**UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE FÍSICA DE SÃO CARLOS
DEPARTAMENTO DE FÍSICA E INFORMÁTICA**

**"Explicação em Sistemas que Utilizam Diagramas de Influências como
Formalismo de Representação do Conhecimento".**

MARIA INES CASTINEIRA

Tese apresentada ao Instituto de Física de São Carlos, Universidade de São Paulo, para obtenção do título de Doutor em Ciências "Física Aplicada".

Orientadora: Profa.Dra. Maria Carolina Monard

el

SÃO CARLOS

SÃO PAULO

1996

MEMBROS DA COMISSÃO JULGADORA DA TESE DE DOUTORADO DE **MARIA INES CASTINEIRA** APRESENTADA AO INSTITUTO DE FÍSICA DE SÃO CARLOS, UNIVERSIDADE DE SÃO PAULO, EM 18/10/1996.

COMISSÃO JULGADORA:

Profa. Dra. Maria Carolina Monard/ICMSC-USP

Profa. Dra. Maria das Graças Volpe Nunes/ICMSC-USP

Prof. Dr. Marcio Luiz de Andrade Neto/UNICAMP

Prof. Dr. Marcio Rillo/EP-USP

Prof. Dr. Osvaldo Novais de Oliveira Junior/IFSC-USP

A meus amigos.

Às mulheres que lutam para conciliar família e profissão.

Agradecimentos

À Professora Maria Carolina Monard pela amizade, incentivo e orientação que fizeram possível a realização deste trabalho.

Ao Professor Wolfgang Meyer do Departamento Prozeßautomatisierungstechnik da Technische Universität Hamburg-Harburg e seu grupo de apoio e pesquisa pela acolhida e atenção a mim dispensadas.

Ao Professor Jan W. Slaets pela confiança e o apoio dispensado nas primeiras horas.

A João B.S. de Oliveira, pela amizade, e pela paciência na leitura e discussão deste trabalho.

A Dachamir Hotza pela cuidadosa e paciente leitura dos primeiros relatórios.

A Solange Rezende, Cesar e Seisa Zuccala pela amizade e apoio.

Aos amigos de Hamburgo que fizeram nossa estadia na Alemanha mais cálida.

A Daniel e meus filhos Cecília e Gabriel pelo incentivo e pelas alegrias do dia a dia.

Àqueles que de uma ou outra forma contribuíram para a realização deste trabalho.

Ao CNPq pelo apoio financeiro.

DESPEDIDA

Finalmente ele disse-lhe
"Tchau querida ..."
e a sua voz soou
tão suave, tão doce, tão funda
que parecia estar abraçando-a
com as palavras.

Elas ficaram ...
junto aos beijos roubados
no vermelho dos sinais.

M. Quinteiro

NO TE SALVES

No te quedes inmóvil
al borde del camino
no congeles el júbilo
no quieras con desgana
no te salves ahora
ni nunca
no te salves
no te llenes de calma
no reserves del mundo
sólo un rincón tranquilo
no dejes caer los párpados
pesados como juicios
no te quedes sin labios
no te duermas sin sueño
no te pienses sin sangre
no te juzgues sin tiempo

pero si
pese a todo
no puedes evitarlo
y congelas el júbilo
y quieres con desgana
y te salvas ahora
y te llenas de calma
y reservas del mundo
sólo un lugar tranquilo
y dejas caer los párpados
pesados como juicios
y te secas sin labios
y te duermes sin sueño
y te piensas sin sangre
y te juzgas sin tiempo
y te quedas inmóvil
al borde del camino
y te salvas
entonces
no te quedes conmigo.

M. Benedetti

Resumo

O presente trabalho discute a necessidade da representação e manipulação de incertezas na resolução de problemas por sistemas baseados em conhecimento, e como isto pode ser realizado utilizando redes de crenças. Este tipo de representação do conhecimento combina a teoria das probabilidades e teoria da decisão, para representar incertezas, com a teoria dos grafos, esta última apropriada para representar as relações de dependência entre as variáveis do modelo. Os diagramas de inferência — redes de crenças que permitem representar incertezas, decisões e preferências do usuário — são discutidos e adotados neste trabalho para desenvolver um sistema normativo de apoio à decisão.

A problemática da explicação em sistemas bayesianos, relativamente nova quando comparada com a dos sistemas baseados em regras, é abordada. Neste contexto dois mecanismos de explicação para diagramas de influências são propostos: análise de sensibilidade e as redes probabilísticas qualitativas. Estes mecanismos são usados para gerar conclusões genéricas bem como para entender qualitativamente as relações entre as ações e eventos que fazem parte do modelo.

Uma ferramenta gráfica de apoio à decisão baseada em diagramas de influências foi implementada na linguagem Smalltalk. Este aplicativo não só permite representar e avaliar o problema do usuário como também incorpora as facilidades de explicação acima descritas. A possibilidade de observar graficamente o que acontece com o modelo quando os valores das variáveis são modificados — análise de sensibilidade — permite compreender melhor o problema descobrindo quais as variáveis que influenciam as decisões e auxilia a refinar os valores das variáveis envolvidas. Por outro lado as redes probabilísticas qualitativas permitem realizar abstrações e simplificações apropriadas do modelo, i.e., obter as relações qualitativas do modelo a partir do seu nível quantitativo. As conclusões genéricas obtidas servem tanto para limitar o espaço da estratégia ótima quanto para entender qualitativamente as relações entre as ações e eventos que fazem parte do modelo.

Abstract

This work discusses the knowledge representation and uncertainty handling of knowledge based systems that use belief networks for this purpose. This sort of networks combine the theory of probability and decision theory –to represent uncertainty– with graph theory to represent the dependence relations between the model parameters. Systems that use belief networks as knowledge representation are named Bayesian or normative systems. This work investigates and adopts influence diagrams –belief networks that represent uncertainty, decisions and preferences– to develop a normative decision support system.

Comprehensible explanations for probabilistic reasoning systems are a prerequisite for wider acceptance of Bayesian methods. Two schemes for explaining influence diagrams are proposed: sensitivity analysis and qualitative probabilistic networks, aiming to find general conclusions and to qualitatively understand the relations between the actions and events of the model.

A graphical decision support system that represents the user problem as influence diagrams has been implemented in Smalltalk. This system allows to represent and evaluate decision problems and incorporates the explanation facilities mentioned above. The possibility to observe graphically the model as the variables change –sensitivity analysis– permits a better understanding of the problem by finding the significant variables. This process also helps to adjust the variables values. Furthermore, the qualitative probabilistic networks allow to realise model abstractions and simplifications, i.e., to obtain the qualitative relations from the quantitative level. These general conclusions limit the optimal strategy space and allow to qualitatively understand the relations between actions and events in the model.

CONTEÚDO

1 INTRODUÇÃO	1
1.1 Considerações Iniciais	1
1.2 Objetivos da Pesquisa Desenvolvida	2
1.3 Organização do Trabalho	2
2 RACIOCÍNIO PROBABILÍSTICO EM REDES	4
2.1 Introdução	4
2.2 Sistemas Baseados em Conhecimento	4
2.3 Necessidade de Raciocínio Incerto	6
2.4 O porquê usar Raciocínio Probabilístico em SBC	6
2.4.1 A abordagem baseada em regras não é modular para domínios incertos	7
2.4.2 Representação do conhecimento na direção causa-efeito	7
2.4.3 Inferências bidirecionais	8
2.5 Diferentes abordagens no tratamento das incertezas	8
2.6 Raciocínio Probabilístico em redes	10
2.7 Considerações finais	11
3 DECISÃO E INFERÊNCIA	12
3.1 Introdução	12
3.2 Teoria da Decisão/Análise da Decisão	13
3.3 Conceitos Básicos da Teoria da Decisão	14
3.3.1 Critério do Valor Monetário Esperado (VME)	15
3.3.2 Função de Utilidade	16
3.3.3 Determinação da Função de Utilidade	17
3.3.4 Função de Utilidade do Dinheiro	19
3.3.5 Maximização da Utilidade Esperada	19
3.4 Métodos Gráficos na Teoria da Decisão: Árvores de Decisão e Diagramas de Influências	21
3.5 Árvores de Decisão	21
3.6 Considerações Finais	27

4 DECISÃO E INFERÊNCIA: DIAGRAMAS DE INFLUÊNCIA.....	31
4.1 Introdução	31
4.2 Representação do Diagrama de Influência	31
4.3 Definição de um DI.....	35
4.4 Avaliação do Diagrama de Influência	37
4.4.1 Mudança da Direção de um Arco Condicional	38
4.4.2 Absorção de um Nó Probabilístico.....	38
4.4.3 Absorção de um Nó de Decisão	39
4.5 Outros Algoritmos de Avaliação	39
4.6 Considerações Finais.....	41
5 FERRAMENTAS DE APOIO À DECISÃO.....	44
5.1 Introdução	44
5.2 Sistemas de Software que Usam a Abordagem Bayesiana.....	44
5.3 Algoritmos Especializados	45
5.4 Extensões dos Diagramas de Influências.....	45
5.5 Ferramentas	46
5.5.1 Glossário	49
5.6 Aplicações	50
5.7 Considerações Finais.....	51
6 EXPLICAÇÕES EM SAD: UMA PROPOSTA.....	52
6.1 Introdução	52
6.2 Explicação em Sistemas baseados em Regras	53
6.3 Explicação em Redes Bayesianas: Outras Metodologias	55
6.4 Proposta de um Esquema de Explicação para DIs	57
6.4.1 Análise de Sensibilidade.....	58
6.4.2 Redes Probabilísticas Qualitativas	60
6.5 Considerações Finais.....	63
7 UMA FERRAMENTA DE APOIO À DECISÃO	65
7.1 Introdução	65
7.2 A linguagem Smalltalk	65
7.3 Etapas do Desenvolvimento	66
7.3.1 Análise	66
7.3.2 Projeto	68
7.3.3 Implementação	68

7.4	Descrição da Ferramenta.....	69
7.4.1	Descrição Geral.....	69
7.4.2	Facilidades de Explicação	72
7.5	Considerações finais.....	74
8	CONCLUSÕES.....	75
8.1	Considerações Gerais	75
8.2	Conclusões	76
8.3	Sugestões para futuros trabalhos.....	77
	LISTA DE ABREVIACÕES.....	79
	REFERÊNCIAS BIBLIOGRÁFICAS	80

LISTA DE FIGURAS

Figura 2.1 Estrutura básica de um SBC.....	5
Figura 2.2 Sistemas Baseados em Conhecimento e Sistemas Especialistas.....	6
Figura 3.1 Exemplo de uma função de utilidade do dinheiro.....	19
Figura 3.2 Modelos gráficos de decisão: a) Árvore de decisão; b) Diagrama de influência.....	21
Figura 3.3 Nó de decisão	24
Figura 3.4 Um dos primeiros nós de eventos.....	24
Figura 3.5 Estrutura da árvore de decisão.....	25
Figura 3.6 Árvore de decisão com probabilidades e utilidades.....	26
Figura 3.7 Árvore de decisão dobrada.....	26
Figura 4.1 Diagrama de influência genérico.....	32
Figura 4.2 Nós de probabilidades.....	32
Figura 4.3 Diagrama de influência com três nós	33
Figura 4.4 Nós de probabilidades do exemplo considerado.....	34
Figura 4.5 Diagrama de influência parcial do exemplo 4.1.....	34
Figura 4.6 Novo diagrama de influência parcial do exemplo 4.1.....	34
Figura 4.7 Diagrama de influência do exemplo.....	35
Figura 4.8 Diagrama de influência com os arcos invertidos.....	38
Figura 4.9 Inversão de um arco num diagrama de influência.....	38
Figura 4.10 Absorção do nó $n(A)$	39
Figura 4.11 Absorção do nó $n(B)$	39
Figura 4.12 Absorção do nó $n(d2)$	39
Figura 4.13 Redes: a) simplesmente conexa, b) multiconexa.....	39
Figura 6.1 Interface gráfica do sistema IDEX mostrando o exemplo 6.1.....	58
Figura 6.2 Análise de sensibilidade para a variável Resultado do Teste. As duas curvas representam duas possíveis estratégias. Se $P(\text{Resultado do Teste}) < 0.5$ a estratégia ótima é ~ Teste & Perfuração, senão Teste & Perfuração.....	59
Figura 6.3 Modelo de decisão genérico.....	62
Figura 6.4 Modelo de decisão genérico reduzido a três variáveis.....	62

Figura 7.1 Diagrama de classes para a interface gráfica segundo a metodologia de Booch [Booch-91].....	67
Figura 7.2 Diagrama de classes para o modelo de decisão utilizado.....	67
Figura 7.3 Arquitetura da ferramenta desenvolvida segundo as categorias de classes.	68
Figura 7.4 Interface da ferramenta mostrando um exemplo	70
Figura 7.5 Items dos principais menus da ferramenta.....	71
Figura 7.6 Análise de sensibilidade para uma variável do modelo.....	72
Figura 7.7 Rede probabilística qualitativa de um exemplo.....	73
Figura 7.8 Rede probabilística qualitativa reduzida.....	74

LISTA DE TABELAS

Tabela 3.1 Retornos do investidor.....	16
Tabela 3.2 Retornos milionários.....	17
Tabela 3.3 Tabela de utilidades do investidor.	20
Tabela 3.4 Probabilidades dos eventos de interesse no exemplo 3.4.....	23
Tabela 3.5 Retornos esperados em \$ milhões.	24
Tabela 5.1 Ferramentas de livre acesso.	47
Tabela 5.2 Ferramentas comerciais.....	48
Tabela 7.1 Avaliação das características das linguagens de programação	66

CAPÍTULO 1

INTRODUÇÃO

1.1 Considerações Iniciais

Um dos objetivos da pesquisa na área de Inteligência Artificial –IA– é a resolução de problemas utilizando o conhecimento do domínio de aplicação. Isto é realizado através dos Sistemas Baseados em Conhecimento (SBCs) ou Sistemas Especialistas (SEs). Estes últimos são assim denominados por tentar modelar o desempenho do especialista numa determinada área de aplicação utilizando o seu conhecimento sobre o domínio de aplicação. Os procedimentos de decisão do especialista são analisados e transformados em regras ou algum outro formalismo de representação o qual é implementado no sistema.

À medida que a pesquisa nesta área foi evoluindo de pequenos problemas para problemas complexos do mundo real, novos desafios a serem vencidos foram aparecendo. Uma destas questões é tratar explicitamente as incertezas do problema devidas a informação parcial e modelos incompletos. Em alguns domínios a incerteza é um componente inseparável do processo de decisão associado à resolução do problema. Ela pode ser originada por desconhecimento das variáveis intrínsecas, por medições ou valores imprecisos ou ausentes, ou por relações causa-efeito não determinísticas.

A fim de representar e tratar as incertezas do domínio diversas abordagens foram empregadas nos SBCs. Dependendo do domínio, a capacidade de manipular incertezas ainda é um dos problemas em aberto deste tipo de sistemas.

Para resolver essa questão, uma abordagem relativamente nova é a de modelar os problemas da área de aplicação com suas incertezas intrínsecas, em lugar de modelar a forma em que o especialista manipula as incertezas. Isto pode ser realizado empregando a teoria das probabilidades como linguagem de representação formal do problema de aplicação, sendo assim possível calcular o impacto que certa informação exerce sobre as condições decisórias. Na prática, porém, estes cálculos revelam-se de difícil resolução ou de complexidade de ordem exponencial. Com o advento das redes de crenças –também denominadas redes probabilísticas– esta abordagem é hoje viável e amplamente utilizada, sendo apropriada tanto na modelagem quanto para o cálculo associado ao processo de inferência.

Outro fator importante a ser considerado em problemas decisórios, além das incertezas, é a diferença nas predileções de cada pessoa frente a uma determinada situação. Em outras palavras, aplicações complexas também revelam a importância de modelar as preferências humanas e atitudes contra o risco. Diagramas de Influências –DIs– são as

redes de crenças que também permitem representar decisões e preferências além das incertezas do problema. Estes diagramas têm sido utilizados extensamente e com sucesso na última década. Diversos sistemas especialistas têm sido construídos empregando redes de crenças como formalismo de representação.

Um motivo fundamental na construção de um modelo matemático ou estatístico é resolver e explicar problemas complexos. Conseqüentemente, explicações e descrições dos mecanismos que constituem o modelo formam uma parte importante da validação, exploração e uso do modelo. Estudos prévios realizados em sistemas especialistas indicam a necessidade de explicações detalhadas. Os usuários de SEs acham que entender o porquê de uma particular decisão ou conselho gerado pelo sistema é tanto ou mais importante quanto a conclusão atingida. Para uma recomendação ser aceita pelo usuário, o sistema de apoio à decisão deve ser capaz de mostrar o modelo de decisão e fornecer uma justificativa para o conselho gerado. A problemática da explicação em sistemas bayesianos –sistemas baseados em redes de crenças– é relativamente nova quando comparada com a dos sistemas baseados em regras. São poucos os sistemas bayesianos que incorporam esta característica. A explicação da linha de raciocínio do sistema dentro dessa abordagem é atualmente objeto de pesquisa na área.

1.2 Objetivos da Pesquisa Desenvolvida

O objetivo deste trabalho é desenvolver facilidades de explicação para sistemas de apoio à decisão que utilizam diagramas de influências como representação do conhecimento. Com tal finalidade foi investigado o raciocínio probabilístico baseado em redes. Os conceitos básicos da teoria da decisão, assim como dos Diagramas de Influências, foram analisados. Ferramentas de apoio à decisão baseadas neste paradigma foram comparadas analiticamente, permitindo verificar uma carência generalizada a respeito de facilidades de explicação por parte dessas ferramentas.

Este trabalho apresenta um esquema de explicação para sistemas de apoio à decisão baseados em DIs –também denominados redes bayesianas de decisão. Este esquema foi implementado como parte de uma ferramenta genérica para representar gráfica e numericamente o problema decisório como um DI, o qual pode ser automaticamente avaliado apresentando ao usuário as decisões ótimas. Esta ferramenta serve tanto para representar como para refinar o problema do usuário. A ênfase deste trabalho é dada na abordagem de explicação –fundamentada na análise de sensibilidade e redes probabilísticas qualitativas– a qual fornece métricas para julgar a importância de cada variável no modelo bem como das relações entre estas, apresentando as conclusões de uma forma gráfica ou através de descrições verbais.

1.3 Organização do Trabalho

O trabalho está organizado da seguinte forma:

No capítulo 2 é discutida a necessidade de representar e manipular incertezas em sistemas baseados em conhecimento, isto é, a necessidade de raciocínio incerto. Para resolver esta questão as redes de crenças são apresentadas como um formalismo que utiliza a teoria das probabilidades em redes como linguagem de representação do

problema. Outras metodologias empregadas para o tratamento das incertezas são também apresentadas.

No capítulo 3 é apresentada a teoria da decisão, a qual estende a linguagem das probabilidades para permitir realizar afirmações sobre as decisões e o quanto os correspondentes resultados são avaliados ou desejados. Dois métodos gráficos de representação são utilizados nesta área: árvores de decisão e diagramas de influências. As árvores de decisão são descritas e exemplificadas neste capítulo.

No capítulo 4 são descritos os diagramas de influências, uma representação esquemática do problema decisório, que consiste de um grafo orientado que descreve a estrutura do modelo. Esta abordagem permite representar problemas de maior envergadura do que as árvores de decisão, apresentando a vantagem de poder ser diretamente utilizada como o mecanismo de representação do conhecimento do sistema especialista.

O capítulo 5 apresenta diversas ferramentas para processar modelos gráficos de probabilidades em redes, como por exemplo redes de Bayes e diagramas de influências, bem como alguns sistemas implementados para aplicações em domínios específicos. Através do exposto neste capítulo pode ser evidenciada a maturidade das ferramentas de apoio à decisão baseadas nesta abordagem assim como algumas deficiências na parte de explicação e justificativa ao usuário..

O capítulo 6 apresenta a base da proposta do esquema de explicação para sistemas baseados em diagramas de influências, a saber: análise de sensibilidade e redes probabilísticas qualitativas. Outras metodologias desenvolvidas para oferecer explicações em sistemas bayesianos também são apresentadas.

No capítulo 7 é descrito o projeto e a implementação de uma ferramenta gráfica para a representação e avaliação de redes bayesianas de decisão assim como dos mecanismos de explicação discutidos no capítulo 6.

No capítulo 8 são apresentadas as conclusões, questões em aberto e algumas sugestões para trabalhos futuros.

CAPÍTULO 2

RACIOCÍNIO PROBABILÍSTICO EM REDES

2.1 Introdução

Este capítulo discute a necessidade de representar e manipular incertezas em sistemas baseados em conhecimento, i.e., a necessidade de realizar inferências em domínios onde os dados são imprecisos —procedimento também denominado *Raciocínio Incerto*.

