SISTEMAS & GESTÃO
Revista Eletrônica

AGREGAÇÃO AUTOMÁTICA DE CLASSES NOS ATRIBUTOS DE DECISÃO EM APLICAÇÕES DE ROUGH SETS COM DOMINÂNCIA

AUTOMATIC AGGREGATION OF CLASSES IN THE DECISION ATTRIBUTES IN APPLICATIONS OF DOMINANCE-BASED ROUGH SETS

Annibal Parracho Sant'Anna[a]; Roberto Malheiros Moreira Filho[b]

[a] Universidade Federal Fluminense (UFF) - Niterói, RJ, Brasil - Programa de Pós-Graduação em Engenharia de Produção

[b] Universidade Federal de Juiz de Fora (UFJF) - Juiz de Fora, MG, Brasil - Departamenteo de Engenharia de Produção e Mecânica

Resumo

A Teoria dos Conjuntos Aproximativos tem seu já elevado potencial de utilização ampliado quando se introduz uma abordagem probabilística nas aplicações com relação de dominância (DRSA). Aqui são propostos métodos para elevar a qualidade da aproximação, o número de reduções e a variedade de regras de decisão em DRSA, baseados na agregação de classes segundo o atributo de decisão. As formas de agregação propostas podem ser usadas em conjunto com as outras alternativas à DRSA já conhecidas. Enquanto essas outras estratégias, para elevar o valor do índice de qualidade da aproximação, levam em conta, apenas, as cardinalidades das classes e das regiões de fronteira, a abordagem aqui desenvolvida emprega também a distância entre elas. Duas propostas de agregação são apresentadas, uma baseada em densidade de probabilidade e outra em probabilidade de atingir valores extremos.

Palavras-chave: Conjuntos Aproximativos – Qualidade da Aproximação – Dominância – Probabilidade – Agregação de Classes

Abstract

Rough Sets Theory has its already large utility enhanced when a probabilistic approach is introduced in the applications with attributes presenting dominance relations (DRSA). Methods are here proposed to improve the quality of approximation, the number of reducts and the variety of decision rules in DRSA, based on the aggregation of classes of values of the decision attribute. The forms of aggregation proposed may be used in conjunction with other already known alternatives to DRSA. While these other strategies to elevate the value of the index of quality of approximation take into account only the cardinality of classes and border regions, the approach developed here employs also the distance between them. Two aggregation proposals are presented, one based on probability density and the other on probability of attaining extreme values.

Keywords: *Rough Sets – Quality of Approximation – Dominance – Probability – Aggregation of Classes*

1. INTRODUÇÃO

A Teoria dos Conjuntos Aproximativos – TCA se baseia na busca de discernir classes identificadas por um conjunto de atributos de decisão por regras alternativas envolvendo um outro conjunto de atributos, chamados atributos de condição.

A caracterização da aspereza (*roughness*) na identificação das unidades observadas tem se concentrado no exame do vetor das avaliações de cada unidade de observação segundo os diversos atributos, mas envolve também a precisão na medição de cada variável. Este último aspecto influencia a qualidade na seleção de atributos e consequentemente nas regras de decisão criadas.

Uma etapa importante da aplicação da TCA é a identificação das reduções (na tradução de Couto e Gomes, 2010; *reducts*, em inglês) do conjunto de atributos de condição, conjuntos mínimos de variáveis de condição capazes de oferecer a mesma qualidade da aproximação do conjunto de todos os atributos de condição para identificar os conjuntos de observações indiscerníveis pelas variáveis

PROPPI / LATEC
DOI: 10.7177/sg.2013.v8.n1.a6

Revista Eletrônica Sistemas & Gestão
Volume 8, Número 1, 2013, pp. 68-76
DOI: 10.7177/sg.2013.v8.n1.a6

de decisão. Dependendo da aspereza na medição de cada variável, o número de reduções será maior ou menor, assim como o número de variáveis em cada redução.

Extensões da TCA são apresentadas por diversos autores. O desenvolvimento detalhado de diversas dessas extensões pode ser acompanhado em DEMBCZYNSKI *et al*. (2009), FORTEMPS *et al*. (2008) e SLOWINSKI *et al*. (2004, 2012). Um desenvolvimento em particular será objeto de atenção aqui. Diz respeito às situações em que os atributos, incluindo os de decisão, possuem, em suas classificações, uma ordenação natural como, por exemplo, a ordenação em uma escala de preferência ou em uma escala de dominância (Pawlak, 1991). Quando se lida com atributos com escalas de dominância, tem-se o que Greco *et al*. (2001) denominaram *Dominance-based Rough Sets Approach* – DRSA.

A Abordagem para a Teoria dos Conjuntos Aproximativos com Dominância e Consistência Variável – VC-DRSA é desenvolvida por Greco *et al*. (2001a) para permitir um relaxamento nas regras de entrada de objetos nas aproximações, desprezando contradições avaliadas como pouco frequentes. Novas propostas no mesmo sentido são aqui apresentadas baseadas na união de classes próximas segundo os atributos de decisão. Enquanto o VC-DRSA e variantes dele derivadas se baseiam no tamanho das classes envolvidas, a abordagem aqui desenvolvida se baseia da distância entre seus valores.

