**Funções-valor
na
alocação de tarefas
em
sistemas robóticos**

Antonio Luiz Basile

DISSERTAÇÃO APRESENTADA
AO
INSTITUTO DE MATEMÁTICA ESTATÍSTICA
DA
UNIVERSIDADE DE SÃO PAULO
PARA
OBTENÇÃO DO TÍTULO DE MESTRE
EM
CIÊNCIAS

Área de Concentração: Ciência da Computação
Orientador: Prof. Dr. Flávio Soares Corrêa da Silva

São Paulo, dezembro de 2006

**Funções-valor
na
alocação de tarefas
em
sistemas robóticos**

Este exemplar corresponde à redação final
da dissertação devidamente corrigida e
defendida por Antonio Luiz Basile
e aprovada pela Comissão Julgadora.

São Paulo, 23 de dezembro de 2006

Banca examinadora:

Prof. Dr. Flávio Soares Corrêa da Silva - IME-USP
Profa. Dra. Leliane Nunes de Barros - IME-USP
Prof. Dr. Paulo Eduardo Santos - IAAA-FEI

Abstract

An important aspect of robotic's research is task allocation: how to select one (or more) robot(s) to accomplish one (or more) tasks, in order to maximize overall performance. A crucial, but very ignored issue of robot task allocation is utility/value functions.

This dissertation deals with main theoretical aspects of measurement theory and decision theory that involve the determination of a function value for multirobotic systems.

Resumo

Um importante aspecto da pesquisa robótica é a alocação de tarefas: como selecionar um (ou mais) robô(s) para realizar uma (ou mais) tarefas, visando maximizar o desempenho geral do sistema. Um ponto crucial, mas muito ignorado, da alocação de tarefas para sistemas robóticos são as funções utilidade e as funções-valor.

Essa dissertação trata os principais aspectos teóricos da teoria da medição e da teoria da decisão, envolvidos na determinação de uma função valor para sistemas robóticos.

Sumário

1	Introdução	1
1.1	Assuntos Não Tratados Nessa Dissertação	2
1.2	Alocação em Sistemas Robóticos Situados	3
2	Notação e Considerações Preliminares	5
2.1	Problema da Perseguição-Evasão	5
2.1.1	O Problema	5
2.1.2	Notação	8
2.1.3	Triangularização de Polígonos	9
2.2	Algoritmo e Simulador para o Problema da Perseguição-Evasão	10
2.2.1	Características Operacionais	13
2.3	Hierarquia de Tarefas	15
3	Atributos e Escalas	19
3.1	Considerações Iniciais	19
3.2	Conjunto de Atributos Relevantes	20
3.2.1	Conjunto de Atributos Completo	22
3.2.2	Conjunto de Atributos Operacional	25
3.2.3	Conjunto de Atributos Decomponível	25
3.2.4	Conjunto de Atributos Não-Redundante	26
3.2.5	Conjunto de Atributos Minimal	26
3.3	Teoria da Medição: Escalas	28
3.3.1	Medição Ordinal	33

3.3.2	Medição de Diferenças	34
3.3.3	Medição Extensiva	36
3.3.4	Medição Conjunta	41
3.3.5	Normalização	43
3.4	Escalas e Modelagem de Problemas de Otimização	44
3.4.1	Otimização Combinatória	46
3.4.2	Programação 0 – 1	48
4	Teoria da Decisão Multiatributo	49
4.1	Determinação dos Pesos	49
4.1.1	Método Entrópico	50
4.2	Taxonomia para Teoria da Decisão	53
4.3	Métodos sem Informação	54
4.3.1	Dominância	55
4.3.2	Maximin	56
4.4	Métodos com Informação Sobre Atributos	56
4.4.1	Métodos de Nível Mínimo	57
4.4.2	Métodos Ordinais	58
4.4.3	Métodos Cardinais	60
4.4.4	Taxa Marginal de Substituição	72
4.5	Métodos com Informação Sobre Alternativas	73
4.5.1	Preferência Par a Par	73
4.5.2	Ordem de Proximidade Par a Par	74
4.6	Teoria da Utilidade Multiatributo	74
4.6.1	Modelo Ordinal	74
4.6.2	Modelo Cardinal Multiplicativo	75
4.6.3	Modelo Cardinal Aditivo	77
4.7	Conclusão	77
5	Trabalhos Futuros	79
6	Conclusão	81

6.1 Contribuições Originais	81
Referências Bibliográficas	82

Lista de Figuras

2.1	Visibilidade entre pontos de um polígono	6
2.2	Polígono não-simples.	6
2.3	Triangularização de um polígono.	7
2.4	Polígonos convexos e não-convexos	7
2.5	Tipos de particionamento convexo	8
2.6	Descontaminação de um polígono	9
2.7	Uma diagonal particionando um polígono em dois polígonos menores	10
2.8	Orelha de um polígono	11
2.9	Dual de uma triangularização.	12
2.10	Triângulos adjacentes	12
2.11	Árvore de busca	13
2.12	Etapas do algoritmo	14
2.13	Características operacionais do simulador	15
2.14	Possível hierarquia para o robô-fugitivo no problema da perseguição-evasão.	17
2.15	Hierarquia de objetivos abstratos.	18
3.1	Função-Valor adequada ao problema	23
3.2	Conjunto de atributos não-completo.	24
3.3	Atributo irrelevante é considerado.	26
3.4	Atributo redundante.	27
3.5	Grafo dos caminhos possíveis ao robô	45
4.1	Taxonomia para teoria da decisão multiatributo	54
4.2	Hierarquia Z-Y-X	63

4.3	Escolha de método para decisão multiatributo	78
-----	--	----

Lista de Tabelas

3.1	Determinação do grau de importância relativa dos atributos.	21
3.2	Atributos dos robôs.	23
3.3	Atributos dos robôs—normalizados $[0,1]$	23
3.4	Atributos dos robôs com um atributo irrelevante.	25
3.5	Atributos dos robôs com um atributo irrelevante—normalizados $[0,1]$	25
3.6	Atributos redundantes.	27
3.7	Atributos redundantes—normalizados $[0,1]$	27
3.8	Definições de medição encontradas na literatura e citadas por Roberts em [Rob79].	30
3.9	Tipos de escalas segundo Stevens.	32
3.10	Preferência viola ordem simples estrita.	39
3.11	Operações estatísticas permitidas por escala.	43
3.12	Temperaturas em graus Célsius.	45
3.13	Temperaturas em Fahrenheit.	45
4.1	Eixos da Taxonomia MADM.	53

Capítulo 1

Introdução

A pesquisa em sistemas robóticos—sistemas formados por robôs situados (inseridos no mundo físico) ou simulados (softbots)—embora com foco em aspectos diversos, tais como colaboração, coordenação, cooperação e comunicação, freqüentemente trata (principalmente em sistemas centralizados) o tema da alocação de tarefas, que é um tema central para a área. Esse tema tem grande importância para a área porque um sistema robótico, geralmente, está associado à execução de tarefas. Se houver disponibilidade de dois ou mais robôs aptos, deveremos escolher o robô mais adequado para realizar a tarefa (sistema de robô único) ou o grupo de robôs mais adequado para realizar a tarefa (sistemas multirobóticos).

O problema da alocação de tarefas em sistemas robóticos consiste em, dado um conjunto de tarefas, um conjunto de robôs e uma função que associa um valor escalar a cada par (robô, tarefa), encontrar a melhor distribuição de robôs para as tarefas com o intuito de otimizar a soma dos valores dos pares escolhidos (pode haver robôs que executem mais de uma tarefa simultaneamente, da mesma forma como pode haver tarefas que precisem de mais de um robô para serem terminadas). Embora a definição abranja casos de tarefas que precisem de um grupo de robôs para serem realizadas—o caso multirobótico—estudaremos neste documento apenas o caso em que apenas um robô é alocado por tarefa¹ e em que há apenas uma tarefa a ser realizada.

Um estudo mais aprofundado sobre o tema pode ser encontrado em [Ger03], onde o autor taxonomiza a área sob o enfoque da alocação de tarefas. Em seu trabalho o autor destaca abordagens dadas por diversas outras áreas, tais como a economia, a pesquisa operacional, escalonamento, fluxo em redes e otimização combinatória. O problema é estudado, em cada um de seus casos, com enfoque nos algoritmos e heurísticas para a alocação de tarefas.

A pesquisa em sistemas robóticos, porém, raramente trata de um tema muito importante para a área, visto sua relevância para a própria alocação de tarefas: a função-valor (caso determinístico) ou função-utilidade (quando há incerteza envolvida). Isso se deve, dentre outras, a

¹Portanto, casos em que haja vários robôs disponíveis para a alocação serão considerados casos de robô único, visto que apenas um robô será alocado.

quatro características encontradas na modelagem dessas funções: (1) a subjetividade² que envolve o tema, (2) a dificuldade em se obter funções-valor (utilidade) ótimas, (3) a diversidade de domínios de aplicação, impossibilitando a generalização dessas funções e (4) o conhecimento prévio, para diversos domínios, de funções que trocam a precisão por simplicidade. Neste documento trataremos do caso determinístico, ou seja, discutiremos aspectos referentes à composição de funções-valor e não funções utilidade.

Como veremos, muitos aspectos distintos dizem respeito à criação de uma função-valor. A subjetividade, sem dúvida, é um deles. Se solicitarmos, por exemplo, a um grupo de pessoas citar quais são os atributos relevantes que um robô deve possuir para realizar uma dada tarefa, é provável que sejam apresentadas várias soluções. Porém, a subjetividade não é o único aspecto (e provavelmente nem o mais importante) e como veremos, há algumas diretivas na literatura em teoria da decisão e da utilidade, que auxiliam sua abordagem.

A dificuldade na obtenção de funções-valor ótimas é um argumento válido. Porém, não estamos procurando funções-valor ótimas, mas funções-valor melhores que as usualmente adotadas. Dispomos para isso do auxílio de técnicas de diversas áreas que estudam há décadas o tema—como, por exemplo, a teoria da decisão, a teoria da utilidade e a teoria da medição.

A diversidade de domínios provavelmente impossibilite a obtenção de uma função-valor genérica. Não é claro o tratamento que deve ser dado à função-valor, para que esta seja válida, simultaneamente, em domínios tais como vigilância, recolhimento de lixo tóxico e carga e descarga, por exemplo. Robôs adequados para o domínio da perseguição-evasão, provavelmente não seriam os mais indicados para recolhimento de lixo tóxico. Porém, isso não é obstáculo para esse trabalho, que não tem a pretensão de generalizar a função-valor.

O último argumento, da troca da precisão pela simplicidade e pelo conhecimento prévio é válido. Cabe, portanto, à pesquisa por funções-valor em sistemas robóticos fornecer funções novas e mais precisas, que futuramente farão parte do arcabouço de funções conhecidas, simplificando o trabalho.

Esta dissertação tem como objetivo salientar a importância da função-valor para alocação em sistemas robóticos, bem como propor melhorias e alternativas a suas formulações mais comuns. Para isso estudaremos a teoria da decisão e a teoria da medição, com ênfase em funções multiatributo. Também ilustraremos, através de um problema único, que nenhum atributo individualmente discrimina sempre o melhor robô para uma determinada tarefa.

1.1 Alocação em Sistemas Robóticos Situados

A utilização de sistemas robóticos vai desde tarefas que envolvam grande perigo aos humanos, como vigilância e recolhimento de lixo tóxico, até tarefas triviais e repetitivas, como carga e descarga de mercadorias.

²Ou seja, diante de um problema muito complexo, se dá soluções diversas e muitas vezes intuitivas. O problema pode admitir mais de uma solução correta.

A má distribuição dos robôs pode causar atrasos, aumentar o custo do sistema (por exemplo, alocando robôs de alto custo de operação para tarefas triviais) e até inviabilizar a tarefa toda, caso não seja efetuada realocação (como por exemplo na alocação de robôs com problemas operacionais ou de robôs inaptos).

Como dissemos, para que a alocação seja possível precisaremos comparar os robôs e para isto associaremos valores escalares a cada par (robô, tarefa). Entretanto, não poderemos associar arbitrariamente esses valores escalares, pois desejamos que eles expressem adequadamente a realidade, de preferência criando uma relação de ordem entre os robôs, para cada tarefa.

No capítulo 2 apresentaremos o problema da perseguição-evasão³ que ilustrará vários aspectos discutidos ao longo da dissertação. A perseguição-evasão é uma variação do problema da galeria de arte, mais especificamente o problema da galeria de arte com guardas móveis. O problema da galeria de arte (com guardas fixos) trata da determinação do número mínimo de guardas necessários para cuidar de uma galeria de arte. É um problema geométrico de decomposição de polígonos [Joh03] e foi inicialmente tratado por Klee e Chvátal [Chv75]. A idéia é usar um polígono como representação geométrica da galeria e através de sua decomposição, tornar esse problema mais tratável. O problema da perseguição-evasão trata de como guiar um ou mais perseguidores para: (1) encontrar ou (2) capturar um ou mais fugitivos e tem aplicações diversas, tais como a procura por intrusos em prédios, controle de tráfego e estratégia militar [Che03]. No nosso caso, o tratamento dado ao problema será de robô único, pois mesmo dispondo de mais de um robô apto para a alocação apenas um robô será incumbido de realizar a tarefa de um perseguidor encontrar um fugitivo.

Como no nosso caso utilizaremos um simulador em que o usuário faz o papel de fugitivo e o computador (algoritmo) faz o papel de perseguidor, adotamos o papel de fugitivo. Também alteramos o problema da perseguição-evasão (baseada em visibilidade). No problema original o fugitivo tinha como objetivo conseguir se esconder indefinidamente do perseguidor (não ser visto por ele). No problema alterado o fugitivo deve permanecer o maior tempo possível sem ser encontrado. Essas alterações puderam ser efetuadas em razão da tarefa em si não ser o foco principal do nosso trabalho.

No Capítulo 3 apresentaremos importantes aspectos referentes a escalas—estudados na teoria da medição—que influenciam a determinação de funções-valor.

No Capítulo 4 apresentaremos o método entrópico para a determinação dos pesos de uma função valor. Também serão descritos e ilustrados vários modelos da teoria da decisão. Para isso nos basearemos principalmente na taxonomia de [HY81], que descreve a área a partir de três eixos: métodos sem informação, métodos com informação sobre atributos e métodos com informação sobre alternativas.

Nessa dissertação pretendemos, portanto, discutir aspectos pertinentes à tomada de decisão na alocação em sistemas robóticos, seja através de modelos da teoria da decisão, seja através da determinação de funções-valor corretas sob o ponto de vista da teoria da decisão e da teoria da

³Do inglês “pursuit-evasion”.

medição.

1.2 Assuntos Não Tratados Nessa Dissertação

Há outras áreas que têm algumas etapas análogas as descritas nessa dissertação. Uma delas é “Descoberta de Conhecimento em Banco de Dados”⁴, que consiste em identificar padrões em dados visando a tomada de decisão ou aumentar o entendimento de um problema específico. Entre suas etapas encontram-se integração, transformação e mineração de dados⁵. A primeira lida com a combinação de diferentes bases de dados e a segunda visa modificar ou transformar dados em formatos apropriados à mineração [HK00]. A mineração de dados visa selecionar os métodos a serem utilizados para localizar padrões nos dados, seguida da busca propriamente dita por padrões relevantes numa dada forma de representação [HK00]. Em todas elas há pontos em comum com o que estudamos nesse documento, como por exemplo, a seleção de atributos na mineração de dados ou a normalização na transformação de dados. Ressalta-se portanto que esse documento não tratará de funções-valor em sistemas robóticos sob o ponto de vista da descoberta de conhecimento em bancos de dados, pois adotará o ponto de vista da teoria da decisão. A escolha pela teoria da decisão deve-se ao fato de que—através de vários modelos, desenvolvidos ao longo de muitas décadas—ela estuda a tomada de decisão sob diversos pontos de vista na representação do problema, o que nos permite lidar com uma variedade grande de domínios—que é característica inerente do estudo em sistemas robóticos.

⁴Do inglês “Knowledge Discovery in Databases”.

⁵Do inglês “Data Mining”.

Capítulo 2

Notação e Considerações Preliminares

Neste capítulo apresentaremos o problema da perseguição-evasão, posteriormente utilizado para ilustrar importantes aspectos da teoria da decisão e da teoria da medição (através de simulações).

Na Seção 2.1 o problema da perseguição-evasão com um perseguidor e um fugitivo será descrito e formalizado. Também será apresentada uma técnica para a decomposição de um polígono simples em triângulos. Na Seção 2.2 apresentaremos o algoritmo utilizado nas simulações, proposto em [IKK04], bem como o simulador propriamente dito.

2.1 Problema da Perseguição-Evasão

A tarefa escolhida para ilustrar os pontos tratados sobre funções-valor será a perseguição-evasão. Conforme mencionamos na Seção 1.2, perseguição-evasão é similar ao problema da galeria de arte com guardas móveis. O problema da galeria de arte (com guardas fixos) trata da determinação do número mínimo de guardas necessários para cuidar de uma galeria de arte. É um problema geométrico de decomposição de polígonos [Joh03] e foi inicialmente tratado por Klee e Chvátal [Chv75]. A idéia é usar um polígono como representação geométrica da galeria e, através de sua decomposição, tornar esse problema mais tratável. O problema da perseguição-evasão trata de como guiar um ou mais perseguidores para: (1) encontrar ou (2) capturar um ou mais fugitivos. Esse problema tem aplicações diversas, tais como a procura por intrusos em prédios, controle de tráfego e estratégia militar [Che03].

2.1.1 O Problema

Trataremos aqui o *problema da perseguição-evasão baseado na visibilidade*, ou seja, o perseguidor tem como tarefa encontrar (ver) o fugitivo. Como veremos, isso é obtido se ambos estiverem simultaneamente dentro de uma mesma região de visibilidade do polígono, ou seja,

uma região em que quaisquer dois pontos u e v possam ser unidos por um segmento de reta uv que esteja inteiramente contido dentro do polígono. A Figura 2.1 ilustra tanto uma situação em que dois pontos são visíveis um ao outro, quanto uma situação em que não são.

Figura 2.1: Visibilidade entre pontos de um polígono. Podemos unir os pontos u e v através de um segmento de reta uv inteiramente contido no polígono. Nesse caso dizemos que u e v são visíveis um ao outro ou que estão na mesma região de visibilidade. Já os pontos u e x são exemplos de dois pontos que não são visíveis um ao outro, pois se os unirmos através de um segmento de reta ux , parte desse segmento será externo ao polígono.

Como lidaremos apenas com polígonos simples, sempre que nos referirmos a polígonos, estaremos nos referindo a polígonos simples. Em 2.1 é dada uma definição para polígono simples [O'R98] e a Figura 2.2 ilustra polígonos não-simples através de dois exemplos.

Definição 2.1 (Polígono Simples). *Sejam $v_0, v_1, v_2, \dots, v_{n-1}$, n pontos no plano. Sejam $(a_0 = v_0v_1), (a_1 = v_1v_2), \dots, (a_i = v_iv_{i+1}), \dots, (a_{n-1} = v_{n-1}v_0)$, n segmentos conectando os pontos. Então esses segmentos limitarão um polígono se e somente se:*

1. *A intersecção de cada par de segmentos adjacentes em ordem cíclica for o único ponto compartilhado por eles: $(\forall i = 0, \dots, n-1)(a_i \cap a_{i+1} = v_{i+1})$.*
2. *Segmentos não-adjacentes não se interseccionarem: $(\forall j \neq i+1)(a_i \cap a_j = \emptyset)$.*

Figura 2.2: Polígono não-simples.

Tanto o problema da galeria de arte original como o problema da perseguição-evasão têm a visibilidade como tema central, pois como vimos, o problema da galeria de arte trata da busca por

um conjunto de guardas que cubra um polígono simples, ou seja, que cada ponto desse polígono seja visível a pelo menos um guarda.

Um dos processos mais comuns para subdividir polígonos em regiões de mesma visibilidade é a triangularização, que decompõe o polígono em triângulos sem acrescentar vértices, como ilustrado na Figura 2.3. Detalharemos a triangularização na Seção 2.1.3.

Figura 2.3: Triangularização de um polígono.

Particionar o polígono em triângulos é interessante, porque discretiza o problema através de regiões de mesma visibilidade, no caso triângulos. Triângulos são regiões de mesma visibilidade, pois são polígonos convexos [Dag97], conforme a Definição 2.2.

Definição 2.2 (Polígono Convexo). *Um polígono é dito convexo se quaisquer dois pontos internos a ele puderem ser ligados por um segmento que não interseccione os limites do polígono.*

A Figura 2.4(b) ilustra um polígono não-convexo. Note que esse polígono possui pontos que contrariam a Definição 2.2. Os pontos u e v exemplificam isso, pois o segmento uv que liga esses pontos interseccionará os limites do polígono.

Figura 2.4: Em (a) temos um exemplo de *polígono convexo*. Note que um segmento que ligue quaisquer dois pontos internos ao polígono nunca interseccionará seus limites. Em (b) temos um exemplo de *polígono não-convexo* e uma possível violação para a Definição 2.2.

Embora a triangularização seja um importante passo—incorporado por muitos algoritmos na geometria computacional (assim como o algoritmo que utilizaremos)—não é a única maneira de se particionar um polígono em polígonos convexos (usualmente chamado de particionamento

convexo¹), mas apenas um caso particular de particionamento convexo.

Há duas importantes formas de se tratar o problema do particionamento convexo. Uma delas se preocupa em obter o menor número de polígonos convexos possível e a outra em terminar o processo no menor tempo possível. Também há duas maneiras de se particionar o polígono: partição através de diagonais ou partição através de segmentos. O que as distingue são os pontos extremos². Conforme ilustrado na Figura 2.5, no particionamento através de diagonais os pontos extremos devem ser vértices, enquanto no particionamento através de segmentos podem ser quaisquer pontos do polígono. No caso (b) da Figura 2.5 os pontos u e v são exemplos de pontos internos ao polígono usados para o particionamento.

Figura 2.5: Particionamento convexo através de (a)Diagonais e (b)Segmentos.

Em nosso problema, utilizaremos apenas um perseguidor e um fugitivo. A posição do fugitivo é desconhecida e a posição do perseguidor é conhecida. A tarefa é encontrar um caminho para o perseguidor que garanta achar o fugitivo, qualquer que seja a velocidade e o trajeto do fugitivo. A seguir será apresentada a notação que adotaremos para esse problema, que é a mesma encontrada em [IKK04].

2.1.2 Notação

Seja P um espaço livre poligonal, V seu conjunto de vértices e $|V| = n$. O comprimento do menor caminho de u a v , que permanece dentro de P , é denotado por $d(u, v)$. Também atribui-se a $d(u, v)$ a seguinte propriedade: $d(u, v)$ é um caminho poligonal cujos vértices internos são vértices de P .

Uma estratégia (caminho) do fugitivo é uma função contínua $f : [0, \infty) \rightarrow P$, tal que $f(t)$ denota a posição do fugitivo no tempo $t > 0$. A estratégia do perseguidor é definida de forma similar. Para qualquer ponto $x \in P$, chamaremos de $Vis(x)$ o conjunto de todos os pontos em P visíveis de x [PLC01]. Se o perseguidor tiver visão 360° podemos dizer que ele pode ver todos os pontos em $Vis(p(t))$ no instante t . Diz-se que um polígono P é viável³ (quanto à

¹Do inglês “convex partitioning”.

²Do inglês “endpoints”.

³Do inglês “searchable”.

busca do fugitivo) se houver um caminho p tal, que para cada função contínua $f : [0, \infty) \rightarrow P$ há um instante no tempo $t \in [0, \infty)$ tal que $f(t) \in \text{Vis}(p(t))$, ou seja, o fugitivo será visto pelo perseguidor, independente de qual estratégia (caminho) adotar. Qualquer região que possa conter um fugitivo não detectado é chamada *contaminada*. Caso contrário é chamada *limpa*. Inicialmente o perseguidor deve *limpar* todo o polígono P —considerado inicialmente inteiramente contaminado. Como o fugitivo pode se locomover com velocidade arbitrária, qualquer região disjunta de $\text{Vis}(p(t))$ está ou inteiramente *limpa* ou inteiramente *contaminada*. Note que o fugitivo poderá *recontaminar* áreas já limpas pelo perseguidor. A Figura 2.6 ilustra o perseguidor descontaminando um polígono.

Figura 2.6: Descontaminação de um polígono. O trajeto efetuado pelo perseguidor. A área sombreada representa a região que ainda está contaminada, ou seja, pode abrigar um intruso.

Em [PLC01] podemos encontrar condições necessárias e suficientes para que um polígono seja viável.

