



FEUC FACULDADE DE ECONOMIA
UNIVERSIDADE DE COIMBRA

Métodos de Classificação Multicritério com classes parcialmente ordenadas

Maria Clara da Silva Pereira Rocha

Orientador : Professor Doutor Luís Miguel Cândido Dias

Doutoramento em Gestão - Ciência Aplicada à Decisão

Coimbra
2011

Agradecimentos

Ao Professor Doutor Luís Cândido Dias quero expressar o meu maior agradecimento, não só pela orientação e acompanhamento deste trabalho, mas também pelo entusiasmo e ânimo que sempre me incutiu ao longo do trabalho. Agradeço, ainda, a leitura atenta e sugestões realizadas na revisão do texto final.

À Fundação para a Ciência e Tecnologia agradeço o financiamento concedido através do Programa de apoio à formação avançada de docentes do Ensino Superior Politécnico (PRO-TEC) e através do projecto Classificação e avaliação de desempenho com informação QUalitativa/ ORdinal Utilizando Múltiplos critérios (QUORUM) do programa POCI 2010 (POCI/EGE/ 58371/2004).

Ao INESC de Coimbra agradeço o apoio financeiro que me permitiu a participação em conferências.

Quero, por fim, expressar a minha gratidão a todos os meus amigos e familiares que de alguma forma contribuíram e me ajudaram neste trabalho, em particular ao Aurélio pelo constante incentivo e apoio, bem como por me ter ajudado a ter o tempo e a disponibilidade necessários para a realização deste trabalho.

E a ti Beatriz, por tudo.

Resumo

Neste trabalho propusemo-nos estudar a utilização de modelos matemáticos no apoio à avaliação de um conjunto de alternativas definido em extensão, centrando-nos em metodologias para problemas de classificação multicritério. Propusemos modelos onde o conjunto de classes é parcialmente ordenado e com uma estrutura que pode ser desconhecida *a priori* ou pode ser identificada pelo decisor.

Quando o objectivo é identificar grupos (*clusters*) de alternativas com características similares, e cuja estrutura é desconhecida *a priori*, o problema é denominado *Clustering* (ou Análise de Agrupamentos). Formalmente, *clustering* com uma estrutura de classes parcialmente ordenada consiste em encontrar uma partição onde a relação de ordem parcial é definida pelas classes: uma classe pode ser melhor ou pior em comparação com outras classes, mas também pode ser incomparável a outras classes. Este é um tema inovador numa literatura onde os problemas abordados são predominantemente os de ordenação completa e *clustering* sem preferências entre classes.

Desenvolvemos dois tipos de métodos que fazem a junção do *clustering* com a classificação multicritério. No primeiro tipo desenvolvemos um modelo muito simples cujo objectivo é detectar primeiro categorias homogêneas de alternativas e formar de seguida uma estrutura de ordem parcial com essas categorias detectadas. Neste modelo, os *clusters* são detectados com base numa medida de similaridade independente das preferências do decisor. No segundo tipo de métodos, as preferências do decisor são tidas em conta, ao longo de todo o processo de decisão. Desenvolvemos para tal dois processos de optimização: um utiliza uma heurística e o outro programação matemática. A noção de similaridade entre duas classes é substituída, neste segundo tipo de métodos, por uma avaliação da qualidade da partição obtida. Nos modelos desenvolvidos, a estrutura das classes é definida à custa de relações de prevalência e a escolha da partição desejada pelo decisor pode ser feita com base na qualidade da partição, transitividade, e intuição do decisor.

Serão ainda discutidos princípios que devem ser seguidos na avaliação da qualidade de candidatos a partições e de como ajustar uma partição intransitiva.

Abstract

In this work we intend to investigate the use of mathematical models to support the evaluation of a set of alternatives defined in extension, focusing on methodologies for multicriteria classification problems. We have proposed models where the set of classes is partially ordered and with a structure that may be unknown *a priori* or can be identified by the decision maker.

When the objective is to identify groups (clusters) of alternatives with similar characteristics, and whose structure is unknown *a priori*, the problem is called *Clustering*. Formally, clustering with a structure of partially ordered classes consists in finding a partition where the partial order relation is defined by the classes: one class can be better or worse compared with other classes, but can also be incomparable to other classes. This is an innovative issue in the literature where the problems addressed are predominantly those of complete ranking and clustering without any preference among classes.

We developed two types of methods that make the fusion of clustering and multicriteria classification. In the first type we developed a very simple model whose purpose is to detect first homogeneous categories of alternatives and then form a partial order structure with these detected categories. In this model, clusters are detected based on a similarity measure independent of the decision maker's preferences. In the second type of methods, the decision maker's preferences are taken into account throughout the decision process. We developed two optimization models: one process uses an heuristic and the other uses mathematical programming. The notion of similarity between two classes is replaced, in this second type of methods for an evaluation of the quality of the obtained partition. In the developed models, the class structure is defined at the expense of outranking relations and the choice of the desired partition can be made based on the quality of the partition, transitivity, and intuition of the decision maker.

We will also discuss principles to be followed in assessing the quality of candidates to partitions and how to adjust an partition.

Índice Geral

1	Introdução	1
2	Apoio Multicritério à Decisão	13
2.1	Nota introdutória	13
2.2	Processo de decisão	16
2.3	Conceito de actor no processo de decisão	18
2.4	Conceito de acção	19
2.5	Critérios	20
2.5.1	Escalas	22
2.5.2	Critério como função e desempenhos	23
2.6	Problemáticas de ajuda à decisão	24
2.7	Relação de dominância multicritério	26
2.8	Sistema de Relações de Preferência	26
2.9	Métodos de agregação MCDA	29
3	Métodos de Classificação Multicritério	35
3.1	Nota introdutória	35
3.2	Métodos ELECTRE	39
3.2.1	ELECTRE TRI	41
3.2.2	Fases do Método	42
3.2.3	Definição de relação de prevalência	42
3.2.4	Regra de afectação	45
3.3	Métodos Clustering tradicionais	47
3.3.1	Medidas de proximidade	52
3.3.2	Algoritmos de agrupamento de dados	54
3.4	Clustering Multicritério	63
3.5	Notas conclusivas	64
4	Discussão de medidas de atribuição	67
4.1	Estrutura de Preferências das classes	67

4.2	Princípios de atribuição	69
4.2.1	STrong ($S\tau$)-Consistency (Consistência forte)	69
4.2.2	Formas de relaxamento de $S\tau$ -Consistency	72
4.3	Notas conclusivas	77
5	Método Clustering-Sorting	79
5.1	Notas introdutórias	79
5.2	Formalização	80
5.3	Método Clustering-Sorting	82
5.4	Exemplos ilustrativos com casos reais da literatura	88
5.4.1	Primeiro exemplo ilustrativo	89
5.4.2	Segundo exemplo ilustrativo	93
5.5	Notas conclusivas	98
6	CHOP - Clustering Hierárquico com Ordenação Parcial	99
6.1	Notas introdutórias	99
6.2	Estrutura de preferências das classes	100
6.2.1	Estrutura de Preferências 1 - definida à custa dos centróides das classes	100
6.2.2	Estrutura de Preferências 2 - definida à custa das relações de prevalência entre classes	101
6.3	Medida de similaridade entre classes	103
6.4	Intransitividades da partição	104
6.5	Método Aglomerativo para CHOP	106
6.5.1	Exemplo ilustrativo	108
6.6	Exemplos adicionais com casos reais da literatura	113
6.6.1	Primeiro exemplo	114
6.6.2	Segundo exemplo	116
6.7	Notas conclusivas	119
7	Optimização	121
7.1	Notas introdutórias	121
7.2	Procedimento de afectação a uma estrutura conhecida <i>versus</i> desconhecida <i>a priori</i>	122
7.3	Modelo de optimização	123

7.3.1	Formulação matemática	124
7.4	Dimensão do problema	127
7.5	Exemplo ilustrativo do PLI	129
7.6	Exemplo - comparação dos resultados do PLI com os do CHOP	133
7.7	Notas conclusivas	135
8	Estudo de um problema	139
8.1	Notas introdutórias	139
8.2	Apresentação do problema	141
8.3	Clustering-Sorting	141
8.4	CHOP	147
8.5	Decisão final	149
8.6	Conclusões	150
9	Conclusões	153
A	Formulação matemática com o princípio $S\tau$-Consistency	179
A.0.1	Formulação matemática	179

Índice de Figuras

3.1	Comparação da Classificação e Clustering, adaptado de Doumpos e Zopounidis (2002, pg.5)	37
3.2	Relações de preferência do critério g_j para o par de acções (a,b) (adaptado de Dias (2000))	43
3.3	Índice de concordância para o critério g_j (adaptado de Dias (2002))	44
3.4	Índice de discordância para o critério g_j (adaptado de Dias (2002))	45
3.5	Diferentes tipos de clustering para um mesmo grupo	49
3.6	Representação gráfica dos grupos	55
3.7	Diferentes medidas de proximidade	58
4.1	Representação do princípio $S\tau$ - Consistency.	71
4.2	Impossibilidade de cumprir $S\tau$ - Consistency	71
4.3	Representação do princípio S-Consistency: três soluções que satisfazem as condições	73
5.1	Ilustração do método Clustering-Sorting.	81
5.2	Distância entre clusters.	90
5.3	Partição final	91
5.4	Distância entre clusters.	94
5.5	Partição final obtida com o método (A)	96
5.6	Partição final obtida com o método (B)	96
5.7	Partição final obtida com o método (B)	97
6.1	Partição final	103
6.2	Estrutura de preferências de \mathcal{P}	106
6.3	Soluções possíveis para as intransitividades de \mathcal{P}_s	107
6.4	Soluções possíveis para as intransitividades de $S1$	107
6.5	Classes obtidas na partição final com Estrutura 1	110
6.6	Partição final com Estrutura 1	111
6.7	Classes obtidas na partição final com Estrutura 2	113
6.8	Partição final com Estrutura 2	113

6.9	Qualidade das partições.	115
6.10	Partição final	116
6.11	Qualidade das partições.	117
7.1	Estruturas possíveis para k categorias, determinadas com o Triângulo de Pascal	128
7.2	Estrutura pré-definida	130
8.1	Distância entre clusters e R-squared	143
8.2	Partição final pelo método (A)	145
8.3	Partição final pelo método (B)	146
8.4	Qualidade das partições.	147
8.5	Partição final do CHOP para $\lambda=0.6$	148
8.6	Pontuações médias do conflito intragrupal em cada cluster	150
8.7	Pontuações médias das estratégias de gestão de conflitos	151

Índice de Tabelas

2.1	Tabela de desempenhos	23
3.1	Algoritmo do ELECTRE TRI pessimista	46
3.2	Algoritmo do ELECTRE TRI optimista	47
3.3	Valores dos coeficientes da equação (3.9) para cada um dos algoritmos	59
4.1	Combinações proibidas por S-consistency (S) e por τ -consistency (τ).	75
5.1	Inconsistências segundo $SS\tau$ -Consistency	86
5.2	Desempenhos das vinte alternativas do sector comercial	89
5.3	Limiares de Indiferença, Preferência e Veto.	91
5.4	Grau de prevalência entre as acções (a_i, a_j), $i,j=1,\dots,20$	92
5.5	Grau de Credibilidade	93
5.6	Número de pares de alternativas para cada relação de preferência.	93
5.7	Número de pares de alternativas que verificam as condições de P, I e R antes da partição e o número de inconsistências pós-partição segundo o $SS\tau$ -Consistency	93
5.8	Limiares e pesos associados aos critérios.	95
5.9	Grau de credibilidade entre centróides.	95
5.10	Número de pares de alternativas para cada relação de preferência.	96
5.11	Número de pares de alternativas que verificam as condições de P, I e R antes da partição e o número de inconsistências pós-partição segundo o $SS\tau$ -Consistency	97
5.12	Número de pares de alternativas para cada relação de preferência.	97
6.1	Valores do grau de prevalência s_{ij} entre as classes (C^i, C^j), $i,j=1,\dots,5$	102
6.2	Grau de prevalência s_{ij} de C^i em C^j , $i,j=1,\dots,5$	106
6.3	Desempenhos das alternativas	109
6.4	Limiares e pesos dos critérios	109
6.5	Prevalências entre alternativas (a_i, a_j), $i,j=1,\dots,16$	109
6.6	Agregações das classes e Qualidade obtida com Estrutura 1	110
6.7	Valores do grau de prevalência entre os centróides	111
6.8	Agregações das classes e Qualidade obtida com Estrutura 2	112

6.9	Valores do grau de prevalência s_{ij} entre as classes (C^i, C^j) , $i, j=1, \dots, 5$	112
6.10	Agregações das classes e Qualidade obtida	114
6.11	Valores do grau de similaridade s_{ij} entre as classes (C^i, C^j) , $i, j=1, \dots, 5$	116
6.12	Comparação das incoerências das soluções obtidas com os métodos Clustering-Sorting e CHOP, segundo o $SS\tau$ -Consistency, para $k=4$	116
6.13	Qualidade das partições obtidas com o Algoritmo 2.	118
6.14	Número de pares de alternativas que verificam as condições de P, I e R antes da partição e o número de inconsistências pós-partição segundo o $SS\tau$ -Consistency	118
6.15	Qualidade das partições obtidas com o Algoritmo 2	118
7.1	Variáveis binárias que representam a situação de incoerência para cada par de alternativas	124
7.2	Dimensão de (PLI 1)	129
7.3	Dimensão de (PLI 2)	130
7.4	Dimensão de PLI com 100 alternativas e 3 classes	131
7.5	Grau de credibilidade entre alternativas.	131
7.6	Restrições da formulação matemática.	132
7.7	Qualidade obtida com o CHOP e otimização	133
7.8	Restrições da formulação matemática.	135
7.9	Qualidade obtida para as diferentes partições de $k=3$	135
8.1	Desempenhos das 77 alternativas - equipas de trabalho.	142
8.2	Limiares de Indiferença, Preferência e Veto.	144
8.3	Grau de credibilidade.	145
8.4	Número de pares de alternativas para cada relação de preferência.	146
8.5	Número de pares de alternativas que verificam as condições de P, I e R antes da partição e o número de inconsistências pós-partição segundo o $SS\tau$ -Consistency	146
8.6	Agregações das classes e Qualidade obtida	147
8.7	Valores do grau de similaridade s_{ij} entre as classes (C^i, C^j) , $i, j=1, \dots, 7$	148
8.8	Resultados para $\lambda=0.6$	148

Introdução

A Investigação Operacional (IO) tem por objectivo estabelecer bases racionais para a ajuda à decisão de problemas reais. Tal é conseguido através da compreensão e estruturação desses problemas, no sentido de prever comportamentos e melhorar os respectivos desempenhos. Para Antunes e Tavares (2000), os problemas reais são quase sempre sentidos com intensidade, mas conhecidos com dificuldade, o que para Antunes e Dias (2007) leva a que se deva considerar três paradigmas:

- adoptar uma formulação sistémica com o objectivo de representar cada problema de forma tão simples mas tão completa quanto possível;
- potenciar a interdisciplinaridade como espaço de sinergia entre diferentes ramos do conhecimento;
- evitar construir princípios e modelos universais e apostar antes numa metodologia orientada para a resolução de problemas.

Tsoukiàs (2007) defende que o que caracteriza a ajuda à decisão, de um ponto de vista científico e profissional, é o uso de linguagens formais, que reduzem a ambiguidade da comunicação humana, e o uso de linguagens que são independentes de um discurso de domínio específico.

Nos últimos quarenta anos, novas metodologias para a ajuda à decisão foram desenvolvidas por muitos investigadores e em diversas áreas. Hoje em dia, o estudo desta área faz parte de muitas disciplinas como psicologia, engenharia, investigação operacional, economia e gestão, confirmando a sua valiosa contribuição para as diversas áreas de conhecimento. Na verdade, numerosas são também as aplicações de interesse prático e de investigação dos problemas de decisão, nomeadamente no campo da medicina, gestão de recursos humanos, marketing, entre outros (Zopounidis e Doumpos, 2002). Esta particular atenção por parte da comunidade científica ao processo de ajuda à decisão deve-se a diversos factores tais como as suas fortes

bases teóricas, a relativa facilidade com que o decisor os entende e utiliza, a implementação dos seus algoritmos com uma apresentação agradável e de fácil comunicação entre as pessoas envolvidas no processo de tomada de decisão, e a possibilidade de resolução de problemas reais (Figueira et al., 2004).

Com uma sociedade cada vez mais complexa, a ajuda à decisão para problemas com critérios conflituosos torna-se particularmente importante. Na verdade, tem sido frequentemente reconhecido que há vantagens em considerar mais do que um critério na modelação de problemas, no sentido de aproximar os modelos à realidade. Note-se contudo que a relevância de uma abordagem multicritério vai para além do argumento “realista”, possuindo um valor acrescentado intrínseco no processo de modelação e da análise do modelo, apoiando a reflexão e a criatividade face a um conjunto mais vasto de potenciais soluções (Bouyssou, 1993).

Na presença de múltiplos critérios, o conceito de solução óptima dá lugar à noção de conjunto de soluções eficientes (não dominadas). Uma solução admissível diz-se eficiente se não existir uma outra qualquer solução admissível que melhore um dos critérios sem piorar, pelo menos, um dos outros. A existência de um conjunto vasto de soluções eficientes pode contudo ser um problema do ponto de vista prático o que leva a que o processo de decisão não se resuma ao cálculo de soluções eficientes. Na verdade, torna-se necessário considerar as preferências do decisor para a selecção de uma solução de compromisso final de entre o conjunto das soluções eficientes. Essa solução de compromisso pode ser, por exemplo, a identificação da alternativa preferida ou uma ordenação total ou parcial das alternativas.

O nosso trabalho enquadra-se na *ajuda* à decisão, uma vez que interessa a forma como os intervenientes de um processo de decisão podem ser ajudados a progredir de acordo com os seus conhecimentos e preferências. Por outro lado, reporta-se à ajuda *multicritério* à decisão uma vez que se foca o uso de modelos de avaliação onde se explicitam múltiplos critérios. O objectivo da ajuda à decisão multicritério (MCDA- Multicriteria Decision Aiding) é apresentar uma variedade de ferramentas e metodologias que possam ajudar os intervenientes do processo de decisão a fazer uma avaliação de um conjunto finito de alternativas definidas à custa de dois ou mais critérios, os quais estão geralmente em conflito (Belton e Stewart, 2002; Bouyssou et al., 2000; Doumpos e Zopounidis, 2002).

Podemos distinguir dois grandes grupos de problemas multicritério a que se associam as correspondentes abordagens metodológicas: a análise multicritério sobre um conjunto finito de soluções explicitamente conhecido - que se designa habitualmente por análise multicritério

- e a programação matemática multicritério (ou multiobjectivo) em que o conjunto das soluções admissíveis é definido implicitamente por restrições analíticas. O nosso estudo reporta-se principalmente à análise multicritério embora numa segunda parte do trabalho se utilize programação matemática multiobjectivo.

O primeiro e mais importante passo no estudo de problemas de decisão multicritério é a identificação da problemática, a qual foi introduzida em MCDA por Roy (1996: 57):

Given the set A of potential actions, the analyst must now determine in what terms he will pose the problem. What types of results does he envision and how does he see himself fitting into the decision process to aid in arriving at these results? Towards what will he direct his investigation? What form does he foresee his recommendation taking? ...We use the word problematic to describe the analyst's conception of the way he envisions the aid he will supply in the problem at hand based on answers to these questions.

Roy (1996) propôs quatro problemáticas gerais de referência em MCDA: Problemática de Selecção, Problemática de Descrição, Problemática de Classificação (ou de Afectação) e Problemática de Ordenação. Por vezes, pretende-se simplesmente descrever as acções possíveis e as consequências que a escolha de cada uma delas acarreta (*descrição*). Outras vezes procura-se classificar cada uma de acordo com um conjunto de categorias pré-definido (*afectação*). Finalmente, o objectivo poderá ser ordenar as diferentes acções por ordem de preferência, sendo necessário compará-las entre si (*ordenação*).

A problemática de classificação refere-se, geralmente, à *avaliação absoluta* das acções, confrontando-as com padrões pré-definidos. Tais padrões podem ser outras acções (designadas *acções de referência*), que podem ser elementos-tipo de uma categoria ou elementos de fronteira entre categorias. As problemáticas de selecção e ordenação referem-se, geralmente, à *avaliação relativa* das acções, confrontando-as entre si. Enquanto que na avaliação absoluta se utilizam conceitos como “mau” e “bom”, na avaliação relativa utilizam-se conceitos como “pior” e “melhor”. Assim, pode acontecer que a melhor acção seja má ou que a pior acção seja boa. O nosso trabalho centra-se na Problemática de Classificação mas com avaliação relativa das acções.

Existem diferentes formas de resolução de problemas de classificação nas diferentes áreas de investigação tal como investigação operacional, estatística e *data mining*. A selecção do método a utilizar está fortemente relacionada com a informação inicial disponível e com o con-

junto de objectivos que se pretendem alcançar nos resultados da classificação. São dois os tipos habituais de problemas de classificação multicritério que diferem na forma como as classes estão definidas: *classificação ordinal* e *classificação nominal*. Nos problemas de classificação ordinal as diferentes classes estão ordenadas segundo uma ordem graduada, começando por aquelas que incluem as acções mais preferidas até às que incluem as acções menos preferidas¹. Nos problemas de classificação nominal as classes não estão ordenadas. Por outro lado, nos problemas de classificação ordinal as acções não são apenas definidas de forma descritiva pelos critérios (como acontece nos problemas de classificação nominal) mas também com informação preferencial. Por exemplo, num diagnóstico médico, a doença do paciente pode ser classificada de acordo com os seus sintomas (as diferentes doenças conhecidas são as classes e os sintomas são os critérios). No entanto, não se pode fazer uma classificação ordinal pois não se consegue estabelecer uma ordem de preferência entre as doenças. Por outro lado, a avaliação de projectos é um problema ordinal, onde as classes poderiam ser, por exemplo, C^1 ="rejeitar", C^2 ="averiguar", C^3 ="aceitar" (C^1 é a pior classe e C^3 a melhor). Quando a classificação é nominal, geralmente as classes são definidas através de protótipos e as acções a classificar são atribuídas de acordo com a similaridade com esses protótipos. Quando a classificação é ordinal, podem-se também definir as classes à custa de protótipos, mas em geral são definidas à custa de limites (fronteiras) que separam as classes.

Na resolução de diversos problemas práticos uma outra técnica bastante referenciada é a Análise de agrupamentos (ou *Clustering Analysis*). Tal como a classificação em MCDA segundo B. Roy, a análise de agrupamentos pode enquadrar-se na *ajuda* multicritério à decisão e o seu objectivo é agrupar um conjunto de acções, avaliadas segundo um conjunto de atributos, em *grupos*. A principal forma de distinguir estas duas técnicas reside na definição dos grupos: na classificação multicritério os grupos são definidos *a priori* e no segundo caso os grupos (geralmente denominados clusters) são definidos *a posteriori*, estando geralmente relacionados com preocupações descritivas.

Seja com o objectivo de simplificação, exploração ou confirmação, a análise de clusters é uma poderosa ferramenta analítica que tem uma vasta gama de aplicações e em diversas áreas. Desde a obtenção de taxonomias em biologia para agregar os organismos vivos, de classificações psicológicas baseadas em traços de personalidade e outros, à análise de segmentação de administradores, a análise de clusters sempre teve uma forte tradição de agrupar indivíduos.

¹ Geralmente assume-se que a ordem representa uma relação de preferência, porém, pode tratar-se de outra relação transitiva e assimétrica, qualquer que seja o seu significado semântico

Essa tradição foi estendida à classificação de acções, incluindo a estrutura de mercado, a análise de semelhanças e diferenças entre novos produtos e avaliações de desempenho de empresas para identificar clusters com base nas estratégias de empresas (Jain et al., 1999). O resultado foi uma explosão de aplicações em diversas áreas de investigação gerando não só conhecimento, mas também a necessidade de uma melhor compreensão da técnica para minimizar a sua má utilização.

Tradicionalmente os algoritmos de análise de agrupamentos podem ser classificados em *hierárquicos* (envolvem a construção de uma hierarquia tipo árvore) e *não-hierárquicos* ou de particionamento. Os algoritmos hierárquicos podem ainda ser subdivididos em *aglomerativos* e *divisivos* (Jain e Dubes, 1988; Kaufman e Rousseeuw, 1990). A análise de agrupamentos aglomerativa inicia com todos os clusters com uma única acção e recursivamente vai juntando dois ou mais clusters de forma apropriada. Quanto à análise de agrupamentos divisiva, esta inicia com um único cluster com todas as acções em estudo e vai particionando recursivamente os clusters, parando quando verificar um critério de paragem (frequentemente é o número k de clusters fixado).

Nos algoritmos não-hierárquicos pode-se tentar identificar os clusters fazendo mudanças das acções de forma iterativa entre os clusters ou tentando identificar os clusters através de áreas altamente densas da base de dados, separando regiões de objectos de alta e baixa densidade (Han et al, 2001).

Geralmente, num problema clustering devem-se seguir os seguintes passos (Doreian et al., 2005):

1. encontrar o conjunto das acções definidas à custa de critérios, os quais podem ser normalizados;
2. definir uma medida apropriada de similaridade entre as acções e muitas das vezes também entre acções e clusters e até entre clusters;
3. escolher o algoritmo de agrupamento mais adequado;
4. avaliar a(s) solução(ões) obtida(s) para ver se têm uma estrutura sustentável.

O conceito de *similaridade* é fundamental na análise de agrupamentos. A similaridade de acções é uma medida de correspondência ou semelhança entre acções a serem agrupadas. Aqui as características de similaridade são em primeiro lugar especificadas, sendo que cada um dos

diferentes métodos utilizados para medir a similaridade representa uma perspectiva particular de similaridade, dependendo dos objectivos, do tipo de dados e do problema em estudo.

Como referido anteriormente, muitos são os trabalhos de classificação e clustering multi-critério que nos últimos anos a comunidade científica tem desenvolvido, e embora estas duas formas de resolução de problemas sejam bastante diferentes à primeira vista, elas podem-se complementar. Foi esta a motivação que levou recentemente ao desenvolvimento de alguns trabalhos de investigação tendo por objectivo a junção das duas diferentes abordagens (De Smet e Montano, 2004; Figueira et al., 2004; Nemery e De Smet, 2005; Fernandez et al., 2010). Tais trabalhos têm por objectivo não só detectar grupos de acções similares como também as relações de preferência existentes nos grupos encontrados. Não é só no detectar das relações de preferência entre grupos que reside a originalidade destes trabalhos, mas também na definição das medidas de similaridade que se baseiam nas relações binárias de preferência entre as acções a agrupar.

Dado que partilhamos o reconhecimento da pertinência do estudo da junção das duas abordagens, classificação MCDA e análise de agrupamentos, dadas as suas potenciais aplicações e constatada a relativamente escassa atenção que esta problemática tem tido em termos de investigação, pretendemos com esta dissertação contribuir para encontrar o melhor uso desta junção de técnicas desenvolvendo para tal novas abordagens metodológicas, tendo em vista apoiar eficazmente o decisor envolvido no problema. Assumimos, neste trabalho, a existência de um só decisor.

Os trabalhos desenvolvidos até agora apenas abordam a atribuição de acções a grupos totalmente ordenados. Contudo, diversos são os problemas reais onde muitas das acções a classificar são incomparáveis. Por exemplo, no diagnóstico de hiperactividade em crianças com pouca atenção (TDAH-Transtorno do Deficit de Atenção com Hiperactividade) (Tannock, 1998), podemos encontrar quatro grupos homogéneos: A) crianças sem sintomas relevantes, B) crianças com TDAH com predominância do tipo desatento, C) crianças com TDAH com tipo predominantemente hiperativo-impulsivo, e D) TDAH do tipo combinado (exibindo ambos os tipos de sintomas). Aqui, apesar de ser possível considerar que A é preferível a B e C, e os últimos serem preferíveis a D, pode-se considerar que as categorias B e C não são comparáveis.

Nestes casos, a estrutura pretendida é a de uma ordenação parcial das classes: uma classe pode ser melhor ou pior que algumas classes, mas também pode ser incomparável a outras

classes.

O principal objectivo deste trabalho é o de contribuir com novas metodologias para o apoio à decisão em problemas de classificação, com base em múltiplos critérios, a um conjunto de classes parcialmente ordenado. No cumprimento deste objectivo, consideramos que a classificação ordinal constitui uma forma fraca de ordenar um conjunto de alternativas e consideramos modelos multicritério com uma estrutura de pré-ordem parcial². Tais modelos fazem a junção da classificação e do clustering multicritério, sendo três as abordagens propostas e estudadas neste trabalho e que passamos a explicitar:

- (i) **Clustering - Sorting** : a estrutura é desconhecida *a priori* e o objectivo é, em primeiro lugar, detectar categorias homogéneas de alternativas e formar de seguida uma estrutura de ordem parcial com essas categorias detectadas. Numa primeira fase precisamos detectar apenas os clusters com base numa medida de similaridade independente das preferências do decisor. Em seguida, pretende-se detectar potenciais relações de ordem parcial que possam existir entre eles. Destacamos que a eliciação de tais valores é feita somente depois de o decisor identificar os agrupamentos e considerar que essas categorias fazem sentido.
- (ii) **CHOP - Clustering Hierárquico de Ordenação Parcial** : a estrutura, tal como no primeiro método que apresentamos, é desconhecida *a priori*, mas contrariamente ao modelo anterior, pretendemos ter em consideração as preferências do decisor ao longo de todo o processo. Assim, propomos um modelo heurístico que é uma extensão do método hierárquico aglomerativo, cuja medida de similaridade entre clusters é baseada nas relações de preferência definidas pelo decisor: dois grupos são tanto mais similares quanto menor o número de incoerências das relações de Preferência, Indiferença e Incomparabilidade entre as acções classificadas, segundo o mesmo princípio de atribuição das alternativas aplicado no Clustering-Sorting. Para tal serão definidos e discutidos princípios que permitem avaliar as partições obtidas.
- (iii) **Optimização** Os métodos Clustering-Sorting e CHOP serão considerados aqui como um módulo de pesquisa inicial que antecede um método “exacto”. A intenção é dar ao decisor num primeiro momento, caso seja necessário, um conhecimento inicial do problema que o ajude a estabelecer e rever as suas preferências a um relativamente baixo custo

²Ordenações das classes com eventuais *ex-aequo* e eventual presença de classes incomparáveis.

computacional, tendo por resultado a estrutura de preferências das classes. Num segundo momento, pretende-se melhorar os resultados da partição obtida pelo Clustering-Sorting ou pelo CHOP ou, caso o decisor tenha ideia da(s) estrutura(s) das classes, obter a melhor partição. A nova abordagem proposta é tecnicamente diferente das heurísticas propostas, mas tendo por base os mesmos objectivos e exigências do decisor, aplicando-se assim o mesmo princípio de atribuição.

Mesmo conscientes das dificuldades computacionais inerentes à optimização de um problema com estrutura desconhecida *a priori*, propomos ainda um modelo exacto para problemas desta natureza.

As duas primeiras abordagens que propomos são abordagens heurísticas, uma vez que não ambicionam obter a melhor partição parcialmente ordenada de entre todas as partições possíveis. Pretendem ser abordagens simples e transparentes para o decisor, cuja assistência é solicitada para ajudar a moldar os resultados. Na verdade, modelos e métodos de apoio à decisão não têm que ser vistos como um operador que produz resultados compatíveis com os parâmetros fornecidos ou combinações destes. A utilização de modelos simples e mais rápidos, que tornam o processo de decisão mais facilmente entendido pelo decisor e que exigem juízos simples por parte deste é defendida por alguns autores como Edwards e Barron (1994). Note-se que no caso dos métodos Clustering-Sorting e CHOP, para além da simplicidade e rapidez procurada nos métodos, tínhamos consciência de que os métodos “exactos” que poderiam ser desenvolvidos teriam que recorrer à programação inteira (PI). Contudo, as dificuldades inerentes à complexidade da PI são agravadas, no nosso caso, pelo desconhecimento da estrutura de preferências dessas classes, o que pode tornar os problemas de maiores dimensões intratáveis.

Segundo as metodologias heurísticas, após obter um resultado, deve-se verificar se o erro potencial é grande para o problema que está a ser estudado. Neste trabalho, estudaremos para tal indicadores da partição para formalizar a noção de “compatível” com a informação dada pelo decisor. As violações das condições de preferência, indiferença e incomparabilidade são então analisadas considerando as relações inter-classes e intra-classes de todos os pares de acções, segundo os princípios que serão definidos.

A classificação ordenada das acções é dependente da forma como o decisor constrói as suas preferências (métodos de prevalência, teoria da utilidade, AHP, etc). Na verdade, quando se aplicam métodos MCDA, os perfis das acções podem ficar restringidos apenas às relações de preferência e de indiferença, como por exemplo na teoria da utilidade, ou podem ainda

considerar a relação de incomparabilidade, como é o caso dos métodos de prevalência. Neste trabalho os perfis das acções serão definidos à custa das relações de preferência, indiferença e de incomparabilidade, e como tal serão usados métodos de prevalência no cálculo das medidas de similaridade e dos indicadores das partições.

Para ilustrar a metodologia proposta são utilizados dados de dois exemplos reais descritos na literatura. Para ilustração dos vários métodos apresentamos ainda um caso real no Capítulo 8. As ferramentas utilizadas são a aplicação SPSS (*Statistical Package for Social Sciences*) versão 17.0 e a aplicação *MatLab 7.0*.

Concluimos este capítulo introdutório com uma breve descrição da forma como se encontra organizada esta dissertação.

No *Capítulo 1 (Introdução)* começa-se por inserir este trabalho na ajuda multicritério à decisão. Em particular, este trabalho visa integrar a classificação multicritério e o clustering por forma a classificar acções em grupos desconhecidos a priori, bem como definir a sua estrutura de relações que possa reflectir as preferências do decisor.

No *Capítulo 2 (Apoio Multicritério à Decisão)* são introduzidos e definidos alguns conceitos preliminares que serão utilizados no resto do trabalho, nomeadamente o processo de decisão, actor no processo de decisão, acção, critérios, problemáticas de apoio à decisão, relação de dominância e sistemas de relações de preferência. No final do capítulo far-se-à uma breve referência a métodos de agregação.

No *Capítulo 3 (Métodos de Classificação)* começamos por fazer uma breve introdução onde destacamos as principais diferenças entre duas formas distintas de classificação multicritério, a classificação MCDA e o *clustering*. De seguida são apresentados os métodos mais relevantes neste trabalho destes dois tipos de classificação: o método ELECTRE TRI da família dos métodos ELECTRE pertencente à classificação MCDA e os modelos Hierárquico e de Partição pertencentes ao *clustering*.

No *Capítulo 4 (Discussão de medidas de atribuição para Problemas de Classificação)* são apresentados e discutidos diferentes princípios de atribuição de um conjunto de alternativas,

avaliadas por múltiplos critérios, a classes parcialmente ordenadas. Tais princípios são baseados nas relações de prevalência das alternativas a serem atribuídas.

No *Capítulo 5 (Classificação Multicritério com classes ordinais e nominais: Método Clustering-Sorting)* propomos uma nova abordagem heurística para classificar um conjunto de alternativas descritas por um conjunto de múltiplos critérios, conjugando para tal, a classificação MCDA e o *clustering*. Numa primeira fase detectam-se grupos homogêneos e numa segunda fase detectam-se as relações de preferência existentes entre os grupos com base nas relações de prevalência das alternativas e num dos princípios propostos no capítulo 4.

No *Capítulo 6 (CHOP - Clustering Hierárquico com Ordenação Parcial)* é proposto um modelo heurístico para problemas de classificação que é uma extensão do método hierárquico aglomerativo. Propomos uma medida de similaridade entre clusters baseada na qualidade da partição: dois grupos são tanto mais similares quanto menor for o número de violações das preferências do decisor de acordo com um dos princípios sugeridos no capítulo 4.

No *Capítulo 7 (Classificação Multicritério com classes ordinais e nominais: modelo de Optimização)* propomos um método exacto para problemas de classificação multicritério, onde o conjunto de categorias a que se pretende atribuir as alternativas é parcialmente ordenado e com ou sem uma estrutura definida *a priori*. Tal estrutura pode ser fornecida pelo decisor ou, caso não tenha qualquer conhecimento prévio sobre o número de classes a escolher nem das suas relações de preferência, pode ser aplicado previamente a heurística CHOP proposta no capítulo anterior. O algoritmo de optimização tem por objectivo obter a melhor partição definida com base nas relações de prevalência das alternativas e num dos princípios propostos no capítulo 4.

No *Capítulo 8 (Classificação Multicritério de grupos de trabalho a classes desconhecidas)*, é descrito um caso de estudo referente à aplicação das metodologias desenvolvidas nos capítulos anteriores, a um conjunto de dados com o tipo e estratégias de gestão de conflitos intra-grupos de grupos de trabalho do departamento de produção de empresas industriais. O intuito é a criação e caracterização dos perfis de funcionamento dos grupos de trabalho, bem como a detecção de potenciais relações que possam existir entre os *clusters* detectados.

No *Capítulo 9 (Conclusões)* procuramos apresentar as principais conclusões deste trabalho e algumas vias de investigação futura.

Apoio Multicritério à Decisão

2.1 Nota introdutória

A tomada de decisão sempre atraiu os pensadores de várias épocas. Na verdade, tomar decisões é uma actividade que todas as pessoas fazem constantemente, desde a decisão mais simples («O que devo almoçar?») à mais complexa («Que tipo investimentos fazer?») e a todos os níveis: individual («Onde devo passar as férias?»), organizacional («Como vamos fazer os horários da Escola?») e inter-organizacional («Como fazer a mobilidade de docentes?»). Diversas vezes, durante esses processos de decisão solicita-se ajuda, aconselhamento ou apoio aos amigos, peritos, empresas consultoras, etc. (Tsoukiàs, 2007)

A ajuda para obter respostas às questões que se colocam aos intervenientes de um processo de decisão é denominada *ajuda à decisão* que, segundo Roy (1996), é uma actividade baseada em modelos claramente explicitados mas não necessariamente e completamente formalizados, e que normalmente recomenda ou simplesmente favorece um comportamento que articule a evolução, os objectivos e os valores envolvidos no processo de decisão. Assim, podemos concluir que o objectivo da ajuda à decisão não é resolver problemas, mas sim ajudar a obter respostas.

Bana e Costa (1986) define *problema de decisão* como um problema em que, face a um conjunto de objectivos, há a considerar um conjunto de possíveis soluções, alternativas, programas - acções potenciais - explícita ou implicitamente definidas. Na verdade, tal como defendido por Keeney e Raiffa (1976), considerando todos os objectivos, raramente alguma opção de acção é mais atractiva ou preferível às demais.

A resolução de um problema de decisão passa assim pela escolha/elaboração de um *modelo* que tenha em conta os parâmetros que concorrem para o mesmo e que, na medida do possível, abarque todos os aspectos relevantes e equacionáveis do mesmo. Roy (1985) define *modelo* como um esquema que, para um conjunto de questões, é tomado como a representação de uma classe de fenómenos, mais ou menos habilmente decompostos do seu contexto por um

observador, para servir de suporte à investigação e/ou à comunicação.

Grandes são as vantagens em considerar mais do que um critério na modelação de problemas, uma vez que apoia a reflexão e a criatividade face a um conjunto mais vasto de possíveis soluções, além de aproximar os modelos à realidade que geralmente é complexa. Segundo Gomes et al. (2002), os problemas complexos, de um modo geral, possuem pelo menos algumas das seguintes características:

- para a resolução do problema, existem no mínimo dois critérios os quais são conflituosos;
- tanto os critérios como as alternativas não são claramente definidos (há uma fronteira difusa) e as consequências da escolha de determinada alternativa, com relação a pelo menos um critério, podem não ser claramente compreendidas, havendo muitas vezes relações de dependência;
- os critérios podem estar interligados, de tal forma que um critério parece reflectir parcialmente outro critério. O mesmo se pode passar com as alternativas quando estas não são mutuamente exclusivas: a eficácia da escolha de uma alternativa pode depender de outra alternativa ter sido, ou não, também escolhida;
- a solução do problema pode envolver um conjunto de pessoas, cada uma das quais tendo o seu próprio ponto de vista, muitas vezes conflituosos com os restantes;
- as restrições do problema não são bem definidas, podendo mesmo haver alguma dúvida a respeito do que é critério e do que é restrição;
- nem todos os critérios são quantificáveis, sendo que alguns só o são por meio de julgamentos de valor efectuados usando uma escala;
- a escala para um dado critério depende dos dados disponíveis e da sua própria natureza;

Várias outras complicações podem surgir num problema real de tomada de decisão, mas estes sete aspectos anteriores caracterizam a complexidade de tal problema. Winterfeldt e Edwards (1986) defendem que os problemas complexos devem ser desagregados em vários sub-problemas, o que facilita a resolução dos mesmos. A aplicação de várias técnicas a cada uma das partes, consoante a problemática, possibilita a resolução do problema como um todo.

Em muitos casos da actividade de ajuda à decisão, especialmente quando os problemas são complexos, será mais transparente e menos arbitrário recorrer a múltiplos critérios de avaliação, a que se denomina *ajuda multicritério à decisão*¹. Roy e Bouyssou (1993) defendem que a abordagem multicritério favorece:

- a constituição de uma base para o diálogo entre os intervenientes utilizando diversos pontos de vista compreendidos por todos;
- uma modelação mais fácil da incerteza, da arbitrariedade ou da imprecisão associada aos dados, segundo cada ponto de vista e, sobretudo
- o encarar de cada decisão como um compromisso entre objectivos em conflito.

Roy (1996) salienta ainda que a ajuda multicritério à decisão visa construir modelos que façam intervir de forma explícita múltiplos critérios de avaliação, constantemente sujeitos à análise crítica dos intervenientes no processo de decisão, capazes de ajudar cada interveniente a moldar as suas preferências e a encontrar argumentos para sustentar em conformidade os seus múltiplos objectivos (frequentemente conflituosos). Desta forma, a metodologia da ajuda multicritério à decisão não pretende constituir um método capaz de fornecer a “decisão (acção) óptima”, já que são raras as situações em que uma única acção se apresenta como a melhor segundo todos os pontos de vista dos decisores. No contexto multicriterio, os investigadores consideram *soluções de compromisso*, ou seja, soluções que são “*globalmente boas*” de acordo com os diferentes critérios (não sendo necessariamente a melhor em todos os critérios) e não são muito más em nenhum dos critérios.

Este capítulo tem como objectivo fazer um enquadramento da ajuda multicritério à decisão, nomeadamente fazer uma breve revisão de conceitos preliminares que serão utilizados ao longo deste trabalho. Assim, nas duas secções seguintes falaremos do processo de decisão e de modelos de ajuda à decisão. De seguida são apresentados alguns conceitos de ajuda multicritério à decisão. No final do capítulo far-se-à um breve inventário de alguns métodos de agregação, sendo desenvolvidos um pouco aqueles mais relevantes para a presente dissertação no Capítulo 3.

¹Designada por MCDA - Multicriteria Decision Aid, ou por MCDM - Multiple Criteria Decision-Making, na literatura Anglo-saxónica e por Analyse Multicritère ou Aide Multicritère à la Décision em francês. Na literatura portuguesa é, também comum, ser designada por ADCM - Ajuda à Decisão com Critérios Múltiplos.

2.2 Processo de decisão

Por *decisão* entende-se em geral qualquer processo de que resulte uma escolha², cuja implementação possa conduzir à obtenção de determinado(s) objectivo(s), definido(s) à custa de um ou vários critérios. Neste sentido, Roy e Bouyssou (1993) definem *processo de decisão* (ou processo de tomada de decisão) como um mecanismo interactivo de aprendizagem resultante do desenvolvimento e evolução das confrontações e interacções entre os diferentes intervenientes, regulado por mecanismos próprios do domínio do problema em causa.

Quanto aos propósitos do processo de decisão, Winterfeldt e Edwards (1986) resumem-nos em cinco:

1. identificação de opções ou reformulação de opções já existentes;
2. definição de objectivos a serem alcançados. Segundo Roy (1985), este ponto é particularmente difícil devido ao conflito de objectivos dentro de um sistema/ contexto e das relações que nele existem;
3. fornecimento de uma linguagem comum para comunicação e consenso de opiniões;
4. especificação de critérios/ atributos e quantificação de variáveis subjectivas;
5. desenvolvimento de índices de valores relevantes.

Na caracterização dos processos de decisão, Dias (2000) distingue dois tipos de problemas: *problemas estruturados e não estruturados*. Para os primeiros existem dados, procedimentos de resolução e resultados completamente especificados, por vezes à custa de regras bem definidas. São problemas geralmente de rotina, como por exemplo definir o escalão a que um aluno pertence. Os problemas não estruturados são exactamente o oposto, como por exemplo escolher a local de funcionamento duma nova empresa. Há ainda referências a problemas *semi-estruturados*, por se encontrarem entre estes dois pólos opostos ou por terem fases onde são estruturados e outras onde são não estruturados (ver Silver, 1991:67).

O processo de decisão, em muitos casos, consiste no esforço para resolver o dilema de critérios ou objectivos conflituosos, conduzindo, por vezes, à procura de uma solução de compromisso e não à procura da óptima (Zeleny, 1982; Lai et al., 1994).³ Esta ideia é reforçada

²Mais adiante ver-se-à outros tipos de problemáticas de decisão além da problemática da escolha

³Tal pode ocorrer pelo facto de não haver informação suficiente, por exemplo, por questões de tempo e/ou falta de recursos.

por Vincke (1992) e por Guitouni e Martel (1998), segundo os quais, é bem aceite nos nossos dias que o processo de tomada de decisão seja mais amplo que o modelo clássico que o caracterizava: optimização de uma única função objectivo para um conjunto de soluções realizáveis. De facto, muitos aspectos conflituosos devem ser equacionados ao mesmo tempo e, assim, a decisão passa a ser multicritério, deixando por vezes de ser óptima para ser satisfatória.

A análise de decisão multicritério pode ser vista como um processo recursivo, construído, geralmente, em quatro fases⁴(Guimarães, 1979; Roy, 1985; Guitouni e Martel, 1998):

1. Estruturação do problema de tomada de decisão;
2. Articulação e modelação de preferências;
3. Agregação das avaliações das alternativas (preferências) e;
4. Elaboração de recomendações.

Também Simon (1977) define quatro fases no processo de decisão, mas da seguinte forma: *inteligência, estruturação, avaliação e revisão*. A fase de *inteligência* é aquela onde se detecta a necessidade de decidir, resultante quer de uma procura activa de situações em que seja necessário decidir, quer da confrontação de um problema que surge. A fase de *estruturação* (ou concepção) compreende a modelação da realidade na sua complexidade não estruturada, ao passo que a fase de *avaliação* (ou escolha) considera o modelo construído na fase de estruturação como o ponto de partida para chegar a uma recomendação final. Assim, a avaliação corresponde à tomada de decisão propriamente dita, frequentemente com o apoio de técnicas matemáticas, algoritmos ou regras de carácter heurístico. Aqui, testam-se ou simulam-se possíveis acções de forma a cumprir o(s) objectivo(s) do processo, de acordo com os critérios previamente definidos. Nesta fase, de acordo com o objectivo do processo de decisão, poder-se-á, por exemplo, escolher a melhor alternativa (ou curso de acção) ou então ordenar as alternativas. Finalmente, a fase de *revisão* consiste na implementação da decisão tomada e no acompanhamento das suas consequências. É nesta fase que se podem detectar incorrecções sendo uma fase em que se pode aprender com esses eventuais erros.

Para Dias (2000), nesta sequência, o processo pode-se considerar *dinâmico*, no sentido em que se pode voltar às etapas anteriores. Na verdade, o modelo pode ser revisto sempre que surjam dados relevantes que possam redefinir ou até contrariar os seus pressupostos. Isto

⁴Segundo Bana e Costa (1992), a metodologia para a implementação de um modelo multicritério pode resumir-se a três fases: estruturação; avaliação; elaboração de recomendações.

conduz a um aprofundamento do conhecimento da realidade e uma eventual reformulação dos critérios de escolha. Por outro lado, pode ser desenvolvido, em paralelo, trabalho relativo a etapas distintas.

A *fase de estruturação* do problema de tomada de decisão é considerada, por vários autores, como a fase mais importante e crucial de todo processo decisional (Ackoff, 1978; Roy, 1985; Winterfeldt e Edwards, 1986; Checkland e Scholes, 1990; Keeney, 1992; Checkland e Holwell, 1993; Bana e Costa, 1993; Brans, 1994; Belton e Ackermann, 1997; Guitouni e Martel, 1998). Winterfeldt e Edwards (1986: 26) referem que a estruturação deve ser vista numa perspectiva mais alargada, que se traduz num conjunto de etapas que pretende dar resposta a um conjunto largo de questões, como exemplo, na identificação do problema (Qual a natureza do problema e qual o seu ambiente? Qual o espaço temporal da decisão? Quem são os decisores? Quais os valores dos decisores? Quem conhece as consequências da decisão? Quais são as classes genéricas de opções? Quais os grupos afectados pela decisão? Qual o propósito da análise? Etc.) Estas questões são discutidas em maior detalhe em Zeleny (1982). Silver (1991: 32) define a estruturação como uma actividade onde se decide como decidir (Quais os dados? Que modelos aplicar?). Embora esta seja uma fase bastante importante, não nos alongaremos na sua descrição dado o âmbito do nosso trabalho.

De seguida serão apresentados em maior detalhe alguns elementos das fases de estruturação e de avaliação, nomeadamente os conceitos de actor, acção, critério, relações de preferência e problemáticas de ajuda à decisão, os quais são cruciais nestas fases. Note-se que estas duas fases podem-se entrelaçar uma vez que na fase de estruturação podem-se fazer, implicitamente, algumas avaliações, e na fase de avaliação pode-se sentir a necessidade de acrescentar ou mesmo alterar algo à estruturação. Por exemplo, na fase de estruturação podem-se ignorar acções que são irrealistas ou até sem grande interesse para o problema em estudo e na fase de avaliação pode-se descobrir que o conjunto de critérios definidos na fase de estruturação afinal não considerava um aspecto tido como importante.

2.3 Conceito de actor no processo de decisão

O processo de decisão pode ser condicionado por *actores*, quer de forma directa, pelo facto de intervirem no processo de decisão, quer de forma indirecta, pelo facto das suas opiniões serem tidas em conta por aqueles que intervêm directamente no processo de decisão.

O *decisor* terá como função coordenar os actores que intervêm de forma directa, *interve-*

nientes ou *participantes*, e possui poder para resolver eventuais divergências entre os mesmos pela imposição do seu julgamento. Ao nível da proposta de uma recomendação, o decisor é quem indica objectivos, afere o que é ou não exequível e expressa preferências que podem ter em consideração as opiniões dos restantes actores. Ao nível de colocar em prática uma recomendação, o decisor nem sempre deterá esse poder, nomeadamente nos casos em que a recomendação do processo de decisão tenha que ser ratificada por outrem.(Dias, 2002)

Mas a tomada de decisão nem sempre é feita de forma individual, sendo cada vez mais frequentes as situações de tomada de *decisão em grupo*. Nestas situações, existem múltiplos decisores que coordenam conjuntamente as actividades, agindo com base no resultado de votações ou através da formação de consensos. Em outros casos temos situações de negociação, onde os intervenientes representam partes com objectivos diferentes e notoriamente divergentes, podendo a negociação, a arbitragem e a mediação ser utilizadas para definir um acordo entre as partes.(Dias, 2002)

O especialista em ajuda à decisão, *analista*, também contribuirá para o avanço do processo de decisão, podendo ser um interveniente ou mesmo o decisor, ou então um consultor externo a esse sub-sistema.

O pouco contacto directo entre o analista e o decisor levam Roy (1985) a considerar ainda um outro actor do processo de decisão denominado *cliente*. É o cliente que requer o estudo e é responsável pela alocação dos meios necessários para o conduzir, fazendo ainda a ligação entre o analista e o decisor.

2.4 Conceito de acção

Intuitivamente, as acções ou alternativas são os projectos de investimento, planos de produção, planos de desenvolvimento, candidatos, entre outros, sobre os quais recai a decisão. Mais formalmente,

Definição 2.4.1 (Roy, 2000) *Uma acção é um termo genérico usado para designar o que constitui o objecto de decisão, ou para o que a tomada de decisão é dirigida.*

O conjunto das acções, geralmente representado por \mathcal{A} , pode ser:

- *estável* caso possa ser definido a priori e não possa sofrer alterações ao longo do processo de decisão;

- *evolutivo* se, pelo contrário, poder sofrer alterações ao longo do processo de decisão: as características de algumas acções podem ficar alteradas ao longo do tempo; algumas acções podem ser inventadas durante o processo de decisão; algumas acções podem deixar de estar disponíveis, ou deixar de ser consideradas interessantes; algumas acções podem passar a estar disponíveis, ou passar a ser consideradas interessantes; o âmbito do processo de decisão pode ser reformulado (Dias,2002);
- definido em *extensão* - o cardinal de \mathcal{A} pode ser finito e relativamente pequeno e, como consequência, os seus elementos são explicitamente enumerados;
- definido em *compreensão* - o cardinal de \mathcal{A} pode ser infinito ou relativamente elevado e, como consequência, \mathcal{A} pode ser definido por um conjunto de restrições a que obedecem os seus elementos.

