Reductos

Los **reductos**, desde el punto de vista de la teoría de conjuntos rugosos, son subconjuntos de atributos que conservan la capacidad de discernibilidad del conjunto total de atributos [1].

Introduciremos algunas definiciones de la Teoría de Conjuntos Rugusos (TCR) importantes para la definición de reducto.

Definition 1 Sistemas de Información

Los ***sistemas de información*** *(SI) son estructuras que permiten una representación básica en el marco de la TCR. Consisten en una tabla cuyas filas corresponden a objetos mientras que las columnas son variables o atributos que caracterízan a los objetos. Un SI es formalmente definido como un par de $SI = (U, A)$ donde U es un conjunto no vacío de objetos $U = \{O_1, O_2, ..., O_m\}$ y $A = \{x_1, x_2, ..., x_n\}$ es un conjunto no vacío de atributos. Para cada atributo x_i de A existe un conjunto de valores V_{x_i}. A cada objeto O_j de U se le hace corresponder una tupla del conjunto $V_{x_1} \times V_{x_2} \times ... \times V_{x_n}$.*

Definition 2 Relación de indiscernibilidad

Una relación de indiscernibilidad para un subconjunto de atributos $B \subseteq A$ se define como:

$$IND\,(B) = \{(O_j, O_k) \in U \times U | x_i\,(O_j) = x_i\,(O_k)\, \forall x_i \in B\}$$

donde $x_i\,(O_j) \in V_{x_i}$ y es el valor que toma el objeto O_j en el atributo x_i. Esto significa que si dos objetos pertenecen a $IND\,(B)$, entonces no son diferentes en ninguno de los atributos de B y no es posible diferenciarles o identificarles.

La relación de indiscernibilidad es una relación de equivalencia y por tanto induce una partición $U/IND\,(A)$ sobre U.

Los atributos de A se pueden dividir en dos conjuntos: el conjunto C de los **atributos de condición** y el conjunto D de **atributos de decisión**. Se cumple que $A = C \cup D$ y $C \cap D = \emptyset$.

Definition 3 Sistema de Decisión

Un SI con $D \neq \emptyset$ es llamado un ***sistema de decisión*** *(SD). En muchos sistemas de decisión D está compuesto por un único atributo que denotaremos por d y a su conjunto de valores V_d. Cada conjunto $L \in U/IND\,(\{d\})$ se denomina* ***clase*** *de SD.*

Example 4 *Ejemplo de Sistema de Decisión*

	x_1	x_2	x_3	x_4	d
O_1	*Normal*	X	0	*Líquido*	*Positivo*
O_2	*Normal*	X	1	*Líquido*	*Positivo*
O_3	*Escaso*	Y	0	*Líquido*	*Negativo*
O_4	*Escaso*	Y	0	*Gaseoso*	*Negativo*
O_5	*Normal*	Z	0	*Gaseoso*	*Negativo*
O_6	*Abundante*	Y	1	*Líquido*	*Positivo*
O_7	*Normal*	X	1	*Líquido*	*Negativo*
O_8	*Abundante*	Y	1	*Líquido*	*Positivo*
O_9	*Normal*	X	0	*Líquido*	*Negativo*

En este ejemplo existen dos clases: $\{O_1, O_2, O_6, O_8\}$ y $\{O_3, O_4, O_5, O_7, O_9\}$.

Podemos apreciar que los sistemas de decisión de la TCR corresponden a las matrices de aprendizaje en la teoría de testores.

Definition 5 Relación de indiscernibilidad relativa

Una ***relación de indiscernibilidad relativa*** *para un subconjunto de atributos $A_t \subseteq C$ se define como:*

$$IND\,(A_t|\{d\}) = IND\,(A_t) \cup IND\,(\{d\})$$

La relación de indiscernibilidad relativa para A_t es un conjunto formado por la unión de todos los pares de objetos que pertenecen a una misma clase, con todos los pares de objetos de clases diferentes que toman valores iguales en todos los atributos de A_t. Esto significa que no son distinguibles si se considera la parte de su descripción que contine solo los atributos de A_t.

En muchos problemas de clasificación que parten de un SD nos interesa diferenciar claramente a los objetos que pertenecen a clases diferentes. Es en este contexto que adquiere relevancia el concepto de reducto.

Definition 6 Reducto *El conjunto $A_t \subseteq C$ es un* **reducto** *para de un sistema de decisión DS si:*

1. $IND\,(A_t|D) = IND\,(C|D)$
2. $\forall x_i \in A_t, IND\,(A_t - \{x_i\}\,|D) \neq IND\,(C|D)$

De acuerdo a la definición, un reducto es un subconjunto de A que es capaz de mantener la capacidad de diferenciación que posee A y en el que todos los atributos son indispensables para mantener esa capacidad. Es decir, la eliminación de un rasgo de un reducto altera la relación de discernibilidad para el conjunto resultante. Si un conjunto cumple la condición 1 pero no la 2, se denomina **super-reducto**.