2.1.3 Triangularização de Polígonos

Um importante processo usado pelo algoritmo é a decomposição do polígono. A idéia é a discretização do problema, o que no nosso caso será feita através da decomposição em triângulos. Há alternativas a esse tipo de decomposição, como, por exemplo, a decomposição em células de visibilidade, descrito em [GLL⁺99]. A seguir descreveremos sucintamente o processo da decomposição de um polígono em triângulos [Dag97]. Antes, porém, descreveremos alguns conceitos importantes citados em [O'R98]:

Definição 2.3 (Diagonal de Polígono). *Uma diagonal de um polígono P é um segmento de reta entre dois de seus vértices a e b que são visíveis um ao outro.*

Como efetuaremos a decomposição do polígono em triângulos, deve-se provar que esta decomposição é possível. O lema a seguir, cuja prova será omitida aqui, é importante para isso.

Lema 2.1 (Diagonal). *Todo polígono de $n \geq 4$ vértices possui uma diagonal.*

E finalmente por indução pode-se provar que todo polígono pode ser triangularizado.

Teorema 2.1 (Triangularização). *Todo polígono P de n vértices pode ser particionado em triângulos pela adição de (zero ou mais) diagonais.*

Demonstração. Por indução. Para $n = 3$ é trivial, pois o polígono é um triângulo. Para $n > 3$, seja ab uma diagonal de P , garantida pelo Lema 2.1. Como ab somente intersecciona o polígono em seus limites, ela particiona P em dois polígonos, cada um com menos vértices que n (nenhum vértice foi adicionado no processo e há pelo menos um vértice em cada uma das partes, além de a e b (veja a Figura 2.7). A aplicação da hipótese de indução aos dois sub-polígonos completam a prova [O'R98]. $\square$

Figura 2.7: Uma diagonal particionando um polígono em dois polígonos menores.

Na prática poderemos triangularizar um polígono através de um importante conceito que veremos a seguir.

Definição 2.4 (Orelha de Polígono). *Dizemos que três vértices consecutivos a , b e c , de um polígono P , formam um orelha do polígono em b se ac for uma diagonal.*

Ou seja, o triângulo formado por b e seus vértices adjacentes é interno ao polígono e não contém nenhum vértice de P dentro de si, pois nesses casos não teríamos uma diagonal em ac .

Teorema 2.2 (Teorema das duas orelhas de Meister's [O'R98]). *Qualquer polígono simples com pelo menos quatro vértices possui pelo menos duas orelhas não-sobrepostas⁴.*

Logo, existe um vértice v no polígono P , onde há uma orelha de P . Podemos, então, remover essa orelha e adicionar uma aresta entre os vizinhos mais próximos de v , v_1 e v_2 , obtendo um novo polígono P_1 . Repetimos o processo até só restar um triângulo. As arestas usadas nesse processo serão as diagonais usadas na triangulação do polígono.

2.2 Algoritmo e Simulador para o Problema da Perseguição-Evasão

Conforme mencionamos em 1.2, ilustraremos cada característica desejável (ou a ausência dela) de uma função-valor através da simulação do problema da perseguição-evasão. Para isso,

⁴Duas orelhas são não-sobrepostas se os seus triângulos interiores forem disjuntos.

Figura 2.8: Orelha de um polígono. (a) O polígono possui uma orelha no vértice b . (b) Neste caso o polígono não possui uma orelha em v , pois o triângulo formado por v e seus vértices adjacentes contém um vértice do polígono (nesse caso u). (c) Este polígono também não possui uma orelha em v , pois o triângulo formado por v e seus vértices adjacentes não é interno ao polígono.

utilizaremos um simulador que emprega um algoritmo similar ao proposto em [IKK04]. Trata-se de um algoritmo probabilístico para perseguição-evasão baseado em visibilidade. O perseguidor nada sabe sobre o fugitivo. O fugitivo, porém, a qualquer tempo t tem acesso à localização $p(t)$ do perseguidor.

Nesse simulador o usuário faz o papel de fugitivo e o computador (algoritmo) faz o papel de perseguidor. Logo, por este motivo operacional, adotamos o papel de fugitivo. Também alteramos o problema da perseguição-evasão (baseada em visibilidade). No problema original o fugitivo tinha como objetivo conseguir se esconder indefinidamente do perseguidor (não ser visto por ele). No problema alterado o fugitivo deve permanecer o maior tempo possível sem ser encontrado. Daqui para frente utilizaremos a denominação *problema da perseguição-evasão alterado*. Ressaltamos que essas alterações puderam ser efetuadas em razão da tarefa em si não ser o foco principal do nosso trabalho.

Inicialmente, o perseguidor efetua a triangularização do polígono. Ou seja, o polígono de visibilidade adotado pelo algoritmo, e portanto pelo simulador, é o triângulo. O resultado disso é que o perseguidor localiza o fugitivo somente se estiver no mesmo triângulo que ele. Como mencionamos na Seção 2.1 a triangularização é apenas uma das formas de particionamento de polígonos em polígonos convexos. Também mencionamos que o particionamento em polígonos convexos é interessante, pois discretiza o problema através de polígonos de mesma visibilidade (convexos).

A triangularização é particularmente interessante para esse algoritmo, pois ele utiliza uma estrutura de dados do tipo árvore para representar os possíveis caminhos do perseguidor (cada nó da árvore representa um triângulo). A raiz da árvore é o nó inicial do perseguidor. Cada nó filho da raiz representa um triângulo adjacente a ela. Um conceito que será bastante útil nesse caso é o de Dual de uma triangularização, descrito na Definição 2.5 [O'R98].

Definição 2.5 (Dual de Triangularização de Polígono [O'R98]). *O dual T da triangu-*

larização de um polígono é um grafo com um nó associado a cada triângulo e um arco entre dois nós se e somente se os seus triângulos compartilharem uma diagonal (conforme ilustrado na Figura 2.9).

Figura 2.9: Dual de uma triangularização.

Teorema 2.3 ([O'R98]). *O dual T de uma triangularização é uma árvore, com cada nó de grau no máximo três.*

Portanto, como o algoritmo lidará somente com triângulos e de acordo com o Teorema 2.3 essa representação poderá ser feita através de uma árvore ternária.

A Figura 2.10 ilustra os três casos possíveis para o polígono da Figura 2.9.

Figura 2.10: Triângulos adjacentes: três casos possíveis para polígonos simples. (a) As áreas hachuradas representam triângulos com apenas um triângulo adjacente. (b) Aqui a área hachurada representa um triângulo que possui dois triângulos adjacentes. (c) Neste último caso o triângulo da área hachurada possui triângulos adjacentes aos seus três lados.

Sua estratégia (caminho) é dividida em rodadas de tamanho máximo igual a $diam(P)$ ⁵. O perseguidor considera, então, sua árvore de triangularização com raiz no triângulo referente à sua posição inicial (em relação à rodada). Seja $T(t)$ a sub-árvore de T com raiz no triângulo t e sejam $t_1, \dots, t_k, k \leq 3$ os filhos do triângulo t , conforme a Figura 2.11. Note que se o perseguidor

⁵Como o perseguidor move-se à velocidade unitária, o diâmetro é igual ao tempo máximo que o perseguidor leva entre dois pontos em P .

estiver no triângulo t e o fugitivo estiver em um triângulo contido em $T(t_j)$, então enquanto o perseguidor estiver em t , o fugitivo não pode entrar em qualquer triângulo contido em $T(t_i)$, $i \neq j$ sem ser visto pelo perseguidor, ou seja, o triângulo t é um separador para as sub-árvores $T(t_i)$.

Figura 2.11: Árvore de busca [IKK04]. Cada vértice desta árvore corresponde a um triângulo na árvore dual de triangularização (ver Figura 2.9). $T(t)$ denota a sub-árvore com raiz no nó t e $l(t)$ denota o número de folhas da sub-árvore $T(t)$.

Ou seja, o fugitivo não conseguirá mudar de sub-árvore sem que se encontre com o perseguidor, conforme ilustrado na Figura 2.12.

Usando essa observação, o perseguidor escolherá aleatoriamente⁶ uma das sub-árvores e irá através dela. Então o perseguidor repetirá recursivamente essa estratégia, até chegar a um triângulo representado por uma folha na árvore de triangularização. Nesse instante terminará essa etapa. O perseguidor repetirá essas etapas até localizar o fugitivo.

O tempo esperado para a captura do fugitivo, com um perseguidor e um fugitivo, é no máximo $n \cdot (\text{diam}(P))$, onde n é o número de vértices de P .

2.2.1 Características Operacionais

A seguir descreveremos algumas características operacionais do simulador, que pode ser encontrado em [Isl04] e emprega um algoritmo similar ao proposto em [IKK04].

Um polígono simples qualquer deve ser desenhado pelo usuário, com no máximo 100 vértices. Para terminar o desenho, o usuário deve coincidir o seu último vértice sobre seu vértice inicial. Quando isso ocorrer, uma triangularização do polígono do usuário é gerada através do algoritmo de triangularização descrito em [Dag97]. O fugitivo é controlado pelo usuário e o perseguidor é controlado pelo algoritmo.

⁶Se $l(t)$ denota o número de folhas da sub-árvore $T(t)$ e $L = \sum_{i=1}^k l(t_i)$, o perseguidor escolherá um filho t_i com probabilidade $\frac{l(t_i)}{L}$.

Figura 2.12: Etapas do algoritmo. (a) Inicialmente o algoritmo triangulariza o polígono. (b) Em seguida determina a árvore dual de triangularização, baseada na triangularização efetuada. (c) Posição inicial do perseguidor e do fugitivo nesta rodada. (d) Sub-árvores com raiz na posição do perseguidor. (e) Não há como o fugitivo mudar de sub-árvore sem passar por t .

O usuário deve escolher uma posição inicial para o fugitivo (um triângulo qualquer gerado pela triangularização). Assim que o fizer, uma posição inicial é sorteada para o perseguidor (um triângulo diferente do escolhido para o fugitivo). Usuário e algoritmo irão se movimentar alternadamente. Quem inicia a movimentação é o fugitivo. Então o perseguidor efetua sua primeira movimentação.

O perseguidor só vê o fugitivo se estiver no mesmo triângulo que ele. Quando isso ocorre, termina a simulação.

A *eficiência* do perseguidor será determinada pela duração da simulação, ou seja, no nosso caso (problema da perseguição-evasão alterado, conforme descrito no início da Seção 2.2), será determinada pelo número de épocas (Definição 2.6) que durou a simulação.

Definição 2.6 (Época). Denotaremos por f_i o i -ésimo movimento do fugitivo e por p_i o i -ésimo movimento do perseguidor, com $i \in \mathbb{N}^*$ e finito. Chamaremos de (i -ésima) época o par ordenado (f_i, p_i) .

Uma época qualquer inicia, portanto, com um movimento do fugitivo e termina com o próximo movimento do perseguidor. Por exemplo, o par (f_2, p_2) denota o segundo movimento do fugitivo e o segundo movimento do perseguidor, ou seja, denota a segunda época.

A Figura 2.13 ilustra o funcionamento do simulador sobre um polígono simples.

Figura 2.13: Características operacionais do simulador. (a) Posições iniciais do perseguidor (área hachurada em preto) e do fugitivo (área hachurada em cinza escuro). Áreas hachuradas em cinza claro representam os possíveis próximos passos do perseguidor. (b) Segunda época. (c) Terceira época. (d) Quarta época. (e) Quinta época. (f) O fugitivo é encontrado na sexta época e a simulação é encerrada.

A escolha de qual robô deverá ser alocado se dará sobre um grupo de dois ou mais robôs heterogêneos, não necessariamente dispostos na mesma região, ou seja, ao compararmos quaisquer dois robôs desse grupo, verificaremos que diferem em pelo menos um de seus atributos.

O local será representado por um polígono simples. Assumiremos que o perseguidor possui visão de 360° e que o fugitivo está em posição desconhecida e é imprevisível. Após a alocação, os outros robôs não participarão da simulação, ou seja, não se trata de uma abordagem multirobótica.

2.3 Hierarquia de Tarefas

Há muitos casos em que uma função-valor envolve não apenas um objetivo, mas vários objetivos. Primeiro vamos definir “objetivos” e “tarefas”:

Definição 2.7 (Objetivo [HY81]). *Trata-se de algo a ser perseguido completamente. Indica a direção da mudança desejada.*

Exemplo: uma fábrica de automóveis que deseja maximizar a autonomia ou minimizar o custo de produção.

Definição 2.8 (Tarefas [HY81]). *Valores ou níveis de aspiração.*

Exemplo: atingir 20 km/litro de consumo.

Em relação ao problema da perseguição-evasão podemos desejar um robô que não apenas encontre o fugitivo, mas que o faça com o menor custo. Isto é justificável se alocamos diariamente e em vários períodos do dia um robô para descontaminar uma andar de um prédio, por exemplo.

Nessa situação um robô de menor custo parece adequado, mesmo se for mais lento que outros candidatos, pois não há suspeita de invasão. Por outro lado, numa situação em que haja suspeita de invasão (disparo de alarme, violação de porta ou janela etc.⁷) parece mais adequado alocar um robô mais rápido, mesmo que seu custo seja elevado em comparação a outros candidatos. O exemplo acima é bastante simples, pois envolve apenas dois objetivos⁸, tanto no primeiro caso como no segundo. Observe que mesmo tratando-se de um exemplo simples com dois casos similares (galeria de arte com guarda(s) móvel(is)), o problema acima exigirá funções-valor muito distintas entre si, para cada um de seus casos.

Para o nosso caso—problema da perseguição-evasão alterado (Seção 2.2)—devemos alocar um robô-fugitivo que evite ser encontrado, pelo robô-perseguidor, o maior período de tempo que puder. Logo, não se trata de um caso de redução de custo, como o primeiro caso do exemplo anterior. Provavelmente um atributo mais importante a ser considerado é a velocidade. Tampouco devemos descartar o custo “a priori”. Considere, por exemplo, uma situação em que dois robôs empatam em todos os atributos relevantes ao problema (inclusive na velocidade), mas suponha que um deles possua menor custo. Uma função-valor que (também) considere o custo parece mais adequada, nesse caso.

O que notamos nos exemplos acima é o que Raiffa denominou *natureza hierárquica dos objetivos* [KR76]. São propostas linhas gerais para a construção e uso de uma hierarquia, como forma de estruturar os objetivos de maneira compreensível (principalmente no trato de problemas complexos). Nessa hierarquia os objetivos são subdivididos em objetivos de baixo nível (mais detalhados), evidenciando o significado do objetivo mais geral. Segundo Raiffa, esses objetivos de baixo nível podem ser tratados como meios para o fim, onde o fim é o objetivo de mais alto nível da hierarquia. Esse objetivo (no topo da hierarquia) é um objetivo genérico que contém todos os outros objetivos e que indica a razão de interesse no problema, mas que em muitos casos é muito vago para servir a propósitos operacionais.

No caso do problema da perseguição-evasão alterado, o objetivo mais genérico do robô-fugitivo é *permanecer o maior tempo possível sem ser encontrado*. Quando analisamos esse problema, percebemos que por trás desse objetivo genérico havia, por exemplo, o interesse em robôs mais rápidos. Uma possível hierarquia é apresentada na Figura 2.14.

Salientamos, porém, que uma função-valor não precisa necessariamente ser determinada através de objetivos de baixo nível. Ilustraremos através de um exemplo inspirado pelo próprio Raiffa em [KR76]. Na Figura 2.15 encontramos dez objetivos de baixo nível ($Z_1, \dots, Z_{10}$). A partir disso, podemos modelar o problema através de uma função-valor $f(Z_1, \dots, Z_{10})$ —uma função-valor num espaço deca-dimensional, o que não é desejável. Porém, há a alternativa de quantificar preferências em níveis mais altos de agregação. Por exemplo, poderíamos trabalhar diretamente com os atributos X_1 e X_2 , onde X_1 é uma composição subjetivamente avaliada de Z_1 até Z_3 e

⁷Há muita discussão sobre o uso ou não da vírgula antes de “etc.”. Há os que o consideram como um elemento a mais da enumeração—logo adotam a vírgula. Há os que consideram seu sentido etimológico—do latim “et cetera” (e outras coisas)—logo dispensam a vírgula. Nessa dissertação optamos pela interpretação etimológica, ou seja, omitiremos a vírgula.

⁸Há problemas que envolvem tantos objetivos, que inviabilizam a determinação de uma função-valor.

Figura 2.14: Possível hierarquia para o robô-fugitivo no problema da perseguição-evasão.

X_2 de Z_4 até Z_{10} . Determinaríamos uma função-valor não mais a partir da tupla $(Z_1, \dots, Z_{10})$, mas através de (X_1, X_2) . Estamos quantificando nossas preferências no nível (X_1, X_2) . Justificasse, porém, descer na hierarquia até os níveis $(Z_1, \dots, Z_{10})$ porque a estrutura qualitativa dos objetivos associados com X_1 e X_2 nos ajuda a pensar mais claramente sobre X_1 e X_2 .

Figura 2.15: Hierarquia de objetivos abstratos.

Capítulo 3

Atributos e Escalas

3.1 Considerações Iniciais

O que queremos em alocação de tarefas em sistemas robóticos é decidir qual robô deve realizar qual tarefa. Para isso devemos comparar os robôs, se possível atribuindo valores escalares a eles. Esses valores visam representar uma ordem de preferência entre os robôs para uma dada tarefa. Para isso podemos usar funções-valor—quando o problema for determinístico ou funções utilidade—quando houver incerteza envolvida.

A utilidade foi inicialmente proposta por Bernoulli (1738), retomada em 1944 por von Neumann e Morgenstern com a teoria dos jogos e estendida por Keeney e Raifa [KR76] para casos com vários atributos. É estudada, entre outras, na teoria da utilidade, teoria da decisão, teoria dos jogos, economia e teoria da medição.

Valores e utilidades são valores escalares que devem expressar preferência relativa. O problema é como encontrar de forma inequívoca esses valores, ou seja, um conjunto de valores que expressem uma ordem de preferência que permita a tomada de decisão.

Porém, encontrar uma função-valor, se existir, nem sempre é trivial. Para a obtenção de uma função-valor alguns pontos devem ser observados. (1) Devemos inicialmente *escolher quais atributos são relevantes ao problema*. Embora seja de grande importância, a determinação de um conjunto de atributos relevantes, por si só, também não é tarefa fácil. Há sugestões na literatura sobre algumas propriedades que um conjunto de atributos deva possuir, como por exemplo as sugestões encontradas em [KR76], porém não há uma teoria normativa sobre como definir esse conjunto [Rob79]. A determinação desse conjunto de atributos, muitas vezes, envolve aspectos subjetivos¹ e até contraditórios (como por exemplo em um problema que exija menor custo e maior eficiência). Finalmente, como ilustrado em [KR76], um conjunto de atributos relevantes

¹ Adotaremos a seguinte interpretação para informação subjetiva: supomos que uma informação seja subjetiva caso possa ter diferentes—e igualmente aceitáveis—avaliações e/ou interpretações para diferentes receptores dessa mesma informação.

para um dado problema não é necessariamente único. (2) Devemos, então, *encontrar uma escala comum a todos os atributos*, pois geralmente os atributos estão em escalas diversas. Para isso precisamos, dentre outras coisas, levar em consideração aspectos pertinentes a cada escala, pois como veremos através da teoria da medição, algumas operações permitidas a determinadas escalas, perdem o sentido em outras. E finalmente (3) precisamos *determinar qual o impacto que cada atributo exerce na função* (quando pertinente) e (4) *determinar como eles se relacionam*.

Nas próximas seções iremos discorrer sobre esses quatro pontos, ilustrando-os através de simulações do problema da perseguição-evasão.

Na Seção 3.2 serão descritas sugestões sobre propriedades relevantes para um conjunto de atributos. Na Seção 3.3, através de conceitos da teoria da medição (tais como escala, transformação admissível e significância), citaremos alguns cuidados que devemos tomar quando trabalhamos numa dada escala. Finalmente a Seção 3.4 tratará de aspectos concernentes às escalas na modelagem de problemas de otimização. O objetivo é evitar perda de informação no uso da função-valor durante a alocação dos robôs.

3.2 Conjunto de Atributos Relevantes

A teoria da decisão costuma abster-se de definir atributo, pois não há uma definição universal para atributos (embora apresente exemplos e sinônimos para explicá-los). Uma característica consensual, no entanto, é que um atributo deva ser capaz de prover uma forma de avaliar os níveis de um objetivo². Alguns sinônimos citados para atributos são parâmetros de performance, fatores, características, propriedades e componentes [HY81].

Na formulação de uma função-valor uma das mais importantes etapas é a determinação do conjunto de atributos. Nesta seção estudaremos algumas propriedades importantes de um conjunto de atributos. Nos basearemos principalmente nas sugestões apresentadas por Raiffa em [KR76].

Como mencionamos na Seção 1.2, ilustraremos os problemas que podem surgir na determinação de funções-valor através de um simulador do problema da perseguição-evasão baseado em visibilidade.

Inicialmente adotaremos uma função-valor bastante simples. Os aspectos descritos em seções posteriores também podem ser fontes de problemas, mas iremos supor nesta seção uma situação quase ideal, ou seja, uma situação em que os atributos são as únicas fontes de problemas na determinação de uma função-valor.

O objetivo do fugitivo no problema da perseguição-evasão alterado é permanecer o maior tempo possível sem ser encontrado. Levando isso em conta, consideraremos como atributos relevantes os seguintes atributos: (1) a onisciência³ do robô-fugitivo em relação à posição do

²Um objetivo é algo a ser atingido, como maximizar a velocidade, minimizar o custo etc.

³No nosso caso um robô-fugitivo é onisciente se sabe onde se encontra seu perseguidor.

perseguidor, representada por $o(R)$, (2) a velocidade⁴ do robô, representada por $v(R)$ e (3) a autonomia⁵ do robô, representada por $a(R)$. Supomos, também, que esses três atributos formam um conjunto de atributos ideal para o problema em questão.

Como não dispomos ainda de qualquer ferramenta que nos auxilie na determinação da função-valor usaremos o bom senso: compararemos esses atributos entre si em situações extremas, com a intenção de classificá-los quanto à importância. Podemos supor, por exemplo, situações extremas para os atributos (velocidade igual a zero, não-onisciência e autonomia igual a zero). A Tabela 3.1 resume essas situações.

<i>Vel.</i>	<i>Oni.</i>	<i>Aut.</i>	<i>Situação extrema</i>	<i>Consequência extrema</i>
Ok	Ok	Ok		
Não	Ok	Ok	Fugitivo parado: $v(fugit) = 0$.	Encontrado rapidamente ($v(per)$).
Ok	Não	Ok	Fugitivo erra o caminho e vai de encontro ao perseguidor, pois não sabe onde ele está.	Encontrado muito rapidamente ($v(per) + v(fug)$).
Ok	Ok	Não	Fugitivo parado: $v(fug) = 0$ e $a(fug) = 0$.	Encontrado rapidamente ($v(per)$).
Não	Não	Ok	Fugitivo parado: $v(fug) = 0$.	Encontrado rapidamente ($v(per)$).
Ok	Não	Não	Fugitivo parado: $v(fug) = 0$.	Encontrado rapidamente ($v(per)$).
Não	Ok	Não	Fugitivo parado: $v(fug) = 0$ e $a(fug) = 0$.	Encontrado rapidamente ($v(per)$).
Não	Não	Não	Fugitivo parado: $v(fug) = 0$ e $a(fug) = 0$.	Encontrado rapidamente ($v(per)$).

Tabela 3.1: Determinação do grau de importância relativa dos atributos.

Ao analisarmos as diversas situações constatamos que a pior situação ocorre quando o fugitivo não possui onisciência e possui velocidade e autonomia diferentes de zero. O que poderá ocorrer neste caso (situação extrema) é ele ir de encontro ao perseguidor. Estará usando seus próprios atributos (velocidade e autonomia) para piorar a situação.

Usando essa constatação para determinar o grau de importância relativa dos atributos concluímos que a onisciência deve possuir peso maior que a velocidade e a autonomia e estas podem possuir o mesmo peso entre si. Ressalta-se que apenas utilizamos este método (uma análise superficial) por falta de outro melhor. Também não podemos afirmar que a relação que determinamos entre os pesos dos atributos seja a ideal, embora iremos supor que sim.

⁴Que para o caso discretizado será medida pelo número de passos por época.

⁵Tempo de duração da bateria do robô.

Portanto, vamos supor que a onisciência seja o mais importante dos três atributos. Além disso, vamos supor que ela seja um atributo mandatário, ou seja, criaremos uma função-valor em que a onisciência possua um peso tal, que mesmo que os outros dois atributos assumam valores máximos, não serão suficientes para superá-la.

Como veremos posteriormente, isto é possível se: (1) estivermos trabalhando com valores de atributos em escalas compatíveis, isto é, que permitam as operações utilizadas na função-valor que determinaremos e (2) normalizarmos os valores para, por exemplo, o intervalo $[0, 1]$. Esses conceitos serão detalhados na Seção 3.3. Por enquanto iremos supor que as escalas sejam compatíveis, bem como a normalização para o intervalo $[0, 1]$ seja um procedimento correto e o mais adequado para a situação.