A quantidade de valores disponíveis para o atributo de decisão pode ser uma das causas da imprecisão. A análise da quantidade de valores para o atributo de decisão pode se assemelhar ao estudo do número de categorias da escala de Likert, em que um grande número de categorias pode levar a interpretações distintas e alterar o resultado da pesquisa (Alexandre *et al*., 2003). A ideia aqui explorada é a de reduzir a precisão na identificação segundo os atributos de decisão empregando um número menor de possíveis valores para eles.São aqui propostos dois métodos de união de classes baseados na aplicação de algoritmos automáticos de aumento da discretização, um baseado na suavização da distribuição de probabilidades empírica determinada pelos atributos de decisão e outro baseado na transformação dos vetores de observações iniciais de cada um desses atributos em um vetor de probabilidades de as unidades observadas apresentarem valor extremo.

Será considerado aqui o caso mais frequente, de um único atributo de decisão. A extensão dos métodos propostos para o caso de mais de um atributo de decisão é simples, desde que se extraia uma relação de dominância única segundo o conjunto de todos os atributos de decisão reunidos. Para gerar essa relação, pode-se usar a composição de probabilidades de preferências (Sant'Anna & Sant'Anna, 2001).

2. A TEORIA DOS CONJUNTOS APROXIMATIVOS

A aplicação da TCA dá-se após a obtenção de um conjunto de dados iniciais. Estes dados são organizados em um sistema de informação, que pode ser representado por uma matriz contendo, nas linhas, as medidas de todos os atributos dos objetos individuais observados e, nas colunas, os valores dos atributos, divididos em dois grupos: atributos de condição e atributos de decisão. Os objetos do sistema de informação original são agrupados segundo seus valores nesses dois grupos de atributos. Se dois objetos possuem os mesmos valores para todos os atributos de um conjunto de atributos de condição, eles são considerados indiscerníveis. O que gera inconsistência em TCA é a possibilidade de objetos classificados como indiscerníveis possuírem classificações diferentes segundo os atributos de decisão.

Para cada conjunto de atributos de condição P e cada objeto x do universo de objetos observados U, denote-se por P(x) o conjunto dos objetos de U indiscerníveis de x segundo P, isto é,

$$P(x) = \{y \in U \ / \ x \ e \ y \ são \ indiscerníveis \ segundo \ P\} \qquad (1)$$

Para todo X, subconjunto de U, e todo conjunto de atributos de condição são definidas duas aproximações: aproximação inferior de X segundo P, $\underline{P}(X)$, e aproximação superior de X segundo P, $\overline{P}(X)$, definidas por

$$\underline{P}(X) = \{x \in U \mid P(x) \subseteq X\} \ \text{ e } \ \overline{P}(X) = \{x \in U \mid P(x) \cap X \neq 0\}. \qquad (2)$$

Uma região de fronteira pode ser definida como

,

$$F_P(X) = \overline{P}(X) - \underline{P}(X) \qquad (3)$$

ou seja, a diferença entre as aproximações superior e inferior (é importante notar que, por definição $\underline{P}(X) \subseteq P(X) \subseteq \overline{P}(X)$). Pawlak (1982) define um conjunto X como áspero (*rough*) se a sua região de fronteira for não vazia e como preciso (*crisp*) se a sua região de fronteira for um conjunto vazio.

Pawlak (1991) define também o índice γ para medir a qualidade da aproximação pelo conjunto de atributos de condição P da partição $\{X_i\}_{i=1,\dots,n}$ de U determinada pelo conjunto de atributos de decisão D. O índice γ é definido como a proporção de objetos cuja classificação segundo os atributos de decisão não é contrariada por indiscernibilidades segundo os atributos de condição considerados. Pode ser calculado, complementarmente, a partir da proporção dos objetos em região de fronteira. Formalmente, o índice de qualidade de aproximação de D por P é

Revista Eletrônica Sistemas & Gestão
Volume 8, Número 1, 2013, pp. 68-76
DOI: 10.7177/sg.2013.v8.n1.a6

$$\gamma_P(D) = 1 - \sum_i \frac{|F_P(X_i)|}{|U|} \quad (4)$$

O objetivo da TCA é estabelecer regras de decisão: critérios, descritos em termos de valores dos atributos de condição, para pertinência às classes determinadas pelos atributos de decisão. Quanto mais simples, mais fáceis de usar as regras de decisão. Se um conjunto de atributos de condição P tem o mesmo índice de qualidade de aproximação de outro maior, isto é, $\gamma_P(D) = \gamma_Q(D)$ para $\mathrm{P} \subseteq \mathrm{Q}$, isto indica que não há nenhuma preferência entre dois objetos x e y segundo os elementos de Q\P que torne falsa a relação de indiscernibilidade entre esses objetos que ampliava alguma fronteira segundo P. Portanto, podem ser usados apenas os atributos em P nas regras de decisão em que se usem, além dos atributos em P, atributos em Q\P.

Pawlak (1991) define como uma redução de um conjunto de atributos de condição qualquer subconjunto dele com o mesmo valor de γ. Quanto menor seja o número de atributos na redução mais simplesmente se podem formular as regras de decisão.

Se os atributos considerados determinam relações de dominância entre os objetos do universo U, isto é, se determinam uma ordem em U, fica mais fácil medir a qualidade da aproximação. A especialização da TCA para o caso de dominância como exposta a seguir foi desenvolvida por Greco *et al*. (2001).