Segundo Roy (1999), as acções são representações que não implicam, *a priori*, uma existência real. Distinguem-se assim *acções reais*, se puderem ser uma decisão final, de *acções fictícias*, se puderem estar presentes ao longo do processo de decisão mas não tiverem uma existência real. Estas últimas podem ainda ser divididas em *realistas* e *irrealistas*, consoante se anteveja ou não a viabilidade da sua existência. As acções podem ainda ser *globalizadas* caso excluam qualquer uma das outras, ou *fragmentadas*, caso seja necessário fazer combinações, atendendo a eventuais interdependências entre elas, para responder a situações de decisão⁵. Utiliza-se ainda a designação de *acções potenciais* para designar acções reais ou fictícias mas realistas.

Assim que os elementos de \mathcal{A} sejam determinados, estes devem ser caracterizados de acordo com os diferentes pontos de vista, geralmente contraditórios. Tal é formalizado pela noção de critério apresentada na secção seguinte.

2.5 Critérios

Um dos aspectos difíceis de estruturar é o dos valores dos decisores, que reflectem não só os seus interesses e preferências, como os valores subjacentes à cultura da organização. Keeney (1992: 6) refere-se aos “objectivos” do decisor como a forma de explicitar os valores, valores esses que define como princípios usados na avaliação das acções e das suas consequências. Já

⁵Os conceitos de acções reais vs. fictícias, realistas vs. irrealistas e globais vs. fragmentárias são os apresentados por Roy (ver Roy,1985; Roy e Bouyssou, 1993; Vincke, 1992)

Bana e Costa (1992) refere-se a *pontos de vista* em vez de objectivos, os quais interligam as características das acções a avaliar e os valores dos decisores.

A construção de um modelo para avaliar alternativas envolve a procura de um conjunto de *pontos de vista fundamentais*, caracterizados por serem aceitáveis por todos os actores e por serem isoláveis entre si (ou seja, a avaliação das alternativas segundo um ponto de vista fundamental não depende de nenhum outro). Tal conjunto de pontos de vista originará o conjunto de critérios a usar na avaliação das acções. (Roy, 1996) define *critério* como um modelo que contém a informação necessária para constituir uma base de avaliação relativa (determinar se uma acção é melhor ou pior que outra) ou absoluta (determinar se uma acção é melhor ou pior que algum standard ou padrão).

Segundo Keeney (1992), o conjunto de critérios deve obedecer às seguintes propriedades: deve indicar apenas as consequências relevantes para o processo de decisão em questão (*essencial*), deve referir-se às consequências da escolha feita no conjunto das acções no contexto da decisão (*controlável*), deve evitar a duplicação de consequências (*não redundante*), deve incluir todos os aspectos fundamentais das várias acções (*completo*), deve definir os objectivos a atingir e especificar os meios para o conseguir (*mensurável*), deve fornecer a informação requerida para uma análise consistente, considerando o tempo e o esforço disponíveis (*operacional*), deve conter o número mínimo de objectivos necessários (*conciso*), deve permitir a separação dos diferentes objectivos em análise (*decomponível*) e deve facilitar a geração e comunicação dos critérios (*compreensível*).

Bana e Costa et al. (2000) reforça que a *isolabilidade* (independência) é crucial para a avaliação multicritério. Esta é a razão pela qual, muitas vezes, diversos pontos de vista, identificados como “fins”, têm que ser agregados num mesmo critério de avaliação. A definição de diferentes critérios que, afinal, representam um mesmo ponto de vista fundamental, introduz no modelo redundância, tendo como consequência a sobrevaloração do que deveria ser um único critério. Por outro lado, a exigência de que a família de critérios seja exhaustiva (ou completa) pode, de certo modo, funcionar contra a sua concisão e a isolabilidade de cada critério. O número de critérios não deve ser reduzido ao ponto de se correr o risco de deixar de fora alguns aspectos fundamentais de avaliação (Bana e Costa et al., 2000: 344).

Para se obter uma *família coerente* de critérios tem que se ter em conta as seguintes exigências respeitantes à interligação dos critérios (Roy, 1996): exhaustividade, coesão e não redundância.

- *Exaustividade*, por se garantir que todos os critérios para a tomada de decisão estejam presentes; caso duas acções sejam iguais em todos os critérios mas não se possa concluir que elas são indiferentes, então determinado elemento de apreciação não foi considerado, ou seja, a família de critérios não é exaustiva.
- *Coesão*, porque mantendo-se todo o resto igual, uma melhoria segundo um critério é sempre considerada desejável; considerando, por exemplo, duas acções *a* e *b* que são iguais em todos os critérios menos num, no qual *a* é melhor, então pode-se afirmar que *a* não é pior que *b*.
- *Não redundância*, por se evitar a duplicação dos critérios; o conjunto de critérios deve ser tal que a eliminação de um deles leve à não verificação de pelo menos uma das exigências anteriores.

2.5.1 Escalas

A construção de critérios implica a definição de uma escala de níveis de desempenho a atribuir às acções. A escala pode ser definida segundo quatro tipos de medidas ou escalas de medidas: escala nominal, escala ordinal, escala de intervalos e escala de razão ou proporção (rácio). A *escala nominal* representa o nível de medida mais baixo. Uma escala diz-se nominal quando se usam números apenas para classificar os indivíduos numa modalidade qualitativa precisa (Ex. O grupo sanguíneo: 1- tipo O; 2- tipo A; 3- tipo B; 4- tipo AB). A seguir tem-se a *escala ordinal* que difere da escala nominal pelo facto de permitir ordenar as diferentes modalidades segundo uma ordem graduada. Os valores numéricos que sejam utilizados para designar as categorias são arbitrários, mas devem respeitar uma ordem lógica que permita situar os sujeitos entre si. (Ex. Classe social: 1- baixa; 2- média; 3- alta). A *escala de intervalos* difere da escala ordinal pelo facto de que, além da ordenação das modalidades de um atributo, pode-se dizer quanto valem exactamente as diferenças entre estas modalidades. Nesta escala, o zero é um valor arbitrário que não corresponde à ausência da modalidade. (Ex. Pode-se dizer que a diferença entre 30°C e 10°C é exactamente de 20°C e é a mesma entre 80°C e 60°C. No entanto, 0°C não indica ausência de calor, não se podendo afirmar que 9°C representa 3 vezes "mais calor" que 3°C). Uma *escala de rácio* difere de uma escala de intervalos, porque o zero tem existência real e não arbitrária, permitindo o cálculo de rácios. (Ex. A idade que é contada a partir do nascimento (0 anos) em qualquer tempo e lugar. Uma pessoa com 40 anos tem o dobro da idade de uma com 20 anos). Esta escala representa o nível de medida mais alto. De

uma escala de um determinado nível pode-se passar para a de um nível anterior, no entanto, perde-se informação.

2.5.2 Critério como função e desempenhos

As acções de um problema de decisão serão analisadas e avaliadas de acordo com o ponto de vista e preferências do decisor. Isso é matematicamente formalizado pela noção de critério.

Definição 2.5.1 (Vincke, 1992) *Um critério é uma função g , cujo domínio é o conjunto \mathcal{A} de acções a serem estudadas e o contradomínio é, em geral, o conjunto de níveis de desempenho que compõem a escala do critério respectivo, os quais representam as preferências do decisor de acordo com alguns pontos de vista.*

$$g: \mathcal{A} \longrightarrow \mathcal{V}, \text{ onde } \mathcal{V} \text{ é um conjunto totalmente ordenado}$$

O conjunto de todos os critérios será representado por $\mathcal{G} = \{g_1, \dots, g_n\}$. A notação $g_j(\cdot)$ ($j=1, \dots, n$) será aqui utilizada para representar os n critérios de avaliação. No caso do conjunto das acções estar definido em extensão ($\mathcal{A} = \{a_1, \dots, a_m\}$), pode-se construir uma tabela de dupla entrada, designada **tabela de desempenhos** ou matriz de decisão (Tabela 2.1). Cada elemento $g_j(a_i)$ da tabela representa o desempenho da acção a_i segundo o j -ésimo critério.

	$g_1(\cdot)$	$g_2(\cdot)$... $g_n(\cdot)$
a_1	$g_1(a_1)$	$g_2(a_1)$... $g_n(a_1)$
a_2	$g_1(a_2)$	$g_2(a_2)$... $g_n(a_2)$
...
a_m	$g_1(a_m)$	$g_2(a_m)$... $g_n(a_m)$

Tabela 2.1: Tabela de desempenhos

Regra geral, o desempenho dado por cada uma destas funções será um número ou uma menção qualitativa. Ao longo deste trabalho será utilizado um número que pode ser meramente ordinal, no caso de escalas qualitativas.

A um critério está associado um sentido de preferência que pode ser crescente, caso o critério seja a maximizar, ou decrescente, caso o critério seja a minimizar. Este aspecto deve ser tido em consideração na construção da função critério e deve constar na tabela de desempenhos. Os critérios podem ser definidos segundo Roy (1985) e Vincke (1992) como:

- *Verdadeiro critério* (critério na sua forma habitual), se a estrutura de preferência subjacente é uma estrutura de pré-ordenação completa (modelo tradicional), ou seja, uma acção é considerada preferível a outra num determinado critério se o seu desempenho for melhor do que a de outra, mesmo que por uma diferença mínima, e só serão consideradas indiferentes segundo esse critério caso os seus desempenhos sejam exactamente os mesmos;
- *Semi-critério*, se a estrutura de preferência subjacente é uma estrutura de semi-ordenação completa: modelo com valor de indiferença que é um valor positivo, abaixo do qual o decisor não consegue dar a sua preferência. Assim, uma acção é estritamente preferível a outra num determinado critério se a diferença de desempenhos for significativa (não inferior a um limiar de preferência) associado ao critério. As duas acções serão consideradas indiferentes segundo esse critério caso a diferença dos desempenhos seja pequena (não excede um limiar de indiferença) associado ao critério. Este valor de indiferença, de acordo com o método utilizado pode ser fixo ou um valor de consenso;
- *Critério de intervalo*, se a estrutura de preferência subjacente é uma estrutura de ordenação em intervalo (modelo com valor de indiferença variável);
- *Pseudo-critério*, se a estrutura de preferência subjacente é uma pseudo estrutura de ordenação, ou seja, uma acção é estritamente preferível a outra num determinado critério se a diferença de desempenhos for significativa (não inferior a um limiar de preferência) associado ao critério. As duas acções serão consideradas indiferentes segundo esse critério caso a diferença dos desempenhos seja pequena (não excede um limiar de indiferença) associado ao critério. Como os limiares de preferência e indiferença não têm que ter o mesmo valor, pode surgir uma situação de hesitação entre a preferência e a indiferença, designada por preferência fraca.

2.6 Problemáticas de ajuda à decisão

A formulação da questão que se pretende responder é uma das actividades da estruturação. Nem sempre o objectivo de um processo de decisão consiste na escolha da melhor alternativa (*selecção*). Por vezes, pretende-se simplesmente descrever as alternativas possíveis e as consequências que a escolha de cada uma delas acarreta (*descrição*). Outras vezes procura-se

classificar cada uma de acordo com um conjunto de categorias pré-definido (*afecção*). Finalmente, o objectivo poderá ser ordenar as diferentes acções por ordem de preferência, sendo necessário compará-las entre si (*ordenação*). Assim, quatro problemáticas gerais de referência devem ser consideradas⁶:

- *Problemática de selecção (P.α)*: consiste em apresentar o problema em termos da escolha de uma acção a implementar, de entre um conjunto de alternativas, mas sem impor a evidência duma solução como óptima;

- *Problemática de descrição (P.δ)*: consiste em apresentar o problema em termos da descrição das possíveis acções, bem como das consequências que delas podem resultar. Nestes casos, o decisor pretende usufruir de uma melhor compreensão de uma situação complexa e estabelecer uma linguagem para a comunicação entre actores;

- *Problemática de afectação* (ou de classificação) (*P.β*): consiste em apresentar o problema em termos da afectação das acções a categorias (classes) definidas *a priori*, que podem ser ordenadas (afecção ordinal) ou não ordenadas (afecção nominal), caso exista ou não uma ordem de preferência entre as categorias;

- *Problemática de ordenação (P.γ)*: consiste em apresentar o problema em termos de ordenação das acções, da mais preferida para a menos preferida, por forma a ser possível determinar acções que possam ser consideradas desejáveis ou suficientemente satisfatórias.

A problemática de afectação refere-se à *avaliação absoluta* das acções, confrontando-as com padrões pré-definidos. Tais padrões podem ser outras acções (designadas *acções de referência*), que podem ser elementos-tipo de uma categoria ou elementos de fronteira entre categorias. As problemáticas de selecção e ordenação referem-se à *avaliação relativa* das acções, confrontando-as entre si. Enquanto que na avaliação absoluta se utilizam conceitos como “mau” e “bom”, na avaliação relativa utilizam-se conceitos como “pior” e “melhor”. Assim, pode acontecer que a melhor acção seja má ou que a pior acção seja boa (Dias, 2002).

Uma vez determinados o conjunto de acções e o conjunto de critérios, tem que se analisar ainda as preferências do decisor. Este passo permitirá construir um modelo que, tanto quanto possível, representará correctamente a sua decisão.

⁶ver Roy e Bouyssou(1993), Roy(1996), Bana e Costa(1992)

2.7 Relação de dominância multicritério

Seja \mathcal{A} um conjunto de acções e \mathcal{G} uma família coerente de critérios resultante da fase de estruturação. Dadas duas acções a_i e a_j de \mathcal{A} , diz-se que a acção a_i *domina* (em termos multicritério) a_j se e só se os seus desempenhos não forem piores em nenhum critério e forem estritamente melhores em pelo menos um. Nesse caso, segundo Dias (2000), a_i deve ser considerada objectivamente superior a a_j , independentemente da importância que se atribua a cada critério, se não poder-se-á pôr em causa a coerência da família de critérios. Uma acção de \mathcal{A} diz-se *não dominada* caso não exista nenhuma acção em \mathcal{A} que a domine.

Note-se no entanto que este conceito de dominância não implica qualquer valorização das acções em termos absolutos, sendo apenas relativo à definição de \mathcal{A} . Assim, uma acção não dominada pode ser considerada má, enquanto uma acção dominada (por uma acção diferente da anterior) pode ser considerada boa (Dias, 2000). Note-se ainda que conceito de dominância é apenas importante caso se pretenda determinar a melhor acção de \mathcal{A} , a qual deverá ser obviamente não dominada.

2.8 Sistema de Relações de Preferência

As situações de preferência são formalmente relações binárias⁷ sobre o conjunto das acções. Roy(1996) admite quatro situações fundamentais de preferência na comparação entre duas acções a_i e $a_j \in \mathcal{A}$:

- *indiferença* ($a_i \mathbf{I} a_j$): corresponde à existência de razões que justifiquem uma indiferença entre as acções a_i e a_j . Nesta situação o decisor não consegue manifestar uma preferência, ainda que ligeira, sobre qualquer uma das acções;
- *preferência estrita* ($a_i \mathbf{P} a_j$): corresponde à existência de razões que justifiquem uma preferência significativa a favor da acção a_i ;
- *preferência fraca* ($a_i \mathbf{Q} a_j$): corresponde a uma hesitação entre **I** e **P**; não existem razões que justifiquem uma preferência significativa a favor da acção a_i . Contudo, há razões claras para não existir indiferença, havendo preferência, ainda que fraca, a favor de a_i ;
- *incomparabilidade* ($a_i \mathbf{R} a_j$): corresponde à ausência de razões que justifiquem alguma

⁷ \mathcal{R} é uma relação binária em $\mathcal{A} \iff \mathcal{R} \subseteq \{(a_i, a_j) : a_i, a_j \in \mathcal{A}\}$. O facto de (a_i, a_j) pertencer à relação \mathcal{R} será representado por $a_i \mathcal{R} a_j$.

das situações anteriores.

Estas relações podem-se assumir desde que satisfaçam as seguintes condições:

$$\left\{ \begin{array}{ll} a_i P a_j \Rightarrow a_j \bar{P} a_i & : \text{ P assimétrica} \\ a_i \bar{P} a_i & : \text{ P irreflexiva} \\ a_i Q a_j \Rightarrow a_j \bar{Q} a_i & : \text{ Q assimétrica} \\ a_i \bar{Q} a_i & : \text{ Q irreflexiva} \\ a_i I a_j \Rightarrow a_j I a_i & : \text{ I simétrica} \\ a_i I a_i & : \text{ I reflexiva} \\ a_i R a_j \Rightarrow a_j R a_i & : \text{ R simétrica} \\ a_i \bar{R} a_i & : \text{ R irreflexiva} \end{array} \right. \quad (2.1)$$

Em muitos métodos multicritério, como por exemplo as funções utilidade multiatributo, apenas duas situações são consideradas na comparação de pares de acções - preferência ou indiferença - as quais são assumidas como transitivas. Contudo, muitos foram os autores que criticaram estas suposições, como por exemplo Luce (1956).

Caso o decisor seja capaz de atribuir uma das quatro situações básicas a todos os pares de acções, de forma exhaustiva - qualquer par de acções verifica uma daquelas quatro relações de preferência - e mutuamente exclusiva - entre duas acções nunca se verifica a ocorrência em simultâneo de dois ou mais tipos de relações de preferências -, o modelo considerado é, segundo Roy (1996), um *Sistema Básico de Relações de Preferências*. Caso o decisor encontre duas ou mais relações básicas que considere como as suas preferências, apesar da mútua exclusividade, torna-se necessário considerar um segundo sistema relacional para modelar as preferências, que admita a verificação simultânea de mais do que uma situação básica. A ocorrência de mais que uma situação básica não é uma possibilidade real, mas traduz a impossibilidade de o decisor ter de optar por uma delas, podendo mesmo ser prematuro ou inútil ter de assumir essa opção de forma definitiva.

Os casos mais relevantes de agrupar as relações de preferência são:

- *não-preferência* ($a_i \sim a_j$): corresponde à não existência de razões que justifiquem qualquer tipo de preferência a favor de qualquer das acções:

$$a_i \sim a_j \iff a_i I a_j \vee a_i R a_j$$

• *preferência* ($a_i \succ a_j$): corresponde à existência de razões claras que justifiquem alguma preferência, estrita ou fraca, a favor de uma das acções:

$$a_i \succ a_j \iff a_i \mathbf{P}a_j \vee a_i \mathbf{Q}a_j$$

• *preferência-J* ($a_i \mathbf{J}a_j$): corresponde à existência de razões evidentes que justifiquem a preferência fraca (ainda que muito fraca) ou, caso contrário, a indiferença mútua:

$$a_i \mathbf{J}a_j \iff a_i \mathbf{Q}a_j \vee a_i \mathbf{I}a_j$$

• *preferência-K* ($a_i \mathbf{K}a_j$): corresponde à existência de razões evidentes que justifiquem a preferência estrita ou então as acções não são comparáveis, não sendo possível definir uma separação clara entre as duas situações:

$$a_i \mathbf{K}a_j \iff a_i \mathbf{P}a_j \vee a_i \mathbf{R}a_j$$

• *prevalência* ($a_i \mathbf{S}a_j$): corresponde à existência de razões claras que justifiquem a preferência, fraca ou estrita, a favor de uma das acções ou, caso contrário, indiferença mútua, mas sempre comparáveis entre si:

$$a_i \mathbf{S}a_j \iff a_i \mathbf{P}a_j \vee a_i \mathbf{Q}a_j \vee a_i \mathbf{I}a_j.$$

O conjunto das nove relações binárias definidas sobre um conjunto de acções ($I, R, \sim, P, Q, \succ, J, K, S$) definem, segundo Roy (1996), um *Sistema Consolidado de Relações de Preferência* (SCRIP), o qual é exaustivo e mutuamente exclusivo.

No caso da prevalência $a_i \mathbf{S}a_j$, que se pode traduzir por “ a_i é pelo menos tão boa quanto a_j ” podemos definir uma estrutura de preferência dado por:

$$\forall a_i, a_j \in \mathcal{A} \begin{cases} a_i \mathbf{S}a_j \wedge a_j \not\mathbf{S}a_i \iff a_i \mathbf{P}a_j \\ a_j \mathbf{S}a_i \wedge a_i \not\mathbf{S}a_j \iff a_j \mathbf{P}a_i \\ a_i \mathbf{S}a_j \wedge a_j \mathbf{S}a_i \iff a_i \mathbf{I}a_j \\ a_i \not\mathbf{S}a_j \wedge a_j \not\mathbf{S}a_i \iff a_i \mathbf{R}a_j \end{cases} \quad (2.2)$$

Numa estrutura de preferências de *ordem total*⁸, a relação R de incomparabilidade é vazia ($R=\emptyset$) e a relação de indiferença I é limitada a pares de elementos iguais ($a_i \mathbf{I}a_i$). A relação de

⁸Bouyssou e Vincke (2006) definem tal estrutura como transitiva, assimétrica e completa (uma relação binária φ em \mathcal{A} é completa caso $a_i \varphi a_j$ ou $a_j \varphi a_i \forall a_i, a_j \in \mathcal{A}$)

preferência estrita P é conexa⁹ e transitiva. Caso não seja imposta a condição de assimetria, ou seja, caso se permita a relação de indiferença, e se S for transitiva a estrutura de preferências é de *pré-ordem total*.

2.9 Métodos de agregação MCDA

Na secção anterior apresentamos noções básicas, definições e conceitos relativos às problemáticas de ajuda à decisão. Nesta secção apresentar-se-à uma breve revisão dos métodos de agregação multicritério referentes às problemáticas apresentadas na Secção 2.6. Obviamente que se torna impossível apresentar todos os métodos existentes, pelo que consideraremos apenas aqueles que têm alguma expressão na prática. Na verdade, podem-se distinguir três famílias de métodos: métodos da teoria da utilidade multiatributo, métodos de prevalência e métodos interactivos. Eles diferenciam-se pela forma como se consideram os critérios e preferências dados pelo decisor.

Os métodos da *teoria da utilidade multiatributo* (MAUT ou MAVT¹⁰) (Keeney e Raiffa, 1976) caracterizam-se por procurar atribuir um valor de desempenho global a cada uma das acções, agregando os seus desempenhos nos vários critérios através de uma função de utilidade. Assim, a decisão final do decisor é suportada, pelas preferências dos decisores em cada critério, de uma forma quantitativa, e baseada nas seguintes assunções: existência de uma função utilidade U , relação de preferência transitiva e completa (Vincke, 1992; Siskos e Spyridakos, 1999). Durante o processo de decisão o decisor tenta maximizar tal função e as suas preferências são expressas em forma de pesos atribuídos aos critérios. Os pesos representam taxas de compensação entre critérios.

A MAUT permite obter uma estrutura de preferências sobre o conjunto de acções tal que, uma acção é preferível a outra se e somente se o seu desempenho global for superior e são indiferentes se e somente se os seus desempenhos globais forem iguais. Desta forma, as relações de preferência e de indiferença são transitivas (ou seja, se $a_i P a_j$ e $a_j P a_t$ então $a_i P a_t$).

A MAUT pode ser utilizada no contexto de três problemáticas: na problemática de escolha, bastando para tal seleccionar a acção com melhor desempenho global; na problemática de

⁹uma relação binária φ em \mathcal{A} é conexa caso $a_i \neq a_j \implies a_i \varphi a_j \vee a_j \varphi a_i, \forall a_i, a_j \in \mathcal{A}$

¹⁰Da literatura anglo-saxónica multi-attribute value theory - MAVT.

ordenação, bastando ordenar as acções pelo seu desempenho global; na problemática de afectação ordinal, associando as categorias a intervalos de desempenho global. São exemplo de procedimento destes métodos as funções de utilidade/ valor aditivas e o método MACBETH¹¹ (Bana e Costa e Vansnick, 1994, 1995, 1997, 1999, 2000). O método AHP¹² de Saaty (1980; 1996), e os operadores de agregação *fuzzy* tais como o OWA¹³ ou os integrais *fuzzy* (ver revisões apresentadas por Grabisch, 1996; Roubens, 1997; Chen e Hwang, 1992) são exemplo de outros métodos que agregam desempenhos num único valor de síntese.

Relativamente à classificação ordinal, os métodos mais amplamente divulgados constroem a função utilidade quer de forma directa (ver, por exemplo, Winterfeldt e Edwards, 1986), quer indirectamente como por exemplo, os métodos UTA¹⁴ (Jacquet-Lagrèze e Siskos, 1982), UTA-DIS¹⁵ (Jacquet-Lagrèze, 1995; Zopounidis e Doumpos, 1999) e suas variantes (Doumpos e Zopounidis, 1998).

Os *métodos de prevalência*¹⁶ caracterizam-se por construir e posteriormente explorar uma ou várias relações binárias entre pares de acções. Estes métodos são menos exigentes em termos de informação requerida e podem modelar aspectos como a existência de preferências intransitivas ou incomparabilidade, contrariamente aos métodos anteriores. Na verdade, nos métodos anteriores, o decisor deve sentir-se capaz de indicar, perante duas acções, qual a melhor ou concluir que são indiferentes. Em contrapartida, as recomendações a extrair dos resultados dos métodos de prevalência poderão não ser tão directas como no caso dos métodos de utilidade/ valor, daí que se encontrem métodos dedicados apenas a uma problemática (Dias, 2002).

Muitos dos métodos de prevalência, tal como o nome indica, baseiam-se na *relação de prevalência* entre pares de acções (a_i, a_j) , a qual corresponde à reunião das relações de indiferença e de preferência. Tal relação é desenvolvida com base na força dos argumentos que suportam a afirmação “ a_i é pelo menos tão boa quanto a_j ” (concordância) e a força dos argumentos que suportam a afirmação contrária (discordância). Exemplos típicos destes métodos aplicados a

¹¹Measuring Attractiveness by a Categorical Based Evaluation Technique

¹²Analytic Hierarchy Process

¹³Ordered Weighted Average

¹⁴UTilitéé Additive Method.

¹⁵UTilités Additives DIScrimi-nantes

¹⁶Tradução possível para “surclassment”, nome dado por Bernard Roy (1985), e conhecidos na literatura Anglo-saxónica como *Outranking Relations*.

problemas de classificação ordinal são os métodos da família ELECTRE¹⁷ (Roy, 1989; Roy e Bouyssou, 1993) como o Trichotomic segmentation (Moscarola e Roy, 1977; Roy, 1981), que foi o primeiro método desta família desenvolvido para este tipo de problemas. Este método baseia-se no princípio de concordância e não-discordância (Roy, 1996) e é limitado a apenas três classes. Posteriormente, surgiram os métodos N-TOMIC (Massaglia e Ostanello, 1991) e ELECTRE TRI (Yu, 1992; Roy e Bouyssou, 1993) com um número arbitrário de classes ordenadas. Uma outra família de métodos de prevalência que atingiu alguma notoriedade é a família PROMETHEE¹⁸ (Brans e Vincke, 1985). Outros métodos menos conhecidos desta classe são o MAPPAC, MELCHIOR, ORESTE, PRAGMA, REGIME e TACTIC (ver com maior detalhe em Vincke (1999) e Roy e Bouyssou (1993)).

O ELECTRE TRI é sensível às variações dos valores dos parâmetros o que levou ao surgimento de uma extensão sua - o método SMAA-TRI (Tervonen et al., 2007). Este método analisa a estabilidade dos parâmetros do ELECTRE TRI arbitrariamente distribuídos num espaço finito para encontrar o suporte na atribuição de alternativas a cada classe. O método SMAA-TRI não serve apenas para classificação mas também para análise de sensibilidade e robustez dos parâmetros do ELECTRE TRI (Tervonen et al., 2007).

Vários têm sido também os métodos de prevalência desenvolvidos para classificação nominal, como por exemplo MC Filtering (Perny, 1998), PROAFTN (Belacel, 2000), *fuzzy integrals* (Montano Guzman, 2004) e TRINOMFC (Léger e Martel, 2002).

Cada um destes dois métodos distintos - teoria da utilidade multiatributo e prevalência - apresenta vantagens e desvantagens. A construção de uma função utilidade é mais restritiva já que possui muitas condições a serem satisfeitas e requer um esforço cognitivo considerável por parte do decisor (não necessariamente intuitivo). Contudo, permite obter um resultado profícuo e é axiomáticamente bem fundamentada. Os métodos de prevalência são bastante mais flexíveis já que têm poucas condições a serem satisfeitas, mas podem conduzir a resultados pobres e difíceis de justificar de um ponto de vista axiomático (Tsoukiàs, 2008).

Muitos autores têm mencionado que, se existem diferenças entre os métodos, estas dependem mais da forma de implementação do processo de decisão do que do método especificamente adoptado (ver por exemplo Belton e Stewart, 2002; Daellenbach, 1994; Roy, 1994; Roy, 1993; Keeney, 1992). Bouyssou e Pirlot mostraram ainda que é possível fornecer uma axi-

¹⁷Elimination and (Et) Choice Translating REality.

¹⁸Preference Ranking Organisation METHod of Enrichment Evaluation.

omática comum para todos esses métodos (ver Bouyssou e Pirlot, 2002, 2004, 2004a, 2005, 2005a). Assim, a divisão em categorias dos métodos supra referidos poderá ser enganadora.

Os métodos utilizados em MCDA necessitam de informação preferencial em forma de parâmetros do modelo. Contudo, nem sempre é fácil para o decisor expressar as suas preferências de forma precisa através da quantificação dos parâmetros. Por conseguinte, poderá ser benéfico questionar indirectamente o decisor, pedindo-lhe para especificar as suas preferências com base em exemplos, permitindo inferir determinados parâmetros do modelo a utilizar. Assim, numa terceira família de métodos, designada por *métodos interactivos*, os métodos são baseados em aproximações de Desagregação-Agregação, que segundo Vincke (1992) alternam fases de cálculos (comportando fases sucessivas de compromisso), com fases de diálogo (fontes de informação extra sobre as preferências dos decisores).

Para os métodos de classificação que se baseiam em julgamentos absolutos, a regra de classificação baseia-se em muitos casos na comparação das acções a classificar com as acções de referência representativas dos limites das classes. Uma aproximação alternativa é baseada em comparações entre as acções a classificar e acções que constituem exemplos típicos de cada classe. Os exemplos mais representativos são os algoritmos do vizinho mais próximo (*nearest-neighbor*) (Duda et al., 2001) e redes probabilísticas neurais (*probabilistic neural networks*) (Specht, 1990). Em ambos os exemplos são utilizadas medições de distância para estimar a similaridade das acções a classificar com o conjunto das acções de referência onde a classificação foi pré-especificada. Mais recentemente foi aplicada esta aproximação a problemas de classificação desenvolvidos com base neste esquema, tornando-se conhecida em métodos baseados na função utilidade como o método UTADIS (Jacquet-Lagrèze, 1995; Zopounidis e Doumpos, 1999), o método de Koksalan e Ulu (2003), PAIRCLAS (Doumpos, Zopounidis, 2004) e MHDIS (Zopounidis e Doumpos, 2000). Também nos métodos baseados nas relações de prevalência foi aplicada tal aproximação alternativa, como é o caso do ELECTRE TRI Assistant (Mousseau e Slowinski, 1998; Mousseau et al., 2000), o Electre Tri (Ngo The e Mousseau, 2002; Dias et al, 2002; Dias e Mousseau, 2003) e a abordagem de Rocha e Dias (2008).

Note-se que a taxinomia que acabámos de descrever não pretende definir três classes de métodos nem sequer apresentar os métodos de forma exaustiva, mas apenas tipificar as suas características.

Em MCDA, recorre-se muitas vezes à formulação de programação linear (PL) para inferir os parâmetros dos modelos de decisão à custa de um conjunto de alternativas de referência onde a classificação é pré-especificada. Por exemplo, no método UTADIS (Doumpos e Zopounidis, 2002) a formulação PL é usada para a construção de funções de utilidade. Mousseau et al. (1999), Dias et al. (2002) bem como Rocha e Dias (2008) recorreram à PL para estimar pesos dos critérios em modelos de classificação com relações de prevalência, desenvolvidos à custa do método ELECTRE (Roy e Bouyssou, 1993). De forma similar, Mousseau e Slowinski (1998), assumindo a não existência de veto, usaram a formulação de programação não-linear para estimar além dos pesos dos critérios, os limiares de preferência e de indiferença.

A inferência dos parâmetros à custa de exemplos é igualmente utilizada em modelos de inteligência artificial, os quais são representados, por exemplo, quer por regras de decisão do tipo “se...então...” ou árvores de decisão, quer à custa de redes artificiais neurais. Foi ainda proposta uma metodologia para a construção de regras de decisão dos modelos de preferência, baseada na teoria *rough sets* (Pawlak, 1982; Pawlak e Slowinski, 1994). Uma outra forma de aliviar o decisor de ter que indicar as suas preferências através dos parâmetros do modelo é utilizar métodos interactivos de análise de decisão verbal. Durante o processo de diálogo com o decisor, este vai classificando algumas acções seleccionadas pelo método, mas não todas. Neste tipo de métodos foram desenvolvidos por exemplo o ORCLASS (Larichev e Moskovich, 1994, 2000), SAC (Larichev et al., 2002) e um novo método Dichotomic Classification (Yevseyeva et al., 2008).

Métodos de Classificação Multicritério

3.1 Nota introdutória

Originalmente, os métodos MCDA foram desenvolvidos para as problemáticas da Seleção e Ordenação, os quais têm por base a medição do grau de preferências entre alternativas de um dado conjunto, e só mais tarde para problemas de Classificação. No contexto da MCDA, classificação supõe a atribuição de um conjunto finito de acções, descritas à custa de um conjunto de critérios. Há problemas de classificação feitos de forma ordinal e outros de forma nominal. As principais diferenças entre estes dois tipos de problemas de classificação são duas: nos problemas de classificação ordinal, os critérios de avaliação fornecem não apenas a descrição das acções, mas também informação preferencial para ajudar na identificação da melhor à pior das acções, e as classes são definidas de forma ordinal e não de forma nominal.

Os métodos para a problemática da Classificação para problemas de Classificação ordinal surgem em MCDA como um relaxamento natural de um problema de ordenação introduzindo a condição de que várias alternativas podem ser atribuídas em cada nível de classificação. Mais tarde, surgem nas metodologias MCDA a classificação do tipo nominal para a problemática da classificação bem como os problemas de agrupamento de dados (*Clustering problems*).

Os métodos de classificação em MCDA pretendem avaliar cada acção em termos absolutos, afectando cada acção a uma classe pré-definida. Este tipo de classificação é aplicado na prática de várias formas, nomeadamente no campo da (Zopounidis e Doumpos, 2002):

- medicina: realização de diagnósticos médicos através da classificação dos doentes com base em grupos de doenças com os mesmos sintomas (Stefanowski and Slowinski, 1998; Tsumoto, 1998; Belacel, 2000; Michalowski et al., 2001);
- gestão de recursos humanos: classificação das pessoas em grupos de acordo com as suas qualificações (Rulon et al., 1967; Gochet et al., 1997);
- marketing: medição da satisfação dos clientes, análise das características dos diferentes

grupos de clientes, desenvolvimento de estratégias para entrar no mercado (Dutka, 1995; Siskos et al., 1998);

- identificação de padrões: avaliação de características físicas de objectos ou pessoas e sua classificação em classes apropriadas (Ripley, 1996; Young and Fu, 1997; Nieddu and Patrizi, 2000);
- ambiente e gestão de energia, ecologia: análise e medição dos impactos ambientais das diferentes políticas de energia, investigação da eficiência das políticas da energia de países (Diakoulaki et al., 1999; Rossi et al., 1999; Flinkman et al., 2000);
- gestão de sistemas de produção e diagnóstico técnico: monitorização de sistemas de produção complexos para o diagnóstico de erros (Nowicki et al., 1992; Catelani and Fort, 2000; Shen et al., 2000);
- gestão financeira e económica: previsão de falência de empresas, avaliação de risco de crédito para firmas e consumidores, avaliação e classificação de acções cotadas na bolsa, avaliação de risco para o país, taxa de título da dívida (Altman et al., 1981; Slowinski and Zopounidis, 1995; Zopounidis, 1998; Doumpos and Zopounidis, 1998; Greco et al., 1998; Zopounidis et al., 1999).

Esta vasta gama de aplicações dos problemas de classificação constitui a principal motivação para o desenvolvimento de metodologias para a construção de novos modelos. De um ponto de vista técnico, tais modelos MCDA podem ser agrupados em duas categorias, com base na forma como o conjunto de alternativas potenciais é explorado (Tsoukiàs, 2008):

1. estabelecimento de uma função utilidade, sintetizando os diferentes critérios;
2. utilização de procedimentos para comparação entre pares e regras de maioria para estabelecer uma recomendação final.

Como referido no capítulo anterior, na primeira categoria temos métodos baseados na construção de uma função utilidade multi-atributo e na segunda categoria temos os métodos conhecidos como métodos de prevalência nos quais as relações de preferência são representadas pelo conceito “pelo menos tão bom como”.

Tal como a classificação em MCDA, o *agrupamento de dados* (ou clustering) tem por objectivo agrupar um conjunto de objectos (por exemplo, pessoas, produtos ou outras entidades), avaliadas segundo um conjunto de atributos, em *grupos*. A principal forma de distinguir

estas duas técnicas reside na definição dos grupos: na Classificação, os grupos (geralmente denominados *classes* ou *categorias*) são definidos *a priori*, respondendo, em certos casos, à clarificação de decisões ou à elaboração de recomendações ou de prescrições; no segundo caso os grupos (geralmente denominados *clusters*) são definidos *a posteriori*, sendo pois uma forma de aprendizagem por observação e geralmente relacionados com preocupações descritivas. Por outras palavras, a Classificação é considerada aprendizagem supervisionada, enquanto que o agrupamento de dados é um exemplo de aprendizagem não supervisionada. A figura 3.1 apresenta de forma gráfica estas diferenças.

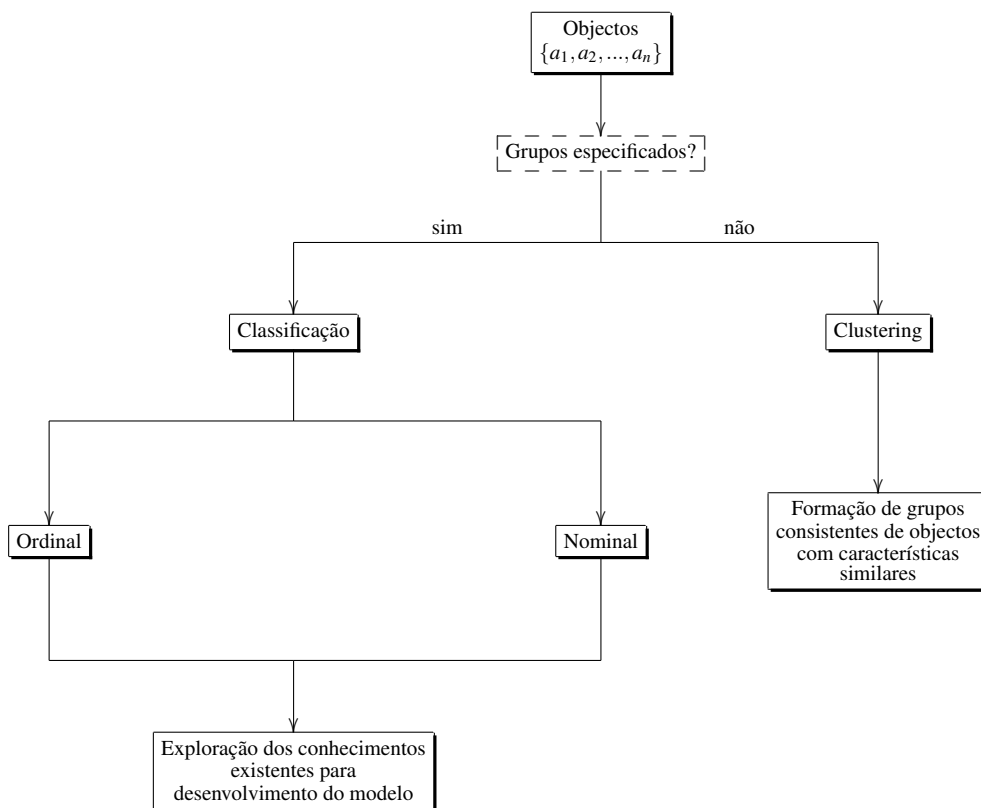


Figura 3.1: Comparação da Classificação e Clustering, adaptado de Doumpos e Zopounidis (2002, pg.5)

No agrupamento de dados, os agrupamentos resultantes de objectos devem exibir elevada homogeneidade interna (ou seja, elevadas similaridades intra-grupo) e elevada heterogeneidade externa (ou seja, elevadas dissimilaridades inter-grupos). Assim, se a classificação for bem sucedida, os objectos dentro dos agrupamentos estarão próximos, de acordo com uma medida de similaridade definida, e diferentes agrupamentos estarão distantes.

O agrupamento de dados tem as suas raízes em diversas áreas que incluem a estatística, o *data mining*, a biologia e a aprendizagem automática. Apesar de ser já bastante antigo,

este tipo de problemas ainda hoje é frequentemente encontrado em diversas áreas, sendo de grande importância na nossa sociedade. Tal é confirmado pelo grande número de artigos que se encontram na literatura, os quais apresentam uma grande variedade de aplicações tais como marketing (De Smet e Montano, 2004), agricultura (Jain et al. 1999), genética (Lu e Han, 2003), processamento de imagem (Jain e Flynn, 1996), aplicações Web (Cooley et al. 1997, 1999; Heer e Chi, 2001; Foss et al. 2001), análises de DNA em biologia computacional (Bendor e Yakhini 1999). Assim, termos mais ou menos sinónimos são utilizados no clustering tais como aprendizagem sem supervisão (Jain e Dubes, 1988), taxonomia numérica (Sneath e Sokal, 1973), vector quantization (Oehler e Gray, 1995), e aprendizagem por observação (Michalski et al., 1983). Várias são também as ferramentas de análise de grupos que têm sido construídas e incluídas em pacotes de software de análise estatística, tais como, S-Plus, SPSS, SAS, Genstat, SYSTAT, BMDP, Statistica, Stata, Minitab, Clustan e ClustanGraphics.

Tradicionalmente os algoritmos podem ser classificados em *hierárquicos* (envolvem a construção de uma hierarquia tipo árvore) e *não-hierárquicos* (ou de particionamento) (não envolvem o processo de construção em árvore; designam acções a clusters assim que o número de agregados a serem formados tenha sido especificado).

Além das técnicas estatísticas de análise de cluster hierárquica e não-hierárquica, outras técnicas como algoritmos evolutivos (Jain et al., 1999), agrupamentos fuzzy (Klawonn e Höppner, 2003), redes neurais (mapas de Kohonen) (Kohonen, 1990; Kohonen, 2001), clustering baseado em restrições para problemas com limitações específicas (Han et al., 2001; Tung et al., 2001) entre outras, podem ser utilizadas para formação de agrupamentos.

Na literatura pode-se encontrar um grande número de algoritmos de clustering (Jain e Dubes, 1988; Jain et al., 1999; Theodoridis e Koutroumbas, 1999; Everitt et al., 2001; Berkhin, 2002; Han e Kamber, 2006). Contudo, neste trabalho na Secção 3.3 serão apresentadas apenas as principais taxonomias existentes no clustering e as adoptadas neste trabalho, bem como, os algoritmos mais representativos e referenciados na literatura.

Tal como foi já referido, muitos são os trabalhos de classificação e clustering multicritério que nos últimos anos a comunidade científica tem desenvolvido, e embora estas duas formas de resolução de problemas sejam bastante diferentes à primeira vista, elas podem-se complementar. Anuska Ferligoj foi a primeira investigadora a abordar esta complementaridade a que denominou de problemática *multicriteria clustering* (Ferligoj, 1997). Recentemente novos trabalhos de investigação de *multicriteria clustering* têm sido desenvolvidos, sendo neste tipo de

trabalhos que estamos interessados.

Por serem os mais relevantes para o presente trabalho, serão apresentados de seguida e de forma bastante sucinta a família dos métodos ELECTRE, em particular o método ELECTRE TRI. De seguida serão apresentados os algoritmos de clustering mais representativos e referenciados na literatura, e por fim, apresentamos a motivação que tem levado ao desenvolvimento de alguns trabalhos de clustering multicritério.

3.2 Métodos ELECTRE

Os métodos ELECTRE (Roy, 1991) resultam da abordagem operacional de agregação de critérios através da relação de prevalência (ver secção 2.8), modelando a incomparabilidade entre acções. Estes métodos são relevantes em situações em que o decisor pretende incluir no modelo pelo menos três critérios. Contudo, os procedimentos de agregação são melhores para casos com pelo menos cinco critérios (até vinte ou trinta) (Roy, 1991; Roy e Bouyssou, 1993). Estes métodos devem ainda verificar pelo menos uma das situações seguintes (Figueira et al., 2005):

- as acções são avaliadas em pelo menos um critério de escala ordinal ou intervalar, as quais não são as mais indicadas para fazer comparações de diferenças. Assim, é difícil definir uma fórmula que faça sentido em termos de diferenças de preferência dos rácios $\frac{g_j(a)-g_j(b)}{g_j(c)-g_j(d)}$, onde $g_j(s)$ é a avaliação da acção s no critério g_j .
- critérios bastante heterogéneos que não permitem a sua agregação num único critério com uma escala comum.
- a compensação da perda num critério como ganho noutra critério pode não ser bem aceite pelo decisor, pelo que se torna necessário o uso de métodos de agregação não compensatórios.
- pequenas diferenças de desempenho num determinado critério, que não são significativas em termos de preferência, podem tornar-se significativas quando acumuladas. Torna-se assim importante o recurso a informação preferencial intra-critérios constituída por limiares de indiferença e preferência (q_j e p_j): dado um critério $g_j(\cdot)$, o limiar de indiferença q_j especifica a maior diferença da avaliação $g_j(a) - g_j(b)$ que mantém a indiferença en-

tre a e b (aI_jb); o limiar de preferência p_j especifica a menor diferença $g_j(a) - g_j(b)$ compatível com a situação de preferência a favor de a (aP_jb). Tal leva a uma estrutura de preferências cujas relações de indiferença são intransitivas.

Os métodos ELECTRE validam ou invalidam a asserção de prevalência a_sSa_t , para cada par de acções (a_s, a_t), que significa “ a_s é pelo menos tão boa quanto a_t ”. Para que a asserção a_sSa_t seja válida, duas condições devem ser verificadas:

- *concordância*: para que a prevalência a_sSa_t seja aceite, uma maioria “suficiente” de critérios deve ser a favor desta asserção (atendendo aos seus pesos);

- *não-discordância* : quando a condição de concordância se verifica, nenhum critério em minoria se deve opor à asserção a_sSa_t "de uma forma muito forte", ou seja, não existe em nenhum critério uma diferença de desempenhos tão significativa que se oponha à prevalência. Note-se que um mau desempenho de uma acção num critério pode não ser compensado por uma bom desempenho noutra critério, ou seja, o princípio da compensação total não se verifica aqui.

Para cada par de acções a_s, a_t , os métodos decidem se a_s prevalece sobre a_t (a_sSa_t) ou não, à custa dos desempenhos daquelas acções e dos seguintes parâmetros:

- o conjunto de *pesos* $\{k_1, k_2, \dots, k_n\}$ que indica a importância relativa dos critérios na agregação. Estes valores não devem ser vistos como taxas de compensação/conversão entre critérios, como acontece nos métodos compensatórios de agregação AHP, MACBETH e MAUT.

- o conjunto dos *limiares de veto* $\{v_1, v_2, \dots, v_n\}$ que representam a maior diferença da avaliação $g_j(a_t) - g_j(a_s)$ compatível com a asserção a_sSa_t .

- o *limiar de concordância ou corte* λ que representa o nível da maioria necessária (mas não suficiente) para a conclusão a_sSa_t .

A primeira fase dos métodos ELECTRE consiste na construção de uma relação de prevalência (ou várias) sobre o conjunto de acções (excepto no método ELECTRE TRI, no qual se comparam as acções a afectar com as de referência). Posteriormente, essas relações são exploradas no âmbito da problemática a que o método se dedica.

Os métodos ELECTRE I e IS (Roy e Skalka, 1984) dedicam-se à problemática da escolha, os métodos ELECTRE II (Roy e Bertier, 1971), ELECTRE III (Roy, 1978) e ELECTRE IV (Roy e Hugonnard, 1982), dedicam-se à problemática da Ordenação, e por último o ELECTRE TRI (Yu, 1992) dedica-se à problemática de classificação ordinal.

O método ELECTRE TRI (Yu, 1992; Roy e Bouyssou, 1993) é um método de afectação ordinal de um conjunto de acções a classes pré-definidas, com base em múltiplos critérios. Face a um conjunto de classes, cada uma delimitada por duas acções de referência, o método afecta cada acção a uma classe, por comparação com cada acção de referência.

Apresentaremos de seguida, de forma sucinta, o método ELECTRE TRI, nas suas variantes pessimista e optimista. Para tal, retomaremos aqui a notação introduzida no Capítulo 2.

3.2.1 ELECTRE TRI

Seja $\mathcal{C} = \{C^1, C^2, \dots, C^k\}$ um conjunto de k categorias (ou classes de preferência) pré-definidas e ordenadas, às quais se afectam as acções de \mathcal{A} . A ordem de preferência das classes é crescente, sendo C^1 a pior e C^k a melhor. Por exemplo, pode-se definir as classes $C^1 = \text{"Mau"}$, $C^2 = \text{"Insuficiente"}$, $C^3 = \text{"Suficiente"}$, $C^4 = \text{"Bom"}$, $C^5 = \text{"Muito Bom"}$.

As k classes consecutivas são separadas por acções de referência dadas pelo conjunto $B = \{b^0, b^1, \dots, b^k\}$. Cada classe C^j ($j=1, \dots, k$) é limitada por duas acções de referência, b^{j-1} e b^j , sendo os seus limites inferior e superior respectivamente. Cada acção de referência b^j ($j=1, \dots, k-1$) constitui simultaneamente o limite inferior da classe C^{j+1} e o limite superior da classe C^j .

As acções de referência devem obedecer às seguintes condições (Yu, 1992):

- b^0 deve ser em todos os critérios pior que qualquer das acções a classificar, o que implica que a_i prevalece sobre a pior acção de referência:

$$a_i S b^0, \forall a_i \in \mathcal{A}$$

- b^k deve ser em todos os critérios melhor que qualquer das acções a classificar, por forma a que a_i não prevaleça sobre a melhor acção de referência:

$$b^k P a_i, \forall a_i \in \mathcal{A}$$

- cada b^h deve ser melhor que b^{h-1} em todos os critérios, isto é:

$$g_j(b^h) > g_j(b^{h-1}), j \in \{1, \dots, n\} \quad (b^h \text{ domina } b^{h-1})$$

- $b^k P b^{k-1} P \dots P b^0$ (Relação de preferência, considerando os limiares de preferência e indiferença¹)
- cada acção $a_i \in \mathcal{A}$ pode ser indiferente no máximo a uma acção de referência.

3.2.2 Fases do Método

O ELECTRE TRI divide-se basicamente em duas fases:

- 1ª fase:** estabelecer uma relação de prevalência entre cada acção a_i ($i=1, \dots, m$) e as acções de referência b_j ($j=0, \dots, k$) que distinguem as diferentes classes de preferência;
- 2ª fase:** processar a afectação das acções às classes à custa das relações de prevalência estabelecidas.

A primeira fase é constituída normalmente por quatro etapas:

- cálculo dos índices de concordância $c_j(a_s, a_t)$ e do índice de concordância global $C(a_s, a_t)$;
- cálculo dos índices de discordância $d_j(a_s, a_t)$ e do índice de discordância global $D(a_s, a_t)$;
- cálculo do índice de credibilidade $S(a_s, a_t)$;
- estabelecimento da relação de prevalência através do limiar de corte.

3.2.3 Definição de relação de prevalência

No ELECTRE TRI a relação de prevalência entre pares acções é decidida considerando os seguintes parâmetros:

- k_j : coeficiente de importância (peso) do critério g_j , o qual é sempre positivo;
- Δ_j : vantagem da acção a_s sobre a acção a_t no critério g_j

$$\Delta_j = \begin{cases} g_j(a_s) - g_j(a_t) & \text{se } g_j \text{ é a maximizar} \\ g_j(a_t) - g_j(a_s) & \text{se } g_j \text{ é a minimizar} \end{cases}$$

- λ : limiar de corte;

¹Segundo o ELECTRE, uma acção pode dominar outra e não tem que lhe ser preferível; tal depende da amplitude da diferença de desempenho quando comparada com os limiares de indiferença e de preferência.