Para essa ilustração foram considerados, então, quatro robôs A, B, C e D , que possuem as características descritas nas Tabelas 3.2 e 3.3. Para o atributo da onisciência, 1 significa que o robô é onisciente e 0 que não é. Note que a onisciência está representada por valores binários e os outros atributos não. O que à primeira vista parece ser um problema, ao contrário, facilitará a determinação da função-valor, pois: (1) definimos que a onisciência deverá superar a soma de todos os outros atributos, (2) a onisciência sempre possuirá valores binários (mesmo após a normalização), pois ela encontra-se no extremos do intervalo $[0, 1]$, (3) qualquer outro atributo, após a normalização para o intervalo $[0, 1]$, só assumirá um valor binário se inicialmente (antes da normalização) seu valor fosse zero ou tendesse a infinito, o que raramente ocorrerá. Logo, se desejarmos que a onisciência seja absoluta, basta que seu peso tenha valor igual ou maior à soma dos pesos dos outros atributos. Portanto, se compararmos um candidato a robô-fugitivo onisciente com qualquer candidato não-onisciente, em condições normais⁶, ele sempre vencerá a disputa (mesmo que ele possua valores mínimos para os outros atributos e os outros possuam valores máximos).

Uma função-valor que atende a esses preceitos é a Função 3.1.

$$f(R) = v(R) + 2o(R) + a(R). \quad (3.1)$$

Note que o peso atribuído à onisciência na Função 3.1 (neste caso 2) serve para torná-la absoluta (nunca superada pela soma dos outros atributos)⁷. Com isso, robôs não-oniscientes são eliminados, a não ser que todos os robôs sejam não-oniscientes.

Portanto os valores dos robôs, segundo a Função 3.1 e em relação à aos valores normalizados (Tabela 3.3) serão: $f(A) = 2.84$, $f(B) = 2.69$, $f(C) = 1.02$ e $f(D) = 2.59$.

3.2.1 Conjunto de Atributos Completo

Segundo Raiffa, um conjunto de atributos é completo se é adequado na indicação do grau de objetivo satisfeito. A seguir ilustraremos uma possível consequência para a alocação de tarefas,

⁶Se os outros atributos não possuírem valores iniciais nulos ou tendendo a infinito.

⁷O peso 2 nesse caso é suficiente: a onisciência sempre está em um dos extremos do intervalo $[0, 1]$ enquanto outros atributos raramente assumem esses valores.

<i>Robô (R)</i>	<i>Velocidade (v)</i>	<i>Onisciência (o)</i>	<i>Autonomia(a)</i>
<i>A</i>	20	1	55
<i>B</i>	10	1	55
<i>C</i>	40	0	45
<i>D</i>	10	1	45

Tabela 3.2: Atributos dos robôs.

<i>Robô (R)</i>	<i>Velocidade (v)</i>	<i>Onisciência (o)</i>	<i>Autonomia(a)</i>
<i>A</i>	0.29	1	0.55
<i>B</i>	0.14	1	0.55
<i>C</i>	0.57	0	0.45
<i>D</i>	0.14	1	0.45

Tabela 3.3: Atributos dos robôs—normalizados [0,1].

ao não escolhermos um conjunto de atributos completo—neste caso, a Função 3.2 representa a Função 3.1 alterada para que não leve em consideração a *velocidade*. O peso da onisciência foi alterado para mantê-lo como um atributo mandatório.

$$f(R) = o(R) + a(R) \quad (3.2)$$

Teremos, portanto, $f(A) = 1.55$, $f(B) = 1.55$, $f(C) = 0.45$ e $f(D) = 1.45$.

Como houve empate entre os robôs *A* e *B*, a escolha do robô *B* (com metade da velocidade do robô *A*) torna-se possível. Podemos simular essa situação com o fugitivo permanecendo parado uma a cada duas épocas.

Figura 3.1: Caso em que a função-valor é adequada ao problema. A região em que o perseguidor está é representada pela cor preta e a que o fugitivo está é representada pelo cinza escuro. As próximas regiões que podem ser visitadas pelo perseguidor são representadas pelo cinza claro. Assume-se que o perseguidor só consiga enxergar e capturar o fugitivo se estiver na mesma região (triângulo) que ele. Quando o fugitivo for encontrado (j), prevalece a cor preta—do perseguidor—para a região e a simulação é encerrada.

Na comparação, o fugitivo permaneceu mais tempo sem ser encontrado, quando possuía a mesma velocidade do perseguidor (Figura 3.1). Ao desconsiderarmos o atributo velocidade (relevante neste caso) podemos alocar um fugitivo mais lento, que pode ser encontrado mais facilmente (Figura 3.2).

Figura 3.2: Conjunto de atributos não-completo.

3.2.2 Conjunto de Atributos Operacional

Um conjunto de atributos é operacional se puder ser usado de forma significativa na tomada de decisão. Não queremos, por exemplo, que atributos irrelevantes causem distorção na função-valor.

Vamos supor que um atributo irrelevante para o problema seja incluído na função, por exemplo, um atributo referente à capacidade de carga do robô. Suponhamos que o robô D possua cinco vezes mais capacidade de carga que os outros, conforme as Tabelas 3.4 e 3.5.

<i>Robô (R)</i>	<i>Velocidade (v)</i>	<i>Onisciência (o)</i>	<i>Autonomia (a)</i>	<i>Capac. Carga (c)</i>
A	20	1	55	1
B	10	1	55	1
C	40	0	45	1
D	10	1	45	5

Tabela 3.4: Atributos dos robôs com um atributo irrelevante.

<i>Robô (R)</i>	<i>Velocidade (v)</i>	<i>Onisciência (o)</i>	<i>Autonomia (a)</i>	<i>Capac. Carga (c)</i>
A	0.29	1	0.55	0.17
B	0.14	1	0.55	0.17
C	0.57	0	0.45	0.17
D	0.14	1	0.45	0.83

Tabela 3.5: Atributos dos robôs com um atributo irrelevante—normalizados $[0,1]$.

Supondo que a função-valor seja dada pela Função 3.3 (novamente o peso do atributo da onisciência foi alterado para mantê-lo como um atributo mandatário).

$$f(R) = v(R) + 3o(R) + a(R) + c(R) \quad (3.3)$$

teremos $f(A) = 4.01$, $f(B) = 3.86$, $f(C) = 1.19$ e $f(D) = 4.42$. Portanto, o robô D é o escolhido, conforme ilustra a Figura 3.3.

3.2.3 Conjunto de Atributos Decomponível

Um problema envolvendo n atributos também envolve a formulação de uma função-valor com n atributos. Isto pode ser bastante complexo, muitas vezes inviável, a não ser que os atributos envolvidos sejam decomponíveis.

Um conjunto de atributos é decomponível se permite ser tratado em partes de menor dimensão, ou seja, se pudermos subdividir um problema com muitos atributos em vários sub-problemas

Figura 3.3: Atributo irrelevante é considerado.

com menos atributos cada um. Formalmente se o conjunto de atributos possuir tamanho n e pudermos subdividir o problema em k partes, com cada parte contendo $x_i < n$ atributos, com i variando de 1 a k .

Este conceito está relacionado com o conceito de independência entre os atributos (detalhado no Capítulo 4), que por sua vez está relacionado com a simplicidade (ou não) na formulação da função-valor.

3.2.4 Conjunto de Atributos Não-Redundante

Um conjunto de atributos é não-redundante se evita múltiplo impacto de um atributo no resultado. Ilustraremos eventuais consequências incluindo atributos redundantes na função-valor. Seja a Função 3.4 nossa função-valor para o problema (descrito na Seção 2.2):

$$f(R) = v_x(R) + v_y(R) + v_z(R) + p(R) + a(R) + o(R), \quad (3.4)$$

onde v_x é a velocidade do robô quando se desloca paralelo ao eixo x do plano cartesiano e similarmente para v_y e v_z (por simplificação, vamos supor que o deslocamento dos robôs seja paralelo aos eixos x e y). Equivocadamente a velocidade está representada de quatro formas distintas na função-valor—direta (velocidades) ou indiretamente (potência). Se considerarmos os robôs da Tabela 3.6, teremos $f(A) = 2.30$ $f(B) = 1.67$ $f(C) = 2.30$ e $f(D) = 1.77$. O robô C empata com o robô A e pode ser o escolhido. Suponhamos que seja esse o caso. Então a captura do robô C é imediata, conforme a Figura 3.4.

3.2.5 Conjunto de Atributos Minimal

Um conjunto de atributos é minimal se torna a dimensão do problema a menor possível, desde que preservando as propriedades anteriores.

Apenas a opinião de um autor (Raiffa) foi considerada na determinação das características relevantes de um conjunto de atributos. Provavelmente as características descritas nesta seção

<i>Robô (R)</i>	<i>Vel₁ (v_x)</i>	<i>Vel₂ (v_y)</i>	<i>Vel₃ (v_z)</i>	<i>Pot (p)</i>	<i>Onisc (o)</i>	<i>Auton (a)</i>
<i>A</i>	20	20	20	200	1	60
<i>B</i>	10	10	10	100	1	30
<i>C</i>	40	40	40	400	0	60
<i>D</i>	10	10	10	100	1	50

Tabela 3.6: Atributos redundantes.

<i>Robô (R)</i>	<i>Vel₁ (v_x)</i>	<i>Vel₂ (v_y)</i>	<i>Vel₃ (v_z)</i>	<i>Pot (p)</i>	<i>Onisc (o)</i>	<i>Auton (a)</i>
<i>A</i>	0.25	0.25	0.25	0.25	1	0.30
<i>B</i>	0.13	0.13	0.13	0.13	1	0.15
<i>C</i>	0.50	0.50	0.50	0.50	0	0.30
<i>D</i>	0.13	0.13	0.13	0.13	1	0.25

Tabela 3.7: Atributos redundantes—normalizados [0,1].

Figura 3.4: Atributo redundante.

não esgotem a questão. As características citadas, porém, são bastante adequadas para ilustrar alguns erros que podem ser cometidos na modelagem de problemas através de funções valor.

3.3 Teoria da Medição: Escalas

Nesta seção estudaremos formalmente as escalas, bem como uma classificação para elas. Uma função-valor pode possuir atributos que não apenas estão em escalas diferentes, mas em escalas que diferem quanto ao tipo. Nesse caso é interessante que a normalização desses atributos respeite algumas restrições, que dizem respeito às operações permitidas a cada um desses tipos de escalas. Outro ponto de interesse refere-se às transformações que são (ou não) admitidas para cada uma das classificações. Isso pode nos auxiliar no teste de viabilidade de determinadas escalas, quando comparadas com escalas naturais (se houver) para o mesmo fenômeno empírico, por exemplo, quando modelamos o problema matematicamente (programação linear, otimização combinatória etc.).

A seguir trataremos da teoria da medição e alguns de seus mais importantes conceitos, tais como representação, unicidade e significância. Então, estudaremos alguns resultados obtidos em [Pek96] sobre escalas e otimização.

A teoria da medição⁸ cria sistemas matemáticos que preservem certas relações observadas⁹. Isso pode ser conseguido, por exemplo, através da atribuição de números que reflitam um determinado fenômeno. A relação binária “mais alto que” em um grupo de pessoas e a atribuição de números a estas pessoas (que correspondam a esta relação) é um exemplo disso. Referências para a área são encontradas em [Rob79, Pfa71, LKST71, LKST89, LKST90, Nar85], mas nos basearemos principalmente em [Rob79]. A teoria da medição será útil na avaliação das escalas adotadas para cada atributo considerado. Isso tem impacto na estruturação da função-valor, pois teremos, por exemplo, que convergir todas as escalas para uma escala comum. Duas avaliações são necessárias: (1) avaliar se a escala adotada (inicial) para cada atributo é adequada, ou seja, se auxilia o atributo a ser compreensivo e operacional e (2) avaliar se as transformações para uma escala final são admissíveis. A seguir serão apresentadas a notação e algumas definições básicas da teoria da medição, de acordo com [Rob79].

Definição 3.1 (Sistema Relacional). *Um sistema relacional é a $(p + q + 1)$ -tupla*

$$\mathcal{U} = (A, R_1, R_2, \dots, R_p, o_1, o_2, \dots, o_q) \quad (3.5)$$

e $R_1, R_2, \dots, R_p$ são relações (não necessariamente binárias) em um conjunto A e $o_1, o_2, \dots, o_q$ são operações binárias em A .

Definição 3.2 (Tipo de Sistemas Relacionais). *Dado um sistema relacional*

$$\mathcal{U} = (A, R_1, R_2, \dots, R_p, o_1, o_2, \dots, o_q),$$

⁸Do inglês “Measurement Theory”.

⁹Chamaremos de sistema observável ou empírico a um sistema que contenha tais relações.

o tipo de $\mathcal{U}$ é uma seqüência $(r_1, r_2, \dots, r_p; q)$ de tamanho $p + 1$, onde r_i vale m se R_i é uma relação m -ária.

O sistema relacional $\mathcal{U} = (\mathbb{R}, >, +)$, por exemplo, tem tipo $(2; 1)$ e o sistema $\mathcal{U} = (\mathbb{R}, >, \leq, +, \times)$ tem tipo $(2, 2; 2)$.

Particularmente, classificaremos os sistemas relacionais em sistemas relacionais empíricos e sistemas relacionais numéricos. Por exemplo, seja $\mathcal{U} = (A, R_i, o_j)$, onde $i = 1 \dots p$ e $j = 1 \dots q$, um sistema relacional. Se $\mathcal{U}$ for um sistema relacional empírico, então A será um conjunto não vazio de objetos que chamaremos empíricos ou observáveis (os quais desejamos medir), R_i representará as p relações empíricas sobre A (“mais leve ou igual a”, “menor custo que”, “mais alto que” etc.) e o_j representará as q operações sobre os objetos de A (“concatenação”, “agrupamento” etc.). Por outro lado, se $\mathcal{U}$ for um sistema relacional numérico, então A deverá representar um conjunto não-vazio de objetos que chamaremos formais ou numéricos (os naturais, os reais, vetores etc.), R_i representará as p relações sobre A (tais como “menor que”, “maior ou igual a” etc.) e o_j representará as q operações binárias sobre A (tais como adição e multiplicação).

O que buscamos com a medição é um mapeamento de um sistema relacional empírico para um sistema relacional numérico que preserve todas as relações e operações do primeiro, conforme a Definição 3.3.

Definição 3.3 (Homomorfismo de Sistemas Relacionais). *Uma função $f : A \rightarrow B$ de $\mathcal{U} = (A, R_1, \dots, R_p, o_1, \dots, o_q)$ em $\mathcal{B} = (B, R'_1, \dots, R'_p, o'_1, \dots, o'_q)$, onde $\mathcal{U}$ e $\mathcal{B}$ têm o mesmo tipo, é chamada homomorfismo se para todo $a_1, \dots, a_{r_i} \in A$*

$$R_i(a_1, \dots, a_{r_i}) \iff R'_i(f(a_1), \dots, f(a_{r_i})) \text{ com } i = 1, \dots, p \quad (3.6)$$

e para todo $a, b \in A$,

$$f(a \circ_i b) = f(a) \circ'_i f(b), \text{ com } i = 1, \dots, q \quad (3.7)$$

Intuitivamente, um homomorfismo pode ser apresentado da seguinte forma. Suponha que saibamos efetuar uma operação (a soma, por exemplo) sobre um conjunto A (de bananas, por exemplo), mas que não saibamos efetuar essa operação sobre um conjunto B (de maçãs, por exemplo). Se houver uma função relacionando um conjunto ao outro que preserve essa operação, poderíamos levar as maçãs até o conjunto de bananas, efetuar a soma lá e devolver em maçãs a resposta [Cab91].

Por exemplo, se estivermos medindo massa, podemos escrever $\mathcal{U} = (A, H, o)$ e $\mathcal{B} = (\mathbb{R}, >, +)$ e então buscarmos um mapeamento de $\mathcal{U}$ em $\mathcal{B}$ que preserve a relação H (mais pesado que, por exemplo) e a operação o (aproximação de dois corpos, por exemplo).

Um problema básico para a teoria da medição é o problema da representação, onde dado um sistema relacional numérico particular $\mathcal{B}$, deve-se encontrar condições em um sistema relacional observável $\mathcal{U}$ necessárias e suficientes para a existência de um homomorfismo de $\mathcal{U}$ para $\mathcal{B}$. Esse conceito trata da existência ou não de escalas, logo para casos em que a escala já é conhecida, ele perde sua importância prática.

Para entendermos porque escalas e homomorfismos estão tão relacionados¹⁰ devemos primeiro entender o que é medição. Algumas definições de medição encontradas na literatura e citadas por Roberts em seu livro[Rob79] são transcritas na Tabela 3.8.

Russell (1938)	“Medição de grandezas é, no sentido mais geral, qualquer método pelo qual uma única e recíproca correspondência é estabelecida entre todas ou algumas das grandezas de um tipo e todos ou alguns dos números, inteiros, racionais ou reais, conforme o caso.”
Campbell (1938)	“Medição é a atribuição de números para representar propriedades de sistemas materiais (que não números) em virtude das leis que governam essas propriedades.”
Stevens (1951)	“Medição é a atribuição de números a objetos ou eventos de acordo com regras.”
Torgerson (1958)	“Medição de uma propriedade . . . envolve a atribuição de números a sistemas para representar aquela propriedade.”

Tabela 3.8: Definições de medição encontradas na literatura e citadas por Roberts em [Rob79].

Mesmo sem uma definição universal para medição, é consenso que medição envolve a atribuição de números que *correspondem a*, ou *representam*, ou *preservam* certas relações observadas. No caso da massa, a relação preservada é a relação *mais pesada que*.

O segundo problema básico é o problema da unicidade, que tenta determinar se um dado homomorfismo f é único. A unicidade é responsável por delimitar as operações matemáticas admitidas numa escala. Veremos a seguir que a unicidade é usada para caracterizar escalas. Antes, porém, citaremos algumas definições importantes ([Rob79]):

Definição 3.4 (Escala). *Uma escala é a tripla $(\mathcal{U}, \mathcal{B}, f)$, onde $\mathcal{U}$ e $\mathcal{B}$ são sistemas relacionais e f é um homomorfismo de $\mathcal{U}$ para $\mathcal{B}$.*

Quando não houver perda de informação chamaremos f de escala (ao invés de $(\mathcal{U}, \mathcal{B}, f)$, por exemplo).

Definição 3.5 (Escala Regular). *Suponha que $(\mathcal{U}, \mathcal{B}, f)$ denote qualquer escala e $(\mathcal{U}, \mathcal{B}, g)$ denote qualquer outra escala, então em alguns casos será possível mapear uma escala em outra, ou seja, será possível encontrar uma função $\phi : \mathcal{B} \rightarrow \mathcal{B}$ tal que $g = \phi \circ f$.*

Se para cada escala $(\mathcal{U}, \mathcal{B}, g)$ houver uma transformação $\phi : \mathcal{B} \rightarrow \mathcal{B}$ tal que $g = \phi \circ f$, então deveremos chamar a escala $(\mathcal{U}, \mathcal{B}, f)$ regular.

Para tornarmos esse conceito mais intuitivo, vamos supor que dois objetos quaisquer possuam o mesmo valor—para uma propriedade qualquer—em uma mesma escala. Se também possuírem o

¹⁰Veremos que, na verdade, escalas se confundem com homomorfismos.

mesmo valor—para a mesma propriedade—em quaisquer outras escalas, homomórficas à primeira, diremos que a primeira escala é regular.

Teorema 3.1 (Roberts [Rob79]). *Uma escala $(\mathcal{U}, \mathcal{B}, f)$ é regular se e somente se para cada outro homomorfismo g de $\mathcal{U}$ para $\mathcal{B}$ e para qualquer $a, b \in A$, $f(a) = f(b)$ implicar em $g(a) = g(b)$.*

Outros importantes conceitos envolvendo escalas são os conceitos de transformação admissível e significância [Rob79].

Definição 3.6 (Transformação Admissível). *Suponha que f seja um homomorfismo de um sistema relacional $\mathcal{U}$ em um sistema relacional $\mathcal{B}$, e suponha que A seja o conjunto subjacente a $\mathcal{U}$ e B seja o conjunto subjacente a $\mathcal{B}$. Seja ϕ uma função que mapeia o contra-domínio de f , o conjunto*

$$f(A) = \{f(a) : a \in A\}$$

no conjunto B . Então a composição $\phi \circ f$ é uma função de A em B . Se $\phi \circ f$ for um homomorfismo de $\mathcal{U}$ em $\mathcal{B}$, chamamos ϕ uma transformação admissível de escala.

Trata-se da transformação de uma escala regular em outra, como por exemplo a transformação admissível de graus Célsius em graus Fahrenheit. Nesse caso teríamos $\phi(x) = (9/5)x + 32$.

Definição 3.7 (Significância). *Uma declaração envolvendo escalas numéricas é significativa se e somente se seu valor verdade for invariante sob todas as transformações admissíveis.*

O conceito de escala regular desempenha um importante papel para o conceito de significância, pois este último pressupõe escalas regulares.

Stevens propõe uma classificação de escalas a partir de transformações admissíveis [Ste46]. Essa classificação não pretende ser completa, mas é atual e bastante utilizada. Cinco escalas são descritas:

A *escala nominal* que apenas classifica os dados, como em masculino(M) e feminino(F) ou em sangue tipo A^- , B^+ etc. Note que essas letras e números podem ser trocados sem ressalvas, pois não desempenham o papel de ordenar os dados. Não faz sentido dizermos que $M < F$ ou que $A^- > B^+$, no caso do tipo sanguíneo. Podemos, no entanto, dizer se alguém pertence a M ou a F. Essa escala permite, portanto, qualquer função ϕ injetora ou transformação uma para um, ou seja, cada elemento do contra-domínio está associado a apenas um elemento do domínio. Um exemplo é $f(n) = 2n$, $n \in \mathbb{Z}$.

A *escala ordinal* que nos permite além de classificar os dados, também ordená-los. Não faz sentido, porém, usar essa escala com intuito de quantificar os dados. Como exemplos temos as notas no mestrado (A, B, C e R), bem como a classificação de satisfação do cliente (1-ruim, 5-médio, 10-ótimo). Podemos dizer que $10 > 5$ ou que $A > B$, mas não são definidas operações aritméticas envolvendo A e B ou 5 e 10, por exemplo. Este tipo de escala permite transformações monotonicamente crescentes ($x \geq y \iff \phi(x) \geq \phi(y)$).

O terceiro tipo de escala é a *escala intervalo* que classifica, ordena e quantifica, mas não possui um “zero natural”, ou seja, nela o zero é arbitrário. Portanto, podemos somar e subtrair, mas não faz sentido a divisão (proporção), por exemplo. Datas, temperaturas em graus Célsius e temperaturas em graus Fahrenheit são exemplos para esta escala. Note que se ontem a temperatura estava 10°C e se hoje está 20°C , isto significa que hoje está 10°C mais quente que ontem, mas não significa que hoje está duas vezes mais quente. A escala intervalo admite qualquer transformação linear positiva, ou seja, $\phi(x) = \alpha x + \beta$, $\alpha > 0$.

A penúltima escala, segundo a classificação de Stevens, é a *escala de razão* que admite além do que já era admitido na escala intervalo, também o conceito de proporcionalidade, pois possui o “zero natural”. Exemplos desse tipo de escala são: temperatura em Kelvin, intervalo de tempo, massa e comprimento etc. Podemos, portanto, afirmar que alguém mede o dobro de outra pessoa. A escala de razão admite quaisquer transformações lineares de similaridade, ou seja, $\phi(x) = \alpha x$, $\alpha > 0$.

Finalmente temos a escala mais restritiva quanto à transformação: a *escala absoluta*. Probabilidade e contagem são exemplos desta escala, que só admite a transformação identidade ($\phi(x) = x$). Note que numa escala absoluta as propriedades dos números refletem por completo as propriedades do atributo. Como exemplos temos, por exemplo, o número de filhos de um casal. Resumidamente, temos as escalas apresentadas na Tabela 3.9.

	<i>Propriedades</i>	<i>Operações</i>	<i>Transformação</i>
<i>Nominal</i>	Classificação	Teste de pertinência etc.	um-para-um
<i>Ordinal</i>	Classif. e ordenação	$>, \geq, <, \leq, =$	$x \geq y \iff \phi(x) \geq \phi(y)$
<i>Intervalo</i>	Classif., orden. e quantif.	$>, \geq, <, \leq, =, +, -$	$\phi(x) = \alpha x + \beta, \alpha > 0$
<i>Razão</i>	Classif., orden. e quantif.	$>, \geq, <, \leq, =, +, -, \times, /$	$\phi(x) = \alpha x, \alpha > 0$
<i>Absoluta</i>	Classif., orden. e quantif.	$>, \geq, <, \leq, =, +, -, \times, /$	$\phi(x) = x$

Tabela 3.9: Tipos de escalas segundo Stevens.

A seguir utilizaremos esse conhecimento para nos auxiliar na escolha de atributos (compreensivos e operacionais) através de um estudo de caso.

Definição 3.8 (Escala Mais Forte). Dizemos que uma escala f é mais forte que uma escala g se $AT(f) \subseteq AT(g)$, onde $AT(.)$ é o conjunto de todas as transformações admissíveis de uma dada escala.

