A New Discernibility Matrix Based on Distribution Reduction

HaiTao Wu
Huanghuai University, Zhumadian, China
Email: wwwht09@163.com

Abstract—In inconsistent information systems, the reduction based on positive region are not equivalent to that based on distribution reduction. The problem of discernibility matrix in algebraic view is analyzed, and a new discernibility matrix based on distribution reduction is proposed in this paper. This algorithm considers whether the objects compared are consistent, analyses in detail the degree of inconsistency and the distributing proportion of their conditional equivalent classes in decision classes, and the reduction in distribution reduction view is acquired finally. The theoretical analysis and simulation instance shows that this algorithm is feasible and effective in practice.

Index Terms—rough sets, distribution reduction, discernibility matrix

I. INTRODUCTION

Rough set theory is a kind of new data analysis method to deal with vagueness and uncertainty[1].In recent years, it has been rapidly developed in the field of intelligent information processing[2]. Attribute reduction can cuts down the whole account of attribute, and remains changeless classification capability of system information. So attribute reduction is a foundational and key issue of knowledge acquiring in theoretical study of rough set.

The discernibility matrix was proposed by Skowron[3], this method provides a normal and precise form to acquire core and reduction. Hu improved it, and put forward a fact that an element of matrix belonged to core if and only if it is a single attribute[4]. The literature[5] pointed out that HU method is application only to consistent decision table, then improved it again.

This paper analyses discernibility matrices mentioned above in detail, and proposes a succinct method based on distribution reduction to acquire reduction. This method not only considers if two objects compared are conflicting, but also analyses the inconsistent degree and the distribution proportion of their condition class among all decision classes, and gains all reductions of decision on basis of distribution reduction. The theoretic analysis shows that this algorithm is applicable to consistent and inconsistent decision table, and be equal to information theory view.

II. DISCERNIBILITY MATRIX BASED ON POSITIVE REGION

An decision information system (also called decision system, decision table, or information system) is denoted by $IS=\langle U,A,V,f \rangle$, where $U=\{x_1,x_2,\cdots,x_n\}$ is the universe which is a finite set of objects, $A=C \cup D$ is a finite set of attributes which is further classified into two disjoint subsets, condition attributes $C$ and decision attributes $D$. $V=\bigcup_{x_i \in U}V_{x_i}$ and $V_{x_i}$ is a domain of attribute $a$. $f:U \times A \rightarrow V$ is a total function such that $f(x_i,a) \in V_{x_i}(x_i \in U, a \in A)$.

Definition 1 Let $R(R \subseteq A), x_i,x_j \in U$, define a binary relation $IND(R)=\{(x_i,x_j) \in U^2 : \forall a \in R, f(x_i,a)=f(x_j,a)\}$.

It can be said that $x_i$ and $x_j$ are indiscernible by a set of attributes $R$ in $A$ if and only if $f(x_i,a)=f(x_j,a)$ for $\forall a \in R$. $IND(R)$ is an equivalence relation on $R$ for every $\forall a \in R$. For any element $x_i$ of $U$, the equivalence class of $x_i$ in relation $R$ is represented as $[x_i]_{IND(R)}$.

Suppose the partition of the universe induced by $C$ and $D$ be: $U/IND(C)=\{X_1,X_2,\cdots,X_n\}$, $U/IND(D)=\{Y_1,Y_2,\cdots,Y_m\}$. $U/IND(C)$ is called condition classification sets, and $U/IND(D)$ is called decision classification set.

Definition 2 Let $X \subseteq U$, a pair of approximation operators $R(X), \overline{R}(X)$ are defined as below:

$\overline{R}(X)=\bigcup\{X_i : X_i \in U, U/IND(R) \land X_i \subseteq X\}$

$R(X)=\bigcup\{X_i : X_i \in U, U/IND(D) \land X_i \cap X \neq \emptyset\}$

The lower approximation $R(X) = U/IND(R)$ is the greatest definable set contained in $X$, and the upper approximation $\overline{R}(X) = U/IND(D)$ is the least definable set containing $X$.

Definition 3 Let an information system $IS=\langle U,A,V,f \rangle$, positive region of the partition $U/D$ with respect to $C$ is defined by: $POS_C(D) = \bigcup_{x_i \in C}C(x_i)$. Positive region is the set of all elements of $U$ that can be uniquely classified to blocks of the partition $U/D$, by means of $C$.