- q_j : limiar de indiferença para o critério g_j ;
- p_j : limiar de preferência para o critério g_j ;
- $c_j(a_s, a_t)$: índice de concordância para a asserção " $a S b$ ", considerando o critério g_j ;
- $C(a_s, a_t)$: índice de concordância global para a asserção " $a_s S a_t$ ", considerando todos os critérios;
- $d_j(a_s, a_t)$: índice de discordância para a asserção " $a_s S a_t$ ", considerando o critério g_j ;
- v_j : limiar de veto associado ao critério g_j ;
- u_j : limiar de discordância associado ao critério g_j ;

A introdução dos limiares de indiferença q_j e de preferência p_j , para cada critério $j (j=1, \dots, n)$, $p_j \geq q_j \geq 0$, permite uma transição gradual entre a indiferença e a preferência (o ELECTRE TRI utiliza **pseudo-critérios**²) como mostra a Figura 3.2:

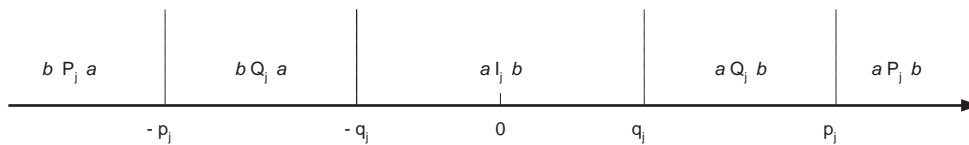


Figura 3.2: Relações de preferência do critério g_j para o par de acções (a,b) (adaptado de Dias (2000))

Obtém-se assim uma estrutura de de preferências (P_j, Q_j, I_j) , na qual Q_j representa uma *preferência fraca*, correspondente a uma hesitação entre a preferência (estrita) P_j e a indiferença I_j .

Para cada critério, o índice de concordância indica quanto o critério está de acordo com a hipótese $a_s S a_t$, como se apresenta de seguida:

$$c_j(a_s, a_t) = \begin{cases} 1 & \text{se } \Delta_j \geq -q_j \\ 0 & \text{se } \Delta_j < -p_j \\ \frac{p_j + \Delta_j}{p_j - q_j} & \text{caso contrário} \end{cases}$$

A concordância é máxima, ou seja 1, quando a acção a_s é melhor que a_t ou é pior mas por

²ver secção 2.5.2

uma pequena diferença (superior a q_j). É mínima, ou seja 0, quando essa diferença se torna igual ou superior a p_j . Finalmente, quando a_j é pior que a_t , a concordância começa a decrescer gradualmente entre 1 e o mínimo 0, à medida que a diferença a favor de a_t se vai tornando cada vez maior que q_j e mais próximo de p_j . Graficamente tem-se (Figura 3.3):

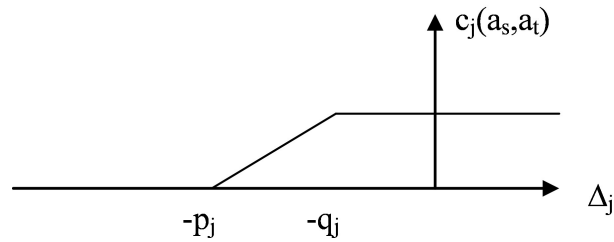


Figura 3.3: Índice de concordância para o critério g_j (adaptado de Dias (2002))

Os índices de concordância de cada um dos n critérios são agregados num *índice global de concordância* (multicritério), considerando os coeficientes de importância k_j de cada critério:

$$C(a_s, a_t) = \frac{\sum_{i=1}^n k_i c_i(a_s, a_t)}{\sum_{i=1}^n k_i}$$

Por outro lado, para cada critério, o índice de discordância indica quanto o critério discorda com a hipótese $a_s Sa_t$, como é apresentado abaixo:

$$d_j(a_s, a_t) = \begin{cases} 0 & \text{se } -\Delta_j \leq u_j \\ \frac{-\Delta_j - u_j}{v_j - u_j} & \text{se } u_j < -\Delta_j \leq v_j \\ 1 & \text{se } -\Delta_j > v_j \end{cases}$$

Graficamente tem-se (Figura 3.4):

A discordância é zero (*ausência de veto*) quando a_s é melhor que a_t ou é pior mas por uma diferença superior a u_j (com $u_j \geq p_j$). Quando essa diferença se torna igual ou superior a v_j , então a discordância é máxima (1) o que irá *vetar a hipótese* $a_s Sa_t$. Por fim, quando a_s é pior que a_t , a discordância cresce gradualmente entre 0 e 1, à medida que a diferença a favor de a_t se afasta mais de u_j e se aproxima de v_j . O veto será neste caso apenas *parcial*.

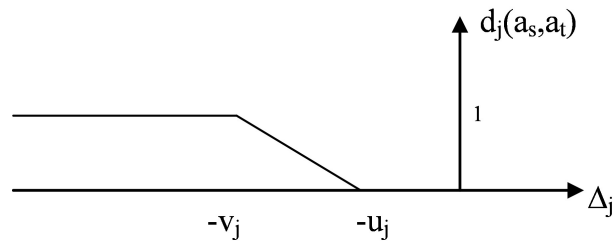


Figura 3.4: Índice de discordância para o critério g_j (adaptado de Dias (2002))

Os índices de discordância para os vários critérios são agregados num *índice global de discordância*, considerando o máximo desses valores:

$$d^{max}(a_s, a_t) = \max_{j \in \{1, \dots, n\}} d_j(a_s, a_t)$$

Finalmente a credibilidade da conclusão $a_s Sa_t$ é fornecida pelo *índice de credibilidade*, o qual pode ser calculado pela variante original de Yu (1992) e de Roy e Bouyssou (1993) dada por:

$$S(a_s, a_t) = C(a_s, a_t) \times \prod_{\substack{j \in \{1, \dots, n\} \\ d_j(a_s, a_t) > c_j(a_s, a_t)}} \frac{1 - d_j(a_s, a_t)}{1 - c_j(a_s, a_t)}.$$

Note-se que outras fórmulas podem ser utilizadas, como por exemplo, a variante de Mousseau e Dias (2004) dada por:

$$S(a_s, a_t) = C(a_s, a_t) \times [1 - d^{max}(a_s, a_t)].$$

Por fim, o *limiar de corte* λ é estabelecido para efectuar a dicotomia entre prevalência e não prevalência:

$$a_s Sa_t \iff S(a_s, a_t) \geq \lambda$$

Neste trabalho a relação de prevalência que aplicamos é a de Mousseau e Dias (2004).

3.2.4 Regra de afectação

Segundo Yu (1992), “Um processo de afectação é composto por um conjunto de mecanismos que, apoiando-se sobre modelos de preferência explícitos, permitem afectar todas as acções de \mathcal{A} às classes previamente definidas”.

No ELECTRE TRI, a afectação de uma dada acção $a_i \in \mathcal{A}$ processa-se comparando-a às sucessivas acções de referência. Este processo difere de acordo com a regra aplicada: pessimista (a mais aplicada na prática) ou optimista. A designação destas regras surge do facto de a classe a que uma acção é atribuída pela regra pessimista ser sempre igual ou pior à classe a que seria atribuída pela regra optimista (Yu, 1992). No caso de ser atribuída a uma classe pior pela regra pessimista, tal indica que a acção atribuída é incomparável a pelo menos uma acção de referência.

No ELECTRE TRI pessimista uma dada acção a_i é atribuída à classe mais elevada tal que $a_i S b^{h-1}$, ou seja:

$$a_i \in C^h \iff (a_i S b^{h-1}) \wedge (a_i \not S b^h) \iff S(a_i, b^{h-1}) \geq \lambda \wedge S(a_i, b^h) < \lambda$$

O algoritmo de afectação pessimista pode ser dado pelo apresentado na Tabela 3.1.

ELECTRE TRI pessimista
$\% C^k$ é a melhor classe
1. Compare a_i sucessivamente com b^r , $r=k-1, k-2, \dots, 0$.
2. Localizar o primeiro limite b^{h-1} tal que $a_i S b^{h-1}$.
3. Afectar a_i à classe C^h .

Tabela 3.1: Algoritmo do ELECTRE TRI pessimista

Quanto ao ELECTRE TRI *optimista*, uma dada acção a_i é atribuída à pior classe tal que $b^h P a_i$ e $b^{h-1} \not P a_i$, ou seja:

$$a_i \in C^h \iff (b^h S a_i \wedge a_i \not S b^h) \wedge (b^{h-1} \not S a_i \vee a_i S b^{h-1}) \iff (S(b^h, a_i) \geq \lambda \wedge S(a_i, b^h) < \lambda) \wedge (S(b^{h-1}, a_i) < \lambda \vee S(a_i, b^h) \geq \lambda)$$

Neste caso, o algoritmo de afectação optimista pode ser dado pelo apresentado na Tabela 3.2.

ELECTRE TRI optimista

% C^1 é a pior classe

1. Compare a_i sucessivamente com b^r , $r=1, 2, \dots, k-1$.
 2. Localizar o primeiro limite b^h tal que $b^h > a_i$.
 3. Afectar a_i à classe C^h .
-

Tabela 3.2: Algoritmo do ELECTRE TRI optimista

3.3 Métodos Clustering tradicionais

A escolha do conjunto de regras ou o algoritmo de clustering a utilizar, dado um determinado problema, não é uma questão simples. Porém, o critério essencial de todos os algoritmos é que eles tentam maximizar as diferenças entre clusters relativamente à variação dentro dos mesmos. Para Doreian et al.(2005), a escolha de um algoritmo de agrupamento deve considerar, entre outros, os seguintes aspectos:

- *objectivo da aplicação*: se a aplicação envolve a escolha do melhor local para construir um Centro de Saúde, o algoritmo de agrupamento escolhido deve ser um que permita encontrar a menor distância entre o local sugerido e os utentes. Neste caso a aplicação envolve a menor distância entre o centro do cluster em relação aos demais objectos pelo que se sugere que sejam usados os algoritmos K-Means e K-Medoids (ver secção 3.5). Já em aplicações onde os dados espaciais são imagens de satélite ou mapas de vegetação, uso do solo, etc, é aconselhável o uso de algoritmos baseados em densidade (ver secção 3.5), os quais criam clusters mais uniformes em termos de densidade, cor, pixels, entre outros.
- *tipo de critérios usados*: numéricos, ordinais, binários, etc;
- *validade da partição obtida*: esta deve ser analisada já que diferentes métodos de agrupamento podem levar a diferentes soluções para um mesmo problema. Jain et al. (1999) definem *cluster-validity* da seguinte forma: *a estrutura clustering é válida se não for razoável ter ocorrido ao acaso ou como um artefacto de um algoritmo de clustering*. Existem três tipos de estudo para validação: *avaliação externa* de resultados onde se compara a priori com uma estrutura; *análise interna* onde um exame intrínseco da estrutura é dado para ver se é apropriada; *teste relativo* que compara duas estruturas e mede os seus méritos relativos;

- fixar ou não a priori o número de grupos;
- a possibilidade de trabalhar com uma base de dados de enormes dimensões, ou seja, com acções caracterizadas por um enorme número de critérios e/ou um enorme número de objectos ;
- a interpretabilidade dos resultados e a visualização dos clusters (Jain et al, 1999).

De acordo com estes critérios existem várias propostas de classificação dos algoritmos de clustering. A classificação mais antiga e mais referida é baseada no tipo de resultado do agrupamento (Jain and Dubes, 1988): abordagens de agrupamento *Hierárquicas* versus *Partição*. Outras classificações separam os algoritmos de agrupamento de acordo com o método adoptado para definir os grupos (Berkhin, 2002; Han and Kamber, 2006): abordagens de partição, hierárquicas, baseados em densidade, baseados em grelha e baseadas em modelos.

Hierárquico O clustering hierárquico é uma técnica relativamente antiga comparativamente com muitos outros algoritmos de clustering, mas continua a ser uma das mais utilizadas. Esta técnica pode ser subdividida em dois tipos de algoritmos: *aglomerativos* e *divisivos* (Jain e Dubes, 1988; Kaufman e Rousseeuw, 1990). O clustering aglomerativo inicia com todos os clusters com uma única acção e recursivamente vai juntando dois ou mais clusters de forma apropriada. Quanto ao clustering divisivo, este inicia com um único cluster com todas as acções em estudo e vai particionando recursivamente os clusters, parando quando verificar um critério de paragem (frequentemente é o número k de clusters fixado). É de notar que nos algoritmos hierárquicos não existe uma função objectivo a ser minimizada. Por exemplo, na Figura 3.5(a-d)³, podem-se obter agrupamentos com 2, 4 ou 6 clusters aplicando o clustering hierárquico em cada nível: o grupo inicial Figura 3.5(a) pode dar origem a dois clusters (Figura 3.5(b)) cujos elementos de cada grupo são os mais similares entre si mas se fixarmos $k=6$, continuando a divisão pode obter-se os seis clusters da Figura 3.5(d).

Partição Os algoritmos de agrupamento de partição criam um agrupamento inicial com k grupos e iterativamente realocam os objectos de dados de um grupo para outro, de forma a otimizar uma determinada função de custo global, também conhecida como *função de similaridade* - entre os elementos do mesmo grupo - ou *função de dissimilaridade* - entre os elementos de grupos diferentes. Cada grupo é representado por um ou mais protótipos e cada

³adaptada de <http://www-users.cs.umn.edu/kumar/dmbook/index.php> visto em 20/06/2011

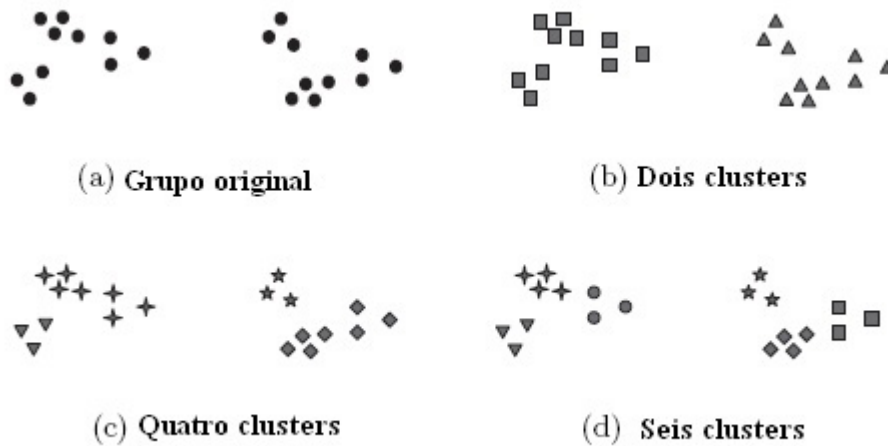


Figura 3.5: Diferentes tipos de clustering para um mesmo grupo

objecto é atribuído ao protótipo mais similar. Estes algoritmos encontram apenas grupos com forma esférica e raramente descobrem grupos com formas arbitrárias já que agrupam os objectos com base na distância entre eles. Por exemplo, usando individualmente cada grupo de clusters (Figura 3.5(b-d)) com o número fixo de grupos em 2, 4 ou 6, tem-se clustering de partição.

Densidade Os algoritmos de agrupamento baseados na noção de densidade consideram tipicamente os grupos como regiões densas de objectos no espaço dos dados que são separados por regiões de densidade significativamente mais baixa. Os objectos pertencentes a estas regiões de densidade significativamente mais baixa são considerados *ruído* ou *valores isolados*. A ideia geral destes algoritmos consiste em fazer crescer um determinado grupo enquanto houver objectos no grupo com densidade (um número de objectos) na sua “vizinhança” (num raio) superior a um determinado número limite. Assim, estes algoritmos têm sido desenvolvidos com o intuito de descobrir grupos com formas arbitrárias determinadas pela densidade. São normalmente eficazes para filtrar valores isolados e em conjuntos de dados de baixa dimensão. A quantidade de dados requeridos para estimar a densidade local aumenta exponencialmente com o aumento do número de atributos dos dados, não sendo por isso apropriados para conjuntos de dados de centenas de atributos.

Estes algoritmos usam uma função a otimizar local já que usam estruturas locais, as regiões densas, para formar os grupos.

Exemplos de algoritmos baseados em densidade são o DBSCAN⁴ (Ester et al., 1996) e o algoritmo DENCLUE⁵ (Hinneburg and Keim, 1998).

Algoritmos baseados em grelha quantificam o espaço de objectos num número finito de células que formam uma estrutura em grelha com múltiplas dimensões. A ideia é separar os valores possíveis de cada atributo num número de intervalos contíguos, criando um conjunto de células em grelha com múltiplas dimensões. Cada objecto situa-se numa célula cujo intervalo de valores para cada atributo contém os valores do objecto. Todas as operações de agrupamento são realizadas na estrutura em grelha existindo diferentes formas para a realização do agrupamento apesar de a maior parte das abordagens serem baseadas em densidade. O agrupamento é realizado pela junção de grupos de células densas adjacentes. Exemplos destes algoritmos são o algoritmo STING⁶ (Wang et al., 1997) e o algoritmo CLIQUE⁷ (Agrawal et al., 1998)

Os **algoritmos baseados em modelos** pretendem otimizar o ajuste entre os dados e um modelo matemático. Inicialmente é definido um modelo para cada um dos grupos e em seguida os modelos são adaptados de acordo com os objectos que lhes são atribuídos. Tais algoritmos são frequentemente baseados no pressuposto de que os dados são gerados por uma mistura de distribuições de probabilidade subjacente e os modelos são obtidos construindo uma função de densidade que reflecte a distribuição espacial dos objectos de dados. Como exemplos destes algoritmos temos o algoritmo EM⁸ (Dempster et al., 1977), o COBWEB (Fisher, 1987), CLASSIT (Gennari et al., 1989) e o AutoClass (Cheeseman and Stutz, 1996). Este último algoritmo é popular pois usa a análise estatística Bayesiana para estimar o número de grupos. Esta abordagem também é referida como abordagem estatística ou de inteligência artificial. A abordagem de rede neuronal pertence também a este grupo de algoritmos baseados em modelos, a qual tende a representar cada grupo como um exemplar (neurónio). Um exemplar actua como um protótipo do grupo e não tem necessariamente que corresponder a um exemplo de dados particular. Usando uma medida de distância, um novo objecto pode ser atribuído ao grupo cujo exemplar é mais semelhante. O estudo das redes neuronais artificiais (RNAs) (Hertz et al., 1991) foi inspirado pela constatação de que os sistemas de aprendizagem neurobiológicos

⁴Density Based Spatial Clustering of Applications with Noise

⁵DENsity-based CLUstEring

⁶STatistical INformation Grid

⁷CLustering In QUEst

⁸Expectation-Maximization

são constituídos por redes muito complexas de neurónios interligados.

Os algoritmos de agrupamento podem ainda dividir-se em *Exclusive* versus *Overlapping* versus *Fuzzy*, e ainda *Completo* versus *Parciais*:

Exclusivos versus Overlapping versus Fuzzy Na figura 3.5, o clustering é sempre **exclusivo**, ou seja, cada objecto é atribuído a um só cluster. O clustering **overlapping** (não-exclusivo) é usado para reflectir o facto de uma acção poder pertencer em simultâneo a mais do que um cluster (por exemplo, numa universidade, um individuo pode ser em simultâneo docente e estudante de doutoramento), ou quando uma acção está “entre” dois ou mais clusters e pode ser razoável atribui-la a qualquer um desses clusters.

No *fuzzy clustering*, faz-se corresponder cada acção a todos os clusters tendo associado um grau de pertença que varia entre 0 (não pertence de certeza) e 1 (pertença de certeza). Tal significa que cada cluster é tratado como um conjunto fuzzy⁹. De forma similar, as técnicas clustering probabilísticas calculam a probabilidade que cada elemento tem de pertencer a cada cluster, devendo a soma das probabilidades ser igual a 1. Na prática, estes dois métodos são geralmente convertidos no final em clustering exclusivos, atribuindo cada acção ao cluster que tiver associado o maior grau de pertença ou a maior probabilidade.

Completo versus Parcial No clustering completo todas as acções são atribuídas enquanto que no parcial tal não acontece. A motivação para o uso do clustering parcial é que determinadas acções podem não se enquadrar nos clusters definidos. Muitas das vezes, existem acções na base de dados que representam outliers ou “acções sem interesse”. Por exemplo, uma determinada patologia pode ter apenas interesse ser estudada em indivíduos idosos, pelo que os mais jovens devem ser excluídos do estudo.

Apesar da vastidão de algoritmos de agrupamento existente, nenhum é aplicável com êxito a todas as possíveis estruturas presentes nos conjuntos de dados multidimensionais. Para um determinado conjunto de dados com grupos de diferentes formatos, nem todos os algoritmos de agrupamento identificam correctamente os diferentes grupos existentes. Isto deve-se ao facto de, frequentemente, os algoritmos adoptarem suposições implícitas relativamente à forma dos

⁹Em Matemática, um conjunto fuzzy é um conjunto onde cada elemento pertence a qualquer conjunto com um grau que varia entre 0 e 1.

grupos, com base nas medidas de similaridade e nos critérios de agrupamento usados. Este facto explica o grande número de algoritmos de agrupamento de dados existentes na literatura, em que cada um, pela sua especificidade, tem um melhor ou pior desempenho num determinado conjunto de dados dependendo da distribuição específica dos objectos de dados. (Duarte,2008)

Por serem as de maior interesse neste trabalho, apresentamos nas secções seguintes as duas classificações de agrupamentos mais antigas: abordagens de agrupamento de partição e hierárquicas. Começaremos por apresentar na secção seguinte uma breve revisão das medidas de proximidade (dissimilaridade/similaridade) usadas nos algoritmos de agrupamento e que serão utilizadas nalgumas partes deste trabalho.

3.3.1 Medidas de proximidade

Um agrupamento pode ser convencional ou conceptual. O *agrupamento convencional* mede as similaridades a partir dos valores dos atributos que descrevem os objectos, sendo para tal usadas, geralmente, medidas de distância geométricas entre os valores dos atributos. No *agrupamento conceptual*, um grupo de objectos forma uma classe somente se esta puder ser descrita por um conceito comum a todos eles, ou seja, os objectos são agrupados de acordo com conceitos descritivos e não apenas a partir de simples medidas de similaridade. O agrupamento conceptual é realizado em duas fases: descoberta das classes apropriadas e criação de descrições para cada classe, tal como acontece na classificação.

No agrupamento convencional, na tentativa de identificação de grupos nos dados, é de primordial importância saber qual a similaridade ou distância entre esses dados.

Dois objectos consideram-se “próximos” quando a sua dissemelhança ou distância é pequena ou, o mesmo será dizer, quando a sua similaridade é grande. A similaridade entre acções pode ser medida de diferentes formas, sendo que cada um dos diferentes métodos utilizados representa uma perspectiva particular de similaridade, dependendo dos objectivos, do tipo de dados e do problema em estudo. Existe um vasto leque de medidas de similaridade/dissimilaridade que podem ser usadas para aferir a proximidade dos dados, devendo-se fazer uma escolha criteriosa já que existe uma grande variedade de tipo de atributos e escalas. O mais comum é calcular a dissimilaridade entre dois objectos utilizando uma medida de

distância definida no espaço de atributos.

Nesta secção são apresentadas medidas de proximidade apropriadas para dados contínuos por serem estes os mais relevantes para este trabalho. As medidas para este tipo de dados são tipicamente calculadas a partir de uma medida de *distância* a qual é generalizada pela matemática através do conceito de métrica. Estudos mais exaustivos de medidas de dissimilaridade podem ser encontradas em (Jain and Dubes, 1988), (Anderberg, 1973), (Diday and Simon, 1976), (Gower, 1985), (Gower and Legendre, 1986).

Distância

Doreian et al.(2005) começa por definir *medidas de dissimilaridade* as quais verificam as seguintes propriedades:

$$\forall a_i, a_j \in \mathcal{A}: d(a_i, a_j) \rightarrow \mathfrak{R} \text{ com}$$

1. $d(a_i, a_j) = 0$
2. $d(a_i, a_j) \geq 0$, não-negatividade
3. $d(a_i, a_j) = d(a_j, a_i)$, simetria

Quando as seguintes condições se verificam, as medidas de dissimilaridade são denominadas de *distância*:

4. $d(a_i, a_j) = 0 \Rightarrow a_i = a_j$
5. $\forall a_s : d(a_i, a_j) \leq d(a_i, a_s) + d(a_s, a_j)$ (desigualdade triangular)

Quando a base de dados é numérica, pode-se usar uma *distância de Minkowski* (L_p) dada por

$$d(a_i, a_j) = \| a_i - a_j \|_p = \left(\sum_{k=1}^m | a_{ik} - a_{jk} |^p \right)^{1/p}, \quad 1 \leq p \leq \infty \quad (3.1)$$

Quanto maior é o valor de p , menor é o valor atribuído às diferenças. Para diferentes valores de p temos diferentes distâncias, como por exemplo:

- $p=1$ obtemos *distância de Manhattan*
- $p=2$ obtemos *distância Euclidiana*
- $p=\infty$ obtemos a *distância de Chebychev* (valor máximo das diferenças absolutas entre as coordenadas: $\max | a_{ik} - a_{jk} |$)

O inconveniente do uso das métricas de Minkowski reside no facto destas métricas apresentarem uma tendência para que os atributos de maior escala dominem os restantes. Soluções para este problema incluem a normalização das escalas dos atributos contínuos.

Uma outra medida de distância é dada pela *distância de Mahalanobis*, dada por

$$d(a_i, a_j) = ((a_i - a_j)C_o^{-1}(a_i - a_j)^t)^{\frac{1}{2}} \quad (3.2)$$

em que C_o é a matriz de covariância dos objectos e t a transposta. Esta distância atribui diferentes ponderações aos diferentes atributos com base nas suas variâncias e nas correlações lineares entre os atributos. O custo computacional para obter a matriz de covariância com muitos atributos é muito elevado, sendo essa a razão porque esta medida não é tão usada. Por outro lado, a correlação linear entre os atributos pode também distorcer as medidas de distância. Esta distorção pode ser minorada usando o quadrado da distância de Mahalanobis (Mahalanobis, 1936):

$$d(a_i, a_j) = (a_i - a_j)C_o^{-1}(a_i - a_j)^t \quad (3.3)$$

Existem algumas medidas de distância na literatura que têm em conta o efeito dos objectos vizinhos (Gowda and Krishna, 1977; Jarvis and Patrick, 1973). Um exemplo destas medidas é a MND¹⁰ proposta por (Gowda and Krishna, 1977) e dada por:

$$MND(a_i, a_j) = NN(a_i, a_j) + NN(a_j, a_i) \quad (3.4)$$

em que $NN(a_i, a_j)$ é o número do vizinho a_j de a_i no ranking dos vizinhos mais próximos.

A medida MND não é métrica porque não satisfaz a desigualdade triangular. No entanto, foi aplicada com sucesso em várias aplicações (Gowda and Diday, 1992), suportando o ponto de vista de que uma medida de dissimilaridade não necessita ser métrica.

3.3.2 Algoritmos de agrupamento de dados

1. Clustering Hierárquico

A representação gráfica dos grupos mais usada nos algoritmos hierárquicos é uma árvore binária de grupos, chamada *dendograma* (Figura 3.6 (a)). A raiz do dendograma representa a totalidade do conjunto de dados, as folhas representam os objectos do conjunto de dados

¹⁰Mutual Neighbor Distance

e os nós internos a união de todos os objectos existentes nas suas sub-árvores. A altura de um nó interno representa a dissimilaridade (ou distância) entre si e os seus nós-filhos. Se um grupo não se juntar a outro ao longo de um considerável comprimento do eixo vertical, pode-se considerar um grupo estável. Cada nível da árvore representa um agrupamento do conjunto de dados com um determinado número de grupos. Outra forma possível, não tão utilizada e mais incompleta de representar os grupos hierarquicamente, consiste em usar um *gráfico de Venn* (Nested cluster diagram). A figura 3.6 (b) apresenta um gráfico de Venn típico para o mesmo conjunto de dados da figura 3.6 (a).

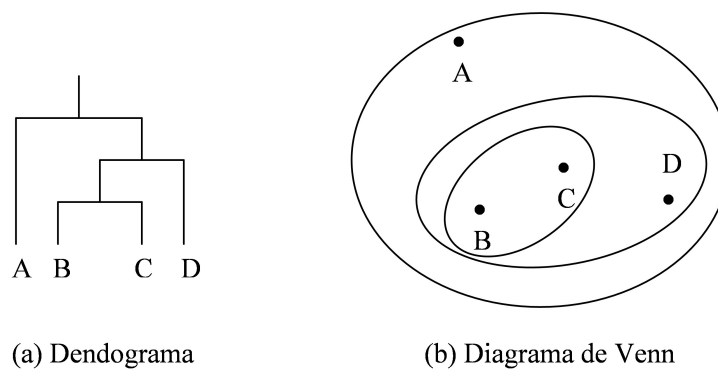


Figura 3.6: Representação gráfica dos grupos

Um algoritmo de agrupamento hierárquico clássico encontra frequentemente dificuldades na selecção dos grupos a juntar ou dividir e sofre da incapacidade de realizar ajustamentos uma vez feita a junção ou divisão dos grupos. Embora esta rigidez acarrete custos computacionais mais baixos já que não implica o teste de todas as combinações possíveis de outras escolhas, um dos seus principais problemas e desvantagens é não poderem corrigir decisões erradas.

Para compensar as medidas de similaridade simplistas e a rigidez das junções e divisões de grupos nos algoritmos hierárquicos clássicos que serão apresentados de seguida, surgiram outros algoritmos hierárquicos que usam princípios mais complexos e mais robustos. Exemplos desse tipo de algoritmos são os algoritmos CURE¹¹ (Guha et al., 1998) e o Chameleon (Karypis et al., 1999). Outros algoritmos integram a aglomeração hierárquica com algoritmos de realocação iterativa, como é o caso do algoritmo Birch (Zhang et al., 1996).

¹¹Clustering Using REpresentatives

1. Algoritmos aglomerativos

De um modo geral, os métodos aglomerativos utilizam um algoritmo padrão, conforme descrito no Algoritmo 1.

Algoritmo 1 - Clustering Hierárquico Aglomerativo

1. Iniciar com n grupos, contendo um elemento em cada grupo e uma matriz de dissemelhança D (ou a matriz de semelhanças) entre todos os pares de grupos;

2. **Repetir**

Localizar a menor distância d_{ij} (maior similaridade) entre grupos i e j ;
 Actualizar a matriz D , juntando os grupos i e j ;

Até critério de paragem.

O que distingue os algoritmos clássicos aglomerativos são as medidas de similaridade entre os grupos, também chamadas métricas de ligação (também denominada *Linkage*). O tipo de métrica de ligação usada pode afectar significativamente o resultado do algoritmo hierárquico, já que reflecte um conceito particular de proximidade e conectividade. Existem dois tipos de métricas de ligação: as que envolvem todos os pares de objectos nos dois grupos (um objecto em cada grupo) e as que envolvem um objecto central de cada grupo. As principais métricas de ligação do primeiro tipo são: ligação simples (Single-Link, SLINK) ou Vizinho Mais Próximo (*nearest neighbour*), ligação completa (Complete-Link, CLINK) ou Vizinho Mais Distante (*furthest neighbour*) e a média do grupo (Average-Link, ALINK). As principais métricas de ligação do segundo tipo são: ligação centróide e a ligação de Ward (também denominado Ward-Link ou mínima variância).

Segue-se uma breve descrição das métricas de ligação acima referidas representadas na Figura3.7.

Considerem-se dois grupos, G e H , para medir a semelhança/dissemelhança entre elas, n_G e n_H o número de objectos pertencentes aos grupos G e H respectivamente, e d_{ij} a distância entre os objectos i e j :

- *Ligação Simples ou Vizinho Mais Próximo* (Sneath and Sokal, 1973): considera que a distância entre dois grupos é a menor distância entre um elemento dum grupo e um elemento do outro grupo.

$$dist_{GH} = \min_{k \in G, l \in H} d_{kl} \quad (3.5)$$

Este método permite detectar grupos de formas não-elípticas; apresenta pouca tolerância a outliers, tendo tendência a incorpora-los num grupo já existente; em geral, grupos muito próximos podem não ser identificados; apresenta bons resultados tanto para distâncias Euclidianas quanto para outras distâncias; tem tendência a formar longas cadeias (encadeamento), ou seja, há um primeiro grupo de um ou mais elementos que passa a incorporar, a cada iteração, um grupo de apenas um elemento (Anderberg, 1973).

- *Ligação completa ou Vizinho Mais Distante* (King, 1967): considera que a distância entre dois subgrupos é a maior distância entre um elemento dum subgrupo e um elemento do outro subgrupo.

$$dist_{GH} = \max_{k \in G, l \in H} d_{kl} \quad (3.6)$$

Este método apresenta bons resultados tanto para distâncias Euclidianas quanto para outras distâncias; tem tendência a formar grupos compactos, mas tem tendência a dividir grupos naturais grandes; os outliers demoram a ser incorporados ao grupo (Kaufmann e Rousseeuw, 1990).

Note-se que estes dois métodos trabalham em direcções opostas. Assim, se eles apresentam resultados semelhantes, significa que o grupo está bem definido no espaço, ou seja, o grupo é real. Mas se ocorre o contrário, os grupos provavelmente não existem. (Romesburg, 1984)

- *Médias entre Grupos* (Sneath and Sokal, 1973): considera que a distância entre duas classes é a média de todas as distâncias entre pares de elementos (um em cada classe)

$$dist_{GH} = \frac{1}{n_G \cdot n_H} \sum_{k \in G} \sum_{l \in H} d_{kl} \quad (3.7)$$

- Método dos Centróides (Sneath and Sokal, 1973): considera-se a distância entre duas classes como sendo a distância entre os centros de gravidade, ou outros pontos considerados “representativos” (centróides) dos grupos. Este método não separa correctamente grupos com tamanhos diferentes e formas não esféricas.
- *Método de Ward* (Ward, 1963; Murtagh, 1984): minimiza o aumento da soma dos quadrados do erro após a fusão de dois grupos H e W num grupo G. A distância entre dois grupos H e W é igual à diferença entre a soma dos quadrados das distâncias entre os objectos e o centro do novo grupo $\overline{a_i^G}$ e a soma dos quadrados das distâncias entre os objectos dos grupos H e W e os respectivos centros $\overline{a_i^H}$ e $\overline{a_i^W}$:

$$dist_{HW} = \sum_{a_i \in G} (a_i - \overline{a_i^G})^2 - \left[\sum_{a_i \in H} (a_i - \overline{a_i^H})^2 + \sum_{a_i \in W} (a_i - \overline{a_i^W})^2 \right] \quad (3.8)$$

Os resultados da aplicação do algoritmo de ligação de Ward são por vezes próximos dos resultados obtidos pela aplicação dos algoritmos da média do grupo e de ligação centróide mas em muitas situações revelam-se melhores que estes.

Note-se que as diferentes métricas de ligação geralmente só produzem o mesmo resultado se os grupos forem compactos, esféricos, de tamanhos semelhantes e bem afastados uns dos outros.

A maior parte dos algoritmos hierárquicos clássicos são variantes dos algoritmos SLink, CLink e de ligação de Ward, sendo os algoritmos SLink e CLink os mais populares.

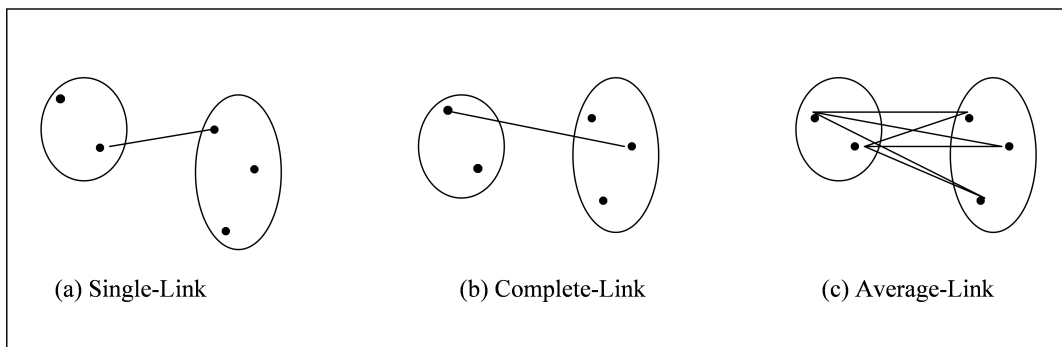


Figura 3.7: Diferentes medidas de proximidade

Qualquer das definições de distâncias entre grupos acima referidas pode ser vista como o

caso particular duma única expressão para a distância entre os grupos K e G \cup H formado pela junção dos grupos G e H¹²:

$$dist_{K,G\cup H} = \alpha_G dist_{KG} + \alpha_H dist_{KH} + \beta dist_{GH} + \gamma |dist_{KG} - dist_{KH}| \quad (3.9)$$

onde α_G , α_H , β e γ são coeficientes correspondentes a uma métrica de ligação particular (ver em Lance and Williams (1967)). A tabela 3.3 sumaria os coeficientes para os algoritmos hierárquicos de agrupamento mais usados, obtidos como instâncias da equação (3.9).

Métrica de Ligação	α_G	α_H	β	γ
SLink	1/2	1/2	0	-1/2
CLink	1/2	1/2	0	1/2
ALink	$\frac{n_G}{n_G+n_H}$	$\frac{n_H}{n_G+n_H}$	0	0
Ward	$\frac{n_G+n_W}{n_G+n_H+n_W}$	$\frac{n_H+n_W}{n_G+n_H+n_W}$	$-\frac{n_W}{n_G+n_H+n_W}$	0
Método dos centróides	$\frac{n_G}{n_G+n_H}$	$\frac{n_H}{n_G+n_H}$	$-\frac{n_G n_H}{(n_G+n_H)^2}$	0

Tabela 3.3: Valores dos coeficientes da equação (3.9) para cada um dos algoritmos

2. Algoritmos divisivos

Os algoritmos clássicos divisivos distinguem-se pela forma de selecção do grupo a ser dividido (por exemplo, o que tiver o maior número de objectos, a maior variância ou o maior erro quadrático médio) bem como pela forma de divisão (entre outras possibilidades, pode ser usada a média ou a mediana numa das direcções do vector de atributos ou a perpendicular à direcção com a maior variância). Um dos algoritmos hierárquicos divisivos mais conhecidos é o DIANA¹³ (Kaufman and Rousseuw, 1990).

Os algoritmos divisivos são pouco mencionados na literatura, pois exigem uma maior capacidade computacional que os algoritmos aglomerativos (Kaufman and Rousseuw, 1990), já que

¹²Este resultado deve-se a Lance e Williams(1966)

¹³**D**ivisive **A**NALysis

necessitam de um maior número de iterações. Nos algoritmos aglomerativos são consideradas todas as uniões possíveis de dois elementos, sendo a complexidade temporal da ordem $O(n^2)$, em que n é o número de elementos. Já os algoritmos divisivos, começam por considerar todas as divisões em dois grupos, com pelo menos um elemento em cada grupo, sendo por isso, a complexidade temporal da ordem de $O(2^n)$.

2. Clustering de Partição

Num algoritmo de agrupamento de partição cada grupo tem uma função de custo associada e o objectivo do algoritmo é obter um agrupamento que minimize a soma dos custos de todos os grupos. A função de custo mais utilizada é o erro quadrático, dada por:

$$e^2 = \sum_{j=1}^k \sum_{s=1}^{n_j} \|a_s^j - c_j\|^2 \quad (3.10)$$

em que $\|a_s^j - c_j\|$ é a distância Euclidiana entre o objecto a_s^j (s -ésimo objecto do j -ésimo grupo), c_j é o centróide (centro de gravidade) do j -ésimo grupo, n_j é o número de elementos desse grupo j e k ($k \leq n$) o número de grupos que se pretende obter.

A procura exaustiva de todos os possíveis agrupamentos, de forma a obter o valor óptimo da função de custo, requer um esforço computacional muitíssimo elevado já que a sua complexidade é exponencial. Alternativamente, adopta-se uma de duas heurísticas ou variações a essas heurísticas: o algoritmo K-Means ou K-Médias (Algoritmo 2), e o algoritmo *K-Medóides* (Algoritmo 3). No K-Means cada grupo é representado pelo valor médio dos objectos no grupo enquanto que no K-Medóides cada grupo é representado por um dos objectos localizados próximo do centro do grupo. Estas heurísticas optimizam uma função de custo usando um procedimento iterativo (Luenberger, 1984).

O critério de convergência normalmente verifica-se quando não existe nenhuma troca de objectos entre grupos ou a diminuição da função objectivo de uma iteração para a seguinte for menor que um valor mínimo atribuído à partida. No entanto, existem outros critérios de convergência menos usados, nomeadamente, a não alteração dos centróides ou o atingir de um limite máximo de iterações.

O algoritmo tem uma complexidade temporal $O(npkt)$, em que n é o número total de ob-

Algoritmo 2 - K-Means

Passo 1: seleção de forma arbitrária de k acções de \mathcal{A} para serem centróides iniciais dos k clusters.

Passo 2: (re)atribuir cada acção a um cluster, para o qual a dissimilaridade entre a acção e o centróide do cluster é menor que as demais.

Passo 3: os centróides c_k são (re)calculados através do valor médio dos objectos para cada grupo ($c_{kj} = \frac{1}{n_k} \sum_{i=1}^{n_k} a_{ij}$).

Passo 4: retorna ao passo 2 até ser atingido o critério de convergência.

jectos, p é o número de dimensões, k é o número de grupos e t é o número de iterações. Normalmente, $k \ll n$ e $t \ll n$. O algoritmo obtém bons resultados quando os grupos são nuvens esféricas de tamanhos semelhantes, compactas e bem separadas umas das outras (Jain et al., 2000).

Apesar da sua grande popularidade justificada pela sua simplicidade, compreensibilidade e pelas suas bases sólidas oriundas da análise da variância, o algoritmo K-Means sofre de algumas limitações que podemos resumir como:

- necessidade de especificar à priori, o número k de grupos a identificar;
- o agrupamento a obter depende dos centróides inicialmente escolhidos;
- aplicável apenas a dados não categóricos, pois só assim se podem calcular médias;
- obtém frequentemente um mínimo local da função objectivo caso os centróides não sejam bem escolhidos inicialmente;
- sensível a ruído e a dados isolados, já que um pequeno número deste tipo de dados pode influenciar substancialmente o valor médio;
- não apropriado para descobrir grupos com forma não convexa, com tamanhos muito diferentes ou mal separados.

Uma solução para tentar reduzir este efeito de mínimo local é executá-lo múltiplas vezes e escolher o agrupamento de dados que minimize o erro quadrático.

Dadas as limitações do algoritmo K-Means, surgiram muitas variantes que tentam superar algumas das suas limitações (Anderberg, 1973). Estas podem diferir na selecção dos K centróides iniciais, no cálculo da dissimilaridade, nas estratégias para calcular as médias dos grupos e na inclusão de procedimentos que permitem dividir ou juntar grupos no processo iterativo. Exemplo deste tipo de algoritmos são o ISODATA (Ball and Hall, 1965) e o K-Modas (Huang, 1998). O primeiro algoritmo tenta seleccionar um bom conjunto de centróides iniciais, para que o algoritmo chegue ao mínimo global da função objectivo. O segundo algoritmo faz agrupamento de dados categóricos por substituição das médias dos grupos por modas.

Os algoritmos K-Medóides, contrariamente ao K-Means que considera o valor médio dos objectos (centróide) como ponto de referência, usam o objecto mais centralmente localizado no grupo, chamado de *medóide*. Tal como o K-Means, estes algoritmos usam o princípio da minimização da soma das dissimilaridades entre cada objecto e o seu objecto de referência, isto é, a minimização da função do erro quadrático. A estratégia essencial dos algoritmos K-Medóides passa por primeiro, arbitrariamente, encontrar um objecto representativo (medóide) para cada grupo. Cada um dos restantes objectos é agrupado com o medóide com o qual é mais similar. A estratégia consiste em substituir iterativamente um dos medóides por um não medóide quando a qualidade do agrupamento resultante é melhorada. A qualidade dos grupos é estimada usando a função objectivo ou de custo que mede a dissimilaridade média entre os objectos e o medóide do seu grupo.

Os passos básicos do algoritmo K-Medóides são:

Algoritmo 3 - K-Medóides

Passo 1: selecção de forma arbitrária de k acções de \mathcal{A} para serem medóides iniciais dos k clusters.

Passo 2: (re)atribuir cada acção a um cluster, para o qual a dissimilaridade entre a acção e o medóide do cluster é menor que as demais.

Passo 3: Seleccionar aleatoriamente um objecto não medóide, $a_{aleatorio}$.

Passo 4: Trocar o medóide pelo $a_{aleatorio}$ caso a qualidade do agrupamento resultante seja melhor.

Passo 4: retorna ao passo 2 até não haver mudança de medóides.

O algoritmo K-Medóides é mais robusto que o algoritmo K-Means em presença de ruído e valores isolados pois o medóide é menos influenciado por ruído e valores isolados do que o centróide. Contudo, o seu processamento é mais dispendioso do que o algoritmo K-Médias.

O algoritmo **PAM**¹⁴ (Kaufman e Rousseuw, 1987) foi um dos primeiros algoritmos K-Medóides. Outros algoritmos desenvolvidos foram os algoritmos **CLARA** (Clustering LARge Applications) (Kaufman e Rousseuw, 1990) e o **CLARANS**¹⁵ (Raymond e Han, 1994, 2002).

3.4 Clustering Multicritério

Segundo Ferligoj (1997) os algoritmos clássicos de *clustering* podem não ser os mais apropriados para resolver alguns problemas de clustering pois requerem a otimização de mais do que um critério. Em geral, as soluções óptimas de cada critério em particular não são idênticas. Neste sentido, esta autora aproximou os problemas *clustering* com base nos métodos “tradicionais” mas tendo em conta a otimização de mais do que um critério em simultâneo e não a otimização de um critério de cada vez, tornando-se a primeira investigadora a abordar esta complementaridade a que denominou de problemática *multicriteria clustering* (Ferligoj, 1997).

Para Doreian et al. (2005) um problema *clustering* multicritério pode ser aproximado de diversas formas, entre elas, tratando-o como um problema *clustering* com um único critério obtido à custa de uma combinação dos critérios dados, ou *clustering* baseado em restrições onde um critério é seleccionado para ser considerado o critério do *clustering* e todos os outros são tratados como restrições (Batagelj e Ferligoj, 1998).

Embora, muitos sejam os trabalhos de classificação e *clustering* multicritério que nos últimos anos a comunidade científica tem desenvolvido, na verdade, poucos são aqueles que focam a complementaridade destas duas abordagens e, muito poucos são os estudos sobre o *clustering* multicritério que utilizam as preferências do decisor. Mas na verdade, é neste último tipo de trabalhos que estamos interessados, no qual a similaridade entre grupos é definida em termos de preferências.

O primeiro estudo que fez a integração das preferências do decisor com o *clustering* foi

¹⁴Partitioning Around Medoids

¹⁵Clustering Large Applications based upon RANdomized Search

desenvolvido por De Smet e Montano Guzman (2004), o qual é uma extensão de um dos métodos mais conhecidos do *clustering*, o método K-Means. É na definição de distância entre alternativas que o método é original, a qual se baseia nas relações binárias de preferência entre alternativas. Neste trabalho, os autores consideram que duas alternativas são similares caso sejam preferíveis, indiferentes e incomparáveis mais ou menos às mesmas alternativas. Desta forma, a natureza multicritério do problema é tida em conta.

Outros trabalhos que focam as preferências do decisor no *clustering* multicritério são, por exemplo, o trabalho de Figueira et al. (2004) que é uma extensão do método Promethee Tri à problemática de clustering, o trabalho de Nemery e De Smet (2005) que foi o primeiro método bastante próximo da bem conhecida Problemática de Ordenação de MCDA, cujo objectivo é ajudar o decisor a obter “clusters ordenados”, e o trabalho de De Smet e Eppe (2009) que estuda as possíveis relações de preferência entre os clusters obtidos com uma extensão do K-Means, relações essas baseadas numa matriz de relações de prevalência. O nosso trabalho segue a mesma direcção do trabalho de De Smet e Eppe (2009) uma vez que também ele se baseia nas relações de prevalência das alternativas. Contudo, exigiremos que as relações de preferência entre os clusters sejam transitivas.

Um dos trabalhos mais recentes do clustering multicritério é de Fernandez et al. (2010) onde se propõe uma ordenação completa dos grupos encontrados à custa das suas relações de preferência. Tais relações de preferência são estabelecidas à custa de valores de indiferença fuzzy, inspirados nos métodos de prevalência.

3.5 Notas conclusivas

Neste capítulo apresentamos dois tipos de classificação multicritério, classificação ordinal e agrupamento de dados (ou *clustering*) cujo objectivo comum é agrupar um conjunto de objectos em grupos. A definição dos grupos e dos objectos são as principais diferenças entre estas duas técnicas: na classificação os grupos são definidos *a priori* e os objectos são descritas à custa de critérios que fornecem informação preferencial; no agrupamento de dados os grupos não são definidos *a priori* e os objectos são definidos à custa de atributos que descrevem apenas os dados sem informação preferencial. Tal como foi referido ao longo deste capítulo, muitos são os trabalhos de classificação e multicritério que nos últimos anos a comunidade científica tem desenvolvido. Contudo, como o objectivo deste trabalho não é apresentar todos os métodos desenvolvidos de uma forma exaustiva, apresentamos de forma resumida os métodos que têm

alguma expressão na prática e em maior detalhe apenas os métodos mais relevantes para o trabalho: dos métodos de classificação ordinal apresentamos o método ELECTRE TRI da família dos métodos ELECTRE e dos métodos de agrupamento de dados apresentamos os métodos de partição e os hierárquicos.

Recentemente, alguns investigadores começaram a estudar a construção de *clusters* num contexto MCDA. Dado que partilhamos o reconhecimento da pertinência do estudo da junção das duas abordagens, classificação multicritério e *clustering*, e constatada a reduzida atenção que esta problemática tem tido em termos de investigação, apresentaremos nos próximos capítulos novas abordagens metodológicas baseadas nos modelos de classificação multicritério e *clustering*.

Discussão de medidas de atribuição para Problemas de Classificação

Este capítulo tem por objectivo apresentar e discutir diferentes princípios de atribuição de um conjunto de alternativas a classes parcialmente ordenadas. Tais princípios são baseados nas relações de prevalência das alternativas a serem atribuídas, e serão aplicados nos capítulos seguintes.

Nos métodos de classificação multicritério, as classes são geralmente definidas à custa de perfis, ou seja, limites de separação das classes ou protótipos delas. Aqui, os princípios que serão discutidos são para afectação a classes que não estão definidas através de duas acções de referência (ver cap.3) para as delimitar. Assim, a regra de afectação de uma alternativa a_i a uma classe C^j deixa de ser feita à custa da sua relação sobre as alternativas de referência dessa classe, mas sim com as restantes alternativas a serem atribuídas, ou seja, far-se-à uma avaliação relativa das alternativas e não uma avaliação absoluta.

As relações binárias aqui aplicadas são as usadas pelos métodos de prevalência. Tal como referido no capítulo anterior, a modelação da relação de incomparabilidades é uma das principais características das técnicas da teoria da prevalência. Para além das incomparabilidades, estes métodos podem ainda modelar aspectos como a existência de preferências intransitivas. A escolha destes métodos neste trabalho prende-se com o facto de considerarmos não só o caso em que as classes têm uma estrutura de ordem total, mas também outras estruturas que permitem que as classes sejam incomparáveis.

4.1 Estrutura de Preferências das classes

Pretende-se obter uma partição $\mathcal{P} = \{C^1, C^2, \dots, C^k\}$ de ordenação parcial de \mathcal{A} em k classes, definida da seguinte forma:

- $C^i \neq \emptyset, i = 1, \dots, k$

- $\mathcal{A} = \cup_{i=1, \dots, k} C^i$
- $C^i \cap C^j = \emptyset, i \neq j.$

Formalmente, a *classificação com classes parcialmente ordenadas* consiste em encontrar uma partição de \mathcal{A} onde a relação de ordem parcial é definida pelas classes. A estrutura de preferências da partição é dada pelo terno (\succ, \perp, \approx) onde “ $C^i \succ C^j$ ” significa que a classe C^i é melhor do que a classe C^j , “ $C^i \perp C^j$ ” significa que as classes C^i e C^j são incomparáveis, e “ $C^i \approx C^j$ ” significa que as classes C^i e C^j são indiferentes. Contudo, o propósito do modelo que pretendemos desenvolver é o de obter uma partição final tal que a estrutura de preferências não possua relações do tipo “ \approx ”, ou seja, não existam classes indiferentes. A partição parcialmente ordenada final, tem assim uma estrutura de preferências dada pelo par (\succ, \perp) , onde \succ é uma relação binária de ordem parcial e \perp é uma relação binária simétrica, ou seja:

- $C^i \succ C^j \Rightarrow C^j \not\succ C^i$: \succ é assimétrica
- $C^i \succ C^j \wedge C^j \succ C^k \Rightarrow C^i \succ C^k$: \succ é transitiva
- $C^i \perp C^j \Rightarrow C^j \perp C^i$: \perp é simétrica
- $C^i \perp C^j \vee C^i \succ C^j \vee C^j \succ C^i$: \approx é vazia.

