

Rough Sets – Técnica de Redução de Atributos e Geração de Regras para Classificação de Dados

Cristian Mara M. M. Patrício **João Onofre Pereira Pinto**

Departamento de Engenharia Elétrica, DEE, UFMS,
Cidade Universitária s/n,
79070-900, Campo Grande, MS
E-mail: cristian_mara@yahoo.com.br, jpintp@del.ufms.br

Celso Correia de Souza

Universidade para o Desenvolvimento do Estado e da Região do Pantanal - UNIDERP,
Rua Cerará, 333
79003-010, Campo Grande, MS
E-mail: celsocorreia@mail.uniderp.br

1 Introdução

A quantidade de informações disponíveis cresceu exponencialmente nos últimos anos com o advento da evolução tecnológica, o que levou a problemas opostos de 20 anos atrás. Hoje, a quantidade de dados é tão grande que se torna mandatório a elaboração de modelos que auxiliem na tomada de decisões gerenciais.

Normalmente, num banco de dados, tem-se um número muito grande de atributos armazenados dos quais partes são relevantes e parte são irrelevantes, desnecessários, para a tomada de uma específica decisão.

Por exemplo, uma decisão gerencial precisa ser tomada em uma empresa, com base nas informações existentes em seu banco de dados. Dos dados ali armazenados, nem todos são necessários para a tomada da decisão, pois não alterarão no resultado. Estas informações que não modificam os resultados com relação à decisão, são denominadas como atributos irrelevantes. Então, pode-se dizer que para obter um modelo que auxilie em uma decisão gerencial específica, apenas uma quantidade reduzida de atributos são fundamentais na avaliação, ou seja, necessários para se chegar a uma decisão. Deve-se portanto, separar os atributos fundamentais dos irrelevantes de forma a reduzir a quantidade de atributos a serem utilizados na tomada da decisão e, conseqüentemente, no modelo a ser definido neste processo. A Figura 1 representa esta redução.

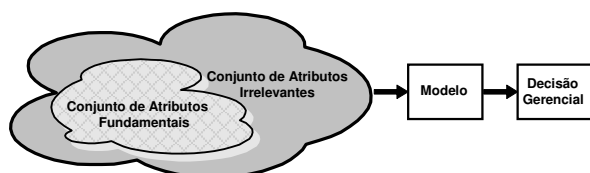


Figura 1 – Representação simplificada do processo de modelagem.

A teoria de Rough Sets, que foi introduzida primeiramente por Zdzislaw Pawlak em 1982 [1], possui propriedades que permitem eliminar variáveis

ou atributos irrelevantes através do processo de redução do sistema de informação, baseando-se na definição de redutos, os quais são subconjuntos de atributos capazes de manter as mesmas propriedades da representação de conhecimento quando esta é feita utilizando todos os atributos. Este procedimento de eliminação de atributos irrelevantes é uma das características dessa teoria.

Rough Sets também proporciona grande habilidade na classificação de objetos. Os objetos contidos em um sistema, de acordo com suas características, são agrupados em classes. Os objetos agrupados em uma mesma classe são indiscerníveis entre si. Como na maioria dos bancos de dados há informações imprecisas, a teoria de Rough Sets é capaz de administrar estas imprecisões, informações ruidosas e incompletas presentes nestes sistemas. Assim, objetos que não podem ser especificados através dos dados disponíveis são classificados por esta teoria através de dois conceitos: as aproximações inferior e superior, que serão abordados mais adiante.

A fundamentação matemática desta teoria permite a descoberta de padrões ocultos na base de dados. Sua utilidade no campo de mineração de dados, em processamento de sinais de som e imagem, clusterização, entre outros, podem ser comprovadas pelo crescente número de aplicações divulgadas com estes conteúdos [1, 2].

Neste trabalho, o assunto será abordado da seguinte forma: após uma breve introdução sobre a teoria de Rough Sets, seus conceitos e descrições serão apresentados seguidos de exemplos para facilitar o entendimento. Primeiramente será apresentado um sistema de informação seguido da definição de indiscernibilidade, aproximações dos conjuntos, bem como a qualidade destas aproximações. A seção seguinte tratará da definição da redução da informação que será subdividida em três partes: matriz de discernibilidade, função de discernibilidade e reduto. Após a redução da informação tratar-se-á da geração de regras e por fim, serão apresentadas as considerações finais sobre a utilização de Rough Sets em classificação de banco de dados.