Supõem-se ordenadas as classes da partição $\{Cl_t\}_{t=1,\ldots,n}$ determinada pelos atributos de decisão de modo que se $r > s$, os objetos pertencentes a Cl_r são preferíveis em relação aos objetos de Cl_s. Então, basta trabalhar com as classes acumuladas, superiores ou inferiores:

$$Cl_t^{\geq} = \bigcup_{s \geq t} Cl_s \quad ou \quad Cl_t^{\leq} = \bigcup_{s \leq t} Cl_s \quad ; \quad t = 1,\ldots,n \quad (5)$$

Assim, o objeto x pertence a $Cl_t^{\geq}$ se e só se pertence à classe Cl_t ou é preferível aos elementos desta classe. Analogamente para $x \in Cl_t^{\leq}$.

Quanto aos atributos de condição, para que x domine y em relação ao conjunto de atributos, o que se denota por x P-domina y, y é P-dominado por x ou xD_Py, exige-se que a preferência por x seja maior ou igual á preferência por y segundo todos eles.

Para relações de dominância, pode-se definir um conjunto de objetos dominando um dado objeto x (conjunto P-dominante):

$$D_P^+(x) = \{y \in U : yD_Px\} \quad (6)$$

e um conjunto de objetos dominados por x (conjunto P-dominado):

$$D_P^-(x) = \{y \in U : xD_Py\} \quad (7)$$

Uma vez definidas as classes e toda a lista de conjuntos de objetos P-dominantes e P-dominados, é possível definir as aproximações superior e inferior. Em relação à definição da seção anterior, a diferença é que, agora, são usadas relações de dominância ao invés de indiscernibilidade.

As aproximações inferior e superior podem ser definidas, respectivamente, para as classes acumuladas superiores, por:

$$\underline{P}(Cl_t^{\geq}) = \{x \in U : D_P^+(x) \subseteq Cl_t^{\geq}\} \quad \text{e}$$
$$\overline{P}(Cl_t^{\geq}) = \{x \in U : D_P^-(x) \cap Cl_t^{\geq} \neq \varnothing\} \quad (8)$$

Analogamente, para as classes acumuladas inferiores:

$$\underline{P}(Cl_t^{\leq}) = \{x \in U : D_P^-(x) \subseteq Cl_t^{\leq}\} \quad \text{e}$$
$$\overline{P}(Cl_t^{\leq}) = \{x \in U : D_P^+(x) \cap Cl_t^{\leq} \neq \varnothing\} \quad (9)$$

As regiões de fronteira e o índice de qualidade de aproximação podem ser definidos para o caso de dominância como antes, sendo bastante considerar as fronteiras das classes acumuladas. Da mesma forma, todas as regras de decisão podem ser extraídas de regras determinando a pertinência a classes acumuladas.

3. VC-DRSA

A VC-DRSA, como proposta por Greco *et al*. (2001a), é baseada em um relaxamento no critério de entrada da aproximação inferior, controlado por um nível de consistência *l*, que pode variar entre zero e um. No caso de uma classe inferior, para que um objeto *x* da classe pertença a sua aproximação inferior, em vez de exigir que nenhum objeto exterior à classe domine *x*, permite-se que alguns não pertençam, desde que a proporção dos que pertençam seja superior a *l*. Analogamente, para uma classe superior $Cl_t^{\geq}$, a aproximação inferior de nível de consistência *l* será:

$$\underline{P}^l(Cl_t^{\geq}) = \left\{x \in Cl_t^{\geq} : \frac{|Dp^+(x) \cap Cl_t^{\geq}|}{|Dp^+(x)|} \geq l\right\} \quad (10)$$

Revista Eletrônica Sistemas & Gestão
Volume 8, Número 1, 2013, pp. 68-76
DOI: 10.7177/sg.2013.v8.n1.a6

Assim, se o nível de consistência for fixado em 1, o critério para a entrada será o mesmo da definição original, mas, para *l* < 1, mais alguns objetos podem fazer parte da aproximação inferior.

Esse relaxamento visa a enfrentar a dificuldade decorrente do excessivo rigor da exigência de nenhuma inconsistência entre as duas ordenações. O seguinte exemplo mostra o problema. Nele, supõe-se que se deseja obter regras de decisão sobre a despesa com certo remédio (D) usando, como variáveis de condição, idade (p_1) e rendimentos de salário (p_2). Para isso, supõe que se dispõe dos dados da Tabela 1 sobre os valores médios dessas variáveis em seis municípios de uma região.

Tabela 1. Dados de Despesas, Idade e Salário Médios com Despesas Próximas

Objetos	D	P1	P2
u	25	40	100
v	20	45	200
w	20	50	150
x	80	55	300
y	80	60	250
z	75	65	350

Como o município u é $\{p_1,p_2\}$ – dominado pelos municípios *v* e *w* e tem valor maior que o desses dois municípios no atributo de decisão, a classe de despesa mais baixa formada por estes dois últimos municípios tem como fronteira

$$F_{\{p_1,p_2\}}(\{v,w\}) = \{u,v,w\} \tag{11}$$

Da mesma forma, como o município *z* domina os municípios *x* e *y* e tem valor menor que o desses dois no atributo de decisão, a classe superior formada por *x* e *y* tem como fronteira

$$F_{\{p_1,p_2\}}(\{x,y\}) = \{x,y,z\} \tag{12}$$

Assim, todos os objetos pertencem a alguma fronteira e a qualidade da aproximação é a pior possível.