Denotemos $C(a_i)$ a classe da alternativa a_i . Consideremos um conjunto \mathcal{P} de classes e a definição da relação de prevalência entre essas classes. Seja τ a relação binária sobre \mathcal{P} tal que $C(a_i) \tau C(a_j)$ significa que “a classe de a_i é pelo menos tão boa quanto a classe de a_j ”, ou seja,

$$C(a_i) \tau C(a_j) \iff C(a_i) \succ C(a_j) \vee C(a_i) = C(a_j) \vee C(a_i) \approx C(a_j)$$

Dada a relação de prevalência $C(a_i) \tau C(a_j)$, podemos definir uma estrutura de preferências dada por:

$$\forall C^i, C^j \in \mathcal{P} \begin{cases} C^i \tau C^j \wedge C^j \not\tau C^i \iff C^i \succ C^j \\ C^j \tau C^i \wedge C^i \not\tau C^j \iff C^j \succ C^i \\ C^i \tau C^j \wedge C^j \tau C^i \iff C^i \approx C^j \\ C^i \not\tau C^j \wedge C^j \not\tau C^i \iff C^i \perp C^j \end{cases} \quad (4.1)$$

Note-se que, face ao objectivo de obter uma estrutura de classes parcialmente ordenadas, pretende-se que a relação \approx seja vazia.

Definição 4.1.1 *A estrutura de preferências da partição \mathcal{P} diz-se uma ordem parcial fraca se a relação binária τ é transitiva.*

Definição 4.1.2 *Se todos os pares de classes (C^i, C^j) tais que $C^i \tau C^j$ e $C^j \tau C^i$ são fundidos $(C^j \cup C^i)$, então τ induz uma ordem parcial do conjunto de classes.*

Dada uma relação binária de \mathcal{P} é bastante útil a sua representação à custa de um grafo. Para cada par $(C^i, C^j) \in \mathcal{P}^2$, representaremos por um arco a relação $C^i \succ C^j$ ($C^i \longrightarrow C^j$). Quando as classes são incomparáveis ($C^i \perp C^j$), não existe qualquer arco a ligar as duas classes. A transitividade da relação \succ ($C^i \succ C^j \wedge C^j \succ C^k \Rightarrow C^i \succ C^k$) será representada apenas por dois arcos ($C^i \longrightarrow C^j \longrightarrow C^k$) sendo omitido o arco que liga C^i a C^k , tornando o grafo mais simples.

4.2 Princípios para problemas de classificação multicritério com classes parcialmente ordenadas

Para agrupar um conjunto de alternativas, dada a relação de prevalência entre classes a qual permite definir a ordem parcial da estrutura, apresentamos quatro modelos diferentes que são apresentados nas secções seguintes. Começaremos por apresentar um princípio que, embora desejável, facilmente se encontram exemplos que mostram que as suas exigências são demasiado fortes. De seguida são apresentados dois princípios que resultam do relaxamento do primeiro princípio, e por fim, o conceito de consistência semi-forte será apresentado no último princípio descrito.

4.2.1 Strong ($S\tau$)-Consistency (Consistência forte)

Uma possibilidade interessante para a atribuição de alternativas é ter por base o princípio **Strong consistency ($S\tau$ - Consistency)**: “uma alternativa a_i prevalece sobre a_j se e somente

se a classe de a_i é pelo menos tão boa quanto a classe de a_j ”, ou seja,

$$a_i S a_j \Leftrightarrow C(a_i) \tau C(a_j) \quad (S\tau\text{-Consistency})$$

Embora este princípio não seja respeitado pelo ELECTRE TRI (Roy e Bouyssou, 1993), consideramos que é um princípio interessante dado que $a_i S a_j$ significa que “ a_i é pelo menos tão boa quanto (ou não é pior do que) a_j ”.

Os corolários seguintes resultam do princípio $S\tau$ - Consistency (prova trivial pelo que é omitida):

- (1) $a_i P a_j$ sse a_i pertence a uma classe melhor que a classe de a_j , ou seja
 $(a_i S a_j \wedge a_j \not S a_i) \Leftrightarrow C(a_i) \succ C(a_j)$.
- (2) $a_i I a_j$ sse a_i e a_j pertencem à mesma classe, ou seja
 $(a_i S a_j \wedge a_j S a_i) \Leftrightarrow C(a_i) = C(a_j)$.
- (3) $a_i R a_j$ sse a_i e a_j pertencem a classes incomparáveis, ou seja
 $(a_i \not S a_j \wedge a_j \not S a_i) \Leftrightarrow C(a_i) \perp C(a_j)$.

O segundo corolário resulta do facto de $a_i I a_j$ se e só se a_i e a_j prevalecerem uma sobre a outra. O último corolário resulta do facto de $a_i R a_j$ se e só se $C(a_i) \not\prec C(a_j)$ e $C(a_j) \not\prec C(a_i)$. Um exemplo que satisfaz as condições de $S\tau$ - Consistency é dado pela Figura 4.1. As relações binárias para cada par $(a_i, a_j) \in \mathcal{A}^2$ são aqui representadas de acordo com as seguintes convenções: representamos $a_i P a_j$ por $a_i \xrightarrow{\quad} a_j$, $a_i I a_j$ por $a_i \xrightarrow{\quad} a_j$ e $a_i R a_j$ não se apresenta qualquer arco de ligação entre as alternativas. A relação binária “ \succ ” para cada par $(C^i, C^j) \in \mathcal{P}^2$ é aqui representada por $C^i \longrightarrow C^j$, não havendo qualquer arco a ligar duas classes incomparáveis.

A única possibilidade de cumprir integralmente o princípio $S\tau$ - Consistency é ter uma relação S que seja transitiva. Por exemplo, para a relação de prevalência na Figura 4.2 não é possível cumprir integralmente tal princípio: a_1 deveria pertencer a uma classe melhor do que a de a_2 e a_2 deveria pertencer a uma classe melhor que a de a_3 , mas não conseguimos concluir que a primeira classe seja melhor do que a última uma vez que é necessário que a_1 e a_3 sejam

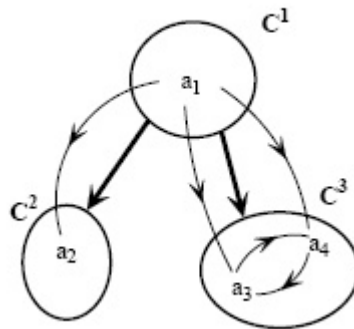


Figura 4.1: Representação do princípio $S\tau$ - Consistency.

atribuídas a classes incomparáveis.



Figura 4.2: Impossibilidade de cumprir $S\tau$ - Consistency

Para a otimização do problema, o objectivo poderá ser procurar soluções que minimizem o número total de violações dos pares de alternativas, ou que considere a violação mais forte dada a relação de prevalência S . Dias e Lamboray (2010) mostram como obter o conjunto óptimo de soluções recorrendo à “minimização” ou à “minimização lexicográfica” dos critérios do problema de acordo com uma estratégia de minimização da violação mais forte. Esses autores também mostram que é fácil encontrar exemplos em que a maior violação é realmente demasiado forte. Assim, apresentamos nas secções seguintes formas menos exigentes de consistência e discutiremos qual pode ser interessante seguir.

4.2.2 Formas de relaxamento de $S\tau$ -Consistency

Relaxando o princípio forte $S\tau$ -Consistency ($S \iff \tau$), obtemos duas formas fracas, a que chamaremos S -Consistency ($S \implies \tau$) e τ -Consistency ($\tau \implies S$), resultantes da divisão do princípio da equivalência em dois princípios de implicação. Começaremos por discutir estas duas formas de relaxamento, e de seguida apresentaremos a proposta de uma terceira forma de relaxamento, que denominamos por $SS\tau$ -Consistency, baseados numa lógica diferente.

A. S-Consistency

Seja **S-Consistency** o princípio de atribuição de alternativas que tem como base “se uma alternativa a_i prevalece sobre outra alternativa a_j então a classe de a_i deve ser pelo menos tão boa quanto a classe de a_j ”, ou seja

$$a_i S a_j \implies C(a_i) \tau C(a_j) \quad (S\text{-Consistency})$$

Os corolários seguintes resultam do princípio S-Consistency (prova trivial pelo que é omitida):

- (1) Se $a_i P a_j$ então a_i deve pertencer a uma classe pelo menos tão boa quanto a classe de a_j , ou seja

$$(a_i S a_j \wedge a_j \not S a_i) \implies C(a_i) \tau C(a_j).$$

- (2) Se $a_i I a_j$ então a_i e a_j devem pertencer à mesma classe, ou seja

$$(a_i S a_j \wedge a_j S a_i) \implies C(a_i) = C(a_j).$$

- (3) Se $a_i R a_j$ então a_i e a_j podem pertencer a qualquer classe.

O terceiro corolário deriva do facto de que se nenhuma das alternativas prevalece sobre a outra, então nada se pode concluir.

Dada uma ordenação parcial de classes, o princípio S-Consistency pode ser rephraseado da seguinte forma: "alternativas pertencentes a uma dada classe não podem prevalecer sobre alternativas pertencentes a classes melhores ou incomparáveis", ou seja,

$$(C(a_i) \not\prec C(a_j) \Rightarrow a_i \not\mathcal{S} a_j) \iff [(C(a_j) \succ C(a_i) \vee C(a_i) \perp C(a_j)) \Rightarrow a_i \mathcal{S} a_j]$$

Como consequência, obtêm-se os seguintes corolários:

- (4) Se $C(a_i) \succ C(a_j)$ então pode-se ter $a_i P a_j$ ou $a_j R a_i$.
- (5) Se $C(a_i) \perp C(a_j)$ então deve-se ter $a_i R a_j$.
- (6) Se $C(a_i) = C(a_j)$ então pode-se ter $a_i P a_j$, $a_i I a_j$, ou mesmo $a_i R a_j$.

De acordo com o princípio S-Consistency, as inconsistências seguintes podem ocorrer:

$$(I.1) \ a_i P a_j \wedge (C(a_j) \succ C(a_i) \vee C(a_j) \perp C(a_i)) \quad (S\text{-Inconsistências})$$

$$(I.2) \ a_i I a_j \wedge (C(a_i) \succ C(a_j) \vee C(a_j) \succ C(a_i) \vee C(a_j) \perp C(a_i))$$

Note-se que colocando todas as alternativas na mesma classe, pelo último corolário, obtém-se uma solução trivial para o princípio S-Consistency (Figura 4.3). Assim, o relaxamento do princípio $S\tau$ -Consistency não é totalmente satisfatório, mesmo permitindo uma maior liberdade na busca de um resultado que não viole muito a relação S inicial.

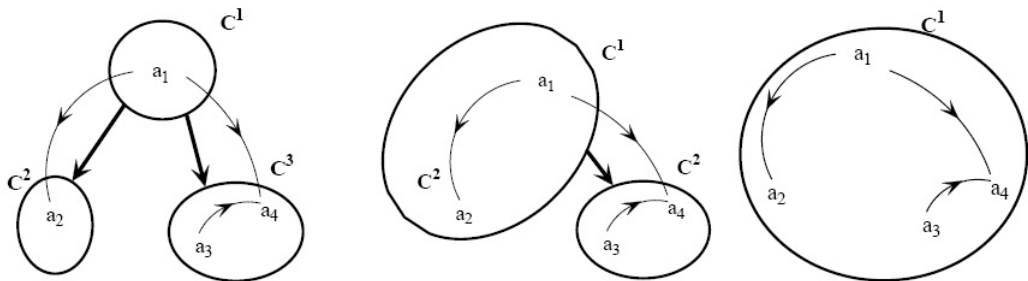


Figura 4.3: Representação do princípio S-Consistency: três soluções que satisfazem as condições

Caso exista um ciclo nas relações de prevalência $(a_i S a_j S a_k S \dots S a_i)$, então o princípio S-Consistency também implica que todas as alternativas implicadas no ciclo devem ser atribuídas à mesma classe. Tal segue a filosofia do ELECTRE I e é o princípio proposto por Rocha e Dias (2008) para um algoritmo interativo para problemas de classificação ordinal.

B. τ - Consistency

Considerando o princípio τ - Consistency, a atribuição de alternativas é baseada no princípio “se a classe de a_i é pelo menos tão boa quanto a classe de a_j então a alternativa a_i prevalece sobre a alternativa a_j ”:

$$C(a_i) \tau C(a_j) \Rightarrow a_i S a_j \quad (\tau - Consistency)$$

Os corolários seguintes resultam do princípio τ -Consistency:

- (1) se $C(a_i)$ é melhor que $C(a_j)$ então temos $a_i P a_j$ ou $a_i I a_j$.
- (2) se $C(a_i)$ e $C(a_j)$ são incomparáveis então podemos ter $a_i P a_j$, $a_j P a_i$, $a_i R a_j$, ou mesmo $a_i I a_j$.
- (3) se $C(a_i) = C(a_j)$ então devemos ter $a_i I a_j$.

O princípio τ -Consistency pode ser rephraseado da seguinte forma “alternativas que não prevaleçam sobre outras, pertencem a classes piores ou incomparáveis”. Como consequência, obtemos os três corolários seguintes:

- (4) se $a_i P a_j$ então a_i deve pertencer tanto a classes incomparáveis como melhores que a classe de a_j .
- (5) se $a_i I a_j$ então a_i e a_j podem pertencer a qualquer classe.
- (6) se $a_i R a_j$ então a_i e a_j devem pertencer a classes incomparáveis.

De acordo com o princípio τ -Consistency, as inconsistências seguintes podem ocorrer:

$$(I.1) \quad a_i P a_j \wedge (C(a_j) \succ C(a_i) \vee C(a_j) = C(a_i)) \quad (\tau\text{-Inconsistências})$$

$$(I.2) \quad a_i R a_j \wedge (C(a_j) = C(a_i) \vee C(a_i) \succ C(a_j) \vee C(a_j) \succ C(a_i))$$

Note-se que pelo corolário (2), se cada classe tiver apenas uma acção e as classes forem consideradas incomparáveis então esta pode ser uma solução trivial do princípio τ -Consistency. Assim, e mais uma vez, este relaxamento do princípio $S\tau$ -Consistency não é totalmente satisfatório.

C. Semi-Strong (SS τ) - Consistency

Dadas duas alternativas a_i e a_j pertencentes às classes $C(a_i)$ e $C(a_j)$, a Tabela 4.1 resume as combinações possíveis que são inconsistentes com S -consistency ou τ -Consistency. Note-se que todas as combinações proibidas por S -Consistency ou τ -Consistency são igualmente proibidas pelo princípio da consistência forte $S\tau$ - Consistency.

É evidente a partir da Tabela 4.1 que o princípio S -Consistency não coloca quaisquer restrições à atribuição de duas alternativas na mesma classe, e que τ -Consistency não coloca quaisquer restrições para a atribuição de duas alternativas a duas classes incomparáveis. No primeiro caso, uma solução trivial (e má) é colocar todas as alternativas numa mesma classe, e no segundo caso é ter classes incomparáveis com uma só alternativa. Tal leva-nos à exploração de um outro princípio que não permita a existência destas soluções triviais, o qual será apresentado de seguida.

		$a_i P a_j$		$a_j P a_i$		$a_i I a_j$		$a_i R a_j$	
		$a_i S a_j$	$a_j \not S a_i$	$a_i \not S a_j$	$a_j S a_i$	$a_i S a_j$	$a_j S a_i$	$a_i \not S a_j$	$a_j \not S a_i$
$C(a_i) \succ C(a_j)$	$C(a_i) \tau C(a_j)$			τ				τ	
	$C(a_j) \not\tau C(a_i)$				S		S		
$C(a_i) = C(a_j)$	$C(a_i) \tau C(a_j)$			τ				τ	
	$C(a_j) \tau C(a_i)$		τ						τ
$C(a_i) \perp C(a_j)$	$C(a_i) \not\tau C(a_j)$	S				S			
	$C(a_j) \not\tau C(a_i)$				S		S		

Tabela 4.1: Combinações proibidas por S -consistency (S) e por τ -consistency (τ).

Recorde-se que, de acordo com o princípio $S\tau$ -Consistency, as inconsistências seguintes podem ocorrer:

(I.1) $C(a_i) \succ C(a_j)$, mas $a_i Pa_j$ (i.e, $a_i \not\> a_j$ ou $a_j Sa_i$).

(I.2) $C(a_i) = C(a_j)$, mas $a_i Ia_j$ (i.e, $a_i \not\> a_j$ ou $a_j \not\> a_i$). ($S\tau$ -Inconsistências)

(I.3) $C(a_i) \perp C(a_j)$, mas $a_i Ra_j$ (i.e, $a_i Sa_j$ ou $a_j Sa_i$).

Dada a relação existente entre duas classes, o princípio da consistência forte impõe condições tanto a $a_i Sa_j$ como a $a_j Sa_i$. Um outro princípio que pode ser usado para relaxar o princípio da consistência forte (mais forte do que os princípios S-Consistency e τ -Consistency) é aceitar que uma dessas duas condições sobre S seja violada. Vamos denotar esse relaxamento de consistência forte por *Semi-Strong Consistency*, ou $SS\tau$ -Consistency. De acordo com este novo princípio, as inconsistências seguintes podem ocorrer:

(1) $C(a_i) \succ C(a_j)$ e $a_j Pa_i$ ($a_i \not\> a_j$ e $a_j Sa_i$).

(2) $C(a_i) = C(a_j)$ e $a_i Ra_j$ ($a_i \not\> a_j$ ou $a_j Sa_i$). ($SS\tau$ -Inconsistências)

(3) $C(a_i) \perp C(a_j)$ e $a_i Ia_j$ ($a_i Sa_j$ e $a_j Sa_i$).

A Tabela 4.1 apresenta-nos também as combinações que são proibidas pela consistência semi-forte. Todas as combinações que são proibidas ou por S-Consistency ou τ -Consistency, mas que afectem apenas um par ordenado, isto é, uma direcção, de a_i para a_j ou de a_j para a_i são agora aceites: se $C(a_i) \succ C(a_j)$, este princípio admite que eles são indiferentes ou incomparáveis (além de aceitar $a_i Pa_j$). Se $C(a_i) = C(a_j)$ ou $C(a_i) \perp C(a_j)$ este princípio admite que uma das alternativas é preferível à outra (além de aceitar que $a_i Ia_j$ ou $a_i Ra_j$, respectivamente). Por outro lado, este princípio rejeita combinações em que os dois pares ordenados (isto é, direcção de a_i para a_j e de a_j para a_i) estão errados, apesar de tais situações serem aceites por S-Consistency e τ -Consistency: não aceita alternativas indiferentes em classes incomparáveis (τ -Consistency aceita), e não aceita alternativas incomparáveis na mesma classe (que S-Consistency aceita).

Idealmente, após a afectação de todas as alternativas de \mathcal{A} deveremos encontrar as seguintes situações:

- $a_i Pa_j \wedge (C(a_i) \succ C(a_j) \vee C(a_j) \perp C(a_i) \vee C(a_j) = C(a_i))$

- $a_i I a_j \wedge (C(a_i) \succ C(a_j) \vee C(a_j) \succ C(a_i) \vee C(a_j) = C(a_i))$
- $a_i R a_j \wedge (C(a_i) \succ C(a_j) \vee C(a_j) \succ C(a_i) \vee C(a_j) \perp C(a_i))$

4.3 Notas conclusivas

Neste capítulo foi proposto um princípio de consistência forte para problemas de classificação com classes parcialmente ordenadas, baseado nas relações de prevalência do conjunto de alternativas a serem atribuídas. Este princípio a que denominamos $S\tau$ -Consistency pode ser traduzido da seguinte forma: “uma alternativa a_i prevalece sobre a_j se e somente se a classe de a_i é pelo menos tão boa quanto a classe de a_j ”.

Este princípio de consistência forte é desejável mas existem demasiadas relações de prevalência que tornam impossível cumprir este requisito, devido à intransitividade de S . Assim, foram propostos dois princípios, S -Consistency e τ -Consistency, que relaxam a consistência forte substituindo a equivalência $S \Leftrightarrow \tau$ por uma implicação: $S \Rightarrow \tau$ (S -Consistency) ou $\tau \Rightarrow S$ (τ -Consistency).

O princípio S -Consistency pode ser traduzido da seguinte forma: “se uma alternativa a_i prevalece sobre outra alternativa a_j então a classe de a_i deve ser pelo menos tão boa quanto a classe de a_j ”. O problema deste princípio S -Consistency é que não coloca quaisquer restrições para a atribuição de duas alternativas à mesma classe, sendo possível encontrar uma solução trivial desinteressante, em que todas as alternativas estão numa única classe.

Por outro lado, o princípio τ -Consistency pode ser traduzido da seguinte forma: “se a classe de a_i é pelo menos tão boa quanto a classe de a_j então a alternativa a_i prevalece sobre a alternativa a_j ”. Também este princípio pode ter um problema ao não colocar quaisquer restrições à atribuição de duas alternativas a duas classes incomparáveis, podendo-se encontrar uma solução trivial desinteressante, em que, caso as alternativas sejam incomparáveis, cada alternativa é uma classe diferente incomparável com as outras.

Apresentamos ainda o conceito de consistência semi-forte, que parece superar os problemas acima mencionados podendo ser uma boa opção para um procedimento de otimização. O princípio apresentado com esta consistência, denominado $SS\tau$ -Consistency, não impõe condições de prevalência a (a_i, a_j) e a (a_j, a_i) em simultâneo como o princípio $S\tau$ -Consistency, mas

aceitando que uma dessas duas condições sobre S seja violada.

Nos próximos capítulos serão apresentados algoritmos que, aplicando o último princípio aqui descrito, procuram atribuir um conjunto de alternativas a classes parcialmente ordenadas.

Classificação Multicritério com classes ordinais e nominais: Método Clustering-Sorting

5.1 Notas introdutórias

Neste capítulo propomos uma abordagem para classificar um conjunto de alternativas descritas por um conjunto de múltiplos critérios, conjugando, para tal, a análise de agrupamentos (clustering) e a classificação multicritério. Aqui, contrariamente à classificação em MCDA (Roy e Bouyssou, 1993) e de forma semelhante ao clustering, as categorias são desconhecidas à partida e não podem ser identificadas pelo decisor. Por outro lado, contrariamente ao clustering, pretende-se obter uma estrutura das categorias que pode ser de ordem parcial.¹

Tal como referido no Capítulo 3, foram desenvolvidos alguns trabalhos de investigação tendo por objectivo a conjugação da análise de agrupamentos e ferramentas tradicionais de classificação multicritério. Estes trabalhos estão direccionados principalmente para encontrar categorias que têm significado em termos de preferências estabelecidas a priori. No entanto, há uma variedade de problemas do mundo real que, embora os objectos sejam descritos por atributos ordinais (denominados critérios), numa primeira fase pretende-se detectar grupos com base numa medida de similaridade independente das preferências. Além disso, o trabalho desenvolvido para abordar a problemática multicritério de agrupamentos tem focado essencialmente a atribuição de alternativas a categorias totalmente ordenadas. No entanto, existem problemas onde muitas alternativas não são comparáveis.

A via que aqui propomos é construir um modelo que faz a junção do clustering com a classificação multicritério. O objectivo é detectar primeiro categorias homogéneas de alternativas

¹Parte deste Capítulo é uma adaptação do artigo publicado nos Proceedings of the 25th Mini-EURO Conference on Uncertainty and Robustness in Planning and Decision Making (Rocha e Dias, 2010).

e formar de seguida uma estrutura de ordem parcial com essas categorias detectadas. Assim, primeiro precisamos detectar apenas os clusters com base numa medida de similaridade independente das preferências do Decisor. Em seguida, pretende-se detectar potenciais relações de ordem parcial que possam existir entre eles, de acordo com os valores subjectivos do decisor. Vamos enfatizar que a eliciação de tais valores é feita somente depois de o Decisor identificar os agrupamentos e considerar que essas categorias fazem sentido.

A abordagem que propomos é uma abordagem heurística, uma vez que não tenta obter a melhor partição parcialmente ordenada de entre todas as partições possíveis. Mais do que isso, pretende ser uma abordagem simples e transparente para o decisor, cuja assistência é solicitada para ajudar a moldar os resultados.

Começaremos por apresentar a notação básica que será utilizada. A Secção 5.3 apresenta a abordagem *Clustering-Sorting*, que é ilustrada com dois exemplos da literatura na secção seguinte. Por último, algumas conclusões são apresentadas.

5.2 Formalização

Seja $\mathcal{A} = \{a_1, a_2, \dots, a_m\}$ um conjunto de alternativas representadas por um conjunto $\mathcal{G} = \{g_1, g_2, \dots, g_n\}$ de n critérios, onde $g_t(a_i)$ representa o desempenho da alternativa a_i no critério g_t . Sem perda de generalidade, consideramos que quanto maior é o valor do desempenho $g_t(a_i)$, melhor é a alternativa naquele critério. Pretende-se obter uma partição \mathcal{P} de ordenação parcial de \mathcal{A} em k categorias $\mathcal{P} = \{C^1, C^2, \dots, C^k\}$ definida da seguinte forma:

- $A = \cup_{i=1, \dots, k} C^i$
- $C^i \cap C^j = \emptyset, i \neq j$

Formalmente, a *classificação com categorias parcialmente ordenadas* consiste em encontrar uma partição de \mathcal{A} onde a relação de ordem parcial é definida pelas categorias: uma categoria pode ser melhor ou pior em comparação com determinadas categorias, mas também pode ser incomparável a outras categorias. A estrutura de preferências da partição é dada pelo par (\succ, \perp) onde “ $C^i \succ C^j$ ” significa que a categoria C^i é melhor do que a categoria C^j e “ $C^i \perp C^j$ ” significa que as categorias C^i e C^j são incomparáveis. Tal estrutura verifica as seguintes condições:

- $C^i \succ C^j \Rightarrow \neg(C^j \succ C^i) : \succ$ é assimétrica

- $C^i \succ C^j \wedge C^j \succ C^k \Rightarrow C^i \succ C^k$: \succ é transitiva
- $C^i \perp C^j \Rightarrow C^j \perp C^i$: \perp é simétrica.

Em termos muito gerais, para a obtenção de uma partição deste tipo, numa primeira etapa, as alternativas semelhantes são atribuídas a categorias homogêneas, usando um algoritmo de agrupamento tradicional. Com o algoritmo de agrupamento encontramos potenciais categorias tais que as alternativas dentro de uma determinada categoria possuem uma maior semelhança entre si (elevada homogeneidade) do que com as alternativas de outras categorias (elevada heterogeneidade). Numa segunda etapa, as categorias são ordenadas de acordo com as preferências do decisor. A condição de preferência é abordada apenas na segunda fase.

Para ilustrar graficamente este método, consideremos a Figura 5.1. Neste caso, as alternativas similares são primeiro agrupadas em categorias homogêneas (Figura 5.1(a)) e, em seguida, as categorias resultantes são ordenadas de forma parcial com base, por exemplo, nos seus centróides (Figura 5.1 (b)).

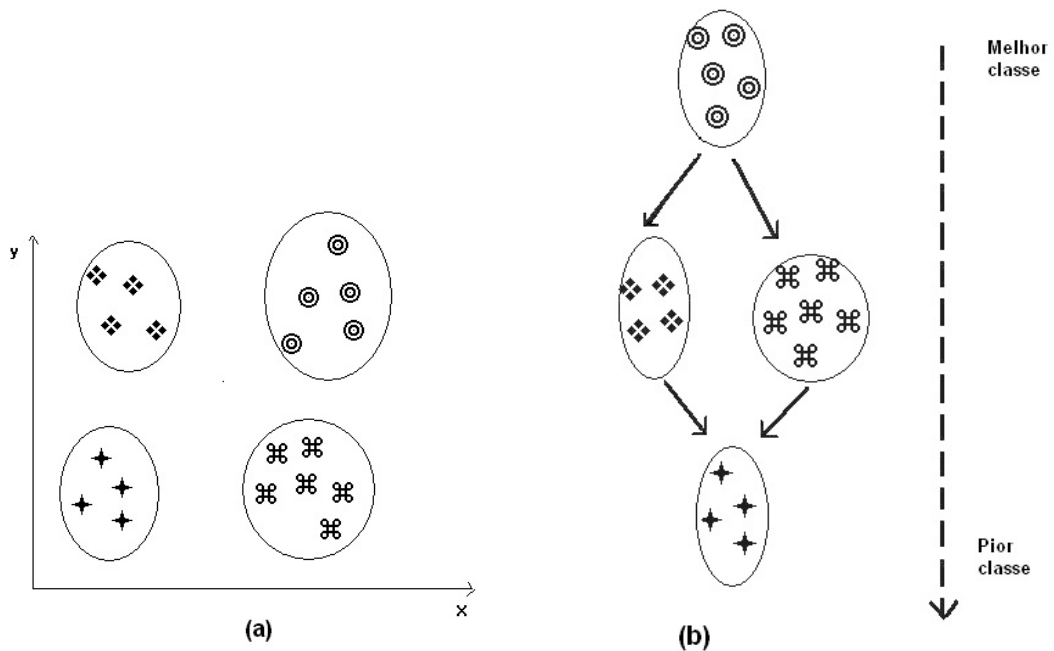


Figura 5.1: Ilustração do método Clustering-Sorting.

Assim, inicialmente resultam quatro grupos homogêneos de alternativas: o grupo cujas alternativas possuem um mau desempenho nos dois critérios x e y de avaliação; o grupo cujas alternativas possuem um bom desempenho nos dois critérios; o grupo cujas alternativas possuem um mau desempenho no critério x e bom desempenho no critério y e o grupo cujas

alternativas possuem um mau desempenho no critério y e bom desempenho no critério x . Numa segunda fase estes quatro grupos são ordenados à custa dos seus centróides: dois centróides incomparáveis - um com mau desempenho no critério x e bom desempenho no critério y e outro com bom desempenho no critério x e mau desempenho no critério y ; um centróide preferível aos restantes - com um bom desempenho nos dois critérios - e um centróide preterido relativamente aos restantes - com um mau desempenho nos dois critérios. A estrutura resultante é a apresentada na Figura 5.1 (b).

5.3 Método Clustering-Sorting

O método Clustering-Sorting é dividido em quatro fases que serão descritas nesta secção.

1. *Primeira fase - Clustering*: as alternativas similares são agrupadas em categorias homogêneas, à custa de um método de Clustering, antes de o decisor fornecer as suas preferências. Vamos usar o método K-Means com base numa distância euclidiana. A incerteza sobre o número das categorias é estudada usando métodos hierárquicos (para mais detalhes ver Jain et al., 1999).
2. *Segunda Fase - Eliciação*: Depois de definir o número de categorias que podem ser formadas, o decisor pode dar informações sobre o modelo multicritério de agregação, tais como os pesos dos critérios (consideramos neste trabalho a agregação de acordo com o ELECTRE (Roy e Bouyssou, 1993)).
3. *Terceira fase - Classificação Ordinal*: construção de uma ordem parcial do conjunto das categorias obtidas na primeira etapa. Para tal, os centróides das categorias são comparados ou a relação de preferência mais comum entre as alternativas das categorias é usada. Em ambos os casos é utilizada uma relação de prevalência do tipo ELECTRE (Roy e Bouyssou, 1993).
4. *Quarta fase - Qualidade*: estudo da qualidade da partição obtida nas fases anteriores, através do cálculo de um índice que indica as atribuições coerentes com o princípio $SS\tau$ -Consistency.

1. Primeira fase - Clustering

Na primeira fase do método é aplicado o *K-Means*. Este exige a definição prévia do número de clusters e do posicionamento do centro de cada cluster no espaço de atributos (ver Algoritmo 2, Secção 3.3.2).

Neste trabalho, para o cálculo da menor distância entre cada alternativa $a_i = (a_{i1}, a_{i2}, \dots, a_{in})$ e cada centróide $c_t = (c_{t1}, c_{t2}, \dots, c_{tn})$ (Passo 2 do Algoritmo 2), é utilizada a distância Euclidiana definida por

$$\sqrt{\sum_{j=1}^n (a_{ij} - c_{tj})^2}$$

Relativamente ao cálculo do valor de k , ou seja, do número de clusters a formar, aplicamos a *Distância entre Clusters*, onde o número de grupos é sugerido quando a inclinação da recta que une a distância entre dois grupos é relativamente pequena. Em caso de dúvida relativamente à escolha, é ainda aplicado o critério R-quadrado (*R-squared*), que é a medida de quão diferente cada um dos grupos ou cluster é em cada passo do algoritmo (Jonhson and Wichern, 2002). O *R-squared* é dado pela razão entre a soma dos quadrados entre os grupos (Soma dos Quadrados do Factor na Anova² one-way) e a soma dos quadrados totais (Soma dos Quadrados Totais na Anova one-way) para cada uma das variáveis usadas no estudo. Desta forma, o *R-squared* é uma medida da percentagem da variabilidade total que é retida em cada uma das soluções dos Clusters. No caso de haver um único cluster, é óbvio que a variabilidade entre os clusters é zero, e no caso de existirem tantos clusters quantas as alternativas, a variabilidade entre os clusters é igual à variabilidade total. Interessa pois encontrar o número mínimo de clusters que retenha uma percentagem significativa da variabilidade total (por exemplo, superior a 80 %).

2. Segunda fase- Eliciação

Ao considerar a análise de decisão multicritério, como pretendemos fazer na próxima fase, incorporamos também os parâmetros dos modelos relacionados com as preferências do decisor. Para transformar as preferências do decisor em valores numéricos é necessário elicitar os valores

²ANalysis Of VAriance

dos parâmetros sobre as suas preferências. Um processo de eliciação resulta da interacção entre o decisor e o analista na qual o decisor expressa as suas preferências relativas a determinado processo de agregação (Figueira et al., 2005).

Segundo Figueira et al. (2005), podem-se distinguir duas técnicas para eliciar valores: directa e indirecta. Nas técnicas directas o decisor pode fornecer directamente informação sobre os valores dos parâmetros de preferência. O maior problema destas técnicas reside na dificuldade em compreender o significado das afirmações do decisor. Quanto às técnicas indirectas, os valores dos parâmetros de preferência não são obtidos por respostas directas do decisor, mas sim através da interpretação das respostas dele no processo de agregação.

Neste trabalho, a aplicação de um método de prevalência na terceira fase exige que o decisor especifique vários parâmetros, tais como limiares de preferência, indiferença e de veto, pesos dos critérios e limiar de corte.

3. Terceira fase- Classificação Ordinal

Na terceira fase, para a ordenação das categorias, apresentamos duas possibilidades:

- (A) ordenação dos centroides das categorias;
- (B) aplicação da relação de preferência mais comum entre as alternativas das categorias.

Em ambos os casos, as relações de Preferência, Indiferença e Incomparabilidade serão expressas pela relação de prevalência. Uma das relações de prevalência para comparar pares de alternativas é usada por métodos tais como ELECTRE III (Roy, 1978) e ELECTRE TRI (Yu, 1992, Roy e Bouyssou, 1993). Dado um par de alternativas $(a_i, a_j) \in \mathcal{A}^2$, pode-se calcular o valor $S(a_i, a_j)$ que indica o grau de prevalência de a_i sobre a_j (ver mais detalhes na Secção 3.2). Assim, da comparação de duas alternativas a_s e a_t resulta uma pré-ordem parcial: a_s é preferível a a_t , a_t é preferível a a_s , a_s é indiferente a a_t , a_s é incomparável a a_t .

(A) Ordenação dos centróides das categorias

Os centróides c_i e c_j das categorias C^i e C^j , respectivamente, são comparados aos pares. Para cada par de centróides (c_i, c_j) o método indica se o primeiro prevalece sobre o segundo ($c_i S c_j$) ou não. A estrutura de preferência da partição é definida à custa das relações dos

centróides, $\{P, I, R\}$, podendo-se obter as seguintes situações:

$$\left\{ \begin{array}{l} C^i \succ C^j \text{ se } c_i P c_j \\ C^j \succ C^i \text{ se } c_j P c_i \\ C^i \approx C^j \text{ se } c_i I c_j \\ C^i \perp C^j \text{ se } c_i R c_j \end{array} \right. \quad (5.1)$$

Assim, a categoria C^i é preferível a C^j ($C^i \succ C^j$) caso o centróide c_i da categoria C^i seja preferível ao centróide c_j da categoria C^j , são indiferentes ($C^i \approx C^j$) caso os seus centróides sejam indiferentes e são incomparáveis ($C^i \perp C^j$) se os seus centróides também o forem.

(B) *A relação de preferência mais comum entre as alternativas das categorias*

A ideia base da segunda proposta para a construção duma estrutura de preferências é escolher a relação de preferência mais comum entre as alternativas das categorias, com base nas relações de prevalência do tipo ELECTRE. Tal pode ser caracterizado pelo vector p_{ij} que indica o número de pares de alternativas de C^i que prevalecem sobre as alternativas de C^j , bem como o número de alternativas de C^i que são prelevadas, que são incomparáveis e que são indiferentes relativamente às alternativas de C^j (definição 5.3.1).

Definição 5.3.1 *Seja $C^i, C^j \in \mathcal{P}$. O perfil p_{ij} do par $(C^i, C^j) \in \mathcal{P}^2$ é definido pelo quarteto $(|P_{ij}|, |P_{ji}|, |R_{ij}|, |I_{ij}|)$ onde:*

- $P_{ij} = \{(a_s, a_t) \in (C^i, C^j) : a_s P a_t\}$
- $P_{ji} = \{(a_s, a_t) \in (C^j, C^i) : a_s P a_t\}$
- $R_{ij} = \{(a_s, a_t) \in (C^i, C^j) : a_s R a_t\}$
- $I_{ij} = \{(a_s, a_t) \in (C^i, C^j) : a_s I a_t\}$

Assim, a estrutura de preferência final da partição é definida como se segue:

$$\left\{ \begin{array}{l} C^i \succ C^j \text{ se } \max p_{ij} = |P_{ij}| \\ C^j \succ C^i \text{ se } \max p_{ij} = |P_{ji}| \\ C^i \approx C^j \text{ se } \max p_{ij} = |I_{ij}| \\ C^i \perp C^j \text{ se } \max p_{ij} = |R_{ij}| \end{array} \right. \quad (5.2)$$

No caso (raro) de haver empate na obtenção do valor máximo de p_{ij} , mais de uma possibilidade pode ser mantida para futuras análises. A escolha final dependerá da robustez das soluções possíveis (apresentada na quarta fase).

Se a relação \succ contém um ciclo (por exemplo, $C^i \succ C^j, C^j \succ C^t$ e $C^t \succ C^i$), então a partição final obtida não é uma ordem, nem uma ordem fraca ou uma pré-ordem parcial. Quando isso acontece, pode-se remover a preferência mais fraca (por exemplo, de acordo com os graus de credibilidade), à semelhança do método ELECTRE IS (Roy e Bouyssou, 1993). Por outro lado, no caso de existirem categorias indiferentes, podem-se unir essas categorias ou transformar a relação de indiferença em relação de preferência, escolhendo a direcção do maior valor dos graus de credibilidade.

4. Quarta fase- Qualidade

Neste trabalho, para avaliar a Robustez da partição, propomos um índice que indica a percentagem de pares de alternativas coerentes com o princípio $SS\tau$ -Consistency apresentado na Secção 4.2.2. Segundo este princípio, as inconsistências na atribuição das alternativas podem ser três que denominaremos por inconsistências nas condições de Preferência, Indiferença e Incomparabilidade (Tabela 5.1).

Sejam $a_i \in C^s$ e $a_j \in C^t$.

Condições	Violação das condições
Preferência: $C^s \succ C^t \Rightarrow \neg(a_j P a_i)$	$C^s \succ C^t \wedge (a_j P a_i) \Leftrightarrow C^s \succ C^t \wedge (a_j S a_i \wedge \neg(a_i S a_j))$
Indiferença: $C^s \perp C^t \Rightarrow \neg(a_i I a_j)$	$C^s \perp C^t \wedge a_i I a_j \Leftrightarrow C^s \perp C^t \wedge (a_i S a_j \wedge a_j S a_i)$
Incomparabilidade: $C^s = C^t \Rightarrow \neg(a_i R a_j)$	$C^s = C^t \wedge a_i R a_j \Leftrightarrow C^s = C^t \wedge (a_i S a_j \vee a_j S a_i)$

Tabela 5.1: Inconsistências segundo $SS\tau$ -Consistency

Assim, uma partição pode ser avaliada por três indicadores distintos: v_R mede o rácio de pares que não verificam a condição de incomparabilidade, v_I mede o rácio de pares que não verificam a condição de indiferença e v_P mede o rácio de pares que não verificam a condição de preferência.

O índice v_R é dado por (5.3) e contabiliza os pares de alternativas incomparáveis e pertencentes à mesma categoria.

Definição 5.3.2 *Seja dada a partição $\mathcal{P} = \{C^1, C^2, \dots, C^k\}$ do grupo de alternativas \mathcal{A} . Sejam ainda a_i e a_j elementos de $C^t \in \mathcal{P}$ e $R_{ij} = \begin{cases} 1 & \text{se } a_i R a_j \\ 0 & \text{caso contrário} \end{cases}$ o valor que indica a existência ou não de uma relação de incomparabilidade entre as alternativas. O índice v_R é então definido por*

$$v_R = \frac{\sum_{\forall C^t \in \mathcal{P}} \sum_{(i,j): a_i, a_j \in C^t, i < j} R_{ij}}{n(n-1)/2} \quad (5.3)$$

O calculo dos índices v_I e v_P para as incoerências das condições de Indiferença entre pares de alternativas de grupos incomparáveis e de Preferência entre pares de alternativas pertencentes a grupos em que um é melhor que outro respectivamente são dados por (5.4) e (5.5).

Definição 5.3.3 *Sejam dadas duas categorias diferentes, C^p e C^q da partição \mathcal{P} e ainda a_i e a_j quaisquer elementos de C^p e C^q respectivamente.*

Seja $I_{ij} = \begin{cases} 1 & \text{se } a_i I a_j \\ 0 & \text{caso contrário} \end{cases}$ o valor que indica a existência ou não de uma relação de indiferença entre as alternativas a_i e a_j . O índice v_I é então definido por

$$v_I = \frac{\sum_{\forall (C^p, C^q) \in \mathcal{P}^2: C^p \perp C^q} \sum_{(i,j): a_i \in C^p, a_j \in C^q} I_{ij}}{n(n-1)/2} \quad (5.4)$$

Definição 5.3.4 *Sejam dadas duas categorias C^p e C^q da partição \mathcal{P} tais que uma é melhor que outra, $C^q \succ C^p$, e ainda a_i e a_j quaisquer elementos de C^p e C^q respectivamente. Seja $P_{ij} = \begin{cases} 1 & \text{se } a_i P a_j \\ 0 & \text{caso contrário} \end{cases}$ o valor que indica a existência ou não de uma relação de preferência a_i sobre a_j . O índice v_P é então definido por*

$$v_P = \frac{\sum_{\forall (C^p, C^q) \in \mathcal{P}^2: C^q \succ C^p} \sum_{(i,j): a_i \in C^p, a_j \in C^q} P_{ij}}{n(n-1)/2} \quad (5.5)$$

Será assim definido um *vector de inconsistências* $ID_{\mathcal{P}}$ associado à partição \mathcal{P} , o qual contém o rácio de pares de alternativas que são incompatíveis com as condições de preferência, indiferença e incomparabilidade, dado por (5.6).

$$ID_{\mathcal{P}} = (v_P, v_I, v_R) \quad (5.6)$$

O índice $Q(\mathcal{P})$ que é um indicador para os pares de alternativas coerentes com o princípio $SS\tau$ -Consistency é dado por 5.7.

Definição 5.3.5 *Sejam α_P , α_I , α_R e α_D os pesos atribuídos pelo decisor respectivamente às violações das condições de Preferência (v_P), Indiferença (v_I), Incomparabilidade (v_R). A Qualidade $Q(\mathcal{P})$ da partição \mathcal{P} é dada por*

$$Q(\mathcal{P}) = 1 - \alpha_P v_P - \alpha_I v_I - \alpha_R v_R \quad (5.7)$$

onde $0 \leq \alpha_P, \alpha_I, \alpha_R \leq 1$.

Se $\alpha_P = \alpha_I = \alpha_R$ então as violações das condições de Preferência, Indiferença e Incomparabilidade serão consideradas equivalentes. Contudo, o decisor pode “jogar” com estes pesos para estudar taxas de compensação (*trade-offs*) entre o número de pares que violam cada condição. Por exemplo, caso o decisor pretenda estudar a probabilidade de que “ $C^i \succ C^j$ ” não seja contestada, ou seja, a probabilidade inversa de encontrar inconsistências relativas a preferências invertidas, quando um par de alternativas (uma de cada categoria) é escolhido aleatoriamente, que pode ser obtida fazendo $\alpha_I = \alpha_R = 0$.

5.4 Exemplos ilustrativos com casos reais da literatura

Nesta secção são apresentados dois exemplos de classificação para ilustrar o método Clustering-Sorting proposto. Consideraram-se equivalentes as violações das condições de Preferência, Indiferença e Incomparabilidade ($\alpha_P = \alpha_I = \alpha_R = 1$) dado não termos qualquer informação sobre as exigências de um decisor.

As ferramentas utilizadas na obtenção de resultados foram o *Statistical Package for Social Sciences 17.0* na primeira fase do método Clustering-Sorting e o *MatLab 7.0* nas terceira e quarta fases do método proposto.

5.4.1 Primeiro exemplo ilustrativo

Consideremos um exemplo baseado em Hurson e Zopounidis (1997), que se refere à avaliação de alternativas cotadas na Bolsa de Atenas. O estudo destes autores abrangeu vinte alternativas do sector comercial e outras vinte do sector industrial, avaliadas segundo seis critérios. Neste exemplo será feita apenas a avaliação das vinte alternativas do sector comercial, cujos desempenhos em 6 critérios são apresentados na Tabela 5.2, sendo os critérios os seguintes:

- g_1 : Retorno [a maximizar].
- g_2 : Liquidez [a maximizar].
- g_3 : Beta [a minimizar].
- g_4 : Dividendo/ Preço [a maximizar].
- g_5 : Crescimento do dividendo[a maximizar]
- g_6 : "Teste ácido"[a maximizar].

Hurson e Zopounidis aplicaram o ELECTRE TRI para afectar as vinte alternativas a três categorias: C^3 ("a rejeitar") é a pior categoria, a seguir C^2 ("a averiguar") e por fim C^1 ("a aceitar") que é a melhor categoria. No entanto, o nosso objectivo é afectar as vinte alternativas ignorando a estrutura usada por Hurson e Zopounidis e partindo do princípio que não é conhecida a priori. Para tal aplicamos a abordagem Clustering-Sorting que pode sugerir uma estrutura final alternativa com categorias incomparáveis, uma vez que, tal como veremos de seguida, o número de alternativas incomparáveis é bastante elevado, fazendo todo o sentido poderem existir categorias que sejam incomparáveis.

	g_1	g_2	g_3	g_4	g_5	g_6		g_1	g_2	g_3	g_4	g_5	g_6
a_0	0,82	0,45	0,26	-4,7	-100	0,45	a_{10}	0,8	0,58	0,62	13,7	34,6	1,54
a_1	0,41	0,63	0,03	2,28	-20	2,04	a_{11}	1,23	0,37	0,64	8,97	45,9	0,96
a_2	0,57	0,2	0,1	6,08	-33,3	1,08	a_{12}	0,24	0,28	0,73	-1,75	0	0,72
a_3	0,24	0,02	0,08	2,41	-53,5	0,62	a_{13}	0,26	0,65	0,58	4,88	7,14	0,9
a_4	0	0,46	0,62	5,04	-76,5	3,02	a_{14}	1,1	0,76	0,54	0,29	0	0,73
a_5	0,93	0,02	0,14	2,82	6,38	0,72	a_{15}	1,79	0,55	0,73	5,88	-100	2,69
a_6	0,01	0,69	0,77	7,55	-40	3,23	a_{16}	1,02	1,06	0,82	5,5	6,38	0,73
a_7	0,86	0,86	0,86	4,28	3,71	0,57	a_{17}	1,32	1,12	0,94	12,06	-61	2,69
a_8	2,16	0,6	0,12	2,11	56,3	0,51	a_{18}	1,36	0,04	1,02	1,79	110	2,31
a_9	1,24	0,12	0,62	11,65	12,5	1,17	a_{19}	0,57	0,17	0,23	-11,5	0	0,52

Tabela 5.2: Desempenhos das vinte alternativas do sector comercial

1. Primeira fase - Clustering

Na primeira fase, as alternativas são agrupadas em categorias homogêneas, à custa do método não hierárquico *K-Means*, fazendo-se previamente a análise do número de clusters através da aplicação do critério da distância entre clusters (quadrado da distância Euclidiana) aplicando o método *Complete Linkage* (Figura 5.2). A distância entre grupos é relativizada (entre 0 e 1).

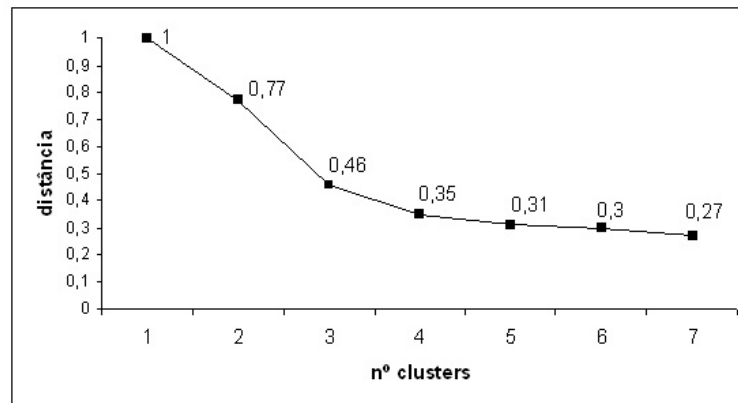


Figura 5.2: Distância entre clusters.

Observou-se que os declives das rectas que unem as distâncias entre dois clusters é relativamente pequeno a partir de 4 clusters. Na ausência de um decisor, aceita-se a solução com o menor número de clusters, ou seja, 4 clusters. Os clusters resultantes são formados pelas seguintes alternativas de acordo com o método K-Means:

- $C^1 = \{a_0, a_1, a_2, a_3, a_5, a_{12}, a_{19}\}$,
- $C^2 = \{a_9, a_{10}, a_{11}, a_{18}\}$,
- $C^3 = \{a_4, a_6, a_{15}, a_{17}\}$
- $C^4 = \{a_7, a_8, a_{13}, a_{14}, a_{16}\}$.

2. Segunda fase- Eliciação

Na segunda fase, vamos usar os valores originais (Hurson e Zopounidis, 1997) para os limiares de indiferença e preferência (Tabela 5.3), mas não vamos usar limiares de veto tão apertados como os originais. Quanto ao limiar de corte λ , este assume o valor 0,6.

Uma vez que os dados originais (Hurson e Zopounidis, 1997) não indicam qualquer informação sobre o peso dos critérios, consideramos que todos os critérios têm o mesmo peso, ou seja, $k_i = 1/6, i \in \{1, \dots, 6\}$.

	g_1	g_2	g_3	g_4	g_5	g_6
q_j	0.05	0.05	0	0.1	8.72	0.05
p_j	0.25	0.2	0.2	0.5	10	0.25
v_j	20	10	10	100	180	2.75

Tabela 5.3: Limiares de Indiferença, Preferência e Veto.

Analisando o grupo de alternativas segundo os seus desempenhos, coeficientes de importância, limiares de indiferença, preferência e veto, e para $\lambda = 0,6$, as relações de preferência entre as alternativas foram obtidas à custa das suas relações de prevalência (Tabela 5.4) obtidas pela aplicação do método ELECTRE. Dos 190 pares de alternativas estudadas ($\frac{n(n-1)}{2}$, $n = 20$), 121 (63,7%) verificam a relação de Preferência, 8 (4,2%) são Indiferentes, e 61(32,1%) são Incomparáveis.

3. Terceira fase- Classificação Ordinal

Na terceira fase, para ordenação das categorias, apresentamos as duas sugestões descritas na Secção anterior: ordenação dos centróides das quatro categorias obtidas na primeira fase e aplicação da relação de preferência mais comum.

Para tal, calculamos o centróide de cada uma das categorias obtidas (c_1 , c_2 , c_3 e c_4), e, em seguida, classificamos os centróides com base nas suas relações de prevalência (Tabela 5.5). Assim, para $\lambda = 0,6$ temos $c_2 P c_4$, $c_2 P c_1$, $c_4 P c_1$, $c_3 P c_1$, $c_3 R c_2$, e $c_3 R c_4$. A estrutura final obtida é pois de ordem parcial: $C^2 \succ C^4$, $C^2 \succ C^1$, $C^4 \succ C^1$, $C^3 \succ C^1$, $C^2 \perp C^3$ e $C^3 \perp C^4$ (Figura 5.3).