2 Sistema de Informação

A forma mais comum para representação dos dados na abordagem de Rough Sets é através de um sistema de informação que contém um conjunto de objetos, sendo que cada objeto tem uma quantidade de atributos. Esses atributos são os mesmos para cada um dos objetos, mas seus valores nominais podem diferir. Portanto, um sistema de informação pode ser representado pela Tabela 1 e indicado por:

$$A = (U, C) \quad (3.1)$$

sendo U o conjunto de objetos e C os atributos disponíveis na base de dados.

U	Atributos Condicionais				
Criança	Cor	Tamanho	Tato	Textura	Material
1	Azul	Grande	Duro	Indefinido	Plástico
2	Vermelho	Médio	Moderado	Liso	Madeira
3	Amarelo	Pequeno	Macio	Áspero	Pelúcia
4	Azul	Médio	Moderado	Áspero	Plástico
5	Amarelo	Pequeno	Macio	Indefinido	Plástico
6	Verde	Grande	Duro	Liso	Madeira
7	Amarelo	Pequeno	Duro	Indefinido	Metal
8	Amarelo	Pequeno	Duro	Indefinido	Plástico
9	Verde	Grande	Duro	Liso	Madeira
10	Verde	Médio	Moderado	Liso	Plástico

Tabela 1: Sistema de informação

Em muitos casos é importante a classificação dos objetos considerando um atributo de decisão (D) que informa a decisão a ser tomada. Pode-se assumir, portanto, que um sistema de informação munido de um atributo de decisão é denominado sistema de informação, denotado por \mathfrak{R} , sendo

$$\mathfrak{R} = (U, C \cup D) \quad (3.2)$$

Na Tabela 2 tem-se um exemplo de classificação de dados de um sistema de informação incluindo um atributo de decisão. Os dados são referentes a crianças e algumas características de brinquedos e deseja-se um sistema de informação capaz de fornecer a atitude das crianças com relação aos brinquedos [3].

U	Atributos Condicionais					Atributo de Decisão
Criança	Cor	Tamanho	Tato	Textura	Material	Atitude
1	Azul	Grande	Duro	Indefinido	Plástico	Negativa
2	Vermelho	Médio	Moderado	Liso	Madeira	Neutra
3	Amarelo	Pequeno	Macio	Áspero	Pelúcia	Positiva
4	Azul	Médio	Moderado	Áspero	Plástico	Negativa
5	Amarelo	Pequeno	Macio	Indefinido	Plástico	Neutra
6	Verde	Grande	Duro	Liso	Madeira	Positiva
7	Amarelo	Pequeno	Duro	Indefinido	Metal	Positiva
8	Amarelo	Pequeno	Duro	Indefinido	Plástico	Positiva
9	Verde	Grande	Duro	Liso	Madeira	Neutra
10	Verde	Médio	Moderado	Liso	Plástico	Neutra

Tabela 2 – Sistema de informação incluindo o atributo de decisão.

Para o Sistema de informação descrito na Tabela 2, temos:

- conjunto de objetos ou registros: $U = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$
- conjunto de atributos condicionais: $C = \{\text{Cor}, \text{Tamanho}, \text{Tato}, \text{Textura}, \text{Material}\}$
- atributo de decisão: $D = \{\text{Atitude}\}$

Assim, como pode haver mais de uma característica

ou valor nominal para cada atributo condicional, pode haver mais de uma opção de característica para o atributo de decisão.

Os valores nominais dos atributos considerados sistema de informação da Tabela 2, estão representados na Tabela 3.

	Atributos	Valores Nominais.
Atributos Condicionais	Cor	Azul, Vermelho, Amarelo, Verde.
	Tamanho	Grande, Médio, Pequeno.
	Tato	Duro, Moderado, Macio.
	Textura	Liso, Áspero, Indefinido.
	Material	Plástico, Madeira, Pelúcia, Metal.
Atributos de Decisão	Atitude	Neutra, Negativa, Positiva.

Tabela 3: Valores nominais dos atributos.

3 Indiscernibilidade

Considerando os Atributos Condicionais (C), para todo subconjunto de atributos $B \subseteq C$ do sistema de informação A (3.1), uma relação de equivalência $IND_A(B)$ é associada, chamada relação de indiscernibilidade, definida como:

$$IND_A(B) = \{(x, y) \in U^2 \mid \forall c \in B, c(x) = c(y)\} \quad (3.3)$$

Assim, x e y são indiscerníveis entre si para todo atributo de B .