Se uma dada escala f é mais forte que outra escala g , então podemos interpretar que a escala f é mais informativa que a escala g . Escalas de razão e intervalo, por exemplo, contêm mais informação que escalas ordinais.

A seguir veremos formalmente três problemas básicos de representação—medição ordinal, medição de diferenças e medição extensiva—indicando as condições pelas quais a medição pode

ser executada. A primeira será útil na determinação de funções-valor ordinais e as outras duas na determinação de funções-valor cardinais.

3.3.1 Medição Ordinal

Nesta seção estudaremos uma formalização para o primeiro problema de representação—a medição ordinal—que está relacionada com escalas ordinais. Escalas ordinais são capazes de classificar e ordenar, mas não conseguem quantificar. Exemplos de escalas ordinais são as notas no mestrado (A, B, C e R) e a classificação de satisfação do cliente (1-ruim, 5-médio, 10-ótimo).

Como dissemos anteriormente, o problema da representação trata da existência ou não de escalas—nesse caso a existência ou não de uma escala ordinal. E será através da medição ordinal que chegaremos às funções-valor ordinais. Também estudaremos formalmente o problema da unicidade, que como comentamos anteriormente, é responsável por delimitar as operações matemáticas admitidas numa escala—nesse caso as operações permitidas às escalas ordinais.

Representação

O problema da representação pode ser formulado como: dado um sistema relacional numérico particular $\mathcal{B}$, devemos encontrar condições em um sistema relacional observável $\mathcal{U}$ necessárias e suficientes para a existência de um homomorfismo de $\mathcal{U}$ para $\mathcal{B}$.

A seguir, apresentaremos formalmente o problema da representação *ordinal*.

O Teorema 3.2 discute o caso finito (nosso interesse).

Teorema 3.2 (Roberts [Rob79]). *Suponha um conjunto finito A e uma relação binária R sobre A . Então existe uma função f real, sobre A que satisfaz*

$$aRb \iff f(a) > f(b) \quad (3.8)$$

se e somente se (A, R) for uma ordem fraca estrita¹¹. Uma ordem fraca estrita é uma relação binária R sobre um conjunto A que é assimétrica e negativamente transitiva¹².

Se R é preferência, então à função f denominamos função-valor ordinal. Logo, para produzirmos uma função-valor ordinal através de preferências, devemos checar se estas preferências satisfazem as condições para uma ordem fraca estrita. Isso pode ser obtido através de comparações par a par¹³.

¹¹Do inglês “strict weak order”.

¹²Assimétrica: $(\forall a, b \in A)(aRb \Rightarrow \neg bRa)$. Negativ. transitiva: $(\forall a, b, c \in A)(\neg aRb \wedge \neg bRc \Rightarrow \neg aRc)$.

¹³A teoria da decisão com frequência faz referência a um “tomador de decisão” que, dados dois objetos a e b , deve escolher se prefere a , b ou se é indiferente. A esse tipo de comparação denomina-se comparação par a par (pairwise comparison).

Unicidade

Agora formalizaremos o problema da unicidade para a representação ordinal 3.8. Com isso delimitaremos a classe de operações admissíveis (Definição 3.6) para escalas ordinais.

Segundo Hays, sempre poderemos efetuar operações sobre números. A questão é *se após efetuarmos tais operações ainda poderemos deduzir resultados significativos sobre os objetos que estão sendo medidos* [Rob79].

Teorema 3.3 ([Rob79]). *Suponha um conjunto finito A , uma relação binária R sobre A e f uma função real sobre A satisfazendo 3.8, então $\mathcal{U} = (A, R) \rightarrow \mathcal{B} = (\mathbb{R}, >)$ é uma representação regular e $(\mathcal{U}, \mathcal{B}, f)$ é uma escala ordinal.*

Ou seja, escalas ordinais têm como transformações admissíveis a classe das funções monotonicamente crescentes ($x \geq y \iff \phi(x) \geq \phi(y)$). Podemos comparar e classificar os dados. Não faz sentido, porém, efetuarmos operações aritméticas sobre eles. No caso da classificação de satisfação do cliente onde 1 significa pouco satisfeito, 5 significa medianamente satisfeito e 10 significa muito satisfeito, qual o significado de $10 + 5 + 1$, por exemplo?

3.3.2 Medição de Diferenças

Como veremos a seguir, a medição de diferenças está relacionada com escalas intervalo. Como vimos, esse tipo de escalas além de classificar e ordenar, também é capaz de quantificar. Exemplos clássicos de escalas intervalo são a temperatura em graus Célsius e Fahrenheit, as datas e os testes de inteligência.

A seguir trataremos as escalas intervalo do ponto de vista formal, ou seja, o problema da representação e da unicidade (da medição) de diferenças. Apresentaremos o problema da representação de diferenças (ou medição algébrica de diferenças), cinco axiomas que formam a base da definição da estrutura algébrica de diferenças. Quatro desses axiomas serão necessários para o problema da representação e todos eles serão necessários para o problema da unicidade.

Representação

Seja A um conjunto e seja D uma relação $D(a, b, s, t)$ ou equivalentemente $abDst$ sobre A .

$$abDst \iff f(a) - f(b) > f(s) - f(t). \quad (3.9)$$

Se estivermos falando de preferências, podemos interpretar $abDst$ como *preferir (mais) a sobre b a s sobre t* . No caso de temperaturas, denota que a diferença entre as temperaturas de a e b é maior que a diferença de s e t através da Equação 3.9. A essa representação chamamos medição algébrica de diferenças.

Para tratarmos o problema da representação de diferenças, adotaremos a notação descrita nas Equações 3.10 e 3.11.

$$abEst \iff [\neg abDst \text{ e } \neg stDab]. \quad (3.10)$$

$$abWst \iff [abDst \text{ ou } abEst]. \quad (3.11)$$

A seguir apresentaremos a estrutura algébrica de diferenças através de cinco axiomas (Definição 3.9). Também apresentaremos a prova da necessidade dos três primeiros.

Definição 3.9 (Estrutura Algébrica de Diferenças [Rob79]).

Axioma 3.1. *Suponha que R seja definido sobre $A \times A$ por $(a, b)R(s, t) \iff abDst$. Então $(A \times A, R)$ é fraca estrita¹⁴.*

Seja $g : A \times A \rightarrow \mathbb{R}$ e $g(a, b) = f(a) - f(b)$. Então $(a, b)R(s, t) \iff g(a, b) > g(s, t)$. $\square$

Axioma 3.2. *Para todo $a, b, s, t \in A$, se $abDst$, então $tsDba$.*

Demonstração. Suponha que vale $abDst$. Então $f(a) - f(b) > f(s) - f(t)$, que pode ser reescrita como $f(t) - f(s) > f(b) - f(a)$, equivalente a $tsDba$. $\square$

Axioma 3.3. *Para todo $a, b, c, a', b', c' \in A$, se $abWa'b'$ e $bcWb'c'$, então $acWa'c'$.*

Demonstração. Suponha que valem $abWa'b'$ e $bcWb'c'$, o que equivale a $f(a) - f(b) \geq f(a') - f(b')$ e a $f(b) - f(c) \geq f(b') - f(c')$, respectivamente (Equação 3.11). Compondo essas inequações, teremos $f(a) - f(b) + f(b) - f(c) \geq f(a') - f(b') + f(b') - f(c')$, ou seja, $f(a) - f(c) \geq f(a') - f(c')$. $\square$

Axioma 3.4. *Para todo $a, b, s, t \in A$, se valem $abWst$ e $stWxx$, então há $u, v \in A$ tal que $auEst$ e $vbEst$ ¹⁵.*

Axioma 3.5. *Cada seqüência padrão estritamente limitada é finita.*

Teorema 3.4 (Krantz. et. al. 1971). *Se (A, D) é uma estrutura algébrica de diferenças, então há uma função real f sobre A tal que para todo $a, b, s, t \in A$,*

$$abDst \iff f(a) - f(b) > f(s) - f(t). \quad (3.12)$$

¹⁴Ou seja, assimétrica: $(\forall a, b \in A)(aRb \Rightarrow \neg bRa)$ e negativ. transitiva: $(\forall a, b, c \in A)(\neg aRb \wedge \neg bRc \Rightarrow \neg aRc)$.

¹⁵Este axioma não é necessário para a representação, mas para a unicidade.

Unicidade

A seguir formalizaremos o problema da unicidade para a representação de diferenças 3.8 com o intuito de delimitar a classe de operações admissíveis (Definição 3.6) para escalas intervalo.

Teorema 3.5. *Suponha um conjunto A não vazio, uma relação D sobre A e uma função real f sobre A satisfazendo*

$$abDst \iff f(a) - f(b) > f(s) - f(t) \quad (3.13)$$

É comum ouvirmos declarações não significativas envolvendo escalas intervalo (principalmente sobre temperaturas). A intuição falha porque escalas intervalo não possuem um “zero natural”, ou seja, nelas o zero é arbitrário. Isto significa que se na semana passada estava 30°C e hoje está 15°C , não faz sentido afirmar que hoje está duas vezes mais frio¹⁶ (afirmação não significativa).

3.3.3 Medição Extensiva

O conceito de medição extensiva relaciona-se com escalas de razão. Como vimos esse tipo de escala, além de classificar, ordenar e quantificar, também admite o conceito de proporcionalidade. A medição extensiva é o modelo clássico para diversos atributos físicos, tais como massa e tamanho, temperatura em Kelvin e intervalo de tempo.

Como a medição extensiva e a álgebra abstrata possuem muitos pontos em comum, revisaremos o conceito de grupos e a propriedade arquimediana para, então, tratarmos o problema da representação e da unicidade (da medição) extensiva.

A seguir veremos o problema da medição extensiva.

Problema da Medição Extensiva

Dado um sistema relacional $(A, R, \circ)$, onde R é uma relação binária sobre A e $\circ$ é uma operação binária, queremos encontrar uma função f que satisfaça:

$$aRb \iff f(a) > f(b) \quad (3.14)$$

e

$$f(a \circ b) = f(a) + f(b). \quad (3.15)$$

Trata-se do *problema da medição extensiva*, ou seja, o problema de se encontrar condições sobre $(A, R, \circ)$, necessárias e suficientes, para a existência de um mapeamento homomórfico em $(\mathbb{R}, >, +)$.

Em seguida revisaremos o conceito de grupo da álgebra abstrata, pois há muitos pontos em comum entre grupos e a medição extensiva.

¹⁶É significativo, porém, afirmar que faz 15°C a menos.

Definição 3.10 (Grupo $(A, \circ)$). Chamamos de grupo o conjunto não-vazio A , fechado sob uma operação binária $\circ$, para o qual valem os axiomas:

Axioma 3.6. Associatividade de $\circ$: Para quaisquer $a, b, c \in A$ temos

$$(a \circ b) \circ c = a \circ (b \circ c).$$

Axioma 3.7. Identidade e para $\circ$: Existe $e \in A$, tal que para qualquer $a \in A$ temos

$$e \circ a = a \circ e = a.$$

Axioma 3.8. Inverso a^{-1} de a : Para cada $a \in A$, existe um elemento $a^{-1} \in A$ tal que

$$a \circ a^{-1} = a^{-1} \circ a = e.$$

A medição extensiva também usa uma importante propriedade dos reais, a propriedade arquimediana [Rob79].

Definição 3.11 (Propriedade Arquimediana).

$$(\forall a > 0)(\forall b \in \mathbb{R})(\exists n \in \mathbb{Z}^+) \mid na > b.$$

Ou seja, se a e b são números reais e $a > 0$, então há um inteiro positivo n tal que $na > b$. Não importa se a é muito pequeno e b muito grande: se $a > 0$, um número de cópias suficiente de a superará b . Essa propriedade é de grande importância para a medição, pois nos permite efetuar comparações aproximadas entre magnitudes relativas de quaisquer duas quantidades a e b (com $b > a$), ao determinarmos quantas cópias de a são necessárias para obtermos b .

Precisamos, porém, traduzir essa propriedade dos reais, para um axioma arquimediano em $(A, R, \circ)$, possibilitando axiomatizar a representação $(A, R, \circ)$ em $(\mathbb{R}, >, +)$. Isso se dará através da Definição 3.12 (grupo arquimediano ordenado¹⁷) e do Teorema 3.6 (teorema de Hölder) [Rob79].

Definição 3.12 (Grupo Ordenado Arquimediano). Um sistema relacional $(A, R, \circ)$ é um grupo ordenado arquimediano se satisfizer os seguintes axiomas:

Axioma 3.9. $(A, \circ)$ é um grupo.

Axioma 3.10. Monotonicidade: Para quaisquer $a, b, c \in A$ temos

$$(aRb) \iff (a \circ c)R(b \circ c) \iff (c \circ a)R(c \circ b).$$

Axioma 3.11. (A, R) é uma ordem simples estrita¹⁸.

¹⁷Do inglês “Archimedean Ordered Group”.

¹⁸Uma ordem estrita simples é assimétrica: $(\forall a, b \in A)(aRb \Rightarrow \neg bRa)$, transitiva: $(\forall a, b, c \in A)(aRb \wedge bRc \Rightarrow aRc)$ e completa: $(\forall a, b, c \in A)(aRb \vee bRa)$.

Axioma 3.12. *Arquimediana:* $(\forall a, b \in A) aRe \Rightarrow (\exists n \in \mathbb{Z}^+) \mid naRb$, onde e é a identidade em $(A, \circ)$.

E o Teorema de Hölder [Rob79].

Teorema 3.6 (Hölder). *Cada grupo ordenado arquimediano (Definição 3.12) é homomórfico a $(\mathbb{R}, >, +)$.*

Testaremos esses axiomas com alguns exemplos [Rob79]. No caso da massa, por exemplo, se A for um conjunto de objetos, aRb é interpretado como “ a é mais pesado que b ” e $a \circ b$ é interpretado como a combinação dos dois objetos.

O Axioma 3.9 requer que $(A, \circ)$ seja um grupo. Para que isso seja possível, testaremos a associatividade, a identidade e o inverso. A associatividade parece fazer sentido nesse caso, pois a combinação de dois objetos $a \circ b$ com um terceiro objeto c resulta igual a combinação de $b \circ c$ com o objeto a (em relação a massa)¹⁹. Também é possível, ao menos idealmente, imaginar um objeto sem massa que sirva de identidade (e). Não faz sentido, porém, falarmos em inverso. Não há objeto b que se combinado com o objeto a resulte na identidade e , um objeto sem massa. Logo para chegarmos a um teorema da representação extensiva, deveremos modificar a Definição 3.12.

O Axioma 3.10 requer monotonicidade. Então, se julgarmos que a é mais pesado que b , a adição (ou retirada) de um mesmo objeto c ao objeto a e ao objeto b não modificará nosso julgamento. A ordem em que isso é feito também não importa. A monotonicidade, portanto, parece correta.

O Axioma 3.11 requer que (A, R) seja uma ordem estrita simples. A assimetria parece razoável, visto que se um objeto a for mais pesado que outro objeto b isso implica que o b não seja mais pesado que a . A transitividade também parece correta: se um objeto a for mais pesado que b e se b for mais pesado que c , podemos concluir que o objeto a é mais pesado que o objeto c . A completude é a única condição que pode causar problemas. Dois objetos distintos podem possuir massa tão próxima que talvez não possamos distingui-los.

O Axioma 3.12 requer a propriedade arquimediana, ou seja, dado um objeto a , mais pesado que o objeto sem massa (a identidade e), e um outro objeto b , se combinarmos suficientes cópias do objeto a criaremos um objeto mais pesado que b . Esse axioma não é passível de ser testado na prática, no sentido que qualquer experimento finito não poderia reunir dados suficientes para refutá-lo.

No caso da preferência A denota um conjunto de objetos ou alternativas, $\circ$ denota combinação, e aRb denota que “ a é estritamente preferível a b ”. A associatividade será refutada se houver interação física: suponha que a represente um chama, b represente uma roupa e c represente água. Combinar a e b primeiro e em seguida combinar c difere bastante de combinar b e c primeiro e só então combinar a .

¹⁹Se $(A, \circ)$ é um grupo, associações tais como $a \circ (a \circ a)$ devem ser válidas. Para isso iremos supor que o conjunto A contém infinitas cópias de cada objeto.

Quanto ao inverso podemos supor que o oposto de ter um objeto é devê-lo a alguém. Se tenho um objeto a e devo um objeto a , o resultado é nulo.

Pode-se questionar, porém, o fato da preferência ser uma ordem estrita simples. Suponha que dois importantes atributos para os cliente de um dada região, são—em ordem de importância—o preço e a qualidade. Ou seja, se uma dada mercadoria a for mais barata que outra equivalente b , a primeira será a escolhida mesmo que a mercadoria b seja melhor. Suponha, então, a seguinte situação: uma loja dessa região possui três mercadorias equivalentes a , b e c . A mercadoria a é muito mais barata que b e um pouco mais barata que c e a mercadoria b equivale a a em qualidade, mas a mercadoria c é muito melhor que a , conforme a Tabela 3.10. Provavelmente, nesse caso, um cliente preferirá c a a .

	a	b	c
<i>Preço (em R\$)</i>	5000	9800	5002
<i>Índice de qualidade (intervalo: 0 a 10)</i>	3	3	10

Tabela 3.10: Preferência viola ordem simples estrita.

A completude, a monotonicidade e o axioma arquimediano também podem ser questionados, ao lidarmos com preferência.

A completude, porque é possível que sejamos indiferentes entre duas alternativas.

Em relação à monotonicidade, considere o seguinte exemplo: entre café e barra de chocolate, posso preferir a barra de chocolate. Agora considere a combinação de ambos com açúcar. O café com açúcar parece preferível à barra de chocolate com açúcar.

Finalmente, também é possível refutar o axioma arquimediano: é provável que uma pessoa mentalmente saudável não aceite trocar sua vida ou sua saúde, por qualquer quantidade de dinheiro, por maior que seja.

Esses exemplos ilustram a necessidade de se alterar a Definição 3.12 para a determinação da representação extensiva, o que será feito a seguir.

Representação

Para determinarmos um conjunto de axiomas necessários e suficientes para a medição extensiva, consideremos, por exemplo, um homomorfismo f de $(A, R, \circ)$ para $(\mathbb{R}, >, +)$ e os axiomas da Definição 3.12. Novamente usaremos os exemplos do Roberts [Rob79]. Podemos afirmar que

$$[f(a) + f(b)] + f(c) = f(a) + [f(b) + f(c)],$$

mas a representação não implica em

$$(a \circ b) \circ c = a \circ (b \circ c),$$

mas somente que

$$[(a \circ b) \circ c]E[a \circ (b \circ c)],$$

onde E é definido por

$$xEy \iff \neg xRy \wedge \neg yRx.$$

A identidade e não é necessária para a representação, desde que não haja nenhum elemento e em A , tal que $f(e) = 0$. O mesmo em relação ao inverso.

Também (A, R) não precisará ser uma ordem simples estrita, pois de acordo com o Teorema 3.2 desde que $aRb \iff f(a) > f(b)$, então (A, R) será fraca estrita.

A monotonicidade é necessária, pois

$$aRb \iff f(a) > f(b) \iff f(a) + f(c) > f(b) + f(c) \iff (a \circ c)R(b \circ c).$$

Finalmente o axioma arquimediano não será necessário como foi colocado, já que não é necessária uma identidade e em A . A premissa aRe equivalente a dizer que $f(a)$ é positivo. Mas isso é o mesmo que dizer que $2f(a) > f(a)$, e desde que f seja aditivo e preserve R , é o mesmo que dizer que $2aRa$. Então, o axioma arquimediano pode ser reescrito da seguinte maneira: se a e b estão em A e $2aRa$, então existe um inteiro positivo n , tal que $naRb$.

Há um caso, porém, que invalida esse axioma da forma que está. Suponha que A denote os reais negativos suplementados com um elemento $-\infty$, que R denote $>$ e $\circ$ denote $+$. Mas $(-\infty) \circ x$ e $x \circ (-\infty)$ resultam sempre $-\infty$.

Então, as condições necessárias e suficientes para a medição extensiva são dadas pelos axiomas abaixo.

Axioma 3.13. *Associatividade fraca:* $[(a \circ (b \circ c))E[(a \circ b) \circ c]]$.

Axioma 3.14. (A, R) é uma ordem fraca estrita.

Axioma 3.15. *Monotonicidade:* $(\forall a, b, c \in A)[(aRb) \iff (a \circ c)R(b \circ c) \iff (c \circ a)R(c \circ b)]$.

Axioma 3.16. *Arquimediana:* Se aRb , então $(\forall a, b, c, d \in A)(\exists n \in \mathbb{Z}^+) \mid (na \circ c)R(nb \circ d)$.

O axioma 3.16 é arquimediano e é necessário. Note que aRb implica em $f(a) > f(b)$ ou $f(a) - f(b) > 0$. Então, pela propriedade arquimediana dos reais, existe um inteiro positivo n , tal que $n[f(a) - f(b)] > f(d) - f(c)$, ou seja, $nf(a) + f(c) > nf(b) + f(d)$. Desde que f seja aditiva, segue que $f(na \circ c) > f(nb \circ d)$. Um sistema $(A, R, \circ)$ que satisfaça os axiomas 3.13, 3.14, 3.15 e 3.16 é chamado de estrutura extensiva.

Teorema 3.7 (Roberts e Luce). *Suponha A um conjunto, R uma relação binária sobre A e $\circ$ uma operação binária sobre A . Então há uma função real sobre A que satisfaz as equações*

$$aRb \iff f(a) > f(b) \tag{3.16}$$

e

$$f(a \circ b) = f(a) + f(b), \quad (3.17)$$

se e somente se $(A, R, \circ)$ for uma estrutura extensiva.

Algumas dúvidas que tínhamos na aplicação da medição extensiva para preferência persistem (associatividade e monotonicidade) e mesmo o axioma arquimadiano pode ser questionado. Suponha que a represente um moeda de ouro, b represente nenhum dinheiro, c represente ter uma grave enfermidade por toda a vida e d represente ter uma vida saudável. Provavelmente não haja n tal que $(na \circ c)R(nb \circ d)$.

O indicado é a aplicação da medição extensiva para fenômenos físicos, embora em alguns casos seja possível tal aplicação—com certa dificuldade—para casos envolvendo preferência [Rob79].

Unicidade

Prova-se que uma representação para medição extensiva é única sob transformações similaridade, ou seja, a medição deverá se dar sob uma escala de razão (ou mais forte, Definição 3.8), pois se uma função-valor satisfaz as equações 3.16 e 3.17, é significativo dizer que o valor de uma alternativa é metade do outro, o triplo do outro etc. Portanto, iniciamos o uso de funções-valor para tomada de decisão quantitativa e não apenas qualitativa.

Teorema 3.8 (Roberts). *Suponha A um conjunto não-vazio, R uma relação binária sobre A e $\circ$ uma operação binária sobre A e uma função real f sobre A que satisfaz as equações 3.16 e 3.17.*

Então $\mathcal{U} = (A, R, \circ) \rightarrow \mathcal{B} = (\mathbb{R}, >, +)$ é uma representação regular e $(\mathcal{U}, \mathcal{B}, f)$ é uma escala de razão.

3.3.4 Medição Conjunta

A medição conjunta²⁰ estuda modelos para casos onde os objetos são avaliados de acordo com várias de suas dimensões, ou seja, quando $\succeq$ for definida sobre um produto de conjuntos.

Portanto, ela tenta encontrar as condições necessárias para a obtenção de uma representação numérica de $\succeq$ e tenta estabelecer a unicidade dessa representação. Tem como modelo central a função-valor aditiva.

Se $(A, R) = (A_1 \times A_2, R)$ é uma relação binária, podemos dizer que é uma estrutura conjunta aditiva se satisfizer os axiomas [Rob79]:

Axioma 3.17. (A, R) é uma ordem fraca estrita²¹.

²⁰Do inglês “conjoint measurement”.

²¹Ou seja, assimétrica: $(\forall a, b \in A)(aRb \Rightarrow \neg bRa)$ e negativ. transitiva: $(\forall a, b, c \in A)(\neg aRb \wedge \neg bRc \Rightarrow \neg aRc)$.

Axioma 3.18. (A, R) satisfaz independência. Dizemos que uma relação binária $(A_1 \times A_2, R)$ satisfaz independência se satisfaz independência sobre o primeiro componente

$$(\forall x, y \in A_1)(\forall q, r \in A_2)[(x, q)R(y, q) \iff (x, r)R(y, r)]$$

e sobre o segundo componente

$$(\forall x, y \in A_1)(\forall q, r \in A_2)[(x, q)R(x, r) \iff (y, q)R(y, r)].$$

Axioma 3.19. (A, R) satisfaz a condição de Thomsen:

$$(\forall x, y, z \in A_1)(\forall q, r, s \in A_2)[(x, s)E(z, r) \wedge (z, q)E(y, s) \Rightarrow (x, q)E(y, r)].$$

Axioma 3.20. Cada seqüência padrão estritamente limitada em cada componente é finita.