É possível que a inversão verificada no extremo superior da despesa seja explicável por um benefício concedido aos idosos, de isenção de pagamento de remédios. É também possível que a inversão no extremo inferior decorra de esse remédio ser um bem inferior, cujo consumo cresce se o indivíduo tem renda muito baixa. Se este é o caso, a adição de mais dois atributos de condição, o indicador de idade inferior a 65 e o indicador de renda salarial inferior a 150, que tornam indiscerníveis, respectivamente, o município u dos municípios *v* e *w* e o município *z* dos municípios x e *y*, esvaziará todas as fronteiras e elevará o valor do índice de qualidade da aproximação de 0 para 1. Restará, então, avaliar se a inclusão desses dois detalhes na análise é um esclarecimento necessário.

E é também possível que as inversões observadas nos extremos decorram simplesmente de imprecisão das medidas que muitas vezes afeta a coleta de dados menos comuns. Note-se que a diferença de 5 unidades no consumo que provocou essa inconsistência é pequena, relativamente à variação de 50 unidades entre os valores centrais de 25 e 75.

O critério adotado em VC-DRSA conduz a que um objeto *x* seja aceito na aproximação inferior de uma classe inferior enquanto outros objetos dominados por *x*, por dominarem menos objetos, não são aceitos. Ora, ser dominado e não dominar seria maior motivo para pertencer à aproximação inferior da classe inferior. Trocando a condição de ser dominado por dominar, idêntica incoerência ocorre nas aproximações superiores das classes superiores. Diversas alternativas (Blasczynski *et al*., 2006; Inuiguchi e Ishioka, 2006; Inuiguchi *et al*. 2009, Deng *et al*, 2011) têm sido propostas para evitar essa dificuldade, mas todas têm preservado o princípio de estabelecer o nível de consistência em termos de proporção de inconsistências, sem levar em conta os espaçamentos com que as dominâncias se verificam.

4. AGREGAÇÃO DE CLASSES

Neste trabalho, se propõe uma alteração nos dados originais com um aumento da *roughness* das medidas. A ideia básica é encontrar classes de valores pouco frequentes e alocá-las a uma classe imediatamente superior ou inferior, diminuindo assim os valores disponíveis para o atributo de decisão e, consequentemente, propiciando o aumento da qualidade da aproximação da TCA e o aumento do número de reduções.

É fácil provar que, ao aumentar o tamanho dos grãos do atributo de decisão juntando valores próximos e manter os valores dos atributos de condições, se eleva a qualidade da aproximação e se aumenta o número de reduções.

De fato, a qualidade da aproximação aumenta sempre que se reduza a cardinalidade de regiões de fronteira, isto é, sempre que se reduza o número de objetos indiscerníveis de algum objeto *x* situados fora da classe em que o atributo de decisão situa esse objeto. Se não se alterarem os valores

Revista Eletrônica Sistemas & Gestão
Volume 8, Número 1, 2013, pp. 68-76
DOI: 10.7177/sg.2013.v8.n1.a6

dos atributos de condição, os objetos indiscerníveis de *x* continuam os mesmos. Assim, ao aumentar-se a classe em que o atributo de decisão o situa, só é possível diminuir o número de objetos indiscerníveis de *x* fora dessa classe.

Quanto ao aumento do número de reduções, pode ser provado em duas etapas. Primeiro, é fácil ver que nenhuma redução se perde ao unir classes segundo o atributo de decisão. De fato, se P_1 é uma redução do conjunto de atributos P, é porque nenhuma indiscernibilidade, segundo atributos em P/P_1, exclui objetos de alguma fronteira. Com a ampliação do tamanho das classes pelos atributos de decisão e a permanência das dominâncias segundo os atributos de condição, não há como suscitar a ocorrência de novas exclusões.

Por outro lado, ao agregar classes pelo atributo de decisão abre-se a possibilidade de extrair, de qualquer redução, reduções com menos atributos. De fato, se, antes da agregação, P2 era redução e P1 $\subseteq$ P2 não era, é porque indiscernibilidade de pelo menos um par de objetos (*x*, *y*) segundo os elementos de P2 que colocava um desses objetos na fronteira de alguma classe é eliminada por dominância segundo pelo menos um atributo em $P_2 \setminus P_1$. Com a agregação das classes, todos os objetos que por tal motivo iam para uma fronteira podem passar a pertencer à classe do outro elemento do par.

O exemplo a seguir demonstra concretamente como a qualidade da aproximação e o número de reduções crescem com a união de classes.

Tabela 2. Agregação de Dados de Despesas, Idade e Salário Médios

Objetos	*D*	*P1*	*P2*	*D2*
u	20	40	150	25
v	25	40	200	25
w	50	50	250	50
x	60	55	250	60
y	70	60	300	70
z	80	65	350	80

Na Tabela 2, D_2 é o resultado da junção das duas classes no extremo inferior do vetor de valores de D. Com D como atributo de decisão, P_1 e P_2 são necessários para atingir a qualidade de aproximação máxima, igual a 1. Já com D_2 substituindo D, é fácil ver que P_1 sozinho atinge a qualidade de aproximação igual a 1.