Attribute reduction and core is the important research issue of rough set, so it received more attention of researchers in rough set field. The literature [3] proposed a method to get core and reduction in consistent decision table, on basis of Skowron discernibility matrix. The discernibility matrix in literature [3] was denoted as below:

$$m_{ij} = \begin{cases} \#\{a \in C \land f(x_i,a) \neq f(x_j,a)\} & \text{if } (f(x_i,D) \neq f(x_j,D)) \\ \emptyset & \text{otherwise} \end{cases}$$
The literature [4] pointed out that the literature only take consistent condition into account when objects belonged to different decision attribute value were compared, so this method is applicable only to consistent decision table. The literature [4] improved this method, defined improved discernibility matrix as below:

\[
m_y = \begin{cases}
\frac{\mu_a \cap \mu_b \cap \mu_c}{\mu_a \cap \mu_b \cap \mu_c} & \text{otherwise}
\end{cases}
\]

Where \([x_i]_{IND(C)}\) is the equivalence class of \(x\) with respect to \(C\), \([x_i]_{IND(C)}\in U/IND(C)\). \(f([x_i]_{IND(C)})\) is the value set of \([x_i]_{IND(C)}\) on \(D\). \(|Y|\) is the number of element contained by \(Y\), the literature [4] drew a conclusion that \(mij\) belongs to cores when it is a single attribute. Essentially, this method come down to the fact that elements in discernibility matrix must be capable of discerning object of positive region from other objects, therefore this discernibility matrix may be considered as discernibility matrix based on changeless positive region.

III. THE PRINCIPLE OF DISTRIBUTION REDUCTION

In order to describe new discernibility matrix based on distribution reduction, In this section we describe the basic principle of distribution reduction and their main properties.

Definition 4 an information system \(IS = \langle U, A = C \cup D, V, f \rangle\), \(U = \{x_1, x_2, \ldots, x_n\}\), \(U/IND(D) = \{Y_1, Y_2, \ldots, Y_m\}\), \(B \subseteq C\), \(x \in U\), \(\mu_b(x) = \{\mu_1 \cap [x_1], \mu_2 \cap [x_2], \ldots, [x_n]\}\) is distribution function of \(x\) with regard to \(B\). Apparently, \(\mu_b(x)\) denotes probability distribution of elements of \([x_i]\) among \(U/IND(D)\). If \((\forall x \in U)(\mu_b(x) = \mu_a(x))\), then \(B\) is called as distribution consistent set. Furthermore, if \(B\) is distribution consistent set, and not exists \(B' \subset B\) is distribution consistent set, then \(B\) is called as distribution reduction.

Definition 5 an information system \(IS = \langle U, A = C \cup D, V, f \rangle\), \(U = \{x_1, x_2, \ldots, x_n\}\), \(a \in C\), the distribution core of \(IS\) is denoted as \(CORE_a(C)\), \(a \in CORE_a(C)\), if and only if \((\exists x \in U)(\mu_b(x) = \mu_a(x))\).

\(CORE_a(C)\) are indispensable components which retain stabilization of subjection degree of all objects with regard to every decision class. If \(a \in CORE_a(C)\) is deleted, then there must be change of subjection degree of some objects with regard to some decision class.

IV. DISCERNIBILITY MATRIX BASED ON DISTRIBUTION REDUCTION

The discernibility matrix that Skowron and Hu proposed, which use object as basic comparison unit, has some disadvantages such as excessive comparison times, much storage space, and only keeps positive region unchanged. This section designs an algorithm of discernibility matrix based on distribution, this algorithm considers whether the objects compared are consistent, analyses in detail the degree of inconsistency and the distributing proportion of their conditional equivalent classes in decision classes, and the reduction in distribution reduction view is acquired finally. The discernibility matrix is improved beforehand, the rows and columns are reduced to the equivalent classes induced by \(A = C \cup D\), and the rule of no converse comparison is set up.