Axioma 3.21. Vale a solucionabilidade restrita²² sobre cada componente, onde a solucionabilidade restrita sobre o primeiro componente, por exemplo, possui a seguinte forma:

$$(x \in A_1)(q, r \in A_2)(\exists y)[(x, q)E(y, r)].$$

Axioma 3.22. Cada componente é essencial, ou seja, dada uma ordem independente fraca estrita $(A_1 \times A_2, R)$, dizemos que o i -ésimo componente ($i = 1, 2$) é essencial se há $x, y \in A_i$ tal que $\neg xE_i y$.

Condições necessárias e suficientes para a medição extensiva aditiva, no caso em que A_i é finito, aplicado à relação binária de preferência ($aSb \iff \neg bRa$):

Axioma 3.23. $(\forall a_1, b_1 \in A_1)(\forall a_2, b_2 \in A_2)[(a_1, a_2)S(b_1, b_2) \vee (b_1, b_2)S(a_1, a_2)]$.

Axioma 3.24. Suponha $x_0, \dots, x_{k-1} \in A_1$ e $y_0, \dots, y_{k-1} \in A_2$ e suponha que π e σ são permutações sobre $\{0, 1, \dots, k-1\}$:

$$(\forall i = 1, 2, \dots, k-1)[(x_i, y_i)S(X_{\pi(i)}, Y_{\sigma(i)}) \Rightarrow (X_{\pi(0)}, Y_{\sigma(0)})S(x_0, y_0)]$$

Teorema 3.9 (Scott). Suponha que A_1 e A_2 sejam conjuntos finitos e que R seja uma relação binária sobre $A_1 \times A_2$. Então os dois axiomas acima serão necessários e suficientes para a existência de funções reais f_1 sobre A_1 e f_2 sobre A_2 , tal que

$$(\forall (a_1, a_2), (b_1, b_2) \in A_1 \times A_2)[(a_1, a_2)R(b_1, b_2) \iff f_1(a_1) + f_2(a_2) > f_1(b_1) + f_2(b_2)]. \quad (3.18)$$

²²Do inglês "restricted solvability".

3.3.5 Normalização

Os valores dos atributos deverão estar em escalas comparáveis. Caso não estejam, precisaremos normalizar estes valores.

Dependendo do modelo adotado, a normalização pode ser um estágio separado ou pode fazer parte do processo de determinação dos pesos.

Escolhido o grupo de atributos que melhor representará os robôs quanto à tarefa, devemos convergir esses atributos para uma escala comum. Não podemos nos esquecer, no entanto, de convergir as escalas levando em consideração se a transformação usada para isto é ou não significativa. Não se trata, portanto, de qual escala está sendo transformada e sim de qual o seu tipo.

Uma normalização muito utilizada é convergir todas as escalas para $[0, 1]$. Porém, qualquer que seja a normalização adotada, ela não altera o tipo das escalas dos atributos. A Tabela 3.11 ilustra algumas operações estatísticas (moda²³, mediana etc.) permitidas para as escalas:

	<i>Operações Estatísticas</i>
<i>Nominal</i>	Moda, frequência
<i>Ordinal</i>	Mediana, porcentagem
<i>Intervalo</i>	Média aritmética, desvio padrão
<i>Razão</i>	Média geométrica, variação de coeficiente
<i>Absoluta</i>	Qualquer

Tabela 3.11: Operações estatísticas permitidas por escala.

Abaixo estão descritas algumas alternativas para a normalização $[0, 1]$ [Kar05].

Genericamente, podemos usar:

$$x'_i = \frac{1}{2} \left(1 - \frac{x_i}{\sqrt{x_i^2 + a}} \right), \quad a > 0.$$

Se soubermos o máximo e o mínimo:

$$x'_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}.$$

Se qualquer $x_i \in \mathbb{R}_+ \cup \{\}$:

$$x'_i = \frac{x_i}{\sum_{i=1}^n x_i}.$$

²³Termo que designa o elemento que aparece com maior frequência, num conjunto de observações (caso exista).

Se houver valores negativos:

$$x'_i = \frac{|x_i|}{\sum_{i=1}^n |x_i|}.$$

ou

$$x'_i = \frac{x_i^2}{\sum_{i=1}^n x_i^2}.$$

Outras:

$$x'_i = n \frac{(\prod_{i=1}^n x_i)^{\frac{1}{n}}}{\sum_{i=1}^n x_i}.$$

Se estivermos trabalhando com uma escala ordinal, devemos primeiro obter o rank de seus valores ($x = 1$ até n). Em seguida normalizamos o rank:

$$x' = \frac{x - 1}{n - 1}.$$

3.4 Escalas e Modelagem de Problemas de Otimização

Como veremos em ??, o problema da alocação de tarefas para sistemas robóticos pode ser caracterizado através das restrições que lhe são impostas, ou seja, de seus casos particulares. Também veremos que há várias abordagens na modelagem de cada um desses casos (grafos, programação linear etc.).

Na Seção 3.3 relacionamos escalas e funções-valor para os robôs, ressaltando que muitos equívocos podem ser cometidos durante a modelagem dessas funções, caso sejam ignorados aspectos concernentes às escalas, o que no pior caso, pode resultar em alocações equivocadas.

Entretanto, embora os cuidados citados na Seção 3.3 sejam necessários, não são suficientes para garantir a não variação de resultados, pois um aspecto importante foi deixado de lado: a modelagem/adequação do problema para o algoritmo de alocação.

Esta seção tratará de alguns resultados obtidos em [Pek96], onde o autor demonstra que pode haver perda de informação, concernente às escalas, quando modelamos problemas de otimização. Em [Pek96] são estudados casos em otimização combinatória, programação linear, programação inteira, escalonamento, teoria dos jogos, entre outros, através do conceito de significância.

Vamos supor que devemos escolher um robô para um problema de “menor caminho”, dispondo de um grupo de robôs, todos com condições de realizar a tarefa. Também, consideraremos que o problema seja encontrar uma rota que minimize a soma das temperaturas encontradas. Logo, as arestas de um grafo que represente esse problema deverão exprimir as temperaturas dos caminhos (e não suas distâncias), conforme a Figura 3.5, com origem no ponto A e destino no ponto D .

Porém, notou-se discordância entre os robôs quanto ao menor caminho. Alguns robôs descreviam a temperatura de cada trecho conforme a Tabela 3.12, usando a escala em graus Célsius.

Figura 3.5: Grafo dos caminhos possíveis ao robô. Duas formas de ir de A para D: consecutivamente por ABCD ou através do atalho AC.

T_{AB}	T_{BC}	T_{CD}	T_{AC}
3	4	10	20

Tabela 3.12: Temperaturas em graus Célsius.

Postulavam que o $Trecho_{A-B-C-D}$ era o melhor. De fato, se somarmos as temperaturas de cada caminho temos $Trecho_{A-B-C-D} = 17^{\circ}\text{C}$ e $Trecho_{A-C-D} = 30^{\circ}\text{C}$. Concluímos com isto que o trecho $A - B - C - D$ é o melhor caminho para um robô que, primordialmente, deve procurar por baixas temperaturas em seus caminhos. Logo, esses robôs estava corretos.

Porém, outros robôs descreviam as temperaturas dos trechos em graus Fahrenheit. Como vimos na Seção 3.3, para escalas intervalo podemos usar a transformação $\phi(x) = \alpha x + \beta$, $\alpha > 0$. Então, para obtermos a tabela de temperaturas usada por esses robôs, basta aplicarmos à Tabela 3.12 a transformação admissível $\phi(x) = 1.8x + 32$. A Tabela 3.13 descreve essa situação.

T_{AB}	T_{BC}	T_{CD}	T_{AC}
37.4	39.2	50	68

Tabela 3.13: Temperaturas em Fahrenheit.

Esses robôs, portanto, postulavam que o $Trecho_{A-C-D}$ era o melhor, pois somando-se os trechos temos $Trecho_{A-B-C-D} = 126.6^{\circ}\text{F}$ e $Trecho_{A-C-D} = 118^{\circ}\text{F}$.

Percebemos que as conclusões dos robôs *não são conclusões significativas*. Isso ocorreu porque

resolveram um problema de otimização (menor caminho), sem levar em consideração aspectos concernentes às escalas.

Ao conjunto de todas as transformações admissíveis da escala $(\mathcal{A}, \mathcal{B}, f)$ denominaremos $AT(f)$. Prova-se que $AT(f)$ é o conjunto de todos os homomorfismos de $f(A)$ em B . Se $AT(f) \subseteq AT(g)$ dizemos que a escala f é mais forte que a escala g .

Também convencionou-se que todos os parâmetros de um problema serão representados por algum homomorfismo regular.

A seguir estudaremos alguns resultados obtidos em [Pek96] sobre escalas e otimização.

3.4.1 Otimização Combinatória

Problema 3.1 (Problema de otimização).

$$\max_{\mathbf{x} \in F} P(\mathbf{x}) \quad (3.19)$$

ou o equivalente

$$\min_{\mathbf{x} \in F} P(\mathbf{x}) \quad (3.20)$$

onde as variáveis $\mathbf{x}$ são de algum conjunto X , o conjunto de soluções factíveis F é um subconjunto de X e a função objetivo P é uma função real, bem definida em F . Uma solução ótima para o problema da maximização acima é a solução factível $\mathbf{x}^*$ satisfazendo

$$P(\mathbf{x}^*) \geq P(\mathbf{x}) \quad \forall \mathbf{x} \in F. \quad (3.21)$$

Convencionamos usar o negrito para denotar vetores nesta seção (a mesma notação utilizada por [Pek96]).

A função linear objetivo $P(\mathbf{x}) = \mathbf{w}^T \mathbf{x}$, onde $\mathbf{w}^T = (w_1, \dots, w_n)$, será escrita na forma $P(\mathbf{x}; \mathbf{w})$ para enfatizar que P depende dos parâmetros $\mathbf{w}$, quando for este o caso.

Geralmente, parâmetros w_i ($\mathbf{w}^T = (w_1, \dots, w_n)$) denotam algum tipo de medida, tal como unidade de custo da i -ésima variável ou um peso relativo a i -ésima variável etc. Nesses casos, a escolha de w_i depende da escolha de uma forma aceitável de medir tais parâmetros. Assumiremos que os parâmetros $\mathbf{w}$ representam valores obtidos por um homomorfismo regular f .

Inicialmente, trataremos modelagens com funções objetivos que dependem de parâmetros medidos numa única escala de medida, ou seja, analisaremos a conclusão de otimalidade para o problema:

Problema 3.2 (Problema de otimização—parâmetros em escala única).

$$\max_{\mathbf{x} \in F} P(\mathbf{x}; \mathbf{w}) \quad (3.22)$$

Teorema 3.10 ([Pek96]). *Suponha que exista uma função $m : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ tal que $P(\mathbf{x}; \lambda \mathbf{w}) = m(\lambda)P(\mathbf{x}; \mathbf{w})$ para qualquer $\lambda > 0$. Então cada solução ótima para o problema definido em 3.2 será significativa se $w_1, \dots, w_n$ forem medidos numa escala de razão.*

Teorema 3.11 ([Pek96]). *Suponha que exista $aY \in \mathbb{R}_+^n$, tal que (i) $\lambda \rightarrow G(\lambda(y))$ seja uma função não-decrescente sobre $\mathbb{R}_+$ e (ii) exista $\lambda^* > 1$, tal que $G(\lambda^*(y)) > G(y)$. Então G será homogêneo de grau r para algum $r > 0$ se e somente se existir uma função $m : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, tal que $P(\mathbf{x}; \lambda \mathbf{w}) = m(\lambda)P(\mathbf{x}; \mathbf{w})$ para qualquer $\lambda > 0$.*

O Teorema 3.11 ilustra que, sob determinadas condições, as funções homogêneas são as únicas que satisfazem as condições do Teorema 3.10, lembrando que uma função é homogênea de grau k , se quando cada um de seus argumentos é multiplicado por qualquer número $t > 0$, o valor da função é multiplicado por t^k . Formalmente:

Definição 3.13 (Função Homogênea). *Uma função f de n variáveis é homogênea de grau k se*

$$(\forall (x_1, \dots, x_n) \wedge t > 0)(f(tx_1, \dots, tx_n) = t^k f(x_1, \dots, x_n)).$$

Teorema 3.12 ([Pek96]). *Seja*

$$P(\mathbf{x}; \mathbf{w}) = \sum_{i=1}^m w_i g_i(x) \quad (3.23)$$

e suponha que exista uma constante M , tal que

$$(\forall x \in F)(\sum_{i=1}^m g_i(x) = M).$$

Então cada solução ótima do problema 3.2 é significativa se $w_1, \dots, w_n$ são medidos numa escala intervalo ou mais forte.

Teorema 3.13 ([Pek96]). *Seja a Equação 3.23 e suponha que*

$$\sum_{i=1}^m g_i(x)$$

não seja uma constante sobre F . Se $w_1, \dots, w_n$ forem medidos numa escala intervalo, então o problema 3.2 não possui solução ótima significativa.

Corolário 3.1 ([Pek96]). *Seja a Equação 3.23 e seja $\mathbf{x}^*$ uma solução ótima para o problema 3.2. Se $w_1, \dots, w_n$ forem medidos numa escala intervalo então as seguintes declarações são equivalentes:*

1. $\mathbf{x}^*$ é uma solução ótima.
2. $\sum_{i=1}^m g_i(\mathbf{x})$ é uma constante sobre F .

Teorema 3.14 ([Pek96]). *Seja a Equação 3.23 e suponha que $w_1, \dots, w_n$ sejam medidos numa escala intervalo restrita a $\mathbb{R}_+$. Então $\mathbf{x}^*$ é uma solução ótima significativa para o problema 3.2 se e somente se*

1. $\mathbf{x}^*$ for uma solução ótima para o problema 3.2 com pesos $w'_i = w_i - w_{\min}, i = 1, \dots, m$.

2. Para qualquer $x \in F$:

$$\sum_{i=1}^m g_i(\mathbf{x}^*) \geq \sum_{i=1}^m g_i(\mathbf{x}).$$

Corolário 3.2 ([Pek96]). *Seja a Equação 3.23 e suponha que $w_1, \dots, w_n$ sejam medidos numa escala intervalo restrita a $\mathbb{R}_+$. Então $\mathbf{x}^*$ é uma solução ótima significativa de minimização para o problema 3.2 se e somente se*

1. $\mathbf{x}^*$ for uma solução ótima para o problema 3.2 com pesos $w'_i = w_i - w_{\min}, i = 1, \dots, m$.

2. Para qualquer $x \in F$:

$$\sum_{i=1}^m g_i(\mathbf{x}^*) \leq \sum_{i=1}^m g_i(\mathbf{x}).$$

A seguir veremos um caso especial de programação inteira: a programação 0 – 1.

3.4.2 Programação 0 – 1

Muitos problemas de otimização combinatória podem ser formulados como um problema de programação 0 – 1, que é um caso especial de programação inteira. A seguir apresentaremos a notação utilizada.

Um importante conceito usado em [Pek96] é o de hipergrafos. Um hipergrafo é uma generalização para grafos onde o conjunto de arestas E representa uma coleção arbitrária de subconjuntos, ou seja, uma coleção de subconjuntos não-vazios de um conjunto X finito, $\mathcal{H} \subseteq 2^X$ é chamada hipergrafo.

Problema 3.3 (Problema de programação inteira).

$$\max P(\mathbf{x}; \mathbf{w}) = \mathbf{w}^T \mathbf{x} = w_1 x_1 + \dots + w_n x_n$$

sujeito a

$$(x_1, \dots, x_n) = \mathbf{x} \in F \subseteq \mathbb{Z}^n,$$

ou seja, com $F \subseteq \{0, 1\}^n$ o problema acima pode ser descrito como:

Problema 3.4 (Problema de programação 0 – 1). $\max\{\mathbf{w}^T \mathbf{x} : \mathbf{x} \in \mathcal{H} \subseteq \{0, 1\}^n\}$.

Capítulo 4

Teoria da Decisão Multiatributo

Neste capítulo, estudaremos o problema de alocação em sistemas robóticos sob o ponto de vista da teoria da decisão. Apresentaremos a taxonomia para a teoria da decisão apresentada no survey de Hwang et.al. [HY81]. Constataremos através desse estudo que a própria escolha de método é um problema de decisão, dada a variedade de abordagens possíveis.

Inicialmente, discutiremos um método para determinação dos pesos de uma função-valor. Em seguida apresentaremos a taxonomia de Hwang et.al. através de descrições resumidas e exemplos numéricos para os modelos citados.

Finalmente, na Seção 4.6 estudaremos três modelos da teoria da utilidade multiatributo (MAUT¹): um modelo para o caso ordinal, um modelo cardinal aditivo e um modelo cardinal multiplicativo.

4.1 Determinação dos Pesos

Alguns métodos de decisão requerem informação sobre a importância relativa de cada atributo. Isso geralmente é obtido através de um conjunto normalizado de pesos cuja soma é igual a 1 [HY81]. Para n atributos, temos:

$$w^T = (w_1, w_2, \dots, w_j, \dots, w_n)$$

$$\sum_{j=1}^n w_j = 1$$

Como em muitos casos o problema de decisão multiatributo é representado por uma matriz de decisão (Definição 4.1), apresentaremos esse conceito a seguir.

¹Iniciais de seu equivalente em inglês “Multi-Attribute Utility Theory”. Usaremos o termo MAUT por ser amplamente adotado no Brasil.

Definição 4.1 (Matriz de Decisão). *Uma matriz de decisão é uma matriz $m \times n$ cujo elemento X_{ij} indica o valor assumido pela alternativa i , A_i , para o atributo j , X_j .*

Note que no nosso caso as alternativas representam robôs.

Por exemplo, vamos supor um sistema composto por quatro robôs, denotados por R_i , onde $i = 1, \dots, 4$ e três atributos (custo em R\$, velocidade em Km/h e capacidade de carga em Kg), denotados por X_j , onde $j = 1, \dots, 3$. Podemos representar esses sistema através da matriz da Figura 4.1.

$$\begin{array}{c} X_1 \quad X_2 \quad X_3 \\ \begin{array}{l} R_1 \\ R_2 \\ R_3 \\ R_4 \end{array} \begin{bmatrix} 20k & 10 & 10 \\ 22k & 15 & 10 \\ 27k & 10 & 20 \\ 31k & 17 & 15 \end{bmatrix} \end{array} \quad (4.1)$$

A seguir apresentaremos em detalhes o método entrópico—um dos quatro métodos citados por Hwang et.al. para determinação dos pesos de uma função-valor (os outros métodos apresentados são: o método Eigenvetor, o método dos mínimos quadrados ponderados e o método LINMAP, sendo os dois primeiros indicados para casos em que devemos estimar a matriz de decisão através de comparações par a par). Como para um sistema robótico geralmente lidamos com características físicas previamente conhecidas (tais como velocidade, autonomia etc.), ilustraremos a determinação dos pesos através de um método que utiliza a matriz de decisão como parte da entrada, ou seja, requer conhecimento prévio da matriz.

4.1.1 Método Entrópico

O método entrópico é indicado quando os valores da matriz de decisão já são conhecidos (não precisam ser estimados). A seguir apresentaremos a entropia sob a ótica da teoria da decisão. Em seguida apresentaremos o método entrópico para determinação de atributos propriamente dito.

Origem da Entropia

A entropia é um importante conceito estudado na física e na teoria da informação, entre outras. A teoria da informação trata, por exemplo, de como representar a informação de forma compacta e isto é obtido através de processos de compressão e descompressão de dados [Say96]. Esses processos relacionam-se com aspectos estruturais dos dados (por exemplo estatísticos) ou com aspectos relativos aos receptores dos dados (podemos tirar partido da faixa de frequências que o receptor da informação consegue ver ou ouvir, por exemplo).

Antes de estudarmos a entropia, porém, trataremos de um importante conceito relacionado a ela: a auto-informação.

Definição 4.2 (Auto-Informação). *Dado um evento A que representa um conjunto de resultados de um experimento aleatório, define-se auto-informação como [Say96]:*

$$i(A) = \frac{1}{\lg P(A)} \quad (4.2)$$

ou seja, quanto maior a probabilidade de um evento, menos informativo ele é. Por exemplo, é mais provável que um mau aluno reclame do nível de exigência de uma prova que um bom aluno. Logo, esse comportamento é pouco informativo se advindo de um mau aluno. Suponha, porém, que um bom aluno, que raramente reclama das avaliações, diga ao professor que achou sua última prova difícil. Isso provavelmente será um bom indício de que o professor tenha exigido demais naquela prova. Essa situação é, portanto, mais informativa que a primeira.

De posse do conceito de auto-informação, podemos tratar a entropia. A entropia tem origem em estudos de termodinâmica, onde caracteriza o grau de desordem de um sistema. Na teoria da informação (Shannon) quantifica o grau de incerteza de uma distribuição de probabilidades (ex. distribuição de um dado normal e de um dado viciado). Para uma distribuição discreta de probabilidades, temos:

$$H = -k \sum_{i=1}^n p_i \lg p_i \quad (4.3)$$

onde a constante k denota apenas a escolha da unidade de medida. Note que, de acordo com a Equação 4.3, a entropia será máxima quando as probabilidades de todos os eventos forem iguais.

Método Entrópico

O método entrópico baseia-se nas seguintes observações: (1) um critério cujas alternativas apresentem saídas similares é de pouca utilidade² e (2) a entropia é uma boa ferramenta para determinação de contrastes entre conjuntos de dados.

Considere a Matriz 4.12.

$$\begin{array}{c} \begin{array}{cccc} & X_1 & X_2 & \dots & X_n \\ A_1 & x_{11} & x_{12} & \dots & x_{1n} \\ A_2 & x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ A_m & x_{m1} & x_{m2} & \dots & x_{mn} \end{array} \end{array} \quad (4.4)$$

²No caso extremo é desconsiderado um critério cujas alternativas apresentem saídas iguais.

O resultado esperado do atributo (critério) j , P_{ij} , vale:

$$P_{ij} = \frac{x_{ij}}{\sum_{i=1}^m x_{ij}}, \quad \forall i, j \quad (4.5)$$

A Equação 4.7 baseia-se no fato de que a relevância de P_{ij} é determinada pela diferença de resultados entre todas as alternativas. Portanto, P_{ij} deve refletir a média da informação gerada pelo conjunto de alternativas, através do critério j [HY81].

A entropia E_j do conjunto de resultados esperados do atributo j é dada por:

$$E_j = -k \sum_{i=1}^m P_{ij} \ln P_{ij}, \quad \forall j \quad (4.6)$$

onde $k = 1/\ln m$, visando garantir que $0 \leq E_j \leq 1$.

O grau de diversificação de informação d_j dado pelos resultados do atributo j é definido por $d_j = 1 - E_j$, para qualquer j .

Se não houver razão para preferir um critério sobre um outro, o princípio da insuficiência da razão sugere que cada um deva ser igualmente preferido. Então o melhor conjunto de pesos é

$$w_j = \frac{d_j}{\sum_{j=1}^n d_j}, \quad \forall j \quad (4.7)$$

Exemplo Numérico

Considere o exemplo da Matriz 4.1. Após usarmos a Equação 4.7, temos:

$$\begin{array}{c} X_1 \quad X_2 \quad X_3 \\ R_1 \left[\begin{array}{ccc} 0.20 & 0.192 & 0.182 \end{array} \right] \\ R_2 \left[\begin{array}{ccc} 0.22 & 0.288 & 0.182 \end{array} \right] \\ R_3 \left[\begin{array}{ccc} 0.27 & 0.192 & 0.364 \end{array} \right] \\ R_4 \left[\begin{array}{ccc} 0.31 & 0.327 & 0.273 \end{array} \right] \end{array} \quad (4.8)$$

Como temos 4 alternativas, adotaremos $k = \ln 4 = 1.3863$.

A entropia E_1 do primeiro atributo é dada por:

$$\begin{array}{c} X_1 \quad X_2 \quad X_3 \\ E_j \left[\begin{array}{ccc} 0.9810 & 0.9794 & 0.9683 \end{array} \right] \\ d_j \left[\begin{array}{ccc} 0.0190 & 0.0206 & 0.0317 \end{array} \right] \\ w_j \left[\begin{array}{ccc} 0.2665 & 0.2889 & 0.4446 \end{array} \right] \end{array} \quad (4.9)$$

Onde E_j é a entropia, d_j é o grau de diversificação e w_j é o peso normalizado.

Portanto, o atributo mais importante, de acordo com o método entrópico, é x_3 . A seqüência de pesos—do mais importante, para o menos importante—é w_3, w_2, w_1 .

Visão Crítica

Para entendermos melhor as motivações do método entrópico voltemos às suas premissas: (1) um critério cujas alternativas apresentem saídas similares é de pouca utilidade e (2) a entropia é uma boa ferramenta para determinação de contrastes entre conjuntos de dados.