5. ALGORITMOS DE AGREGAÇÃO

São apresentados aqui dois algoritmos automáticos para reduzir o número de classes. Como a redução de classes não impede a subsequente aplicação de qualquer método de extração de regras, eles podem ser aplicados em conjunto com qualquer uma das técnicas desenvolvidas para elevar a qualidade da aproximação a partir do VC-DRSA.

A primeira proposta envolve encontrar uma melhor distribuição de probabilidades que suavize a variação do atributo de decisão. É baseada na identificação, com o auxílio do software Arena® 12.0, da distribuição de probabilidade que melhor se ajuste a um histograma.

Se as classes tiverem a mesma participação, uma distribuição uniforme será a mais adequada (pois o histograma desta distribuição é uma reta paralela ao eixo dos x). Porém, outras distribuições podem representar melhor situações em que se observem concentrações de valores. Nessas distribuições com concentração de valores, a importância das classes será considerada.

Se o sistema possui *n* objetos e *t* classes de valores para o atributo de decisão, uma distribuição uniforme sugeriria uma frequência de *n/t* valores em cada classe. Havendo concentração de valores em determinadas classes, algumas classes ficarão abaixo da frequência *n/t*. Essas classes serão candidatas a unirem-se às classes vizinhas para formar uma classe com maior representação em termos de frequência.

As distribuições normal e triangular, por exemplo, sugerem que as classes iniciais e finais têm pouca participação na frequência, a primeira com simetria e a segunda podendo representar decréscimo mais suave de um lado (dependendo dos parâmetros da distribuição triangular). Uma distribuição exponencial pode corresponder a classes finais de um dos lados com muito pouca participação.

O critério para que uma classe se una à classe vizinha (imediata superior ou inferior) será o impacto adicional desta classe na função acumulada de probabilidade. Se for uma contribuição pequena, a classe é candidata a esta união. Se há concentração dos valores do atributo de decisão nos valores altos, uma distribuição assimétrica com concentração de valores à direita será escolhida e serão unidas classes de valores pequenos. Inversamente se a concentração é do outro lado. Se os valores do atributo de decisão tiverem uma concentração central, as classes sujeitas à união estarão tanto à esquerda como à direita.

Desta união de classes, espera-se que as regiões de fronteira tenham suas cardinalidades diminuídas e, naturalmente, que a qualidade da aproximação aumente. Assim como em VC-DRSA (em que o índice *l* para a entrada de objetos na aproximação inferior não é fixo) não haverá um único valor fixo para a união de classes. O índice a ser apresentado, nomeado de *l**, poderá ser alterado de acordo com a conveniência do pesquisador.

Da mesma maneira que em VC-DRSA, um valor muito baixo de *l* representa um relaxamento muito grande para a entrada de valores nas aproximações inferiores, nesta

Revista Eletrônica Sistemas & Gestão
Volume 8, Número 1, 2013, pp. 68-76
DOI: 10.7177/sg.2013.v8.n1.a6

proposta um valor alto de *l** possibilita a união de muitas classes. Portanto, embora um limite para o valor de *l** não seja fixado, deve-se ter em mente que um alto valor deste índice pode permitir a união de classes que não necessariamente estejam na cauda da função densidade de probabilidade.

A proposta para a redução de classes no atributo de decisão pode ser precisamente apresentada da seguinte forma para os três tipos de concentração:

- Se a concentração de valores for à direita: as frequências baixas ficam à esquerda (ou para valores baixos do atributo de decisão). As classes candidatas à união são as que incluem os valores abaixo do limite s = s(*l**), tal que

$$\int_{-\infty}^{s} f(x)\,dx = l^* \tag{13}$$

- Se a concentração de valores for à esquerda: as frequências baixas ficam à direita (ou para valores altos do atributo de decisão). As classes candidatas à união são as que incluem os valores acima do limite s = s(*l**), tal que

$$\int_{s}^{+\infty} f(x)\,dx = l^* \tag{14}$$

- Se a concentração de valores for central: as frequências baixas ficam à esquerda e à direita. As classes candidatas à união são as que incluem os valores abaixo de um limite $s_1 = s_1$ (*l**) e as que incluem os valores acima de outro limite $s_2 = s_2$ (*l**).

$$\int_{-\infty}^{s_1} f(x)\,dx = \frac{l^*}{2} \; e \int_{s_2}^{+\infty} f(x)\,dx = \frac{l^*}{2} \tag{15}$$

A outra proposta de agregação de classes envolve atribuir um erro de medida do atributo de decisão para cada objeto. Levando em conta a distribuição de probabilidades deste erro, se transformam os valores exatos do atributo de decisão em distribuições de probabilidades. Como, agora, os valores são funções de probabilidade, é possível calcular a probabilidade de determinado objeto apresentar o maior (ou o menor) valor em uma replicação qualquer.

A combinação de TCA com transformações probabilísticas já foi realizada em Sant'Anna (2004). É difícil definir qual a precisão adequada para medir cada variável, visto que suas escalas naturais podem ser muito diferentes. A transformação probabilística torna possível reduzir todas as variáveis a uma mesma escala natural, qual seja a escala da probabilidade de atingir uma posição extrema, isto é, de a unidade de observação apresentar o maior ou o menor valor em replicações aleatórias da mesma medição.