Se a relação de indiscernibilidade existe entre dois objetos, significa que todos os valores de seus atributos são idênticos com respeito ao subconjunto de atributos B considerado, ou seja, não podem ser diferenciados entre si.

O conjunto de todas as classes de equivalência determinadas por $IND_A(B)$ é representado por $U/IND_A(B)$, denominado conjunto quociente de U pela relação $IND_A(B)$.

Para a base de dados da Tabela 1, alguns dos possíveis subconjuntos dos Atributos Condicionais são: $\{\text{Cor}\}$, $\{\text{Tamanho}\}$, $\{\text{Tato}\}$, $\{\text{Textura}\}$, $\{\text{Material}\}$, $\{\text{Cor}, \text{Tamanho}\}$, $\{\text{Cor}, \text{Tato}\}$, $\{\text{Cor}, \text{Textura}\}$, $\{\text{Cor}, \text{Material}\}$, $\{\text{Tamanho}, \text{Tato}\}$, $\{\text{Tamanho}, \text{Textura}\}$, $\{\text{Tamanho}, \text{Material}\}$, $\{\text{Tato}, \text{Textura}\}$, $\{\text{Tato}, \text{Material}\}$, $\{\text{Textura}, \text{Material}\}$, entre outras possíveis combinações, num total de 31 (trinta e um) subconjuntos não vazios. Considerando o subconjunto $B = \{\text{Tamanho}\}$, os objetos 1, 6 e 9 estão na mesma classe de equivalência e são indiscerníveis, assim como os objetos 2, 4 e 10, e os objetos 3, 5, 7 e 8. Assim, $U/IND_A(B)$ para alguns dos possíveis subconjuntos $B \subseteq C$ são:

$$U/IND_A(\{\text{Tamanho}\}) = \{\{1, 6, 9\}, \{2, 4, 10\}, \{3, 5, 7, 8\}\};$$

$$U/IND_A(\{\text{Material}\}) = \{\{1, 4, 5, 8, 10\}, \{2, 6, 9\}, \{3\}, \{7\}\};$$

$$U/IND_A(\{\text{Cor}, \text{Tamanho}\}) = \{\{1\}, \{2\}, \{3, 5, 7, 8\}, \{4\}, \{6, 9\}, \{10\}\}.$$

Para cada subconjunto de $B \subseteq C$ os objetos são agrupados e os grupos consistem de objetos que são

indiscerníveis entre si. De acordo com a teoria de Rough Sets, cada um desses grupos é uma classe de equivalência. Por exemplo, as classes de equivalências para o subconjunto {Tamanho} estão representadas na Tabela 4. A classe CI1 originou-se dos objetos 1, 6 e 9, a classe CI2 dos objetos 2, 4 e 10 e a classe CI3 dos objetos 3, 5, 7 e 8.

É importante relembrar que não se analisa o Atributo de Decisão para estes agrupamentos. Os objetos de uma mesma classe podem possuir diferentes valores para o Atributo de Decisão. Por exemplo, na classe CII o objeto 1 tem como Atributo de Decisão a Atitude igual à Negativa, assim como para o objeto 6 a Atitude é Positiva e para o objeto 9 a Atitude é Neutra.

	Atributo Condicional
Classes	Tamanho
C11	Grande
C12	Médio
C13	Pequeno

Tabela 4 – Classes para $B = \{\text{Tamanho}\}$

4 Aproximação de Conjuntos

A Tabela 5 apresenta o agrupamento dos objetos do sistema de informação (Tabela 2), conforme as características do atributo de decisão.

Inicialmente, seja a seguinte questão: quais as características dos atributos condicionais que definem as atitudes das crianças com relação aos brinquedos como sendo Negativa, Neutra ou Positiva? Pode-se notar que não há uma resposta única para esta pergunta, pois as crianças 6 e 9 apresentam as mesmas características nos Atributos Condicionais, mas diferenciam-se no Atributo de Decisão.

Pode-se dizer com certeza, conforme Tabela 5, que qualquer criança com características iguais as das crianças 1 ou 4 terão atitude Negativa, assim como qualquer criança com características iguais as crianças 2, 5 ou 10 terão atitudes Neutra e qualquer criança com características iguais as das crianças 3, 7 ou 8 terão atitudes Positiva. Nada se pode afirmar para crianças com características iguais as das crianças 6 e 9. São nesses casos que a noção de Rough Sets emerge.