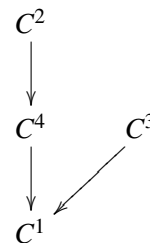


Figura 5.3: Partição final

Aplicando agora a relação de preferência mais comum entre as alternativas das categorias (Tabela 5.6) obtém-se a mesma estrutura final (Figura 5.3).

a_1	0	0.19	0.37	0.42	0.07	0.35	0	0.44	0.14	0.33	0.27	0.2	0.47	0.33	0.17	0.2	0.21	0.17	0	0.47
a_2	0.83	0	0.74	0.99	0.67	0.5	0.62	0.33	0.61	0.5	0.5	0.5	0.83	0.67	0.58	0.5	0.33	0.33	0.29	0.74
a_3	0.67	0.44	0	0.98	0.32	0.67	0.67	0.24	0.5	0.5	0.47	0.18	0.37	0.8	0.67	0.5	0.46	0.5	0.33	0.83
a_4	0.67	0.36	0.19	0	0.14	0.5	0.06	0.33	0.41	0.28	0.17	0.17	0.63	0.33	0.45	0.27	0.28	0.27	0.1	0.56
a_5	0.67	0.37	0.33	0.51	0	0.5	0.37	0.5	0.28	0.5	0.41	0.34	0.61	0.48	0.43	0.62	0.35	0.33	0	0.5
a_6	0.83	0.57	0.49	0.95	0.18	0	0.1	0.67	0.48	0.44	0.33	0.18	0.83	0.56	0.73	0.31	0.63	0.31	0.45	0.89
a_7	0.67	0.5	0.67	0.68	0.88	0.5	0	0.53	0.49	0.38	0.38	0.39	0.65	0.51	0.48	0.8	0.5	0.5	0.18	0.5
a_8	0.83	0.51	0.5	0.83	0.12	0.73	0.04	0	0.5	0.32	0.33	0.17	0.81	0.5	0.58	0.25	0.45	0.25	0.4	0.83
a_9	1	0.49	0.65	0.83	0.1	0.7	0.01	0.66	0	0.67	0.67	0.67	0.87	0.67	0.74	0.23	0.53	0.23	0.38	1
a_{10}	0.67	0.5	0.8	0.83	0.36	0.83	0.28	0.83	0.33	0	0.33	0.67	0.88	0.8	0.77	0.49	0.83	0.49	0.49	0.83
a_{11}	0.83	0.67	0.83	0.83	0.51	0.77	0.42	0.83	0.5	0.83	0	0.67	1	0.94	0.62	0.64	0.69	0.5	0.5	0.83
a_{12}	0.8	0.5	0.78	0.83	0.28	0.83	0.19	0.83	0.33	0.68	0.48	0	1	0.78	0.75	0.41	0.83	0.41	0.56	0.83
a_{13}	0.53	0.23	0.33	0.67	0.18	0.5	0.1	0.5	0.17	0.24	0.08	0.22	0	0.43	0.34	0.31	0.5	0.31	0.33	0.67
a_{14}	0.67	0.58	0.39	0.83	0.25	0.67	0.17	0.67	0.5	0.5	0.33	0.49	1	0	0.73	0.38	0.5	0.33	0.45	0.67
a_{15}	0.83	0.5	0.5	0.67	0.18	0.67	0.1	0.78	0.33	0.43	0.5	0.45	1	0.73	0	0.32	0.67	0.32	0.33	0.83
a_{16}	0.83	0.59	0.63	0.67	0.58	0.43	0.4	0.45	0.14	0.4	0.27	0.2	0.47	0.43	0.47	0	0.43	0.5	0	0.47
a_{17}	0.83	0.58	0.5	0.83	0.18	0.83	0.1	1	0.5	0.36	0.33	0.23	0.93	0.73	0.81	0.32	0	0.32	0.45	0.83
a_{18}	0.83	0.67	0.67	0.83	0.67	0.66	0.53	0.68	0.37	0.63	0.5	0.43	0.67	0.66	0.67	0.67	0.66	0	0.05	0.67
a_{19}	0.67	0.5	0.54	0.67	0.33	0.67	0.33	0.53	0.41	0.63	0.5	0.5	0.67	0.5	0.67	0.17	0.5	0.43	0	0.74
a_{20}	0.5	0.33	0.56	0.67	0.1	0.47	0.02	0.5	0.24	0.33	0.18	0.17	0.64	0.5	0.37	0.23	0.37	0.23	0.33	0

Tabela 5.4: Grau de prevalência entre as ações (a_i, a_j) , $i, j=1, \dots, 20$

	c_1	c_2	c_3	c_4
c_1	-	0.3333	0.2884	0.3333
c_2	0.8333	-	0.5348	0.7158
c_3	0.6667	0.3522	-	0.4814
c_4	0.7158	0.4767	0.2120	-

Tabela 5.5: Grau de Credibilidade

$C^2 \succ C^1$	$C^1 \mathbf{P} C^2: 0$	$C^2 \mathbf{P} C^1: 24$	$C^1 \mathbf{R} C^2: 4$	$C^1 \mathbf{I} C^2: 0$
$C^3 \succ C^1$	$C^1 \mathbf{P} C^3: 2$	$C^3 \mathbf{P} C^1: 16$	$C^1 \mathbf{R} C^3: 10$	$C^1 \mathbf{I} C^3: 0$
$C^4 \succ C^1$	$C^1 \mathbf{P} C^4: 3$	$C^4 \mathbf{P} C^1: 23$	$C^1 \mathbf{R} C^4: 6$	$C^1 \mathbf{I} C^4: 3$
$C^2 \perp C^3$	$C^2 \mathbf{P} C^3: 1$	$C^3 \mathbf{P} C^2: 1$	$C^2 \mathbf{R} C^3: 14$	$C^2 \mathbf{I} C^3: 0$
$C^2 \succ C^4$	$C^2 \mathbf{P} C^4: 13$	$C^4 \mathbf{P} C^2: 3$	$C^2 \mathbf{R} C^4: 4$	$C^2 \mathbf{I} C^4: 0$
$C^4 \perp C^3$	$C^4 \mathbf{P} C^3: 0$	$C^3 \mathbf{P} C^4: 4$	$C^4 \mathbf{R} C^3: 16$	$C^4 \mathbf{I} C^3: 0$

Tabela 5.6: Número de pares de alternativas para cada relação de preferência.

4. Quarta fase- Qualidade

Depois de obter a partição final, analisamos a percentagem de pares de alternativas incoerentes com o princípio $SS\tau$ -Consistency. Os resultados são apresentados na Tabela 5.7.

Condições	Relação S	Clustering-Sorting
Preferência	121 (63.7%)	8 (4.2%)
Indiferença	8 (4.2%)	0(0.0%)
Incomparabilidade	61(32.1%)	7 (3.7%)
Total	190	15 (7.9%)
$Q(\mathcal{P})$	-	92.10%

Tabela 5.7: Número de pares de alternativas que verificam as condições de P, I e R antes da partição e o número de inconsistências pós-partição segundo o $SS\tau$ -Consistency

A qualidade $Q(\mathcal{P})$ da partição \mathcal{P} obtida com o *Clustering-Sorting* é dada por:

$$Q(\mathcal{P}) = 1 - \alpha_{PV_P} - \alpha_{IV_I} - \alpha_{RV_R} = 1 - 0.042 - 0.0 - 0.037 = 0.921$$

5.4.2 Segundo exemplo ilustrativo

O segundo exemplo é baseado nos dados de um exemplo de Yu (1992), sobre a avaliação de 100 alternativas a serem classificadas, com base nos seus desempenhos, em 7 critérios a minimizar. Tal como no exemplo anterior, neste exemplo Yu utiliza o ELECTRE TRI para classificar as alternativas em duas categorias pré-ordenadas. Aqui apresentaremos uma abor-

dagem *Clustering-Sorting* que pode sugerir uma estrutura alternativa de categorias exibindo incomparabilidades, já que, tal como no exemplo anterior, o número de alternativas incomparáveis é elevado.

1. Primeira fase - Clustering

O gráfico das distâncias inter-clusters (obtido com o método *Complete Linkage*) fornece informação ao decisor sobre o número de clusters a ser considerado (Figura 5.4), para poder aplicar o método *K-Means*.

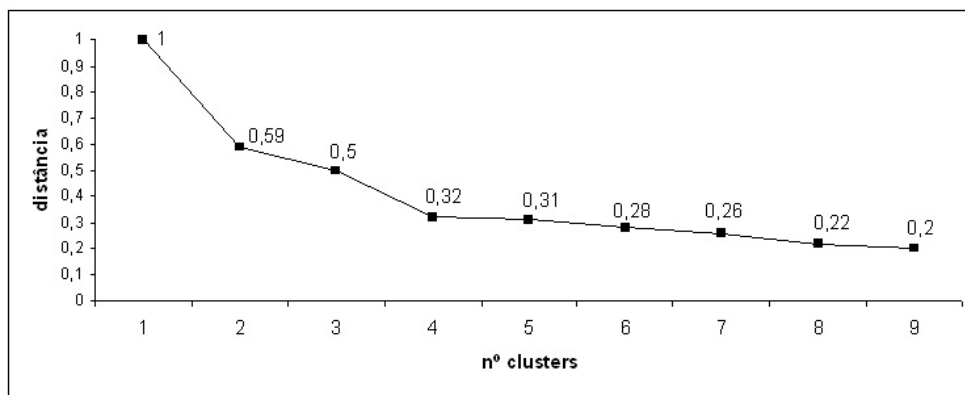


Figura 5.4: Distância entre clusters.

Tal como se pode ver na Figura 5.4, devemos ter pelo menos 4 clusters, já que os declives das rectas que unem as distâncias entre dois clusters são relativamente pequenos a partir de 4 clusters. Considerando $k=4$, o algoritmo *K-Means* forneceu 4 grupos homogéneos - C^1, C^2, C^3 e C^4 - com 17, 63, 16 e 4 alternativas respectivamente:

- $C^1 = \{a_{20}, a_{22}, a_{23}, a_{24}, a_{25}, a_{50}, a_{64}, a_{65}, a_{72}, a_{73}, a_{74}, a_{75}, a_{81}, a_{82}, a_{86}, a_{87}, a_{89}\}$
- $C^2 = \{a_1, a_2, a_3, a_4, a_6, a_7, a_8, a_9, a_{10}, a_{11}, a_{12}, a_{13}, a_{14}, a_{15}, a_{16}, a_{17}, a_{26}, a_{27}, a_{28}, a_{29}, a_{30}, a_{31}, a_{32}, a_{33}, a_{34}, a_{35}, a_{36}, a_{37}, a_{38}, a_{39}, a_{40}, a_{41}, a_{42}, a_{43}, a_{44}, a_{45}, a_{46}, a_{47}, a_{48}, a_{49}, a_{51}, a_{52}, a_{53}, a_{54}, a_{55}, a_{56}, a_{57}, a_{58}, a_{59}, a_{60}, a_{61}, a_{62}, a_{63}, a_{66}, a_{67}, a_{68}, a_{69}, a_{70}, a_{71}, a_{76}, a_{78}, a_{79}, a_{83}\}$
- $C^3 = \{a_{90}, a_{91}, a_{77}, a_{80}, a_{84}, a_{85}, a_{88}, a_{92}, a_{93}, a_{94}, a_{95}, a_{96}, a_{97}, a_{98}, a_{99}, a_{100}\}$
- $C^4 = \{a_5, a_{21}, a_{18}, a_{19}\}$.

2. Segunda fase- Eliciação

Na segunda fase, vamos usar os valores originais (Yu, 1992) para os pesos e limiares de indiferença e preferência associados a cada critério, mas à semelhança do exemplo anterior não vamos usar limiares de veto tão apertados como os originais (Tabela 5.8). O limiar de corte usado é $\lambda=0,6$.

	g_1	g_2	g_3	g_4	g_5	g_6	g_7
q_j	0.65	1.7	1.7	1.7	1.7	1.7	2.3
p_j	1.31	3.45	3.45	3.45	3.45	3.45	4.7
v_j	1.95	100.4	100.4	100.4	100.4	100.5	18.7
k_j	0.24	0.12	0.12	0.12	0.12	0.12	0.16

Tabela 5.8: Limiares e pesos associados aos critérios.

Analisando o grupo de alternativas segundo os seus desempenhos, coeficientes de importância e limiares, dos 4950 pares de alternativas estudadas ($\frac{n(n-1)}{2}, n = 100$), 2967 (59.9%) verificam a relação de Preferência, 364 (7.4%) são Indiferentes, e 1619 (32.7%) são Incomparáveis.

3. Terceira fase- Classificação Ordinal

Na terceira fase, calculamos o centróide de cada uma das categorias obtidas, e, em seguida, classificamos os centróides c_1, c_2, c_3, c_4 com base nas suas relações de prevalência (Tabela 5.9), dando origem a uma estrutura parcialmente ordenada das categorias (Figura 5.5).

centróides	c_1	c_2	c_3	c_4
c_1	-	0.4454	0.0	0.6281
c_2	0.0	-	0.0	1.0
c_3	1.0	0.9575	-	1.0
c_4	0	0.2186	0.0	-

Tabela 5.9: Grau de credibilidade entre centróides.

Definindo a estrutura de preferências entre as categorias à custa da relação de preferência mais comum entre as alternativas das categorias (Tabela 5.10), obtém-se a estrutura apresentada na Figura 5.6. Note-se que a relação de incomparabilidade entre C^1 e C^4 é definida não só pela relação com um maior número de alternativas, mas também pela relação que leva à melhor qualidade da partição resultante, dado que houve um empate entre as relações de Preferência e Incomparabilidade.

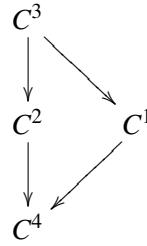


Figura 5.5: Partição final obtida com o método (A)

$C^3 \succ C^1$	$C^1 \mathbf{P} C^3: 10$	$C^3 \mathbf{P} C^1: 135$	$C^1 \mathbf{R} C^3: 123$	$C^1 \mathbf{I} C^3: 4$
$C^3 \succ C^2$	$C^2 \mathbf{P} C^3: 4$	$C^3 \mathbf{P} C^2: 629$	$C^2 \mathbf{R} C^3: 374$	$C^2 \mathbf{I} C^3: 1$
$C^3 \succ C^4$	$C^3 \mathbf{P} C^4: 34$	$C^4 \mathbf{P} C^3: 0$	$C^3 \mathbf{R} C^4: 30$	$C^3 \mathbf{I} C^4: 0$
$C^2 \perp C^1$	$C^1 \mathbf{P} C^2: 376$	$C^2 \mathbf{P} C^1: 117$	$C^1 \mathbf{R} C^2: 547$	$C^1 \mathbf{I} C^2: 31$
$C^2 \perp C^4$	$C^2 \mathbf{P} C^4: 77$	$C^4 \mathbf{P} C^2: 25$	$C^2 \mathbf{R} C^4: 147$	$C^2 \mathbf{I} C^4: 3$
$C^1 \perp C^4$	$C^1 \mathbf{P} C^4: 31$	$C^4 \mathbf{P} C^1: 5$	$C^4 \mathbf{R} C^1: 31$	$C^4 \mathbf{I} C^1: 1$

Tabela 5.10: Número de pares de alternativas para cada relação de preferência.

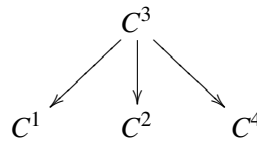


Figura 5.6: Partição final obtida com o método (B)

4. Quarta fase - Qualidade

Após a obtenção das partições com os métodos (A) e (B), analisamos a percentagem de pares de alternativas incoerentes com o princípio $SS\tau$ -Consistency em cada um dos casos. As estruturas obtidas com os métodos (A) e (B), embora diferentes, conduzem a resultados muito idênticos como mostra a Tabela 5.11.

Perante duas estruturas diferentes com uma robustez muito idêntica, a escolha da estrutura final deve pertencer ao decisor, optando pela estrutura que faça mais sentido perante o problema em estudo.

Considerando como solução do problema a estrutura obtida com o método (B), a qualidade $Q(\mathcal{P})$ da partição \mathcal{P} é dada por:

$$Q(\mathcal{P}) = 1 - \alpha_{PV_P} - \alpha_{IV_I} - \alpha_{RV_R} = 1 - 0.008 - 0.006 - 0.074 = 0.912.$$

Condições	Relação S	C-S com (A)	C-S com (B)
Preferência	2967 (59.9%)	44 (0.9%)	39 (0.8%)
Indiferença	364 (7.4%)	31 (0.6%)	32 (0.6%)
Incomparabilidade	1619 (32.7%)	367 (7.4%)	367(7.4%)
Total	4950	442 (8.9%)	438 (8.8%)
$Q(\mathcal{P})$	-	91.1%	91.2%

Tabela 5.11: Número de pares de alternativas que verificam as condições de P, I e R antes da partição e o número de inconsistências pós-partição segundo o $SS\tau$ -Consistency

Analisando agora o grupo de alternativas segundo os seus desempenhos, coeficientes de importância e limiares originais ($v_j=(1.95, 100.4, 100.4, 100.4, 100.4, 100.5, 18.7)$), bem como o limiar de corte ($\lambda=0.86$), dos 4950 pares de alternativas estudadas, 1803 (36.42%) verificam a relação de Preferência, 8 (0.16%) são Indiferentes, e 3139 (63.41%) são Incomparáveis. Neste caso, com o dobro das incomparabilidades, a qualidade da estrutura obtida é bastante menor (68.18%). Tal estrutura foi obtida com o método (B) (a obtida com o método (A) era bastante pior), após análise das relações de preferência apresentadas na Tabela 5.12. A estrutura obtida está representada na Figura 5.7.

$C^1 \perp C^2$	$C^1 \mathbf{P} C^2$: 314	$C^2 \mathbf{P} C^1$: 77	$C^1 \mathbf{R} C^2$: 679	$C^1 \mathbf{I} C^2$: 1
$C^1 \perp C^3$	$C^1 \mathbf{P} C^3$: 5	$C^3 \mathbf{P} C^1$: 121	$C^1 \mathbf{R} C^3$: 146	$C^1 \mathbf{I} C^3$: 0
$C^1 \perp C^4$	$C^1 \mathbf{P} C^4$: 28	$C^4 \mathbf{P} C^1$: 2	$C^1 \mathbf{R} C^4$: 38	$C^1 \mathbf{I} C^4$: 0
$C^3 \succ C^2$	$C^2 \mathbf{P} C^3$: 2	$C^3 \mathbf{P} C^2$: 568	$C^2 \mathbf{R} C^3$: 488	$C^2 \mathbf{I} C^3$: 0
$C^2 \perp C^4$	$C^2 \mathbf{P} C^4$: 56	$C^4 \mathbf{P} C^2$: 17	$C^2 \mathbf{R} C^4$: 179	$C^2 \mathbf{I} C^4$: 0
$C^3 \perp C^4$	$C^3 \mathbf{P} C^4$: 30	$C^4 \mathbf{P} C^3$: 0	$C^4 \mathbf{R} C^3$: 34	$C^4 \mathbf{I} C^3$: 0

Tabela 5.12: Número de pares de alternativas para cada relação de preferência.

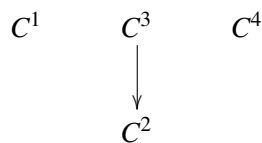


Figura 5.7: Partição final obtida com o método (B)

5.5 Notas conclusivas

Neste trabalho, propusemos uma nova abordagem para a classificação de um conjunto \mathcal{A} de alternativas, com base em vários critérios, a um conjunto de categorias parcialmente ordenado, com uma estrutura desconhecida a priori. Numa primeira fase, as alternativas semelhantes são atribuídas às categorias homogêneas, usando o método *K-Means*. Para determinar o número de categorias a formar, podem ser utilizados dois indicadores - distância entre clusters e índice *R-squared*.

Em seguida, as categorias são classificadas com base nos seus centróides ou na relação de preferência mais comum entre as alternativas das categorias, usando um método de classificação ordinal, como o ELECTRE, com base nas preferências que o decisor suscitou nesta fase. Após a obtenção de uma partição do conjunto de alternativas com base numa das abordagens propostas, pode-se analisar a sua qualidade: propusemos o índice $Q(\mathcal{P})$ que é um indicador dos pares de alternativas coerentes com o princípio *SS τ -Consistency*. Este princípio proíbe a existência de alternativas incomparáveis dentro da mesma classe, alternativas indiferentes em classes incomparáveis e alternativas melhores que outras alternativas e pertencentes a classes piores que as últimas (preferências invertidas). Para o cálculo da qualidade de cada partição foram considerados três indicadores distintos: v_P indica o rácio de pares de alternativas com preferência invertida, v_I indica o rácio de pares de alternativas de grupos incomparáveis e v_R indica o rácio de pares de alternativas incomparáveis dum mesmo grupo. O decisor pode razoavelmente decidir que algumas violações são mais graves do que outras, podendo colocar mais peso sobre essas violações.

Este método foi testado em dois exemplos reais da literatura. Em ambos os exemplos, o método *Clustering-Sorting* apresentou estruturas parcialmente ordenadas alternativas às estruturas totalmente ordenadas sugeridas na literatura e com resultados para a coerência segundo o princípio *SS τ -Consistency* bastante bons.

Esta abordagem *Clustering-Sorting*, embora desenvolvida para ser bastante simples, na verdade pode conduzir a resultados menos satisfatórios para as condições de Preferência e Indiferença uma vez que, numa primeira fase, as alternativas são avaliadas por atributos e não critérios. Para além disso, não garante uma partição “ótima” de minimização das incoerências do princípio *SS τ -Consistency*. O desenvolvimento de métodos mais sofisticados de optimização utilizando programação matemática ou heurísticas para coincidir com as preferências de um decisor é pois o tópico importante a desenvolver nos próximos capítulos.

CHOP - Clustering Hierárquico com Ordenação Parcial

6.1 Notas introdutórias

O objectivo deste capítulo consiste em apresentar um modelo que é uma *extensão do método hierárquico aglomerativo* (ver detalhes na secção 3.3.2). De um modo geral, os métodos aglomerativos utilizam um algoritmo padrão, conforme descrito no Algoritmo 1 (ver Secção 3.3.2). De forma bastante sumária, o método inicia com m clusters, cada um contendo uma das m alternativas de um conjunto \mathcal{A} de alternativas a atribuir. De seguida, e de forma iterativa, as alternativas cuja similaridade seja bastante elevada são agrupadas num mesmo cluster homogéneo. A diferença entre os métodos ocorre na definição de similaridade entre clusters a qual é definida de acordo com cada método.

Neste trabalho, para quantificar a medida de similaridade, recorreremos à estrutura de preferências (Preferência, Indiferença e Incomparabilidade) definida pelo decisor. Propomos, no entanto, uma nova medida de similaridade: dois grupos são tanto mais similares quanto maior for a qualidade da partição obtida com a sua junção, ou seja, quanto menor for o número de violações das relações de Preferência, Indiferença e Incomparabilidade entre as alternativas classificadas. Para a avaliação da qualidade da partição será usado o princípio *Semi-Strong consistency* ($SS\tau$ -Consistency) apresentado no Capítulo 4, o qual proíbe a existência de alternativas incomparáveis dentro da mesma classe, alternativas indiferentes em classes incomparáveis e alternativas melhores que outras alternativas e pertencentes a classes piores que as últimas (a que chamaremos preferências invertidas).

Este trabalho distingue-se dos trabalhos clustering multicritério que se têm desenvolvido, não só na definição de similaridade mas também pela estrutura que se pretende obter, a qual pode ser *parcialmente ordenada* e não necessariamente totalmente ordenada. Apresentamos

duas formas de construção de tal estrutura: uma definida à custa das relações binárias prevalência dos centróides representativos das classes e outra à custa das relações binárias prevalência das alternativas atribuídas. A comparação e contabilização das relações de todos os pares de alternativas pertencentes a cada par de classes da partição é a proposta para a construção da estrutura de preferências em cada iteração do Algoritmo 1.

O objectivo do método que é aqui apresentado é obter uma partição transitiva de um dado conjunto de alternativas tendo informação da matriz dos valores de prevalência das alternativas mas não tendo qualquer informação a priori sobre a estrutura de preferências das classes.

Note-se que as mesmas ideias se poderiam aplicar, por exemplo, com o princípio $S\tau$ -Consistency. Já os princípios S-Consistency e τ -Consistency, pelas razões apresentadas no capítulo anterior, consideramo-los de facto menos interessantes.¹

6.2 Estrutura de preferências das classes

Neste trabalho apresentamos duas formas distintas de obter a estrutura de preferências das classes: uma obtida à custa das relações de prevalência entre os centróides das classes e outra definida pela relação de prevalência das alternativas pertencentes às classes. Passaremos a descrever cada uma das construções das estruturas de preferências.

6.2.1 Estrutura de Preferências 1 - definida à custa dos centróides das classes

Cada classe pode ser representada por um ponto cujas coordenadas são a média de cada um dos critérios dos sujeitos que fazem parte da classe, ou seja, pelo seu *centróide*. Assim, em cada iteração s do Algoritmo 1 (ver secção 3.3.2), obtém-se uma partição $\mathcal{P}_s = \{C^1, C^2, \dots, C^r\}$ das alternativas de \mathcal{A} e um conjunto $\mathcal{C}_s = \{c_1, c_2, \dots, c_r\}$ dos centróides que caracterizam as classes de \mathcal{P}_s .

Definição 6.2.1 *Dada uma qualquer classe C^k com n_k alternativas $a_i = (a_{i1}, a_{i2}, \dots, a_{in})$ avaliadas em n critérios, o seu centróide $c_k = (c_{k1}, c_{k2}, \dots, c_{kn})$ é definido da seguinte forma:*

¹Foram realizadas algumas experiências, que aqui não são relatadas, com os princípios S-Consistency e τ -Consistency. Em tais experiências verificamos que o princípio S-Consistency conduz a partições com uma das categorias com quase todas as alternativas. Já o princípio τ -Consistency conduz a partições com muitas categorias incomparáveis.

$$c_{kj} = \frac{1}{n_k} \sum_{i=1}^{n_k} a_{ij}, \quad \forall j = 1 \dots n$$

A estrutura de preferências de cada partição, obtida nas iterações do Algoritmo 1, é aqui obtida à custa das relações $\{P, I, R\}$ dos centróides, as quais são definidas pelas relações de prevalência S de acordo com (2.2), podendo-se obter as seguintes situações:

$$\left\{ \begin{array}{l} C^r \succ C^t \quad \text{se } c_r P c_t \\ C^t \succ C^r \quad \text{se } c_t P c_r \\ C^t \perp C^r \quad \text{se } c_t R c_r \\ C^t \approx C^r \quad \text{se } c_t I c_r \end{array} \right. \quad (6.1)$$

Assim, a classe C^i é preferível a C^j ($C^i \succ C^j$) quando o centróide c_i da classe C^i for preferível ao centróide c_j da classe C^j , ou seja, quando $c_i S c_j$ e $c_j \not S c_i$. Por outro lado, as classes C^i e C^j são indiferentes ($C^i \approx C^j$) quando os seus centróides são indiferentes, ou seja, quando $c_i S c_j$ e $c_j S c_i$. Por fim, a classe C^i é incomparável a C^j ($C^i \perp C^j$) quando os seus centróides são incomparáveis, ou seja, quando $c_i \not S c_j$ e $c_j \not S c_i$.

6.2.2 Estrutura de Preferências 2 - definida à custa das relações de prevalência entre classes

Uma outra forma proposta para a construção da estrutura de preferências é definir para cada par de classes $(C^i, C^j) \in \mathcal{P}^2$ a sua relação de prevalência. Considerar-se-á que uma classe C^i prevalece sobre C^j se pelo menos 50% das relações dos pares de alternativas $(a_s, a_t) \in (C^i, C^j)$ forem de prevalência de a_s sobre a_t .

Definição 6.2.2 *Seja $C^i, C^j \in \mathcal{P}$, onde n_i e n_j são o número de alternativas de C^i e C^j respetivamente. Diz-se que o grau de prevalência de C^i sobre C^j é*

$$s_{ij} = \frac{\sum_{a_s \in C^i} \sum_{a_t \in C^j} S(a_s, a_t)}{n_i \times n_j}, \quad \text{com } S(a_s, a_t) = \begin{cases} 1 & \text{se } a_s S a_t \\ 0 & \text{caso contrário} \end{cases}.$$

Definição 6.2.3 *Seja dado o par $(C^i, C^j) \in \mathcal{P}^2$. Diz-se que C^i **prevalece sobre** C^j ($C^i \tau C^j$) se e só se o grau de prevalência de C^i sobre C^j é de pelo menos 0.5, ou seja, $s_{ij} \geq 0.5$.*

Assim, a estrutura de preferências de cada partição, definida à custa das relações de prevalência entre classes, pode ser definida da seguinte forma:

$$\left\{ \begin{array}{l} C^i \succ C^j \iff C^i \tau C^j \wedge C^j \not\tau C^i \\ C^j \succ C^i \iff C^j \tau C^i \wedge C^i \not\tau C^j \\ C^i \approx C^j \iff C^i \tau C^j \wedge C^j \tau C^i \\ C^i \perp C^j \iff C^i \not\tau C^j \wedge C^j \not\tau C^i \end{array} \right. \quad (6.2)$$

A classe C^i é preferível a C^j quando o grau de prevalência de C^i sobre C^j é de pelo menos 50% e o grau de prevalência de C^j sobre C^i é inferior a 50%. As classes C^i e C^j são indiferentes quando o grau de prevalência de C^i sobre C^j e de C^j sobre C^i são de pelo menos 50%. Por fim, as classe C^i e C^j são incomparáveis quando ambos os graus de prevalência são inferiores a 50%.

Exemplo 6.2.1 *Consideremos a partição $\mathcal{P} = \{C^1, C^2, C^3, C^4, C^5\}$ cujas relações de prevalência são apresentadas na Tabela 6.1.*

	C^1	C^2	C^3	C^4	C^5
C^1	0	0.3333	0.2000	0.1333	0.1000
C^2	0.3333	0	0.6	0	0.1250
C^3	0.4000	0.5833	0	0.3333	0.3750
C^4	0.4667	0.4222	0	0	0.5833
C^5	0.4000	0.9167	0	0	0

Tabela 6.1: Valores do grau de prevalência s_{ij} entre as classes (C^i, C^j) , $i, j=1, \dots, 5$

A estrutura de preferências resultante é apresentada na Figura 6.1.

Independentemente da forma como se constrói a estrutura de preferências das classes, o objectivo deste modelo é obter uma partição final cuja estrutura de preferências não tenha classes indiferentes e seja transitiva. Tal é conseguido fazendo uma revalidação da partição obtida,

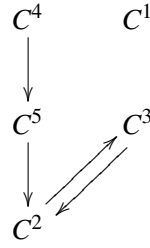


Figura 6.1: Partição final

tal como será apresentado na Secção 6.4. Assim, a partição final transitiva e parcialmente ordenada que se pretende obter, tem uma estrutura de preferências (\succ, \perp) definida da seguinte forma:

- $C^i \succ C^j \Rightarrow C^j \not\succeq C^i$: \succ é assimétrica
- $C^i \succ C^j \wedge C^j \succ C^k \Rightarrow C^i \succ C^k$: \succ é transitiva
- $C^i \perp C^j \Rightarrow C^j \perp C^i$: \perp é simétrica
- $C^i \perp C^j \vee C^i \succ C^j \vee C^j \succ C^i$: \approx é vazia

6.3 Medida de similaridade entre classes

Dada a natureza multicritério do tipo de problemas para os quais se pretende desenvolver o modelo aqui desenvolvido, o Algoritmo 1 terá em conta a natureza multicritério das alternativas. Para tal, em cada iteração s do algoritmo considerar-se-á que o par de classes mais similar da partição \mathcal{P}_s será aquele que ao ser fundido leva à melhor partição \mathcal{P}_{s+1} de \mathcal{A} , ou seja, aquele que leva ao menor número de pares de alternativas incoerentes com o princípio $SS\tau$ -Consistency (ver Secção 4.2.2).

Será assim definido um *vector de inconsistências* $ID_{\mathcal{P}_s} = (v_P, v_I, v_R)$ associado à partição \mathcal{P}_s , onde v_R (dado por (5.3)) mede o rácio de pares que não verificam a condição de incomparabilidade, v_I (dado por (5.4)) mede o rácio de pares que não verificam a condição de indiferença e v_P (dado por (5.5)) mede o rácio de pares que não verificam a condição de preferência (ver Secção 5.3). O objectivo é minimizar este vector, ou seja, minimizar o rácio de pares de alternativas que são incompatíveis com as condições de Preferência, Indiferença e Incomparabilidade. Para tal, é definido o índice $Q(\mathcal{P}_s)$, dado por (5.7). Quanto maior é o valor de $Q(\mathcal{P}_s)$, melhor é a partição \mathcal{P}_s .

Assim, o método aglomerativo que propomos tem por objectivo encontrar, em cada iteração s do Algoritmo 1, a partição parcialmente ordenada \mathcal{P}_s^* tal que:

$$Q(\mathcal{P}_s^*) > Q(\mathcal{P}_s), \forall \mathcal{P}_s \text{ determinada na iteração } s.$$

Exemplo 6.3.1 *Sejam \mathcal{P}_1 e \mathcal{P}_2 duas partições cujos vectores de índices de desajustamento (v_P, v_I, v_R) são respectivamente $(0.03, 0.073, 0.03)$ e $(0.03, 0.083, 0.01)$.*

Considerando $\alpha_P = \alpha_I = \alpha_R = \alpha_D = 1$ tem-se $Q(\mathcal{P}_1) = 0.876$ e $Q(\mathcal{P}_2) = 0.877$. Assim, a melhor partição é \mathcal{P}_2 pois é a que tem uma melhor qualidade.

6.4 Intransitividades da partição

Geralmente, uma relação de prevalência é um mecanismo para modelar e representar as preferências de um decisor. Esta relação proporciona modelos e representação de situações onde a transitividade não se verifica. Um exemplo bem conhecido de ausência de transitividade é o apresentado por Luce(1956) (ver Roy e Vincke, 1981): obviamente que não conseguimos diferenciar um café com αgr de açúcar e um com $\alpha + 0.01 gr$; assim, existe uma relação de indiferença entre estas duas situações. De forma similar, existe uma relação de indiferença entre $\alpha + 0.01 gr$ e $\alpha + 0.02 gr$ de açúcar. Se a relação de indiferença é transitiva, então αgr e $\alpha + 0.02 gr$ de açúcar devem ser consideradas indiferentes. Seguindo este raciocínio de inferência, pode-se deduzir que não há diferença entre um café com αgr de açúcar e um café com a chávena cheia de açúcar, independentemente de α . Obviamente que é uma conclusão errada, mostrando que de facto existem situações onde a transitividade não se verifica.

Formalmente, considerando três quaisquer alternativas a_i, a_j e a_k , a propriedade de transitividade de uma relação \mathfrak{S} é expressa da seguinte forma:

$$a_i \mathfrak{S} a_j \wedge a_j \mathfrak{S} a_k \implies a_i \mathfrak{S} a_k \quad (6.3)$$

O objectivo do método proposto neste capítulo é determinar uma partição parcialmente ordenada e compatível com as preferências do decisor, mas que tenha também a relação τ entre

as classes transitiva. Contudo, a relação que definimos entre as classes é de prevalência, relação essa que modela situações que podem ser intransitivas. Assim, em cada iteração s do Algoritmo 1, a partição \mathcal{P}_s construída com relações de prevalência intransitivas (por exemplo, relação baseadas no método ELECTRE), pode conter classes tais que:

$$\exists\{C^i, C^j, C^k\} \in \mathcal{P}_s: C^i \tau C^j \wedge C^j \tau C^k \wedge C^i \not\tau C^k$$

ou seja, obtém-se o seguinte diagrama

$$C^i \longrightarrow C^j \longrightarrow C^k$$

onde cada arco $C^i \longrightarrow C^j$ representa a existência da relação $C^i \tau C^j$.

Note-se que no caso da estrutura de preferências das classes ser definida pelos centróides, as relações de prevalência são definidas pelos centróides.

Para tornar a relação τ transitiva dado o terno $\{C^i, C^j, C^k\}$, propomos a “negação” de um dos seus arcos não transitivo pertencente ao grupo de alternativas. Assim, propomos três soluções:

1. Exigir a relação $C^i \tau C^k$: $C^i \longrightarrow C^j \longrightarrow C^k$
2. Eliminar a relação $C^i \tau C^j$: $C^i \quad C^j \longrightarrow C^k$
3. Eliminar a relação $C^j \tau C^k$: $C^i \longrightarrow C^j \quad C^k$

Caso não exista mais nenhuma intransitividade na partição, deve-se avaliar a qualidade das três possíveis partições e a que originar a melhor qualidade será a partição final escolhida. Caso ainda existam intransitividades, estas devem ser solucionadas da mesma forma (processo recursivo) e só depois avaliar cada uma das partições originadas.

Note-se que, em particular, se na relação de prevalência entre as classes existe um ciclo $(C^i \tau C^j \tau \dots \tau C^i)$ então devem ser avaliadas todas as partições obtidas pela eliminação de uma

das relações do ciclo, e escolher a que leva à melhor qualidade final. Em particular, para duas classes indiferentes ($C^i \tau C^j \wedge C^j \tau C^i$), estas podem-se fundir.

Como exemplo, consideremos os graus de prevalência entre as classes apresentados na Tabela 6.2. A relação correspondente entre as classes é apresentada na Figura 6.2.

	C^1	C^2	C^3	C^4	C^5
C^1		0.333	0.6	0.133	0.100
C^2	0.333		0.6	0.583	0.125
C^3	0.400	0.183		0.333	0.375
C^4	0.167	0.333	0		0.583
C^5	0.4	0.183	0	0.917	

Tabela 6.2: Grau de prevalência s_{ij} de C^i em C^j , $i, j=1, \dots, 5$

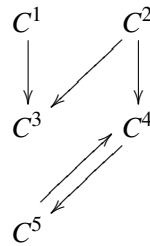


Figura 6.2: Estrutura de preferências de \mathcal{P}

Nesta estrutura existem dois problemas: a intransitividade do terno (C^2, C^4, C^5) e a indiferença do par (C^4, C^5) .

Para obter transitividade, e começando pelo par indiferente, existem três possíveis soluções: S1, S2 e S3 (Figura 6.3). S2 e S3 são partições transitivas, mas S1 não verifica a transitividade. Assim, as três possíveis soluções para se obter transitividade para S1 são as partições S1a, S1b e S1c (Figura 6.4). Finalmente, deve-se avaliar a qualidade das cinco possíveis partições transitivas e aquela que leva a uma melhor qualidade será a partição final.

6.5 Método Aglomerativo para CHOP

Adaptando o clustering aglomerativo tradicional a problemas multicritério e com classes parcialmente ordenadas, inicialmente, à semelhança do Algoritmo 1, formam-se tantas classes quantas as alternativas de \mathcal{A} , onde cada classe é formada por uma alternativa. Constrói-se de

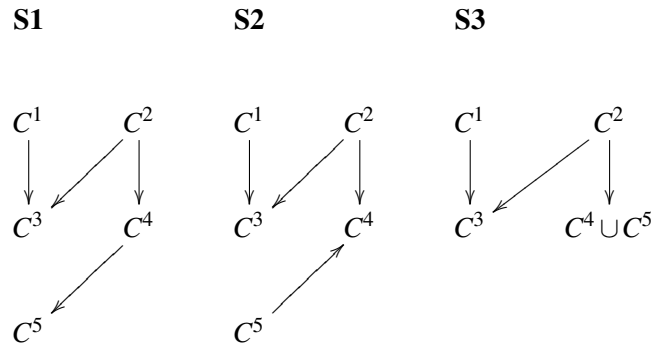


Figura 6.3: Soluções possíveis para as intransitividades de \mathcal{P}_s

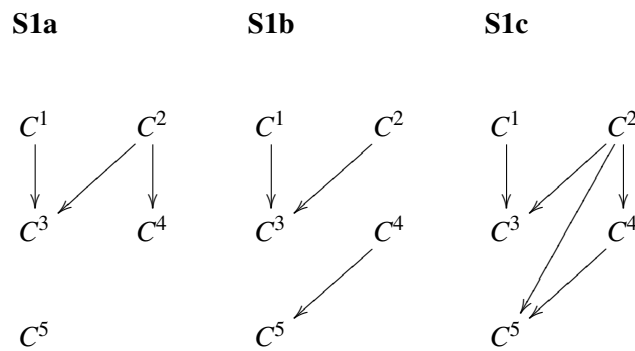


Figura 6.4: Soluções possíveis para as intransitividades de S1

seguida a estrutura de preferências desta primeira partição. Em cada iteração s do algoritmo que propomos, selecciona-se da partição \mathcal{P}_s o par de classes mais similar para ser fundido, ou seja, o par de classes que leva à melhor qualidade da partição \mathcal{P}_{s+1} tornando-a na partição mais compatível com a informação das prevalências fornecidas como *input*. De seguida, actualiza-se a estrutura de preferências da nova partição obtida (Algoritmo 2).

Note-se que neste método proposto, o decisor não necessita especificar o número de classes, inicialmente, podendo ser feito no final com a ajuda dos resultados do método. Na verdade, o decisor pode, eventualmente, decidir o número de classes com base na violação das condições de indiferença, preferência e incomparabilidade ou mesmo no tamanho relativo das classes.

O objectivo deste algoritmo é obter uma *partição transitiva parcialmente ordenada* de um dado conjunto de alternativas. Assim, para garantirmos a transitividade da partição final que se pretende, propomos a reavaliação da qualidade das partições intransitivas que se transformam em transitivas à custa da eliminação ou introdução de arco(s) na partição.

Nas próximas secções serão apresentados três exemplos para ilustrar o algoritmo proposto, para os quais, para se obter as relações de prevalência, calculamos os graus de credibilidade

Algoritmo 2 - Método aglomerativo - extensão

1. Considerar a matriz de relações de prevalência S ($S_{ij} = S(a_i, a_j)$) como *input*.
 2. $s=0$ (iteração)
 3. $\mathcal{P}_s = \{C^1, C^2, \dots, C^m\}$
 - While** existirem pelo menos duas classes em \mathcal{P}_s **do**
 4. Seleccionar o par de vizinhos $(C^q, C^p) \in \mathcal{P}_s^2$ tal que minimize $Q(\mathcal{P}_{s+1})$ considerando a nova classe $C^r = C^q \cup C^p$ em vez de C^q e C^p
 5. Fundir as classes C^q e C^p formando uma nova classe $C^r = C^q \cup C^p$
 6. Actualizar estrutura de preferências para \mathcal{P}_{s+1} .
 7. $s=s+1$
 - end while**
 8. Verificar as partições que têm o número de classes indicadas pelo decisor quanto à transitividade, e fazer as correcções necessárias.
-

do ELECTRE III (Roy, 1978) e ELECTRE-TRI (Yu, 1992; Roy e Bouyssou, 1993), conforme definido na variante proposta por Mousseau e Dias (2004).

6.5.1 Exemplo ilustrativo

Para ilustrar o Algoritmo 2 é apresentado um exemplo com 16 alternativas avaliadas por dois critérios a maximizar. Os valores dos limiares dos critérios, bem como os seus pesos, são apresentados nas Tabelas 6.3 e 6.4, e a matriz S dos valores de prevalência das alternativas é apresentada na Tabela 6.5. Para um limiar de corte de valor 0,6 tem-se que 10% dos pares de alternativas são Indiferentes, 15,83% são Incomparáveis e 74,17% têm uma relação de Preferência.

Aplicando o Algoritmo 2, para cada par de classes $(C^i, C^j) \in \mathcal{P}_s^2$, calcula-se a qualidade da partição caso C^i e C^j se fundissem e vão-se fundindo sucessivamente as classes que geram uma melhor qualidade para $\alpha_P = \alpha_I = \alpha_R = 1$.

Inicialmente temos 16 classes sendo que cada alternativa representa uma classe. A Tabela 6.6 indica a ordem de agregação das classes bem como o vector de índices de desajustamento (v_P, v_I, v_R) e a qualidade da partição resultante, em cada iteração do Algoritmo 2 aplicando

	g_1	g_2		g_1	g_2
a_1	0.57	0.2	a_9	0.24	0.28
a_2	0.24	0.02	a_{10}	0.26	0.65
a_3	0.93	0.02	a_{11}	1.1	0.76
a_4	0.01	0.69	a_{12}	1.79	0.55
a_5	0.86	0.86	a_{13}	1.02	1.06
a_6	2.16	0.6	a_{14}	1.32	1.12
a_7	1.24	0.12	a_{15}	1.36	0.04
a_8	0.8	0.58	a_{16}	0.57	0.17

Tabela 6.3: Desempenhos das alternativas

	g_1	g_2
q_j	0.25	0.1
p_j	0.5	0.2
v_j	2	1.5
k_j	0.5	0.5

Tabela 6.4: Limiares e pesos dos critérios

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}
a_1	0,00	1,00	0,78	0,50	0,42	0,00	0,50	0,50	1,00	0,50	0,00	0,00	0,10	0,00	0,50	1,00
a_2	0,44	0,00	0,50	0,50	0,00	0,00	0,50	0,00	0,50	0,50	0,00	0,00	0,00	0,00	0,50	0,59
a_3	0,60	1,00	0,00	0,50	0,50	0,00	0,88	0,50	0,50	0,50	0,50	0,00	0,35	0,22	0,64	0,75
a_4	0,50	1,00	0,50	0,00	0,15	0,00	0,50	0,50	1,00	1,00	0,50	0,15	0,00	0,00	0,43	0,50
a_5	1,00	1,00	1,00	1,00	0,00	0,47	0,74	1,00	1,00	1,00	1,00	0,50	0,50	0,08	0,50	1,00
a_6	1,00	1,00	1,00	1,00	0,50	0,00	1,00	1,00	1,00	1,00	0,70	1,00	0,50	0,5	1,00	1,00
a_7	1,00	1,00	1,00	0,50	0,50	0,00	0,00	0,50	0,70	0,50	0,50	0,00	0,43	0,38	1,00	1,00
a_8	1,00	1,00	1,00	0,95	0,50	0,43	0,62	0,00	1,00	1,00	0,50	0,50	0,50	0,00	0,50	1,00
a_9	0,84	1,00	0,50	0,50	0,00	0,00	0,50	0,00	0,00	0,50	0,00	0,00	0,00	0,00	0,5	0,84
a_{10}	0,88	1,00	0,50	1,00	0,00	0,07	0,50	0,50	1,00	0,00	0,45	0,31	0,00	0,00	0,50	0,88
a_{11}	1,00	1,00	1,00	1,00	1,00	0,50	1,00	1,00	1,00	1,00	0,00	0,50	0,50	0,50	0,98	1,00
a_{12}	1,00	1,00	1,00	0,80	0,50	0,76	1,00	1,00	1,00	1,00	0,50	0,00	0,50	0,50	1,00	1,00
a_{13}	1,00	1,00	1,00	1,00	1,00	0,50	1,00	1,00	1,00	1,00	1,00	0,50	0,00	0,90	0,82	1,00
a_{14}	1,00	1,00	1,00	1,00	1,00	0,50	1,00	1,00	1,00	1,00	1,00	0,56	1,00	0	1,00	1,00
a_{15}	0,70	1,00	1,00	0,50	0,50	0,00	1,00	0,50	0,50	0,50	0,50	0,14	0,37	0,32	0,00	0,85
a_{16}	1,00	1,00	0,78	0,50	0,42	0,00	0,50	0,50	0,95	0,50	0,00	0,00	0,10	0,00	0,50	0,00

Tabela 6.5: Prevalências entre alternativas (a_i, a_j) , $i, j=1, \dots, 16$

a Estrutura 1. Neste exemplo todas as partições obtidas em cada iteração do algoritmo são transitivas.

Para um hipotético decisor que considere como solução a melhor partição com o menor número de classes, pode-se considerar, como solução final a partição com 4 clusters apresentada na Figura 6.5. A qualidade desta partição é 98,33% resultante de 0% de incoerências das condições de Preferência e Incomparabilidade e 1,67% de incoerências da condição de Indiferença (2/120).

Passo	Nº classes	Classes agregadas	Qualidade	(v_P, v_I, v_R)
0º	16	-	100%	(0 0 0)
1º	15	$\{a_1\} \cup \{a_2\}$	100%	(0 0 0)
2º	14	$\{a_1, a_2\} \cup \{a_3\}$	100%	(0 0 0)
3º	13	$\{a_1, a_2, a_3\} \cup (\{a_8\})$	100%	(0 0 0)
4º	12	$\{a_1, a_2, a_3, a_8\} \cup \{a_7\}$	99.17%	(0.0083 0 0)
5º	11	$\{a_4\} \cup \{a_9\}$	100%	(0 0 0)
6º	10	$\{a_1, a_2, a_3, a_8\} \cup \{a_{16}\}$	99.17%	(0 0.0083 0)
7º	9	$\{a_4, a_9\} \cup \{a_{10}\}$	98.33%	(0 0.0167 0)
8º	8	$\{a_5\} \cup \{a_{11}\}$	98.33%	(0 0.0167 0)
9º	7	$\{a_{13}\} \cup \{a_5, a_{11}\}$	98.33%	(0 0.0167 0)
10º	6	$\{a_5, a_{11}, a_{13}\} \cup \{a_{14}\}$	98.33%	(0 0.0167 0)
11º	5	$\{a_6\} \cup \{a_{12}\}$	98.33%	(0 0.0167 0)
12º	4	$\{a_6, a_{12}\} \cup \{a_{15}\}$	98.33%	(0 0.0167 0)
13º	3	$\{a_1, a_2, a_3, a_8, a_{16}\} \cup \{a_4, a_9, a_{10}\}$	94.17%	(0 0 0.0583)
14º	2	$\{a_1, a_2, a_3, a_4, a_8, a_9, a_{10}, a_{16}\} \cup \{a_5, a_{11}, a_{13}, a_{14}\}$	94.17%	(0 0 0.0583)
15º	1	-	84.17%	(0 0 0.1583)

Tabela 6.6: Agregações das classes e Qualidade obtida com Estrutura 1

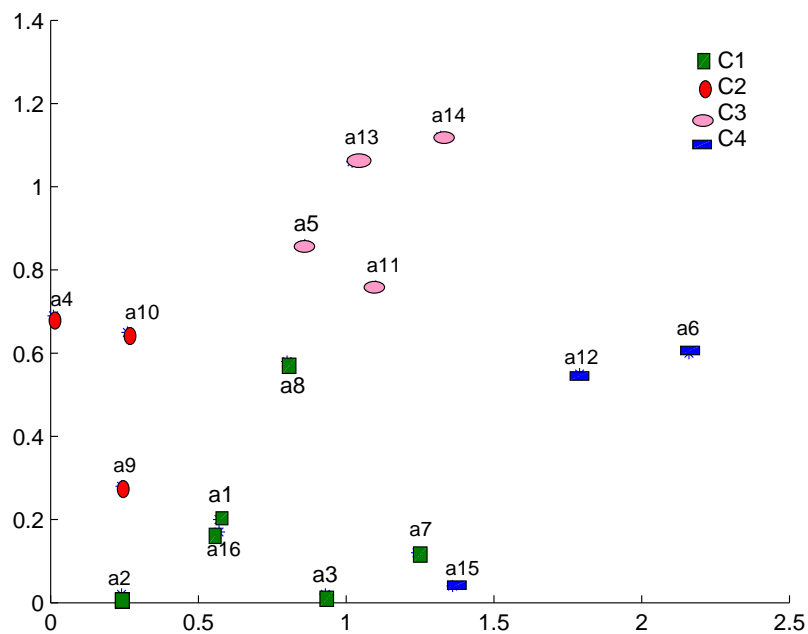


Figura 6.5: Classes obtidas na partição final com Estrutura 1

As classes foram ordenadas à custa das relações de prevalência dos seus centróides (Tabela 6.7), obtendo-se a partição final ordenada e transitiva da Figura 6.6.

	c_1	c_2	c_3	c_4
c_1	0	0.50	0.30	0
c_2	0.50	0	0	0.27
c_3	1.00	1.00	0	0.50
c_4	1.00	0.78	0.50	0

Tabela 6.7: Valores do grau de prevalência entre os centróides

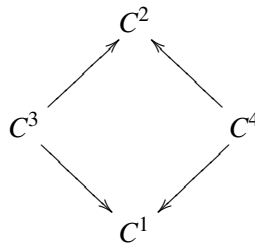


Figura 6.6: Partição final com Estrutura 1

Note-se que as incoerências resultam de $a_9 \in C^2$ ser indiferente a $a_1 \in C^1$ e a $a_{16} \in C^1$ e $C^1 \perp C^2$. É de notar ainda que para obter uma partição com uma qualidade de 100%, o menor número de classes possível é $k=11$ o que, dado o número reduzido de alternativas, é muito elevado.

Aplicando agora as relações de prevalência entre classes (Estrutura 2) para construir a estrutura de preferências das classes em cada iteração do algoritmo proposto, os resultados obtidos na agregação das classes são apresentados na Tabela 6.8.

Para um hipotético decisor que considere como solução a melhor partição com o menor número de classes, pode-se considerar, como solução final a partição com 5 clusters com uma qualidade de 100% apresentada na Figura 6.7. Neste caso não existem pares de alternativas que não verifiquem as condições de Preferência, Indiferença ou Incomparabilidade.