Existem domínios ou aplicações onde incerteza é parte inerente do problema devido a dados ausentes ou imprecisos, a relações causa-efeito não determinísticas ou também em aplicações de prognóstico. Esta característica deve ser apropriadamente considerada pelo sistema baseado em conhecimento, embora abordagens tradicionais utilizadas na área de IA para realizar raciocínio em domínios incertos apresentem diversas restrições. Para resolver esta questão a metodologia descrita e adotada neste trabalho utiliza a teoria das probabilidades em redes como linguagem de representação do problema —também denominada *raciocínio em redes de Bayes, redes de crença ou redes probabilísticas*. Algumas das vantagens de se utilizar esta abordagem serão aqui discutidas, assim como uma breve descrição de outras metodologias.

O capítulo está organizado da seguinte forma: a seção 2.2 introduz os sistemas baseados em conhecimento, a seção 2.3 discute a necessidade de se utilizar raciocínio incerto em sistemas baseados em conhecimento, enquanto a seção 2.4 aborda algumas das vantagens de se utilizar raciocínio probabilístico para domínios incertos. Outras metodologias utilizadas para o tratamento das incertezas são descritas na seção 2.5 e o raciocínio probabilístico em redes é discutido na seção 2.6. Finalmente a seção 2.7 compreende as considerações finais.

2.2 Sistemas Baseados em Conhecimento

No contexto deste trabalho, o termo *Sistemas Baseados em Conhecimento* —SBC— denomina sistemas computacionais que utilizam conhecimento (representado explicitamente) para resolver problemas que geralmente requereriam inteligência humana [Hayes-Roth-92]. Assim, a representação de conhecimento e o processo de resolução de problemas são pontos centrais na estrutura de um SBC. Como mostrado na figura 2.1, a estrutura básica destes sistemas pode-se dividir em três módulos principais:

1. *Base de Conhecimento (BC)*: contém o conhecimento específico do domínio da aplicação, i.e., fatos e regras que descrevem as relações das variáveis do problema, bem como métodos e heurísticas para a resolução de problemas no domínio
2. *Motor de Inferência (MI)*: mecanismo responsável pelo processamento do conhecimento, usando para isso uma determinada linha de raciocínio. Implementa as estratégias de inferência e controle.
3. *Interface com o Usuário*: módulo responsável pela comunicação entre o usuário e o sistema. Deve fornecer também justificativas e explicações referentes às conclusões obtidas, bem como do raciocínio utilizado.

Figura 2.1 Estrutura básica de um SBC

O conjunto constituído pelo Motor de Inferência e a Interface com o Usuário é chamado de *shell*¹ do SBC. É conveniente lembrar que os chamados *Sistemas Especialistas* —SEs— constituem uma classe específica dos sistemas baseados em conhecimento. Em geral, pode-se dizer que:

- SBCs são sistemas que exibem comportamento inteligente através da habilidade no uso de heurísticas e resolvem problemas usando conhecimento sobre o domínio específico. Este conhecimento é representado numa base, separado dos mecanismos de inferência e controle,
- SEs são SBCs que resolvem problemas os quais requerem considerável habilidade, conhecimento e heurísticas se resolvidos por um ser humano.

A figura 2.2 [Waterman-86] mostra as características desses sistemas no contexto de Inteligência Artificial. Na literatura os termos SE e SBC são usados indistintamente, e o presente trabalho adota a mesma abordagem.

Um SBC deve ser capaz de explicar seu comportamento e suas decisões ao usuário, ou seja, mostrar *porque* e *como* chegou a uma determinada conclusão. Geralmente as perguntas *por que* referem-se a qual conhecimento respalda a conclusão; as perguntas *como* referem-se aos passos de raciocínio seguidos para determinar a solução do problema. Esta característica é especialmente necessária quando o SBC lida com domínios incertos. A explicação pode aumentar o grau de confiança que o usuário deposita no sistema, ou então, ajudá-lo a encontrar alguma falha no raciocínio deste.

¹ Alguns termos técnicos serão mantidos em inglês, uma vez que o seu uso nesta forma é aceito pela comunidade da área.

A aquisição de conhecimento consiste basicamente na transferência do conhecimento —informações e formas de condução do raciocínio— do especialista humano à base de conhecimento do SBC envolvendo essencialmente extração, análise e interpretação do conhecimento que o especialista humano usa quando resolve problemas. Muitas vezes, devido à complexidade de problemas reais, a compilação do conhecimento específico do domínio em um conjunto de fatos e regras é uma tarefa difícil. Esse processo, denominado *aquisição do conhecimento*, é muitas vezes uma das tarefas na construção do sistema que exige um maior investimento em tempo e esforço. Por este motivo diversas técnicas são utilizadas para auxiliar neste processo. Entre elas podem ser mencionados desde técnicas manuais —que tipicamente consistem de entrevistas e análise de protocolos [Boose-89]— até modalidades baseadas em Aprendizado de Máquina —as quais utilizam técnicas baseadas em indução [Castiñeira-91] [Castiñeira-90] [Castiñeira-90a] [Nicoletti-92] ou generalizam situações específicas em conceitos [Nicoletti-94].

Figura 2.2 Sistemas Baseados em Conhecimento e Sistemas Especialistas

2.3 Necessidade de Raciocínio Incerto

O raciocínio com informações incompletas ou incertas é determinante na resolução de problemas. A incerteza pode ser originada por relações causa-efeito desconhecidas ou não determinísticas, por ignorância de algumas das variáveis intrínsecas do problema, ou por valores imprecisos ou ausentes.

Em alguns domínios, como por exemplo em aplicações médicas, as relações causa-efeito são muitas vezes inexatas. Um sintoma pode ser originado por diversas causas, levando isso a conclusões incertas. Por exemplo, um médico pode não ter certeza da presença de algum sintoma; ou a presença deste pode não ser uma prova conclusiva do paciente possuir uma determinada doença. Muitas vezes o médico atinge conclusões das quais não está totalmente certo.

Em muitos problemas não é possível prever exatamente as conseqüências de cada possível alternativa, já que cada uma delas envolve riscos e incertezas. Por exemplo, devido à avançada idade de o paciente o tratamento com medicamentos pode ser escolhido como mais adequado do que uma operação, embora uma operação sem nenhum problema pós-operatório possa ser mais eficiente do que o tratamento. A

Analogamente, para atingir conclusões semelhantes um SBC deve possuir a capacidade de lidar com incertezas, isto é, realizar *raciocínio incerto* ou *raciocínio plausível*.

O *raciocínio probabilístico em redes de inferências* utiliza a teoria das probabilidades para tratar com incertezas. Esta seção discute algumas das vantagens de se empregar esta abordagem para realizar raciocínio incerto num SBC.

Uma metodologia frequentemente utilizada na construção de SEs é a representação baseada em regras. Nestes sistemas o conhecimento é representado com regras da forma

onde A e B são proposições lógicas. A propriedade de poder adicionar novas regras ou remover as já existentes sem a necessidade de modificar outras regras da BC é denominada *modularidade*. Uma vantagem da abordagem baseada em regras é a possibilidade de poder construir sistemas modulares.

No processo de aquisição de conhecimento para a construção de SEs o conhecimento é geralmente representado na direção na qual será aplicado, isto é, na direção das evidências observáveis para as hipóteses a ser inferidas. Por exemplo, com regras do tipo

SE <evidência> ENTÃO <hipótese> .

Para o especialista, porém, é geralmente mais fácil raciocinar no sentido inverso, da hipótese para as evidências:

SE <o paciente tem úlcera –hipótese> ENTÃO <os sintomas –evidências– são ...>

pois essa direção reflete a relação causa-efeito. Utilizando alguns dos mecanismos da abordagem probabilística, como por exemplo *redes de Bayes* ou *diagramas de influências*, a BC pode ser construída mantendo a direção natural de raciocínio do especialista e depois revertida para o uso. Essa escolha na representação facilita a aquisição do conhecimento em domínios determinísticos e é essencial quando o problema envolve incertezas. Uma vez construídas, as relações podem facilmente ser revertidas para realizar a inferência e diagnose [Shachter-87].

2.4.3 Inferências bidirecionais

A habilidade de utilizar informação para realizar prognose e diagnose é um componente importante do raciocínio plausível. Isto é, dada a regra A implica B é possível raciocinar em ambas as direções: de A para B (se A é verdade então B é verdade) e de B para A (se B é verdade então A é mais provável). Esse último procedimento é denominado padrão indutivo [Polya-1954].

O ser humano utiliza esse tipo de raciocínio bidirecional. Por exemplo, "fogo implica fumaça" faz com que a presença de fumaça leve a considerar a possibilidade de existir fogo. Para um sistema baseado em regras atingir tal comportamento é necessário que a segunda regra ("fumaça pode implicar fogo") esteja explícita e, que a primeira regra seja removida. Isto é, a inferência pode ser realizada numa única direção. Essa limitação pode ser superada utilizando raciocínio probabilístico.

Resumindo, a utilização de raciocínio probabilístico permite representar fielmente o problema para domínios incertos, independentemente da topologia da rede de inferências associada ao modelo. A possibilidade de representar o conhecimento na direção causa-efeito facilita o processo de aquisição do conhecimento e permite realizar inferências bidirecionais.

2.5 Diferentes abordagens no tratamento das incertezas

O *raciocínio probabilístico em redes de inferências* emprega a teoria das probabilidades para tratar com incertezas. Atualmente essa abordagem é amplamente aceita e utilizada, embora diversas outras técnicas já tenham sido propostas e usadas para lidar com o problema da incerteza. De uma forma geral as abordagens existentes podem ser divididas em duas categorias: *quantitativa* e *qualitativa*, segundo a caracterização das incertezas [Bonissone-90].

Modelos baseados na abordagem *qualitativa* geralmente são utilizados para tratar o problema da incerteza devido às informações *incompletas* empregando técnicas não numéricas, principalmente lógica não monotônica.

Na abordagem *quantitativa* são utilizadas técnicas numéricas; os diferentes modelos estão baseados ou na teoria das probabilidades ou em lógicas polivalentes. Entre alguns dos modelos mais destacados podem ser mencionados:

- modelo probabilístico em redes (abordagem utilizada neste trabalho) [Pearl-88],
- teoria da confirmação ou fatores de certeza [Buchanam-84],
- teoria de Dempster-Shafer [Dempster-67], [Shafer-76],
- teoria das evidências [Lowrance-86],
- lógica nebulosa [Yager-87], [Zadeh-78].

Outras revisões das diferentes abordagens no tratamento das incertezas podem ser encontradas em [Clark-90] e [Sheridan-91].

Pearl [Pearl-88] julga que a classificação acima descrita considera somente as variações sintáticas entre os diferentes modelos e propõe uma outra classificação onde são contempladas as variações semânticas das diferentes abordagens para o tratamento das incertezas: *abordagens extensionais* e *abordagens intencionais*².

Abordagens Extensionais versus Abordagens Intencionais

Na *abordagem extensional*, também conhecida como sistema de produção ou sistema baseado em regras, a incerteza é considerada como um valor de verdade atribuído às fórmulas e, seguindo a escola da lógica clássica, calcula-se a incerteza de uma fórmula em função da incerteza das suas sub-fórmulas. Um representante típico desta linha de pensamento são os fatores de certeza utilizados em MYCIN [Shortliffe-76]. Sistemas extensionais são computacionalmente convenientes mas semanticamente imprecisos.

Na *abordagem intencional*, também conhecida como declarativa, a incerteza é ligada ao estado dos possíveis eventos ou subconjuntos de possíveis mundos. Um representante é a teoria das probabilidades, onde medidas de certezas são atribuídas a conjuntos de eventos. Em sistemas intencionais a sintaxe consiste de declarações afirmativas sobre estados de eventos. A manipulação de inferências bidirecionais e evidências correlatas é realizada naturalmente, mas esta sintaxe declarativa não especifica procedimentos operativos para a sua manipulação, a qual pode ser intratável. Isto é, essa abordagem é semanticamente bem definida mas computacionalmente difícil.

Independentemente do tipo de notação utilizada, pesquisas na área tentam atingir um ponto médio entre a necessária clareza semântica e a eficiência computacional. Consequentemente é possível usar a teoria das probabilidades extensionalmente (como em PROSPECTOR [Duda-76]) ou intencionalmente (como em MUNIN [Andreassen-87]). Similarmente acontece com as outras abordagens, como por exemplo com a teoria de Dempster-Shafer.

Em cada uma destas abordagens as regras cumprem um papel diferente. As regras num sistema extensional permitem realizar certas operações lógicas. Por exemplo, a regra

$$A \rightarrow B \text{ (m)}$$

² Também denominadas *abordagens sintáticas* e *semânticas* respectivamente.

pode significar:

"Se A é observado então a certeza de B acontecer pode ser atualizada numa certa quantidade que é função de m".

Isto é, regras descrevem a forma de atuar frente a um determinado acontecimento. Computacionalmente esta regra significa

"Se a certeza de A é modificada numa quantidade δ_a então independentemente dos outros fatos contidos na Base de Conhecimento e independentemente de como δ_a foi ativado a certeza de B pode ser modificada numa quantidade δ_b que pode depender de m, δ_a , e da certeza atual de B".

Nos sistemas intencionais, por outro lado, regras somente têm poder descritivo. Por exemplo, no formalismo Bayesiano a mesma regra é interpretada como a probabilidade condicional $P(B|A)=m$. Esta interpretação não permite realizar nenhuma ação, e ainda se A é verdade na BC nada pode ser afirmado sobre B ou $P(B)$, pois o significado desta declaração é

"Se A é verdade e nenhum outro fato relevante existe na BC, então B tem probabilidade m".

Assim que outros fatos K forem introduzidos na BC não é mais possível deduzir $P(B)=m$, e agora deve-se considerar $P(B|A,K)$. Isto é, a notação probabilística não permite iniciar nenhum cálculo a menos que seja verificado que todos os outros fatos na Base de Conhecimento são irrelevantes.

Assim é necessário construir mecanismos para os sistemas intencionais que permitam converter as declarações de entrada em procedimentos que respondem perguntas. Tais mecanismos são oferecidos por técnicas baseadas em grafos —*redes de crença*—, abordagem utilizada neste trabalho.

2.6 Raciocínio Probabilístico em redes

O objetivo da abordagem baseada em redes é tornar os sistemas intencionais operativos, ressaltando as relações de relevância entre as variáveis. Isto é possível de ser obtido fazendo com que as relações que não podem ser ignoradas sejam facilmente identificadas e acessíveis. Redes de crenças utilizam teoria das probabilidades para representar incertezas e codificam relevância entre variáveis através da vizinhança dos nós de um grafo. Dessa forma consultando-se os nós vizinhos tem-se todas as informações que podem ser relevantes para o item ou variável; os outros nós podem ser ignorados.

Teoria das probabilidades é uma linguagem apropriada para descrever incertezas e para realizar inferências a partir de conhecimento incompleto, isto é, seus métodos são aplicáveis para toda tarefa requerendo inferências a partir de informações incompletas.

Com uma representação formal apropriada do problema de aplicação é possível utilizar teoria das probabilidades para calcular o impacto que certa informação exerce sobre as

condições do problema. Na prática, porém, estes cálculos revelam dificuldades intratáveis ou de ordem exponencial. Uma restrição no cálculo de probabilidades é que o tamanho da distribuição conjunta, isto é, a função que descreve as probabilidades do conjunto de variáveis do problema, cresce exponencialmente com o número de variáveis. Dessa forma pode ser muito complexo adquirir, armazenar e calcular todos os valores de uma distribuição a não ser para pequenos domínios. Com o advento de abordagens probabilísticas baseadas em redes (*diagramas de influências*, *redes de Bayes*, *cadeias de grafos*, *modelos recursivos* e *redes de Markov*) essa dificuldade foi abordada exitosamente [Keshavan-93]. *Redes de crenças* é a expressão empregada para enquadrar todos os tipos de redes acima mencionados.

As abordagens baseadas em redes utilizam a habilidade de representar todas as dependências existentes entre as variáveis do domínio, assumindo independência no caso contrário, sempre que isto seja consistente com o problema considerado.

Esta metodologia ainda se encontra em constante desenvolvimento [Matzkevich-95], embora a teoria que sustenta este modelo [Neapolitan-90], [Pearl-88], [Spiegelhalter-93] já seja suficientemente sólida para fundamentar as numerosas aplicações e ferramentas desenvolvidas tanto comercialmente como para pesquisa. Algumas dessas ferramentas serão discutidas no capítulo 5.

2.7 Considerações finais

Neste capítulo foi discutida a utilização da teoria das probabilidades associada à teoria dos grafos —abordagem bayesiana ou redes de crenças— como meio de representar e manipular incertezas em SBC. Várias controvérsias são originadas quando discute-se se o cálculo de probabilidades é adequado para o tratamento das incertezas e se ele é a melhor metodologia. Pode-se afirmar que ele é uma ferramenta válida e apropriada para determinados tipos de problemas em SBC. Essa abordagem é hoje viável e amplamente utilizada. Redes de crenças são apropriadas tanto no processo de modelagem quanto no cálculo associado ao processo de inferência. A possibilidade de representar explicitamente as dependências do problema permite criar modelos mais realistas, enquanto a suposição de independência, sempre que possível, reduz as complexidades de tempo de computação e espaço de armazenamento.

CAPÍTULO 3

DECISÃO E INFERÊNCIA

3.1 Introdução

Uma decisão é a escolha de uma determinada ação entre uma lista de alternativas. Em geral, sucessivas decisões devem ser tomadas mesmo quando os futuros acontecimentos ou estados são desconhecidos e, por conseguinte, o resultado das correspondentes escolhas é incerto. Um problema decisório envolve ao longo do tempo escolhas que transformam um estado incerto num outro estado. As incertezas dificultam uma descrição formal da situação a ser avaliada, e isso leva a um dilema: por um lado existe a necessidade de atuar e resolver o problema sem demora, por outro, a obtenção de mais dados auxiliaria a entender melhor a situação.

Quem toma a decisão aplica modelos e usa informação como recurso para a ação na solução de problemas. Dependendo da qualidade da informação disponível, suas decisões levam a um plano de ação ótimo, subótimo ou simplesmente factível.

Nos últimos anos, tem-se incrementado consideravelmente o nível de sofisticação das ferramentas computacionais de apoio à decisão. Atualmente elas oferecem conselho não só em escolhas rotineiras, como também na tomada de decisões intuitivas e interpretativas (sistemas de comunicação, sistemas especialistas). Ferramentas de apoio à decisão que operam em contextos pouco estruturados são denominados *Sistemas de Apoio a Decisão* —SAD.

A abordagem dos SAD não é apenas representar a estrutura do problema e automatizá-la, mas oferecer apoio a uma variedade de processos de decisão não estruturados. Sob essa perspectiva há uma grande motivação para oferecer auxílio na tomada de decisão, através dos SADs e Sistemas Especialistas —SEs—, nos níveis de planejamento estratégico e gerenciamento. SADs oferecem facilidades para auxiliar a tomada de decisões e são ainda sistemas gerais, isto é, podem ser aplicados em diferentes áreas.

Mecanismos especiais de representação devem ser utilizados quando no domínio de aplicação do sistema especialista ou sistema de apoio à decisão existem incertezas (ou extensões dos mecanismos tradicionais, como descrito no capítulo anterior). Em particular, são desejados mecanismos que permitam representar as variáveis do problema (com suas respectivas incertezas), as relações entre estas, as decisões, seus resultados e as informações disponíveis. Dentro do contexto das redes de crenças ou redes de Bayes, os *Diagramas de Influências* —DIs— são as redes que explicitamente consideram as decisões e são o mecanismo de representação utilizado neste trabalho. A teoria que sustenta este modelo é a *Teoria da Decisão* —TD—, a qual estende a linguagem das

probabilidades para permitir realizar afirmações sobre as ações alternativas e o quanto os correspondentes resultados são valorados ou desejados.

Neste capítulo serão apresentados os conceitos básicos da teoria da decisão e os seus mecanismos de representação: árvores de decisão e diagramas de influências. As árvores de decisão têm sido utilizadas tradicionalmente na área da análise da decisão, enquanto os diagramas de influências representam uma abordagem relativamente nova. Esta abordagem permite representar problemas de maior envergadura apresentando a vantagem de poder ser diretamente utilizada como o mecanismo de representação do conhecimento do SE.

O capítulo está estruturado da seguinte forma: a seção 3.2 apresenta as diferenças entre a teoria da decisão e a análise da decisão, a seção 3.3 introduz os conceitos básicos da teoria da decisão. Dois métodos gráficos utilizados para representar e resolver o problema decisório são apresentados na seção 3.4, enquanto a seção 3.5 descreve um destes métodos: as árvores de decisão. A seção 3.6 compreende as considerações finais.