U	Atributos Condicionais					Atributo Decisão
Criança	Cor	Tamanho	Tato	Textura	Material	Atitude
1	Azul	Grande	Duro	Indefinido	Plástico	Negativa
4	Azul	Médio	Moderado	Áspero	Plástico	Negativa
2	Vermelho	Médio	Moderado	Liso	Madeira	Neutra
5	Amarelo	Pequeno	Macio	Indefinido	Plástico	Neutra
9	Verde	Grande	Duro	Liso	Madeira	Neutra
10	Verde	Médio	Moderado	Liso	Plástico	Neutra
3	Amarelo	Pequeno	Macio	Áspero	Pelúcia	Positiva
6	Verde	Grande	Duro	Liso	Madeira	Positiva
7	Amarelo	Pequeno	Duro	Indefinido	Metal	Positiva
8	Amarelo	Pequeno	Duro	Indefinido	Plástico	Positiva

Tabela 5 – Atributo de decisão agrupado por seus valores nominais.

Seja $A = (U, C)$ o sistema de informação, $B \subseteq C$ e $X \subseteq U$, onde X é o conjunto de objetos ou registros com respeito a B , isto é, X é obtido através das

informações dos atributos de B . Assim, defini-se Aproximação Inferior de X em relação a B , denotado por $\underline{B}(X)$ e Aproximação Superior de X em relação a B , denotado por $\overline{B}(X)$, como:

$$B(X) = \{x \in U \mid U / IND_A(B) \subseteq X\} \quad (3.4)$$

$${}^e \bar{B}(X) = \{x \in U \mid U / IND_A(B) \cap X \neq \emptyset\} \quad (3.5)$$

Os objetos ou registros da Aproximação Inferior $\underline{B}(X)$ são classificados com certeza como membros de X , utilizando o conjunto de atributos B , enquanto que os objetos da Aproximação Superior $\overline{B}(X)$ podem ser classificados como possíveis membros de X , utilizando o mesmo conjunto B . Portanto, obtém-se uma Região de Fronteira de X pela diferença de $\overline{B}(X)$ para $\underline{B}(X)$, denotado por $RF(X)$, ou seja,

$$RF(X) = \bar{B}(X) - B(X) \quad (3.6)$$

que consiste de objetos impossíveis de serem classificados em X . Ainda é possível definir como Fora da Região de X o conjunto $U - \overline{B}(X)$, ou seja, consiste de objetos que não pertencem a X , considerando o mesmo conjunto B .

Um conjunto X é definido como *rough* (impreciso) se sua Região de Fronteira é diferente do conjunto vazio ($RF(X) \neq \emptyset$), e é definido como *crisp* (preciso) se o conjunto for vazio ($RF(X) = \emptyset$).

Seja $B = \{\text{Cor, Tamanho, Tato, Textura, Material}\}$, $U/IND_A(B) = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6, 9\}, \{7\}, \{8\}, \{10\}\}$ e $X = \{3, 6, 7, 8\}$, o conjunto formado por criâncias que correspondem no atributo de decisão a Atitude = Positiva, tem-se as seguintes aproximações, da Tabela 2:

- Aproximação Inferior: $\underline{B}(X) = \{\{3\}, \{7\}, \{8\}\}$
- Aproximação Superior: $\overline{B}(X) = \{\{3\}, \{6, 9\}, \{7\}, \{8\}\}$
- Região de Fronteira: $RF(X) = \{6, 9\}$
- Fora da Região: $U - \overline{B}(X) = \{\{1\}, \{2\}, \{4\}, \{5\}, \{10\}\}$

A Figura 2 exibe as aproximações definidas pelo exemplo acima em forma de conjuntos.

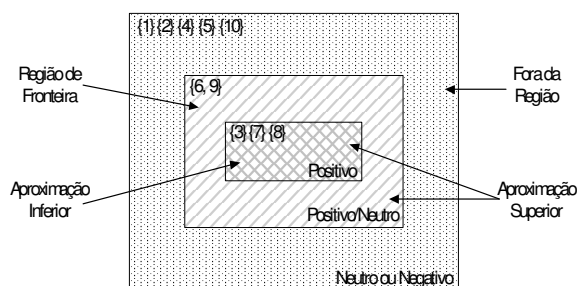


Figura 2 – Aproximações em forma de conjuntos.