Nesta proposta, os valores do atributo de decisão são tratados como médias de distribuições normais. O uso da distribuição normal para cada classificação do atributo de decisão permite considerar que esta classificação tem uma imprecisão. A variância de cada uma dessas distribuições normais pode ser estimada ou arbitrada.

Por exemplo, pode-se assumir que todas as classificações possuem a mesma variância e usar, como estimador dessa variância, a variância da amostra de valores observados para o atributo. Sendo assim, quanto maior for a variação do atributo de decisão em torno de seu valor médio, maiores são as chances de uma classificação inferior vir a ser, de fato, maior que uma outra classificação superior (uma vez que os valores são transformados em normais com média igual a própria avaliação e com certa variância).

Uma vez pensados os valores absolutos como médias de distribuições independentes, pode-se calcular a probabilidade de, em uma replicação da avaliação, cada uma dessas distribuições apresentar um valor mais alto que todas as outras, se o objetivo for encontrar a probabilidade de uma determinada classificação ser a maior entre todas do universo de objetos. Se o objetivo for encontrar a probabilidade de uma determinada classificação ser a menor, deve-se calcular a probabilidade desta distribuição apresentar o menor valor entre todas as distribuições.

Portanto, cada valor do atributo terá uma probabilidade de maximizar (ou minimizar) o valor no universo agora transformado em um conjunto de funções de probabilidade. Como algum objeto sempre será o máximo (ou o mínimo) em uma replicação, a soma de todas as probabilidades do vetor criado desse modo será sempre 1.

A probabilidade de maximizar de um objeto x, com valor igual ao de outro objeto y no atributo de decisão, é a mesma. Porém, nunca será a mesma de, por exemplo, um objeto z que pertença a uma classe vizinha superior. A probabilidade de um objeto com média maior que outro ser o máximo em uma replicação qualquer é maior do que a probabilidade do objeto com média menor. O raciocínio se inverte quando a transformação é baseada na probabilidade de ser o mínimo.

O que se explora nessa transformação é o fato de a probabilidade de maximizar (ou minimizar) ser tão próxima do valor zero em algumas classes que elas podem se unir para representar o conjunto de objetos com probabilidades de ser o máximo (ou o mínimo) em uma replicação abaixo de um certo limite. O critério aqui adotado será a aproximação até a terceira casa decimal, ou seja, probabilidades menores

Revista Eletrônica Sistemas & Gestão
Volume 8, Número 1, 2013, pp. 68-76
DOI: 10.7177/sg.2013.v8.n1.a6

que 0,0005 (que quando arredondadas para três casas decimais ficam iguais a zero) serão consideradas equivalentes a zero; portanto, identificarão classes de objetos candidatas à união.

Se o objetivo for explorar a probabilidade de uma classe apresentar o máximo em uma replicação dos valores, as probabilidades menores serão das classes à direita (ou com valores baixos para o atributo de decisão). De maneira análoga, se o objetivo for explorar a probabilidade de uma classe ser o mínimo em uma replicação dos valores, as probabilidades menores serão das classes à esquerda (ou com valores baixos para o atributo de decisão).

6. EXEMPLOS

Considere-se o sistema de informações tratado em Moreira Filho (2012) composto da seguinte maneira:

• Atributos de condição: acabamento, compatibilidade e durabilidade (todos com dominância com valores de 1 a 5);

• Atributo de decisão: avaliação global (com dominância com valores de 1 a 5).

A aplicação da Teoria dos Conjuntos Aproximativos foi feita com o auxílio do software jMAF® (Blaszczynski *et al*, 2009). O total de 58 objetos resultou em uma qualidade de aproximação de 0,483 (28 dos 58 objetos são classificados de forma consistente pelas regras de decisão).

A análise dos valores do atributo de decisão permitiu a união de algumas classes para aumentar a qualidade da aproximação.

A distribuição sugerida foi normal com média 3,07 e desvio-padrão 1,03. As classes 1 e 5 são menos representativas e acumulam os menores percentuais de frequência. Se a classe do valor 1 se unir à classe do 2 e a classe do 5 se unir à classe do 4, tem-se uma nova definição de valores para o atributo de decisão:

• a: representando os valores 1 e 2 (baixa avaliação geral);

• b: representando o valor 3 (avaliação média);

• c: representando os valores 4 e 5 (avaliação superior).

A classe a, com cardinalidade 16, representa o conjunto das duas classes inferiores no sistema de informação anterior. O mesmo raciocínio pode ser estendido à classe c (com cardinalidade igual a 19, que representa a união das duas classes superiores no sistema de informações anterior).

A qualidade da aproximação passa de 0,483 (28 em 58) para 0,741 (43 em 58).

A outra possibilidade de união de classes do atributo de decisão explorada é a transformação desta variável em probabilidade de o valor de um objeto ser o máximo da classe. Usando como função densidade a distribuição normal previamente apresentada e aproximando com apenas três casas decimais os valores da probabilidade de maximizar, os valores associados aos objetos com valores pequenos para o atributo de decisão passam a ter o mesmo valor.

Com a aproximação de três casas decimais, por este método apenas as classes dos valores 1 e 2 ficam unidas, com probabilidade de maximização zero. Os valores obtidos na maximização do vetor do atributo de decisão são:

• 1 e 2: probabilidade de maximização igual a 0,000;

• 3: probabilidade de maximização igual a 0,002;

• 4: probabilidade de maximização igual a 0,021;

• 5: probabilidade de maximização igual a 0,133.