3.2 Teoria da Decisão/Análise da Decisão

A *teoria da decisão* permite realizar afirmações sobre as incertezas associadas às variáveis do problema, as ações alternativas —*decisões*— e o quanto os seus correspondentes resultados são valorados ou desejados. A teoria da decisão está baseada nos axiomas da teoria das probabilidades e teoria das utilidades. Enquanto a teoria das probabilidades fornece as bases para uma correta atribuição de crenças frente a informações incompletas, a teoria das utilidades introduz um conjunto de princípios para a necessária consistência entre *preferências* e *decisões*.

Uma *decisão* pode ser definida como a alocação irrevogável de recursos que estão sob controle de quem toma a decisão. *Preferências* descrevem um ordenamento relativo da desejabilidade de vários possíveis resultados ou acontecimentos. O resultado chave da *teoria das utilidades* é que, dadas as propriedades fundamentais de crença e ação então existe uma função escalar —a *função de utilidade*— que descreve as preferências do usuário para os possíveis resultados. Esta teoria também permite expressar atitudes frente a incerteza dos valores dos resultados, como aversão ao risco. A avaliação de um resultado pode ser baseada nos atributos tradicionais como dinheiro ou tempo, mas também em outras dimensões de valor, como prazer, dor, anos de vida ou esforço computacional. Baseando-se nos axiomas da teoria da decisão é relativamente fácil demonstrar que devem ser escolhidas as decisões que maximizam a utilidade esperada sobre um conjunto de decisões com resultados incertos.

A teoria da decisão é freqüentemente qualificada como *normativa*, fornecendo um conjunto de critérios para garantir a consistência entre crenças, preferências e escolhas, o qual deve ser usado por um agente que deseja tomar uma decisão racional. Dado um conjunto de crenças e preferências, a teoria prescreve as decisões que devem ser escolhidas: aquelas que maximizam a utilidade esperada. Esta teoria não é *descritiva* pois não é o seu intuito fornecer uma boa descrição de como as pessoas realmente se comportam no momento de tomar decisões em domínios incertos. Este fato tem provocado diversas pesquisas examinando as diferenças entre a forma que a teoria da decisão sugere atuar e a forma que as pessoas realmente atuam.

A teoria das probabilidades e a teoria da decisão apresentam os princípios básicos para realizar inferências racionais e decidir em domínios incertos. Estas teorias porém pouco dizem sobre a aplicação destes princípios em problemas reais. *Análise da decisão* é a ciência que aplica essas idéias para auxiliar de forma prática a tomada da decisão em problemas reais. Ela inclui técnicas para estruturar problemas, para codificar probabilidades e utilidades, para realizar análise de sensibilidade e explicar resultados com a finalidade de promover um maior entendimento do problema.

3.3 Conceitos Básicos da Teoria da Decisão

A seguir serão apresentados sucintamente alguns conceitos relevantes da teoria da decisão (um conhecimento básico da teoria das probabilidades é requerido). Para uma leitura mais aprofundada recomenda-se [Baker-85], [Harnett-82], [Kim-92], [Smith-88].

A *teoria da decisão* fornece ferramentas para tratar situações onde devem ser tomadas decisões em condições de incerteza, isto é, condições que não são predeterminadas nem controláveis. Especificamente, sob estas condições não se podem prever as conseqüências exatas de cada ação; a falta de informação precisa é uma característica muito importante no problema e deve ser explicitamente formulada na análise da situação.

Uma *decisão* é a escolha de uma ação dentre possíveis alternativas (investir em uma única aplicação financeira ou em várias, e em que proporções, por exemplo). É assumido que as conseqüências de uma determinada ação podem ser medidas ou estimadas, em geral o valor da sua desejabilidade é denominado *utilidade*. Se somente o fator econômico é considerado então o valor monetário associado a uma determinada conseqüência é denominado *retorno* ou *ganho* (pay-off). Finalmente, nas situações analisadas o resultado de uma decisão —*ação*— depende de condições incertas denominadas *estados* ou *eventos* (por exemplo, o futuro estado do mercado financeiro). Resumindo, deve-se considerar o retorno associado à ação selecionada; esse retorno depende da ação (que é controlável) e dos possíveis eventos ou incertezas (que são incontroláveis). O ponto chave na análise de decisão é escolher uma ação *antes* de se ter a oportunidade de determinar as incertezas, como por exemplo escolher uma aplicação financeira antes de saber se ela terá o maior rendimento. O problema pode ser definido da seguinte forma:

Definição 3.3.1

Seja Θ o conjunto de possíveis resultados para as incertezas θ do problema (θ discreto), e D o conjunto das decisões viáveis do problema. Então para $\theta \in \Theta$, $d \in D$, temos que a função de probabilidade de massa $p(\theta|d)$ obedece

$$p(\theta|d) \geq 0 \quad \forall \theta \in \Theta \quad \text{e} \quad \sum_{\theta \in \Theta} p(\theta|d) = 1,$$

i.e. $p(\theta|d)$ é a probabilidade de que θ aconteça se for escolhida a decisão d .

Definição 3.3.2

Usando o escopo da definição acima, a função $R(d, \theta)$ é definida como a função de retorno que especifica o ganho (geralmente em termos monetários) correspondente a um resultado θ se foi tomada a decisão d .

A função $R(d, \theta)$ deve ser conhecida para todos os valores de d e θ . Uma vez especificado o problema nestes termos, deseja-se escolher a decisão $d = d_{max}$ que maximiza a função de retorno $R(d, \theta)$. Embora a função de massa e a função de retorno estejam quantificadas no momento de escolher a ação, i.e. na hora de tomar uma decisão d , o valor do resultado θ só será conhecido depois de tomar a decisão. Desta forma não é possível saber com antecedência qual será a decisão d_{max} que maximiza a função de retorno.

Ante a impossibilidade de se achar uma solução determinística para o problema, o procedimento geralmente utilizado é escolher a ação que "em média" leva a um bom valor da função de retorno. Em outras palavras, é escolhida a decisão que maximiza o ganho esperado (expected pay-off) como resultado dessa ação. Isto é, o objetivo pode ser formulado nos seguintes termos.

3.3.1 Critério do Valor Monetário Esperado (VME)

Segundo o critério do VME deve-se escolher uma decisão $d^* \in D$ que maximiza o retorno esperado ao longo do espaço dos resultados Θ . Isto é, deve se determinar o valor d^* que maximiza $R(d)$ para:

$$R(d) = \sum_{\theta \in \Theta} R(d, \theta) p(\theta|d),$$

onde $R(d)$ é o retorno esperado quando a decisão d é tomada.

A decisão d^* que maximiza $R(d)$ é a escolha ótima se o algoritmo do valor monetário esperado captura os reais objetivos de quem realiza a decisão. Como exemplo considere-se a seguinte situação:

Exemplo 3.1

Um investidor com \$10000 dólares deve decidir entre:

- uma aplicação conservadora, do tipo prazo fixo,
- uma aplicação especulativa do tipo ações da bolsa.

O investidor estima que os juros da primeira aplicação serão de 5% ao ano, assim seu ganho será de \$500 anuais. O retorno da bolsa depende do estado do mercado de ações: se as ações subirem ele pode ganhar \$2000, se as ações descenderem ele pode perder até \$1000. Os valores dos correspondentes retornos são mostrados na tabela 3.1.

Decisões	Eventos	
	Ações sobem	Ações descem
Investir a prazo fixo	500	500
Investir em ações	2000	-1000

Tabela 3.1 Retornos do investidor.

Por outro lado o investidor estima que a probabilidade das ações descerem é de 0,4, e a probabilidade de elas subirem é 0,6. Com esses dados pode se calcular o valor monetário esperado (VME) para cada ação. Seja *f* a ação de investir a prazo fixo e *b* a ação de investir na bolsa, então:

$$VME(f) = 500 * 0,6 + 500 * 0,4 = 500$$

$$VME(b) = 2000 * 0,6 + (-1000) * 0,4 = 800$$

Decisões	Eventos		VME
	Ações sobem	Ações descem	
Investir a prazo fixo (f)	500	500	500
Investir em ações (b)	2000	-1000	<u>800</u>

Ou seja, segundo o algoritmo do valor monetário esperado o investidor deve escolher a ação que maximiza este valor, isto é, investir na bolsa de valores.

3.3.2 Função de Utilidade

Em algumas ocasiões o algoritmo do valor monetário esperado não reflete os objetivos do cliente (a pessoa para quem se desenvolve a análise da decisão). Muitas pessoas defrontadas com o problema acima prefeririam ganhar 500 dólares com certeza (ou seja, realizar uma aplicação a prazo fixo) a enfrentar a incerteza de ganhar 2000 ou perder 1000 (mesmo que a probabilidade de ganhar 2000 seja maior). Outras pessoas dividiriam o dinheiro para investir uma parte em cada aplicação, diminuindo assim o risco da perda. Pode ser observado que, um fator que o algoritmo do VME não considera é que o valor do dinheiro é diferente para cada pessoa e que para um determinado indivíduo esse valor pode ser uma função não linear. Em outras palavras, \$10000 dólares podem não ser necessariamente 10000 vezes mais desejáveis do que um dólar (mesmo que eles sejam 10000 vezes mais valiosos em poder de compra). Para esclarecer essa última afirmação considere o exemplo a seguir.

Exemplo 3.2

Um investidor pode escolher entre duas possibilidades a) e b):

- a) ganhar \$1 milhão de dólares com certeza
- b) ganhar \$100 milhões (evento *i* com probabilidade de 90%) ou não ganhar nada (evento *ii* com probabilidade de 10%)

Esses valores com o respectivo VME de cada situação são mostrados na tabela 3.2.

Ações	Eventos		VME
	i)	ii)	
a)	1 milhão	1 milhão	1 milhão
b)	100 milhões	0	90 milhões

Tabela 3.2 Retornos milionários

Apesar da grande diferença nos VMEs muitas pessoas prefeririam a possibilidade a), pois 1 milhão de dólares já é uma quantia satisfatória (é claro que os 100 milhões seriam melhores ... se não implicassem um risco de 10% de ficar com as mãos vazias). Nesse caso 100 milhões não são necessariamente 100 vezes mais desejados do que 1 milhão, senão se aplicaria o critério do VME.

Assim, em determinadas situações o critério do VME não é o mais apropriado, faltando considerar as preferências individuais de cada pessoa com respeito às consequências do problema. A teoria das utilidades oferece um outro critério que cobre essas deficiências: o critério da *utilidade*. A utilidade serve para medir as *preferências* do cliente, i.e., o seu valor relativo frente às consequências ou retornos do determinado problema. A seguir, uma descrição da função de utilidade.

Definição 3.3.3

A função escalar $U(d, \theta)$ é a *função de utilidade* para uma determinada pessoa frente a um problema de decisão, se ela atribui um número, real, para cada resultado θ e cada decisão d indicando a preferência relativa dessa pessoa.

Esta métrica permite considerar não somente atributos monetários como também outras dimensões de valor, por exemplo: fatores de segurança, prazer, dor, anos de vida, prestígio, ou esforço computacional. Desta forma também é possível considerar atributos subjetivos cuja avaliação varia para as diferentes pessoas, ou até, para uma mesma pessoa em diferentes circunstâncias. A determinação da função de utilidade é discutida a seguir.

3.3.3 Determinação da Função de Utilidade

A modelagem de preferências na forma de funções de utilidades é um dos pontos centrais da análise da decisão. Em decisões importantes geralmente existem objetivos conflitantes, via de regra maximizar qualidade e minimizar custos. Ou, por exemplo, na escolha de uma terapia médica devem ser considerados: qualidade de vida, custo da terapia e anos esperados de vida. Para modelar as preferências em tais situações é utilizada a *teoria da utilidade para multi-atributos* [Keeney-76], [vonWinterfeldt-86]. As principais etapas desse processo compreendem:

- identificação dos objetivos ou atributos,
- eliminação de atributos que se superpõem,
- identificação de uma função de valor para cada atributo,
- escolha de uma única função de utilidade que combine as funções de valor de cada atributo.

As referências acima citadas apresentam numerosos resultados teóricos e discutem diversas técnicas de extração para codificar o complexo esquema de preferências de um agente em termos de atributos individuais.

Outro componente importante na codificação das preferências consiste em identificar a atitude frente ao *risco*. Para alguns, riscos significam desafios a enfrentar enquanto que para outros significam problemas a evitar. Existem diversas técnicas para determinar a função de utilidade, várias delas consistem em perguntar ao cliente suas preferências frente a apostas hipotéticas e depois combinar os resultados para verificar inconsistências. A seguir é descrito um método para codificar a função de utilidade para atributos não conflitantes.

Para determinar as utilidades do cliente, primeiro deve-se escolher o retorno mais desejado, M , e o menos desejado, N . A cada um deles pode-se atribuir o valor de utilidade 1 e 0 respectivamente (esses valores são arbitrários). Supondo agora existir um outro retorno, P com ($N < P < M$), do qual se deseja calcular a utilidade, isso pode ser feito da seguinte forma. Sejam as seguintes situações:

- A) receber P com certeza,
- B) receber M com probabilidade p ou
receber N com probabilidade $(1 - p)$

Denominando $U(c)$ a utilidade da consequência c então a utilidade esperada (UE) de cada ação é:

$$\begin{aligned} UE(A) &= U(P). \\ UE(B) &= U(M) p + U(N) (1 - p) \\ &= 1p + 0 (1 - p) \\ &= p \end{aligned}$$

Se $U(P) < p$, então a situação B deve ser escolhida; se $U(P) > p$, a situação A deve ser escolhida. Se $U(P) = p$, a escolha entre as duas situações é numericamente indiferente. Essa propriedade será usada para determinar a utilidade de P . Se a probabilidade p , com a qual é indiferente escolher entre as duas situações, puder ser determinada, então a utilidade de P é igual ao valor de p . Dessa forma, pode-se determinar a utilidade de qualquer consequência ou ganho, uma vez que a consequência mais desejada M e a menos desejada N forem determinadas.

Pode ser observado que as consequências não precisam estar definidas em termos monetários, assim quando a utilidade de uma consequência é calculada podem ser considerados tanto fatores monetários quanto fatores não monetários. Por exemplo, em alguns tipos de problemas é importante considerar fatores de segurança, condições de trabalho, popularidade, prestígio, etc.

3.3.4 Função de Utilidade do Dinheiro

Apesar de os fatores não monetários serem de certa importância, também é interessante determinar a relação entre dinheiro e utilidade. A fim de exemplificar considere-se que se deseja calcular esta relação para o intervalo de valores entre \$10 e -\$10, ou seja,

$M=\$10$ e $N=-\$10$. Usando-se a análise acima definida, pode-se calcular a utilidade para qualquer quantia entre $\$10$ e $-\$10$. Se for o caso que a escolha entre:

- 1) receber $\$5$ com certeza, e
- 2) receber $\$10$ com probabilidade 0,8 ou perder $\$10$ com probabilidade 0,2

é indiferente, então $U(\$5)=0,8$. Fazendo o mesmo para diversos valores entre $\$10$ e $-\$10$, os pontos correspondentes podem ser colocados num gráfico, e pode-se traçar uma curva através deles, a qual é denominada *função de utilidade do dinheiro*, figura 3.1. É importante ressaltar que cada pessoa pode ter uma função de utilidade diferente, e que uma mesma pessoa pode ter diferentes funções de utilidade ao longo do tempo ou frente a diversas condições.

Figura 3.1 Exemplo de uma função de utilidade do dinheiro.

Baseando-se nos axiomas da teoria da decisão é relativamente fácil demonstrar que devem ser realizadas as escolhas que maximizam a *Utilidade Esperada* —*UE*— sobre um conjunto de decisões com resultados incertos.

3.3.5 Maximização da Utilidade Esperada

A melhor decisão $d=d_{max}$ é aquela que maximiza a função de utilidade $U(d, \theta)$. Como já discutido, o valor do resultado θ só será conhecido depois de tomar a decisão. Desta forma não é possível saber com antecedência qual será a decisão d_{max} que maximiza a função de utilidade. Ante a impossibilidade de achar uma solução determinística para o problema é escolhida a decisão que maximiza a utilidade *esperada* como resultado dessa ação. Este objetivo pode ser formulado nos seguintes termos.

Critério da Utilidade Esperada (UE)

Segundo o critério da UE deve-se escolher uma decisão $d^ \in D$ que maximiza a utilidade esperada ao longo do espaço dos resultados Θ . Isto é, deve se determinar o valor d^* que maximiza $U(d)$ para:*

$$U(d) = \sum_{\theta \in \Theta} U(d, \theta) p(\theta | d),$$

onde $U(d)$ é a utilidade esperada quando a decisão d é tomada.

Se o problema considerado envolve somente retornos monetários, então uma vez que a função de utilidade do dinheiro para o correspondente problema for determinada o critério da UE pode ser aplicado. Cada valor da tabela de retornos do problema de decisão pode ser convertido numa utilidade usando a função já definida. Toda tabela de retornos pode assim ser convertida numa tabela de utilidades, como mostrado no seguinte exemplo.

Exemplo 3.3

Considere novamente o exemplo 3.1 da página 15 referente ao investidor com \$10000 que deve decidir entre uma aplicação a prazo fixo ou investir na bolsa de valores (tabela 3.1). O investidor é, neste exemplo, uma pessoa que gosta de evitar riscos. Depois de analisar diversas situações para definir a sua função de utilidade comprova-se que esta pode ser representada por

$$U(D) = 10 - \left(\frac{2000 - D}{500} \right)^2,$$

onde D representa o ganho. Esse exemplo ilustra como funções matemáticas podem ser usadas para representar funções de utilidade (as mais utilizadas são as curvas quadráticas, exponenciais, logarítmica e linear). Como a função de utilidade do investidor não é linear com respeito ao dinheiro, é necessário converter a tabela de retornos (tabela 3.1) a valores de utilidade. Para isso, pode-se utilizar a fórmula acima, por exemplo:

$$U(\$500) = 10 - \left(\frac{2000 - 500}{500} \right)^2 = 1.$$

Analogamente, todos os valores da tabela podem ser convertidos em utilidades. Os valores resultantes são mostrados na tabela 3.3.

Decisões	Eventos	
	Bolsa sobe	Bolsa desce
Investir a prazo fixo (f)	1	1
Investir em ações (b)	10	-26

Tabela 3.3 Tabela de utilidades do investidor.

As utilidades esperadas de cada ação são agora:

$$UE(f) = 1 * 0,6 + 1 * 0,4 = 1$$

$$UE(b) = 10 * 0,6 + (-26) * 0,4 = -4,4$$

Assim, a ação ótima segundo o critério da UE é o investimento a prazo fixo. Nesse caso, o investimento em ações é arriscado demais para o gosto do investidor mesmo que ele possa obter o ganho máximo.

Decisões importantes devem ser tomadas utilizando o critério da UE em lugar do critério do VME. Porém, quando a função de utilidade do dinheiro é linear pode-se maximizar o VME e, nesse caso específico, maximizar o VME é equivalente a maximizar a UE.

O problema de usar utilidades é que, algumas vezes, determinar a função utilidade do cliente não é tarefa simples. Essa curva pode ser formulada matematicamente em muitos casos, mas o grau de dificuldade para fazê-lo pode variar muito.

3.4 Métodos Gráficos na Teoria da Decisão: Árvores de Decisão e Diagramas de Influências

Dois métodos gráficos são freqüentemente utilizados na teoria da decisão, eles são

- árvores de decisão [Raiffa-68], [Smith-88],
- diagramas de influência [Howard-84], [Shachter-86]

Como apresentado na figura 3.2, os dois métodos representam os conceitos básicos dos modelos de decisão: decisões, informação, resultados da decisão e relações entre objetos (equações, regras, ligação causa-efeito, etc.). Ambos são estruturas gráficas para modelar variáveis incertas e decisões revelando a dependência probabilística e o fluxo de informação. Eles podem ser utilizados para formular os problemas tais como eles são percebidos pelo usuário, bem como para incorporar o conhecimento do especialista, fornecendo descrições precisas de informações e procedimentos de decisão que podem ser armazenados e manipulados pelo computador. Os diagramas de influência foram concebidos especialmente como uma estrutura completa ("*front-end*") para sistemas de apoio à decisão.

Árvores de decisão e diagramas de influência são matematicamente bem definidos e podem ser utilizados diretamente em processos teóricos. Estes últimos diagramas podem ser também utilizados para representar o conhecimento da decisão em sistemas especialistas. O conhecimento pode ser codificado como uma parte dos diagramas, para ser diretamente utilizado pela máquina de inferências do SE.

A seguir as árvores de decisão serão apresentadas mais detalhadamente. Os diagramas de influências, representação utilizada neste trabalho, serão discutidos no próximo capítulo.