Suponha que estejamos avaliando candidatos para uma vaga de trabalho. Após muita deliberação e discussão, suponha que a diretoria tenha determinado o seguinte conjunto de atributos relevantes para um eventual candidato: (1) caráter, (2) competência técnica, (3) pró-atividade (4) criatividade e (5) capacidade de trabalhar em equipe. Suponha que seja solicitado, então, à diretoria, em que medida um atributo é mais (ou menos) importante que seus pares.

Vamos supor que as premissas do método entrópico sejam verdadeiras. Nota-se que esse método está preocupado em salientar os contrastes entre as alternativas. Observe que nada é dito sobre o significado de cada atributo. No caso acima, por exemplo, teríamos os pesos dos atributos determinados com base nos seus contrastes e não em suas importâncias para o problema propriamente dito, pouco importando se falamos em caráter ou em criatividade.

Parece correta a observação que “um critério cujas alternativas apresentem saídas similares é de pouca utilidade”. O que parece-nos incorreto, porém, é ignorar quaisquer outros fatores relevantes para a determinação dos pesos dos atributos.

4.2 Taxonomia para Teoria da Decisão

Hwang et.al. taxonomizam a teoria da decisão através de três eixos principais. Os eixos são apresentados na Tabela 4.1 e a taxonomia para a decisão multiatributo é apresentada na Figura 4.1.

<i>Modelo</i>	
(1) <i>Sem informação</i>	Sem indicação sobre preferências na função-valor.
(2) <i>Com informação sobre atributos</i>	Preferências representadas através de atributos.
(3) <i>Com informação sobre alternativas</i>	Preferências representadas através de alternativas.

Tabela 4.1: Eixos da Taxonomia MADM.

Segundo Hwang et.al. [HY81] a informação sobre as preferências de um tomador de decisão pode ser denotada através de atributos e de alternativas. No primeiro caso—o caso mais comum—há consenso que se trata de um problema de seleção ou avaliação sobre um grupo limitado de alternativas. Portanto as alternativas são predeterminadas. Problemas desse tipo

Figura 4.1: Taxonomia para teoria da decisão multiatributo.

são denominados problemas de *tomada de decisão sobre vários atributos* (ou MADM—“Multiple Attribute Decision Making”).

No segundo caso trata-se de um problema de design, ou seja, trata-se de projetar a melhor alternativa, levando em consideração as várias interações entre as restrições de projeto que melhor satisfaçam o tomador de decisão. Problemas desse tipo são denominados problemas de *tomada de decisão sobre vários objetivos* (ou MODM—“Multiple Objective Decision Making”).

Por exemplo, a escolha do melhor carro pode ser efetuada sobre os modelos de carros disponíveis no mercado (MADM), como pode tratar-se do projeto de um carro ideal (MODM).

Na Seção 4.3 descreveremos os métodos sem informação. Métodos com informação sobre atributos e sobre alternativas serão descritos nas seções 4.4 e 4.5, respectivamente.

4.3 Métodos sem Informação

Os métodos descritos a seguir—dominância, maximin e maximax—não requerem informação sobre as preferências do tomador de decisão.

A *eliminação de alternativas por dominância* é uma regra clássica de simplificação de proble-

mas de decisão. Pode ser aplicada quando alternativas dominadas não forem importantes para o problema em questão.

O método *maximax* é indicado sempre que a eficiência for determinada pelo melhor atributo e o método *maximin*, sempre que for determinada pelo pior atributo.

4.3.1 Dominância

Uma alternativa é dita dominada se houver (outra) alternativa que a supere em um ou mais atributos e que, simultaneamente, a iguale nos atributos restantes. Alternativas dominadas serão, então, eliminadas—o que simplifica o problema, pois reduz o número de alternativas.

A técnica de eliminação de alternativas dominadas consiste na comparação par a par de alternativas. Suponha um conjunto A com n alternativas, onde para $i = 1 \dots n$, a_i representa cada alternativa. Por simplificação trabalharemos na seqüência: escolhemos duas alternativas a_1 e a_2 para comparar. Se alguma delas for dominada pela outra, a dominada é eliminada. Em seguida comparamos todas as alternativas do passo anterior, que não foram dominadas, com a_3 e eliminamos as dominadas, se houver alguma. Prosseguimos com este processo até a alternativa a_n . O conjunto que restar é o conjunto de alternativas não dominadas.

Por exemplo, suponha um problema com quatro atributos (X_1, X_2, X_3 e X_4) e quatro alternativas (A_1, A_2, A_3 e A_4), descrita através da matriz de decisão da Figura 4.10.

$$\begin{array}{c} X_1 \quad X_2 \quad X_3 \quad X_4 \\ \begin{array}{l} A_1 \\ A_2 \\ A_3 \\ A_4 \end{array} \left[\begin{array}{cccc} 2 & 3 & 4 & 5 \\ 2 & 3 & 6 & 6 \\ 3 & 4 & 2 & 9 \\ 3 & 4 & 3 & 10 \end{array} \right] \end{array} \quad (4.10)$$

Podemos simplificar o problema através da técnica de eliminação de alternativas dominadas. Para isso, iremos efetuar a comparação par a par de alternativas, a partir das alternativas de menor índice. Portanto, as alternativas A_1 e A_2 são as alternativas iniciais do processo. Após a comparação par a par de A_1 e A_2 eliminamos a alternativa A_1 , pois ela perde para A_2 nos atributos X_3 e X_4 e empata nos demais. Segundo a técnica descrita acima, devemos comparar as alternativas não eliminadas com a próxima, nesse caso, A_3 . Como só restou A_2 , devemos, portanto, comparar A_2 com A_3 . Após a comparação, notamos que nenhuma delas pôde ser eliminada, pois a alternativa A_3 vence para os atributos X_1 , X_2 e X_4 , mas perde para o atributo X_3 . Logo, ambas deverão ser comparadas com A_4 —a próxima alternativa. A comparação entre A_2 e A_4 não elimina nenhuma delas, pois, assim como no caso anterior, A_2 é superada nos atributos X_1 , X_2 e X_4 , mas vence no atributo X_3 . A comparação entre A_3 e A_4 resulta na eliminação de A_3 , que é superada nos atributos X_3 e X_4 e empata nos demais.

Logo, a aplicação da técnica da eliminação de alternativas dominadas resultou em benefício, ou seja, o problema foi simplificado: um tomador de decisão, diante do problema descrito acima,

deverá considerar apenas as alternativas A_2 e A_4 , pois as alternativas A_1 e A_3 foram filtradas no processo.

4.3.2 Maximin

Há situações cuja solução depende do elo mais fraco. Por exemplo, considere um problema que envolva a escolha do carro mais confiável para uma longa viagem. Suponha que após alguma deliberação, um tomador de decisão decida levar em consideração três aspectos para um suposto carro-candidato: a mecânica, a elétrica e os pneus. O carro mais confiável pode ser determinado como aquele cujo atributo mais fraco é mais aceitável que o atributo mais fraco de seus concorrentes. Observe que apenas um atributo é levado em consideração nesse caso (os outros são deixados de lado). Como freqüentemente lidamos com valores de atributos distintos, esse método só poderá ser usado se os atributos puderem ser comparados entre si, ou seja, se puderem ser medidos na mesma escala (normalizados).

Para ilustrar essa técnica, consideraremos a matriz da Figura 4.11, que é a matriz da Figura 4.10 após a normalização para o intervalo $[0, 1]$. A fórmula utilizada para a normalização foi $x'_{ij} = x_{ij}/x_j^{max}$.

$$\begin{array}{c} X_1 \quad X_2 \quad X_3 \quad X_4 \\ \begin{array}{l} A_1 \\ A_2 \\ A_3 \\ A_4 \end{array} \left[\begin{array}{cccc} 0.67 & 0.75 & 0.67 & 0.5 \\ 0.67 & 0.75 & 1.0 & 0.6 \\ 1.0 & 1.0 & 0.33 & 0.9 \\ 1.0 & 1.0 & 0.5 & 1.0 \end{array} \right] \end{array} \quad (4.11)$$

As alternativas A_1 , A_2 , A_3 e A_4 têm seus mínimos (pontos fracos) em X_4 , X_4 , X_3 e X_3 , respectivamente. Dessas quatro alternativas, a que possui o maior mínimo é a alternativa A_2 : seu mínimo vale 0.6 e, portanto, supera os mínimos de A_1 , A_3 e A_4 , que valem 0.5, 0.33 e 0.5, respectivamente.

Maximax

O método maximax é uma variação de Maximin. Sua única diferença é que efetua a escolha com base na comparação entre valores de melhor atributo de cada alternativa.

4.4 Métodos com Informação Sobre Atributos

Os métodos descritos a seguir expressam informações sobre preferências através de atributos, ou seja, tratam-se de abordagens MADM.

4.4.1 Métodos de Nível Mínimo

Os métodos descritos a seguir são indicados para casos em que um ou mais atributos deva atender a algum nível mínimo de exigência. O governo ou uma empresa, pode exigir um nível mínimo de qualidade ou de preço para determinado produto que esteja interessado em adquirir. Outro exemplo comum ocorre em muitas provas de seleção de candidatos (para cursos, empregos etc.): há algum tipo de exigência quanto a nota mínima para a prova.

Método Conjuntivo de Restrições

O método conjuntivo de restrições³ é indicado para casos em que vários atributos devam atender a níveis mínimos de exigência. Não é um método de seleção propriamente dito, mas um método que dicotomiza possíveis candidatos em *aceitáveis* e *não aceitáveis*. Um comandante que deseja atribuir uma tarefa a um soldado que, simultaneamente, saiba pilotar caças e conheça muito bem a área alvo da tarefa, provavelmente rejeitará ótimos pilotos que ignorem a área alvo. É provável que também rejeite outro soldado que tenha morado na área alvo, mas que não saiba pilotar.

Voltando ao exemplo da Matriz 4.10, suponha que os atributos X_1 , X_2 , X_3 e X_4 devam valer no mínimo (3, 3, 2, 7), respectivamente. Apenas as alternativas A_3 e A_4 atendem a essas exigências. As alternativas A_1 e A_2 serão classificadas como não aceitáveis.

Método Disjuntivo de Restrições

O método disjuntivo de restrições⁴ é indicado quando uma alternativa é avaliada com base no maiores valores de um atributo. O futebol americano é um bom exemplo para esse caso [HY81]. Um jogador pode ser escolhido por seu desempenho como chutador ou no passe ou na corrida. A habilidade de correr de um jogador escolhido como chutador pode ser irrelevante. Idem para a habilidade de chutar de um passador ou corredor.

Com base no exemplo da Matriz 4.10, vamos exemplificar. Suponha que um tomador de decisão especifique para os atributos X_1 , X_2 , X_3 e X_4 , respectivamente os valores (5, 5, muito alto, 8). As alternativas A_2 , A_3 e A_4 atendem os requisitos: A_2 é aceita, pois seu atributo X_3 é aceito; A_3 e A_4 são aceitas, pois possuem valores que atendem à exigência para o atributo X_4 . Já a alternativa A_1 é rejeitada, pois nenhum de seus atributos atende a qualquer das exigências.

³Do inglês "conjunctive constraint method".

⁴Do inglês "disjunctive constraint method".

4.4.2 Métodos Ordinais

Estudaremos a seguir três métodos ordinais para denotar a informação sobre preferências de um tomador de decisão: o método lexicográfico, a eliminação por aspectos e o método de permutação.

A importância relativa entre os atributos determinada pela preferência ordinal é menos exigente para o tomador de decisão do que pela preferência cardinal [HY81]. Os modelos ordinais estabelecem um ranking entre as alternativas. Os modelos cardinais vão além: não apenas constatam que a alternativa a supera a alternativa b , como determinam em que grau isso se dá, através da quantificação de suas importâncias relativas.

Método Lexicográfico

O método lexicográfico é indicado para casos de fácil determinação de uma ordem de prioridades para os atributos.

Suponha que diante de um problema com m alternativas $A_1, \dots, A_m$ e n atributos $X_1, \dots, X_n$, possamos facilmente identificar uma ordem de prioridades para os atributos. Suponha que esta ordem de prioridades coincida com $X_1, X_2, \dots, X_n$, ou seja, X_1 é o atributo mais prioritário, X_2 é o segundo atributo mais prioritário etc. e X_n é o atributo menos prioritário de todos.

Inicialmente serão identificadas todas as alternativas que possuam valores maximais para o atributo mais prioritário, nesse caso X_1 . Se apenas uma alternativa foi encontrada, então ela é a escolhida. Caso contrário, serão identificadas todas as alternativas—dentre as alternativas selecionadas anteriormente—que possuam valores maximais para o segundo atributo mais prioritário. O processo é repetido até que uma alternativa seja escolhida ou até que não haja mais atributos a considerar. No segundo caso, todas as alternativas restantes são equivalentes.

Como exemplo numérico, consideramos novamente a Matriz 4.10, supondo que a ordem de prioridades dos atributos é—do mais prioritário, para o menos prioritário— X_1, X_2, X_3, X_4 .

Inicialmente todas as alternativas são consideradas, quanto ao atributo mais prioritário, nesse caso X_1 . Apenas duas alternativas possuem valores maximais: A_3 e A_4 , que empatam com valor 3. Em seguida é considerado o segundo atributo mais prioritário, nesse caso X_2 e novamente A_3 e A_4 empatam, agora com valor 4. Finalmente X_3 —o terceiro atributo mais prioritário—é considerado, e ocorre o desempate: A_3 com valor 2 e A_4 com valor 3. Portanto, segundo o método lexicográfico, a alternativa A_4 será escolhida.

Eliminação por Aspectos

Assim como no método conjuntivo, o tomador de decisão requer um valor mínimo de exigência para cada atributo. Em seguida um atributo é selecionado e todas as alternativas que não atinjam o valor de corte para esse atributo são eliminadas. Então, outro atributo é selecionado, e assim por diante. O processo se repete até que todas as alternativas menos uma sejam eliminadas [HY81].

Suponha, por exemplo, que um tomador de decisão deva escolher um entre vários robôs para transportar blocos de gelo até um sala refrigerada. Suponha que fracionar os blocos esteja fora de questão. Inicialmente o tomador de decisão possui um conjunto R_1 de robôs (suponha que R_1 seja formado por subconjuntos de robôs muito distintos entre si). Ele elimina qualquer robô fixo (que não pode se locomover) de R_1 , obtendo R_2 (se não houver, então $R_1 = R_2$). Em seguida ele elimina de R_2 todos os robôs que não possam carregar os blocos de gelo, e obtém R_3 . Então, o tomador de decisão elimina todos os robôs de R_3 que custem mais do que um determinado valor X (valor que ele julga aceitável para a compra), obtendo R_4 . Em seguida elimina do conjunto R_4 todos os robôs que não atinjam uma velocidade mínima para o transporte e obtém R_5 . O processo de eliminação por aspectos prossegue até que só reste um robô.

Método de Permutação

O método da permutação testa cada possibilidade de ranking de alternativas contra todos os outros. Para isso são efetuadas sucessivas permutações de todos os possíveis ranking de alternativas ($m!$ possibilidades para m alternativas). O método identificará o melhor rank de alternativas e então a alternativa dominante [HY81].

Considere a Matriz 4.12, envolvendo m alternativas e n atributos.

$$\begin{array}{c} \begin{matrix} & X_1 & X_2 & \dots & X_n \\ A_1 & \left[\begin{matrix} x_{11} & x_{12} & \dots & x_{1n} \end{matrix} \right. \\ A_2 & \left[\begin{matrix} x_{21} & x_{22} & \dots & x_{2n} \end{matrix} \right. \\ \vdots & \left[\begin{matrix} \vdots & \vdots & \vdots & \vdots \end{matrix} \right. \\ A_m & \left[\begin{matrix} x_{m1} & x_{m2} & \dots & x_{mn} \end{matrix} \right] \end{matrix} \end{array} \quad (4.12)$$

Suponha um conjunto de pesos w_j , $j = 1, \dots, n$ e $\sum_j w_j = 1$, seja dado para o conjunto de atributos correspondente.

Suponha que haja três alternativas A_1 , A_2 e A_3 . Das 6 possíveis permutações, testaremos (A_2, A_3, A_1) . O conjunto de concordância parcialmente ordenado é $\{A_2 \geq A_3, A_2 \geq A_1, A_3 \geq A_1\}$ e o conjunto de discordância é $\{A_2 \leq A_3, A_2 \leq A_1, A_3 \leq A_1\}$. Se no ranking aparecer $A_k \geq A_l$, $x_{kj} \geq x_{lj}$ será atribuído como w_j e $x_{kh} \leq x_{lh}$ será atribuído como $-w_h$.

Considere a i -ésima permutação:

$$P_i = (\dots, A_k, \dots, A_l, \dots), \quad i = 1, \dots, m! \quad (4.13)$$

Nesse caso A_k é melhor ranqueado que A_l . O critério de avaliação (R_i) da hipótese de escolha (P_i) é dado pela Equação 4.14.

$$R_i = \sum_{j \in C_{kl}} w_j - \sum_{j \in D_{kl}} w_j, \quad i = 1, \dots, m! \quad (4.14)$$

onde

$$C_{kl} = \{j | x_{kj} \geq x_{lj}\}, \quad k, l = 1, \dots, m, \quad k \neq l$$

$$D_{kl} = \{j | x_{kj} \leq x_{lj}\}, \quad k, l = 1, \dots, m, \quad k \neq l$$

O método da permutação é mais indicado para casos ordinais, embora também seja capaz de lidar com problemas cardinais. O método ELECTRE, que será estudado em 4.4.3, é o método análogo ao método da permutação para lidar com casos cardinais. Ambos medem o nível de concordância da ordem de preferências. No método da permutação, porém, isso é feito sobre a ordem completa de preferências. Veremos que o método ELECTRE primeiro mede o nível de concordância da ordem parcial sobre as alternativas para em seguida formar a possível ordem agregada [HY81].

4.4.3 Métodos Cardinais

Como mencionamos anteriormente, modelos cardinais não apenas constatarem que uma determinada alternativa supera outra, como determinam em que grau isso se dá, através da quantificação de suas importâncias relativas.

Estudaremos a seguir cinco métodos cardinais para denotar a informação sobre preferências de um tomador de decisão: o método de atribuição linear, o método de ponderação aditiva simples, o método de ponderação aditiva hierárquica, ELECTRE (Elimination et Choice Translating Reality) e TOPSIS (Technique for Order Preference by Similarity to Ideal Solution).

Método de Atribuição Linear

O método de atribuição linear⁵ parte do pressuposto que uma alternativa que possua muitos atributos bem ranqueados deva ser bem ranqueada.

Por exemplo, com base na Tabela 4.15, determinamos a Tabela 4.16 que ranqueia as alternativas quanto aos atributos, com a suposição—para todos os atributos—de que quanto maior for o valor dos atributos, melhor será.

$$\begin{array}{c}
 \\
 \\
 \\
 \\
 \end{array}
 \begin{array}{cccc}
 & X_1 & X_2 & X_3 & X_4 \\
 \begin{array}{l} A_1 \\ A_2 \\ A_3 \\ A_4 \end{array} & \left[\begin{array}{cccc} 2 & 3 & 4 & 5 \\ 2 & 2 & 9 & 3 \\ 5 & 1 & 2 & 9 \\ 3 & 2 & 3 & 2 \end{array} \right]
 \end{array}
 \quad (4.15)$$

⁵Do inglês “linear assignment method”.

$$\begin{array}{c}
\begin{array}{c} 1^\circ \\ 2^\circ \\ 3^\circ \\ 4^\circ \end{array} \left[\begin{array}{cccc} X_1 & X_2 & X_3 & X_4 \\ A_3 & A_1 & A_2 & A_3 \\ A_4 & A_2, A_4 & A_1 & A_1 \\ A_1, A_2 & & A_4 & A_2 \\ & A_3 & A_3 & A_4 \end{array} \right]
\end{array} \quad (4.16)$$

A matriz produto-atributo π , será uma matriz não negativa, cujos elementos π_{ik} representem o número de vezes que uma alternativa A_i foi ranqueada na k -ésima posição. Se os pesos fossem iguais para todos os atributos, a matriz π para nosso caso seria igual a Matriz 4.17.

$$\begin{array}{c}
\begin{array}{c} 1^\circ \\ 2^\circ \\ 3^\circ \\ 4^\circ \end{array} \left[\begin{array}{cccc} 1^\circ & 2^\circ & 3^\circ & 4^\circ \\ A_1 & 1 & 2 & 1 & 0 \\ A_2 & 1 & 1 & 2 & 0 \\ A_3 & 2 & 0 & 0 & 2 \\ A_4 & 0 & 2 & 1 & 1 \end{array} \right]
\end{array} \quad (4.17)$$

Os empates podem ser equalizados, tomando um atributo extra com a metade do peso original, conforme a Tabela 4.18.

$$\begin{array}{c}
\begin{array}{c} 1^\circ \\ 2^\circ \\ 3^\circ \\ 4^\circ \end{array} \left[\begin{array}{cccccc} X_{11}(w_1/2) & X_{12}(w_1/2) & X_{21}(w_2/2) & X_{22}(w_2/2) & X_3 & X_4 \\ A_3 & A_3 & A_1 & A_1 & A_2 & A_3 \\ A_4 & A_4 & A_2 & A_4 & A_1 & A_1 \\ A_1 & A_2 & A_4 & A_2 & A_4 & A_2 \\ A_2 & A_1 & A_3 & A_3 & A_3 & A_4 \end{array} \right]
\end{array} \quad (4.18)$$

Vamos supor que o conjunto de pesos (normalizado), determinado pelo tomador de decisão, seja $w = \{.4, .3, .2, .1\}$. Então calculamos a matriz π ponderada, conforme a Matriz 4.19.

$$\begin{array}{c}
\begin{array}{c} 1^\circ \\ 2^\circ \\ 3^\circ \\ 4^\circ \end{array} \left[\begin{array}{cccc} 1^\circ & 2^\circ & 3^\circ & 4^\circ \\ A_1 & .3 & .3 & .2 & .15 \\ A_2 & .2 & .15 & .45 & .2 \\ A_3 & .5 & 0 & 0 & .5 \\ A_4 & 0 & .55 & .35 & .1 \end{array} \right]
\end{array} \quad (4.19)$$

onde, $(\pi_{11} = w_{21} + w_{22} = .15 + .15 = .3)$, $(\pi_{12} = w_3 + w_4 = .3)$ e $(\pi_{13} = w_{11} = .2)$ etc.

Método de Ponderação Aditiva Simples

O método de ponderação aditiva simples⁶ é um dos mais adotados e conhecidos métodos MADM. É análogo a encontrar uma função-valor aditiva (veremos alguns métodos para isso na

⁶Do inglês "simple additive weighting method".

Seção 4.6).

Para cada atributo é atribuído um valor de importância ou peso, ou seja, é determinado o conjunto de pesos $w = \{w_1, w_2, \dots, w_n\}$. Também é feita a normalização entre os atributos para uma escala comum, geralmente o intervalo $[0, 1]$. Supomos, também, que os atributos estejam em uma escala de razão (ou mais forte). Então, é determinada a alternativa que maximize f , onde para j atributos e i alternativas, ou seja, a Equação 4.20.

$$f = \frac{\sum_{j=1}^n w_j x_{ij}}{\sum_{j=1}^n w_j} \quad (4.20)$$

Normalmente os pesos são normalizados para que $\sum_{j=1}^n w_j = 1$, que aplicado à Equação 4.20 resulta na Equação 4.21.

$$f = \sum_{j=1}^n w_j x_{ij} \quad (4.21)$$

Para exemplificar, considere a Matriz 4.11. Também suponha que o conjunto de pesos (também normalizado para o intervalo $[0, 1]$) seja dado por $w = \{w_1 = 0.4, w_2 = 0.3, w_3 = 0.2, w_4 = 0.1\}$.

Teremos, portanto, $A_1 = 6.77$, $A_2 = 7.53$, $A_3 = 8.56$ e $A_4 = 9$ e a alternativa A_4 será a escolhida.

Na Seção 4.6 estudaremos a determinação de funções-valor, através da teoria da utilidade multiatributo (MAUT), para o caso ordinal aditivo, para o caso cardinal aditivo e o caso cardinal multiplicativo.

Método de Ponderação Aditiva Hierárquica

Uma hierarquia pode ser vista em termos de conjuntos parcialmente ordenados. Um hierarquia possui níveis e o maior nível consiste de um único elemento. Cada elemento de um dado nível domina ou cobre um ou mais elementos de um nível imediatamente inferior (se houver) [HY81]. Suponha que L_k represente o k -ésimo nível de uma hierarquia com h níveis. Assuma que $Y = \{y_1, y_2, \dots, y_k\} \in L_k$, $X = \{x_1, x_2, \dots, x_{k+1}\} \in L_{k+1}$ e que existe um elemento $z \in L_{k-1}$ tal que Y é coberto por z . Na Figura 4.2 podemos observar uma hierarquia completa⁷ para esse exemplo.

Considere as funções prioridade $w_z : Y \rightarrow [0, 1]$ e $w_y : X \rightarrow [0, 1]$. Podemos construir a função

⁷Uma hierarquia é dita completa se todo elemento $x \in L_k$ for coberto por cada elemento em L_{k-1} , $k = 2, \dots, h$.