5 Qualidade das Aproximações

A qualidade das aproximações obtidas pelas definições dadas previamente pode ser caracterizada

numericamente a partir dos próprios elementos que a definem. O coeficiente para medir essas qualidades é representado por $\alpha_B(X)$, sendo X o conjunto de objetos ou registros com respeito à B , e podem ser realizadas de três formas:

- a) Coeficiente de Imprecisão $\alpha_B(X)$, que pode ser entendido como a qualidade da aproximação de X , dado por:

$$\alpha_B(X) = \frac{|B(X)|}{|\overline{B}(X)|} \quad (3.7)$$

em que $|B(X)|$ e $|\overline{B}(X)|$ denotam a cardinalidade das Aproximações Inferior e Superior, respectivamente e, são conjuntos não-vazios. Obviamente $0 \leq \alpha_B \leq 1$. Se $\alpha_B(X) = 1$, X é *crisp* (preciso) em relação ao conjunto de atributos B . Se $\alpha_B(X) < 1$, X é *rough* (impreciso) em relação ao conjunto de atributos B .

- b) Coeficiente da Qualidade da Aproximação Superior $\alpha_B(\overline{B}(X))$, que pode ser interpretado como sendo o percentual de todos os objetos possivelmente classificados como pertencentes a X , dado por:

$$\alpha_B(\overline{B}(X)) = \frac{|\overline{B}(X)|}{|U|} \quad (3.8)$$

sendo $|U|$ a cardinalidade do conjunto de objetos do Sistema de informação e, $U \neq \emptyset$.

- c) Coeficiente da Qualidade da Aproximação Inferior $\alpha_B(B(X))$, que pode ser interpretado como sendo o percentual de todos os objetos certamente classificados como pertencentes a X , dado por:

$$\alpha_B(B(X)) = \frac{|B(X)|}{|U|} \quad (3.9)$$

Considerando $X = \{3, 6, 7, 8\}$, do exemplo anterior, tem-se:

- $\alpha_B(X) = \frac{|\{3, 7, 8\}|}{|\{3, 6, 7, 8, 9\}|} = \frac{3}{5} = 0.6$, ou seja, 60% de X é preciso, com respeito a B .
- $\alpha_B(\overline{B}(X)) = \frac{|\{3, 6, 7, 8, 9\}|}{|\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}|} = \frac{5}{10} = 0.5$, ou seja, 50% de U possivelmente pertence à X .
- $\alpha_B(B(X)) = \frac{|\{3, 7, 8\}|}{|\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}|} = \frac{3}{10} = 0.3$, ou seja, 30% de U certamente pertence à X .

6 Redução do Sistema de Informação

A disposição dos dados num sistema de informação não deve possuir redundâncias, pois isso aumenta a

complexidade computacional de tal forma que a extração de regras torna-se uma tarefa difícil, desperdiçando tempo e recursos computacionais. Com o intuito de reduzir o sistema de informação, tratando essas redundâncias, é que surge o processo de redução de atributos, sem alterar a relação de indiscernibilidade, denominado redução da informação. Essa redução será realizada pela função de discernibilidade a partir da matriz de discernibilidade.

Matriz de Discernibilidade

Considerando o conjunto de atributos $B = \{\text{Cor, Tamanho, Tato, Textura, Materia}\}$ para o sistema de informação A da Tabela 1, então o conjunto de todos as classes de equivalência determinadas por B sobre A é dado por $U/IND_A(B) = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6, 9\}, \{7\}, \{8\}, \{10\}\}$, que estão representadas Tabela 6.

A matriz de discernibilidade de um sistema de informação A , denotada por $M_D(B)$, é uma matriz simétrica $n \times n$ com:

$$m_D(i, j) = \{b \in B \mid b(Cl(i)) \neq b(Cl(j))\} \quad (3.10)$$

para $i, j = 1, 2, \dots, n$, sendo $1 \leq i, j \leq n$ e $n = |U/IND_A(B)|$. Logo, os elementos da matriz de discernibilidade $m_D(i, j)$ é o conjunto de atributos condicionais de B que diferenciam os objetos das Classes com relação aos seus valores nominais.