Figura 3.2 Modelos gráficos de decisão:
a) Árvore de decisão; b) Diagrama de influência

3.5 Árvores de Decisão

Uma conhecida representação gráfica de um problema decisório é a árvore de decisão. Uma vez desenhada, a árvore pode ser usada para calcular a utilidade esperada associada a cada seqüência de ações ou escolhas. Assim pode ser determinada a seqüência de decisões —*estratégia de decisão*— que maximiza esses valores.

Árvores de decisão são uma poderosa ferramenta de análise que auxiliam a especificar e quantificar as relações existentes entre as decisões e os eventos incertos, assim como a achar a solução ótima. O problema com árvores de decisão é que, à medida que o número de variáveis do problema aumenta, os gráficos gerados podem tornar-se confusos e complexos.

A técnica de desenhar e calcular a árvore será ilustrada utilizando o exemplo a seguir [Smith-88].

Exemplo 3.4

O governo oferece a uma companhia de petróleo a opção de perfurar um campo A ou perfurar um campo B. A probabilidade de existir petróleo no campo A é de 0,4, isto é $P(A)=0,4$, e no campo B de 0,2, $P(B)=0,2$ (esses dois eventos são independentes). O ganho líquido esperado se petróleo for extraído é de \$77 milhões para o campo A, e \$195 milhões para B. As possíveis decisões da companhia são:

- 1) não perfurar nenhum dos campos;
- 2) perfurar o campo A ou o campo B;
- 3) pagar para realizar uma investigação prévia em um dos campos (não em ambos), e segundo o resultado dessa investigação escolher entre as opções 1) e 2) acima.

Deve ser observado que o resultado da investigação prévia não é totalmente preciso. Assim, se existir petróleo em um campo o resultado da investigação aconselhará perfurar com uma probabilidade de 0,8. Se não existir petróleo o resultado aconselhará perfurar com probabilidade 0,4. O custo dessa investigação é de \$6 milhões, e o custo de perfurar qualquer dos campos é de \$31 milhões.

Sejam A e B a possibilidade de existir petróleo nos campos A e B respectivamente. Sejam $-A$ e $-B$ os eventos complementares (não existe petróleo nos respectivos campos). Os eventos rotulados a e b acontecem quando a investigação aconselha perfurar nos campos A e B respectivamente. Os eventos $-a$ e $-b$ são complementares (não se aconselha perfurar os campos A e B respectivamente). Com a notação acima tem-se que:

$$\begin{array}{ll} P(A)=0,4 & P(B)=0,2 \\ P(a|A)=0,8 & P(b|B)=0,8 \\ P(a|-A)=0,4 & P(b|-B)=0,4 \end{array}$$

Como A e B são independentes, a investigação num campo não oferece nenhuma informação a respeito do outro, isto é $P(a|A, B) = P(a|A)$ e $P(b|B, A) = P(b|B)$. Assim, pelo teorema de Bayes temos que:

$$\begin{aligned} P(a \cap A \cap B) &= P(a|A \cap B) P(A \cap B) \\ &= P(a|A) P(A \cap B) \\ &= P(a|A) P(A) P(B) \end{aligned}$$

Todas as probabilidades relevantes ao problema podem ser calculadas e são mostradas na tabela 3.4.

	$A \cap B$	$-A \cap B$	$A \cap -B$	$-A \cap -B$	Probabilidade marginal
a	0,064	0,048	0,256	0,192	0,56
$-a$	0,016	0,072	0,064	0,288	0,44
b	0,064	0,096	0,128	0,192	0,48
$-b$	0,016	0,024	0,192	0,288	0,52
Probabilidade marginal	0,080	0,120	0,320	0,480	

Tabela 3.4 Probabilidades dos eventos de interesse no exemplo 3.4.

No problema existem três decisões terminais, isto é, decisões que podem ser escolhidas depois de haver ou não uma investigação prévia. Elas são: perfurar A, perfurar B, não perfurar. Como mostrado na tabela 3.5 os retornos esperados (pay-off) para cada uma delas, dado o estado de ambos os campos, é o ganho menos o custo da perfuração.

Supondo que a função de utilidade do dinheiro da companhia de petróleo é linear, então maximizar a utilidade esperada é equivalente a maximizar o valor monetário esperado. Isto é, pode se escolher as melhores decisões segundo o critério do VME.

Evento	Ações		
	Perfurar A	Perfurar B	Não perfurar
$A \cap B$	46	164	0
$\neg A \cap B$	-31	164	0
$A \cap \neg B$	46	-31	0
$\neg A \cap \neg B$	-31	-31	0

Tabela 3.5 Retornos esperados em \$ milhões.

A seguir será explicado como desenhar a árvore de decisão e calcular as "melhores decisões" (aquelas que maximizam o VME):

- i) Antes de começar a desenhar a árvore devem ser identificadas as possíveis decisões a serem tomadas antes de poder observar algum evento de interesse. Elas são:

- d1 - investigar o campo A,
- d2 - investigar o campo B,
- d3 - perfurar o campo A sem investigação prévia,
- d4 - perfurar o campo B sem investigação prévia,
- d5 - não investigar nem perfurar nenhum dos campos.

Como apresentado na figura 3.3, a decisão pode agora ser representada como um nó (*nó de decisão*) do qual emergem cinco ramos (*ramos de decisão*) representando as cinco alternativas.

Figura 3.3 Nó de decisão

- ii) Escolhida uma entre as possíveis alternativas (um ramo da árvore), pode ser observado um evento que inicialmente era incerto. Por exemplo, depois de escolher a alternativa d1 (investigar o campo A) pode ser observada uma recomendação positiva (a) ou negativa ($\neg a$) de perfurar o campo A. Como apresentado na figura 3.4, a observação de um evento é representada com um nó circular, denominado *nó de eventos* ou *nó de probabilidade*, do qual emergem tantos ramos quanto os possíveis valores desse evento.

Figura 3.4 Um dos primeiros nós de eventos.

iii) De forma análoga, para cada ramo são introduzidos novos nós de decisão ou nós de eventos, dependendo das decisões que já foram escolhidas anteriormente ou dos eventos que já foram observados. Por exemplo, depois de investigar o campo A e receber uma recomendação positiva para perfurá-lo, deve ser decidido sobre:

- perfurar o campo A (PA),
- perfurar o campo B (PB),
- não perfurar nenhum dos dois campos (NP).

Considerando a seguinte nomenclatura para os correspondentes eventos:

$A|a$ existência de petróleo no campo A dado que foi recomendado perfurá-lo;
 $-A|a$ inexistência de petróleo no campo A dado que foi recomendado perfurá-lo;
 $B|a$ existência de petróleo no campo B dado que foi recomendado perfurar o A;
 $-B|a$ inexistência de petróleo no campo B dado que foi recomendado perfurar o A;
 e assim por diante.

obtém-se a árvore mostrada na figura 3.5. Uma vez representadas todas as possíveis sequências de decisões e eventos, a estrutura da árvore está completa.

Figura 3.5 Estrutura da árvore de decisão.

- iv) Uma vez representadas todas as possíveis decisões e eventos na estrutura da árvore, falta agora adicionar os correspondentes valores numéricos, ou seja, os valores dos retornos em cada nó terminal bem como das probabilidades em cada ramo dos eventos.

Para calcular os retornos correspondentes a cada nó terminal devem ser somados os ganhos e subtraídos os custos originados por cada decisão e evento acontecido no percurso que vai da raiz da árvore até o correspondente ramo terminal. Por exemplo, considerando o ramo (d1, a, PA, A|a): primeiro foi decidido investigar o campo A (custo=6), a investigação aconselhou perfurar esse campo, o campo foi perfurado (custo=31) e petróleo foi achado nele (ganho=77). O retorno r_1 para este ramo é:

$$r_1 = -6 - 31 + 77 = 40$$

Deve-se proceder analogamente com todos os ramos e o valor do retorno é escrito à direita de cada nó terminal.

- v) Os últimos componentes a ser adicionados na árvore são as probabilidades correspondentes a cada ramo de eventos. Por exemplo, considere-se a decisão d1: investigar-se ou não o campo A. No momento dessa decisão, não se sabe se será observada uma recomendação positiva (a) ou negativa (-a) de perfurar o campo A. Porém, da tabela 3.4 podem ser obtidas informações probabilísticas sobre o evento

a , que são $P(a)=0,56$ e $P(-a)=0,44$. Esses valores devem ser escritos ao lado do correspondente ramo. Agora, supondo observar um resultado positivo (a), é resolvido perfurar o campo A (PA). Depois dessa decisão, pode acontecer que o campo A tenha ou não petróleo, considerando as decisões e acontecimentos anteriores. Nesse caso, o único que influencia essa probabilidade é a , ou seja, deve-se calcular $P(A|a)$. Utilizando o teorema de Bayes,

$$P(A|a) = \frac{P(a|A) P(A)}{P(a)},$$

obtem-se $P(A|a)=0,571$ e $P(-A|a)=1-P(A|a)=0,429$. Essas probabilidades podem ser adicionadas nos dois ramos do extremo superior direito. Adicionando o valor das probabilidades em cada ramo probabilístico da árvore obtém-se os valores mostrados na figura 3.6.

A seguir, será descrito o procedimento de *indução "backward"* ou *dobramento da árvore*.

- vi) A árvore contém agora toda a informação necessária para calcular as melhores decisões, isto é, aquelas que maximizam o VME. Deve-se trabalhar das folhas para a raiz, calculando o valor do retorno esperado correspondente a cada nó de eventos terminal (a soma dos retornos multiplicados pela correspondente probabilidade). Por exemplo, para o nó que emerge do ramo (d1, a , PA) este valor é dado por:

$$0,571 * 40 + 0,429 * (-37) = 7$$

- vii) O mesmo deve ser realizado para todos os nós terminais. Depois disso cada nó de decisão terminal tem nos extremos de seus ramos o valor de um retorno (d5) ou um nó de probabilidade (evento) com um valor esperado do retorno. Para cada nó de decisão deve ser escolhido o ramo emergindo deste com o maior valor do retorno ou retorno esperado. Este valor deve ser escrito sobre o correspondente nó de decisão. Por exemplo, como pode ser observado na figura 3.7, para o nó depois de (d1, a) os retornos são 7 (perfurar A), -6 (não perfurar) e 2 (perfurar B). Assim o valor 7, maior valor de retorno, é associado ao nó e o seu ramo é sinalizado (por exemplo com um "/").

Figura 3.6 Árvore de decisão com probabilidades e utilidades.

viii) Os passos vi) e vii) devem ser repetidos até os valores correspondentes serem escritos sobre todos os nós conectados com a raiz da árvore. A árvore aparece como mostrado na figura 3.7.

Figura 3.7 Árvore de decisão dobrada

ix) A estratégia ótima pode agora ser extraída diretamente da árvore. Começando na raiz da árvore deve-se escolher a decisão com o ramo que possui o maior valor esperado do retorno. No exemplo considerado deve ser escolhido d2, investigar o campo B. Depois de se ter investigado o campo B será obtido um conselho positivo (b) ou negativo ($-b$) para perfurar o campo B. Todavia é necessário saber qual é a melhor decisão depois de receber essa informação. Caso a resposta seja b , a melhor decisão (com maior retorno esperado) é agora perfurar o campo B (PB). Se a resposta é negativa ($-b$), a melhor decisão é não perfurar nenhum dos campos (NP). Esse curso de ação ótimo, que especifica o que fazer segundo o resultado dos vários eventos que são incertos na hora da análise, é denominado *regra de decisão* ou *estratégia*.

Se em lugar do algoritmo do VME deseja-se maximizar a utilidade esperada, então todos os retornos do lado direito da árvore (figura 3.6) devem ser convertidos em utilidades antes de iniciar o procedimento de dobrar a árvore. É importante lembrar que a conversão de retornos para utilidades deve ser realizada antes de calcular os valores esperados. Calcular os retornos esperados e depois converter para utilidades é incorreto.

3.6 Considerações Finais

Neste capítulo foi abordada a problemática da representação e manipulação pelo sistema de apoio à decisão dos componentes básicos de uma escolha. Mecanismos especiais de representação devem ser utilizados quando no domínio de aplicação do SE ou SAD existem incertezas. No contexto das redes de Bayes os diagramas de influências são o mecanismo de representação que explicitamente considera as decisões. A teoria que sustenta este modelo é a teoria da decisão.

Árvores de decisão são o mecanismo de representação tradicionalmente utilizado na teoria da decisão. Os conceitos básicos desta teoria e das árvores de decisão foram descritos neste capítulo.

Uma das restrições das árvores de decisão é que, à medida que o número de variáveis do problema aumenta, os grafos gerados podem tornar-se confusos e complexos. Para muitos problemas reais a representação da correspondente árvore de decisão pode ser impraticável. Outra desvantagem é que muitas vezes esta abordagem transforma o problema da forma que ele é percebido pelo usuário numa representação diferente —a árvore de decisão— para poder avaliá-lo. Este fato impede o usuário de participar mais ativamente na modelagem e análise do problema, obtendo assim um menor entendimento deste, e restringindo o seu "feedback". Como discutido por Henrion et al. [Henrion-91], geralmente só apresentar a estratégia da decisão que maximiza a utilidade esperada não é efetivo. Um dos resultados mais importantes da análise não são os números atingidos ou as decisões recomendadas, mas auxiliar o usuário a obter uma melhor percepção do problema. Este entendimento advém de compreender o porquê de uma decisão ser mais recomendada do que outra e quais os fatos e incertezas que são mais críticos à conclusão. Este processo é geralmente promovido através do envolvimento do usuário no processo de modelagem e análise. Por esta razão é importante utilizar um formalismo de representação mais compacto e fácil de ser entendido. Estes objetivos foram atingidos com a concepção de um novo mecanismo de representação: os diagramas de influências.

Os diagramas de influências constituem uma abordagem relativamente nova a qual permite modelar problemas de maior envergadura apresentando a vantagem de poder ser diretamente utilizada como o mecanismo de representação do conhecimento do SE. Este mecanismo de representação será discutido a seguir.

CAPÍTULO 4

DECISÃO E INFERÊNCIA: DIAGRAMAS DE INFLUÊNCIA

4.1 Introdução

Um *Diagrama de Influência* —DI— é uma representação esquemática do problema decisório, a qual consiste de um grafo orientado que descreve a estrutura do modelo. O grafo possui um nó para cada variável do modelo e arcos que representam as relações entre estas. Em cada nó são armazenadas informações detalhadas sobre a correspondente variável.

Uma vez representado o problema na forma de um DI, a solução pode ser calculada através de uma série de transformações no diagrama. Essas transformações mudam a estrutura gráfica do modelo e os dados armazenados em cada nó, mas preservam o valor da solução.

Diagramas de influência foram concebidos como uma ferramenta computadorizada para auxiliar na modelagem de problemas decisórios [Howard-84], [Miller-76]. Olmsted desenvolveu o conceito de resolver o problema de decisão através de operações no diagrama [Olmsted-83], e Shachter desenvolveu o algoritmo que realiza as correspondentes transformações [Shachter-86]. Um estudo teórico da independência condicional nos DIs é formulada por Smith [Smith-89a], [Smith-89b]. Ao longo dos trabalhos acima citados é apresentada uma álgebra formal para a operação nos diagramas de influência, o qual permite que estes sejam utilizados não só como uma representação gráfica da estrutura de dependência das variáveis aleatórias como também na resolução do problema.

Neste capítulo serão apresentados os conceitos básicos dos diagramas de influências. O capítulo está estruturado da seguinte forma: a seção 4.2 descreve os componentes básicos de um DI e através de um exemplo estes conceitos são aplicados. A seção 4.3 compreende uma descrição formal de um DI. A seção 4.4 apresenta as transformações necessárias para avaliar um DI, a seção 4.5 descreve outros algoritmos para realizar esta tarefa e a seção 4.6 apresenta as considerações finais.

4.2 Representação do Diagrama de Influência

Numa descrição informal um diagrama de influência é um grafo orientado, como o apresentado na figura 4.1. Cada nó representa uma variável do modelo e os arcos

simbolizam as relações entre estas variáveis. A cada nó corresponde uma tabela com as informações associadas à variável ali representada. Os nós podem ser de três diferentes tipos:

- *nó de probabilidade* $n(p)$, representado por um círculo;
- *nó de decisão* $n(d)$, representado por um retângulo;
- *nó de valor* $n(v)$, representado por um retângulo com os cantos arredondados.

Figura 4.1 Diagrama de influência genérico.

Nós de probabilidades representam variáveis aleatórias ou quantidades incertas. Nós de decisão representam algum componente do espaço das possíveis decisões. Finalmente, o nó de valor representa a questão a ser resolvida, ou seja, a utilidade esperada condicionada ao valor dos predecessores diretos.

Um nó n_2 é dito um *sucessor direto* do nó n_1 se existe um arco orientado de n_1 para n_2 ; e neste caso n_1 é denominado o *predecessor direto* do n_2 . Os arcos do diagrama podem ser de dois tipos:

- arcos que incidem em nós de probabilidades ou em um nó de valor são denominados *arcos condicionais*;
- arcos que incidem em nós de decisão são denominados *arcos de informação*.

Sejam as variáveis aleatórias X_1 e X_2 , representadas no diagrama pelos nós de probabilidade $n(p_1)$ e $n(p_2)$ respectivamente, então $n(p_1)$ deve ser um predecessor direto de $n(p_2)$ se e somente se a distribuição aleatória de X_2 for condicionada ao valor de X_1 , em outras palavras, se X_1 e X_2 não são independentes. Os arcos entre nós de probabilidades revelam dependência das respectivas variáveis associadas ao nó. Por exemplo, se a função de distribuição conjunta $p(x_1, x_2)$ de duas variáveis aleatórias não independentes X_1 e X_2 é formulada como

$$p(x_1, x_2) = p(x_2|x_1) p(x_1) \quad (4.1)$$

então essa decomposição pode ser representada como na figura 4.2. Se X_1 e X_2 fossem independentes, ou seja, se $p(x_1, x_2) = p(x_2) p(x_1)$, não haveria um arco unindo os dois nós.

Figura 4.2 Nós de probabilidades

Analogamente, se a função de distribuição conjunta $p(x_1, x_2, x_3)$ de três variáveis aleatórias X_1, X_2, X_3 fosse formulada como

$$p(x_1, x_2, x_3) = p(x_3 | x_2, x_1) p(x_2 | x_1) p(x_1) \quad (4.2)$$

então esta decomposição pode ser representada como na figura 4.3. Se porém X_3 e X_2 fossem independentes e condicionados a X_1 , então $p(x_3 | x_2, x_1)$ seria constante em x_2 , ou seja, $p(x_3 | x_2, x_1) = p(x_3 | x_1)$, e o arco unindo $n(p_2)$ a $n(p_3)$ não existiria. Como já foi mencionado, a ausência do arco entre dois nós implica a independência das variáveis associadas aos nós.

Figura 4.3 Diagrama de influência com três nós

A representação das incertezas (os nós de probabilidades) num diagrama de influência não é única. A mesma função de distribuição conjunta $p(x_1, x_2, x_3)$ dada pela equação 4.2 poderia também ser formulada como

$$p(x_1, x_2, x_3) = p(x_1 | x_2, x_3) p(x_2 | x_3) p(x_3).$$

Neste caso, a direção dos arcos da figura 4.3 deve ser invertida. Ambos os diagramas representam porém a mesma situação. Este fato explica-se porque arcos entre nós probabilísticos representam dependência e não (necessariamente) causalidade. Desta forma, o significado do diagrama (a correspondente distribuição conjunta) não muda, independentemente da direção dos arcos

Também deve ser observado que pelas regras de combinação das probabilidades nenhuma decomposição da função de distribuição conjunta pode levar a um circuito fechado no grafo. Se um diagrama admite um circuito fechado então ele não representa a decomposição de uma função de distribuição conjunta, e isto pode ser devido a um erro na formulação do problema.

Como descrito anteriormente, *arcos condicionais* (que incidem em nós probabilísticos) representam a dependência das variáveis. Por outro lado, *arcos de informação* (os que incidem em nós de decisão) indicam que a informação está disponível no momento de realizar a correspondente decisão. Desta forma, a ausência de arcos condicionais denota *independência* e a presença de arcos de informação indica disponibilidade de informação.

Um nó de decisão $n(d)$ é um predecessor direto de um nó de probabilidade $n(p)$ se a distribuição da variável aleatória representada por $n(p)$ depender da escolha ou resultado da decisão rotulada por $n(d)$. Um nó de probabilidade $n(p)$ é um predecessor direto de um nó de decisão $n(d)$ se e somente se o valor da variável rotulada por $n(p)$ for conhecido na hora da decisão e puder influir nesta. Um nó de decisão $n(d1)$ é um predecessor direto de um outro nó de decisão $n(d2)$ se e somente se $d1$ for realizada antes de $d2$ e o valor de $d1$ for considerado ao efetuar a decisão $d2$.