Figura 4.2: Hierarquia Z-Y-X.

prioridade dos elementos de X com respeito a z , denotada w ($w : X \rightarrow [0, 1]$), pela Equação 4.22.

$$w(x_i) = \sum_{j=1}^k w_{y_j}(x_i) w_z(y_j), \quad i = 1, \dots, k+1 \quad (4.22)$$

Ou seja, trata-se do processo de ponderar a influência do elemento y_j sobre a prioridade de x_i pela multiplicação dele com a importância de y_j com respeito a z . Se uma matriz B denotar $w_{y_j}(x_i)$, ou seja, $b_{ij} = w_{y_j}(x_i)$ e se $W_i = w(X_i)$ e $W'_j = w_z(y_j)$, então podemos reescrever a Equação 4.22 na Equação 4.23.

$$W_i = \sum_{j=1}^k b_{ij} W'_j, \quad i = 1, \dots, k+1 \quad (4.23)$$

Falamos, portanto, no vetor prioridade W e na matriz prioridade B , ou seja, em $W = BW'$.

Considere uma hierarquia completa H , com h níveis e com o elemento de menor prioridade b . Seja B_k a matriz prioridade para o k -ésimo nível, $k = 1, \dots, h$. Se W' for o vetor prioridade do p -ésimo nível com respeito a algum elemento z no $p-1$ -ésimo nível, então o vetor prioridade W do q -ésimo nível ($p < q$) com respeito a z é dado pela Equação 4.24 e o vetor de menor nível com respeito ao elemento b é dado pela Equação 4.25 [HY81].

$$W = B_q B_{q-1} \dots B_{p+1} W' \quad (4.24)$$

$$W = B_h B_{h-1} \dots B_2 W' \quad (4.25)$$

Como usualmente L_1 possui apenas um elemento, W' é um escalar (e não um vetor).

Para exemplificar, considere novamente a Matriz 4.10, repetida aqui por comodidade.

$$\begin{array}{c} X_1 \quad X_2 \quad X_3 \quad X_4 \\ \begin{array}{l} A_1 \\ A_2 \\ A_3 \\ A_4 \end{array} \left[\begin{array}{cccc} 2 & 3 & 4 & 5 \\ 2 & 3 & 6 & 6 \\ 3 & 4 & 2 & 9 \\ 3 & 4 & 3 & 10 \end{array} \right] \end{array} \quad (4.26)$$

Como a matriz é dada, não é necessário efetuar comparações par a par. Podemos obter as prioridades através da equação

$$k_j = \frac{x_{ij}}{\sum_{i=1}^4 x_{ij}}, \quad j = 1, 2, 3, 4.$$

Observe que há casos em que um atributo m qualquer refere-se a uma grandeza que deve ser minimizada (como por exemplo o custo). Para esses casos, alteramos a equação anterior para

$$k_m = \frac{1/x_{im}}{\sum_{i=1}^4 (1/x_{im})}$$

e obtemos a matriz B_3 :

$$\left[\begin{array}{cccc} .200 & .214 & .267 & .167 \\ .200 & .214 & .400 & .200 \\ .300 & .286 & .133 & .300 \\ .300 & .286 & .200 & .333 \end{array} \right]$$

Suponha que o conjunto de pesos (normalizado para o intervalo $[0, 1]$) seja dado por $B_2 = [.4, .3, .2, .1]^T$. Então $W = B_3 \times B_2 \times W'$, onde $W' = 1$, pois só há um elemento em L_1 , então obtemos:

$$\begin{array}{c} A_1 \\ A_2 \\ A_3 \\ A_4 \end{array} \left[\begin{array}{c} .133 \\ .064 \\ .211 \\ .218 \end{array} \right] \quad (4.27)$$

O resultado obtido, portanto, será $A_4 > A_3 > A_1 > A_2$.

ELECTRE

O método ELECTRE (Elimination et Choice Translating Reality) usa o conceito de relação outranking (super-classificação). Métodos outranking visam a representação de casos onde o

tomador de decisão não saiba dizer se a alternativa x é melhor ou igual a alternativa y . A comparação entre x e y baseia-se, então, no grau para o qual as avaliações das alternativas e os pesos preferenciais firmam ou não a relação de dominância par a par entre as alternativas [HY81]. A seguir apresentaremos os passos do método ELECTRE e simultaneamente aplicaremos cada passo através de um exemplo.

Passo 1. Cálculo da matriz de decisão normalizada (R). Para normalizarmos a matriz de decisão utilizaremos a Equação 4.28.

$$r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^m x_{ij}^2}} \quad (4.28)$$

Obtemos, então, a matriz normalizada R , conforme 4.31.

$$R = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ r_{21} & r_{22} & \dots & r_{2n} \\ \vdots & & & \\ r_{m1} & r_{m2} & \dots & r_{mn} \end{bmatrix} \quad (4.29)$$

Para exemplificar, utilizaremos a Matriz 4.30. Porém, para tornar o exemplo mais interessante, iremos supor que todos os atributos—menos um—devam ser maximizados. O atributo que deverá ser minimizado será X_3 (denotando, por exemplo, o custo de operação).

$$\begin{array}{c} X_1 \quad X_2 \quad X_3 \quad X_4 \\ \begin{matrix} A_1 \\ A_2 \\ A_3 \\ A_4 \end{matrix} \begin{bmatrix} 4 & 3 & 2 & 2 \\ 2 & 6 & 7 & 3 \\ 5 & 3 & 3 & 2 \\ 3 & 1 & 5 & 4 \end{bmatrix} \end{array} \quad (4.30)$$

E a matriz normalizada

$$R = \begin{bmatrix} .544 & .405 & .214 & .348 \\ .272 & .809 & .750 & .522 \\ .680 & .405 & .322 & .348 \\ .408 & .135 & .536 & .696 \end{bmatrix}$$

Passo 2. Cálculo da matriz de decisão normalizada ponderada (V). Essa matriz pode ser calculada através da multiplicação de cada coluna da matriz R , obtida no passo anterior,

com seu peso associado w_j .

$$V = RW = \begin{bmatrix} w_1 r_{11} & \dots & w_j r_{1j} & \dots & w_n r_{1n} \\ w_1 r_{21} & \dots & w_j r_{2j} & \dots & w_n r_{2n} \\ \vdots & & & & \\ w_1 r_{m1} & \dots & w_j r_{mj} & \dots & w_n r_{mn} \end{bmatrix} \quad (4.31)$$

Para o exemplo, vamos supor que o conjunto de pesos (normalizado para o intervalo $[0, 1]$) é dado por $w = \{w_1 = 0.4, w_2 = 0.3, w_3 = 0.2, w_4 = 0.1\}$. Então teremos

$$V = \begin{bmatrix} .218 & .122 & .043 & .035 \\ .109 & .243 & .150 & .052 \\ .272 & .122 & .064 & .035 \\ .163 & .041 & .107 & .070 \end{bmatrix}$$

Passo 3. Determinação do conjunto de concordância (C_{kl}) e do conjunto de discordância (D_{kl}).

Para cada par k e l ($k, l = 1, 2, \dots, m$ e $k \neq l$) o conjunto do critério de decisão é dividido em dois subconjuntos distintos. O conjunto de concordância C_{kl} de A_k e A_l é composto de todos os critérios para os quais A_k é preferido a A_l , ou seja, $C_{kl} = \{j | x_{kj} \geq x_{lj}\}$. O conjunto de discordância será, portanto, $D_{kl} = \{j | x_{kj} < x_{lj}\}$ ou $J - C_{kl}$. No exemplo teremos (lembrando que X_3 denota custo):

$C_{12} = \{1, 3\}$	$D_{12} = \{2, 4\}$	$C_{21} = \{2, 4\}$	$D_{21} = \{1, 3\}$
$C_{13} = \{2, 3, 4\}$	$D_{13} = \{1\}$	$C_{23} = \{2, 4\}$	$D_{23} = \{1, 3\}$
$C_{14} = \{1, 2, 3\}$	$D_{14} = \{4\}$	$C_{24} = \{2\}$	$D_{24} = \{1, 3, 4\}$
$C_{31} = \{1, 2, 4\}$	$D_{31} = \{3\}$	$C_{41} = \{4\}$	$D_{41} = \{1, 2, 3\}$
$C_{32} = \{1, 3\}$	$D_{32} = \{2, 4\}$	$C_{42} = \{1, 3, 4\}$	$D_{42} = \{2\}$
$C_{34} = \{1, 2, 3\}$	$D_{34} = \{4\}$	$C_{43} = \{4\}$	$D_{43} = \{1, 2, 3\}$

Passo 4. Cálculo da matriz de concordância (C). O valor relativo do conjunto de concordância é medido através do índice de concordância, conforme a Equação 4.32.

$$c_{kl} = \frac{\sum_{j \in C_{kl}} w_j}{\sum_{j=1}^n w_j} \quad (4.32)$$

E para o conjunto normalizado ponderado

$$c_{kl} = \sum_{j \in C_{kl}} w_j \quad (4.33)$$

O índice de concordância reflete a importância relativa de A_k com respeito a A_l e $0 \leq c_{kl} \leq 1$. Um valor alto para c_{kl} indica que A_k é preferido a A_l .

Em nosso exemplo, um elemento C_{12} da matriz C é calculado

$$C_{12} = w_1 + w_3 = .6$$

A matriz de concordância será:

$$C = \begin{bmatrix} - & .6 & .6 & .9 \\ .4 & - & .4 & .3 \\ .8 & .6 & - & .9 \\ .1 & .7 & .1 & - \end{bmatrix}$$

Passo 5. Cálculo da matriz de discordância.

Com o intuito de medir o grau para o qual as avaliações da alternativa A_k são piores que as avaliações de outra A_l foi criado o índice de discordância, conforme a Equação 4.34 ($0 \leq d_{kl} \leq 1$).

$$d_{kl} = \frac{\max_{j \in D_{kl}} |v_{kj} - v_{lj}|}{\max_{j \in J} |v_{kj} - v_{lj}|} \quad (4.34)$$

Um valor alto para d_{kl} implica que A_k é menos favorável que A_l . Em nosso exemplo, teremos

$$\begin{aligned} d_{12} &= \frac{\max_{j \in D_{12}} |v_{1j} - v_{2j}|}{\max_{j \in J} |v_{1j} - v_{2j}|} \\ &= \frac{\max_{j \in \{2,4\}} |v_{1j} - v_{2j}|}{\max_{j \in \{1,2,3,4\}} |v_{1j} - v_{2j}|} \\ &= \frac{\max\{|v_{12} - v_{22}|, |v_{14} - v_{24}|\}}{\max_{j \in \{1,2,3,4\}} |v_{1j} - v_{2j}|} \\ &= \frac{\max\{.121, .017\}}{\max\{.109, .121, .107, .017\}} = \frac{.121}{.121} = 1 \end{aligned}$$

Do mesmo modo

$$d_{13} = \frac{\max_{j \in D_{13}} |v_{1j} - v_{3j}|}{\max_{j \in J} |v_{1j} - v_{3j}|} = \frac{\max_{j \in \{1\}} |v_{1j} - v_{3j}|}{\max_{j \in \{1,2,3,4\}} |v_{1j} - v_{3j}|} = \frac{\max\{.055\}}{\max\{.055, 0, .021, 0\}} = 1$$

e

$$d_{14} = \frac{\max_{j \in \{4\}} |v_{1j} - v_{4j}|}{\max_{j \in \{1,2,3,4\}} |v_{1j} - v_{4j}|} = \frac{\max\{.035\}}{\max\{.055, .081, .064, .035\}} = \frac{.035}{.081} = .432$$

e assim por diante. A matriz de discordância será

$$D = \begin{bmatrix} - & .901 & 1 & .432 \\ .901 & - & 1 & .267 \\ .389 & .742 & - & .321 \\ .79 & 1 & 1 & - \end{bmatrix}$$

Passo 6. Determinação da matriz de concordância dominante (C_{dom}). A matriz de concordância dominante é calculada com a inclusão de um limiar de disparo⁸ para o índice de concordância. Então, a alternativa A_k só conseguirá dominar a alternativa A_l caso seu índice de concordância c_{kl} iguale ou supere um certo valor de limiar de disparo $\bar{c}$, ou seja, $c_{kl} \geq \bar{c}$. O valor limiar de disparo pode ser calculado, por exemplo, como a média do índice de concordância, conforme a Equação 4.35.

$$\bar{c} = \frac{\sum_{k=1}^m \sum_{\substack{l=1 \\ l \neq k}}^m c_{kl}}{m(m-1)} \quad (4.35)$$

Com base no valor do limiar de disparo, construímos uma matriz booleana C_{dom} , cujos elementos são definidos conforme a Equação 4.36.

$$c'_{kl} = \begin{cases} 1 & \text{se } c_{kl} \geq \bar{c} \\ 0 & \text{se } c_{kl} < \bar{c} \end{cases} \quad (4.36)$$

Para o exemplo temos

$$\bar{c} = \frac{\sum_{k=1}^4 \sum_{l=1}^4 c_{kl}}{m(m-1)} = \frac{6.4}{12} = .533$$

e a matriz de concordância dominante vale

$$C_{dom} = \begin{bmatrix} - & 1 & 1 & 1 \\ 0 & - & 0 & 0 \\ 1 & 1 & - & 1 \\ 0 & 1 & 0 & - \end{bmatrix}$$

Passo 7. Determinação da matriz de discordância dominante.

A construção da matriz de discordância dominante é análoga à construção da matriz de concordância dominante, conforme a Equação 4.37.

$$\bar{d} = \frac{\sum_{k=1}^m \sum_{\substack{l=1 \\ l \neq k}}^m d_{kl}}{m(m-1)} \quad (4.37)$$

⁸Do inglês "threshold".

Com base no valor do limiar de disparo, construímos uma matriz booleana D_{dom} , cujos elementos são definidos conforme a Equação 4.38.

$$d'_{kl} = \begin{cases} 1 & \text{se } d_{kl} \leq \bar{d} \\ 0 & \text{se } d_{kl} > \bar{d} \end{cases} \quad (4.38)$$

E para o exemplo temos

$$\bar{d} = \frac{\sum_{k=1}^4 \sum_{l=1}^4 d_{kl}}{m(m-1)} = \frac{8.743}{12} = .729$$

e a matriz de discordância dominante vale

$$D_{dom} = \begin{bmatrix} - & 0 & 0 & 1 \\ 0 & - & 0 & 1 \\ 1 & 0 & - & 1 \\ 0 & 0 & 0 & - \end{bmatrix}$$

Passo 8. Determinação da matriz de agregação dominante (M_{cd}). Em seguida calculamos a matriz de agregação dominante M_{cd} , ou seja, a intersecção de C_{dom} e D_{dom} . Um elemento de M_{cd} qualquer pode ser calculado conforme a Equação 4.39.

$$m_{kl} = c'_{kl} \cdot d'_{kl} \quad (4.39)$$

Em nosso caso teremos

$$M_{kl} = \begin{bmatrix} - & 0 & 0 & 1 \\ 0 & - & 0 & 0 \\ 1 & 0 & - & 1 \\ 0 & 0 & 0 & - \end{bmatrix}$$

Passo 9. Eliminação das alternativas menos favoráveis.

A matriz M_{kl} gera uma ordem parcial entre as alternativas. Se $m_{kl} = 1$, então A_k é preferida a A_l para ambos os critérios—de concordância e de discordância. Há, porém, o risco de que mesmo assim A_k seja dominada por outras alternativas.

Em nosso exemplo, a partir da matriz M_{kl} determinamos as seguintes relações: $A_1 \rightarrow A_2$, $A_1 \rightarrow A_4$, $A_3 \rightarrow A_1$, $A_3 \rightarrow A_2$, $A_3 \rightarrow A_4$. Portanto, de acordo com o método ELECTRE, a alternativa A_3 é a escolhida.

TOPSIS

O método TOPSIS (Technique for Order Preference by Similarity to Ideal Solution) baseia-se no conceito de que a alternativa escolhida deva ser a mais próxima da solução ideal (e mais distante da solução negativa ideal). Considere uma função-valor com todos os atributos monotonicamente crescentes. A solução ideal é aquela composta pelos melhores valores para os atributos. TOPSIS escolherá a alternativa que possua a menor distância (ponderada) para a solução ideal, no senso geométrico.

Passo 1. Construção da matriz de decisão normalizada. O processo de normalização é o mesmo descrito em 4.4.3. Supondo a matriz de decisão 4.41, através da Equação 4.40, podemos obter sua matriz normalizada 4.42.

$$r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^m x_{ij}^2}} \quad (4.40)$$

$$\begin{array}{c} X_1 \quad X_2 \quad X_3 \quad X_4 \\ \begin{array}{l} A_1 \\ A_2 \\ A_3 \\ A_4 \end{array} \begin{bmatrix} 4 & 3 & 2 & 2 \\ 2 & 6 & 7 & 3 \\ 5 & 3 & 3 & 2 \\ 3 & 1 & 5 & 4 \end{bmatrix} \end{array} \quad (4.41)$$

E a matriz normalizada

$$R = \begin{bmatrix} .544 & .405 & .214 & .348 \\ .272 & .809 & .750 & .522 \\ .680 & .405 & .322 & .348 \\ .408 & .135 & .536 & .696 \end{bmatrix}$$

Passo 2. Cálculo da matriz de decisão normalizada ponderada (V).

Para o exemplo, vamos supor que o conjunto de pesos (normalizado para o intervalo $[0, 1]$) é dado por $w = \{w_1 = 0.4, w_2 = 0.3, w_3 = 0.2, w_4 = 0.1\}$. Então teremos (conforme vimos em 4.4.3, passo 2)

$$V = \begin{bmatrix} .218 & .122 & .043 & .035 \\ .109 & .243 & .150 & .052 \\ .272 & .122 & .064 & .035 \\ .163 & .041 & .107 & .070 \end{bmatrix}$$

Passo 3. Determinação da solução ideal (A^) e a solução negativa ideal (A^-).* Criaremos duas alternativas artificiais:

$$A^* = \{(max_i v_{ij} | j \in J), (min_i v_{ij} | j \in J') | i = 1, 2, \dots, m\} = (v_1^*, \dots, v_n^*) \quad (4.42)$$

e

$$A^- = \{(min_i v_{ij} | j \in J), (max_i v_{ij} | j \in J') | i = 1, 2, \dots, m\} = (v_1^-, \dots, v_n^-) \quad (4.43)$$

onde $J = \{j = 1, 2, \dots, n\}$ com j associado ao critério-benefício e $J' = \{j = 1, 2, \dots, n\}$ com j associado ao critério-custo.

Ou seja, A^* indica a alternativa mais preferida e A^- a menos preferida.

Em nosso exemplo teremos

$$A^* = \{(max_i v_{i1}, max_i v_{i2}, max_i v_{i3}, max_i v_{i4}) = (.272, .243, .150, .070)$$

e

$$A^- = \{(min_i v_{i1}, min_i v_{i2}, min_i v_{i3}, min_i v_{i4}) = (.109, .041, .043, .035)$$

Passo 4. Cálculo da medida de separação A separação entre uma alternativa qualquer e a alternativa ideal é dada pela distância Euclidiana n -dimensional, conforme a Equação 4.44.

$$S_i^* = \sqrt{\sum_{j=1}^n (v_{ij} - v_j^*)^2}, \quad i = 1, 2, \dots, m \quad (4.44)$$

Similarmente temos a separação para a negativa ideal

$$S_i^- = \sqrt{\sum_{j=1}^n (v_{ij} - v_j^-)^2}, \quad i = 1, 2, \dots, m \quad (4.45)$$

E aplicado ao exemplo

$$S_i^* = \sqrt{\sum_{j=1}^4 (v_{ij} - v_j^*)^2}, \quad i = 1, 2, 3, 4$$

ou seja,

$$S_1^* = \sqrt{(.218 - .272)^2 + (.109 - .272)^2 + (.163 - .272)^2} = .203$$

$$S_2^* = \sqrt{(.122 - .243)^2 + (.122 - .243)^2 + (.041 - .243)^2} = .265$$

$$S_3^* = \sqrt{(.043 - .150)^2 + (.064 - .150)^2 + (.107 - .150)^2} = .144$$

$$S_4^* = \sqrt{(.035 - .070)^2 + (.052 - .070)^2 + (.035 - .070)^2} = .053$$

e

$$S_i^- = \sqrt{\sum_{j=1}^4 (v_{ij} - v_j^-)^2} \quad i = 1, 2, 3, 4$$

ou seja,

$$S_1^- = \sqrt{(.218 - .109)^2 + (.272 - .109)^2 + (.163 - .109)^2} = .203$$

$$S_2^- = \sqrt{(.122 - .041)^2 + (.243 - .041)^2 + (.122 - .041)^2} = .232$$

$$S_3^- = \sqrt{(.150 - .043)^2 + (.064 - .043)^2 + (.107 - .043)^2} = .126$$

$$S_4^- = \sqrt{(.052 - .035)^2 + (.070 - .035)^2} = .039$$

Passo 5. Cálculo da proximidade relativa para a solução ideal

A proximidade relativa de A_i para A^* é definida pela Equação 4.46

$$C_i^* = \frac{S_i^-}{S_i^* + S_i^-}, \quad 0 < C_i^* < 1, \quad i = 1, 2, \dots, m \quad (4.46)$$

Observe que se $A_i = A^*$ então $C_i^* = 1$ e se $A_i = A^-$ então $C_i^* = 0$.

No exemplo teremos

$$C_1^* = \frac{S_1^-}{S_1^* + S_1^-} = \frac{.203}{.203 + .203} = .5$$

de modo análogo obteremos $C_2^* = .467$, $C_3^* = .467$ e $C_4^* = .424$.

Passo 6. Determinação do rank de preferências Podemos determinar o rank de preferências das alternativas de acordo com a ordem descendente C_i^* .

Portanto, para o exemplo, determinamos o rank: A_1, A_2, A_3, A_4 .

4.4.4 Taxa Marginal de Substituição

Suponha um problema que envolva vários atributos. Qual o aumento que devemos esperar no valor do atributo X_i , com a redução do valor de outro atributo X_j ($i \neq j$)—enquanto os outros atributos permanecem inalterados?

Tradeoffs Hierárquicos

Problemas que envolvam independência entre os atributos podem ser simplificados através do método de tradeoffs hierárquicos. Esse método consiste em estabelecermos, com auxílio de tradeoffs, a equivalência entre alternativas usando todos os atributos menos um e então avaliá-los através desse atributo remanescente. Para isso descartamos os atributos um a um, usando tradeoffs da combinação de dois atributos. Curvas de indiferença facilitam esse processo.

4.5 Métodos com Informação Sobre Alternativas

Os métodos estudados nesta seção requerem que o tomador de decisão indique sua preferência entre duas alternativas, no nosso caso, entre dois robôs. Este tipo de informação é muito mais difícil de se obter (o número de atributos é um grande limitador para essa técnica: problemas que envolvam mais de sete atributos são praticamente inviáveis [HY81]). Métodos com informação sobre alternativas requerem que o tomador de decisão tenha razoável experiência quanto aos seus robôs para uma determinada tarefa—o que raramente é o caso. Pelo exposto acima, daremos apenas uma breve descrição de cada método, dado que raramente serão usados para sistemas robóticos.

4.5.1 Preferência Par a Par

São métodos que envolvem preferência par a par de alternativas como entrada. Os métodos prevêm e tentam minimizar possíveis inconsistências.

LINMAP

O método LINMAP (Linear Programming techniques for Multidimensional Analysis of Preference) denota m alternativas compostas de n atributos através de m pontos num espaço n -dimensional. Também assume que o tomador de decisão possua um ponto ideal que denote a localização de sua alternativa preferida.

Método Aditivo Simples Iterativo

O método iterativo ranqueia as alternativas através de interações com o tomador de decisão. Tem como proposta evitar demanda excessiva de informação sobre as preferências do tomador de decisão.

4.5.2 Ordem de Proximidade Par a Par

Uma representação espacial multidimensional é construída a partir da ordem de proximidade de pares de alternativas do tomador de decisão. As alternativas, então, são representadas por pontos em um espaço com muito menos dimensões que o original. O tomador de decisão localiza a alternativa ideal nesse espaço e a distância até o ponto ideal é determinada [HY81].

Scaling Multidimensional com Ponto Ideal (MDS)

O método MDS é indicado quando desejamos encontrar uma estrutura escondida em dados experimentais, como por exemplo em casos que envolvam muitos atributos ou com atributos vagos e desconhecidos.

4.6 Teoria da Utilidade Multiatributo

Na teoria da utilidade multiatributo para o caso determinístico, três teorias de funções de preferência multiatributo se destacam: (1) o modelo de preferências multiatributo ordinal aditivo, que apenas classifica preferências, sem quantificá-las; (2) o modelo que faz uso do conceito de independência fraca de diferenças (permite quantificação) e (3) o modelo que faz uso do conceito de independência de diferenças. Estudaremos a seguir cada um desses modelos.