Analisando os elementos das classes resultantes apresentados na Figura 6.7 concluímos que o grupo C^3 é constituído por alternativas cuja junção é pouco intuitiva. Contudo, o facto de a_6 e a_{12} serem preferíveis a a_4 e a_{10} permitiu a sua junção no 11^o passo sem provocar incoerências na partição obtida e de acordo com o princípio $SS\tau$ -Consistency. A junção de C^1

Passo	Nº classes	Classes agregadas	Qualidade	(v_P, v_I, v_R)
0º	16	-	100%	(0 0 0)
1º	15	$\{a_1\} \cup \{a_9\}$	100%	(0 0 0)
2º	14	$\{a_3\} \cup \{a_7\}$	100%	(0 0 0)
3º	13	$\{a_{15}\} \cup (\{a_3, a_7\})$	100%	(0 0 0)
4º	12	$\{a_4\} \cup \{a_{10}\}$	100%	(0 0 0)
5º	11	$\{a_1, a_9\} \cup \{a_{16}\}$	100%	(0 0 0)
6º	10	$\{a_5\} \cup \{a_{11}\}$	100%	(0 0 0)
7º	9	$\{a_{13}\} \cup \{a_{14}\}$	100%	(0 0 0)
8º	8	$\{a_6\} \cup \{a_{12}\}$	100%	(0 0 0)
9º	7	$\{a_8\} \cup \{a_5, a_{11}\}$	100%	(0 0 0)
10º	6	$\{a_1, a_9, a_{16}\} \cup \{a_2\}$	100%	(0 0 0)
11º	5	$\{a_4, a_{10}\} \cup \{a_6, a_{12}\}$	100%	(0 0 0)
12º	4	$\{a_1, a_2, a_9, a_{16}\} \cup \{a_5, a_8, a_{11}\}$	98.33%	(0 0.0167 0)
13º	3	$\{a_1, a_2, a_5, a_8, a_9, a_{11}, a_{16}\} \cup \{a_4, a_6, a_{10}, a_{12}\}$	94.17%	(0 0.0167 0.0417)
14º	2	$\{a_1, a_2, a_4, a_5, a_6, a_8, a_9, a_{10}, a_{11}, a_{12}, a_{16}\} \cup \{a_3, a_7, a_{15}\}$	87.50%	(0 0 0.1250)
15º	1	-	84.17%	(0 0 0.1583)

Tabela 6.8: Agregações das classes e Qualidade obtida com Estrutura 2

com C^2 seria mais intuitiva, contudo levaria ao aparecimento de incoerências da condição de Incomparabilidade. Por outro lado, a junção de C^4 com C^5 embora também mais intuitiva na verdade conduziu ao mesmo resultado prático pois as alternativas de C^5 são preferíveis às de C^4 .

De seguida ordenaram-se as classes à custa das suas relações de prevalência (Tabela 6.9), obtendo-se assim a partição final ordenada e transitiva da Figura 6.8.

	C^1	C^2	C^3	C^4	C^5
C^1	0	0.173	0	0	0
C^2	0.75	0	0	0	0
C^3	0.86	0.50	0	0.25	0
C^4	1.00	0.78	0.50	0	0
C^5	1.00	1.00	0.50	1.00	0

Tabela 6.9: Valores do grau de prevalência s_{ij} entre as classes (C^i, C^j), $i, j=1, \dots, 5$

Perante as duas partições apresentadas ao decisor obtidas com Estrutura 1 e 2, um hipotético decisor pode optar por aquela que melhor represente o problema ou, em caso de desconhecimento sobre a melhor representação, pode escolher a partição com a melhor qualidade. Neste caso, a partição resultante da Estrutura 2 é a que tem a melhor qualidade, podendo pois ser uma possível solução.

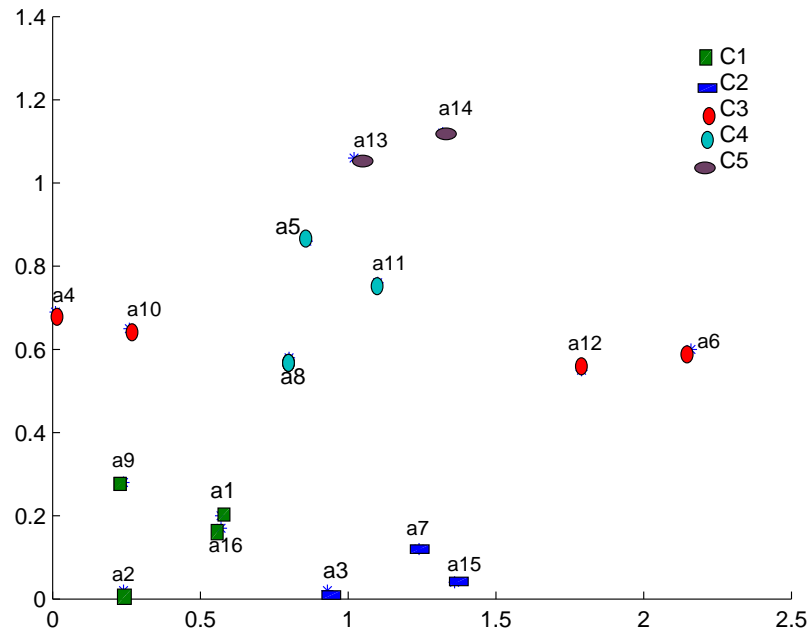


Figura 6.7: Classes obtidas na partição final com Estrutura 2

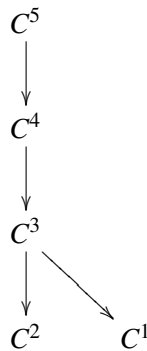


Figura 6.8: Partição final com Estrutura 2

6.6 Exemplos adicionais com casos reais da literatura

Nesta secção serão apresentados mais dois casos reais da literatura para ilustrar o modelo **CHOP** descrito nas secções anteriores. A estrutura final será escolhida de forma mais ou menos “cega”, tendo apenas em consideração os resultados obtidos e a regra de escolher a partição que tenha o menor número de classes e com uma qualidade razoável (ou seja, que não melhora significativamente). No entanto, na prática a escolha da partição final deverá basear-se não só nos resultados obtidos, mas também nas preferências e experiência dos decisores.

A estrutura de preferências utilizada para obter os resultados que serão apresentados de seguida será apenas a Estrutura 2 definida à custa das relações de prevalência entre as classes e optamos por não apresentar os resultados obtidos com a Estrutura 1 definida à custa dos centróides. Tal decisão resulta do facto de termos obtido partições com bastantes intransitividades quando aplicada a Estrutura 1 o que exigia uma reavaliação das soluções. Na prática torna o processo mais moroso e não leva a resultados finais melhores, quando muito iguais aos obtidos com a Estrutura 2.

Os resultados do Algoritmo 2 que são apresentados de seguida consideram no cálculo da qualidade da partição $\alpha_P = \alpha_I = \alpha_R = 1$.

A ferramenta utilizada na obtenção de resultados foi o *MatLab 7.0*.

6.6.1 Primeiro exemplo

Nesta subsecção é apresentado o Exemplo 1 da Secção 5.4.1, cujos desempenhos em 6 critérios são apresentados na Tabela 5.2. Quanto aos valores dos limiares dos critérios, estes são dados, uma vez mais, pela Tabela 5.3. Considerar-se-á que todos os critérios têm o mesmo peso, ou seja, os coeficientes de importância serão todos iguais a $1/6$ ($k_i = 1/6, i \in \{1, \dots, 6\}$).

Pela análise da Figura 6.9 podemos ver que só para partições com pelo menos 12 classes se consegue uma qualidade final $Q^*=100\%$, o que, dado o número de alternativas, poderá não ser uma solução satisfatória. Assim, considerando que o decisor apenas está interessado nas soluções com menos de 7 classes (o que é razoável tendo em atenção que $|\mathcal{A}|=20$), os resultados obtidos, aplicando o Algoritmo 2, são apresentados na Tabela 6.10.

k	Transitividade	Qualidade	(v_P, v_I, v_R)	Estrutura
6	✓	97.89	(0 0 0.0211)	
5	✓	97.37	(0 0.0053 0.0211)	
4	✓	96.32	(0 0.0053 0.0315)	
3	✓	94.21	(0 0.0053 0.0526)	
2	✓	88.95	(0 0.0053 0.1053)	
1	✓	67.89	(0 0 0.3211)	

Tabela 6.10: Agregações das classes e Qualidade obtida

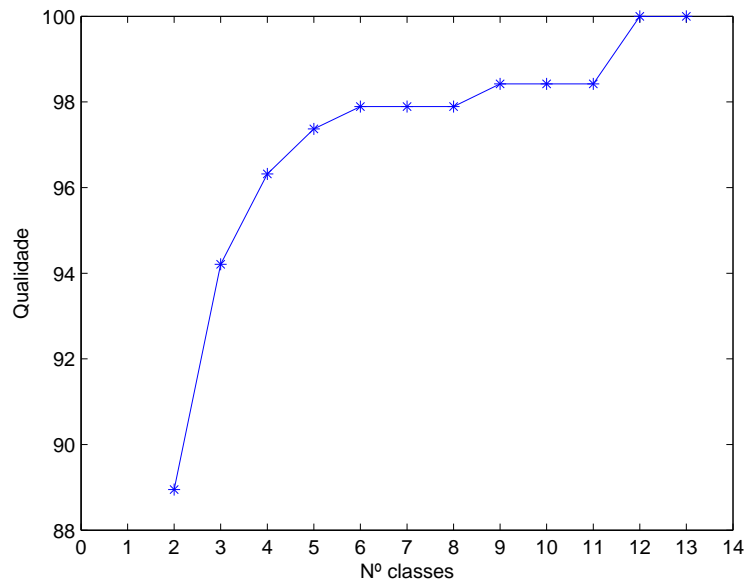


Figura 6.9: Qualidade das partições.

Considerando que o decisor opta por $k=5$ classes como uma boa solução (note-se que a qualidade não melhora significativamente a partir de 5 classes), obtém-se uma partição com uma qualidade de 97.37% resultante de 0% de violações das condições de Preferência e 0.53% de Indiferença, e 1.58% de Incomparabilidade.

As classes resultantes são as seguintes:

- $C^1 = \{a_1, a_2, a_4, a_7, a_{19}\}$,
- $C^2 = \{a_3, a_6, a_8, a_{13}, a_{14}, a_{20}\}$,
- $C^3 = \{a_5, a_{18}\}$,
- $C^4 = \{a_9, a_{11}, a_{16}\}$ e
- $C^5 = \{a_{10}, a_{12}, a_{15}, a_{17}\}$.

Ordenaram-se as classes à custa das suas relações de prevalência (Tabela 6.11), obtendo-se assim a partição final ordenada e transitiva da Figura 6.10.

Note-se que, a solução sugerida pelo método *Clustering-Sorting* apresentado no Capítulo anterior tem quatro categorias e não cinco. Comparando tal solução encontrada com a solução fornecida pelo método CHOP para quatro categorias (ver Tabela 6.10), podemos verificar uma

	C^1	C^2	C^3	C^4	C^5
C^1	0	0.3333	0.2000	0.1333	0.1000
C^2	0.3333	0	0	0	0.1250
C^3	0.4000	0.5833	0	0.3333	0.3750
C^4	0.4667	0.7222	0	0	0.5833
C^5	0.4000	0.9167	0	0	0

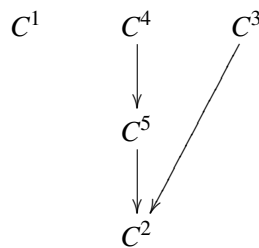
Tabela 6.11: Valores do grau de similaridade s_{ij} entre as classes (C^i, C^j), $i, j=1, \dots, 5$ 

Figura 6.10: Partição final

melhoria significativa na incoerência da condição de Preferência no método CHOP, levando a uma melhoria significativa na qualidade da partição obtida (ver Tabela 6.12).

Condições	Relação S	CHOP	Clustering-Sorting
Preferência	121 (63.7%)	0 (0.0%)	8 (4.2%)
Indiferença	8 (4.2%)	1(0.53%)	0 (0.0%)
Incomparabilidade	61(32.1%)	6 (3.2%)	7(3.7%)
Total	190	7 (3.68%)	15(7.9%)
$Q(\mathcal{P})$	-	96.32%	92.10%

Tabela 6.12: Comparação das incoerências das soluções obtidas com os métodos Clustering-Sorting e CHOP, segundo o $SS\tau$ -Consistency, para $k=4$.

6.6.2 Segundo exemplo

O segundo exemplo é sobre a avaliação de 100 alternativas a serem classificadas, com base nos seus desempenhos, em 7 critérios a minimizar, apresentado na Seção 5.4.2. Consideram-se novamente os valores originais (Yu, 1992) para os limiares de indiferença, preferência e pesos (Tabela 5.8) associados a cada critério bem como os coeficientes de importância, mas com menos incomparabilidades ($v_j=(1.95, 100, 100, 100, 100, 100, 18.7)$) e $\lambda = 0.6$.

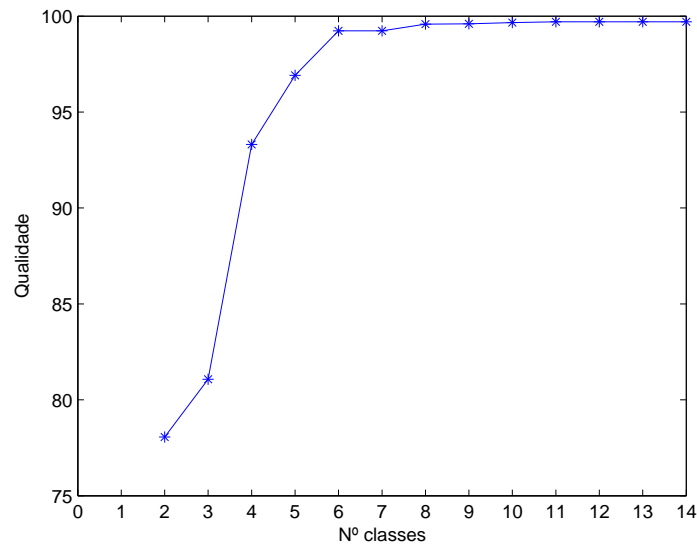


Figura 6.11: Qualidade das partições.

Pela análise da Figura 6.11 podemos ver que uma boa solução será reter 6 classes. Considerando que o decisor só está interessado em obter informações sobre as estruturas das partições com no máximo 7 classes, os resultados obtidos, aplicando o Algoritmo 2, são apresentados na Tabela 6.13.

Caso o decisor opte por $k=6$ classes como uma boa solução, obtém-se uma partição com uma qualidade de 99.23% resultante de 0% de violações das condições de Preferência, 0.016% de Indiferença, e 0.61% de Incomparabilidade.

Note-se que, a solução sugerida pelo método *Clustering-Sorting* apresentado no Capítulo anterior tem quatro categorias e não seis. Comparando tal solução encontrada com a solução fornecida pelo método CHOP para quatro categorias (ver Tabela 6.14), podemos verificar uma melhoria significativa nas incoerências das condições de Preferência, Indiferença e Incomparabilidade no método CHOP, levando a uma melhoria na qualidade da partição obtida.

Analisando agora o grupo de alternativas segundo todos os seus valores originais (nomeadamente $\lambda = 0.86$), os resultados obtidos, aplicando o Algoritmo 2, são apresentados na Tabela 6.15.

Neste caso, uma vez que se tem uma elevada percentagem de alternativas incomparáveis, todas as partições com no máximo 7 classes são consideradas incomparáveis para se obter a melhor partição. À semelhança dos resultados obtidos nos anteriores exemplos, também aqui todas as partições são obviamente transitivas.

k	Transitividade	Qualidade	Incoerências	Estrutura
7	✓	99.23	(0 0.0016 0.0061)	
6	✓	99.23	(0 0.0016 0.0061)	
5	✓	96.91	(0 0.0016 0.0293)	
4	✓	93.31	(0 0.001 0.0659)	
3	✓	81.07	(0 0.0004 0.1889)	
2	✓	78.06	(0 0.0002 0.2192)	
1	-	67.29	(0 0 0.3271)	

Tabela 6.13: Qualidade das partições obtidas com o Algoritmo 2.

Condições	Relação S	Clustering-Sorting	CHOP
Preferência	2967 (59.9%)	39 (0.8%)	0 (0.0%)
Indiferença	364 (7.4%)	32 (0.6%)	5 (0.1%)
Incomparabilidade	1619 (32.7%)	367(7.4%)	326 (6.59%)
Total	4950	438 (8.8%)	331 (6.69%)
$Q(\mathcal{P})$	-	91.2%	93.31%

Tabela 6.14: Número de pares de alternativas que verificam as condições de P, I e R antes da partição e o número de inconsistências pós-partição segundo o $SS\tau$ -Consistency

k	Transitividade	Qualidade	Incoerências
7	✓	85.32	(0 0.0004 0.1465)
6	✓	84.32	(0 0.0004 0.1562)
5	✓	83.74	(0 0.0004 0.1622)
4	✓	80.71	(0 0.0004 0.1925)
3	✓	66.67	(0 0.0004 0.3230)
2	✓	41.23	(0 0.0002 0.5875)
1	-	36.59	(0 0 0.6324)

Tabela 6.15: Qualidade das partições obtidas com o Algoritmo 2

6.7 Notas conclusivas

Neste capítulo propomos um modelo para a classificação de um conjunto \mathcal{A} de alternativas, com base em vários critérios, a um conjunto de classes parcialmente ordenado com uma estrutura desconhecida *a priori*.

Para a obtenção de partições do conjunto \mathcal{A} , propomos uma adaptação do método hierárquico aglomerativo. O método proposto inicia com m clusters, cada um contendo uma das m alternativas de um conjunto \mathcal{A} de alternativas a atribuir. De seguida, e de forma iterativa, os clusters com maior similaridade são fundidos. A noção de similaridade entre duas classes é substituída por uma avaliação da qualidade da partição obtida com a fusão de classes: dois grupos são tanto mais similares quanto menor for o número de violações das preferências do decisor de acordo com o princípio *Semi-Strong consistency* ($SS\tau$ -Consistency) gerado com a sua fusão.

Em cada iteração do algoritmo proposto, obtém-se uma estrutura de preferências *parcialmente ordenada* da partição \mathcal{P} das alternativas, ou seja, uma estrutura onde podem existir classes melhores, piores e até incomparáveis. Sugerimos duas formas de construção da estrutura de preferências das classes, ambas definidas à custa de relações binárias de prevalência, uma entre pares de centróides representativos das classes e outra entre pares de classes $(C^i, C^j) \in \mathcal{P}^2$. No primeiro caso, uma classe é considerada melhor que outra caso o seu centróide prevaleça sobre o centróide da outra classe mas não seja prevalecido, as classes são incomparáveis se nenhum dos centróides prevalecer sobre o outro e por fim são consideradas indiferentes caso os centróides prevaleçam um sobre o outro (Estrutura 1). No segundo caso, considerou-se que uma classe C^i prevalece sobre C^j se pelo menos 50% das relações dos pares de alternativas $(a_s, a_t) \in (C^i, C^j)$ seja de prevalência de a_s sobre a_t . A relação de preferência entre as classes é então definida de forma idêntica à dos centróides, de acordo com as relações de prevalência entre as classes (Estrutura 2).

Os resultados obtidos com a Estrutura 2 mostraram-se no geral melhores que os obtidos com a Estrutura 1 uma vez que levam à obtenção de muito menos intransitividades e com uma qualidade idêntica ou mesmo melhor.

O objectivo do método proposto neste capítulo é determinar uma partição parcialmente ordenada e compatível com as preferências do decisor, mas que seja também transitiva. Contudo, a relação que definimos entre as classes é de prevalência, relação essa que é intransitiva. Assim, em cada iteração do Algoritmo proposto, pode existir pelo menos um terno de classes

que não verificam a transitividade. Propomos re-avaliação da qualidade de tais partições após serem transformadas em transitivas à custa da eliminação ou introdução de arco(s) na partição.²

O desenvolvimento do CHOP pode ser útil como abordagem isolada ou como um módulo de pesquisa inicial que antecede um método “exacto”³. Nesta segunda situação, a intenção do CHOP é dar ao decisor um conhecimento inicial do problema, neste caso, dar um conhecimento inicial sobre a estrutura das classes. O decisor avançará depois para uma re-afecção das alternativas por forma a minimizar a percentagem de incoerências de acordo com o princípio *SS τ -Consistency*, caso não esteja satisfeito com os resultados obtidos com o CHOP.

No próximo capítulo a nossa atenção irá pois centrar-se no desenvolvimento de um método “exacto” para problemas de programação inteira 0-1, com estruturas conhecida e desconhecida *a priori*. Note-se que não nos propusemos a desenvolver apenas uma metodologia “exacta” com desconhecimento da estrutura das classes *a priori* pois estávamos conscientes da necessidade de recorrer à programação inteira (PI) e, como se sabe, a utilização de métodos “exactos” pode tornar os problemas de PI de maiores dimensões intratáveis. Assim, a heurística CHOP desenvolvida assumiu um papel importante para ultrapassar as dificuldades do cálculo, permitindo a busca de soluções boas num tempo computacional razoável (nos exemplos mais pequenos usados, conseguimos obter resultados em segundos e no maior obtivemos resultados em poucas horas).

²Note-se que esta re-avaliação proposta pode ser igualmente aplicada ao método Clustering-Sorting já que este modelo também não garante a transitividade da partição final. Na verdade, tal como no método CHOP, a construção da estrutura das classes tem por base relações de prevalência, o que pode levar a uma estrutura intransitiva.

³Designamos por métodos “exactos” porque calculam soluções eficientes do problema, e não apenas aproximações destas como as heurísticas.

Classificação Multicritério com classes ordinais e nominais: modelo de Optimização

7.1 Notas introdutórias

Neste capítulo apresentamos modelos de otimização para problemas classificação multicritério com classes parcialmente ordenadas, recorrendo à programação inteira 0-1. Contudo, as dificuldades inerentes à complexidade da programação inteira (PI), agravadas pela existência de múltiplos critérios, são ainda agravadas, no nosso caso, pelo desconhecimento da estrutura de preferências dessas classes, o que pode tornar os problemas de maiores dimensões intratáveis.

Assim, caso o problema seja relativamente pequeno e o decisor tenha informação da(s) estrutura(s) das classes, então partir-se-à directamente para o cálculo da partição óptima. Caso contrario, um dos métodos propostos nos capítulos anteriores, o método *Clustering-Sorting* ou o método *CHOP*, serão considerados neste capítulo, como um módulo de pesquisa inicial que antecede um método “exacto”. A intenção é dar ao decisor um conhecimento inicial do problema que o ajude a estabelecer e rever as suas preferências a um relativamente baixo custo computacional.

A nova abordagem descrita neste capítulo é tecnicamente diferente do CHOP, mas tem por base os mesmos objectivos e exigências do decisor:

- (1) traduzir as preferências do decisor através do princípio $S\tau$ - Consistency ou do princípio $SS\tau$ - Consistency,

- (2) minimizar a taxa de erro de classificação de um conjunto de alternativas a classes parcialmente ordenadas, avaliadas à custa das suas relações de prevalência,

(3) traduzir a percepção do decisor quanto às violações que considera mais graves, dando mais peso a essas violações.

O capítulo está organizado da seguinte forma:

Começamos por apresentar as restrições de afectação das alternativas às classes, de acordo com o princípio $SS\tau$ - Consistency. Serão ainda apresentadas as restrições para a estrutura pretendida das classes no caso desta ser desconhecida *a priori*. De seguida é descrito o modelo matemático de optimização para minimizar a taxa de erro de classificação das alternativas. As dimensões dos problemas PLI são apresentados, comparando o caso dos problemas com estrutura desconhecida *a priori* com os problemas cuja estrutura é dada. Ilustramos o modelo apresentado com um pequeno exemplo e ainda com um exemplo aplicado ao CHOP, comparando-se os resultados obtidos nos dois modelos. Finalmente apresentamos algumas conclusões do modelo aqui descrito.

7.2 Procedimento de afectação a uma estrutura conhecida *versus* desconhecida *a priori*

Para o modelo de optimização, para problemas de classificação multicritério, serão conhecidos *a priori* o número de classes (denotado por n_{cat}), bem como a matriz dos valores de prevalência das acções de \mathcal{A} , denotada por S . Caso o decisor tenha conhecimento da estrutura parcialmente ordenada das classes, esta será representada por um grafo dirigido (ou digrafo) $\mathcal{G}=(\mathcal{P},\mathcal{E})$ onde $\mathcal{P} = \{C^1, C^2, \dots, C^{n_{cat}}\}$ é o conjunto de nodos e \mathcal{E} o conjunto de arcos.

O modelo de optimização que apresentamos neste capítulo consiste na atribuição de um conjunto de alternativas a um conjunto de classes parcialmente ordenadas. O objectivo é fazer a atribuição o mais compatível possível com o princípio $SS\tau$ -Consistency¹ segundo o qual se devem verificar os corolários (7.1), (7.2) e (7.3) para a atribuição do conjunto \mathcal{A} de alternativas, que denominaremos corolários de Preferência, Indiferença e Incomparabilidade respectivamente.

Sejam $a_i \in C^s$ e $a_j \in C^t$.

¹Em apêndice é apresentada a formulação matemática para a atribuição das alternativas segundo o princípio $S\tau$ -Consistency.

(1) Se uma acção a_j é preferível a a_i então não deve pertencer a uma classe pior:

$$(a_j Pa_i \wedge a_i \in C^s \wedge a_j \in C^t) \Rightarrow C^s \succ C^t, \text{ i.e.}$$

$$C^s \succ C^t \Rightarrow (a_j \not Pa_i \vee a_i \notin C^s \vee a_j \notin C^t) \quad (7.1)$$

(2) Acções com preferência similar (indiferentes) não devem pertencer a classes incomparáveis:

$$(a_i Ia_j \wedge a_i \in C^s \wedge a_j \in C^t) \Rightarrow C^s \not\perp C^t.$$

i.e

$$C^s \perp C^t \Rightarrow (a_j \not Ia_i \vee a_i \notin C^s \vee a_j \notin C^t), \quad (7.2)$$

(3) Acções incomparáveis não devem pertencer à mesma classe:

$$(a_i Ra_j \wedge a_i \in C^s \wedge a_j \in C^t) \Rightarrow s \neq t, \text{ i.e.}$$

$$s = t \Rightarrow (a_j \not Ra_i \vee a_i \notin C^s \vee a_j \notin C^s) \quad (7.3)$$

Entre todas as possíveis partições \mathcal{P} de \mathcal{A} , procuramos a partição mais compatível com as três condições anteriores, ou seja, que minimizam as incoerências das condições de Preferência, Indiferença e Incomparabilidade, e que tenha um número conhecido de classes cuja estrutura pode ou não ser conhecida *a priori*. A partição óptima pode não ser única, ou seja, procuramos uma das partições $\mathcal{F}_{\mathcal{P}}^*$ da família de todas as partições possíveis de \mathcal{A} que minimizam as incoerências.

7.3 Modelo de optimização

O objectivo do modelo de optimização que propomos é minimizar a taxa de erro de classificação. A taxa de erro refere-se à razão entre o número de pares de alternativas mal atribuídas de acordo com o principio $SS\tau$ -consistency e o número total de pares de alternativas de \mathcal{A} , $\frac{m(m-1)}{2}$. Estes pares de alternativas que não verificam as condições de $SS\tau$ -consistency podem ser representados pelas variáveis binárias apresentadas na Tabela 7.1.

Variável erro	$a_i \in C^l, a_k \in C^j$
$\varepsilon_{P_{ik,jl}}$	$a_i P a_k \quad C^j \succ C^l$
$\varepsilon_{I_{ik,jl}}$	$a_i I a_k \quad C^j \perp C^l$
$\varepsilon_{R_{ik,j}}$	$a_i R a_k \quad C^j = C^l$

Tabela 7.1: Variáveis binárias que representam a situação de incoerência para cada par de alternativas

- $\varepsilon_{P_{ik,jl}}$ é uma variável binária 0-1, definida para indicar a existência de incoerência do par de alternativas (a_i, a_k) , onde $a_i P a_k$ e $a_k \in C^j$, $a_i \in C^l$ e $C^j \succ C^l$: $\varepsilon_{P_{ik,jl}}=1$ se existir uma relação de preferência invertida entre o par (a_i, a_k) , e $\varepsilon_{P_{ik,jl}}=0$ caso contrário.
- $\varepsilon_{I_{ik,jl}}$ é uma variável binária 0-1, definida para indicar a existência de incoerência na relação de indiferença entre as alternativas indiferentes $a_i \in C^j$ e $a_k \in C^l$ se estas forem atribuídas a classes incomparáveis, $\varepsilon_{I_{ik,jl}}=1$, e $\varepsilon_{I_{ik,jl}}=0$ caso contrário.
- $\varepsilon_{R_{ik,j}}$ é uma variável binária 0-1, definida para indicar a existência de incoerência na relação de incomparabilidade do par de alternativas incomparáveis (a_i, a_k) se estas forem atribuídas à mesma classe, $\varepsilon_{R_{ik,j}}=1$, e $\varepsilon_{R_{ik,j}}=0$ caso contrário.

Assim, uma medida adequada da taxa de erro de classificação é expressa da seguinte forma:

$$\gamma = \frac{1}{m(m-1)} \sum_{i,k=1}^m \sum_{j,l=1}^{n_{cat}} (\alpha_P \times \varepsilon_{P_{ik,jl}} + \alpha_I \times \varepsilon_{I_{ik,jl}} + \alpha_R \times \varepsilon_{R_{ik,j}}) \quad (7.4)$$

onde $0 \leq \alpha_P, \alpha_I, \alpha_R \leq 1$ são os custos de cada erro de violação das condições de preferência, indiferença e incomparabilidade, respectivamente, indicados pelo decisor.

7.3.1 Formulação matemática

Tendo por objectivo a minimização da taxa de erro de classificação esperada (7.4), tal é conseguido através de técnicas de programação matemática. Seja x_{ij} uma variável binária 0-1 definida tal que $x_{ij} = 1$ se a alternativa $a_i \in C^j$ e $x_{ij} = 0$ se a alternativa $a_i \notin C^j$ ($i=1, \dots, m$; $j=1, \dots, n_{cat}$). Seja ainda p_{ij} uma variável binária 0-1 definida tal que $p_{ij} = 1$ se $C^i \succ C^j$ e $p_{ij} = 0$ caso contrário ($i, j=1, \dots, n_{cat}$).

Assim, na atribuição das alternativas a uma estrutura com classes parcialmente ordenadas temos as restrições que se seguem:

1. Seja $C^s, C^t \in \mathcal{P}$ e seja $a_j P a_i$. Por (7.1), caso $C^s \succ C^t$ ($p_{st} = 1$), a_j não deve pertencer a C^t ($x_{jt} = 0$) ou a_i não deve pertencer a C^s ($x_{is} = 0$). Logo,

$$x_{jt} + x_{is} + p_{st} \leq 2 \quad (7.5)$$

2. Seja $C^s, C^t \in \mathcal{P}$ e seja $a_j I a_i$. Por (7.2), caso $C^s \perp C^t$ ($p_{st} = 0 \wedge p_{ts} = 0$), a_j não deve pertencer a C^t ($x_{jt} = 0$) ou a_i não deve pertencer a C^s ($x_{is} = 0$)

$$x_{jt} + x_{is} - p_{ts} - p_{st} \leq 1 \quad (7.6)$$

3. Seja $C^s \in \mathcal{P}$ e $a_j R a_i$. Por (7.3), a_j não deve pertencer a C^s ($x_{js} = 0$) ou a_i não deve pertencer a C^s ($x_{is} = 0$). Assim,

$$x_{js} + x_{is} \leq 1 \quad (7.7)$$

4. Cada alternativa a_i deve pertencer a uma e uma só classe C^s . Assim,

$$\sum_{s=1}^{n_{cat}} x_{is} = 1 \quad (7.8)$$

5. Cada par de classes (C^i, C^j) deve ser apenas incomparável ou ter uma relação de preferência. Assim,

$$p_{ij} + p_{ji} \leq 1 \quad (7.9)$$

6. Cada terno de classes (C^i, C^j, C^k) deve ser transitivo. Assim,

$$p_{ik} \geq p_{ij} + p_{jk} - 1 \quad (7.10)$$

7. Caso se deseje, pode-se ainda exigir que cada classe C^s da partição não fique vazia. Logo

$$\sum_{i=1}^m x_{is} \geq 1 \quad (7.11)$$

O Problema Linear Inteiro (PLI 1) seguinte pode ser resolvido para encontrar o valor mínimo γ na atribuição de um conjunto de alternativas \mathcal{A} a serem atribuídas a uma determinada estrutura de classes parcialmente ordenadas \mathcal{P} , **desconhecida a priori** (mas com o número máximo de categorias, n_{cat} , especificado):

$$((PLI1)) \quad \min \gamma \quad (7.12)$$

sujeito a :

$$x_{ij} + x_{kl} + p_{jl} \leq 2 + \epsilon_{P_{ik,jl}}, \forall a_i, a_k \in A : a_i P a_k, \forall j, l \in \{1, \dots, n_{cat}\} \quad (7.13)$$

$$x_{ij} + x_{kl} - p_{jl} - p_{lj} \leq 1 + \epsilon_{I_{ik,jl}}, \forall a_i, a_k \in A : a_i I a_k, \forall j, l \in \{1, \dots, n_{cat}\} \quad (7.14)$$

$$x_{ij} + x_{kj} \leq 1 + \epsilon_{R_{ik,j}}, \forall a_i, a_k \in A : a_i R a_k, i < k, \forall j \in \{1, \dots, n_{cat}\} \quad (7.15)$$

$$\sum_{j=1}^{n_{cat}} x_{ij} = 1, \forall i \in \{1, \dots, m\} \quad (7.16)$$

$$p_{ij} + p_{ji} \leq 1, \forall i, j \in \{1, \dots, n_{cat}\} \quad (7.17)$$

$$p_{ik} \geq p_{ij} + p_{jk} - 1, \forall i, j, k \in \{1, \dots, n_{cat}\} \quad (7.18)$$

$$x_{ij}, p_{ij}, \epsilon_{P_{ik,jl}}, \epsilon_{I_{ik,jl}}, \epsilon_{R_{ik,j}} \in \{0, 1\}, \forall i, k \in \{1, \dots, m\}, j, l \in \{1, \dots, n_{cat}\}$$

e opcionalmente,

$$\sum_{i=1}^m x_{is} \geq 1, \forall i \in \{1, \dots, n_{cat}\} \quad (7.19)$$

A funç o objectivo (7.12) do (PLI 1) corresponde   f rmula (7.4). As restriç es (7.13)-(7.15) correspondem  s condiç es (7.5)-(7.7) respectivamente, representando as condiç es de atribuiç o do conjunto de alternativas segundo o princ pio $SS\tau$ - Consistency. A restriç o (7.16) diz respeito   unicidade da atribuiç o. Quanto  s restriç es (7.17) e (7.18), estas correspondem a (7.9) e (7.10) respectivamente, e dizem respeito  s condiç es de inexist ncia de classes indiferentes e de transitividade da estrutura de classes pretendida.

Caso a estrutura das classes seja **conhecida a priori** (sendo uma das estruturas poss veis da Figura 7.1), o (PLI 1) fica consideravelmente mais simples, com bastantes menos restriç es e vari veis, resultante da inexist ncia de condiç es para a estrutura das classes:

- nas restriç es de (7.13), $p_{jl} = 1 \forall j, l : C^l \succ C^j$
- nas restriç es de (7.14), $p_{jl} = p_{lj} = 0 \forall j, l : C^l \perp C^j$
- as restriç es (7.17) e (7.18) deixam de existir.

Obtém-se assim o (PLI 2).

$$(PLI2) \quad \min \gamma \quad (7.20)$$

sujeito a :

$$x_{ij} + x_{kl} \leq 1 + \varepsilon_{P_{ik,jl}}, \forall a_i, a_k \in A : a_i P a_k \wedge C^l \succ C^j, \forall j, l \in \{1, \dots, n_{cat}\}, j \neq l \quad (7.21)$$

$$x_{ij} + x_{kl} \leq 1 + \varepsilon_{I_{ik,jl}}, \forall a_i, a_k \in A : a_i I a_k \wedge C^l \perp C^j, \forall j, l \in \{1, \dots, n_{cat}\}, j \neq l \quad (7.22)$$

$$x_{ij} + x_{kj} \leq 1 + \varepsilon_{R_{ik,j}}, \forall a_i, a_k \in A : a_i R a_k, \forall j \in \{1, \dots, n_{cat}\} \quad (7.23)$$

$$\sum_{j=1}^{n_{cat}} x_{ij} = 1, \forall i \in \{1, \dots, m\} \quad (7.24)$$

$$x_{ij}, \varepsilon_{P_{ik,jl}}, \varepsilon_{I_{ik,jl}}, \varepsilon_{R_{ik,j}} \in \{0, 1\}, \forall i, k \in \{1, \dots, m\}, j, l \in \{1, \dots, n_{cat}\}$$

e opcionalmente,

$$\sum_{i=1}^m x_{is} \geq 1, \forall i \in \{1, \dots, n_{cat}\} \quad (7.25)$$

Note-se que uma alternativa ao problema linear inteiro (PLI 1) é o (PLI 2) com a restrição de não existência de classes vazias (7.25) aplicado a cada estrutura possível para k classes, escolhendo-se para solução final do problema a(s) partição(ões) com melhor solução γ . Sabendo o número k de classes que se pretende ter na estrutura final, as estruturas possíveis podem ser determinadas mais facilmente com a ajuda do Triângulo de Pascal (Figura 7.1). Para tal, basta fixar k linhas consecutivas do Triângulo de Pascal, sendo que a primeira linha contém k elementos consecutivos, a segunda linha contém $(k-1)$ elementos, até à última linha que deverá conter apenas um elemento. De seguida, fixando k elementos e considerando as relações existentes entre eles, vão-se obtendo as diferentes estruturas possíveis.

No caso do problema ser de dimensões consideráveis, tanto em termos de número de alternativas como de classes, as dimensões do (PLI 1) podem mesmo reduzir drasticamente, como se verá na secção seguinte.

7.4 Dimensão do problema

Dado o conjunto \mathcal{A} com m alternativas e dada a sua matriz S de prevalências, facilmente se deduz o número de pares de alternativas que são indiferentes, incomparáveis ou que têm uma relação de preferência. A representação desses números é aqui dada por:

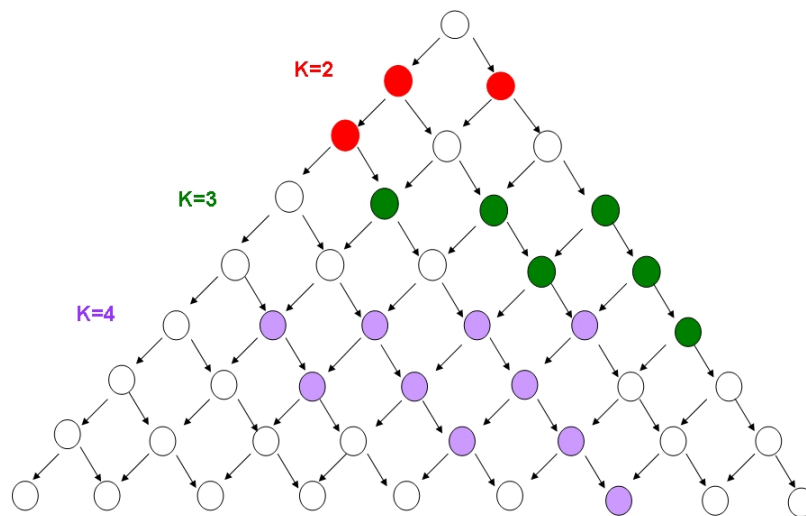


Figura 7.1: Estruturas possíveis para k categorias, determinadas com o Triângulo de Pascal

- n_I : número de pares de alternativas indiferentes
- n_R : número de pares de alternativas incomparáveis
- n_P : número de pares de alternativas com uma das alternativas melhor que a outra

O número de restrições e o número de variáveis do (PLI 1), ou seja, a dimensão do problema é dada na Tabela 7.2.

onde ${}^a C_b$ é o coeficiente binomial e representa o número de subconjuntos com b elementos que se podem formar num conjunto com a elementos, podendo ser calculado por $\frac{a!}{b!(a-b)!}$.

No caso de se conhecer a estrutura parcialmente ordenada das classes onde

- n_{\perp} : número de pares de categorias incomparáveis
- $n_{>}$: número de pares de categorias com uma das categorias melhor que a outra

o número de restrições e o número de variáveis do (PLI 2) é dado na Tabela 7.3.

Este número de restrições facilmente se torna muito elevado quando o número de alternativas é elevado o que pode tornar o modelo impraticável. Por exemplo, consideremos uma

Restrições		Número
Classificação	(7.13)-(7.15)	$(n_P + n_I) \times n_{cat}(n_{cat} - 1) + n_R \times n_{cat}$
Unicidade da atribuição	(7.16)	m
$C^i \not\approx C^j, \forall i, j$	(7.17)	$n_{cat} C_2$
Transitividade	(7.18)	$n_{cat}(n_{cat} - 1)(n_{cat} - 2)$
$C^i \neq \emptyset, \forall i$	(7.19)	n_{cat}

Variáveis	Número
Erro $\varepsilon_I, \varepsilon_R, \varepsilon_P$	$(n_P + n_I) \times n_{cat}(n_{cat} - 1) + n_R \times n_{cat}$
Atribuição x_{ij}	$n_{cat} \times m$
Estrutura p_{ij}	$(n_{cat} - 1) \times n_{cat}$

Tabela 7.2: Dimensão de (PLI 1)

estrutura com três classes e 100 alternativas correspondentes a 4.950 pares de alternativas. Consideremos ainda que, por exemplo, $n_I=364$, $n_R=1619$, $n_P=2967$. Neste caso, as dimensões do (PLI 1) são de facto muito maiores que as do (PLI 2), considerando por exemplo uma estrutura fixa *a priori* com 2 classes incomparáveis e uma classe pior do que as outras 2 classes (Tabela 7.4).

7.5 Exemplo ilustrativo do PLI

Nesta secção ilustraremos o modelo proposto neste capítulo através da sua aplicação a um conjunto de dados. Para a obtenção dos resultados que serão apresentados, foi utilizado o *bintprog* do software MATLAB, que usa um algoritmo de programação linear com base no branch-and-bound para resolver problemas de programação inteira binária.

Consideremos $\mathcal{A}=\{a_1, a_2, a_3, a_4, a_5, a_6, a_7\}$ e as suas relações de prevalência apresentadas

Restrições		Número
Classificação	(7.21)-(7.23)	$n_P \times n_{>} + n_I \times 2 \times n_{\perp} + n_R \times n_{cat}$
Unicidade da atribuição	(7.24)	m
$C^i \neq \emptyset, \forall i$	(7.25)	n_{cat}
Variáveis		Número
Erro $\varepsilon_I, \varepsilon_R, \varepsilon_P$		$n_P \times n_{>} + n_I \times 2 \times n_{\perp} + n_R \times n_{cat}$
Atribuição x_{ij}		$n_{cat} \times m$

Tabela 7.3: Dimensão de (PLI 2)

na Tabela 7.5 e as consequentes relações binárias apresentadas na coluna 2 da Tabela 7.6 considerando $\lambda = 0.6$. Considere-se ainda 3 categorias (C^1 , C^2 e C^3), onde $C^1 \succ C^3$ e $C^2 \succ C^3$ (Figura 7.2).

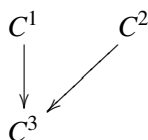


Figura 7.2: Estrutura pré-definida

Dos 21 pares de alternativas de \mathcal{A} , 2 pares são indiferentes, 5 pares são incomparáveis, e os restantes pares verificam a relação de preferência.

Tal como apresentado na Tabela 7.6, o par de alternativas (a_1, a_5) são indiferentes pelo que não devem ser atribuídas às categorias incomparáveis C^1 e C^2 (a_1 ser atribuída a C^1 e a_2 ser atribuída a C^2 ou a_2 a C^1 e a_1 a C^2), originando assim quatro restrições. O mesmo acontece ao par de alternativas indiferentes (a_6, a_7) .

Alternativas incomparáveis não devem pertencer à mesma categoria. Por exemplo, a_1 e a_6 não devem ser atribuídas em simultâneo a C^1 , originando a restrição $x_{11} + x_{61} \leq 1 + \varepsilon_{R_{16,1}}$. O mesmo se aplica à atribuição de a_1 e a_6 a C^2 e a C^3 . Assim, por par de alternativas incomparáveis obtém-se três restrições para o problema (PLI 2) resultando neste exemplo um total de 15 restrições.

Restrições	PL1	PL2
Classificação	28.608	12.247
Unicidade da classificação	100	100
$C^i \neq \emptyset, i = 1, 2, 3$	3	3
$C^i \not\approx C^j, \forall i, j = 1, 2, 3$	3	-
Transitividade	6	-
Total	28.720	12.350

Variáveis	PL1	PL2
x_{ij}	300	300
Erro ε	28.608	12.247
p_{ij}	6	0
Total	28.914	12.547

Tabela 7.4: Dimensão de PLI com 100 alternativas e 3 classes

	a_1	a_2	a_3	a_4	a_5	a_6	a_7
a_1	-	1.00	1.00	0.63	1.00	0.50	0.50
a_2	0.43	-	0.93	0.29	0.50	0.31	0.00
a_3	0.07	0.50	-	0.30	0.31	0.00	0.00
a_4	0.50	1.00	1.00	-	0.50	0.50	0.31
a_5	0.64	1.00	0.83	0.50	-	0.50	0.50
a_6	0.50	1.00	1.00	0.97	0.50	-	0.69
a_7	0.50	1.00	1.00	1.00	0.53	1.00	-

Tabela 7.5: Grau de credibilidade entre alternativas.

Finalmente, uma alternativa fortemente preferível a outra não deverá ser atribuída a uma categoria pior. Por exemplo, para o par (a_1, a_2) onde $a_1 Pa_2$, se a_1 é atribuída à pior categoria C^3 , a_2 não deverá ser atribuída a uma das melhores categorias, C^1 ou C^2 . Tal origina duas restrições: $x_{13} + x_{21} \leq 1 + \varepsilon_{p_{12,31}}$ e $x_{13} + x_{22} \leq 1 + \varepsilon_{p_{12,32}}$. Para as condições de preferência, neste exemplo, tem-se 28 restrições.

A unicidade da atribuição de cada alternativa origina ainda as restrições seguintes:

$$x_{i1} + x_{i2} + x_{i3} = 1, i = 1, \dots, 7 \quad (7.26)$$

Finalmente, caso queiramos que todas as categorias possuam pelo menos uma alternativa, são necessárias mais três restrições:

$$x_{1j} + x_{2j} + x_{3j} + x_{4j} + x_{5j} + x_{6j} + x_{7j} \geq 1, j = 1, 2, 3 \quad (7.27)$$

Relações	Pares	Restrições	
Indiferença	(a_1, a_5)	$x_{1i} + x_{5j} \leq 1 + \epsilon_{I_{15,ij}}$,	$(i,j)=(1,2),(2,1)$
	(a_6, a_7)	$x_{6i} + x_{7j} \leq 1 + \epsilon_{I_{67,ij}}$,	$(i,j)=(1,2),(2,1)$
Incomparabilidade	(a_1, a_6)	$x_{1i} + x_{6i} \leq 1 + \epsilon_{R_{16,i}}$,	$i=1,2,3$
	(a_1, a_7)	$x_{1i} + x_{7i} \leq 1 + \epsilon_{R_{17,i}}$,	$i=1,2,3$
	(a_4, a_5)	$x_{4i} + x_{5i} \leq 1 + \epsilon_{R_{45,i}}$,	$i=1,2,3$
	(a_5, a_6)	$x_{5i} + x_{6i} \leq 1 + \epsilon_{R_{56,i}}$,	$i=1,2,3$
	(a_5, a_7)	$x_{5i} + x_{7i} \leq 1 + \epsilon_{R_{57,i}}$,	$i=1,2,3$
Preferência	(a_1, a_2)	$x_{1i} + x_{2j} \leq 1 + \epsilon_{P_{12,ij}}$,	$(i,j)=(3,1),(3,2)$
	(a_1, a_3)	$x_{1i} + x_{3j} \leq 1 + \epsilon_{P_{13,ij}}$,	$(i,j)=(3,1),(3,2)$
	(a_1, a_4)	$x_{1i} + x_{4j} \leq 1 + \epsilon_{P_{14,ij}}$,	$(i,j)=(3,1),(3,2)$
	(a_2, a_3)	$x_{2i} + x_{3j} \leq 1 + \epsilon_{P_{23,ij}}$,	$(i,j)=(3,1),(3,2)$
	(a_4, a_2)	$x_{4i} + x_{2j} \leq 1 + \epsilon_{P_{42,ij}}$,	$(i,j)=(3,1),(3,2)$
	(a_5, a_2)	$x_{5i} + x_{2j} \leq 1 + \epsilon_{P_{52,ij}}$,	$(i,j)=(3,1),(3,2)$
	(a_6, a_2)	$x_{6i} + x_{2j} \leq 1 + \epsilon_{P_{62,ij}}$,	$(i,j)=(3,1),(3,2)$
	(a_7, a_2)	$x_{7i} + x_{2j} \leq 1 + \epsilon_{P_{72,ij}}$,	$(i,j)=(3,1),(3,2)$
	(a_4, a_3)	$x_{4i} + x_{3j} \leq 1 + \epsilon_{P_{43,ij}}$,	$(i,j)=(3,1),(3,2)$
	(a_5, a_3)	$x_{5i} + x_{3j} \leq 1 + \epsilon_{P_{53,ij}}$,	$(i,j)=(3,1),(3,2)$
	(a_6, a_3)	$x_{6i} + x_{3j} \leq 1 + \epsilon_{P_{63,ij}}$,	$(i,j)=(3,1),(3,2)$
	(a_7, a_3)	$x_{7i} + x_{3j} \leq 1 + \epsilon_{P_{73,ij}}$,	$(i,j)=(3,1),(3,2)$
	(a_6, a_4)	$x_{6i} + x_{4j} \leq 1 + \epsilon_{P_{64,ij}}$,	$(i,j)=(3,1),(3,2)$
	(a_7, a_4)	$x_{7i} + x_{4j} \leq 1 + \epsilon_{P_{74,ij}}$,	$(i,j)=(3,1),(3,2)$
Unicidade da atribuição	$x_{i1} + x_{i2} + x_{i3} = 1,$	$i=1, \dots, 7$	
$C^i \neq \emptyset$	$x_{1j} + x_{2j} + x_{3j} + x_{4j} + x_{5j} + x_{6j} + x_{7j} \geq 1,$	$j=1,2,3$	

Tabela 7.6: Restrições da formulação matemática.

Calculando o valor óptima γ (taxa de erro de classificação), tendo em consideração 57 restrições e 68 variáveis, obtém-se, por exemplo, a seguinte solução óptima: $x_{11} = 1$ e $x_{51} = 1$, ou seja, $\{a_1, a_5\} \leftarrow C^1$; $x_{62} = 1$ e $x_{72} = 1$, ou seja, $\{a_6, a_7\} \leftarrow C^2$; $x_{23} = 1$, $x_{33} = 1$ e $x_{43} = 1$, ou seja, $\{a_2, a_3, a_4\} \leftarrow C^3$. O valor óptimo obtido é zero, ou seja, não existem alternativas mal atribuídas.

7.6 Exemplo - comparação dos resultados do PLI com os do CHOP

À semelhança do exemplo anterior será aqui utilizado o *bintprog* do software MATLAB para obter os resultados que se apresentam de seguida.

Consideremos o exemplo baseado em Hurson e Zopounidis (1997) (Exemplo 1 da Secção 5.4.1) cuja matriz dos desempenhos é dada pela Tabela 5.2 e os valores dos limiares dos critérios são dados pela Tabela 5.3. Considerar-se-á uma vez mais que todos os critérios têm o mesmo peso, ou seja, os coeficientes de importância serão todos iguais a $1/6$ ($k_i = 1/6, i \in \{1, \dots, 6\}$).

Numa tentativa de melhorar as partições deste exemplo onde os resultados obtidos com o modelo CHOP foram melhores que os obtidos com o Clustering-Sorting (ver Capítulo 6), aplicou-se o modelo de optimização (PLI 2) utilizando não só a estrutura com 5 classes sugerida pelo CHOP como uma boa solução do problema, mas também as estruturas com $2 \leq k \leq 6$ obtidas também com o CHOP. Os resultados obtidos apresentam-se na Tabela 7.7.

k	Qualidade CHOP	Qualidade Optimização	Estrutura
6	97.89%	100%	
5	97.37%	100%	
4	96.32%	99.47%	
3	94.21%	97.89%	
2	88.95%	94.74%	

Tabela 7.7: Qualidade obtida com o CHOP e optimização

Em qualquer das partições com $2 \leq k \leq 6$ (supondo que um hipotético decisor só está interessado em ter no máximo 6 classes) conseguiu-se melhorar as partições finais. No caso das partições com $k=5$ classes ou $k=6$ classes conseguiu-se obter uma qualidade de 100% (antes era de 97.89% e 97.37% respectivamente), ou seja, não existem pares de alternativas que sejam incoerentes com o princípio $SS\tau$ -Consistency. Para $k \leq 4$ classes, embora não se consiga atingir uma qualidade de 100%, consegue-se melhorar a partição.