Os conceitos acima descritos serão ilustrados através do seguinte exemplo:

Exemplo 4.1

Considere-se o problema dos poços de petróleo do capítulo 3, exemplo 3.4. As variáveis incertas (nós de probabilidades) são:

- $n(A)$ - existência ou não de petróleo no campo A;
- $n(B)$ - existência ou não de petróleo no campo B;
- $n(T)$ - resultado do teste, caso este seja realizado em algum dos dois campos.

Como mencionado anteriormente, a existência de petróleo no campo A é um evento independente da existência de petróleo em B, porém o resultado do teste depende da existência de petróleo no campo examinado. Essa situação pode ser representada como na figura 4.4: observe-se que não existe um arco unindo $n(A)$ a $n(B)$.

Figura 4.4 Nós de probabilidades do exemplo considerado.

As decisões do exemplo (nós de decisão) são:

- $n(d1)$ - realizar-se um teste no campo A, no campo B, ou não realizar-se teste nenhum.
- $n(d2)$ - perfurar-se o campo A, o B ou não perfurar-se nenhum deles.

No problema considerado a única variável incerta que depende de uma decisão é $n(T)$. Isto é, o resultado do teste $n(T)$ depende da decisão $n(d1)$ de realizar ou não o teste e, em caso afirmativo, em qual campo. Assim, o nó $n(d1)$ deve ser um predecessor do nó $n(T)$. Por outro lado, a existência de petróleo nos campos não é afetada pela decisão de

realizar algum teste nestes campos; por conseguinte, não existe ligação entre o nó $n(d1)$ e $n(A)$ ou $n(B)$, como ilustrado na figura 4.5.

Figura 4.5 Diagrama de influência parcial do exemplo 4.1.

A decisão $n(d2)$ de perfurar um campo deve ser efetuada depois da decisão de realizar uma investigação prévia $n(d1)$ e à luz do resultado dessa investigação. Isto é, os nós $n(d1)$ e $n(T)$ devem ser predecessores diretos do nó $n(d2)$, como apresentado na figura 4.6.

Figura 4.6 Novo diagrama de influência parcial do exemplo 4.1.

Um nó de valor representa a utilidade esperada dados os valores de seus predecessores diretos. Ou seja, ele representa uma função cujas variáveis são os resultados das decisões e dos eventos rotulados pelos nós de decisão e nós de probabilidades que são seus predecessores diretos. Arcos direcionados chegam ao nó de valor se a utilidade esperada for funcionalmente dependente das variáveis rotuladas pelos nós predecessores diretos. É possível assumir que todos os problemas podem ser formulados como um diagrama de influência com um único nó de valor sem sucessores.

Assumindo que no exemplo a utilidade exprime o retorno, isto é, pretende-se maximizar o retorno, o diagrama resultante é o mostrado na figura 4.7. O nó $n(d2)$ representa a decisão de perfurar ou não (e, caso afirmativo, se no campo A ou B) e o nó de valor $n(v)$ representa o retorno esperado. Esse valor depende de:

- $d1$, pois se algum campo é investigado o custo da investigação deve ser contabilizado;
- os eventos A e B de haver petróleo (se petróleo for extraído, haverá um ganho);
- $d2$, a decisão de qual campo perfurar (deve-se considerar o custo da perfuração e o fato de existir petróleo em tal campo).

Estes nós são os predecessores diretos do nó de valor $n(v)$. O diagrama representa assim todos os componentes do problema. Pode-se observar como esse diagrama é mais

conciso e fácil de ser interpretado em nível cognitivo do que a correspondente árvore de decisão apresentada na figura 3.6.

Figura 4.7 Diagrama de influência do exemplo.

4.3 Definição de um DI

Um DI é uma rede que consiste de um grafo direcionado $G = (N, A)$, onde N é o conjunto dos nós e A o conjunto dos arcos do grafo. O diagrama possui também um conjunto de funções associadas aos nós do grafo. O conjunto N pode ser particionado em três subconjuntos os quais podem ser denominados V , P e D segundo o tipo de nó. Existe no máximo um *nó de valor* $v \in V$, o qual representa o objetivo cuja esperança deve ser maximizada. Existem zero ou mais *nós probabilísticos* no conjunto P , representando as variáveis aleatórias. Finalmente, o subconjunto D possui zero ou mais *nós de decisão* correspondentes às escolhas disponíveis.

Os arcos do conjunto A têm um significado diferente segundo o tipo de nó que eles associam. Arcos incidindo em nós probabilísticos ou no nó de valor são denominados *condicionais* e representam dependência probabilística. Arcos chegando em nós de decisão são denominados *arcos de informação* e implicam precedência temporal. Todas as incertezas ou escolhas que precedem a decisão considerada devem ser resolvidas antes de esta ser realizada.

Em lugar de considerar arcos do conjunto A é mais conveniente pensar em termos dos predecessores e sucessores de um nó do grafo. O conjunto dos *predecessores diretos* do nó k , $Pred(k)$, pode ser definido como

$$Pred(k) = \{j \in N: (k, j) \in A\}.$$

Associado com cada nó k do grafo existe uma variável X_k e um conjunto Ω_k dos possíveis valores que ela pode assumir. Se k é o nó de valor, então X_k representa a

utilidade esperada e o seu domínio Ω_k é um subconjunto dos números reais. Se k é um nó de probabilidade então Ω_k representa o domínio da variável aleatória X_k . Finalmente, para um nó de decisão k a escolha correspondente X_k pode assumir algum dos valores definidos no conjunto D .

Cada nó do DI tem um mapeamento associado. Para os nós probabilísticos e de valor, este mapeamento compõe-se dos dados de entrada que devem ser atribuídos para poder avaliar o diagrama. À medida que o diagrama é transformado pelo algoritmo de avaliação estes mapeamentos são redefinidos. Para os nós de decisão o algoritmo calcula os respectivos mapeamentos que constituem a estratégia ótima.

O nó de valor $v \in V$ tem associada uma função de utilidade

$$U: \Omega_{\text{Pred}(v)} \rightarrow \Omega_v,$$

a qual representa a utilidade esperada como uma função dos valores dos predecessores do nó de valor. Para cada nó de probabilidades k existe uma distribuição condicional de probabilidades π_k dados os valores de seus predecessores diretos:

$$\pi_k(x_k | x_{\text{Prd}(k)}) = \Pr\{X_k = x_k | X_{\text{Prd}(k)} = x_{\text{Prd}(k)}\}.$$

Se o nó k não tem predecessores diretos então ele representa a distribuição marginal para X_k , ou seja,

$$\pi_k(x_k) = \Pr\{X_k = x_k\}.$$

Para cada nó de decisão k existe uma escolha ótima d_k^* calculada pelo procedimento de avaliação. Este origina uma transformação de $\Omega_{\text{Prd}(k)}$ em Ω_k denotando as alternativas ótimas dada a informação disponível no momento de realizar a escolha.

Com base nos conceitos acima descritos é possível definir um DI da seguinte forma:

Um diagrama de influencia consiste de um grafo direcionado $G = (N, A)$, com nós em N e arcos em A . Os nós são particionados nos subconjuntos V , P e D . Para cada nó k existe um conjunto de valores Ω_k e um mapeamento U , π_k ou d_k^* , dependendo do tipo de nó.

4.4 Avaliação do Diagrama de Influência

Uma vez representado, o DI pode ser avaliado obtendo-se assim a estratégia ótima. O procedimento descrito a seguir reduz o número de nós do DI através de uma série de transformações que preservam o valor da solução [Howard-84], [Olmsted-83], [Shachter-86], [Shachter-88a]. Essas transformações são:

- a mudança da direção de um arco entre dois nós probabilísticos aplicando o teorema de Bayes.
- a remoção de um nó probabilístico através da esperança condicional.

- a remoção de um nó de decisão através da maximização da utilidade esperada.

Antes de avaliar o diagrama algumas condições devem ser verificadas:

- 1) é assumido que existe uma única pessoa que realiza as decisões e portanto estas podem ser ordenadas ao longo do tempo na sequência ($D_1, D_2, \dots, D_m$) tal que D_n seja predecessor direto de D_{n+1} para $1 \leq n < m$;
- 2) o diagrama não contém circuitos fechados;
- 3) o nó de valor é o único nó que não tem sucessor direto;
- 4) se um nó probabilístico $n(p)$ não é predecessor direto de um nó de decisão $n(d)$, então não deve existir um caminho orientado de $n(p)$ a $n(d)$.

Se existe algum nó, exceto o nó de valor, que não tem sucessor direto –condição 3– este nó não influi em nenhum dos outros e pode ser removido do diagrama [Shachter-86].

Para verificar a condição 4 é algumas vezes necessário inverter a direção de algum arco condicional. Por exemplo no problema da figura 4.7 $n(A)$ não é um predecessor direto de $n(d2)$, porém existe um caminho orientado de $n(A)$ a $n(d2)$; similarmente acontece com $n(B)$ - $n(d2)$. Para solucionar esta situação, os arcos $(n(A), n(T))$ e $(n(B), n(T))$ devem ser invertidos. Isso pode ser realizado utilizando-se uma outra representação da distribuição conjunta obtida através do teorema de Bayes. Em lugar de se utilizar a representação associada à figura 4.7, ou seja,

$$P(T, A, B) = P(T | A, B) P(A) P(B),$$

pode ser usada a seguinte representação equivalente, associada à figura 4.8:

$$\begin{aligned} P(T, A, B) &= P(A, B | T) P(T) \\ &= P(A | T) P(B | T) P(T). \end{aligned}$$

Figura 4.8 Diagrama de influência com os arcos invertidos.

De uma forma geral essa transformação é realizada através da mudança na direção de um arco condicional. Esta transformação, assim como as transformações correspondentes à remoção de um nó probabilístico ou de decisão, será descrita a seguir.

4.4.1 Mudança da Direção de um Arco Condicional

Para inverter um arco condicional direcionado de um nó probabilístico i a um nó j deve-se mudar a direção do arco e ambos nós herdam os predecessores condicionais do outro nó [Shachter-86], [Shachter-88a], [Smith-89a], [Smith-89b]. Este processo — a implementação do teorema de Bayes — é ilustrado na figura 4.9, onde $P(i)$ e $P(j)$ são os conjuntos dos nós probabilísticos predecessores diretos de i e j respectivamente.

Figura 4.9 Inversão de um arco num diagrama de influência.

Aplicando essa transformação, e se o problema estiver corretamente modelado, o DI pode verificar as quatro propriedades acima citadas. Neste caso existe pelo menos um nó cujo único sucessor direto é o nó de valor e pode-se iniciar a avaliação do diagrama através da absorção deste nó.

4.4.2 Absorção de um Nó Probabilístico

Dado que um nó probabilístico $n(p)$ precede direta e somente o nó de valor $n(v)$, então $n(p)$ pode ser removido calculando-se a utilidade condicional esperada. Depois disso o nó de valor herda todos os arcos incidentes a $n(p)$.

No exemplo 4.1, figura 4.8, os nós $n(A)$, $n(B)$ e $n(d2)$ têm como único sucessor direto o nó de valor. No caso das variáveis aleatórias A e B , isso significa que elas só influenciam a função de utilidade esperada G associada ao nó de valor. Assim o nó $n(A)$ pode ser eliminado calculando-se a utilidade esperada de G ao longo dos valores de A obtendo-se o retorno $G1$. Na figura 4.10 é mostrado o correspondente diagrama obtido. O nó de valor $n(v)$ é renomeado $n1(v)$ com a correspondente função de utilidade associada $G1$.

Um procedimento análogo é realizado com o nó $n(B)$, obtendo-se a utilidade esperada $G2$ associada ao nó $n2(v)$, processo ilustrado na figura 4.11. Em cada passo são somente afetados o nó absorvido e o nó de valor, as informações e as relações entre os restantes nós permanecem inalteradas.

Figura 4.10 Absorção do nó $n(A)$.

Figura 4.11 Absorção do nó $n(B)$.

No exemplo considerado, depois das transformações realizadas, o nó de decisão $n(d2)$ é o nó que tem como único sucessor direto o nó de valor $n(v)$. A seguir é descrito o processo de absorção de um nó de decisão.

4.4.3 Absorção de um Nó de Decisão

Seja um nó de decisão $n(d)$ que precede diretamente o nó de valor $n(v)$ e todos os outros predecessores condicionais C do nó $n(v)$ são também predecessores do nó $n(d)$. Nestas condições $n(d)$ pode ser removido através da maximização da utilidade esperada, condicional nos valores dos seus predecessores C . As decisões que maximizam a utilidade devem ser registradas na estratégia ótima. O nó $n(v)$ não herda novos predecessores através desta operação.

No exemplo considerado, dada toda a informação disponível quando a decisão $d2$ é realizada, deve-se calcular a decisão $d2^*$ que maximiza a função de utilidade $G2$. Assim, $G3 = G2(d2^*)$ é associado ao novo nó de valor $n3(v)$, e o diagrama resultante é mostrado na figura 4.12 (o nó $n(d2)$ é eliminado e o valor $d2^*$ da decisão $d2$ deve ser registrado).

Figura 4.12 Absorção do nó $n(d2)$.

O processo de eliminar nós continua analogamente até se obter um diagrama com um único nó, o nó de valor, representando a utilidade esperada do problema se a sequência ótima de decisões foi realizada.

4.5 Outros Algoritmos de Avaliação

Redes de Bayes [Pearl-88], [Neapolitan-90], podem ser definidas como DIs que somente possuem nós probabilísticos. Esta representação permite expressar exclusivamente relações probabilísticas entre o estado das variáveis sem considerar explicitamente decisões e preferências. Intensiva pesquisa tem sido realizada sobre este mecanismo de representação e numerosos algoritmos para avaliar redes de Bayes foram desenvolvidos. Cooper mostra como transformar um DI numa rede de Bayes para assim poder avaliar o diagrama utilizando qualquer um dos algoritmos para redes [Cooper-88]. A seguir é apresentada uma breve resenha desses algoritmos.

Todos os *algoritmos exatos* para realizar inferência probabilística em redes de Bayes exploram a característica da independência condicional entre as variáveis do problema, embora de diferentes formas. Por exemplo, o algoritmo de Shachter [Shachter-88a] inverte os arcos da rede até que a probabilidade da variável desejada possa ser diretamente lida do grafo. Neste algoritmo cada inversão de um arco corresponde à aplicação do teorema de Bayes. Pearl [Pearl-86] desenvolveu um esquema de envio de mensagens entre os nós de forma que a distribuição de probabilidades de cada nó é atualizada como consequência da observação dos valores de uma ou mais variáveis. O algoritmo de Lauritzen et al. [Lauritzen-88] cria primeiro um grafo adirecionado a partir da rede bayesiana. A seguir, diversas propriedades matemáticas dos grafos adirecionados são utilizadas para poder realizar a inferência probabilística. Mais recentemente, D'Ambrosio [D'Ambrosio-91] desenvolveu um algoritmo de inferência que simplifica as somas e produtos simbolicamente, como nas transformações descritas por Heckerman [Heckerman-90].

Embora a condição de independência seja utilizada pelos diversos algoritmos para reduzir o esforço computacional (tempo de computação e espaço de armazenamento), deve ser observado que no caso geral a inferência exata numa rede de Bayes arbitrária é NP-completo [Cooper-90].

Existe, contudo, um tipo de redes que pode ser eficientemente resolvido por meio de um algoritmo exato, num tempo que cresce linearmente com o número de nós. Estas são as *redes simplesmente conexas*. Uma rede simplesmente conexa (também denominada *polytree*) é uma rede na qual no correspondente grafo adirecional não existe mais do

que um caminho entre dois nós quaisquer. A rede da figura 4.13.(a) é simplesmente conexa, enquanto que a da figura 4.13.(b) não é. Nesse último caso ela é denominada *rede multiconexa*.

Figura 4.13 Redes: a) simplesmente conexa, b) multiconexa.

Descrições do algoritmo para solucionar redes simplesmente conexas podem ser encontradas em [Neapolitan-90] e [Pearl-88].

Algumas redes multiconexas podem ser transformadas numa rede simplesmente conexa para serem avaliadas com um algoritmo exato [Jensen-89], [Lauritzen-88]. Existem porém redes multiconexas que, pela sua topologia ou número de nós, não podem ser avaliadas com métodos exatos. Nestes casos podem ser utilizados diversos algoritmos para achar uma *solução aproximada* da rede [Charniak-91] [Henrion-91a] [Horvitz-88], como por exemplo o método de Monte Carlo. Mesmo assim, também utilizando procedimentos aproximados o problema continua sendo NP-completo [Dagum-93].

Como já observado, a inferência probabilística em redes de Bayes é um problema NP-completo, tanto se realizada com métodos exatos ou aproximados. Porém para muitas aplicações as redes são suficientemente pequenas ou podem ser simplificadas de forma tal que a inferência pode ser realizada com um esforço computacional aceitável. Para aqueles casos nos quais os métodos usuais de inferência são impraticáveis estão sendo desenvolvidos novos algoritmos apropriados para determinadas topologias da rede ou para questionamentos específicos [Heckerman-95].

4.6 Considerações Finais

Diagramas de influência são uma ferramenta apropriada para o desenvolvimento do modelo de decisão. Eles são concisos e intuitivos, proporcionando uma mesma linguagem para quem desenvolve, analisa e utiliza o modelo. Por outro lado, DIs oferecem uma estrutura conveniente para os procedimentos computacionais de resolução de problemas. A estrutura gráfica do modelo contém suficiente informação mesmo sem considerar as tabelas com os valores da variável associada a cada nó.

Enquanto a árvore de decisão representa todas as possíveis combinações de decisões e eventos, o diagrama de influência representa a relação entre os componentes da decisão e as quantidades incertas ou variáveis aleatórias do problema. Devido a isso o diagrama de influência oferece uma representação mais compacta do problema do que a da árvore de decisão. Outra vantagem é que o diagrama de influência representa minuciosamente

qualquer independência condicional entre variáveis, além da sequência temporal das diversas escolhas que constituem o modelo, bem como a disponibilidade de informação ante cada decisão. DIs permitem também calcular facilmente o valor esperado da informação associada a um nó.

A possibilidade de representar as relações entre as variáveis na direção causa-efeito facilita o processo de aquisição do conhecimento e promove o envolvimento do usuário no processo de modelagem e análise, obtendo-se assim uma melhor percepção do problema.

O atual interesse despertado pelos diagramas de influência pode ser constatado nos trabalhos compilados em [ACM-95] [IEEE-93] [Oliver-90]: desenvolvimento, modelos e aplicações onde os DI são utilizados apresentando resultados promissores. O mesmo acontece nos diversos artigos com desenvolvimentos teóricos ou aplicações práticas como [Ndilikiliksha-94] [Shachter-90] [Smith-93] [Ramoni-95], entre outros.

CAPÍTULO 5

FERRAMENTAS DE APOIO À DECISÃO

5.1 Introdução

Este capítulo apresenta diversas ferramentas para processar modelos gráficos de probabilidades em redes, como por exemplo redes de Bayes e diagramas de influências, bem como alguns sistemas implementados para aplicações em domínios específicos. O propósito desta discussão não é apresentar uma revisão dos softwares disponíveis —já que a situação provavelmente terá evoluído ao momento desta publicação— mas mostrar a viabilidade prática desta abordagem, determinar as relações entre estas ferramentas ou aplicativos e a teoria na qual estão baseados, além de fornecer os endereços eletrônicos onde muitos destes softwares estão disponíveis. Através do exposto neste capítulo pode ser evidenciada a maturidade das ferramentas de apoio à decisão baseadas nesta abordagem.

O capítulo está organizado da seguinte forma: a seção 5.2 classifica os sistemas de software que utilizam a abordagem bayesiana. De acordo com esta classificação cada uma das seções seguintes descreve um tipo de sistemas: a seção 5.3 apresenta sistemas que são a implementação de teorias ou algoritmos direcionados para realizar uma tarefa específica, a seção 5.4 descreve extensões gráficas ou teóricas dos DIs convencionais e a seção 5.5 introduz as ferramentas ou shells enquanto a seção 5.6 discute alguns dos sistemas implementados para resolver problemas em áreas específicas. Finalmente a seção 5.7 compreende as considerações finais.

5.2 Sistemas de Software que Usam a Abordagem Bayesiana

Os sistemas de software que utilizam a abordagem probabilística em redes podem ser classificados em dois grandes grupos:

- sistemas gerais para modelar redes de Bayes ou DIs,
- sistemas que utilizam esta abordagem para resolver problemas específicos de um domínio determinado, também denominados sistemas aplicativos.

Por sua vez os sistemas gerais podem ser discriminados segundo o seguinte critério [Matzkevich-95]:

- implementação de teorias ou algoritmos direcionados para realizar uma tarefa específica,
- extensões gráficas ou teóricas dos DIs convencionais,
- ferramentas ou shells que oferecem um ambiente para entrar, editar, refinar e avaliar DIs ou redes de Bayes.