Classes	Atributos Condicionais				
	Cor	Tamanho	Tato	Textura	Materia
C11	Azul	Grande	Duro	Indefinido	Plástico
C12	Vermelho	Médio	Moderado	Liso	Madeira
C13	Amarelo	Pequeno	Macio	Áspero	Pelúcia
C14	Azul	Médio	Moderado	Áspero	Plástico
C15	Amarelo	Pequeno	Macio	Indefinido	Plástico
C16	Verde	Grande	Duro	Liso	Madeira
C17	Amarelo	Pequeno	Duro	Indefinido	Metal
C18	Amarelo	Pequeno	Duro	Indefinido	Plástico
C19	Verde	Médio	Moderado	Liso	Plástico

Tabela 6 – Classes de equivalências determinadas por B sobre A .

Considerando $\text{Cor} = \text{Co}$; $\text{Tamanho} = \text{Ta}$; $\text{Tato} = \text{To}$; $\text{Textura} = \text{Te}$; $\text{Material} = \text{Ma}$, com a finalidade da construção da matriz de discernibilidade $M_D(B)$, tem-se na Tabela 7 a sua representação.

Função de Discernibilidade

A função de discernibilidade $F_A(B)$ é uma função booleana com m variáveis que determina o conjunto mínimo de atributos necessários para diferenciar qualquer classe de equivalência das demais, definida como:

$$F_A(b_1^*, b_2^*, \dots, b_m^*) = \bigwedge \{ \bigvee m_D^*(i, j) \mid i, j = 1, 2, \dots, n, m_D(i, j) \neq \emptyset \} \quad (3.11)$$

Sendo $m_D^*(i, j) = \{b^* \mid b \in m_D(i, j)\}$.

	C11	C12	C13	C14	C15	C16	C17	C18	C19
C11	\emptyset								
C12	Co,Ta,To, Te,Ma	\emptyset							
C13	Co,Ta,To, Te,Ma	Co,Ta,To, Te,Ma	\emptyset						
C14	Ta,To,Te	Co,Te,Ma	Co,Ta,To, Ma	\emptyset					
C15	Co,Ta,To	Co,Ta,To, Te,Ma	Te,Ma	Co,Ta,To, Te	\emptyset				
C16	Co,Te,Ma	Co,Ta,To	Co,Ta,To, Te,Ma	Co,Ta,To, Te, Ma	Co,Ta,To, Te, Ma	\emptyset			
C17	Co,Ta,Ma	Co,Ta,To, Te,Ma	To,Te,Ma	Co,Ta,To, Te, Ma	To,Ma	Co,Ta,Te, Ma	\emptyset		
C18	Co,Ta	Co,Ta,To, Te,Ma	To,Te,Ma	Co,Ta,To, Te	To	Co,Ta,Te, Ma	Ma	\emptyset	
C19	Co,Ta,To, Te	Co,Ma	Co,Ta,To, Te,Ma	Co,Te	Co,Ta,To, Te	Ta,To,Ma	Co,Ta,To, Te,Ma	Co,Ta,To, Te	\emptyset

Utilizando o método de simplificação de expressões booleanas na função $F_A(B)$, obtém-se o conjunto de todos os implicantes primos dessa função, que determina todos os redutos de A . A simplificação é um processo de manipulação algébrica das funções lógicas com a finalidade de reduzir o número de variáveis e operações necessárias para a sua realização.

A função de discernibilidade $F_A(B)$ é obtida da seguinte forma: para os atributos contidos dentro de cada célula da matriz de discernibilidade, aplica-se o operador “soma”, “or” ou “ \vee ” e, entre as células dessa matriz, utiliza-se do operador “produto”, “and” ou “ \wedge ”, resultando em uma expressão booleana de “Produto – da - Soma”.

A $F_A(B)$ da Tabela 7 é representada por:

$$F_A(B) = (Co \vee Ta \vee To \vee Te \vee Ma) \wedge (Co \vee Ta \vee To \vee Te \vee Ma) \wedge (Ta \vee To \vee Te) \wedge (Co \vee Ta \vee To) \wedge (Co \vee Te \vee Ma) \wedge (Co \vee Ta \vee Ma) \wedge (Co \vee Ta) \wedge (Co \vee Ta \vee To \vee Te) \wedge (Co \vee Ta \vee To \vee Te \vee Ma) \wedge (Co \vee Te \vee Ma) \wedge (Co \vee Ta \vee To \vee Te \vee Ma) \wedge (Co \vee Ta \vee To) \wedge (Co \vee Ta \vee To \vee Te \vee Ma) \wedge (Co \vee Ta \vee To \vee Te \vee Ma) \wedge (Co \vee Ma) \wedge (Co \vee Ta \vee To \vee Ma) \wedge (Te \vee Ma) \wedge (Co \vee Ta \vee To \vee Te \vee Ma) \wedge (To \vee Te \vee Ma) \wedge (To \vee Te \vee Ma) \wedge (Co \vee Ta \vee To \vee Te \vee Ma) \wedge (Co \vee Ta \vee To \vee Te) \wedge (Co \vee Ta \vee To \vee Te \vee Ma) \wedge (Co \vee Ta \vee To \vee Te \vee Ma) \wedge (Co \vee Ta \vee To \vee Te) \wedge (Co \vee Te) \wedge (Co \vee Ta \vee To \vee Te \vee Ma) \wedge (To \vee Ma) \wedge (To) \wedge (Co \vee Ta \vee To \vee Te) \wedge (Co \vee Ta \vee Te \vee Ma) \wedge (Co \vee Ta \vee Te \vee Ma) \wedge (Ta \vee To \vee Ma) \wedge (Ma) \wedge (Co \vee Ta \vee To \vee Te \vee Ma) \wedge (Co \vee Ta \vee To \vee Te)$$

Simplificando esta expressão, utilizando teoremas, propriedades e postulados da Álgebra Booleana, obtém-se a seguinte expressão minimizada:

$$F_4(B) = ((Co \vee Ta \vee Te) \wedge To \wedge Ma),$$

que ainda pode ser escrita na forma de “Soma – do – Produto”, ou seja,

$$F_4(B) = ((Co \wedge To \wedge Ma) \vee (Ta \wedge To \wedge Te \wedge Ma))$$

A função de discernibilidade determinou o termo

mínimo da função, ou seja, determinou o conjunto mínimo de atributos necessários para discernir as Classes formadas por todos as classes de equivalência da relação $IND_A(B)$.

Reduto

Reduto é um conjunto de atributos mínimos necessários para manter as mesmas propriedades de conhecimento de um sistema de informação quando este é feito utilizando todos os atributos. Ou seja, o reduto é capaz de classificar objetos, ou classes, sem alterar a representação do conhecimento.

Um Reduto de B sobre um sistema de informação A é um conjunto de atributos B^* , em que $B^* \subseteq B$, sendo todos os atributos $c \in (B - B^*)$ dispensáveis. Com isso, $U/IND_A(B) = U/IND_A(B^*)$.

O conjunto formado pelo termo mínimo da função de discernibilidade $F_A(B)$ determina os redutos de B .

Considerando o exemplo de $B = \{\text{Cor, Tamanho, Tato, Textura, Materia}\}$ e a função de discernibilidade $F_A(B)$ para este conjunto de atributos, previamente definida como:

$$F_4(B) = ((Co \wedge To \wedge Ma) \vee (Ta \wedge To \wedge Te \wedge Ma))$$

Assim, o conjunto de redutos desta função é:

$$RED_A(B) = \{\{\text{Cor, Tato, Material}, \\ \{\text{Tamanho, Tato, Textura, Material}\}\}$$

Portanto, a $U/IND_A(B^*)$ para cada reduto é:

$$U/IND_A(\text{Cor}, \text{Tato}, \text{Material}) = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6, 9\}, \{7\}, \{8\}, \{10\}\}$$

$$U/IND_A(\text{Tamanho, Tato, Textura, Material}) = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6, 9\}, \{7\}, \{8\}, \{10\}\}$$

Assim, $U/IND_A(B) = U/IND_A(B^*)$, portanto, os redutos encontrados através do procedimento da redução do sistema de informação são redutos para o sistema de informação da Tabela 1.

Podem existir mais de um reduto para um mesmo

conjunto de atributos.

O sistema de informação original considerado em nosso exemplo da Tabela 1 pode ser representado por um de seus redutos. Considerando o Reduto $B^* = \{\text{Cor, Tato, Material}\}$, temos então a redução do sistema de informação original, representado na Tabela 8.

U	Atributos Condicionais		
Criança	Cor	Tato	Material
1	Azul	Duro	Plástico
2	Vermelho	Moderado	Madeira
3	Amarelo	Macio	Pelúcia
4	Azul	Moderado	Plástico
5	Amarelo	Macio	Plástico
6	Verde	Duro	Madeira
7	Amarelo	Duro	Metal
8	Amarelo	Duro	Plástico
9	Verde	Duro	Madeira
10	Verde	Moderado	Plástico

Tabela 8 – Redução do sistema de informação original.