Los pares de objetos que tienen el mismo valor para el atributo de decisión d forman parte de cualquier relación de indiscernibilidad de DS. Para un conjunto $A_t \subseteq C$, la razón por la cual su indiscernibilidad empeoraría con respecto a la de C tiene que ver exclusivamente con la aparición de un par de objetos con valores diferentes de d que no se diferencian en ninguno de los valores que toman en los atributos de A_t.

Los pares de objetos de clases diferentes pertenecen al complemento de $IND(\{d\})$ y los representaremos por $\complement IND(\{d\})$. Nos interesa separar aquellos pares de objetos que tienen descripciones iguales pero pertenecen a clases diferentes. Ese conjunto de pares de objetos se describe en la siguiente definición:

Definition 7 Relación de indiscernibilidad dura

*Una **relación de indiscernibilidad dura** para un subconjunto de atributos $A_t \subseteq C$ se define como*

$$INDD(A_t|\{d\}) = IND(A_t) \cap \complement IND(\{d\})$$

Con el objetivo de resumir la información necesaria para el cálculo de reductos de un DS podemos usar una idea similar al concepto de matriz de diferencias asociado a la teoría de testores. Denominaremos a esta matriz **Matriz de Discernibilidad (MD)**. Cada fila de MD corresponde a un par de objetos de clases diferentes y cada columna está asociada a un atributo. En cada entrada puede haber un 0 o un 1. El valor de la entrada es 0 si el par de objetos de la fila correspondiente a dicha entrada tiene valores iguales en el atributo de la columna correspondiente y un 1 en el caso contrario.

Example 8 *Ejemplo de MD para SD del Ejemplo 1*

	x_1	x_2	x_3	x_4
O_1O_3	1	1	0	0
O_1O_4	1	1	0	1
O_1O_5	0	1	0	1
O_1O_7	0	0	1	0
O_1O_9	0	0	0	0
O_2O_3	1	1	1	0
O_2O_4	1	1	1	1
O_2O_5	0	1	1	1
O_2O_7	0	0	0	0
O_2O_9	0	0	1	0
O_6O_3	1	0	1	0
O_6O_4	1	0	1	1
O_6O_5	1	1	1	1
O_6O_7	1	1	0	0
O_3O_9	1	1	0	0
O_8O_3	1	0	1	0
O_8O_4	1	0	1	1
O_8O_5	1	1	1	1
O_8O_7	1	1	0	0
O_8O_9	1	1	1	0

En [2] se demuestra que se puede reducir la información contenida en la matriz de diferencias sin afectar el conjunto de todos los testores típicos. La matriz reducida a partir de la matriz de diferencias se denomina**matriz básica**.

Siguiendo el mismo argumento, se puede reducir la matriz de discernibilidad con un proceso parecido, pero hay que tener en cuenta que la en la teoría de testores se supone que no deben existir solapamiento de clases en el conjunto total de variables. Esta idea expresada en función de las definiciones introducidas anteriormente significa que partimos de una SD tal que $INDD(C|\{d\}) = \varnothing$, esto significa que:

$$IND(C|\{d\}) = IND(\{d\})$$

En ese caso el proceso para calcular los reductos es similar al descrito en [mencionar la parte del artículo en dónde se calcularán los testores típicos].

Si $INDD(C) \neq \varnothing$ esto implica que aparecerá al menos una fila de 0s en la matriz de diferencias. En este caso, antes de determinar las filas básicas, se eliminan las filas de 0s. Si obsevamos la matriz del ejemplo 2, se eliminan las filas correspondientes a las comparaciones de los objetos 1 y 9 y 2 y 7 antes de eliminar las filas no básicas (ver ejemplo 3).

Example 9 *MD para SD del Ejemplo 8 con eliminación de pares partenecientes a $INDD(C)$.*

	x_1	x_2	x_3	x_4
O_1O_3	1	1	0	0
O_1O_4	1	1	0	1
O_1O_5	0	1	0	1
O_1O_7	0	0	1	0
O_2O_3	1	1	1	0
O_2O_4	1	1	1	1
O_2O_5	0	1	1	1
O_2O_9	0	0	1	0
O_6O_3	1	0	1	0
O_6O_4	1	0	1	1
O_6O_5	1	1	1	1
O_6O_7	1	1	0	0
O_3O_9	1	1	0	0
O_8O_3	1	0	1	0
O_8O_4	1	0	1	1
O_8O_5	1	1	1	1
O_8O_7	1	1	0	0
O_8O_9	1	1	1	0

La siguiente tabla ilustra un ejemplo de cómo obtener una matriz básica a partir de la tabla anterior.

Example 10 *MB para MD del Ejemplo 9.*

	x_1	x_2	x_3	x_4
O_1O_5	0	1	0	1
O_1O_7	0	0	1	0
O_6O_7	1	1	0	0

La defnción de reducto sobre esta matriz básica coincidiría con la definición de testor típico introducida en (citar sitio del artículo). Los reductos son $\{x_1, x_3, x_4\}$ y $\{x_2, x_3\}$.