Caso o objectivo seja encontrar a melhor atribuição das acções de \mathcal{A} à estrutura de classes parcialmente ordenada com 5 classes e sugerida pelo CHOP, uma partição final obtida com a

aplicação do modelo (PLI 2) é a seguinte:

- $C^1 = \{a_{19}, a_{20}\}$
- $C^2 = \{a_6, a_8, a_{13}, a_{14}, a_{15}, a_{17}, a_{18}\}$
- $C^3 = \{a_2, a_5, a_7\}$
- $C^4 = \{a_4, a_{16}\}$
- $C^5 = \{a_1, a_3, a_9, a_{10}, a_{11}, a_{12}\}$

Neste caso, os pares de alternativas indiferentes originam 96 restrições e 96 variáveis binárias de erro. Alternativas Incomparáveis originam 305 restrições e 305 variáveis binárias de erro, e as condições de preferência originam 605 restrições e 605 variáveis erro binárias (Tabela 7.8). Note-se contudo que, neste exemplo, apesar do número de alternativas ser pequeno, as dimensões do problema de optimização (PLI 1) tornam-se bastante maiores quando a estrutura é desconhecida *a priori*, como mostra a Tabela 7.8.

Considerando $k=3$ classes, a Tabela 7.9 fornece todas as possíveis estruturas, de entre as quais estão assinaladas a estrutura sugerida pelo CHOP e a solução encontrada nesse caso pelo (PLI 2), a solução dada pelo (PLI 1), e as restantes soluções foram encontradas pelo (PLI 2) para estruturas fixas, por exemplo por um decisor. Como se pode constatar, a estrutura fornecida pelo modelo CHOP pertence ao conjunto de soluções que levam à melhor solução encontrada pelo (PLI 1), tal como aconteceu para $k \geq 4$ classes.

No Capítulo 6 foi ainda aplicado o método CHOP a um exemplo com 100 alternativas. Neste exemplo escolheu-se como solução uma partição com $k=6$ classes e cuja qualidade final corresponde a 99.23%. Dado que o valor obtido da qualidade da partição é bastante bom, optamos por não optimizar o problema com a estrutura sugerida pelo CHOP uma vez que tem uma dimensão bastante elevada. Na verdade, este problema possui 18.924 restrições e 19.030 variáveis binárias 0-1, o que pode trazer problemas a nível de memória de vulgares computadores pessoais.

Relação	Nº Pares	PLI 2		PLI 1	
		Nº Restrições	Nº variáveis	Nº Restrições	Nº variáveis
Indiferença	8	96	96	160	160
Incomparabilidade	61	305	305	305	305
Preferência	121	484	484	2420	2420
Unicidade da atribuição	-	20	0	20	0
$C^i \not\subset C^j$	-	0	0	0	10
Transitividade	-	0	0	0	24
$C^i \neq \emptyset, \forall i$	-	5	0	5	0
Atribuição (x_{ij})	-	-	100	-	100
Estrutura (p_{ij})	-	-	0	-	20
Total	190	910	985	2910	3039

Tabela 7.8: Restrições da formulação matemática.

γ^*	Qualidade	Estrutura	
2.11% (4/190)	97.89%	○ ○ ○	PLI 2 + Decisor
2.11% (4/190)	97.89%	○ ← ○ → ○	(PLI 1)
4.21% (8/190)	95.79%	○ → ○ ← ○	PLI 2 + Decisor
2.11% (4/190)	97.89%	○ ○ → ○	(PLI 2 + CHOP)
4.74% (9/190)	95.26%	○ → ○ → ○	PLI 2 + Decisor

Tabela 7.9: Qualidade obtida para as diferentes partições de k=3.

7.7 Notas conclusivas

Neste capítulo propusemos um método exacto para a classificação de um conjunto \mathcal{A} de alternativas, com base em múltiplos critérios, a um conjunto de categorias parcialmente ordenadas com uma estrutura que pode ou não estar definida *a priori*. Foi proposto um algoritmo

de optimização para obter a melhor partição definida com base nas relações de prevalência das alternativas e no princípio ($SS\tau$)- Consistency. O objectivo do modelo proposto é minimizar a razão entre o número de pares de alternativas que não verificam o princípio ($SS\tau$)-Consistency e o número total de pares de alternativas existentes.

Quanto à estrutura das classes parcialmente ordenadas, esta pode ser fornecida pelo decisor ou, caso não tenha qualquer conhecimento prévio sobre o número de classes a escolher nem das suas relações de preferência, pode ser aplicado previamente uma das heurísticas, Clustering-Sorting ou CHOP, propostas nos capítulos anteriores. Caso se trate de um problema relativamente pequeno (com muito poucas alternativas), pode-se otimizar em simultâneo a estrutura e a partição correspondente, ou seja, trata-se de um problema cuja estrutura é desconhecida *a priori*. Caso contrário, a aplicação do PLI 1 torna-se impraticável.

Com a aplicação inicial ao problema em estudo de um modelo heurístico, Clustering-Sorting ou CHOP, o esforço computacional é enormemente reduzido pois não obriga à procura da melhor estrutura em simultâneo com a melhor partição. Tal poderá fornecer ao decisor uma solução do seu problema em tempo real, facilitando a comunicação com este. Apesar das vantagens resultantes da aplicação do algoritmo exacto (PLI 2) após um modelo heurístico, a complexidade agrava-se em problemas de grandes dimensões e/ou com muitas classes. Pode mesmo tornar-se impraticável tratar estes problemas através do método exacto apresentado. Na verdade, nas experiências que fizemos com o problema de 20 alternativas, usando o sistema computacional desenvolvido para o CHOP, notámos que o cálculo de cada solução era instantâneo. Mesmo para problemas de maior dimensão verificamos uma relativa rapidez. O mesmo não se pode dizer quando se usou o modelo exacto que com os casos mais pequenos demora minutos, passando rapidamente a demorar muitas horas para exemplos um pouco maiores.

Uma alternativa para esses casos será a de recorrer apenas a técnicas heurísticas, Clustering-Sorting ou CHOP, para calcular soluções aproximadas das soluções exactas em tempos computacionais razoáveis. Embora as heurísticas não garantam a optimalidade, o tempo gasto ser reduzido é um factor bastante importante para o decisor. Note-se ainda que a qualidade das aproximações obtidas foram bastante boas o que torna estas heurísticas bastante apelativas. Na verdade, foi usado um exemplo baseado em Hurson e Zopounidis (1997) para ilustrar o modelo, aplicado no capítulo anterior, e verificou-se que embora houvesse uma melhoria da qualidade das partições obtidas com o modelo CHOP, esta não era significativa. Por outro lado, as estruturas sugeridas pelo CHOP levaram às partições com melhor qualidade quando aplicado o

(PLI 2), qualidade essa confirmada pelo (PLI 1).

Classificação Multicritério de grupos de trabalho a classes desconhecidas

Neste capítulo, apresentamos uma aplicação das heurísticas desenvolvidas nos Capítulos 5 e 6 a um problema real com dados de Dimas (2007) sobre conflitos intra-grupo e estratégias de gestão desses conflitos em grupos de trabalho do departamento de produção de diversas empresas industriais.¹ O objectivo deste estudo é comparar grupos de trabalho: um grupo de trabalho pode ser atribuído a uma classe melhor ou pior em comparação com outros grupos de trabalho, podendo ainda ser considerado incomparável aos outros grupos no que diz respeito à incidência de conflitos, bem como a forma que este grupo de trabalho gere os conflitos.

8.1 Notas introdutórias

O conflito está presente nas relações inter-pessoais (Pruitt e Carnevale, 1993), nas relações intra-grupo e inter-grupo (Jehn, 1995), em áreas estratégicas de tomadas de decisão (Amason, 1996), entre outras situações organizacionais. Tal como muitos outros autores, Jehn (1995) distingue dois tipos de conflitos dentro do grupo: conflitos tarefa e conflitos afectivos. Os *conflitos tarefa* englobam situações de tensão vividas no grupo devido à presença de diferentes perspectivas em relação à execução de uma tarefa. A existência de desacordos entre os elementos de um grupo quanto à melhor forma de alcançar os objectivos comuns é inevitável pois um grupo é constituído por indivíduos que percebem a realidade de diferentes pontos de vista, ideias e opiniões. Os *conflitos afectivos* envolvem situações de incompatibilidade inter-pessoal entre os membros do grupo. De acordo com Amason e Sapienza (1997) este tipo de conflito degrada a qualidade das relações e dificulta a circulação da informação, diminuindo, desta forma, o grau de identificação dos membros com o grupo e com as suas decisões. A energia do grupo

¹Parte deste Capítulo é uma adaptação do artigo publicado nos Proceedings of the 25th Mini-EURO Conference on Uncertainty and Robustness in Planning and Decision Making (Rocha e Dias, 2010).

é dirigida para o desenvolvimento da coesão e para a resolução de problemas interpessoais, o que, na perspectiva de Rentsch e Zelno (2003), pode ter efeitos muito negativos ao nível da eficácia grupal.

As consequências produzidas pelas situações conflituosas ao nível da eficácia do grupo, nomeadamente no que diz respeito à qualidade das decisões, aos níveis de criatividade e inovação e à coesão do grupo, dependem da forma como são geridas no grupo e do momento em que emergem. As *estratégias da gestão de conflitos* podem ser definidas segundo dois pontos de vista: inter-pessoal e intra-grupo. Ao nível inter-pessoal traduzem-se como as reacções individuais à percepção de que existem ideias, opiniões e/ou objectivos divergentes entre as partes envolvidas numa interacção (Deutsch, 1973; Pruitt e Rubin, 1986), divergências essas que geram elevados níveis de tensão. Ao nível intra-grupo, traduzem-se no conjunto de respostas dadas pelos membros de um grupo com o fim de reduzir ou solucionar um determinado conflito (DeChurch e Marks, 2001). Rahim e Bonoma (1979) conceptualizaram as estratégias de gestão de conflitos numa lógica bidimensional: preocupação consigo que é o grau em que o indivíduo procura satisfazer os seus objectivos, e preocupação com os outros que é o grau em que o indivíduo procura satisfazer os objectivos dos outros. Estas duas dimensões quando se combinam entre si dão origem a cinco estilos de gestão de conflitos: integração, acomodação, compromisso, domínio e evitamento².

A estratégia de *integração* caracteriza-se pela partilha de informação entre as partes envolvidas tendo em vista o encontrar de uma solução aceitável para todos. Na estratégia de *acomodação* as diferenças existentes são minimizadas e são sobrevalorizados os pontos de acordo, procurando-se, desta forma, a satisfação dos objectivos da outra parte. Em certa medida, supõe um sacrifício para a parte em conflito que adopta esta estratégia. A estratégia de *evitamento* caracteriza-se por uma atitude de fuga ou de negação do problema existente. Em algumas situações pode constituir um meio de adiar a resolução do conflito para um momento posterior, enquanto noutras permite a retirada de uma situação sentida como ameaçadora. Na estratégia de *domínio* o alcance dos objectivos de uma das partes é considerado prioritário perante os interesses da outra parte. Quando numa situação de conflito grupal, ambas as partes envolvidas adoptam esta linha estratégica, a rigidez na prossecução dos objectivos torna difícil o encontrar de uma solução partilhada.

²Na linha de autores como Pruitt (1983) e Dimas(2007), considera-se que o compromisso não é uma estratégia distinta da integração.

8.2 Apresentação do problema

O exemplo real que é aqui apresentado baseia-se nos dados de Dimas (2007) com 77 equipas de trabalho de 14 empresas caracterizadas pelos seus desempenhos em 6 critérios: 2 critérios para o tipo de conflito - Tarefa (g_1) e Afectivo (g_2) - e 4 critérios para as estratégias de gestão dos conflitos - Integração (g_3), Domínio (g_4), Acomodação (g_5) e Evitamento (g_6). Os critérios avaliam a frequência média com que cada uma das situações apresentadas acontece no grupo (1 = “nunca acontece”; 7 = “acontece sempre”).

Para avaliar as estratégias de gestão de conflitos utilizadas pelos membros do grupo, Dimas realizou uma adaptação da escala ROCI-II³ (Rahim, 1983), integrando 15 itens (4 relativos à dimensão integração, 4 ao domínio, 3 ao evitamento e 4 à acomodação) avaliados pelos membros das equipas numa escala de Likert de 7 pontos, que varia entre “nunca acontece”(1) e “acontece sempre”(7).

Sendo o conflito intra-grupo o tema central do estudo, o nível de análise é o grupo e não o indivíduo. Contudo, os dados foram recolhidos dos membros das equipas de trabalho analisadas, havendo portanto necessidade de obter informação sobre cada equipa de trabalho. Para tal, calcularam-se as pontuações médias obtidas pelos membros do grupo nos instrumentos de medida utilizados. Na verdade, a agregação dos dados permite reduzir o impacto das diferenças individuais dentro de cada equipa, conseguindo-se, desta forma, obter indicadores mais representativos das características dos grupos (Simons e Peterson, 2000). Antes de se proceder à agregação dos resultados foi, contudo, necessário calcular um índice de acordo inter-juízes para garantir que as pontuações médias obtidas fossem representativas das percepções individuais relativas à realidade grupal - índice *Average Deviation*. Os valores resultantes de cada equipa para cada um dos critérios são apresentados na Tabela 8.1.

8.3 Clustering-Sorting

Aplicando o método Clustering-Sorting descrito no Capítulo 5, começamos por identificar grupos de trabalho compactos em relação às suas várias características e, numa segunda fase, identificamos as relações de preferência dos grupos de trabalho obtidos na primeira fase, de acordo com os conflitos dentro do grupo e da forma como são geridos, ou seja, de acordo com o funcionamento do grupo de trabalho.

³Rahim Organizational Conflict Inventory

	g_1	g_2	g_3	g_4	g_5	g_6		g_1	g_2	g_3	g_4	g_5	g_6
a_1	3.4	3.19	5.19	2.88	3.69	5.0	a_{40}	2.7	2.2	4.875	2.25	3.125	3.277
a_2	2.86	2.33	4.33	3.42	3.66	4.33	a_{41}	2.4	1.687	6.189	2.31	4.125	2.833
a_3	2.3	2.5	5.25	2.75	3.75	3.5	a_{42}	2.1	2.625	5.0	2.25	3.5	3.5
a_4	3.68	3.2	5.25	3.5	3.95	3.4	a_{43}	2.24	2.5	6.15	1.9	3.4	3.8
a_5	2.88	2.15	4.85	3.0	4.15	4.266	a_{44}	1.6	2.0	6.125	1.5	2.5	4.5
a_6	3.2	2.667	5.5	2.33	2.92	3.77	a_{45}	2.2	3.625	6.875	2.125	3.5	4.33
a_7	3.5	2.75	4.875	3.375	3.5	3.33	a_{46}	2.2	2.5	5.125	1.5	2.375	2.0
a_8	3.2	2.78	5.66	2.91	3.28	3.46	a_{47}	2.7	2.125	5.75	2.625	2.75	3.58
a_9	2.8	3.0	5.33	3.25	4.08	4.77	a_{48}	3.3	2.0	4.375	1.625	3.25	3.667
a_{10}	2.2	2.375	5.25	2.875	2.875	4.5	a_{49}	1.8	2.375	6.25	2.0	2.75	4.166
a_{11}	2.4	3.625	5.63	2.625	5.0	4.5	a_{50}	1.93	2.08	6.25	2.08	3.0	3.888
a_{12}	2.93	2.86	5.125	3.07	3.75	3.79	a_{51}	2.2	2.0	6.75	1.375	1.875	3.5
a_{13}	3.266	3.166	4.75	3.08	3.25	3.22	a_{52}	3.8	3.75	4.13	4.0	4.13	4.83
a_{14}	2.52	2.85	5.25	2.95	4.2	4.266	a_{53}	3.6	3.5	3.125	3.0	4.37	4.166
a_{15}	3.2	2.625	6.125	2.875	2.25	4.5	a_{54}	3.53	3.166	4.75	3.58	4.08	4.22
a_{16}	2.8	1.87	5.375	3.625	3.125	3.33	a_{55}	4.0	3.375	3.75	3.625	3.25	4.0
a_{17}	3.1	3.44	6.0	3.125	3.125	3.166	a_{56}	3.3	3.875	3.125	4.0	3.88	4.166
a_{18}	3.13	3.08	5.667	2.25	3.75	4.333	a_{57}	3.13	3.58	4.0	3.83	3.666	4.55
a_{19}	3.0	2.125	6.5	2.75	3.63	6.0	a_{58}	3.33	3.0	4.416	3.58	4.08	4.22
a_{20}	3.54	2.96	5.535	2.928	3.785	3.52	a_{59}	3.6	2.5	4.25	3.5	3.94	4.416
a_{21}	3.32	2.65	5.3	2.95	3.15	4.4	a_{60}	2.8	3.13	4.0	3.63	3.63	4.166
a_{22}	3.275	2.06	5.875	2.56	3.656	3.54	a_{61}	4.8	3.92	3.08	3.83	3.42	4.11
a_{23}	3.1	2.75	5.25	3.25	4.0	5.33	a_{62}	3.2	4.0	5.0	3.75	3.25	4.22
a_{24}	3.8	3.0	6.0	2.75	4.125	4.0	a_{63}	3.44	3.25	4.35	4.0	3.6	3.73
a_{25}	2.775	3.03	5.06	2.63	3.53	3.79	a_{64}	3.488	2.527	3.92	3.055	4.14	5.22
a_{26}	2.98	3.045	4.52	3.11	3.43	3.697	a_{65}	3.666	4.08	3.75	3.666	3.58	4.0
a_{27}	1.8	1.625	5.75	1.5	3.25	4.5	a_{66}	3.48	3.29	4.73	3.71	3.85	4.25
a_{28}	2.3	2.125	6.0	1.5	2.75	3.0	a_{67}	4.2	3.89	4.57	3.928	3.57	4.24
a_{29}	2.3	2.5	4.625	2.0	2.875	3.166	a_{68}	3.58	2.84	4.5	3.45	3.43	3.845
a_{30}	2.6	3.125	6.0	1.75	2.0	3.0	a_{69}	3.45	3.75	4.81	3.56	3.44	4.17
a_{31}	2.93	2.33	5.5	1.83	2.33	3.33	a_{70}	3.8	2.75	5.125	4.875	3.75	4.33
a_{32}	1.65	1.375	6.25	1.69	3.5	4.42	a_{71}	3.49	2.977	4.636	3.159	3.59	3.939
a_{33}	2.8	2.75	5.58	1.83	3.58	3.666	a_{72}	3.65	3.08	4.9167	3.81	3.65	3.888
a_{34}	2.34	2.25	5.357	2.0	3.5	3.76	a_{73}	3.91	3.68	4.46	3.21	3.11	3.71
a_{35}	1.466	1.416	5.58	1.92	3.0	4.11	a_{74}	2.5	3.5	3.0	1.63	1.88	2.0
a_{36}	2.53	2.75	5.83	1.916	3.5	3.555	a_{75}	3.4	4.06	4.19	1.63	3.06	4.0
a_{37}	1.8	1.75	5.25	2.75	3.0	3.33	a_{76}	3.7	3.63	4.625	2.63	3.63	3.0
a_{38}	3.0	2.06	5.375	1.25	3.0	2.916	a_{77}	3.75	3.31	4.125	2.5	3.187	3.0
a_{39}	2.45	2.0	5.44	1.875	3.187	3.0							

Tabela 8.1: Desempenhos das 77 alternativas - equipes de trabalho.

1. Primeira fase - Clustering

Na primeira fase, os desempenhos das alternativas (ou seja, dos grupos de trabalho) foram padronizados, já que se usa o método K-Means com base numa distância euclidiana. O número de clusters (desconhecido a priori) foi determinado à custa do critério da distância entre clusters (quadrado da distância Euclidiana) aplicando o método *Complete Linkage*, e à custa do *R-squared* (Figura 8.1).

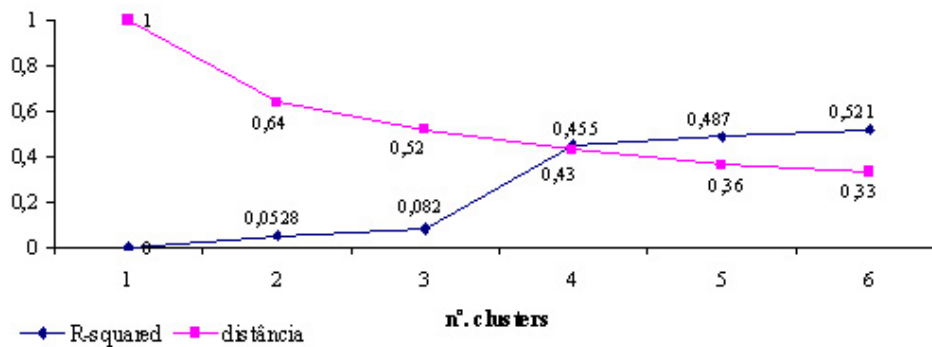


Figura 8.1: Distância entre clusters e R-squared

Pela figura 8.1 observa-se que os declives das rectas que unem as distâncias entre dois clusters é relativamente pequeno a partir de 2 clusters. Contudo, atendendo ao conhecimento empírico do fenómeno em estudo, não é facilmente interpretável a recomendação de 2 clusters. Assim, complementando esta análise com a informação do *R-squared*, verificou-se que a variabilidade retida por mais de 4 categorias é relativamente pequena quando comparada com a evolução de 1 a 4. Assim, uma boa solução é reter 4 categorias: C^4 com 4 alternativas, C^3 com 22 alternativas, C^2 com 25 alternativas, a C^1 com 26 alternativas.

2. Segunda fase- Eliciação

Tal como indica a maior parte dos estudos empíricos desenvolvidos no âmbito do conflito intragrupal, o conflito em si mesmo não é um fenómeno positivo para a eficácia do grupo. Na realidade, um conflito é sempre uma experiência negativa, emocionalmente desagradável, geradora de insegurança e frustração.

Por outro lado, embora o conflito grupal seja perspectivado pela literatura da especialidade como positivo para o desempenho do grupo, desde que centrado nas tarefas, na verdade, os resultados empíricos que têm sido encontrados nas investigações apontam para efeitos negati-

vos (Dimas,2007). Assim, dado que o estudo incidiu sobre equipas de trabalho, que têm sob a sua responsabilidade a execução de um conjunto de tarefas, que têm que cumprir objectivos, muitas vezes exigentes e difíceis e pelo alcance dos quais são avaliadas, considerar-se-à, tal como indica Dimas (2007) no seu estudo, que os conflitos relacionados com tarefas ameaçam os resultados grupais.

Assim sendo, os conflitos serão a *minimizar*, ou seja

g_1 : Conflito tarefa [a minimizar].

g_2 : Conflito afectivo [a minimizar].

Relativamente à forma como os conflitos são geridos, estas podem traduzir ganhos grupais significativos. No estudo de Dimas verificou-se que a utilização de estratégias de gestão de conflitos de carácter integrativo se traduziu num aumento da eficácia grupal. Equipas com o mesmo nível de conflitos, mas que utilizam com maior frequência estratégias integrativas, têm desempenhos superiores e os seus membros apresentam níveis mais elevados de satisfação.

Assim sendo considerar-se-á, neste estudo, que a forma mais construtiva de gerir os conflitos é a estratégia integrativa que será a única que será a *maximizar*. Todas as restantes, serão a *minimizar*. Assim, tem-se:

g_3 : Integração [a maximizar].

g_4 : Domínio [a minimizar].

g_5 : Acomodação [a minimizar]

g_6 : Evitamento [a minimizar].

Como na literatura não há qualquer informação que nos forneça pistas sobre o peso dos critérios, consideramos que todos os critérios têm o mesmo peso, ou seja, ou seja, $k_j=1/6$, $j=1,\dots,6$. Quanto aos limiares dos critérios, estes são apresentados na Tabela 8.2. O limiar de corte λ toma o valor 0.86.

	g_1	g_2	g_3	g_4	g_5	g_6
q_j	0.25	0.25	0.25	0.25	0.25	0.25
p_j	0.75	0.75	0.75	0.75	0.75	0.75
v_j	2	2	2	2	2	2

Tabela 8.2: Limiares de Indiferença, Preferência e Veto.

Analisando o grupo de alternativas segundo os seus desempenhos, coeficientes de importância, limiares de indiferença, preferência e veto e fazendo $\lambda = 0,86$, dos 2926 pares de alternativas estudadas ($\frac{n(n-1)}{2}$, $n = 77$), 1936 (66,16%) verificam a relação de Preferência, 54

(1,85%) são Indiferentes, e 936 (31,99%) são Incomparáveis (Tabela 8.5).

3. Terceira fase - Classificação Ordinal

Na terceira fase, para ordenação das categorias, apresentamos duas sugestões: ordenação dos centróides das quatro categorias obtidas na primeira fase e aplicação da relação de preferência mais comum.

Para tal, calculamos o centróide de cada uma das categorias obtidas (c_1 , c_2 , c_3 e c_4), e, em seguida, classificamos os centróides com base nas suas relações de prevalência (Tabela 8.3) obtidas pela aplicação do método ELECTRE com os parâmetros que foram fixos (método (A)). Para $\lambda=0.86$ obtém-se $c_1Pc_3, c_2Pc_1, c_2Pc_3, c_2Pc_4, c_4Pc_3$ e c_1Rc_4 . Assim a estrutura final obtida é parcial (Figura 8.2), sendo C^2 a categoria melhor, seguida de C^1 e C^4 (as quais são incomparáveis), e a pior categoria é C^3 .

centroide	c_1	c_2	c_3	c_4
c_1	-	0.3483	1.0000	0.5314
c_2	1.0000	-	1.0000	0.8994
c_3	0.4523	0.0590	-	0.3417
c_4	0.4084	0.2360	0.9915	-

Tabela 8.3: Grau de credibilidade.

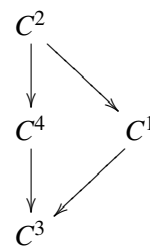


Figura 8.2: Partição final pelo método (A)

Aplicando a relação mais comum entre as alternativas das categorias (Tabela 8.4), método (B), obtemos a estrutura final da Figura 8.3. Comparando a estrutura obtida com a estrutura do método (A), verificamos que C^3 continua a ser considerada a pior categoria, e C^1 e C^4 são incomparáveis. Quanto a C^2 , esta deixa de ser considerada a melhor categoria e passa a ser incomparável a C^1 e C^4 .

$C^1 \perp C^2$	$C^1 \mathbf{P} C^2$: 189	$C^2 \mathbf{P} C^1$: 187	$C^1 \mathbf{R} C^2$: 217	$C^1 \mathbf{I} C^2$: 9
$C^1 \succ C^3$	$C^1 \mathbf{P} C^3$: 253	$C^3 \mathbf{P} C^1$: 106	$C^1 \mathbf{R} C^3$: 165	$C^1 \mathbf{I} C^3$: 12
$C^1 \perp C^4$	$C^1 \mathbf{P} C^4$: 29	$C^4 \mathbf{P} C^1$: 29	$C^1 \mathbf{R} C^4$: 35	$C^1 \mathbf{I} C^4$: 1
$C^2 \succ C^3$	$C^2 \mathbf{P} C^3$: 220	$C^3 \mathbf{P} C^2$: 115	$C^2 \mathbf{R} C^3$: 177	$C^2 \mathbf{I} C^3$: 7
$C^2 \perp C^4$	$C^2 \mathbf{P} C^4$: 26	$C^4 \mathbf{P} C^2$: 31	$C^2 \mathbf{R} C^4$: 34	$C^2 \mathbf{I} C^4$: 1
$C^4 \succ C^3$	$C^3 \mathbf{P} C^4$: 14	$C^4 \mathbf{P} C^3$: 35	$C^4 \mathbf{R} C^3$: 33	$C^4 \mathbf{I} C^3$: 3

Tabela 8.4: Número de pares de alternativas para cada relação de preferência.

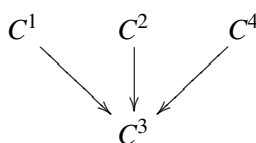


Figura 8.3: Partição final pelo método (B)

4. Quarta fase - Robustez

Depois de obter a partição final, analisamos a percentagem de pares de alternativas inconsistentes com o princípio $SS\tau$ -Consistency. Os resultados são apresentados na Tabela 8.5. No cálculo da qualidade da cada partição, $Q(\mathcal{P})$, consideramos equivalentes as violações das condições de Preferência, Indiferença e Incomparabilidade, ou seja, consideramos $\alpha_P = \alpha_I = \alpha_R = 1$.

Condições	Relação S	Clustering-Sorting (A)	Clustering-Sorting (B)
Preferência	1936 (66.16%)	453 (15.5%)	235 (8.0%)
Indiferença	54 (1.85%)	1 (0.03%)	11 (0.37%)
Incomparabilidade	936 (31.99%)	275 (9.4%)	275 (9.4%)
Total	2926	729 (24.93%)	521 (17.77%)
$Q(\mathcal{P})$	-	75.07%	82.23%

Tabela 8.5: Número de pares de alternativas que verificam as condições de P, I e R antes da partição e o número de inconsistências pós-partição segundo o $SS\tau$ -Consistency

Perante duas soluções diferentes para o problema, o decisor pode optar, por exemplo, pela partição que tem a melhor qualidade, que neste caso é a partição obtida pelo método (B), com uma qualidade de 82.23%.

8.4 CHOP

Aplicando agora o método CHOP ao problema, obtemos os resultados apresentados na Tabela 8.6.

k	Transitividade	Qualidade	(v_P, v_I, v_R)	Estrutura
7	✓	83.11	(0 0.0038 0.1651)	
6	✓	79.35	(0 0.0038 0.2027)	
5	✓	76.04	(0 0.0031 0.2365)	
4	✓	74.23	(0 0.0021 0.2556)	
3	✓	74.13	(0 0.0017 0.2570)	
2	✓	68.66	(0 0.0014 0.3120)	
1	✓	68.01	(0 0 0.3199)	

Tabela 8.6: Agregações das classes e Qualidade obtida

Pela análise da Figura 8.4 podemos ver que a qualidade das partições vai aumentando até $k=7$ classes, mantendo a partir daí uma qualidade bastante similar e aproximadamente igual a 83%. Para $k=13$ volta a melhorar ligeiramente para 86%. Note-se que só a partir de $k \geq 26$ se consegue obter uma qualidade das partições superior a 90%.

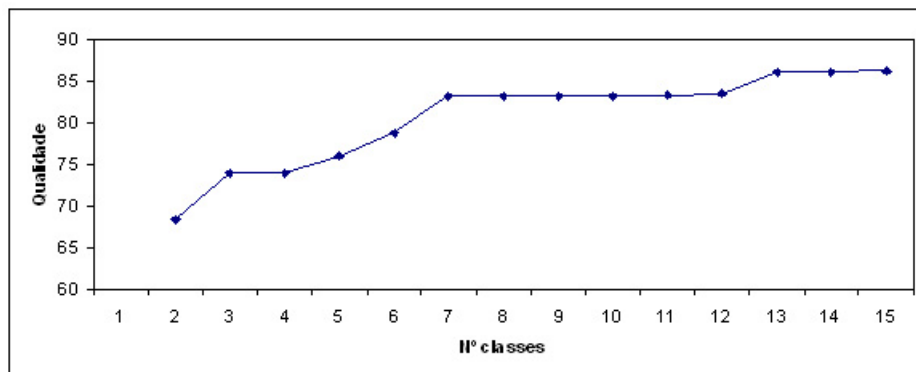


Figura 8.4: Qualidade das partições.

Assim, considerando que o decisor opta por $k=7$ classes como uma boa solução, obtém-se uma partição com uma qualidade de 83.11% resultante de 0% de violações das condições

de Preferência e 0.38% de Indiferença, e 16.51% de Incomparabilidade, cuja partição parcialmente ordenada e transitiva resultou da avaliação das relações de prevalência das classes (Tabela 8.7).

	C^1	C^2	C^3	C^4	C^5	C^6	C^7
C^1	-	0.2381	0.3598	0.2381	0.2487	0.1349	0.4444
C^2	0.1799	-	0.1111	0.0000	0.0000	0.0000	0.0000
C^3	0.1640	0.1111	-	0.0000	0.0000	0.1667	0.0000
C^4	0.3810	0.3333	0.6667	-	0.1667	0.2500	0.0000
C^5	0.4974	0.2222	0.2222	0.1667	-	0.3333	1.0000
C^6	0.2302	0.0000	0.1667	0.0000	-	0.0000	0.5000
C^7	0.3968	0.0000	0.0000	0.0000	0.0000	0.0000	-

Tabela 8.7: Valores do grau de similaridade s_{ij} entre as classes (C^i, C^j), $i, j=1, \dots, 7$

Realizaram-se ainda algumas experiências para $\lambda < 0.86$ e verificou-se alguma sensibilidade face ao λ , o que, no caso do método CHOP, leva à obtenção de estruturas alternativas. Por exemplo, $\lambda=0.6$ conduz aos resultados apresentados na Tabela 8.8 (k=4 classes) e a estrutura sugerida pelo CHOP passa a ser a representada na figura 8.5.

Condições	Relação S	Clustering-Sorting	CHOP
Preferência	2077 (70.98%)	11 (0.37%)	0 (0.0%)
Indiferença	641 (21.92%)	10 (0.34%)	10 (0.34%)
Incomparabilidade	208 (7.09%)	164 (5.6%)	89 (3.52%)
Total	2926	185 (6.31%)	99 (3.86%)
$Q(\mathcal{P})$	-	93.69%	96.14%

Tabela 8.8: Resultados para $\lambda=0.6$

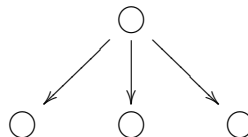


Figura 8.5: Partição final do CHOP para $\lambda=0.6$.

Tal estrutura não identifica uma classe pior que as restantes o que, perante o conhecimento que tínhamos do problema, levou-nos a excluir esta solução. O mesmo aconteceu com as es-

truturas sugeridas pelo CHOP para outros valores de λ , as quais consideramos que não eram as mais indicadas para o problema em estudo. Note-se contudo que, perante decisores com diferentes interpretações do problema, tais estruturas podem ser consideradas mais interessantes que as obtidas com $\lambda=0.86$.

8.5 Decisão final

Comparando os resultados do *Clustering-Sorting* com os do CHOP para $\lambda=0.86$, verificamos que se obtiveram melhores resultados no primeiro método aplicado. Por exemplo, para $k=4$ verificamos que o CHOP melhora significativamente a condição de preferência do princípio $SS\tau$ -Condition, embora piore significativamente a condição de incomparabilidade, levando a uma qualidade da partição bastante mais baixa do que a obtida na partição obtida com o *Clustering-Sorting*.

Como consideramos que todas as condições têm o mesmo peso, uma boa solução poderá ser aquela que conduz à melhor qualidade da partição obtida. Assim, a solução escolhida é a estrutura parcialmente ordenada fornecida pelo método (B) do *Clustering-Sorting* (Figura 8.3), com quatro categorias, tendo uma qualidade final de 82.23%. As categorias obtidas são as seguintes:

- $C^1 = \{a_3, a_5, a_{12}, a_{15}, a_{20}, a_{25}, a_{27}, a_{28}, a_{32}, a_{36}, a_{37}, a_{39}, a_{40}, a_{41}, a_{43}, a_{44}, a_{46}, a_{48}, a_{58}, a_{59}, a_{61}, a_{63}, a_{66}, a_{69}, a_{76}, a_{77}\}$
- $C^2 = \{a_6, a_{13}, a_{14}, a_{16}, a_{17}, a_{18}, a_{19}, a_{21}, a_{22}, a_{24}, a_{30}, a_{31}, a_{33}, a_{34}, a_{35}, a_{42}, a_{45}, a_{50}, a_{52}, a_{65}, a_{67}, a_{68}, a_{70}, a_{71}, a_{72}\}$
- $C^3 = \{a_1, a_2, a_7, a_8, a_{11}, a_{23}, a_{26}, a_{29}, a_{38}, a_{49}, a_{51}, a_{53}, a_{54}, a_{55}, a_{56}, a_{57}, a_{60}, a_{62}, a_{64}, a_{73}, a_{74}, a_{75}\}$
- $C^4 = \{a_4, a_9, a_{10}, a_{47}\}$.

De acordo com a informação dada pelos centróides de cada categoria, verificamos que a categoria C^2 reflecte um contexto grupal caracterizado por uma emergência muito reduzida de

conflitos, quer conflitos afectivos que envolvem situações de incompatibilidade interpessoal entre os membros do grupo, quer conflitos tarefa que surgem com as diferentes perspectivas em relação à execução de uma tarefa. De facto, as pontuações médias situam-se ambas aproximadamente na pontuação 2 (*acontece muito pouco*). Quanto ao grupo C^1 , neste ocorrem em média poucos conflitos, tanto de tarefa como afectivos já que as pontuações médias situam-se ambas aproximadamente na pontuação 3 (*acontece pouco*). Os restantes grupos reflectem um contexto grupal caracterizado por um aumento de conflitos (Figura 8.6).

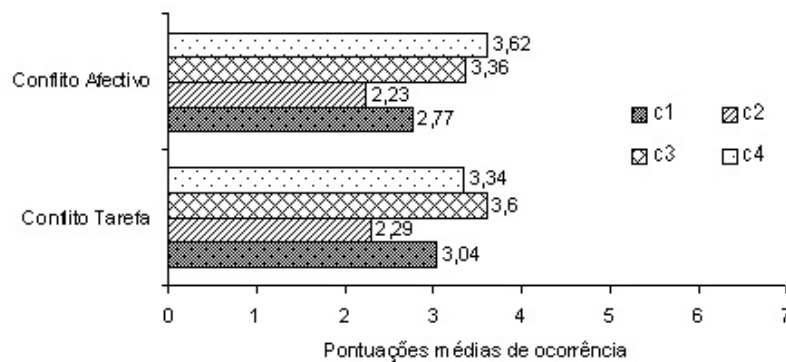


Figura 8.6: Pontuações médias do conflito intragrupal em cada cluster

Quanto à forma de gestão destes conflitos, tal como mostra a Figura 8.7, são as estratégias integrativas as mais utilizadas pelos grupos analisados, ou seja, partilham informação com o objectivo de encontrar uma solução aceitável para todos. Contudo, a sua incidência nos grupos C^1 e C^2 é de facto maior que nos restantes grupos (o grupo C^2 utiliza em média *Muitas vezes* (pontuação média próxima de 6) seguido do grupo C^1 que utiliza *Bastantes vezes* (pontuação média próxima de 5). Por outro lado, as estratégias de domínio são as menos utilizadas, sendo C^3 o grupo que mais a utiliza, ou seja, é o grupo mais marcado pela luta constante entre os membros a respeito dos papéis, objectivos e procedimentos, onde cada um tenta impor as suas perspectivas sobre os problemas e a percepção das semelhanças é baixa.

8.6 Conclusões

Foi feita uma aplicação dos métodos *Clustering-Sorting* e CHOP a dados reais de um conjunto de 77 grupos de trabalho avaliados pelo tipo de conflitos intra-grupais e gestão desses

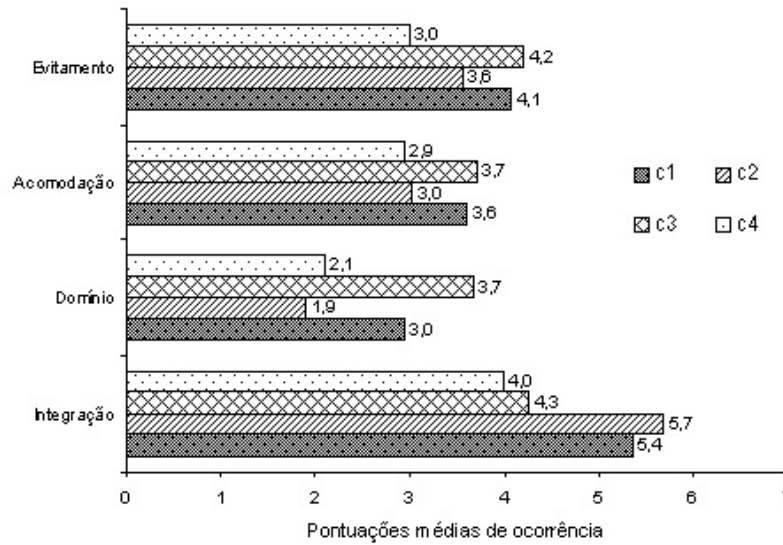


Figura 8.7: Pontuações médias das estratégias de gestão de conflitos

conflitos. Um dos motivos que nos levou a testar os métodos heurísticos por nós desenvolvidos com este caso real foi a existência de poucos estudos que aplicam a análise de cluster às Ciências Organizacionais, em particular, no campo de pesquisa de grupo. Uma das exceções é o estudo desenvolvido por Munduate et al. (1999) onde os autores procuram identificar grupos de trabalho com diferentes padrões relativamente ao tipo de conflitos e estratégias utilizadas. Outro dos motivos que nos levou a testar os métodos heurísticos no referido caso real foi a não existência de estudos para detectar potenciais relações de ordem parcial que possam existir entre os grupos. Por outro lado, achamos que o método exacto de optimização sugerida no Capítulo 7 não será a melhor opção para este caso, dada a dimensão do problema (daria origem a um problema linear com 15735 restrições, 16004 variáveis binárias 0-1, e ainda as *slacks* que são 15658 variáveis. Tal levaria a uma matriz de 498 201 570 elementos). Assim consideramos que os métodos heurísticos desenvolvidos poderão ser uma mais valia para problemas desta natureza e dimensão.

Verificamos alguma sensibilidade face ao λ e escolhemos o valor ($\lambda=0.86$) que nos forneceu as soluções que consideramos mais indicadas para o problema em estudo. Foram apresentadas três possíveis soluções para o problema, duas fornecidas pelo método *Clustering-Sorting* e outra pelo CHOP. Destas soluções, escolhemos para solução final a que apresenta o número de classes que melhor se adequa ao problema e a que apresenta menos incoerências segundo o

princípio $SS\tau$ -Consistency e que é dada pelo *Clustering-Sorting* (método (B)). Nesta solução foi possível conceber que os grupos de trabalho tendem a agrupar-se em 4 diferentes categorias. A estrutura final obtida para essas categorias foi uma ordem parcial: uma categoria pior e três categorias incomparáveis.

Os resultados encontrados apresentam a pior categoria C^3 como disfuncional, estando de acordo com a literatura (De Dreu et al., 2004; Kuhn e Poole, 2000; Friedman et al., 2000; Pneuman e Bruehl, 1982; DeChurch e Marks, 2001). Na verdade, os grupos de trabalho desta categoria têm mais conflitos do que os grupos em C^1 e C^2 . A colaboração com os outros ao tentar encontrar uma solução mutuamente aceitável, é uma abordagem menos utilizada nesta categoria do que nas categorias C^2 e C^1 . Ao mesmo tempo, as estratégias menos adequadas de lidar com o conflito são usadas com mais frequência pelos grupos de C^3 em relação aos grupos de outras categorias.

C^1 e C^2 são as categorias mais eficazes em termos de estratégias de resolução de conflitos intra-grupo. Na verdade, os grupos dessas categorias usam com mais frequência a integração relativamente às outras categorias. Resultados de pesquisas anteriores têm mostrado uma associação positiva entre o uso da integração e os níveis de eficácia da equipa (Kuhn e Poole, 2000) e uma relação negativa entre o uso de domínio e os resultados da equipe (Friedman et al., 2000; Pneuman e Bruehl, 1982).

Por outro lado, C^2 e C^4 são as que menos recorrem ao domínio e acomodação, sendo C^4 o grupo que menos recorre ao evitamento como gestão dos conflitos. Com relação às abordagens não-confronto com conflito - evitamento e acomodação - apesar de poucos estudos terem sido realizados nesta área, há alguma evidência apoiando os seus efeitos negativos nos resultados da equipe (DeChurch e Marks, 2001).

Assim, considerar que C^1, C^2 e C^4 são incomparáveis e melhores que C^3 poderá ser uma boa solução para o problema em estudo.

Conclusões

Neste capítulo revemos as principais contribuições deste trabalho, fazemos uma síntese das conclusões mais significativas e apontamos algumas vias de investigação futura nesta área.

O objectivo central do trabalho foi o de contribuir com novas metodologias para o apoio à decisão em problemas de classificação de um conjunto \mathcal{A} de alternativas, com base em múltiplos critérios, a um conjunto de classes parcialmente ordenadas.

No cumprimento desse objectivo, desenvolvemos dois métodos heurísticos que fazem a junção de duas abordagens diferentes, classificação e clustering multicritério, para o cálculo de soluções aproximadas de partições parcialmente ordenadas. Desenvolvemos ainda modelos para a pesquisa de soluções eficientes recorrendo à programação inteira 0-1.

Formalmente, clustering com classes parcialmente ordenadas consiste em encontrar uma partição de \mathcal{A} onde a relação de ordem parcial é definida pelas classes: uma classe pode ser melhor ou pior que algumas classes, mas também pode ser incomparável a outras classes. Esta é uma inovação numa literatura onde os problemas nesta área ou são de clustering sem qualquer tipo de relação estabelecida entre as classes ou tal relação é de ordenação completa.

O facto de considerarmos não só o caso em que as classes têm uma estrutura de ordem total, mas também outras estruturas que permitem que as classes sejam incomparáveis levou-nos à utilização de métodos de prevalência. A modelação da relação de incomparabilidades é uma das principais características das técnicas da teoria da prevalência.

Os métodos desenvolvidos procuram satisfazer os seguintes requisitos, que julgamos importantes para apoiar eficazmente o decisor na análise de problemas de classificação multicritério:

- apresentar um protocolo facilmente entendido pelo decisor que exija juízos simples por parte deste;
- possibilitar o diálogo com o decisor por forma a que este possa participar no processo de

decisão. O processo de decisão termina quando o decisor considerar que encontrou uma solução de compromisso satisfatória.

- encontrar soluções boas num tempo computacional razoável. O cálculo de soluções eficientes aproximadas através de técnicas heurísticas talvez seja o único processo viável para determinados problemas de grandes dimensões.

Estruturamos o nosso trabalho com três ideias fundamentais. A primeira tem por base detectar grupos homogêneos, considerando os objectos definidos com medidas independentes das preferências do decisor, e só numa segunda fase, ter em consideração as preferências para estabelecer uma estrutura das classes obtidas. Por exemplo, agrupar e descrever pacientes de acordo com os seus sintomas, e apenas numa segunda fase, comparar e ordenar os grupos de acordo com a gravidade dos sintomas. A segunda ideia é a de explicitar as preferências do decisor desde o início do processo de decisão. Nestas duas primeiras ideias, a nossa atenção centrou-se no desenvolvimento de modelos heurísticos, considerando relevante o tempo computacional bem como os recursos exigidos. A terceira ideia surge no seguimento das duas primeiras ideias, tendo por objectivo encontrar soluções exactas e não soluções aproximadas, utilizando para tal programação matemática. Apresentamos para tal várias novas contribuições:

- apresentamos e discutimos princípios que devem ser seguidos na avaliação da qualidade de candidatos a partições de um conjunto de alternativas. A qualidade é definida à custa das relações binárias de prevalência, podendo o decisor participar no seu cálculo ao indicar qual(ais) a(s) sua(s) preferências mais importantes através de pesos.
- definimos duas formas de construção de *relações de preferência entre classes* com base em relações binárias de prevalência: uma entre pares de centróides representativos das classes e outra entre pares de classes.
- definimos relações de prevalência entre classes com base numa maioria fraca de pares de alternativas que são a favor da asserção “uma classe é pelo menos tão boa quanto uma outra classe”.
- definimos a noção *similaridade entre duas classes* com base na qualidade obtida com a fusão das classes.
- discutimos como *ajustar uma partição intransitiva* para a transformar em transitiva;

Começamos por apresentar e discutir diferentes princípios de atribuição de um conjunto de alternativas a classes parcialmente ordenadas. Tais princípios são baseados nas relações de prevalência das alternativas a serem atribuídas, as quais são representadas pelo conceito “pelo menos tão bom como” e usadas pelos métodos de prevalência. É apresentado um princípio de consistência forte que denominamos $S\tau$ -consistency, o qual, embora bastante interessante, nem sempre é possível cumprir integralmente já que a relação de prevalência S é intransitiva. Outro princípio é apresentado, este de consistência semi-forte que denominamos $SS\tau$ -consistency. Também ele é atraente e fácil de entender, mas mais modesto nas suas exigências: pretende-se que a relação de prevalência seja respeitada para cada par de alternativas em pelo menos uma das direcções (a_i, a_j) ou (a_j, a_i) e não nas duas direcções (a_i, a_j) e (a_j, a_i) como o princípio de consistência forte.

De seguida apresentamos um modelo heurístico denominado *Clustering-Sorting* que pretende ser uma abordagem simples e transparente para o decisor, cuja assistência é solicitada para ajudar a moldar os resultados. O objectivo, numa primeira fase, é detectar primeiro categorias homogéneas de alternativas e formar de seguida uma estrutura de ordem parcial com essas categorias detectadas. Assim, primeiro precisamos detectar apenas os clusters com base numa medida de similaridade independente das preferências do Decisor. Em seguida, pretende-se detectar potenciais relações de ordem parcial que possam existir entre eles, de acordo com os valores subjectivos do decisor. A eliciação de tais valores é feita somente depois de o decisor identificar os agrupamentos e considerar que essas categorias fazem sentido.

De seguida a nossa atenção centrou-se no desenvolvimento de dois métodos mais sofisticados de optimização, um utilizando heurísticas e outro programação matemática, para coincidir com as preferências de um decisor.

Começamos por desenvolver a heurística CHOP - **C**lustering **H**ierárquico com **O**rdenação **P**arcial - baseada no método hierárquico aglomerativo mas numa perspectiva multicritério. Sugerimos duas formas de construção da estrutura de preferências das classes, ambas definidas à custa de relações binárias de prevalência, uma entre pares de centróides representativos das classes e outra entre pares de classes $(C^i, C^j) \in \mathcal{P}^2$. No primeiro caso, uma classe é considerada melhor que outra caso o seu centróide prevaleça sobre o centróide da outra classe mas não seja prevalecido, as classes são incomparáveis se nenhum dos centróides prevalecer sobre o outro e por fim são consideradas indiferentes caso os centróides prevaleçam um sobre o outro

(Estrutura 1). No segundo caso, considerou-se que uma classe C^i prevalece sobre C^j se pelo menos 50% das relações dos pares de alternativas $(a_s, a_t) \in (C^i, C^j)$ seja de prevalência de a_s sobre a_t . A relação de preferência entre as classes é então definida de forma idêntica à dos centróides, de acordo com as relações de prevalência entre as classes (Estrutura 2).

Nas experiências que realizamos, os resultados obtidos com a Estrutura 2 mostraram-se no geral melhores que os obtidos com a Estrutura 1 uma vez que levam à obtenção de muito menos intransitividades e com uma qualidade idêntica ou mesmo melhor.

O objectivo do método CHOP é determinar uma partição parcialmente ordenada e compatível com as preferências do decisor, mas que seja também transitiva. Contudo, a relação que definimos entre as classes é de prevalência, relação essa que é intransitiva. Assim, em cada iteração do Algoritmo proposto, pode existir pelo menos um terno de classes que não verificam a transitividade. A nossa proposta nesses casos é a de ajustar as partições intransitivas, à custa da eliminação ou introdução de arco(s) na partição. Cada partição transitiva resultante é então re-avaliada quanto à sua qualidade, escolhendo-se a que leva à melhor qualidade.

Um método que usa técnicas heurísticas pode ser útil como abordagem isolada ou como um módulo de pesquisa inicial que antecede um método exacto, como é o caso dos dois primeiros métodos que desenvolvemos. Nesta segunda situação, a intenção do módulo heurístico é dar ao decisor um conhecimento inicial do problema que o ajude a estabelecer e rever as suas preferências, a um baixo custo computacional. O decisor avançará depois para re-atribuição do conjunto de alternativas a um conjunto de classes cuja estrutura é a sugerida pela heurística.

Assim, julgámos importante poder dispor de uma ferramenta que, dado o número de classes, nos fornecesse a partição óptima. A nossa atenção centrou-se pois no desenvolvimento de modelos exactos para problemas de programação inteira 0-1, para estruturas não só conhecidas *a priori* como também desconhecidas. As estruturas conhecidas *a priori* podem ser sugeridas pelo decisor dado o seu conhecimento sobre o problema em estudo ou, caso não tenha qualquer conhecimento prévio sobre o número de classes a escolher nem das suas relações de preferência, pode ser aplicada previamente uma das heurísticas apresentadas, Clustering-Sorting ou CHOP.