Os subconjuntos de atributos obtidos através da redução do sistema de informação são capazes de manter as mesmas propriedades da representação de conhecimento quando esta é feita utilizando todos os atributos.

7 Geração de Regras

Os atributos encontrados através do método de redução do sistema de informação podem ser descritos na forma de regras, ou seja, as regras de classificação são extraídas do banco de dados reduzido. Para transformar um reduto em regras, deve-se somente unir o atributo de decisão.

Para exemplificar as regras geradas pelos métodos abordados, seja o Reduto $B^* = \{\text{Cor, Tato, Material}\}$ do sistema de informação A da Tabela 1. As regras de decisão que descrevem a dependência de $\{\text{Atitude}\}$ com relação a B^* podem ser representadas na forma de “Se ... então ...” e podem ser vistas no Quadro 1.

As regras R_6 e R_9 têm as mesmas características para os Atributos Condicionais Cor, Tato e Material, mas diferentes decisões. Portanto, aplicando esta regra, não se pode afirmar que a decisão será correta. Regras desse tipo são chamadas de não-determinísticas (*inconsistentes*), enquanto que as regras R_1 , R_2 , R_3 , R_4 ,

R_5 , R_7 , R_8 e R_{10} são chamadas de determinísticas (*consistentes*). Desta forma, apenas as regras consistentes são consideradas como válidas para classificar o sistema de informação.

O ganho desta técnica para o exemplo considerado na Tabela 1, foi à redução de operadores “E” que passou de quatro para dois.

8 Considerações Finais

Neste trabalho, apresentou-se a teoria de Rough Sets baseado no método para aprendizagem e obtenção de conhecimento em banco de dados. Definiu-se primeiramente um sistema de informação, a indiscernibilidade entre objetos, as aproximações dos conjuntos bem como a qualidades de cada aproximação, a redução do sistema de informação e a geração de regras. Pode-se observar que a aproximação de conjuntos está diretamente ligada “a indiscernibilidade ou relação de não-discernibilidade”.

A redução do sistema de informação e consequentemente a geração de regras são o objetivo desta teoria uma vez que o sistema de informação pode ser desnecessariamente grande. Utilizando a teoria de Rough Sets no exemplo proposto na Tabela 1, conclui-se que houve redução de atributos irrelevantes, conforme a Tabela 8. Esta redução implica diretamente na diminuição do custo computacional do sistema. As regras de classificação geradas são determinísticas, uma vez que precisam ser consistentes com todos os dados do sistema de informação.

Referências

- [1] Pawlaz, Z. Rough sets. International Journal of Computer and Information Sciences, p. 341-356, 1982.
- [2] Pawlaz, Z. Rough sets: theoretical aspects of reasoning about data. Kluwer Academic Publishers, 1991.
- [3] Ohrn, A. Rosetta. Technical reference manual. Relatório técnico., Knowledge Systems Group. Department of Computer and Information Science. Norwegian University of Science and Technology, Norway, 1999.

Quadro 1 – Regras de classificação geradas.

R_1 : SE	Cor = Azul	E	Tato = Duro	E	Material = Plástico	ENTÃO	Atitude = Negativa
R_2 : SE	Cor = Vermelho	E	Tato = Moderado	E	Material = Madeira	ENTÃO	Atitude = Neutra
R_3 : SE	Cor = Amarelo	E	Tato = Macio	E	Material = Pelúcia	ENTÃO	Atitude = Positiva
R_4 : SE	Cor = Azul	E	Tato = Moderado	E	Material = Plástico	ENTÃO	Atitude = Negativa
R_5 : SE	Cor = Amarelo	E	Tato = Macio	E	Material = Plástico	ENTÃO	Atitude = Neutra
R_6 : SE	Cor = Verde	E	Tato = Duro	E	Material = Madeira	ENTÃO	Atitude = Positiva
R_7 : SE	Cor = Amarelo	E	Tato = Duro	E	Material = Metal	ENTÃO	Atitude = Positiva
R_8 : SE	Cor = Amarelo	E	Tato = Duro	E	Material = Plástico	ENTÃO	Atitude = Positiva
R_9 : SE	Cor = Verde	E	Tato = Duro	E	Material = Madeira	ENTÃO	Atitude = Neutra
R_{10} : SE	Cor = Verde	E	Tato = Moderado	E	Material = Plástico	ENTÃO	Atitude = Neutra