Como menos classes foram unidas, a qualidade da aproximação ficou em uma posição intermediária com valor de 0,69 (40 em 58).

A análise seguinte é de um sistema de informações tratado em Moreira Filho (2012), com 41 objetos com os mesmos atributos de condição dos exemplos acima e o atributo de decisão com valores de 1 a 7.

O cálculo das aproximações feitas com índice de consistência l de 1, ou seja, nenhum relaxamento sendo permitido na entrada de objetos na aproximação inferior, resulta em qualidade de aproximação de 0,366 (15 dos 41 objetos são consistentes pelas regras de decisão). Se for utilizado o algoritmo de relaxamento de entrada de objetos na aproximação inferior através da utilização do índice l=0,85, a qualidade da aproximação irá aumentar para 0,780 (32 objetos consistentes).

A distribuição de melhor ajuste é exponencial de média 1,72. Como a distribuição tem concentração de valores à esquerda, as classes no extremo superior são candidatas a se unirem. Com nível de consistência *l*=0,85 e *l**= 0,31, as classes superiores com limites inferiores maiores que 2 devem ser unidas. Assim, as novas classes são:

• a: representando a classe 1 original;

• b: representando a classe 2 original;

• c: representado as classes 3, 4, 5, 6 e 7 originais.

A qualidade da aproximação passa de 0,780 (32 em 41) para 0,829 (34 em 41).

Na segunda abordagem, emprega-se a transformação na probabilidade de ser o menor em uma replicação. As probabilidades obtidas com a aproximação até a terceira casa decimal e a reclassificação decorrente são dadas na Tabela 3.

Revista Eletrônica Sistemas & Gestão
Volume 8, Número 1, 2013, pp. 68-76
DOI: 10.7177/sg.2013.v8.n1.a6

Tabela 3. Classes pelas Probabilidades de Minimizar

Inicial	*Probabilidade*	*Nova*
1	0,048	A
2	0,011	B
3	0,002	C
4	0,000	D
5	0,000	D
6	0,000	D
7	0,000	D

Verifica-se na Tabela 3 que a aproximação nula para a probabilidade passa a ocorrer a partir do valor 4 e, em vez de 3 classes como na aproximação anterior, passa-se a ter neste caso 4 classes. Entretanto, com o índice de consistência de *l*=0,85, as mesmas regras de decisão são obtidas e a qualidade da aproximação é elevada também, como na abordagem anterior, para 0,829.

Exemplos com outras distribuições assim como o detalhamento dos sistemas de informações utilizados nesta seção são discutidos em Moreira Filho (2012).

7. CONCLUSÃO

A importância de um sistema de informação que permita uma análise eficiente é crescente. Muitos dados estão disponíveis e o uso racional dos mesmos torna-se de vital relevância. A TCA permite usar um sistema de informação para extrair regras de decisão baseadas nos dados, sem premissas sobre distribuição de probabilidade, independência entre objetos e conhecimento passado sobre os objetos. Isto a torna uma técnica interessante, com uso e divulgação aumentando com o passar dos anos desde o seu desenvolvimento por Pawlak em 1982.

Com o avanço dos estudos empregando a TCA, surgiram técnicas que estudam relações de dominância das avaliações dos atributos do sistema de informação original. As propostas iniciais avaliam a inclusão ou exclusão de objetos em relação a índices e em relação aos demais objetos do sistema de informação original. Nessas propostas, o critério original de entrada de objetos em uma aproximação inferior é rigoroso e muitas inconsistências podem ocorrer, dificultando a extração de regras.

Este estudo investigou o *trade-off* entre a quantidade de classes no atributo de decisão e as possibilidades de inconsistência. Diminuindo as classes de um atributo de decisão com dominância, segundo algum critério matemático, a qualidade da aproximação, relação direta entre a quantidade de objetos bem definidos e o total de objetos, melhora. Na prática, é importante avaliar o significado da união de algumas classes no atributo de decisão.

Duas novas propostas foram aqui apresentadas com resultados diferentes pelo critério de entrada. Ambas exploram uma nova ideia de rever os valores do atributo de decisão. Não há uma proposta melhor do que a outra. O que se deve saber é que, quanto mais relaxados os critérios de união de classes, maiores são as chances de aumento no índice de qualidade da aproximação. O mesmo acontece com a aplicação de VC-DRSA.

As propostas apresentadas seguiram definições probabilísticas, empregando a aproximação por funções de densidade de probabilidade e a transformação na probabilidade de um determinado valor ser o máximo ou o mínimo em uma replicação. Orientando a busca da união por critérios numéricos, tem-se uma base natural para essa avaliação.

Ao aplicar as técnicas de união de classes no atributo de decisão, nenhuma premissa para aplicação de TCA com relação de dominância é violada. Apenas algumas classes do atributo de decisão são unidas. As propriedades das aproximações inferior e superior continuam preservadas.