4.6.1 Modelo Ordinal

Se não houver incerteza envolvida no problema, podemos adotar uma função-valor multiatributo que será incumbida de traduzir em valores os graus de preferência entre os robôs para uma dada tarefa. Adotaremos $\sim$ significando “indiferente a” e $\succ$ significando “preferido a”. Formalmente também temos:

Definição 4.3 ([ZP04, KR76] Função-Valor Multiatributo). *Seja $X = \{X_1, \dots, X_n\}$ um conjunto de atributos, $D_i (1 \leq i \leq n)$ um conjunto de possíveis valores para o atributo X_i e D o espaço de todos os possíveis resultados, ou seja, o produto cartesiano $D_1 \times D_2 \times \dots \times D_n$. Uma função-valor é uma função v , que associa um número real $v(x)$ a cada vetor valor em D , representando a estrutura de preferência, dado que para todo $x', x'' \in D$:*

$$x' \sim x'' \iff v(x') = v(x'') \quad (4.47)$$

e

$$x' \succ x'' \iff v(x') > v(x'') \quad (4.48)$$

Estamos particularmente interessados em determinar sob quais condições temos:

$$(x_1, \dots, x_n) \succeq (y_1, \dots, y_n) \iff v(x_1, \dots, x_n) \geq v(y_1, \dots, y_n) \quad (4.49)$$

Para isto, tudo que precisamos é que as preferências sejam uma ordem fraca sobre os vetores de valores de atributos [FGE05].

Como $v(x), v(y), \dots$ ordenados por $>$ são consistentes com a ordem de $x, y, \dots$ sobre $\succ$, dizemos que v é uma função que preserva a ordem. Como quaisquer transformações monotônicas sobre v também preservam a ordem para esta relação binária de preferência, concluímos que trata-se de uma função ordinal (retorna números reais com o intuito de (apenas) classificar os possíveis resultados).

O desejável para uma função-valor é que ela possa ser descrita da maneira mais simples possível. Uma função do tipo

$$v(x_1, x_2, \dots, x_n) = f[v_1(x_1), v_2(x_2), \dots, v_n(x_n)] \quad (4.50)$$

onde v_i denota uma função-valor sobre um atributo X_i e f é, por exemplo, uma função aditiva, ou seja, para calcularmos o *valor* total, calculamos o *valor* em cada atributo separadamente e então somamos tudo:

$$v(x_1, \dots, x_n) = \sum_{i=1}^n v_i(x_i) \quad (4.51)$$

A condição chave para obtermos a função aditiva descrita acima é que o conjunto de atributos seja mutuamente preferencialmente independente. Formalmente:

Definição 4.4 ([ZP04, KR76] Independência Preferencial). *Seja X o conjunto de atributos e $Y \subset X$. Dizemos que Y é preferencialmente independente de seu complemento $\bar{Y}$ se as ordens de preferência das consequências envolvendo apenas mudanças nos valores dos atributos em Y , não dependem dos valores aos quais os atributos de $\bar{Y}$ são mantidos.*

Definição 4.5 ([KR76] Independência Preferencial Mútua). *Os atributos $X_1, \dots, X_n$ são mutuamente preferencialmente independentes, se cada subconjunto Y destes atributos é preferencialmente independente de seu conjunto complementar de atributos.*

Funções-valor ordinais aditivas são interessantes para aplicações práticas. Entretanto não são de fácil obtenção, pois durante o processo requerem escolhas entre dois atributos. Outro problema é que essas funções possuem interpretação ordinal, não permitindo medir a intensidade da preferência.

4.6.2 Modelo Cardinal Multiplicativo

Como vimos, o modelo ordinal de preferências não alcança o conceito de intensidade de preferências⁹. Nesta seção, com base na notação encontrada em [Dye05], vamos conceituar a inten-

⁹Do inglês “strength of preference”.

sidade de preferência para obtermos um modelo cardinal de funções-valor, alternativo ao modelo ordinal aditivo (multiplicativo e outras formas não-aditivas de funções valor multiatributo).

Definição 4.6 ([Dye05] Notação). *Seja X o conjunto de todos os possíveis resultados em um problema de decisão. Em um problema multiatributo $X = \prod_{i=1}^n X_i$, onde X_i é o conjunto dos possíveis resultados para o i -ésimo atributo. Utilizaremos as letras w, x, y e z para indicar elementos distintos de X . Exemplo: $w \in X$ é representado por $(w_1, \dots, w_n)$, onde w_i é um nível em um conjunto não-vazio de atributos X_i , para $i = 1, \dots, n$. Também seja $I \subset \{1, 2, \dots, n\}$ um subconjunto dos índices dos atributos, define-se X_I como o subconjunto de atributos designados pelos subscritos em I e seja $\bar{X}_I$ o representante do subconjunto complementar dos n atributos. Utilizaremos a notação $(w_i, \bar{w}_i)$ e $(x_i, \bar{w}_i)$ para denotar dois elementos de X que diferem somente no nível do i -ésimo atributo. $X^* = \{wx : w, x \in X\}$ denota um subconjunto não-vazio de $X \times X$ e $\succeq^*$ denota uma ordem fraca sobre X^* . Também, interpretaremos que $wx \succeq^* yz$ denota que a diferença de preferências entre w e x é superior ou igual a diferença de preferências entre y e z . Finalmente, se uma alternativa multiatributo é preferida a outra, diferindo apenas sobre o valor do atributo X_i , então a diferença de preferências entre esta alternativa e alguma referência comum será maior que a diferença entre a alternativa preterida e essa referência.*

Esse modelo baseia-se no conceito de independência fraca de diferenças. Para isso precisaremos de uma estrutura matemática que permita a comparação de diferenças de preferências par a par de alternativas.

Dizemos que X_I possui independência fraca de diferenças de $\bar{X}_I$ se as ordens de preferência das diferenças das consequências envolvendo apenas mudanças nos níveis dos atributos em X_I , não dependem dos níveis aos quais os atributos de $\bar{X}_I$ são mantidos. Formalmente:

Definição 4.7 ([Dye05] Independência Fraca de Diferenças). *O subconjunto de atributos X_I é fracamente diferencialmente independente de $\bar{X}_I$ se, dados quaisquer $w_I, x_I, y_I, z_I \in X_I$ e algum $\bar{w}_I$: se $(w_I, \bar{w}_I)(x_I, \bar{w}_I) \succeq^* (y_I, \bar{w}_I)(z_I, \bar{w}_I)$ então também será considerado $(w_I, \bar{x}_I)(x_I, \bar{x}_I) \succeq^* (y_I, \bar{x}_I)(z_I, \bar{x}_I)$, para qualquer $\bar{x}_I \in \bar{X}_I$.*

Ou seja, a ordem de preferência das diferenças depende apenas dos valores dos atributos X_I e não de valores fixos em $\bar{X}_I$.

Os atributos serão mutuamente fracamente diferencialmente independentes se todo subconjunto próprio desses atributos forem fracamente diferencialmente independentes de seus subconjuntos complementares. Sabemos que se os atributos forem mutuamente fracamente diferencialmente independentes, podemos obter uma função-valor multiatributo da forma:

$$1 + \lambda v(x) = \prod_{i=1}^n [1 + \lambda \lambda_i v_i(x_i)], \quad (4.52)$$

onde v_i é uma função-valor cardinal de atributo simples sobre X_i normalizada no intervalo de 0 a 1, λ_i são constantes de mudança de escala positivas e λ é uma constante de mudança de

escala adicional que quando assume valor 0 reduz a função-valor à forma aditiva:

$$v(x) = \sum_{i=1}^n \lambda_i v_i(x_i), \quad (4.53)$$

onde

$$\sum_{i=1}^n \lambda_i = 1.$$

4.6.3 Modelo Cardinal Aditivo

A condição que permitirá a obtenção de um modelo cardinal aditivo é a independência diferencial [Dye05]. O atributo X_i é diferencialmente independente de $\bar{X}_i$ se, para todo $w_i, x_i \in X_i$ tal que $(w_i, \bar{w}_i)(x_i, \bar{w}_i)$ para algum $\bar{w}_i \in \bar{X}_i$, $(w_i, \bar{w}_i)(x_i, \bar{w}_i) \sim^* (w_i, \bar{x}_i)(x_i, \bar{x}_i)$ para qualquer $\bar{x}_i \in \bar{X}_i$. Ou seja, a independência preferencial entre duas alternativas multiatributo diferindo apenas sobre um atributo não depende sobre valores comuns dos outros atributos. Novamente, os atributos serão mutuamente diferencialmente independentes se todos os subconjuntos próprios desses atributos forem diferencialmente independentes de seus subconjuntos complementares. Para o caso com três ou mais atributos, atributos mutuamente diferencialmente independentes são condição necessária (dentre outras) para garantir que se $wx, yz \in X^*$, então $wx \succeq^* yz$ se e somente se:

$$\sum_{i=1}^n \lambda_i v_i(w_i) - \sum_{i=1}^n \lambda_i v_i(x_i) = \sum_{i=1}^n \lambda_i v_i(y_i) - \sum_{i=1}^n \lambda_i v_i(z_i) \quad (4.54)$$

e $x \succeq y$ se e somente se

$$\sum_{i=1}^n \lambda_i v_i(x_i) = \sum_{i=1}^n \lambda_i v_i(y_i) \quad (4.55)$$

onde v_i é uma função-valor cardinal de atributo simples sobre X_i normalizadas no intervalo de 0 a 1 e $\sum_{i=1}^n \lambda_i = 1$.

4.7 Conclusão

Conforme pudemos observar durante este capítulo, há muitas alternativas para problemas de decisão. Não há um método que consiga lidar com todos os aspectos do problema, mas diversos métodos que lidam com o problema sob diversos ângulos. Como a própria escolha de um método é em si um problema de decisão, incluímos abaixo um diagrama, proposto por Hwang et.al. [HY81], para facilitar esse trabalho.

Figura 4.3: Escolha de método para decisão multiatributo.

A primeira questão indaga se apenas aceitamos a melhor resposta ou se aceitamos qualquer resposta desde que satisfaça um nível mínimo para os atributos. No segundo caso, se houver exigência de nível mínimo para apenas um atributo, usamos o *método disjuntivo de restrições*. Se a exigência for para todos os atributos, usamos o *método conjuntivo de restrições*.

A segunda questão trata da importância ou não de alternativas dominadas. Se alternativas dominadas não forem importantes para o problema, poderão ser excluídas através da *eliminação por dominância*.

A terceira questão é voltada para casos com mais de um tomador de decisão. Se nesses casos houver conflito de preferências, indica-se um método de restrições (conjuntivo ou disjuntivo), como alternativa para o tratamento do conflito.

A quarta questão indaga se a eficiência é determinada pelo melhor ou pior atributo. Se for pelo melhor atributo, usa-se *Maximax* e se for pelo pior, usa-se o *Maximin*.

A quinta questão indaga se estamos modelando o problema como um problema de seleção ou como um problema de design (conforme explicado na introdução da Seção 4.2).

Finalmente as questões seis e sete tratarão sobre o tipo de escala para sua função-valor.

Capítulo 5

Trabalhos Futuros

Além de grande fonte de possíveis incorreções, a determinação de uma função-valor também é terreno fértil e (em alguns casos¹) pouco explorado de trabalho. A seguir citaremos possíveis continuações para o trabalho aqui realizado, observando que algumas das sugestões aqui apresentadas são aplicáveis também à teoria da decisão, ultrapassando, portanto, as fronteiras dos sistemas robóticos. Em seguida citaremos alguns temas que poderiam ser desenvolvidos nesse trabalho.

- Uma das etapas iniciais e das mais importantes para se determinar uma função-valor é a determinação do conjunto de atributos. Se solicitarmos a um grupo de pessoas um conjunto de atributos para um determinado problema, é provável que entre as respostas haja poucas coincidências. Além de envolver subjetividade, essa tarefa também envolve cuidadoso estudo e bom entendimento do problema em questão: (1) atributos redundantes podem ser incluídos na função, pois de forma quase imperceptível, apresentavam-se distintos e necessários, quando na verdade eram “variações sobre um mesmo tema”, (2) importantes aspectos do problema podem ser deixados de lado, levando à determinação de uma função-valor incompleta etc. Poucos autores tratam esse tema (Raiffa é um deles [KR76]) e, portanto, há uma avenida de pesquisa aqui.
- Para funções-valor que envolvam muitos atributos, e de grande complexidade, como no caso de um robô que tome decisões sobre o mercado de renda variável, por exemplo, uma alternativa é a mudança de paradigma. Ao invés de procurarmos por uma função valor que descreva de forma ótima e completa a bolsa de valores (se é que isso é possível, dado o dinamismo da bolsa de valores), procurarmos por modelos de sucesso. Considere, por exemplo, um algoritmo de aprendizado, em redes neurais, que seja suficientemente expressivo (que consiga modelar problemas não linearmente separáveis²). Ao invés de tentarmos modelar a bolsa de valores, podemos tentar modelar um investidor de sucesso, como por exemplo

¹Sistemas robóticos, por exemplo.

²Técnicas como Perceptron multi-camada, Madaline (multi-Adaline) etc.

Warren Buffett³. Nesse caso abrimos mão do tratamento “função-valor ótima e completa” para um tratamento que não é ótimo nem completo, mas muito satisfatório. Resta saber se o problema em questão permite tal tratamento.

- Sabemos que para que a alocação de tarefas em um sistema robótico seja viável, deve ser possível comparar os robôs entre si. Como foi visto anteriormente, essa comparação pode ser feita através de valores gerados por uma função-valor, previamente formulada. Vimos que uma suposta função-valor ideal deve ser capaz de exprimir com exatidão as habilidades relativas dos robôs em relação a tarefa. Também pudemos—ao longo desse trabalho—observar diversos problemas que dificultam, senão inviabilizam, a obtenção de funções-valor ideais.

Acreditamos funções-valor mais expressivas est de uma função-valor a aproxima de uma suposta função-valor ideal. Também postulamos que

Na tentativa de obtermos funções-valor mais próximas do ideal funções-valor

A inclusão de atributos baseados no plano do robô visa obter funções-valor mais expressivas e que denotem aspectos dinâmicos do sistema. Nesta dissertação apresentei algumas idéias iniciais sobre o tema, que devem ser postas a prova através de experimentos envolvendo robôs situados. Muitas modificações sobre minhas idéias, bem como, muitas outras (novas) idéias são possíveis, dada a quase ausência de funções-valor que incluam atributos baseados no plano, na pesquisa em sistemas robóticos. O sistema M+ [BA99, BA98, BA00b, BA00a]. é uma (bem-vinda) exceção. Logo, há outra avenida de pesquisa aqui.

A confirmação ou refutação (formal) de que funções-valor tornam-se mais expressivas com a inclusão de atributos referentes aos planos dos robôs.

Para que essa estimativa seja coerente, todos os robôs devem ser “avaliados” sob uma mesma métrica. Dois exemplos de métricas encontradas em alocação de tarefas em sistemas robóticos são (1) *A estimativa da aptidão* do robô para cada tarefa: vai desde uma aptidão nata (sua velocidade de locomoção, o custo para efetuar a locomoção, comunicação etc.) até seu status corrente (a que distância encontra-se o robô do alvo, se está ou não impossibilitado de sair de uma sala etc.) e (2) *O comprimento simples do plano* gerado pelo robô. Esta técnica pode ser encontrada no sistema M+ [BA99, BA98, BA00b, BA00a].

Com freqüência, em sistemas robóticos, aspectos inerentes aos robôs são incorporados às funções-valor, como por exemplo a velocidade e o custo de operação do robô. Muito mais raros são os sistemas que incorporam aspectos referentes ao dinamismo do domínio, sejam eles descritivos (por exemplo, a posição do robô) ou contingenciais (por exemplo, robô funcionando com algum tipo de problema em algum subsistema). CONTINUA

³ Acreditamos que Warren Buffett seja um bom modelo, dado que é o segundo homem mais rico do mundo (em 1986 era o homem mais rico do mundo)—condição obtida graças ao seu talento como investidor no mercado de renda variável de longo prazo.

Capítulo 6

Conclusão

Esta dissertação teve como objetivos salientar a importância da função-valor para alocação em sistemas robóticos, bem como propor melhorias e alternativas a suas formulações mais comuns. Foi salientado que de nada adianta um algoritmo de alocação ótimo e de complexidade de tempo e espaço polinomial, que rode sobre dados gerados por uma função-valor incorreta.

Em seguida foram evidenciados e ilustrados importantes aspectos causadores de incorreções em funções-valor. Foram estudados aspectos pertinentes à teoria da medição, mais especificamente, aspectos relacionados a escalas, que podem causar distorções (1) sobre a agregação e normalização de atributos numa função-valor, (2) sobre o processo de alocação e (3) de interpretação, ou seja, do significado denotado pelo atributo propriamente dito. Também foram estudados importantes aspectos e modelos da teoria da decisão, tais como (1) a tarefa inicial—e das mais difíceis e importantes de todo o processo—de se determinar um conjunto de atributos relevantes que melhor expresse o problema em questão, (2) a determinação dos pesos dos atributos, ou seja, do grau de importância relativa entre os atributos numa função-valor e (3) métodos da decisão multiatributo. Alguns desses métodos são indicados para casos específicos (variações do mesmo problema) e outros são diferentes modos de se lidar com o mesmo problema.

6.1 Contribuições Originais

Um tema raramente discutido e aprofundado em sistemas robóticos é o da formulação de funções-valor. A formulação de funções-valor está, geralmente, relacionada a alocação de tarefas em sistemas robóticos. O presente trabalho teve como contribuição original discutir a formulação de funções-valor para sistemas robóticos sob o ponto de vista da teoria da decisão e da teoria da medição. Da teoria da decisão adaptamos observações sobre a determinação de um conjunto de atributos relevante, para sistemas robóticos. Vários modelos da teoria da decisão também foram apresentados e utilizados no contexto de um sistema robótico.

Da teoria da medição ressaltamos alguns de seus importantes aspectos formais, tais como

escalas, significância etc., relevantes tanto para a formulação da função-valor em um sistema robótico, bem como para a alocação propriamente dita. Esse trabalho também contribui originalmente com a adaptação do problema do perseguidor-fugitivo para sua formulação alterada.

Referências Bibliográficas

- [BA98] S. Botelho and R. Alami. A distributed scheme for task planning and negotiation in multi-robot systems. In *European Conference in Artificial Intelligence, ECAI98*, number 98416, pages 516–517, Brighton (GB), 1998. 80
- [BA99] S. Botelho and R. Alami. M+: A scheme for multi-robot cooperation through negotiated task allocation and achievement. In *Proc. IEEE International Conference on Robotics and Automation (ICRA '1999)*, pages 1234–1239, Detroit, Michigan, May 1999. 80
- [BA00a] S. Botelho and R. Alami. A multi-robot cooperative task achievement system. In *IEEE International Conference on Robotics and Automation (ICRA '2000)*, volume 3, pages 2716–2721, San Francisco (USA), April 2000. 80
- [BA00b] S. Botelho and R. Alami. Robots that cooperatively enhance their plans. In Lynne E. Parker, George A. Bekey, and Jacob Barhen, editors, *5th International Symposium on Distributed Autonomous Robotic Systems (DARS)*, pages 55–68, Knoxville, Tennessee, USA, 2000. Springer. 80
- [Cab91] M.A.P. Cabral. Álgebra: Um Guia de Estudo. World Wide Web, 1991. <http://www.dma.im.ufrj.br/~mcabral/textos/algebra.pdf>—Último acesso em 22/10/2006. 29
- [Che03] P. Cheng. A short survey on pursuit-evasion games. Technical report, Dept. of Computer Science, University of Illinois at Urbana-Champaign, 2003. 3, 5
- [Chv75] V. Chvátal. A combinatorial theorem in plane geometry. *Journal of Computational Theory (B)*, (41):18:39, 1975. 3, 5
- [Dag97] T. Dagnino. The art gallery problem. World Wide Web Project: Computational Geometry—Submitted to G. Toussaint—School of Computer Science, McGill University, Canada, December 1997. <http://www-cgri.cs.mcgill.ca/~godfried/teaching/cg-projects/97/Thierry/thierry507webprj/artgallery.html>—Último acesso em 21/10/2006. 7, 9, 13

- [Dye05] J.S. Dyer. *Multiple Criteria Decision Analysis: State of the Art Surveys*, volume 78 of *International Series in Operations Research & Management Science*, chapter 7. Springer Media B.V., New York, 2005. 75, 76, 77
- [FGE05] J. Figueira, S. Greco, and M. Ehrgott, editors. *Multiple Criteria Decision Analysis: State of the Art Surveys*, volume 78 of *International Series in Operations Research & Management Science*. Springer Media B.V., New York, 2005. 75
- [Ger03] B.P. Gerkey. *On Multi-Robot Task Allocation*. Phd dissertation, University of Southern California Computer Science Department, August 2003. Also Technical Report CRES-03-012. 1
- [GLL⁺99] L.J. Guibas, J.C. Latombe, S.M. LaValle, D. Lin, and R. Motwani. A visibility-based pursuit-evasion problem. *International Journal of Computational Geometry and Applications*, 9(4/5):471–493, 1999. 9
- [HK00] Jiawei Han and Micheline Kamber. *Data Mining: Concepts and Techniques*. Morgan Kaufmann, 2000. 3
- [HY81] C.L. Hwang and K. Yoon. Multiple attribute decision making — methods and applications — a state-of-the-art survey. *Lecture Notes in Economics and Mathematical Systems*, 186:1–259, 1981. 4, 15, 20, 49, 52, 53, 57, 58, 59, 60, 62, 63, 65, 73, 74, 77
- [IKK04] V. Isler, S. Kannan, and S. Khanna. Locating and capturing an evader in a polygonal environment. In *Sixth International Workshop on the Algorithmic Foundations of Robotics*, 2004. 5, 8, 11, 13
- [Isl04] V. Isler. Randomized visibility-based pursuit-evasion applet. World Wide Web, 2004. <http://www.cis.upenn.edu/~isleri/research/PEapplet/>—Último acesso em 22/10/2006. 13
- [Joh03] E. Johnson. A survey of mobile guards in art galleries. World Wide Web, 2003. Texas A&M University – Computer Science Department. <http://eric.gruver.net/ArtGalleryProblems.html>—último acesso em 22/10/2006. 3, 5
- [Kar05] T. Kardi. Normalization. World Wide Web, 2005. <http://people.revoledu.com/kardi/tutorial/Similarity/Normalization.html>—Último acesso em 22/10/2006. 43
- [KR76] R.L. Keeney and H. Raiffa. *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. Wiley series in probability and mathematical statistics. John Wiley & Sons, New York, 1976. 16, 19, 20, 74, 75, 79
- [LKST71] R.D. Luce, D.H. Krantz, P. Suppes, and A. Tversky. *Foundations of Measurement*, Vol. 1: *Additive and Polynomial Representations*. Academic Press, New York, 1971. 28

- [LKST89] R.D. Luce, D.H. Krantz, P. Suppes, and A. Tversky. *Foundations of Measurement* Vol. 2: *Geometrical, Threshold and Probabilistic Representations*. Academic Press, New York, 1989. 28
- [LKST90] R.D. Luce, D.H. Krantz, P. Suppes, and A. Tversky. *Foundations of Measurement*, Vol. 3: *Representation, Axiomatisation and Invariance*. Academic Press, New York, 1990. 28
- [Nar85] L. Narens. *Abstract Measurement Theory*. MIT Press, Cambridge, MA, 1985. 28
- [O'R98] J. O'Rourke. *Computational Geometry in C*. Cambridge University Press, New York, NY, USA, second edition, 1998. 6, 9, 10, 11, 12
- [Pek96] A. Pekec. *Limitations on conclusions from combinatorial optimization models*. PhD thesis, Rutgers University, 1996. 28, 44, 46, 47, 48
- [Pfa71] J. Pfanzagl. *Theory of Measurement*. Physica Verlag, Würzburg, 2nd revised edition, 1971. 28
- [PLC01] S.M. Park, J.H. Lee, and K.Y. Chwa. Visibility-based pursuit-evasion in a polygonal region by a searcher. *Lecture Notes in Computer Science*, 2076:456–468, 2001. 8, 9
- [Rob79] F.S. Roberts. *Measurement Theory with Applications to Decisionmaking, Utility, and the Social Sciences*, volume 7 of *Encyclopedia of Mathematics and its Applications*. Addison-Wesley, 1979. VI, 19, 28, 30, 31, 33, 34, 35, 37, 38, 39, 41
- [Say96] K. Sayood. *Introduction to Data Compression*. Morgan Kaufmann Publishers, 2929 Campus Drive, Suite 260, San Mateo, CA 94403, USA, 1996. 50, 51
- [Ste46] S. S. Stevens. On the theory of scales of measurement. *Science*, (103):677–680, 1946. 31
- [ZP04] J. Zhang and P. Pu. Survey of solving multi-attribute decision problems. Technical report, School of Computer and Communication Sciences, Swiss Federal Institute of Technology, Lausanne, 2004. 74, 75