Definition 6 Given an information system \(IS = \langle U, A = C \cup D, V, f \rangle\), \(U = \{x_1, x_2, \ldots, x_n\}\), \(a \in C\), \(U/IND(C) = \{x_1, x_2, \ldots, x_n\}\), \(U/IND(D) = \{Y_1, Y_2, \ldots, Y_m\}\), \(U/IND(C) \cup D = \{Z_1, Z_2, \ldots, Z_k\}\), the discernibility matrix of \(IS\) denoted as \(M_{IS} = (m_{ij})\) that is a \(K \times K\) matrix, the element \(m_{ij}\) of matrix is defined as below:

\[
m_{ij} = \begin{cases}
\frac{\mu_a \cap \nu_b \cap \nu_c}{\mu_a \cap \nu_b \cap \nu_c} & \text{otherwise}
\end{cases}
\]

This discernibility matrix designs reduction method based on discernibility matrix by using distribution reduction as theory foundation, 3 kinds of condition would be considered in detail:

1. \(Z_i, Z_j\) both belong to positive region

When \(Z_i, Z_j\) both belong to positive region, if they belong to same decision equivalent class (\(C_i\)), then \(\mu_{C_i}(Z_i) = \mu_{C_i}(Z_j) = (0, \ldots, 0)\), \(\mu_{D_i}(Z_i) = (0, \ldots, 0, 1, \ldots, 0)\), in subsystem \(Z_i U Z_j\), no matter what condition attribute is deleted, the distribution function has only one item \([x_i] \cap [x_j] = 1\) before and after deleted. If they belong to different decision equivalent class, the distribution function of \(Z_i\) and \(Z_j\) would keep unchanged, as long as some condition attribute with different value exists. So when the objects in positive region are discerned, the rule based on distribution reduction is the same with the rule based on positive region, i.e. (i) \(f(Z_i, D) = f(Z_j, D)\) \(\wedge\) \(f([Z_i]_{IND(D)}, D] = 1)\), (ii) \(f(Z_i, D) = f(Z_j, D)\) \(\wedge\) \(f([Z_i]_{IND(C)}, D] = 1)\), the discernibility matrix are:

\[
\{a \in C \wedge f(Z_i, a) = f(Z_j, a)\}
\]

2. \(Z_i, Z_j\) one belongs to positive region and the other doesn’t belong to positive

Let us supposed that \(Z_i\) belongs to positive region and \(Z_j\) doesn’t belong to positive, then \(f(Z_i, D) = f(Z_j, D)\) \(\wedge\) \(f([Z_i]_{IND(C)}, D] = 1)\). If \(f(Z_i, D) = f(Z_j, D)\), then there must be \(Z_i\) belongs to the same equivalent class with \(Z_i\), and \(f(Z_i, D) = f(Z_j, D)\), so as to \(m_{ij} = m_{ij}\), therefore we may be leave out the circumstance of \(f(Z_i, D) = f(Z_j, D)\). Under the
circumstances, the discernibility set still are \([a | a \in C \land f(Z, a) \neq f(Z', a)]\).

(3) \(Z_i, Z_j\) both doesn’t belong to positive region

If \(Z_i, Z_j\) both doesn’t belong to positive region, then
\(|f([x_{\text{INDC}}, D]) > 1 \land (f([x_{\text{INDC}}, D]) > 1)|

When we extract the element \(m_{ij}\) of matrix, if \(\mu_c(Z_i) = \mu_c(Z_j)\), in subsystem \(Z_i \cup Z_j\), no matter what condition attribute is deleted, the distribution function of \(Z_i\) and \(Z_j\) are equal, they are not discernable. If \(\mu_c(Z_i) \neq \mu_c(Z_j)\), in subsystem \(Z_i \cup Z_j\), the distribution function of \(Z_i\) and \(Z_j\) would remain unchanged and discernable as long as some condition attribute with different value exist. We can generalize this circumstances as the fact that when
\(\min\{f([x_{\text{INDC}}, D]) \land (f([x_{\text{INDC}}, D]) > 1\}

(3) may be summarized as the fact that when
\(\min\{f([x_{\text{INDC}}, D]) \land (f([x_{\text{INDC}}, D]) > 1 \land \mu_c(Z_i) \neq \mu_c(Z_j)\). Otherwise, if
\(\min\{f([x_{\text{INDC}}, D]) \land (f([x_{\text{INDC}}, D]) > 1 \land \mu_c(Z_i) = \mu_c(Z_j)

Definition 7 Suppose MIS=(mij) is the discernibility matrix of decision system IS=<U, A, V, D>, the discernibility function of MIS is denoted as the conjunct of every \(\land\) where \(\land\) is the disjunction of each attribute contained in mij. The discernibility function can be transformed to an alternative representation which is called disjunctive normal form, then the most simplest form of the disjunctive normal form is corresponding to all reductions of information system.