Cada uma destas categorias será apresentada a seguir.

5.3 Algoritmos Especializados

Na primeira categoria pode ser mencionado o algoritmo desenvolvido por Russel et.al. para a construção automática de redes de Bayes esparsas [Russel-90]. Este requisita como entrada um modelo probabilístico não estruturado, as variáveis do modelo e informação sobre as relações causais, o qual é determinado por um especialista da área. O algoritmo auxilia o usuário a construir a rede adicionando um nó a cada passo, de forma que o número de arcos introduzidos para cada nó seja mínima. Este algoritmo foi implementado como parte do sistema IDEAL [Srinivas-90].

O sistema TETRAD auxilia na construção do modelo a partir de dados estatísticos. O algoritmo, baseado nos trabalhos de [Glymour-87] [Spirtes-92], sugere a estrutura causal, dentre um conjunto de estruturas disponíveis, que melhor se ajusta aos dados considerados.

O algoritmo K2 [Cooper-91] constrói uma rede de Bayes a partir de uma base de dados que contém casos do domínio realizando uma busca heurística para a estrutura mais provável. Os dados de entrada são: i) o conjunto de nós, ii) o ordenamento dos nós, iii) o número máximo de predecessores diretos para cada nó, e iv) uma base de dados com casos.

Pearl apresenta uma discussão teórica do problema de construir a rede de Bayes a partir de dados estatísticos [Pearl-91]. De uma forma geral algoritmos para a construção automática da rede (sub-área também denominada *aprendizado do modelo probabilístico*) estão ganhando cada vez mais importância [Almond-96a] [Breese-94].

5.4 Extensões dos Diagramas de Influências

Extensões gráficas ou teóricas dos DIs que auxiliam na aquisição e representação do conhecimento são apresentadas em [Geiger-91] e [Heckerman-91]. As *redes probabilísticas de similaridade* propostas por Heckerman utilizam o princípio de divisão e conquista para diagnosticar em domínios onde o problema consiste em discriminar entre um conjunto de hipóteses. Estas redes modelam o problema em diversos DIs (em lugar de um único diagrama) onde somente um par de hipóteses similares são consideradas de cada vez. Este mecanismo permite transformar automaticamente os diversos DIs num único DI que representa o problema discriminando entre todas as possíveis hipóteses. Este trabalho foi motivado pela observação de que para o médico é freqüentemente mais fácil descrever as diferenças entre duas doenças similares do que entre um conjunto de doenças muito diferentes.

5.5 Ferramentas

A terceira categoria de sistemas de software, constituída pelas ferramentas ou shells, foi desenvolvida para auxiliar a automatizar a aquisição do conhecimento e a construção do modelo. A maior parte destes sistemas tem evoluído de produtos de pesquisa acadêmica para produtos comerciais, como HUGIN [Andersen-89] e GRAPHICAL-BELIEF [Almond-95]. De forma geral estes sistemas oferecem um ambiente para realizar a aquisição do conhecimento independente do domínio, representar e realizar inferências no modelo bem como oferecer apoio à decisão. Alguns deles têm facilidades para automatizar e/ou otimizar a inferência, outros automatizam o processo de construção da rede dada uma entrada não estruturada, e todos fornecem uma interface gráfica para representar o modelo. Uma característica importante no desenvolvimento destas ferramentas é a influência —às vezes a participação ativa— da comunidade científica neste processo, a qual tem um impacto positivo na busca de soluções inovadoras para os problemas ou limitações apresentados por estes sistemas.

As tabelas apresentadas a seguir apresentam as principais características das ferramentas consideradas. Também são descritas as características ausentes ou aspectos não cobertos pela implementação, a plataforma na qual os sistemas estão disponíveis, bem como referências bibliográficas/e-mail para contato (quando disponível). Muitas destas ferramentas são de livre acesso se utilizadas com fins didáticos ou de pesquisa. Dos softwares comercializados geralmente é possível obter um sistema demonstrativo. Informações mais detalhadas sobre endereços completos de contato e a forma de obter as ferramentas de livre acesso ou um demonstrativo das ferramentas comerciais podem ser obtidas em [Almond-96]. Após as tabelas encontra-se um glossário com os termos utilizados para a descrição destes sistemas. Uma primeira discriminação das ferramentas foi realizada segundo estas fossem produtos de pesquisa ou comerciais (tabela 5.1 e tabela 5.2, respectivamente).

Uma revisão abrangente de ferramentas de apoio à decisão (não necessariamente utilizando a abordagem bayesiana) pode ser encontrada em [Buede-94].

Nome	Características	Aspectos não cubertos	Plataforma	Referências/Contato
BAYES	probabilidades, RB em polytrees	funções de crença, utilidades, parâmetro de incerteza, redes (qualquer topologia da rede)	MS DOS	Carnegie Mellon University AI Repository (ai+query@cs.cmu.edu)
BELIEF 1.2	probabilidades, funções de crença, parâmetro de incerteza	utilidades, GUI	CL2	[Almond-95], almond@acm.org
IDEAL	probabilidades, utilidades, DIs	funções de crença, parâmetro de incerteza	Common Lisp	[Srinivas-90], ideal-request@rpal.rockwell.com
MacEvidence	probabilidades, funções de crença	utilidades	Macintosh	[Hisa-98], p-shenoy@ukans.edu
MSBN32	probabilidades, utilidades, base de regras padrão, DIs	funções de crença, DI com uma única decisão	Windows 95, Windows NT	dtg-msbn@microsoft.com
Pulcinella	probabilidades, funções de crença, possibilidades, GUI	utilidades	CL1	[Saffioti-91], [Saffioti-94], asaffio@ulb.ac.be
S-ElimBel	probabilidades	funções de crença, utilidades, GUI	MIT Scheme	nthiery@uci.edu
SPI	probabilidades, utilidades, LEL, explicação, modelos dinâmicos, GUI	funções de crença, documentação	Common Lisp	dambrosi@research.cs.orst.edu
TresBel	funções de crença, GUI	probabilidades	CL1	[Xu-94], hongxu@ulb.ac.be
XBaies 1.3	probabilidades, utilidades, cadeias de grafos	funções de crença	HP, SUN, DEC, PC/Windows	xbaies-list@stats.ucl.ac.uk

Tabela 5.1. Ferramentas de livre acesso

Nome	Características	Aspectos não cubertos	Plataforma	Referências/Contato
Baron	probabilidades, algoritmo exato e simulação, GUI	funções de crença, utilidades	Windows 3.0, C++	kchang@gmu.edu
Demos	DIs hierárquicos, parâmetro de incerteza, links para tabela de cálculos	funções de crença,	Macintosh, Windows	support@lumina.com
DXpress	probabilidades, GUI, utilidades, interface de programação API	funções de crença, parâmetro de incerteza	Windows 3.1, NT, 95, C++	[Horvitz-88], ki@kic.com
Ergo	probabilidades,	funções de crença, parâmetro de incerteza, utilidades,	Macintosh, Windows	info@noeticsystems.com
GRAPHICAL-BELIEF 2.0	probabilidades, funções de crença, parâmetro de incerteza, KBMC	utilidades	Unix, Common Lisp	[Almond-94], [Almond-95], almond@acm.org
HUGIN	probabilidades, utilidades, crítica do modelo, parâmetro de incerteza, cadeias de grafos	funções de crença	Unix, Windows	[Lauritzen-88], [Andersen-89], info@hugin.dk
Netica	probabilidades, utilidades, DIs, API	funções de crença, parâmetro de incerteza	API para todas as plataformas	norsys_info@norsys.com
STRATEGIST	probabilidades, utilidades, LEL	funções de crença	Windows 3.1, NT, 95, C++ API para Unix	dambrosi@prevision.com

Tabela 5.2. Ferramentas comerciais

5.5.1 Glossário

API (Application Programming Interface): Interface para a programação de aplicativos.

Base de regras padrão: utiliza o formato de redes proposto pela Microsoft [Microsoft-96]. Este formato não é amplamente utilizado.

Cadeias de grafos: suporta grafos com uma mistura de arcos direcionados e adirecionados [Lauritzen-89].

CL1: Common Lisp baseado em "Common Lisp: the Language, Version 1" [Steele-84], isto significa que o software pode necessitar algumas adaptações se portado para versões mais recentes.

CL2: Common Lisp padrão baseado em "Common Lisp: the Language, Version 2" [Steele-90], onde a maioria das mudanças no Common Lisp propostas pelo comitê ANSI Standard são descritas. Isto indica que o programa pode ser incompatível com versões anteriores.

Crítica do modelo: indica a habilidade de atualizar a estrutura do modelo na presença de novos dados [Cowell-93].

DIs (Diagramas de Influências): o modelo inclui probabilidades e utilidades.

Explicação: contém facilidades para explicar o comportamento do modelo.

Funções de crença: a ferramenta suporta funções de crença (teoria das evidências de Dempster-Shafer [Lowrance-86], [Shafer-76]) como uma representação de incertezas.

GUI (Graphical User Interface): possui uma interface gráfica para o usuário.

KBMC (Knowledge Based Model Construction): facilidades para auxiliar a construir modelos complexos a partir de regras [Almond-94].

LEL (Local Expression Language): o programa pode representar e utilizar durante a inferência assimetrias, noisy ors, contingências e outros tipos especiais de casos do modelo [D'Ambrosio-91].

Modelos dinâmicos: possui facilidades para a criação e modificação de modelos dinâmicos, particularmente importante para modelos de raciocínio temporal.

Parâmetro de incerteza: permite ao usuário especificar uma lei para um parâmetro de uma distribuição de probabilidades ou função de crença [Almond-95] [Spiegelhalter-90].

Possibilidades: Suporta a teoria das possibilidades [Dubois-88] [Zadeh-78] como uma representação alternativa para as incertezas.

Probabilidades: suporta somente probabilidades para representar incertezas, geralmente segundo o formalismo de Pearl [Pearl-88].

Redes: significa que o programa permite toda rede exprimível como um grafo acíclico direcionado, e não somente como árvores.

Redes de Bayes (RB): modelos probabilísticos que utilizam um grafo acíclico direcionado para representar a independência condicional entre as probabilidades associadas a cada variável [Pearl-88].

RB em polytrees: restrito a modelos nos quais a correspondente rede de Bayes pode ser expressa como uma árvore.

Redes de crenças: termo utilizado para abranger todos os modelos probabilísticos baseados em redes, como por exemplo redes de Bayes, diagramas de influências, cadeias de Markov, etc.

Utilidades: inclui uma representação para as preferências além das incertezas. Ou seja, a ferramenta suporta redes de Bayes e também DIs.

5.6 Aplicações

Uma das vantagens da abordagem bayesiana é que este tipo de representação gráfica das probabilidades oferece uma base apropriada para o desenvolvimento de sistemas baseados em conhecimento. SBCs que utilizam DIs ou redes de Bayes são também denominados *sistemas normativos*, por estarem baseados nos princípios da teoria da decisão. Os vários sistemas implementados, muitos deles na área médica, têm mostrado a eficiência da abordagem bayesiana para construir SBCs quando existem incertezas no domínio de aplicação. Alguns dos sistemas mais significativos são descritos a seguir.

Pathfinder: este sistema médico diagnostica doenças do sistema linfático. O diagnóstico de doenças através da biopsia do nódulo linfático é uma das tarefas mais difíceis na patologia cirúrgica. Ao mesmo tempo a discriminação de características benignas ou malignas é crítica para a escolha do tratamento adequado (quimioterapia, radiação ou cirurgia). Isto demonstra a importância da análise e os benefícios possíveis se o desempenho neste tipo de diagnose for melhorado. Durante o projeto do Pathfinder foram propostas e usadas as redes de similaridade [Heckerman-91], já descritas na seção 5.4, para auxiliar a aquisição do conhecimento necessário a fim de construir a variável das doenças com seus 63 possíveis valores. O sistema inclui aproximadamente 30 diferentes tipos de doenças malignas e a mesma quantidade de outros males que podem ser facilmente confundidos com linfomas malignos [Heckerman-92], [Heckerman-92a]. Quase todas as relações na rede são probabilidades condicionais e o sistema utiliza um algoritmo de inferência exato. Como resultado deste projeto de pesquisa foi elaborado o sistema comercial INTELLIPATH o qual integra o raciocínio probabilístico com uma série de ferramentas de apoio, como base de dados com imagens do microscópio, textos com informações sobre as doenças, etc.

MUNIN: este sistema médico serve para diagnosticar disfunções neuromusculares e foi desenvolvido na Universidade de Aalborg, na Dinamarca [Andreassen-87]. MUNIN foi implementado em HUGIN possuindo cerca de 1000 nós, cada um com até 7 possíveis estados e o algoritmo exato de inferência utilizado demora apenas alguns segundos para resolver a rede.

QMR-BN: QMR (Quick Medical Reference) é uma base de dados e ferramenta de diagnóstico para medicina interna que faz parte do projeto Internist-1 [Shwe-91]. QMR-DT (Quick Medical Reference-Decision Theory) é um sistema de regras que complementa a base de dados permitindo decisões a respeito de tratamentos e custos ou modelos de valor. QMR-BN é uma reformulação de QMR-DT utilizando redes de Bayes. Este sistema, junto com MUNIN e Pathfinder, é um dos trabalhos mais significativos na área médica.

Rachel: este sistema possui um DI geral para representar o modelo de casais com problemas de fertilidade [Holtzman-89]. À medida que os valores das variáveis para um determinado casal são introduzidos o sistema gera um DI específico o qual pode ser resolvido pelos algoritmos tradicionais de inferência.

ARCO1: Sistema para prever o preço do petróleo baseado numa rede de Bayes que modela o mercado mundial do petróleo ao longo de um ano. As variáveis econômicas e as variáveis das políticas subjetivas de preços estão relacionadas tanto probabilística como algebricamente. O algoritmo de inferência utiliza simulação de Monte Carlo.

IES: Um sistema para reconhecimento de imagens, detecção de sinais de radar e reconhecimento militar. IES opera formando uma hierarquia de DIs que são incremental e dinamicamente atualizados para interpretar as imagens. Ele utiliza informações visuais como forma de detectar o movimento de tropas e outros objetos militares.

5.7 Considerações Finais

Neste capítulo foram descritos alguns dos sistemas mais representativos que utilizam a abordagem bayesiana para representar o modelo. A rapidez com a qual as redes de Bayes e os DIs têm se firmado como fundamento para a construção de SBCs pode ser constatada pelo recente desenvolvimento dos sistemas aplicativos. Muitos outros sistemas estão sendo construídos em diversas áreas e conseqüentemente a lista de softwares representativos deve crescer rapidamente. Porém, os sistemas descritos neste capítulo pertencem à primeira geração, merecendo portanto menção especial.

Das ferramentas descritas na seção 5.5 é interessante destacar que estas continuam incorporando novas funcionalidades à medida que a teoria subjacente vai sendo depurada. A maior parte destas opera com redes de Bayes e apenas a metade permite representar DIs. Quase todas elas suportam probabilidades para representar as incertezas, entanto menos de um terço destas ferramentas permite também utilizar funções de crenças e apenas Pulcinella permite também utilizar possibilidades. Uma característica importante é a facilidade da ferramenta explicar o modelo; esta propriedade aparece somente em um dos casos considerados e será tema de discussão no próximo capítulo.

CAPÍTULO 6

EXPLICAÇÕES EM SAD: UMA PROPOSTA

6.1 Introdução

Geralmente é difícil explicar sumariamente uma decisão complexa. No entanto, para um sistema de apoio à decisão ser aceito pelo usuário ele deve ser capaz de mostrar o modelo de decisão utilizado e fornecer uma justificativa para a recomendação resultante. Médicos, por exemplo, tendem a não confiar em diagnósticos gerados pelo computador a menos que consigam confirmar o porquê destes diagnósticos [Teach-81].

Muitas decisões devem ser tomadas dentro de um tempo determinado e se baseando em informações incompletas ou não totalmente confiáveis. Sistemas computacionais de apoio à decisão conseguem dar assistência para esse tipo de situação, mas o sistema também deve ser capaz de fornecer um resumo do problema e uma explicação para suas recomendações. Além de incrementar a confiança do usuário no sistema uma facilidade de explicação também pode ser útil como uma ferramenta de depuração para quem desenvolve o sistema. Explicações geradas pela máquina fazem aparentes os erros que podem passar despercebidos na codificação do sistema. Elas também são um componente importante na arquitetura de sistemas tutores.

Na análise da decisão os resultados mais importantes não são os números atingidos ou as decisões recomendadas, mas auxiliar o cliente a obter uma melhor percepção do problema [Henrion-91]. Este entendimento advém de compreender porque uma decisão é mais recomendável que outra e quais os fatos e incertezas que são mais críticos à conclusão. Este processo é geralmente promovido através do envolvimento do usuário no processo de modelagem e análise. Sem um entendimento básico da análise este pode não aceitar como confiável a ação recomendada. Esta perspectiva enfatiza a importância de explicações claras a respeito da análise e as afirmações realizadas sobre o modelo para que estas sejam úteis.

A problemática da explicação em sistemas bayesianos é relativamente nova quando comparada com a dos sistemas baseados em regras. Conseqüentemente, como se pode constatar através do exposto no capítulo 5, são poucos os sistemas bayesianos que incorporam esta característica. Em particular é o objetivo deste trabalho desenvolver facilidades de explicação para sistemas baseados em DIs.

O capítulo está estruturado da seguinte forma: a seção 6.2 aborda o problema da explicação em sistemas baseados em regras. Na seção 6.3 são descritas outras metodologias desenvolvidas para oferecer explicações em sistemas bayesianos, bem

como outros trabalhos que fornecem explicações em sistemas de decisão baseados nos princípios da teoria da decisão. A seção 6.4 apresenta a base da proposta do esquema de explicação para DIs, a saber: análise de sensibilidade e redes probabilísticas qualitativas. A seção 6.5 compreende as considerações finais.

6.2 Explicação em Sistemas Baseados em Regras

Os primeiros trabalhos de explicação em SEs sugeriam a utilização das regras de resolução do problema (contidas na base de conhecimento definida pelo especialista) para explicar o raciocínio do sistema. Durante a década passada diversas formas de usar esta BC para justificar as conclusões e ações do sistema foram pesquisadas. Embora importantes avanços tenham sido realizados estas explicações geradas automaticamente sofrem diversas limitações. A premissa de todos estes trabalhos é que a base da explicação é a descrição da linha de raciocínio do SE.

Entretanto, para as explicações serem efetivas o seu formato deve ser adequado ao tipo de usuário, seu nível de conhecimento e linha de raciocínio. Uma das deficiências da abordagem tradicional de explicação em SEs baseados em regras é o fato das explicações, estarem estritamente ligadas à linha de raciocínio do SE. Deste modo são mais apropriadas para o engenheiro de conhecimento que está depurando a base de conhecimento, mas nem sempre são esclarecedoras para o usuário leigo ou com menor conhecimento de como as informações são processadas pelo motor de inferência [Wick-92]. Outras explicações podem ser excessivamente longas e, embora corretas e compreendidas pelo usuário, não contribuem para aumentar a confiança deste nas conclusões do sistema especialista [Wolverton-95]. Este problema ocorre geralmente quando o usuário é o próprio especialista.

Nesta abordagem a forma de explicar o comportamento do sistema é traduzindo o código e a sua execução em expressões ou frases mais próximas à linguagem do usuário. Isso permite evidenciar as ações tomadas pelo sistema de uma forma consistente com o seu comportamento. Este tipo de explicação tem sido bastante empregada e de forma geral é adequada quando o usuário é o Engenheiro de Conhecimento. Neste caso ela serve para depurar a BC. O problema é que esta abordagem não permite justificar o porque das ações recomendadas serem as mais apropriadas. Ela não descreve o raciocínio que leva à escolha de cada ação ou recomendação. Isto requer conhecimento sobre as decisões de projeto do sistema e sobre o domínio do problema o qual não necessariamente deve estar explicitamente representado para o sistema funcionar corretamente.

A natureza de uma explicação efetiva depende também do tipo de usuário. O engenheiro de conhecimento requer uma facilidade de explicação que mostre quais os passos seguidos pelo sistema para atingir um certo resultado. Esta facilidade, porém, pode ser inapropriada para um usuário leigo no domínio. Uma explicação para este usuário deve incrementar a sua confiança no sistema e auxiliá-lo a entender as consequências e o porquê da conclusão do sistema. Muitas vezes o usuário necessita esclarecimentos do domínio do problema que são independentes das conclusões do sistema. Por exemplo o paciente pode ter dúvidas se realmente é tão importante tratar uma dada doença pois desconhece os sintomas e as consequências desta, independentemente do tratamento indicado pelo SE.