Outra possibilidade, a VC-DRSA, foi proposta por Greco *et al*. (2001) e trabalhada posteriormente por diversos autores. Nela, os dados do sistema de informação não são alterados; porém, os critérios para a classificação de objetos nas aproximações superior e inferior são alterados. O uso da união de classes no atributo de decisão não impede a aplicação simultânea de técnicas com alterações nos critérios de entrada na aproximação inferior. Assim, as estratégias de união de classes de atributos apresentadas neste trabalho podem ser aplicadas em conjunto com técnicas de VC-DRSA.

8. REFERÊNCIAS

ALEXANDRE, J. W. C., ANDRADE, D. F., VASCONCELOS, A. P., ARAUJO, A. M. S. & BATISTA, M. J. Análise do uso de categorias da escala Likert aplicada à gestão da qualidade total através da teoria da resposta ao item. **Anais do XXIII ENEGEP**, 2003.

BLASCZYNSKI, J., GRECO, S. & SLOWINSKI, R. & SZELAG, M. On Variable Consistency Dominance-based Rough Set Approaches. **Lecture Notes in Artificial Intelligence**, v. 4259, Berlin: Springer-Verlag, pp. 191-202, 2006.

BLASZCZYNSKI, J., GRECO, S., SLOWINSKI, R. & SZELAG, M.: Monotonic Variable Consistency Rough Set Approaches. **International Journal of Approximate Reasoning**, v. 50, n. 7, pp. 979-999, 2009.

COUTO, A. B. G. & GOMES, L. F. A. M. G. A Tomada de Decisão em Recursos Humanos com Dados Replicados e Inconsistentes: Uma Aplicação da Teoria dos Conjuntos Aproximativos. **Pesquisa Operacional**, v. 30, n.3, pp. 657-686, 2010.

Revista Eletrônica Sistemas & Gestão
Volume 8, Número 1, 2013, pp. 68-76
DOI: 10.7177/sg.2013.v8.n1.a6

DEMBCZYNSKI K, GRECO S & SLOWINSKI R. Rough set approach to multiple criteria classification with imprecise evaluations and assignments. **European Journal of Operational Research**, v. 198, pp. 626-636, 2009.

DENG, W., WANG, G., YANG, S. & HU, F. A new method for Inconsistent Multicriteria Classification. In ROUGH SETS AND KNOWLEDGE TECHONOLOGY'11, **Proceedings of the 6th International conference on Rough sets and Knowledge Technology,** pp. 600-609, 2011.

FORTEMPS PH, GRECO S & SLOWINSKI R. Multicriteria decision support using rules that represent rough-graded preference relations. **European Journal Operational Research**, v. 188, pp. 206-223, 2008.

GRECO, S., MATARAZZO, B., SŁOWIŃSKI, R. Rough sets theory for multi-criteria decision analysis. **European Journal of Operational Research**, 129, v.1, pp. 1–47, 2001.

GRECO, S., MATARAZZO, B., SLOWINSKI, R. & STEFANOWSKI, J. Variable consistency model of dominance-based rough set approach. In Ziarko, W & Yao, Y. (eds.): ROUGH SETS AND CURRENT TRENDS IN COMPUTING, **Lecture Notes in Artificial Intelligence**, v. 2005, Springer-Verlag, Berlin, pp. 170-181, 2001a.

INUIGUCHI & ISHIOKA, Variable-Precision Dominance-Based Rough Set Approach, ROUGH SETS AND CURRENT TRENDS IN COMPUTING, **Annals of the 5th International Conference on Rough Sets**, pp. 203-213, 2006.

INUIGUCHI, M., YOSHIOKA, Y. & KUSUNOKI, Y. Variable-precision dominance-based rough set approach and attribute reduction. **International Journal of Approximate Reasoning**, v. 50, pp. 1199-1214, 2009.

MOREIRA FILHO, R. M. **Redução de valores no critério de decisão em aplicações de Rough Sets com dominância e seus impactos na qualidade da aproximação**. Tese de doutorado em Engenharia de Produção, Universidade Federal Fluminense, 2012.

PAWLAK, Z. **Rough Sets, International Journal of Computer and Information Science**, v. 11, pp. 341-356, 1982.

PAWLAK, Z. **Rough Sets: Theoretical Aspects of Reasoning about Data**. Berlin: Springer, 1991.

SANT'ANNA, A.P. & SANT'ANNA, L.A.F.P. Randomization as a Stage in Criteria Combining. In: Ribeiro, J. L. D., Fogliatto, F. S. & Guimarães, L. B. M. (eds.) PRODUCTION AND DISTRIBUTION CHALLENGES FOR THE 21ST CENTURY, **Proceedings of the VII ICIEOM**, Salvador: Abepro, pp. 248-256, 2001.

SANT'ANNA A.P. Rough sets in the Probabilistic Composition of Preferences. In De Baets, B., D Caluwe, R, De Tré, G., Fodor, J., Kacprzyk, J & Zadrozny, S. (eds.) CURRENT ISSUES IN DATA AND KNOWLEDGE ENGINEERING, **Proceedings of Eurofuse2004**, p. 407- 414, 2004.

SLOWINSKI, R. GRECO, S. & MATARAZZO, B. **Inducing Robust Decision Rules from Rough Approximations of a Preference Relation**. Springer-Verlag, Berlim, Heildeberg, pp. 118-132, 2004.

SLOWINSKI, R. GRECO, S. & MATARAZZO, B. Rough set and rule-based multicriteria decision aiding. **Pesquisa Operacional**, v. 32, pp.213-270, 2012.