As dificuldades inerentes à complexidade da programação inteira, agravadas pela existência de múltiplos critérios, são ainda agravadas, no nosso caso, pelo desconhecimento da estrutura de preferências dessas classes. As experiências computacionais que efectuamos mostraram-nos

que a aplicação da programação inteira pode tornar os problemas de maiores dimensões intratáveis, quer pelo tempo quer pelos recursos que exigem. Contudo, partindo do conhecimento da estrutura das classes, a complexidade computacional do problema diminui significativamente tornando-se assim praticável a sua utilização pelo menos em problemas que não sejam demasiado grandes.

Em suma, nas experiências que fizemos com o Clustering-Sorting e com o CHOP, usando o sistema computacional desenvolvido, notámos que o cálculo de cada solução era, na maior parte das vezes, praticamente instantâneo. Na verdade, tal como já foi referido, este é um factor importante para o desenrolar do processo interactivo de decisão. Por outro lado, com a aplicação dos modelos exactos, tal não se verificou, tornando-se mesmo bastante difícil obter resultados para o caso da estrutura desconhecida *a priori*, com a agravante de os resultados obtidos não serem significativamente melhores que os obtidos com o Clustering-Sorting ou com o CHOP. Assim, embora o método heurístico não garantam a optimalidade, pelos exemplos que testámos, julgamos que é possível obter um compromisso aceitável entre a qualidade das soluções e o tempo computacional mas, como em qualquer outra abordagem heurística, só experiências mais exaustivas poderiam confirmar ou desmentir esta ideia.

Para ilustrar os vários métodos de classificação multicritério descritos, usamos problemas reais descritos na literatura. Os dois métodos heurísticos foram ainda aplicados a dados reais de um conjunto de 77 grupos de trabalho avaliados pelo tipo de conflitos intra-grupais e gestão desses conflitos. A motivação que nos levou a testar o método com este último caso real foi, por um lado, a convicção de que os métodos por nós desenvolvidos seriam indicados para o problema em causa, e por outro lado, pela existência de poucos estudos que aplicam a análise de clusters para os processos grupais. Os resultados obtidos foram ao encontro do que a literatura relata sobre gestão de conflitos em grupos.

Os novos métodos sugeridos para problemas de classificação multicritério com classes parcialmente ordenadas, basearam-se nos métodos de prevalência ELECTRE. A investigação futura poderá estudar as ideias apresentadas mas baseadas noutros métodos de prevalência (como por exemplo os métodos PROMETHEE) o que poderá suscitar novas ideias e novos conceitos.

A investigação futura poderá estudar novos métodos também na heurística CHOP que de-

envolvemos. As ideias que apresentamos basearam-se no método hierárquico aglomerativo, sendo contudo interessante explorar as ideias apresentadas mas baseadas noutros métodos clustering, nomeadamente o método hierárquico divisivo.

Além disso, as abordagens heurísticas estudadas deveriam ser testadas com mais problemas e problemas de maiores dimensões, o que permitiria uma aferição mais acurada, aspecto importante para a sua utilização futura no estudo de casos.

O estudo de modelos exactos para problemas de programação inteira 0-1 não foi muito aprofundado. No contexto deste método abrem-se ainda outras perspectivas de desenvolvimento futuro. Uma delas consiste em desenvolver uma abordagem baseada em meta-heurísticas que constitua uma extensão de algoritmos genéricos como o *tabu search* e *simulated annealing* num contexto multicritério.

Uma outra via de investigação consiste na aplicação das metodologias propostas no estudo de casos, sendo para tal indispensável a intervenção de peritos da área a que o problema diga respeito. Associada a esta via de investigação está a identificação de situações onde cada um dos métodos poderá ser o mais indicado já que esta não é de todo uma tarefa fácil e lógica, pelo menos em determinadas situações.

Por fim, e no âmbito do desenvolvimento de software, seria interessante construir de raiz software que integrasse as metodologias propostas aos estudos de caso, facilitando e apoiando o utilizador, através de um interface de fácil utilização e específico.

Referências Bibliográficas

- [1] Ackoff R.L. (1978), *The Art of Problem Solving*, A. Wiley.
- [2] Agrawal R. , Gehrke J., Gunopulos D., e Raghavan P. (1998), Automatic Subspace Clustering of High Dimensional Data for Data Mining Applications, *Proceedings of ACM SIGMOD Conference on Management of Data*. 94-105, Seattle, WA.
- [3] Altman E.I., Avery R., Eisenbeis R., Stinkey J. (1981). *Application of Classification Techniques in Business, Banking and Finance*. Contemporary Studies in Economic and Financial Analysis, Vol. 3. JAI Press, Greenwich, CT.
- [4] Amason A.C. (1996), Distinguishing the effects of functional and dysfunctional conflict on strategic decision making: resolving a paradox for top management teams. *Academy of Management Journal*, 39, 123-148.
- [5] Amason A.C., e Sapienza H.J. (1997), The effects of top management team size and interaction norms on cognitive and affective conflict. *Journal of Management*, 23, 495 - 516.
- [6] Anderberg M. (1973), *Cluster Analysis for Applications*, Academic Press, Inc., New York, NY.
- [7] Antunes C.H., Tavares L.V. (2000), *Casos de aplicação de Investigação Operacional*. (Coordenadores). McGraw-Hill Portugal.
- [8] Antunes C.H., Dias L.C (2007), *Decisão Perspectivas interdisciplinares*. (Coordenadores). Imprensa da Universidade de Coimbra.
- [9] Ball G. e Hall D. (1965), *ISODATA: a novel method of data analysis and classification*, Tech. Rep.. Stanford University, Stanford, CA.
- [10] Bana e Costa C.A. (1990), The outranking approach and the foundations of ELECTRE methods, in *Readings in Multiple Criteria Decision Aid*, Ed.Berlin: Springer Verlag, 155-183.

- [11] Bana e Costa C.A. (1992), Structuration, Construction et Exploitation d'un Modèle Multicritère d'Aide à la Décision. Lisboa, Tese de Doutoramento. Universidade Técnica de Lisboa.
- [12] Bana e Costa C.A. (1993), Les Problématiques dans le Cadre de l'activité d'aide à la Décision", LAMSADE. Université de Paris-Dauphine 80.
- [13] Bana e Costa, C. A., Ferreira, A. e Corrêa, É. (2000), Metodologia Multicritério de Apoio à Avaliação de Propostas em Concurso Públicos. Casos de Aplicação da Investigação Operacional, McGrawHill: 337-363.
- [14] Bana e Costa e Vansnick (1997), Applications of the MACBETH Approach in the Framework of an Additive Aggregation Model, Journal of Multi-Criteria Decision Analysis 6(2): 107-114.
- [15] Bana e Costa C.A., e Vansnick J.C. (1994), MACBETH - An interactive path towards the construction of cardinal value functions, International Transactions in Operational Research 1(4): 489-500.
- [16] Bana e Costa C.A., e Vansnick J.C. (1995), General overview of the MACBETH approach, in Advances in Multicriteria Analysis, P.M. Pardalos, Y. Siskos, C. Zopounidis (eds.), Kluwer Academic Publishers, Book Series: Nonconvex Optimization and its Applications, 93-100.
- [17] Bana e Costa C.A., e Vansnick J.C. (1997), A theoretical framework for Measuring Attractiveness by a Categorical Based Evaluation Technique (MACBETH), in Multicriteria Analysis, J. Clímaco (ed.), Springer-Verlag, 15-24.
- [18] Bana e Costa C.A., e Vansnick J.C. (1999), The Macbeth Approach: Basic Ideas, Software, And an Application. Advances in Decision Analysis. N. M. e. M. Roubens, Kluwer Academic Publishers: 131-157.
- [19] Bana e Costa C.A. e Vansnick J.C. (2000), Cardinal value measurement with macbeth, in Decision Making: Recent Developments and Worldwide Applications, S.H. Zanakis, G. Doukidis, C. Zopounidis (eds.), Kluwer Academic Publishers, Book Series: Applied Optimization, 45: 317-329.

- [20] Belton V., e Ackermann F. (1997), Integrated Support from Problem Structuring through to Alternative Evaluation Using COPE and V.I.S.A., *Journal of Multi-Criteria Decision Analysis* 6: 115-130.
- [21] Belacel N. (2000). Multicriteria assignment method PROAFTN: methodology and medical application. *European Journal of Operational Research*, 125(1):175-183.
- [22] Belton V., Stewart T. (2002). *Multiple Criteria Decision Analysis: An Integrated Approach*. Dordrecht: Kluwer Academic Publishers.
- [23] Ben-Dor A., e Yakhini Z. (1999). Clustering gene expression patterns. *Proceedings of the 2nd SIAM International Conference on Data Mining, (ICDM'99)*, Arlington, VA. USA., 420-436.
- [24] Berkhin P. (2002), *Survey of Clustering Data Mining Techniques*, Accrue Software.
- [25] Bouyssou D., Marchant T., Pirlot M., Perny P., Tsoukiàs A., Vincke P.(2000). *Evaluation and decision models: A critical perspective*. Dordrecht: Kluwer Academic Publishers.
- [26] Bouyssou D., e Pirlot M. (2002). Non Transitive Decomposable Conjoint Measurement. *Journal of Mathematical Psychology*, 46:677-703.
- [27] Bouyssou D., e Pirlot M. (2004). Additive difference models without additivity and subtractivity. *Journal of Mathematical Psychology*, 48(3):263-291.
- [28] Bouyssou D., e Pirlot M. (2004a). Preferences for multiattributed alternatives: Traces, Dominance, and Numerical Representations. *Journal of Mathematical Psychology*, 167:427-443.
- [29] Bouyssou D., e Pirlot M. (2005). A characterization of concordance relations, *European Journal of Operational Research*, 167(2):427-443.
- [30] Bouyssou D., e Pirlot M. (2005a). Following the traces: An introduction to conjoint measurement without transitivity and additivity. *European Journal of Operational Research*, 163(2):287-337.
- [31] Bouyssou D., e Vincke Ph. (2006). Relations binaires et modélisation des préférences. In D. Bouyssou, D. Dubois, M. Pirlot, and H. Prade, editors, *Concepts et méthodes pour l'aide à la décision*, vol. 1 : outils de modélisation, pages 71-109. Hermès, Paris.

- [32] Brans J.P. (1994), The PROMCALC & GAIA decision support system for multicriteria decision aid, *Decision Support Systems* 12: 297-310.
- [33] Brans J.P, Vincke Ph. (1985). A preference ranking organization method, *Management Science* 31(6):647-656.
- [34] Catelani M., Fort A. (2000). Fault diagnosis of electronic analog circuits using a radial basis function network classifier. *Measurement* 28 (3), 147-158.
- [35] Checkland P., e Scholes J. (1990), *Soft Systms Methodology in Action*, John Wiley & Sons.
- [36] Checkland P., e Holwell S. (1993), Information management and organizacional processes: an approach through soft systems methodology, *Journal of Information Systems* 3(1): 3-16.
- [37] Cheeseman P., e Stutz J. (1996), Bayesian classification (AutoClass): Theory and Results, *Advances in Knowledge Discovery and Data Mining*, U. M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy, Eds., AAAI Press/MIT Press.
- [38] Chen S.J., Hwang C.L. (1992), *Fuzzy multiple attribute decision making: methods and applications*, Springer, Heidelberg.
- [39] Cooley R., Mobasher B., e Srivastava J. (1999). Data Preparation for Mining World Wide Web Browsing Pattern, *Knowledge and Information Systems*, 1(1):5-32.
- [40] Cooley R., Mobasher B., e Srivastava J. (1997). Web Mining: Information and Pattern Discovery on the World Wide Web. In: *Proceedings of the 9th IEEE International Conference on Tools with Artificial Intelligence*, Newport Beach, CA.
- [41] Daellenbach H.G. (1994), *Systems and Decision Making: A Management Science Approach*, J.Wiley, New York.
- [42] DeChurch L.A., Marks M.A. (2001), Maximizing the benefits of task conflict: the role of conflict management. *Journal of Conflict Management*, 12(1), 4-22.
- [43] De Dreu C.K.W., Dierendonckn D., e Dijkstra M.T.M. (2004). Conflict at work and individual well-being. *International Journal of Conflict Management*, 15, 6-26.

- [44] De Dreu, C. K. W. e Weingert, L. R. (2003). A contingency theory of task conflict and performance in groups and organizational teams. In M. A. West, D. Tjosvold e K. G. Smith (Eds.), *International Handbook of Organizational Teamwork and Cooperative Working* (pp. 150-166). Chichester: John Wiley & Sons, Ltd.
- [45] Dempster A., Laird N., e Rubin D. (1977), Maximum likelihood from incomplete data via the EM algorithm, *Journal of the Royal Statistical Society, Series B*, 39(1), 1-38.
- [46] De Smet Y., Montano G.L. (2004). Towards multicriteria clustering: an extension of the k-means algorithm. *European Journal of Operational Research* 158(2) : 390-398.
- [47] De Smet Y., Eppe S. (2009) Multicriteria Relational Clustering: The case of binary outranking matrices. *Evolutionary multi-criterion optimization Lecture Notes in Computer Science*, volume 5467/2009, 380-392, DOI: 10.1007/ 978-3-642-01020-0-31.
- [48] Deutsch M. (1973), *The resolution of conflict*. New Haven, CT: Yale University Press.
- [49] Diakoulaki D., Zopounidis C., Mavrotas G., Doumpos M. (1999). The use of a preference disaggregation method in energy analysis and policy making. *Energy* 24 (2), 157-166.
- [50] Dias L.C. (2000), *A Informação Imprecisa e os Modelos Multicritério de Apoio à Decisão - identificação e uso de conclusões robustas*, Dissertação de Doutoramento em Organização e Gestão de Empresas, Especialidade de Ciências dos Sistemas nas Organizações, Faculdade de Economia, Universidade de Coimbra.
- [51] Dias, L.C. (2002), *Apontamentos de análise da decisão: como considerar múltiplos critérios*, Secção de Textos da Faculdade de Economia da Universidade de Coimbra, Coimbra.
- [52] Dias L.C., Clímaco J.N. (2000), Additive aggregation with variable interdependent parameters: the VIP Analysis software, *Journal of the Operational Research Society* 51, 1070-1082.
- [53] Dias L.C., Lamboray C. (2010). Extensions of the prudence principle to exploit a valued outranking relation. *European Journal of Operational Research* 201, 828-837.
- [54] Dias L., Mousseau V., Figueira J., Clímaco J. (2002), An aggregation/ disaggregation approach to obtain robust conclusions with ELECTRE TRI, *European Journal of Operational Research*, 138:332-348.

- [55] Dias L., Mousseau V. (2003), IRIS - A DSS for Multiple Criteria Sorting Problems, *Journal of Multi-Criteria Decision Analysis*, 12: 285-298.
- [56] Diday E. e Simon J. (1976), Clustering analysis, In *Digital Pattern Recognition*, K. S. Fu, Ed. Springer-Verlag, Secaucus, NJ, 47-94.
- [57] Dimas I.D (2007), (Re)pensar o conflito intragrupal : níveis de desenvolvimento e eficácia, Faculdade de Psicologia e Ciências da Educação da Universidade de Coimbra, Dissertação de Doutoramento.
- [58] Dimas I.D., Lourenço P.R., Miguez J. (2008). Conflitos, gestão de conflitos e desenvolvimento grupal: contributo para uma articulação. *Psychologica*, 47, 9-28.
- [59] Doreian P., Batagelj V., Ferligoj A. (2005). *Generalized Blockmodeling*. University of Cambridge, New York.
- [60] Doumpos M., Zopounidis C. (1998), The use of the preference disaggregation analysis in the assessment of financial risks. *Fuzzy Economic Review* 3(1), 39-57.
- [61] Doumpos M., Zopounidis C.(2002), *Multicriteria Decision Aid Classification Methods*, Kluwer Academic Publishers, Dordrecht.
- [62] Doumpos M., Zopounidis C. (2004), A multicriteria classification approach based on pairwise comparisons, *European Journal of Operational Research*, 158:378-389.
- [63] Duarte F.J (2008), *Optimização da Combinação de Agrupamentos baseado na Acumulação de Provas pesadas por Índices de Validação e com uso de Amostragem*, Departamento de Engenharias da Universidade de Trás-os-Montes e Alto Douro, Dissertação de Doutoramento.
- [64] Duda R.O., Hart P.E., e Stork D.G. (2001). *Pattern Classification*. Wiley-Interscience, 2nd edition.
- [65] Dutka A. (1995). *AMA Handbook of Customer Satisfaction: A Guide to Research, Planning and Implementation*. NTC Publishing Group, Illinois.
- [66] Edwards W., e Barron FH (1994). SMARTS and SMARTER: Improved simple methods for multiattribute utility measurement. *Organizational Behavior & Human Decision Processes* 60, 306-325.

- [67] Ester M., Kriegel H., Sander J., e Xu X. (1996), A Density-based Algorithm for Discovering Clusters in Large Spatial Databases with Noise, 2nd Int. Conf. on Knowledge Discovery and Data Mining (KDD 96), Portland, USA.
- [68] Everitt B., Landau S., e Leese M. (2001), Cluster analysis, 4nd edition. Arnold Publication.
- [69] Ferligoj, A. (1997). Recent developments in cluster analysis,ISSN 1330-1012, 19th International Conference on INFORMATION TECHNOLOGY INTERFACES ITI '97, pp 253.
- [70] Fernandez E., Navarro J., Bernal S. (2010). Handling multicriteria preferences in cluster analysis. *European Journal of Operational Research* 202 : 819-827.
- [71] Figueira J., Mousseau V., Roy B. (2005). ELECTRE Methods, In *Multiple Criteria Decision Analysis: State of the Art Surveys*, J. Figueira, S. Greco, M. Ehrgott (Editors), Int. Series in Operations Research and Management Science, Vol. 78, 133-172, Springer, New York.
- [72] Figueira J., De Smet Y., Brans JP (2004). Promethee for MCDA Classification and Sorting problems: Promethee TRI and Promethee CLUSTER, Université Libre de Bruxelles. Service de Mathématiques de la Gestion, Working Paper 2004/02. <<http://www.ulb.ac.be/polytech/smg/indexpublications.htm>>.
- [73] Fishburn P.C. (1970), *Utility Theory for Decision Making*. New York, Wiley.
- [74] Fisher D. (1987), Knowledge acquisition via incremental conceptual clustering, *Machine Learning*, 2, 139-172.
- [75] Flinkman M., Michalowski W., Nilsson S., Slowinski R., Susmaga R., Wilk S. (2000). Use of rough sets analysis to classify Siberian forest ecosystem according to net primary production of phytomass. *INFOR* 38 (3), 145-161.
- [76] Foss A., Wang W., e Zanne O., (2001). A non parametric approach to web log analysis. *Proceedings of the 1st SIAM ICDM, Workshop on Web Mining, (WM'01)*, Chicago, Illinois, 51-58.

- [77] Friedman R.A., Tidd S.T., Currall S.C., Tsai J.C. (2000). What goes around comes around: the impact of personal conflict style on work conflict and stress. *International Journal of Conflict Management*, 11, 32-55.
- [78] Gardiner L.R., Vanderpooten D. (1997). Interactive Multiple Criteria Procedures: Some Reflections, 11th International Conference on Multiple Criteria Decision Making, Coimbra, Portugal, Springer, 290-301.
- [79] Gennari J., Langley P., e Fisher D. (1989), Models of incremental concept formation, *Journal of Artificial Intelligence*, 40, 11-61.
- [80] Gochet W., Stam A., Srinivasan V., Chen S. (1997). Multigroup discriminant analysis using linear programming. *Operations Research* 45 (2), 213-225.
- [81] Gomes L., Gomes C. e Almeida A. (2002), *Tomada de Decisão Gerencial: Enfoque Multicritério*. São Paulo, Atlas.
- [82] Gowda K. e Diday E. (1992), Symbolic clustering using a new dissimilarity measure, *IEEE transactions on Systems, Man and Cybernetics*, 22(2), 368-378.
- [83] Gowda K., e Krishna G. (1977), Agglomerative clustering using the concept of mutual nearest neighbourhood, *Pattern Recognition* 10, 105-112.
- [84] Gower J. (1985), Measures of similarity, dissimilarity and distance, *Encyclopaedia of Statistical Sciences*, Volume 5 (S. Kotz, N.L. Johnson and C.B. Read, eds). Wiley, New York.
- [85] Gower J., e Legendre P. (1986), Metric and Euclidean properties of dissimilarity coefficients, *Journal of classification*, 5, 5-48.
- [86] Grabisch M. (1996), The application of fuzzy integrals in multicriteria decision making, *European Journal of Operational Research* 89, 445-456.
- [87] Greco S., Matarazzo B., Slowinski R. (1998). A new rough set approach to evaluation of bankruptcy risk. In: Zopounidis, C. (Ed.), *Operational Tools in the Management of Financial Risks*. Kluwer Academic Publishers, Dordrecht, pp. 121-136.
- [88] Guha S., Rastogi R., e Shim K. (1998), CURE: an efficient clustering algorithm for large databases, *ACM SIGMOD International Conference on Management of Data*, 73-84.

- [89] Guimarães R.C. (1979), Metodologia da Investigação Operacional. GEIN - Gabinete de Gestão e Engenharia Industrial, FEUP - Universidade do Porto.
- [90] Guitouni A., e Martel J.M. (1998), Tentative guidelines to help choosing an appropriate MCDA method, *European Journal of Operational Research* 109(2): 501-521.
- [91] Han J., Kamber M. , Tung A. (2001). Spatial Clustering Methods in Data Mining: a review. In H. J. Miller & J. Han (Eds.), *Geographic data mining and knowledge discovery* (pp. 33-50). London and New York: Taylor and Francis.
- [92] Han J., e Kamber M. (2006), *Data Mining, Concepts and Techniques*, Morgan Kaufmann Publishers, 2nd Edition.
- [93] Heer I., e Chi E. (2001). Identification of web user traffic composition using multi-modal clustering and information scent. *Proceedings of the 1st SIAM ICDM Workshop on Web Mining*, Chicago, IL., 51-58.
- [94] Hertz J., Krogh A., e Palmer R. (1991), *Introduction to the Theory of Neural Computation*, Santa Fe Institute Studies in the Sciences of Complexity lecture notes. Addison-Wesley Longman Publ. Co., Inc., Reading, MA.
- [95] Hinneburg A. e Keim D. (1998), An Efficient Approach to Clustering in Large Multimedia Databases with Noise, *Proc. 4th Int. Conf. on Knowledge Discovery and Data Mining*.
- [96] Huang Z. (1998), Extensions to the k-means algorithm for clustering large data sets with categorical values, *Data Mining and Knowledge discovery*, 2:283-304.
- [97] Hurson Ch., Zopounidis C. (1997), On the use of multicriteria decision aid methods to portfolio selection, in Clímaco, J.(ed), *Multi-Criteria Analysis*, Springer, 496-507.
- [98] Iryna Yevseyeva , Miettinen K., Räsänen P. (2008), Verbal ordinal classification with multicriteria decision aiding, *European Journal of Operational Research*, 185, 964-983.
- [99] Jacquet-Lagrèze, E. (1995), An application of The UTA discriminant model for the evaluation of R&D projects. In: Pardalos, M.P., Siskos, Y., Zopounidis, C.(Eds.), *Advances in Multicriteria Analysis*, Kluwer Academic Publishers, Dordrecht, pp. 203-211.
- [100] Jain A.K., Dubes R.C. (1988). *Algorithms for Clustering Data*. Prentice-Hall advanced reference series. Prentice-Hall,Inc., Upper Saddle River, NJ.

- [101] Jain A. e Flynn P. (1996), Image segmentation using clustering, *Advances in Image Understanding: A Festschrift for Azriel Rosenfeld, N. Ahuja and K. Bowyer, Eds, IEEE Press, Piscataway, NJ, 65-83.*
- [102] Jain A. , Duin R., e Mao J. (2000), Statistical pattern recognition: A review, *IEEE Trans. Pattern Analysis and Machine Intelligence, 22:4-37.*
- [103] Jain A.K., Murty M.N., Flynn P.J. (1999). Data Clustering : A review. *ACM Computing Surveys, 31(3).*
- [104] Jarvis R., e Patrick E. (1973), Clustering using a similarity method based on hared near neighbors, *IEEE Transactions on Computers. C-22, 8, 1025-1034.*
- [105] Jehn K.A. (1995), A multimethod examination of the benefits and detriments of intra-group conflict, *Administrative Science Quarterly, 40, 256-282.*
- [106] Jonhson R.A., e Wichern D.W. (2002). *Applied multivariate statistical analysis. Prentice Hall, New Jersey, 5th edition.*
- [107] Karypis G., Han E.H, e Kumar V. (1999), Chameleon: Hierarchical clustering using dynamic modelling, *IEEE Computer, 32 (8), 68-75.*
- [108] Kaufman L., Rousseeuw P.J. (1987), Clustering by means of medoids, in *Statistical Data Analysis based on the L1 Norm, Y. Dodge, Elsevier/North- Holland, Amsterdam, 405-416.*
- [109] Kaufman L., Rousseeuw P.J. (1990), *Finding Groups in Data: An Introduction to Cluster Analysis. New York: Wiley.*
- [110] Keeney R., e Raiffa H. (1976), *Decisions with Multiple Objectives: Preferences and Value Tradeoffs. John Wiley and Sons, New York .*
- [111] Keeney R.L. (1992), *Value-Focused Thinking: A Path to Criative Decision-Making, Harvard University Press.*
- [112] King B. (1967), Step-wise clustering procedures, *Journal of the American Statistical Association 69, 86-101.*
- [113] Kohonen T. (1990), The Self-Organizing Map, *Proceedings of the IEEE, 78(9): 1464-1480.*

- [114] Kohonen T. (2001), *Self-Organising Maps*, 3rd ed., Springer-Verlag, Springer series in Information Sciences.
- [115] Köksalan M., Ulu C. (2003), An interactive approach for placing alternatives in preference classes. *European Journal of Operational Research* 144, 429-439.
- [116] Kuhn T., Poole M.S (2000). Do conflict management styles affect group decision making? Evidence from a longitudinal study. *Human Communication Research*, 26 (4), 558-590.
- [117] Lai Y.J., Liu T.Y. e Hwang C.L. (1994), TOPSIS for MODM, *European Journal of Operational Research* 76: 486-500.
- [118] Lance G., e Williams W. (1967), A general theory of classification sorting strategies, *Computer Journal*, 9, 373-386.
- [119] Larichev O., e Moshkovich H. (1994). An approach to ordinal classification problems. *International Transactions in Operational Research*, 1(3):375-385.
- [120] Larichev O., e Brown R. (2000). Numerical and verbal decision analysis: Comparison on practical cases. *Journal of Multi-Criteria Decision Analysis*, 9:263-273.
- [121] Larichev O., Kortnev A., e Kochin D.(2002). Decision support system for classification of a finite set of multicriteria alternatives. *Decision Support Systems*, 33(1):13-21.
- [122] Léger J., e Martel J.M. (2002). A multicriteria assignment procedure for a nominal sorting problematic. *European Journal of Operational Research*, 138(2):349-364.
- [123] Lu Y., e Han J. (2003), Cancer Classification using Gene Expression Data, *Information Systems*, 243-268.
- [124] Luce R.D (1956), Semiorders and theory of utility discrimination, *Econometrica*, 24:178-191.
- [125] Luenberger D. (1984), *Introduction to Linear and Nonlinear Programming*, Addison-Wesley.
- [126] Mahalanobis P. (1936), On the generalized distance in statistics, *Proceedings of the National Institute of Sciences of India, New Delhi*, v. 2, 4955.

- [127] MATLAB. The MathWorks, Inc., Natick, MA 01760, <<http://www.mathworks.com/products/matlab/>> (visto a 2009-09-14).
- [128] Massaglia R., Ostanello A. (1991). N-tomic: a support system for multicriteria segmentation problems. In: Korhonen, P., Lewandowski, A., Wallenius, J. (eds), *Multiple Criteria Decision Support*, pp. 167-174. Springer Verlag, LNEMS 356.
- [129] Michalowski W., Rubin S., Slowinski R., Wilk S. (2001). Triage of the child with abdominal pain: A clinical algorithm for emergency patient management. *Paediatrics and Child Health* 6(1), 23-28.
- [130] Michalski R., Stepp R., e Diday E. (1983), Automated construction of classifications: conceptual clustering versus numerical taxonomy, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-5, 4, 396-410.
- [131] Montano Guzman L. (2004), *Fuzzy Measures and Integrals in the MCDA Sorting Problematic. Methodology and Application to the Diagnosis of firms*. Thèse de doctorat, Université Libre de Bruxelles.
- [132] Moscarola J., Roy B. (1977). Procedure automatique d'examen de dossiers fondée sur une segmentation trichotomique en présence de critères multiples. *RAIRO Oper. Res.* 11(2), 145-173
- [133] Mousseau V., Dias L. (2004), Valued outranking relations in ELECTRE providing manageable disaggregation procedures. *European Journal of Operational Research*, 156(2), 467-482.
- [134] Mousseau V., Slowinski R. (1998), Inferring an ELECTRE-TRI model from assignment examples. *Journal of Global Optimization* 12(2), 157-174.
- [135] Mousseau V., Slowinski R., e Zielniewicz P. (1999), *ELECTRE TRI 2.0a : Methodological guide and user's documentation*, Université de Paris Dauphine, Document du LAMSADE 111.
- [136] Mousseau V., Slowinski R., e Zielniewicz P. (2000), A user-oriented implementation of the ELECTRETRI method integrating preference elicitation support. *Computers & Operations Research*, 27: 757-777.

- [137] Munduate L., Ganaza J., Peiró J.M., Euwema M. (1999). Patterns of styles in conflict management and effectiveness. *International Journal of Conflict Management*, 10(1), 5-20.
- [138] Murtagh F. (1984), A survey of recent advances in hierarchical clustering algorithms which use cluster centers, *Computer Journal* 26, 354-359.
- [139] Nemery P., De Smet Y. (2005). *Multicriteria Ordered Clustering*. Technical Report TR/SMG/2005-003, SMG, Université Libre de Bruxelles.
- [140] Ngo The A., Mousseau V. (2002), Using assignment examples to infer category limits for the ELECTRE TRI method, *Journal of Multi Criteria Decision Analysis*, 11(1): 29-43.
- [141] Nieddu L., Patrizi G. (2000). Formal methods in pattern recognition: A review. *European Journal of Operational Research* 120, 459-495.
- [142] Nowicki R., Slowinski R., Stefanowski J. (1992). Evaluation of vibroacoustic diagnostic symptoms by means of the rough sets theory. *Journal of Computers in Industry* 20, 141-152.
- [143] Oehler K., e Gray R. (1995), Combining image compression and classification using vector quantization, *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 17: 461-473.
- [144] Pawlak Z. (1982), Rough sets. *International Journal of Information and Computer Sciences* 11, 341-356.
- [145] Pawlak Z., Slowinski R. (1994), Rough set approach to multi-attribute decision analysis. *European Journal of Operational Research* 72, 443-459.
- [146] Perny P. (1998), Multicriteria filtering methods based on concordance and non-discordance principles. *Annals of Operations Research* 80, 137-167.
- [147] Pneuman R.W., Bruehl M.E. (1982). *Managing Conflict*, Prentice-Hall, Englewood Cliffs, NJ.
- [148] Pruitt D.G., Carnevale P.J. (1993), *Negotiation in social conflict*, Pacific Groves, CA: Brooks/Cole.

- [149] Pruitt D.G, Rubin J.Z. (1986), *Social conflict: escalation, stalemate and settlement*. New York: Random House.
- [150] Rahim M.A. (1983). A measure of styles of handling interpersonal conflict. *Academy of Management Journal*, 26, 368-376.
- [151] Rahim M.A., Bonoma T.V. (1979), *Managing organizational conflict: a model for diagnosis and intervention*. *Psychological Reports*, 44 (3), 1323-1344.
- [152] Raymond T., e Han J. (1994), *Efficient and Effective Clustering Methods for Spatial Data Mining*, *Proceedings of 20th International Conference on Very Large Databases*, Morgan Kaufmann, 144-155.
- [153] Raymond T., e Han J. (2002), *Clarans: a method for clustering objects for spatial data mining*, *IEEE Transactions on Knowledge and Data Engineering*, 14(5): 1003-1016.
- [154] Rentsch J.R., e Zelno J.A. (2003). *The role of cognition in managing conflict to maximize team effectiveness*. In M. A. West, D. Tjosvold e K. G. Smith (Eds.), *International handbook of organizational teamwork and cooperative working* (pp. 130-149). Chichester: John Wiley and Sons, Ltd.
- [155] Ripley B.D. (1996). *Pattern Recognition and Neural Networks*. Cambridge University Press, Cambridge.
- [156] Rocha C., Dias L. (2008), *An Algorithm for Ordinal Sorting Based on ELECTRE with Categories Defined by Examples*. *Journal of Global Optimization*, Vol. 42, No. 2, 255-277.
- [157] Rocha C., Dias L. (2010), *Multicriteria classification of workgroups with unknown number of categories: a clustering-sorting approach*. *Proceedings of the 25th Mini-EURO Conference on Uncertainty and Robustness in Planning and Decision Making (URPDM2010)*, INESC Coimbra - suporte CD-ROM, ISBN: 978-989-95055-3-7.
- [158] Romesburg H.C. (1984). *Cluster Analysis for Researchers*. 1st Edn., Life Time Learning Publications, Belmont, California, 168-195.
- [159] Rossi L., Slowinski R., Susmaga R. (1999). *Rough set approach to evaluation of stormwater pollution*. *International Journal of Environment and Pollution* 12 (2-3), 232-250.

- [160] Roubens M. (1997), Fuzzy sets and decision analysis, *Fuzzy Sets and Systems* 90, 199-206.
- [161] Roy B. (1978), ELECTRE III: Un algorithme de classements fondé sur une représentation floue des préférences de critères multiples. *Cahiers du CERO*, 20(1):3-24.
- [162] Roy B. (1981), A multicriteria analysis for trichotomic segmentation problems, in *Multiple Criteria Analysis: Operational Methods*, P. Nijkamp and J. Spronk, Eds. Farnborough: Gower Press, pp. 245-257.
- [163] Roy B. (1985), *Méthodologie Multicritère D'Aide à la Décision*, Economica.
- [164] Roy B. (1989), The outranking approach and the foundations of ELECTRE methods, *LAMSADE - Laboratoire d'Analyse et Modélisation de Systemes pour l'Aide à la Decision-Université de Paris-Dauphine*(53): 40.
- [165] Roy B. (1991), The outranking approach and the foundations of ELECTRE methods. *Theory and Decision* 31, 49-73.
- [166] Roy B. (1993). Decision science or decision-aid science? *European Journal of Operational Research*, 8(1):184-203.
- [167] Roy B. (1994). On operational research and decision aid. *European Journal of Operational Research*, 73, 23-26.
- [168] Roy B. (1996), *Multicriteria methodology for decision aiding*, Kluwer Academic, Dordrecht (Tradução de: *Méthodologie multicritère d'aide à la décision*, Economica, Paris, 1985).
- [169] Roy B. (1999), Decision-aiding today: What should we expect?, in Tomas Gal, Theodor J. Stewart, Thomas Hanne (eds.), *Multicriteria Decision Making - Advances in MCDM Models, Algorithms, Theory, and Applications*, Kluwer Academic Publishers, 1-1-1-35.
- [170] Roy B. (2000), Réflexion sur le thème quête de l'optimum et aide à la décision, in *Décision, Prospective, Auto-Organisation - Mélanges en l'honneur de Jacques Lesourne*, Textes réunis par J. Thépot, M. Godet, F. Roubelat, A.E. Saad, Paris, Dunod, 61-83.
- [171] Roy B., Bertier P. (1971), *La Méthode ELECTRE II*. Working paper 142, SEMA.

- [172] Roy B., Bouyssou D. (1993), Aide multicritère à la décision : Methodes et cas. Economica, Paris.
- [173] Roy B., Hugonnard J.C. (1982), Ranking of suburban line extension projects on the Paris metro system by a multicriteria method, Transportation Research, Vol. 16A, No. 4, 301-312.
- [174] Roy B., e Skalka J. M. (1984), ELECTRE IS - Aspects méthodologiques et guide d'utilisation, LAMSADE - Laboratoire d'Analyse et Modélisation de Systemes pour l'Aide à la Decision, Université de Paris-Dauphine(Document n° 30).
- [175] Roy B., Vincke Ph. (1981), Multicriteria analysis: Survey and new tendencies, European Journal of Operational Research, 8: 207-218.
- [176] Rulon P.J., Tiedeman D.V., Tatsuoka M.M., Langmuir C.R. (1967). Multivariate Statistics for Personnel Classification. Wiley, New York.
- [177] Saaty T.L.(1980), The analytic hierarchy process, McGraw-Hill, New York.
- [178] Saaty T.L.(1996), Decision making with dependence and feedback: the analytic network process. RWS Publications, Pittsburg.
- [179] Shen L., Tay F.E.H., Qu L., Shen Y. (2000). Fault diagnosis using rough sets theory. Computers in Industry 43, 61-72.
- [180] Silver M.S. (1991), Systems that support decision makers: description and analysis, Wiley, Chichester.
- [181] Simon H. (1977), The new science of management decision, Prentice-Hall, Englewood Cliffs.
- [182] Simons T.L., e Peterson R.S. (2000). Task conflict and relationship conflict in top management teams: the pivotal role of intragroup trust. Journal of Applied Psychology, 85, 102-111.
- [183] Siskos Y., Grigoroudis E., Zopounidis C., Saurais O. (1998). Measuring customer satisfaction using a survey based preference disaggregation model. Journal of Global Optimization 12 (2), 175-195.

- [184] Siskos Y., e Spyridakos A. (1999), Intelligent multicriteria decision support: Overview and perspectives, *European Journal of Operational Research* 113(2): 236-246.
- [185] Slowinski R., Zopounidis C. (1995). Application of the rough set approach to evaluation of bankruptcy risk. *International Journal of Intelligent Systems in Accounting, Finance and Management* 4 (1), 27-41.
- [186] Sneath P., e Sokal R. (1973), *Numerical Taxonomy*, Freeman, London, Uk.
- [187] Specht D.F. (1990), Probabilistic Neural Networks, *Neural Networks*, vol. 3, pp.109-118.
- [188] Stefanowski J., Slowinski R. (1998). Rough set theory and rule induction techniques for discovery of attribute dependencies in medical information systems. *Bulletin of the Polish Academy of Sciences, ser. Technical Sciences* 46 (2), 247-263.
- [189] Tannock R. (1998), Attention Deficit Hyperactivity Disorder: Advances in cognitive, neurobiological, and genetic research, *Journal of the Child Psychology and Psychiatry*, 39, 65-99.
- [190] Tervonen T., Lahdelma R., Dias J.A., Figueira J., Salminen P. (2007), SMAA-TRI: a parameter stability analysis method for ELECTRE-TRI. In *Environmental Security in Harbors and Coastal Areas*, Kiker GA, Linkov I (eds). Springer: Berlin, 217-231.
- [191] Theodoridis S., e Koutroubas K. (1999), *Pattern Recognition*, Academic Press.
- [192] Tsoukiàs A. (2007), On the concept of decision aiding process: an operational perspective. *Annals OR* 154(1): 3-27.
- [193] Tsoukiàs A. (2008), From decision theory to decision aiding methodology, *European Journal of Operational Research*, 187(1):138-161.
- [194] Tsumoto S. (1998). Automated extraction of medical expert system rules from clinical databases based on rough set theory. *Information Sciences* 112, 67-84.
- [195] Tung A.K.H., Hou J., Lakshmanan L.V.S., e Ng R.T. (2001). Constraint-Based Clustering in Large Databases. *Proceedings of the Eighth International Conference on Database Theory*, London, U.K, pp. 405-419.

- [196] Vanderpooten D., Vincke P. (1989), Description and Analysis of some Representative Interactive Multicriteria Procedures, *Mathematical and Computer Modelling*, 12, 1221-1238.
- [197] Vincke P. (1992), *Multicriteria Decision Aid*, John Wiley, New York.
- [198] Vincke P., (1999), Outranking approach, *in GAL*, T. et al.(eds),11.1-11.29.
- [199] Wang W., Yang J., e Muntz R. (1997), *STING: A Statistical Information Grid Approach to Spatial Data Mining*, International Conference on Very Large Databases. Los Angeles, California.
- [200] Ward J. (1963), Hierarchical grouping to optimize an objective function, *Journal of the American Statistical Association* 58, 236-244.
- [201] Winterfeldt D. V., e Edwards W. (1986), *Decision Analysis and Behavioral Research*. New York, Cambridge University Press.
- [202] Young T.Y., Fu K.S. (1997). *Handbook of Pattern Recognition and Image Processing*. Handbooks in Science and Technology. Academic Press, New York.
- [203] Yu W. (1992), *Aide multicritère à la décision dans le cadre de la problématique du tri: concepts, méthodes et applications*. Thèse de doctorat, Université Paris-Dauphine.
- [204] Zhang T., Ramakrishnan R., e Livny M. (1996), BIRCH: an efficient data clustering method for very large databases, *SIGMOD Record*, 25(2):103-114.
- [205] Zeleny M. (1982), *Multiple Criteria Decision Making*. New York, McGraw-Hill.
- [206] Zopounidis C. (1998). *Operational Tools in the Management of Financial Risks*. Kluwer Academic Publishers, Dordrecht.
- [207] Zopounidis C., Doumpos M. (1999), A multicriteria decision aid methodology for sorting decision problems: The case of financial distress. *Computational Economics* 14(3), 197-218.
- [208] Zopounidis C., Doumpos M. (2000), Building additive utilities for multi-group hierarchical discrimination: The M.H.DIS method. *Optimization Methods and Software* 14(3), 219-240.

-
- [209] Zopounidis, C., Doumpos, M. (2002). Multicriteria classification and sorting methods: A literature review. *European Journal of Operational Research* 138 (2), 229-246.
- [210] Zopounidis C., Slowinski R., Doumpos M., Dimitras A.I., Susmaga R. (1999). Business failure prediction using rough sets: A comparison with multivariate analysis techniques. *Fuzzy Economic Review*, 4(1), 3-33.

Formulação matemática com o princípio $S\tau$ -Consistency

A.0.1 Formulação matemática

Com o objectivo de fazer a atribuição das alternativas do conjunto \mathcal{A} o mais compatível possível com o princípio $S\tau$ -Consistency, os corolários (A.1), (A.2) e (A.3) devem-se verificar.

Sejam $a_i \in C^s$ e $a_j \in C^t$.

- (1) Se uma acção a_j é preferível a a_i então deve pertencer a uma classe melhor:

$$(a_j P a_i \wedge a_i \in C^s \wedge a_j \in C^t) \Rightarrow C^t \succ C^s, \text{ i.e.}$$

$$(C^s \succ C^t \vee C^s \perp C^t \vee s = t) \Rightarrow (a_j \not P a_i \vee a_i \notin C^s \vee a_j \notin C^t) \quad (\text{A.1})$$

- (2) Acções com preferência similar (indiferentes) devem pertencer à mesma classe:

$$(a_i I a_j \wedge a_i \in C^s \wedge a_j \in C^t) \Rightarrow s = t.$$

i.e

$$(C^t \succ C^s \vee C^s \succ C^t \vee C^s \perp C^t) \Rightarrow (a_j \not I a_i \vee a_i \notin C^s \vee a_j \notin C^t), \quad (\text{A.2})$$

- (3) Acções incomparáveis devem pertencer a classes incomparáveis:

$$(a_i R a_j \wedge a_i \in C^s \wedge a_j \in C^t) \Rightarrow C^s \perp C^t, \text{ i.e.}$$

$$(C^t \succ C^s \vee C^s \succ C^t \vee s = t) \Rightarrow (a_j \not R a_i \vee a_i \notin C^s \vee a_j \notin C^s) \quad (\text{A.3})$$

Seja x_{ij} uma variável binária 0-1 definida tal que $x_{ij} = 1$ se a alternativa $a_i \in C^j$ e $x_{ij} = 0$ se a alternativa $a_i \notin C^j$ ($i=1, \dots, m; j=1, \dots, n_{cat}$). Seja ainda p_{ij} uma variável binária 0-1 definida tal que $p_{ij} = 1$ se $C^i \succ C^j$ e $p_{ij} = 0$ caso contrário ($i, j=1, \dots, n_{cat}$).

Na atribuição das alternativas a uma estrutura com classes parcialmente ordenadas, de acordo com o princípio $S\tau$ -Consistency, temos assim as restrições que se seguem:

1. Sejam $C^s, C^t \in \mathcal{P}$ e seja $a_j \text{ P } a_i$. Por (A.1), a_j não deve pertencer a C^t ($x_{jt} = 0$) ou a_i não deve pertencer a C^s ($x_{is} = 0$), caso $C^s \succ C^t$ ($p_{st} = 1$), ou $C^s \perp C^t$ ($p_{st} = 0 \wedge p_{ts} = 0$), ou $s=t$. As respectivas restrições são as seguintes:

$$x_{jt} + x_{is} + p_{st} \leq 2 \quad (\text{A.4})$$

$$x_{jt} + x_{is} - p_{ts} - p_{st} \leq 1 \quad (\text{A.5})$$

$$x_{js} + x_{is} \leq 1 \quad (\text{A.6})$$

2. Sejam $C^s, C^t \in \mathcal{P}$ e seja $a_j \text{ I } a_i$. Por (A.2), caso $C^s \perp C^t$ ($p_{st} = 0 \wedge p_{ts} = 0$), ou $C^s \succ C^t$ ($p_{st} = 1$), a_j não deve pertencer a C^t ($x_{jt} = 0$) ou a_i não deve pertencer a C^s ($x_{is} = 0$). No caso de $C^s \succ C^t$, deve-se ainda verificar: a_j não deve pertencer a C^s ($x_{js} = 0$) ou a_i não deve pertencer a C^t ($x_{it} = 0$). As respectivas restrições são as seguintes:

$$x_{jt} + x_{is} - p_{ts} - p_{st} \leq 1 \quad (\text{A.7})$$

$$x_{jt} + x_{is} + p_{st} \leq 2 \quad (\text{A.8})$$

$$x_{js} + x_{it} + p_{st} \leq 2 \quad (\text{A.9})$$

3. Sejam $C^s, C^t \in \mathcal{P}$ e $a_j \text{ R } a_i$. Por (A.3), caso $s = t$ ou $C^s \succ C^t$ ($p_{st} = 1$), a_j não deve pertencer a C^t ($x_{jt} = 0$) ou a_i não deve pertencer a C^s ($x_{is} = 0$). No caso de $C^s \succ C^t$, deve-se ainda verificar: a_j não deve pertencer a C^s ($x_{js} = 0$) ou a_i não deve pertencer a C^t ($x_{it} = 0$). As respectivas restrições são as seguintes:

$$x_{js} + x_{is} \leq 1 \quad (\text{A.10})$$

$$x_{jt} + x_{is} + p_{st} \leq 2 \quad (\text{A.11})$$

$$x_{js} + x_{it} + p_{st} \leq 2 \quad (\text{A.12})$$

4. Cada alternativa a_i deve pertencer a uma e uma só classe C^s . Assim,

$$\sum_{s=1}^{n_{cat}} x_{is} = 1 \quad (\text{A.13})$$

5. Cada par de classes (C^i, C^j) deve ser incomparável ou ter apenas uma relação de preferência. Assim,

$$p_{ij} + p_{ji} \leq 1 \quad (\text{A.14})$$

6. Cada terno de classes (C^i, C^j, C^k) deve ser transitivo. Assim,

$$p_{ik} \geq p_{ij} + p_{jk} - 1 \quad (\text{A.15})$$

7. Caso se deseje, pode-se ainda exigir que cada classe C^s da partição não fique vazia. Logo

$$\sum_{i=1}^m x_{is} \geq 1 \quad (\text{A.16})$$

Seja γ a taxa de erro de classificação segundo o princípio $S\tau$ -Consistency, expressa da seguinte forma:

$$\gamma = \frac{1}{m(m-1)} \sum_{i,k=1}^m \sum_{j,l=1}^{n_{cat}} (\alpha_P \times \epsilon_{P_{ik,jl}} + \alpha_I \times \epsilon_{I_{ik,jl}} + \alpha_R \times \epsilon_{R_{ik,j}}) \quad (\text{A.17})$$

onde $0 \leq \alpha_P, \alpha_I, \alpha_R \leq 1$ são os custos de cada erro de violação das condições de preferência, indiferença e incomparabilidade, respectivamente, indicados pelo decisor.

O Problema Linear Inteiro (PLI 1) seguinte pode ser resolvido para encontrar o valor mínimo γ na atribuição de um conjunto de alternativas \mathcal{A} a serem atribuídas a uma determinada estrutura de classes parcialmente ordenadas \mathcal{P} , **desconhecida a priori** (mas com o número máximo de categorias, n_{cat} , especificado):

$$(PLI1) \quad \min \gamma \quad (A.18)$$

sujeito a :

$$x_{ij} + x_{kl} + p_{jl} \leq 2 + \varepsilon_{P_{ik,jl}}, \forall a_i, a_k \in A : a_i P a_k, \forall j, l \in \{1, \dots, n_{cat}\} \quad (A.19)$$

$$x_{ij} + x_{kl} + p_{jl} \leq 2 + \varepsilon_{I_{ik,jl}}, \forall a_i, a_k \in A : a_i I a_k, \forall j, l \in \{1, \dots, n_{cat}\} \quad (A.20)$$

$$x_{ij} + x_{kl} + p_{jl} \leq 2 + \varepsilon_{R_{ik,jl}}, \forall a_i, a_k \in A : a_i R a_k, \forall j, l \in \{1, \dots, n_{cat}\} \quad (A.21)$$

$$x_{ij} + x_{kl} - p_{jl} - p_{lj} \leq 1 + \varepsilon_{I_{ik,jl}}, \forall a_i, a_k \in A : a_i I a_k, \forall j, l \in \{1, \dots, n_{cat}\} \quad (A.22)$$

$$x_{ij} + x_{kl} - p_{jl} - p_{lj} \leq 1 + \varepsilon_{P_{ik,jl}}, \forall a_i, a_k \in A : a_i P a_k, \forall j, l \in \{1, \dots, n_{cat}\} \quad (A.23)$$

$$x_{ij} + x_{kj} \leq 1 + \varepsilon_{R_{ik,j}}, \forall a_i, a_k \in A : a_i R a_k, \forall j \in \{1, \dots, n_{cat}\} \quad (A.24)$$

$$x_{ij} + x_{kj} \leq 1 + \varepsilon_{P_{ik,j}}, \forall a_i, a_k \in A : a_i P a_k, \forall j \in \{1, \dots, n_{cat}\} \quad (A.25)$$

$$\sum_{j=1}^{n_{cat}} x_{ij} = 1, \forall i \in \{1, \dots, m\} \quad (A.26)$$

$$p_{ij} + p_{ji} \leq 1, \forall i, j \in \{1, \dots, n_{cat}\} \quad (A.27)$$

$$p_{ik} \geq p_{ij} + p_{jk} - 1, \forall i, j, k \in \{1, \dots, n_{cat}\} \quad (A.28)$$

$$x_{ij}, p_{ij}, \varepsilon_{P_{ik,jl}}, \varepsilon_{I_{ik,jl}}, \varepsilon_{R_{ik,j}} \in \{0, 1\}, \forall i, k \in \{1, \dots, m\}, j, l \in \{1, \dots, n_{cat}\}$$

e opcionalmente,

$$\sum_{i=1}^m x_{is} \geq 1, \forall i \in \{1, \dots, n_{cat}\} \quad (A.29)$$

Caso a estrutura das classes seja **conhecida a priori**, o (PLI 1) fica consideravelmente mais simples, com bastantes menos restrições e variáveis, resultante da inexistência de condições para a estrutura das classes. Obtém-se assim o (PLI 2).

$\min \gamma$

sujeito a:

$$x_{ij} + x_{kl} \leq 1 + \varepsilon_{P_{ik,jl}}, \forall a_i, a_k \in A : a_i P a_k \wedge (C^l \succ C^j \vee C^l \perp C^j), \forall j, l \in \{1, \dots, n_{cat}\}$$

$$x_{ij} + x_{kl} \leq 1 + \varepsilon_{I_{ik,jl}}, \forall a_i, a_k \in A : a_i I a_k \wedge (C^l \succ C^j \vee C^l \perp C^j), \forall j, l \in \{1, \dots, n_{cat}\}$$

$$x_{ij} + x_{kl} \leq 1 + \varepsilon_{R_{ik,jl}}, \forall a_i, a_k \in A : a_i R a_k \wedge C^l \succ C^j, \forall j, l \in \{1, \dots, n_{cat}\}$$

$$\sum_{j=1}^{n_{cat}} x_{ij} = 1, \forall i \in \{1, \dots, m\}$$

$$x_{ij}, \varepsilon_{P_{ik,jl}}, \varepsilon_{I_{ik,jl}}, \varepsilon_{R_{ik,j}} \in \{0, 1\}, \forall i, k \in \{1, \dots, m\}, j, l \in \{1, \dots, n_{cat}\}$$

e opcionalmente, $\sum_{i=1}^m x_{is} \geq 1, \forall i \in \{1, \dots, n_{cat}\}$