O estado atual da pesquisa sobre explicação automática de SEs envolve problemas de resolução complexa. Por exemplo, diversos sistemas modelam o tipo de usuário, realizam planejamento da explicação e incorporam interfaces em linguagem natural, entre outros. Porém a maior parte destes sistemas considera a tarefa da explicação como a tarefa de apresentar ao usuário a linha de raciocínio do sistema de uma forma mais natural, isto é, traduzir a linha de raciocínio para uma linguagem mais fácil de ser entendida pelo usuário.

Numerosos esforços foram realizados na tentativa de melhorar este tipo de explicações. Três idéias principais podem ser salientadas neste processo:

1. A descrição da execução do SE pode ser utilizada para fornecer uma explicação da forma na qual o sistema resolve o problema. *Por que e como* são o tipo de perguntas que o sistema consegue responder mostrando e utilizando a sua linha de raciocínio. Estas são as facilidades básicas de explicação que a maior parte dos SBCs incorporam até hoje.
2. Um modelo do domínio pode ser usado para explicar as regras usadas pelo SE. Swartout [Swartout-83] introduz o sistema denominado XPLAIN explicitamente designado para fornecer explicações usando os "princípios do domínio" e o "raciocínio do domínio" para poder gerar as justificativas das regras.
3. Explicações podem ser geradas a diferentes níveis. Por exemplo, o sistema NEOMYCIN [Clancey-84] permite mudar o foco da explicação do domínio do problema para o conhecimento sobre as estratégias usadas na resolução do problema.

Também foi observada a necessidade de processar a linha de raciocínio do sistema antes de apresentá-la como uma explicação. Por exemplo, MYCIN poda as folhas terminais da linha de raciocínio [Wallis-82]. Vários outros projetos também interpretam o processo de explicação como o processo de podar, aumentar ou traduzir a linha de raciocínio do sistema.

Recentemente novas abordagens consideram o problema da explicação não mais como um "traduzir" a linha de raciocínio do sistema, mas como um problema independente a ser resolvido. A explicação consiste agora em construir uma boa justificativa da conclusão do sistema (independentemente de como este atingiu a conclusão). O modelo do domínio é usado para gerar uma justificativa causal que suporte a conclusão dada pelo SE [Ryan-88] [Moore-89].

Numa outra abordagem, na *explicação reconstitutiva*, [Paris-88] [Wick-89] [Wick-92] a explicação envolve a completa reconstrução do raciocínio do sistema para chegar na conclusão. Isto é:

a explicação pode ser considerada como um processo complexo a ser resolvido, bem diferente do processo original de resolução do problema do domínio.

Desta forma uma nova linha de raciocínio para obter a explicação deve ser gerada e isso é realizado a partir de uma outra base de conhecimento. Isto tem maiores custos

associados, pois duas BCs devem ser geradas e mantidas, oferecendo por outro lado a vantagem de poder gerar diversas explicações para uma mesma conclusão dependendo do grau de conhecimento do usuário. Esta abordagem é também mais consistente com a observação de que usualmente quando as pessoas explicam a solução de um problema utilizam um conhecimento diferente daquele usado para resolver o problema.

Uma outra abordagem ainda mais inovadora produz explicações a partir da crítica realizada as conclusões obtidas pelo usuário, *sistemas especialistas de crítica* [Miller-86] [Langlotz-89]. Especificamente esta abordagem foi desenvolvida na área médica [Cooper-84] [Langlotz-83] [Miller-82] [Miller-83] [Rennels-86]. De uma forma geral o sistema primeiro consulta o usuário sobre a sua recomendação para o caso determinado. Se a conclusão do médico coincide com a do sistema nenhuma explicação é gerada, a menos que isso seja explicitamente requerido. Se o tratamento recomendado pelo médico difere daquele gerado pelo sistema então este último gera uma explicação (na realidade uma *crítica* a recomendação do médico e justifica o porquê da sua própria recomendação). Neste caso a explicação se adequa automaticamente às necessidades do usuário pois o sistema não explica toda a sua linha de raciocínio mas somente as diferenças entre as duas conclusões.

6.3 Explicação em Redes Bayesianas: Outras Metodologias

São várias as metodologias desenvolvidas a fim de fornecer explicações em redes bayesianas. Algumas delas, consideradas mais relevantes, bem como outros trabalhos relacionados com o problema de explicação em sistemas baseados na teoria da decisão, são descritas a seguir.

[Jimison-90, 92]

No trabalho de Jimison et al. [Jimison-90, 92] é descrito um modelo de decisão genérico para pacientes com "angina pectoris". Esse sistema auxilia na escolha do tratamento adequado: cirurgia ou tratamento com drogas. O modelo genérico é representado como um diagrama de influência onde as características específicas do paciente, probabilidades e preferências, são tratadas como variáveis aleatórias. A distribuição inicial para essas variáveis é tirada da literatura sobre uma população de pacientes com sintomas de angina. Com os dados de um paciente específico (idade, resultados de testes médicos, tipo de atividades preferidas: sedentárias ou não, etc.) os parâmetros da rede (probabilidades e utilidades) são ajustados individualmente. Este esquema de um DI geral do problema cujos valores são ajustados segundo os dados de entrada do usuário é similar ao já utilizado em Rachel [Holtzman-89]. Essa representação oferece a vantagem de uma seleção automática das probabilidades para cada paciente (segundo os seus dados específicos), como também a vantagem do recurso de explicação na comparação do modelo específico do paciente com o modelo genérico. Três métricas são utilizadas nos recursos de explicação:

- *valor de contribuição*, definido como a contribuição de uma variável para a mudança na utilidade esperada do modelo específico do paciente com respeito ao modelo genérico,
- *desvio da média*, para uma variável no valor específico do paciente com respeito à distribuição genérica,
- *sensibilidade*, considera se a decisão é sensível a mudanças no valor de uma variável na distribuição de probabilidades específica do paciente.

As três métricas são geradas na comparação do modelo específico do paciente com o modelo genérico, e são utilizadas por uma interface gráfica para enfatizar os componentes que contribuem significativamente para a decisão, os que diferem daqueles usualmente observados ou as variáveis que são sensíveis para o tratamento escolhido.

Os outros trabalhos aqui revisados consideram a problemática da explicação em redes bayesianas, e não especificamente em DIs. Geralmente neste caso deseja-se explicitar como as evidências (os valores conhecidos de uma ou mais variáveis aleatórias) influenciam a conclusão obtida através da inferência bayesiana, ou seja, quais os fatos que contribuem, e em que medida, para o sistema atingir uma conclusão determinada. Por se tratar de redes de Bayes, e não de DIs, nem preferências nem decisões são consideradas na explicação. A seguir vários deles são descritos.

[Reggia-85]

Reggia e Perricone implementam um sistema de diagnóstico médico baseado no teorema de Bayes. Esse sistema de classificação ainda não utiliza os princípios dos sistemas Bayesianos baseados em redes: simplesmente o teorema de Bayes é utilizado para calcular a doença mais provável dados os sintomas do paciente. A abordagem de explicação utilizada analisa as probabilidades anteriores e condicionais para calcular quais dos sintomas ou evidências causaram a ordem relativa das doenças mais prováveis.

[Elsaesser-89, 90]

Elsaesser e Henrion utilizam uma abordagem linguística para explicar a inferência bayesiana. A explicação é gerada utilizando os padrões indutivos de Polya [Polya-54] para descrever a atualização da probabilidade posterior. Um único padrão é utilizado, onde as variáveis são preenchidas com frases selecionadas representando as probabilidades ou mudanças no valor das probabilidades do problema.

[Henrion-91a]

Henrion e Druzel consideram a explicação no contexto de grafos direcionais e enfatizam "probabilidades linguísticas", isto é, a conversão das probabilidades em frases tais como "muito provável", "difícil de ocorrer", etc. Eles consideram grafos simplesmente conexos com nós binários e fornecem explicações textuais do comportamento dinâmico da rede. Isto é, o efeito qualitativo da evidência na rede através das diferentes variáveis é descrito textualmente.

[Suermondt-93]

Suermondt e Cooper também consideram grafos direcionados e discutem diversas métricas para a influência e conflitos na evidência. A identificação do fluxo da evidência na rede é semelhante ao posteriormente proposto por Madigan [Madigan-95], mas as explicações são apresentadas num formato textual.

[Madigan-95]

Madigan utiliza uma abordagem visual para explicitar como a evidência se propaga através de uma rede de Bayes. A métrica utilizada para avaliar a propagação da evidência são os "pesos de evidência" (weight of evidence) [Good-85] e o próprio grafo é utilizado para fornecer um cenário para os fluxos de evidência.

Para decidir quais variáveis têm maior influência na conclusão obtida, os pesos de evidência são calculados. Uma planilha com gráfico de barras é utilizada para mostrar o

peso da evidência com que cada variável contribui na variável objetivo. Essas informações também são representadas na rede utilizando uma escala de cores. Cada nó é colorido segundo o peso de evidência com o qual ele contribui.

Para responder porque uma particular evidência tem tanta (ou tão pouca) influência na conclusão, o fluxo da evidência ao longo da rede é calculado e exibido no grafo através de cores e diferentes amplitudes nos vértices da rede.

Os trabalhos descritos a seguir objetivam explicar o resultado de sistemas de decisão, também baseados nos princípios da teoria da decisão, mas que não utilizam a abordagem bayesiana.

[Klein-94]

Klein e Shortliffe apresentam estratégias para explicar automaticamente decisões baseadas na teoria de valor para multiatributos, ramo da teoria da decisão que estuda como modelar funções de valor frente a objetivos ou preferências conflitantes.

[Langlotz-88]

Langlotz et. al. desenvolveram um programa que utiliza técnicas de raciocínio simbólico para gerar explicações qualitativas do resultado da análise da decisão. O método implementado explica as diferenças na utilidade esperada entre os diferentes ramos da árvore de decisão. Primeiro, relações matemáticas inerentes à estrutura da árvore são usadas para determinar assimetrias na estrutura desta ou inequações entre variáveis de decisão análogas, responsáveis pelas diferenças na utilidade esperada. Depois é selecionada uma técnica de explicação e esta é aplicada às variáveis mais significativas, criando uma expressão simbólica que justifique a decisão. Finalmente a expressão simbólica é convertida à linguagem de texto (inglês). A explicação gerada justifica a desejabilidade da escolha com maior utilidade esperada, sem usar fórmulas matemáticas ou incluir valores de probabilidades e utilidades.

6.4 Proposta de um Esquema de Explicação para DIs

Um motivo fundamental na construção de um modelo matemático ou estatístico é explicar sistemas complexos do mundo real. Conseqüentemente as explicações, isto é, descrições dos mecanismos que constituem tal modelo, formam uma parte importante da validação, exploração e uso do modelo. O objetivo deste trabalho é propor e implementar um esquema de explicação para sistemas de apoio à decisão, ou sistemas baseados em conhecimento, que utilizam diagramas de influências como mecanismo de representação do conhecimento.

Pelo exposto na seção anterior somente o trabalho de Jimison [Jimison-90, 92] aborda a problemática da explicação em DIs, e isto é realizado comparando o modelo específico gerado para um determinado paciente com o DI geral (criado com dados obtidos da literatura para uma extensa população de pacientes). Em outras palavras a explicação baseia-se nas diferenças entre um modelo específico e o modelo genérico do problema. Em contraposição, o nosso interesse é originar explicações ainda quando não existe um modelo genérico do problema.

À seguir são descritos os princípios das facilidades de explicação [Castiñeira-96] propostas e implementadas no sistema IDEX –Influence Diagrams with EXplanation–

[Castiñeira-96a], ferramenta gráfica para a representação e avaliação de redes bayesianas de decisão.

6.4.1 Análise de Sensibilidade

Na modelagem e avaliação do problema decisório nem todos os benefícios possíveis são obtidos se o mecanismo de representação somente permite encontrar a estratégia ótima. Facilidades para entender o modelo e as conclusões do sistema também são necessárias. Uma das principais vantagens dos mecanismos de explicação é a capacidade de responder perguntas do tipo *O que aconteceria se...? (What if?)*. Isto pode ser atingido realizando-se uma análise de sensibilidade nas variáveis do modelo, como discutido a seguir através de um exemplo.

Exemplo 6.1

Considere-se um problema clássico da literatura: o da perfuração de um poço de petróleo [Raiffa-68], [Shachter-86]. Um prospector deve decidir se realiza ou não uma perfuração à busca de petróleo. Antes disso ele pode realizar um teste cujo resultado tem boas probabilidades de acerto, mas não é determinístico e depende da quantidade de petróleo existente. O objetivo é maximizar o ganho, e isto depende da quantidade de petróleo achada assim como das decisões de se realizar ou não o teste e a perfuração (os custos associados devem ser contabilizados). A figura 6.1 apresenta o sistema IDEX mostrando a representação do exemplo como um DI.

No exemplo considerado pode ser interessante responder às seguintes questões: *O que aconteceria se a possibilidade de haver petróleo fosse maior? O que aconteceria se o Teste fosse menos preciso?* Tais questões são respondidas realizando-se uma análise de sensibilidade –variando os valores determinados para uma ou mais variáveis de uma forma sistemática e repetindo os cálculos para determinar se a decisão ótima se modifica.

Figura 6.1 Interface gráfica do sistema IDEX mostrando o exemplo 6.1.

Uma possível análise de sensibilidade é feita modificando o valor de uma única variável e recalculando a utilidade esperada de cada estratégia [Pauker-87]. No exemplo acima descrito considere-se a variável Resultado do Teste, a qual pode assumir os valores {sim, não}, ou seja, uma recomendação positiva ou negativa de realizar a perfuração. Das quatro possíveis estratégias:

1. teste & perfuração: realizar o teste e a perfuração,
2. ~teste & perfuração: não realizar o teste mas a perfuração,
3. teste & ~perfuração: realizar o teste mas não perfurar,
4. ~teste & ~perfuração: não realizar o teste nem a perfuração,

considerem-se só as duas primeiras. A figura 6.2 mostra como o valor da utilidade esperada de cada estratégia muda segundo os diferentes valores da variável Resultado do Teste. Nesta figura também pode-se observar que as linhas das diferentes estratégias se intersectam em pontos que determinam os denominados *valores críticos da variável* (threshold values). Se a probabilidade do resultado do teste ser negativo dado que a quantidade de petróleo é nula for representada como $P(RT=não|QP=seco)$, então para $P(RT=não|QP=seco) < 0.5$ a estratégia ótima é ~Teste & Perfuração, caso contrario Teste & Perfuração. O valor 0.5 é um valor crítico da variável considerada, o qual é um resultado lógico: se o resultado do teste tiver menor precisão do que 50% não é aconselhável realizá-lo. De forma geral se a variável for menor do que um valor crítico a estratégia ótima é diferente da obtida para valores maiores que o valor crítico. Isto é, os valores críticos indicam se e quando mudanças numa determinada variável modificam a decisão ótima. Variáveis cuja mudança de valores mudam a estratégia ótima são denominadas *variáveis sensíveis*.

O tipo de análise de sensibilidade acima descrito é denominado *análise de uma via* (one way sensitivity analysis) e serve para determinar as variáveis sensíveis e os seus valores críticos. Isto é, quais variáveis influem no resultado se modificadas e de que forma. No momento de ajustar ou refinar os valores das variáveis especial atenção deve ser dada às variáveis sensíveis.

Figura 6.2 Análise de sensibilidade para a variável Resultado do Teste . As duas curvas representam duas possíveis estratégias. Se a probabilidade $P(\text{Resultado do Teste}) < 0.5$ a estratégia ótima é ~Teste & Perfuração, senão Teste & Perfuração.

No sistema IDEX também foi implementado um outro mecanismo do tipo *What if* que permite modificar arbitrariamente os valores de todas as variáveis do modelo e comparar a estratégia ótima desta versão com respeito aos valores originais.

6.4.2 Redes Probabilísticas Qualitativas

A outra pergunta usual frente a uma conclusão do sistema é *Por que?*: "Por que esta é a decisão ótima?" Na abordagem aqui descrita a primeira instância de tal pergunta implica descrever o modelo normativo da teoria da decisão, i.e. a decisão ótima é aquela que maximiza a utilidade esperada. O usuário, porém, geralmente deseja uma explicação no contexto do problema. Esta facilidade pode ser implementada utilizando Redes Probabilísticas Qualitativas —RPQ.

Redes probabilísticas qualitativas [Wellman-90] [Wellman-90a] são DIs onde os valores numéricos das probabilidades são substituídos por relações qualitativas entre as variáveis. Estas relações podem ser consideradas como limitantes dos valores das probabilidades, expressando inequações que devem ser satisfeitas pelos diversos elementos da rede. Os tipos de nós e as propriedades das dependências topológicas numa RPQ são equivalentes aos dos DIs. Enquanto um DI representa uma particular distribuição conjunta para as variáveis consideradas, uma RPQ representa uma família de distribuições que verifica as correspondentes restrições. O objetivo da inferência em RPQs é obter restrições na forma da estratégia de decisão ótima. Deste modo, esse modelo, com o seu próprio mecanismo de inferência, sugere-se apropriado para descrever qualitativamente o impacto de uma ou mais variáveis sobre outra variável objetivo.

Dois tipos de relações são definidas entre os nós de uma RPQ: *influências qualitativas* e *sinergias qualitativas*. Seja F_c a distribuição de probabilidade acumulada de c em função dos seus predecessores. Cada uma das relações acima mencionadas restringe os valores relativos de F_c segundo as variações de um (dois) dos predecessores de c . Ambas as relações podem ser interpretadas como versões probabilísticas da condição de monotonicidade nas derivadas parciais de primeira e segunda ordem respectivamente.

Uma *influência qualitativa* positiva (negativa) de a em c significa que a distribuição de c dada a é não decrescente (não crescente) para todo a , no sentido de dominância estocástica de primeira ordem. A notação

$$N^\delta(a,c), \delta \in \{+,-,0,?\}$$

é utilizada para denotar tal influencia qualitativa, onde $N^?$ sempre verifica e N^0 verifica quando as variáveis a e c são condicionalmente independentes, ou seja, quando no DI associado não existe um arco entre estas variáveis. Por convenção a influência N^0 não é explicitada na representação gráfica da RPQ.

Definição 6.1

Uma rede probabilística qualitativa G verifica $N^\delta(a,c)$ se e somente se para todo valor c_0 de c , x sendo os predecessores de c em G com $x \neq a$, e $a_1 > a_2$ para a :

$$F_c(c_0 | a_1 x) R_\delta F_c(c_0 | a_2 x),$$

onde R_δ é $\leq, \geq$ ou $=$ quando δ é $+, -,$ ou 0 . Por exemplo, para $N^+(a,c)$ corresponde R_+ que é a relação $\leq$. $R^?$ é a relação completa, ou seja $N^?$ sempre verifica.

A variável a influencia positivamente o nó de valor v , $N^+(a,v)$, se e somente se a função de utilidade u satisfaz (com as mesmas condições para a , e x):

$$u(a_1, x) \geq u(a_2, x)$$

O segundo tipo de relação, *sinergia qualitativa*, descreve a influência de duas variáveis a e b numa terceira c . Esta relação mostra como o incremento numa variável tem maior efeito para maiores valores da segunda. Em outras palavras o efeito combinado de incrementar as variáveis é maior do que o efeito das duas separadamente.

Definição 6.2

Uma rede probabilística qualitativa G verifica a sinergia qualitativa $S^\delta(\{a,b\},c)$ se e somente se para todo valor c_0 de c , x sendo os predecessores de c em G com $x \neq a$ e $x \neq b$, $a_1 > a_2$ para a , e $b_1 > b_2$ para b , satisfaz:

$$[F_c(c_0 | a_1 b_1 x) - F_c(c_0 | a_2 b_1 x)] R_\delta [F_c(c_0 | a_1 b_2 x) - F_c(c_0 | a_2 b_2 x)].$$

A sinergia qualitativa na variável de valor v , $S^\delta(\{a,b\},v)$, é verificada se e somente se a função de utilidade u satisfaz (com as mesmas condições para a , b , e x):

$$[u(a_2, b_1, x) - u(a_1, b_1, x)] R_\delta [u(a_2, b_2, x) - u(a_1, b_2, x)]$$

Exemplo 6.2

A RPQ da figura 6.3 representa um modelo de decisão genérico. Este modelo inclui duas variáveis de decisão, observar e atuar, representando:

- observar: a escolha de se realizar uma observação (por exemplo se realizar ou não um teste prévio ou um exame de diagnóstico, etc.),
- atuar: executar ou não uma ação (realizar-se ou não a perfuração do exemplo 6.1, qual tratamento médico escolher, que tipo de aplicação financeira realizar, etc.).