Existen otras definiciones de reductos que manejan de manera diferente el solapamiento de clases, es decir, el hecho de que aparezcan dos objetos iguales en clases diferentes comparados en el conjunto total de rasgos. Según esta definición se requiere conservar a aquellos objetos que en el conjunto total de rasgos no se han confundido con ningún otro objeto de otra clase. En cambio, no es importante si un objeto que ya tiene uno igual en otra clase se confunde con otro más de esa misma clase. Revisemos una nueva definición de reducto que tiene en cuenta estos criterios. Pare ello introducimos las siguientes definiciones.

Definition 11 Región positiva
Dada una $SD = (U, A)$, y $A_t \subseteq C$, llamaremos región positiva de U con respecto a A_t al conjunto de todos los objetos decada clase que no se confunden con ningún objeto de otra clase.

$$POS\,(A_t) = \{O_i \in U | \forall O_j \in U, i \neq j, (O_i, Oj) \notin INDD\,(A_t)\}$$

Si analizamos el sistema de decisión del ejemplo 4, vemos que $POS\,(C) = U/\{O_1, O_2, O_7, O_9\}$.

Definition 12 Reducto duro *El conjunto $A_t \subseteq C$ es un* ***reducto duro*** *para de un sistema de decisión DS si:*

1. $POS\,(A_t) = POS\,(C)$
2. $\forall x_i \in A_t, POS\,(A_t - \{x_i\}) \neq POS\,(C)$.

Según la definición anterior un reducto duro es un subconjunto mínimo de atributos que conserva la región positiva del conjunto total de atributos.

Definition 13 Par irrelevante
Sea (O_j, O_k) un par de objetos de U tal que $d\,(O_j) \neq d\,(O_k)$. (O_j, O_k) se denomina ***par irrelevante*** *si:*

1. $(O_j, O_k) \notin INDD\,(C|\{d\})$.
2. $O_j \notin POS\,(C) \wedge O_k \notin POS\,(C)$.

Example 14 *Pares de objetos irrelevantes(en negritas) partiendo de la matriz del Ejemplo 9.*

	x_1	x_2	x_3	x_4
O_1O_3	1	1	0	0
O_1O_4	1	1	0	1
O_1O_5	0	1	0	1
$\mathbf{O_1O_7}$	**0**	**0**	**1**	**0**
O_2O_3	1	1	1	0
O_2O_4	1	1	1	1
O_2O_5	0	1	1	1
$\mathbf{O_2O_9}$	**0**	**0**	**1**	**0**
O_6O_3	1	0	1	0
O_6O_4	1	0	1	1
O_6O_5	1	1	1	1
O_6O_7	1	1	0	0
O_3O_9	1	1	0	0
O_8O_3	1	0	1	0
O_8O_4	1	0	1	1
O_8O_5	1	1	1	1
O_8O_7	1	1	0	0
O_8O_9	1	1	1	0

Para determinar los reductos duros a partir de la matriz de discernibilidad es necesario eliminar previamente todos los pares de objetos de $INDD\,(C|\{d\})$ y los pares de objetos irrelevantes como se aprecia en el ejemplo 14.

Example 15 *MD para SD del ejemplo 4 una vez eliminados los pares de $INDD\,(C|\{d\})$ y los pares irrelevantes.*

	x_1	x_2	x_3	x_4
O_1O_3	1	1	0	0
O_1O_4	1	1	0	1
O_1O_5	0	1	0	1
O_2O_3	1	1	1	0
O_2O_4	1	1	1	1
O_2O_5	0	1	1	1
O_6O_3	1	0	1	0
O_6O_4	1	0	1	1
O_6O_5	1	1	1	1
O_6O_7	1	1	0	0
O_3O_9	1	1	0	0
O_8O_3	1	0	1	0
O_8O_4	1	0	1	1
O_8O_5	1	1	1	1
O_8O_7	1	1	0	0
O_8O_9	1	1	1	0

Una vez depurada la matriz se procede al cálculo de la matriz básica con el mismo criterio usado en [citar la parte del artículo donde se ahace esto para el cálculo de testores típicos].

Example 16 *MB para determinar reductos duros.*

	x_1	x_2	x_3	x_4
O_1O_3	1	1	0	0
O_1O_5	0	1	0	1
O_6O_3	1	0	1	0

Los reductos duros son $\{x_1, x_2\}$, $\{x_1, x_4\}$ *y* $\{x_2, x_3\}$.

Referencias

[1] Pawlak, Z. (2012). Rough sets: Theoretical aspects of reasoning about data (Vol. 9). Springer Science & Business Media.

[2] Piza-Davila, I., Sanchez-Diaz, G., Lazo-Cortes, M. S., Rizo-Dominguez, L. (2017). A CUDA-based hill-climbing algorithm to find irreducible testors from a training matrix. Pattern Recognition Letters, 95, 22-28.