Definition 8 Let IS information system and its discernibility matrix MIS, its discernibility set is defined as the set which is made of all non-empty elements from discernibility matrix. The discernibility set is denoted by \(\mathcal{D}\).

Discernibility set forms by extracting and optimizing all elements in discernibility matrix, any two elements in discernibility set have not inclusive relation. Discernibility function can be constructed with discernibility set efficiently.

V. ANALYSIS OF INSTANCE

This section describes the reduction acquiring method in detail. Given an information system in Table 1, where \(C=\{a,b,c,d,e,f\}\), (numx) indicates that equivalent class includes num objects.

<table>
<thead>
<tr>
<th>TABLE II.</th>
<th>THE IMPROVED DISCERNIBILITY MATRIX BASED ON CHANGELESS POSITIVE REGION</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Z_1)</td>
<td>[a,b,d]</td>
</tr>
<tr>
<td>(Z_2)</td>
<td>[a,d,e]</td>
</tr>
<tr>
<td>(Z_3)</td>
<td>[a,f]</td>
</tr>
<tr>
<td>(Z_4)</td>
<td>[f]</td>
</tr>
<tr>
<td>(Z_5)</td>
<td>[a,b,d,f]</td>
</tr>
<tr>
<td>(Z_6)</td>
<td>[d,e,f]</td>
</tr>
<tr>
<td>(Z_7)</td>
<td>[a,e]</td>
</tr>
<tr>
<td>(Z_8)</td>
<td>[a,f]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE III.</th>
<th>THE IMPROVED DISCERNIBILITY MATRIX BASED ON DISTRIBUTION REDUCTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Z_1)</td>
<td>[a,b,d]</td>
</tr>
<tr>
<td>(Z_2)</td>
<td>[a,d,e]</td>
</tr>
<tr>
<td>(Z_3)</td>
<td>[a,f]</td>
</tr>
<tr>
<td>(Z_4)</td>
<td>[f]</td>
</tr>
<tr>
<td>(Z_5)</td>
<td>[a,b,d,f]</td>
</tr>
<tr>
<td>(Z_6)</td>
<td>[d,e,f]</td>
</tr>
<tr>
<td>(Z_7)</td>
<td>[a,e]</td>
</tr>
<tr>
<td>(Z_8)</td>
<td>[a,f]</td>
</tr>
</tbody>
</table>

In IS, the equivalence classes of positive region are \(Z_1, Z_2, Z_3\) and \(Z_4\); the condition equivalence classes are \(Z_1, Z_2, Z_3, Z_4, Z_5, Z_6, Z_7, Z_8\). Then table 2 is the improved discernibility matrix based on changeless positive region, table 3 is the improved discernibility matrix based on distribution reduction.
Based on Table 2, we can construct discernibility set $\psi = \{\{e\}, \{f\}, \{a,b,d\}\}$, the discernibility function $DM(IS) = e \land f \land (a \lor b \lor d) = (a \land e \land f) \lor (b \land e \land f) \lor (d \land e \land f)$, so the reductions based on positive region are $aef$, $bef$ and $def$. Based on Table 3, we can construct discernibility set $\psi = \{\{e\}, \{f\}, \{a\}\}$, the discernibility function $DM(IS) = a \land e \land f$, so the reductions based on distribution reduction are $aef$.

VI. CONCLUSION

In inconsistent information systems, the reduction in algebra view are not equivalent to that in information view. The conclusion is expounded that distribution reduction is equivalent to the reduction in information view, the problem of discernibility matrix in algebra view is analyzed, and an algorithm of discernibility matrix based on distribution is proposed in this paper. This algorithm considers whether the objects compared are consistent, analyses in detail the degree of inconsistency and the distributing proportion of their conditional equivalent classes in decision classes, and the reduction in distribution reduction view is acquired finally. The theoretic analysis and simulation instance shows that this algorithm is feasible and effective in practice.

REFERENCES