31,655</td>
<td>6,587</td>
<td>4,327</td>
<td>585</td>
<td>317</td>
<td>1,250</td>
<td>2,462</td>
<td>1,400</td>
<td>4,566</td>
</tr>
<tr>
<td>2003</td>
<td>15,500</td>
<td>46,602</td>
<td>34,054</td>
<td>6,933</td>
<td>4,519</td>
<td>631</td>
<td>357</td>
<td>1,312</td>
<td>2,607</td>
<td>1,849</td>
<td>4,705</td>
</tr>
<tr>
<td>2004</td>
<td>15,738</td>
<td>48,189</td>
<td>36,639</td>
<td>7,245</td>
<td>4,702</td>
<td>676</td>
<td>399</td>
<td>1,351</td>
<td>2,724</td>
<td>2,368</td>
<td>4,902</td>
</tr>
</tbody>
</table>

Theorem 2 Agents make circles (coming back to node $\Theta$) only when $\text{tabu}_k - \{\Theta\}$ is a partition consistent set in general information system.

Proof 2 $\forall (x_i, x_j) \in \overline{U}, \psi((x_i, x_j), \Theta) = 0$, so, when an agent is in node $a_i$, $\eta_{a_i, \Theta} = 0$, then $\rho_{\Theta, \Theta}(t_i) \neq 0$ if and only if $\forall a \in J_k$, $\eta_{a, a} = 0$ (according to Note 1). So agent $k$ makes circles, i.e., coming back to node $\Theta$ only when $\forall a_j \in J_k = (A - \text{tabu}_k) \cup \{\Theta\}$, $\eta_{a_j, a} = 0$ i.e., only when $\forall a_j \in (A - \text{tabu}_k)$, $\eta_{a, a} = 0$. Then according to Theorem 1, $\text{tabu}_k - \{\Theta\}$ is a partition consistent set.

Remark 2 According to Theorem 2, agents make circles are all partition consistent sets in general information system. Then using ant colony optimization, we can obtain many different partition consistent sets synchronously, and we can obtain the partition reduct possibly. Furthermore, the minimal reduct can be obtained possibly.

4. Experiment

Table 2 is a general information about the status of crops in China (http://www.agri.gov.cn/sjzl/baips/h/WB2005.htm#12). There are 22 objects which represent the years and 11 attributes which represent the outputs of various crops, and denoted as $a_1, a_2, \cdots, a_{11}$.

4.1. Experiment about reducing attributes as maximum as possible

In this experiment, let $\alpha = 1, \beta = 1, Q = 10, \rho = 0.5$, and according to attributes’ value orderly, we disperse their domain into average 4 classes which denoted as 1, 2, 3, 4. Then the continuous value information system can be transformed into a discrete value information system. According to Theorem 2, for each agent in every cycle, there is a partition consistent set (partition reduct) correspondingly. Our experiment has been performed 4 times according to 4 different cases: 10 agents and 5 cycles, 10 agents and 10 cycles, 100 agents and 10 cycles, 10 agents and 100 cycles (See Fig 1). From Fig 1, we can find that: for different cases, the number of partition consistent sets or partition reducts which contain 5 attributes is the maximum. And we have affirmed that the number of attributes in the minimal reducts of this information system is 5 when it is dispersed averagely.
According to the theory of ACO, we may get in the partial optimization sometimes. In this paper, that is partition consistent set or partition reduct with more attributes than other partition reduct. But we can still use this new approach to obtain many results synchronously.

**Remark 3** The time cost: $T_c$ for this approach can be estimated as: $T_c \leq O(m \cdot n \cdot |U|^2 \cdot |A|^2)$, where, $m$ denotes the number of agents and $n$ denotes cycle count.

For the binary discernibility matrix $\langle \bar{U}, A, \{0, 1\}, \psi \rangle$ with respect to a given information system $\langle U, A, V, f \rangle$, $|\bar{U}| \leq \frac{|U|(|U|-1)}{2}$. And for agent $k$, when it is in the original node, i.e., $\Theta$, the time cost: $t_k^1$ for selecting next node is: $t_k^1 \leq O\left(\frac{|U|(|U|-1)}{2}, |A|\right)$. Then, when this agent has select one node, there are $|A| - 1$ attributes left, so, the time cost: $t_k^2$ for selecting next node is: $t_k^2 \leq O\left(\frac{|U|(|U|-1)}{2}, |A| - 1\right)$. And so on, the time cost for agent $k$ coming back to the node $\Theta$ is: $t_k^h \leq \sum_{i=0}^{(|U|-1)} \frac{O\left(\frac{|U|(|U|-1)}{2}\right)}{2} \cdot (|A| - i) \leq \sum_{i=0}^{(|U|-2)} \frac{O\left(\frac{|U|(|U|-1)}{2}\right)}{2} \cdot (|A| - i) = O\left(\frac{|U|(|U|-1)}{2}\right).$ So, for $m$ agents and $n$ cycle count, the total time cost: $T_c \leq n \cdot \sum_{k=1}^{m} t_k \leq O(m \cdot n \cdot |U|^2 \cdot |A|^2)$. From Fig 2., we can obtain affirmance of the time cost in this experiment, and we can reach a conclusion that the time cost is polynomial with numbers of attributes, objects, agents and cycle count, not exponential.

**4.2. Remark about obtaining many different results synchronously and reducing the scope of core attributes**

When using the approach which based on weights of attributes to reduce information system, we can obtain only one result always in this experiment. Based on the discussion above, for kinds of cases(different numbers of agents or cycles), we can obtain at most $a \times b$ partition consistent sets. Where $a$ denotes the number of agents and $b$ denotes the number of cycles. In despite of that there are much of the same results, we can still obtain many different results synchronously through this new approach. Then these different results can be selected to make different decisions.

And according to Definition 3., ordinarily, the more different partition consistent sets(partition reducts) we have, the smaller scope of core attributes is. Because we can obtain many different results(partition consistent sets or partition reducts) using this new approach, then we can reduce the scope of core attributes. Later, we can use the most important attributes in restricted cases.

**5. Conclusion**

In this paper, we present an approach about attributes reduction based on ant colony optimization(ACO). By this approach, we can obtain the results as simply as possible and
reduce the scope of core attributes in order to assist managers. In the future, we can apply this approach into much larger information systems and knowledge reduction of decision information systems.

Acknowledgement

The authors would like to thank the anonymous referees for their very constructive comments. This work is supported by the Young Foundation of Sichuan Province (Grant No.06ZQ026-037) and the Education Department Foundation of Sichuan Province (Grant No.2005A121, 2006A084).

References