Estas duas variáveis podem ser proposicionais (escolher-se ou não a ação) ou podem admitir uma escala de valores. Como consequência da ação existem custos e benefícios (ambos são nós de probabilidades). O nó atuar influencia de forma positiva os custos e benefícios. Por exemplo se atuar significa escolher um tratamento médico então ambas as consequências deste (benefícios, e custos) podem sofrer um incremento (no sentido de N^+) segundo a eficiência do tratamento (segundo o tratamento realizado podem-se obter maiores benefícios e maiores custos). Por definição os custos influenciam negativamente o nó de valor (sempre que possível deseja-se minimizar os custos) enquanto a influência dos benefícios é positiva (o objetivo principal é curar o paciente).

Existe também um estado α desconhecido, por exemplo a gravidade da doença considerada, o qual afeta a avaliação dos custos e benefícios. A relação direta de α na variável de valor é indeterminada ($N^?(\alpha, v)$). Embora a variável α não seja diretamente observável, existe uma outra variável, informação, a qual representa um maior grau de conhecimento sobre α (por exemplo o resultado de alguma análise médica confirmando um determinado sintoma associado à doença ou à gravidade desta). A variável α se relaciona com informação através de uma influência positiva. A linha tracejada ligando-a à variável de decisão observar indica que se a observação é realizada o valor de informação é acessível no momento de decidir a ação. O ato de observar tem custos, mas nenhum outro benefício que o de obter mais informação.

Figura 6.3 Modelo de decisão genérico.

As relações sinérgicas de α com os custos e benefícios é positiva embora a relação direta de α na variável de valor seja indeterminada ($N^?(\alpha, v)$). No caso considerado a sinergia positiva de α e benefícios significa que quanto maior a severidade da doença maior o valor da cura. A sinergia positiva com os custos descarta a possibilidade de α agravar o efeito negativo do tratamento, ou seja, no exemplo da terapia a severidade da doença é no máximo neutra com respeito aos custos.

Este modelo de decisão genérico é inadequado para justificar uma particular escolha da variável atuar, porém o modelo é forte o suficiente para fornecer restrições úteis na estratégia ótima como uma função da informação disponível. Utilizando o algoritmo de inferência gráfica proposto por [Wellman-90] a RPQ da figura 6.3 pode ser reduzida à forma apresentada na figura 6.4.

Figura 6.4 Modelo de decisão genérico reduzido a três variáveis.

Como observado na figura 6.4, o modelo não permite estabelecer uma influência determinística entre as variáveis atuar e valor. O resultado interessante da figura 6.4 é a sinergia positiva $S^+({\text{informação}}, {\text{atuar}})$, a qual permite estabelecer a propriedade de monotonia para atuar e informação. Esta propriedade estabelece que o valor ótimo de atuar é uma função não decrescente de informação. Se a observação é realizada e a informação é disponível, então o nível de atuar deve ser elevado quando a informação indica maiores valores dos benefícios esperados. Por exemplo, se a estratégia selecionada é $\langle \text{informação} = \text{info1} \text{ implica } \text{atuar} = a1 \rangle$ então a propriedade de monotonia requer que para $\text{informação} = \text{info2}$ e $\text{info2} < \text{info1}$ seja escolhido $\text{atuar} = a2$ com $a2 < a1$. Se atuar é uma variável proposicional o resultado da sinergia implica que a estratégia ótima é realizar a ação se e somente se informação excede um valor crítico. Desta forma alguns dos resultados obtidos através das RPQs se complementam com aqueles obtidos pela análise de sensibilidade.

Considere-se novamente o exemplo 6.1 da página 58. Dentro deste contexto e utilizando o algoritmo de inferência gráfica o seguinte tipo de afirmações podem ser geradas: *maior probabilidade de <Petróleo> aumenta a probabilidade de <Resultado do Teste (positivo)>*, *<Perfurar> é uma função não decrescente de <Resultado do Teste>*, i.e. para maiores valores da variável <Resultado do Teste> o valor de <Perfurar> deve ser incrementado. No caso de valores binários a estratégia ótima é realizar <Perfurar> se e somente se <Resultado do Teste> excede um valor crítico.

As conclusões genéricas que as RPQs permitem realizar sobre o problema específico servem para limitar o espaço da estratégia ótima a um alto nível de abstração e para entender qualitativamente as relações entre as ações e eventos que fazem parte do modelo.

6.5 Considerações Finais

Neste capítulo foram descritas diversas abordagens utilizadas para construir facilidades de explicação para sistemas de apoio à decisão baseados em redes de crenças ou na teoria da decisão. Somente o trabalho de Jimison et al. [Jimison-90] aborda a problemática da explicação para sistemas baseados em DIs, ou seja, ele também considera as preferências e decisões no processo de explicação. Neste trabalho porém, as medidas utilizadas para gerar as explicações são desvios do modelo obtido para um paciente específico com respeito a um modelo genérico. Este modelo genérico é constituído por um DI genérico do domínio, construído com dados da literatura sobre uma população de pacientes com a determinada doença. Com os dados do paciente específico este DI genérico é individualmente ajustado para o caso considerado. Ou seja, nesta abordagem é necessário já ter representado, e da forma adequada, o conhecimento do domínio do problema em questão. Este esquema não é adequado para uma ferramenta genérica de apoio à decisão baseada em DIs. Desta forma o presente trabalho tem como objetivo apresentar e mostrar a viabilidade de um esquema de explicação para DIs mesmo quando não existe uma base de conhecimento para o domínio do problema.

Dois mecanismos de explicação para uma ferramenta de apoio à decisão foram aqui descritos: análise de sensibilidade e as redes probabilísticas qualitativas foram apresentadas como meio para gerar conclusões genéricas e para entender

qualitativamente as relações entre as ações e eventos que fazem parte do modelo. A possibilidade de observar graficamente o que acontece com o modelo quando os valores das variáveis são modificados –análise de sensibilidade– permite compreender melhor o problema descobrindo quais as variáveis que influenciam as decisões. Este processo também auxilia a refinar os valores das variáveis envolvidas. Entretanto as RPQs permitem realizar abstrações e simplificações apropriadas do modelo, i.e. permitem obter as relações qualitativas do modelo a partir do seu nível quantitativo. A implementação destas facilidades será descrita posteriormente.

Pelo exposto anteriormente este trabalho propõe uma teoria fortemente fundamentada para auxiliar o usuário a compreender o modelo e as conclusões do sistema quando DIs são utilizados como formalismo de representação. Embora já existam trabalhos que tentam explicar as conclusões obtidas através de uma rede de Bayes, muito pouco tem sido feito para DIs, ou seja, quando também as decisões e função de utilidades são explicitamente modeladas. Desta forma este trabalho pode servir para futuras extensões ou para abordagens divergentes no processo de obter explicações quando diagramas de influências são empregados.

CAPÍTULO 7

UMA FERRAMENTA DE APOIO À DECISÃO

7.1 Introdução

Neste capítulo é apresentada a implementação de uma ferramenta gráfica de apoio à decisão baseada em diagramas de influências. Este aplicativo também apresenta o esquema de explicação proposto.

Por se tratar de uma interface gráfica, durante as etapas de análise, projeto e programação, foi decidido utilizar o paradigma orientado a objetos. O método de análise e projeto segundo Booch [Booch-91], o sistema ROSE [Rational-92], [White-93] (ferramenta CASE que suporta o paradigma orientado a objetos segundo Booch) e a linguagem "Smalltalk/V for Windows" [Digitalk-91] [Mittendorfer-92] foram utilizados no desenvolvimento deste trabalho.

O capítulo está organizado da seguinte forma: na seção 7.2 são descritas as características principais de Smalltalk. Na seção 7.3 são apresentadas as etapas de análise, projeto e implementação realizadas no desenvolvimento da ferramenta. A descrição desta encontra-se na seção 7.4, enquanto a seção 7.5 compreende as considerações finais.

7.2 A Linguagem Smalltalk

Nesta seção serão discutidas algumas das características principais da linguagem Smalltalk e as razões da sua escolha para a implementação desta ferramenta. Smalltalk foi a primeira linguagem puramente orientada a objetos. Desde seu início, nos anos 70, já foi concebida como um sistema de desenvolvimento de software e não só como uma linguagem de computação; ela já possuía ferramentas tais como editor de textos, depurador e "*Browser*". Tudo nela é visto como um objeto, até mesmo números inteiros e classes. Junto com a linguagem Simula é talvez a mais importante linguagem de Programação Orientada a Objetos –POO– pois seus conceitos influenciaram não só as subseqüentes linguagens de POO como também a forma de projetar as interfaces gráficas para o usuário, como por exemplo no Macintosh. Ela obteve uma ampla difusão (especialmente nos Estados Unidos) e serviu como exemplo para outras linguagens orientadas a objetos (C++, CLOS, etc.).

Os aplicativos desenvolvidos em Smalltalk são facilmente transportáveis de um sistema operacional para outro, incluindo tanto as interfaces gráficas do usuário bem como as

interfaces para o sistema operacional. Smalltalk está constituída em torno de dois simples conceitos: tudo é tratado como um objeto e estes se comunicam através do intercâmbio de mensagens.

Smalltalk oferece a vantagem de possuir um bom suporte para a construção de interfaces gráficas, como ilustrado na tabela 7.1, o que a torna apropriada para o desenvolvimento da aplicação requerida.

Características	C++	Smalltalk	C, Lisp, Pascal
Capacidade de adaptação para programadores tradicionais	alta	baixa	muito alta
Velocidade de execução	alta	média	alta
Exigências de memória	baixa	alta	baixa
Esforço para o aprendizado	médio	alto	baixo
Produtividade	média	muito alta	baixa
Reutilização de código	média	muito alta	ruim
Apoio às interfaces gráficas	ruim	boa	ruim

Tabela 7.1 Avaliação das características das linguagens de programação [Bourne-92]

7.3 Etapas do Desenvolvimento

A seguir serão descritas cada uma das etapas seguidas no desenvolvimento da ferramenta.

7.3.1 Análise

A utilização das técnicas de análise e projeto orientado a objetos facilitaram o processo de desenvolvimento da ferramenta. Na etapa de análise, depois de definir o problema e as funções do sistema, as classes principais foram determinadas e especificadas. Estes processos, assim como os seguintes, são repetidos de forma tal que os resultados de cada etapa são refinados à medida que o problema é mais bem compreendido. Depois de delinear as principais classes, as relações entre estas devem ser definidas com a sua correspondente cardinalidade. A seguir devem ser determinadas as operações a serem realizadas por cada classe, assim como os argumentos destas operações, caso neste ponto isso seja possível. Os atributos de cada classe devem estar especificados, assim como as relações de herança entre classes. Alguns dos resultados destes processos são mostrados nas figuras 7.1 e 7.2.

Figura 7.1 Diagrama de classes para a interface gráfica segundo a metodologia de Booch [Booch-91].

Figura 7.2 Diagrama de classes para o modelo de decisão utilizado.

7.3.2 Projeto

Segundo o método de Booch a arquitetura do sistema deve estar organizada em camadas hierárquicas, as quais devem ser divididas em partições acopladas fracamente, fornecendo cada partição um tipo de serviço. Estas partições são denominadas categorias de classes. O grau de exposição do conteúdo de cada categoria de classe para uma outra categoria de classe é denominado visibilidade. As categorias de classes junto com a sua visibilidade definem parte da arquitetura do sistema. A figura 7.3 mostra a arquitetura para a ferramenta desenvolvida e as categorias de classes especificadas são descritas a seguir:

Figura 7.3 Arquitetura da ferramenta desenvolvida segundo as categorias de classes.

ID_GUI: esta categoria constitui a ferramenta em si.

ID_Model: esta partição corresponde ao conhecimento e funções associadas ao domínio de aplicação, no caso a representação e avaliação de um problema como um diagrama de influência.

User_Interface: é a interface gráfica (como as diversas janelas, menus e ações realizadas através do "mouse" são coordenadas para apresentar na tela os diversos aspectos do modelo).

Persistent_Data: os dados persistentes, isto é, como os diversos problemas ou exemplos são armazenados em arquivos.

Data_Structures: estruturas de dados já existentes na linguagem ou implementadas para a aplicação (esta categoria é a única de acesso global, ela é visível para todas as outras categorias).

7.3.3 Implementação

A transição da análise e projeto para a etapa de implementação foi realizada de uma forma natural devido às características da linguagem utilizada. Uma decisão importante de implementação, a ser ressaltada, é a escolha de duas classes principais: InfluenceDiagram e InfluenceManager. Ambas correspondem às categorias de classe ID_Model e User_Interface da arquitetura da ferramenta mostrada na figura 7.3.

A primeira classe, *InfluenceDiagram*, define o conhecimento do domínio, ela representa um DI que possui nós de três diferentes tipos (nós probabilísticos, de decisão e de valor), além de guardar a informação referente às conexões entre estes e outros dados de relevância. As operações associadas à classe (denominadas "métodos" no contexto da linguagem Smalltalk) definem o comportamento do DI (ele pode se avaliar, eliminar ou adicionar nós, desenhar-se numa janela, etc.). O procedimento de avaliação do DI implementado foi o algoritmo de Shachter [Shachter-86].

A segunda classe a ser destacada é *InfluenceManager*, uma subclasse de *ViewManager*, já existente na linguagem Smalltalk. *InfluenceManager* representa a interface gráfica, ou seja, cada componente da aplicação constituído pelos objetos gráficos que aparecem interativos na tela do computador. Esta classe se comunica com o modelo (o DI representado por um objeto da classe *InfluenceDiagram*) de forma que cada vez que o modelo se modifica a interface gráfica modifica o conteúdo das respectivas janelas. É a classe *InfluenceManager* que responde aos movimentos do "mouse" e controla o conteúdo das janelas da aplicação com seus menus, entre outros.

Esta discriminação, mesmo em nível de implementação, entre o conhecimento do domínio (a classe *InfluenceDiagram*) e a interface (a classe *InfluenceManager*) deve-se a conceitos implícitos da linguagem. O conceito de separar o modelo da interface gráfica foi concebido durante a especificação e descrição de aplicativos em Smalltalk-80 [Goldberg-89] e denomina-se "Model-View-Controller". Neste caso o "View" corresponde à interface gráfica (a classe *ViewManager*) e o controlador corresponde ao sistema Smalltalk.

7.4 Descrição da Ferramenta

Em seguida serão descritas as funcionalidades da ferramenta implementada, bem como as facilidades de explicação.

7.4.1 Descrição Geral

A ferramenta permite representar o problema do usuário como um DI. De uma forma geral o diagrama pode ser editado, modificado e avaliado. A figura 7.4 ilustra a interface da ferramenta mostrando o DI do exemplo 6.1, página 56. A representação deste diagrama pode ser gerada interpretando um pequeno programa ou interativamente. À continuação é mostrada a listagem do programa e a sua descrição.

```
"Initialization of a DI"

"Definition of nodes"
N1:=ChanceNode new name: 'Petroleo' position:174@210
    values: #('yes' 'no').
N2:=DecisionNode new name:'Perfurar?' position:207@132
    meaning:'Realizar a perfuracao?' values: #('yes' 'no').
N3:=ChanceNode new name:'Resultado do Teste' position:72@132
    values: #('positivo' 'negativo').
N5:=DecisionNode new name:'Teste?' position:72@45
    values: #('Test' 'noTest').
N6:=ValueNode new name:'Ganho' position:319@132.
```

```

"Read of probability distributions and utility function
for the value node"
N1 distribution: (Distribution readFrom: 'infl\id2\distrbN1.txt').
N3 distribution: (Distribution readFrom: 'infl\id2\distrbN3.txt').
N6 utility: (Utility readFrom: 'infl\id2\utilN6.txt') .

"Initialization of the DI"
Influence:= InfluenceDiagram new initialize.

"Definition of the arcs"
Influence
    connect: N1 to: N3;
    connect: N1 to: N6;
    connect: N2 to: N6;
    connect: N5 to: N6;
    connect: N3 to: N2;
    connect: N5 to: N3.

"Open the DecisionTool and show the DI"
IManager:= InfluenceManager new openOn:Influence.

```

O programa define cada um dos nós que fazem parte do diagrama. A seguir a distribuição de probabilidades para os nós probabilísticos e a função de utilidades para o nó de valor são lidos de arquivos com formato ASCII. O diagrama é então inicializado e especificado através dos nós que o constituem e como estes estão relacionados. Deve ser observado que o método `connect: NA to: NB` não só indica que existe um arco do nó NA que incide em NB como também adiciona os nós NA e NB para o correspondente diagrama quando eles não existiam anteriormente. A última linha do programa chama a interface gráfica.

Figura 7.4 Interface da ferramenta mostrando um exemplo

Como pode ser observado na figura 7.4 a ferramenta possui botões (na lateral esquerda) que podem ser pressionados com o mouse para criar graficamente nós e arcos, ou modificar a disposição destes na tela.

Os itens dos três menus ilustrados esquematicamente na figura 7.5 definem as funcionalidades da ferramenta e serão descritos a seguir.

Nodes	Diagram	Explanation
New ...	New ...	What-if a Node ...
Open ...	Inspect	What-if ...
Remove ...	Barren Nodes	Why ...
Connect ...	Verify no Cycles	
	Redraw	
	Evaluate	

Figura 7.5 Itens dos principais menus da ferramenta.

A partir do menu Nodes, nós podem ser inseridos (New...) ou removidos (Remove...) do diagrama. Novas relações de precedência entre nós podem ser definidas (Connect...), ou um nó pode ser examinado (Open...) o qual também permite modificar os seus valores.

Do menu Diagram as seguintes opções podem ser escolhidas:

- New: um novo diagrama pode ser criado,
- Inspect: o DI pode ser examinado e seus valores modificados,
- Barren Nodes: os "Barren Nodes" (aqueles que não têm sucessor direto, exceto o nó de valor, e portanto são irrelevantes) podem ser eliminados,
- Verify no cycles: pode-se verificar que o diagrama não possui ciclos,
- Redraw: ele pode ser desenhado novamente na janela,
- Evaluate: o DI pode ser avaliado (obtendo-se desta forma a estratégia ótima).

Se o DI representado for avaliado, então a estratégia ótima (best policy) é apresentada na forma de regras, como as mostradas a seguir para o exemplo 3.4, no qual considerava-se a decisão de perfurar dois campos, A e B, à busca de petróleo, sendo possível realizar um teste prévio à perfuração:

```
TestB if b then DrillB. (EV=27.9)
TestB if -b then noDrill. (EV=-6.0)
```

Neste caso aconselha-se realizar o teste no campo B. Se este tiver um resultado positivo recomenda-se perfurar o campo B com um valor esperado de 27.9 (primeira regra). Se o resultado do teste fosse negativo recomenda-se não perfurar (com uma perda esperada de -6).

7.4.2 Facilidades de Explicação

A seguir serão descritas as facilidades de explicação implementadas. Estas podem ser ativadas através dos três itens do menu Explanation, representados na figura 7.5.

Escolhendo-se a opção What-if a Node... do menu Explanation pode-se eleger uma variável para realizar a análise de sensibilidade no modelo. A figura 7.6 ilustra o resultado de uma interação depois de ter escolhido o nó probabilístico Resultado do Teste. Esta análise de sensibilidade foi implementada para variáveis com valores binários. Além de eleger a variável, também é possível controlar quais as estratégias a serem consideradas.

Figura 7.6 Análise de sensibilidade para uma variável do modelo.

O seguinte item do menu (What-if ...) abre uma janela de diálogo através da qual é possível modificar os valores numéricos das distribuições de cada variável. O modelo com estes novos valores pode ser avaliado, obtendo-se assim a nova estratégia ótima que pode ser comparada com a obtida anteriormente.

Estas duas facilidades acima descritas permitem descobrir quais as variáveis que maior influência tem nas conclusões atingidas.

Através da terceira opção do menu, (Why ...), o sistema constrói a correspondente rede qualitativa do modelo como mostrado na figura 7.7. A partir desta rede pode-se obter uma descrição do relacionamento entre duas variáveis (analisando a influência qualitativa que liga os correspondentes nós). Por exemplo, a seguir são transcritas as explicações obtidas no diagrama da figura 7.7 para cada um dos pares de nós escolhidos:

Figura 7.7 Rede probabilística qualitativa de um exemplo.

Tratamento medico? - Custos:

The decision <Tratamento médico?> influences positively the value of <Beneficios>. (For higher values of <Tratamento médico?> higher values for the probability of <Beneficios> are expected).

Resultado Teste - Custos:

The variable <Resultado Teste> has no influence on <Custos>.

Grau da doença - Valor:

Nothing can be concluded about the influence of <Grau da doença> on the variable <Valor>.

Além disso a inferência pode ser realizada na rede qualitativa para transformá-la de forma que as relações de interesse sejam diretas. Por exemplo, a rede da figura 7.7 pode ser reduzida à da figura 7.8 onde as relações qualitativas de Tratamento médico e Resultado Teste sobre o nó de valor estão explícitas. Este processo é realizado especificando uma variável alvo, no caso o nó de valor, e o conjunto de variáveis condicionais, Tratamento médico e Resultado Teste. Desta forma é possível gerar o seguinte tipo de afirmação:

<Tratamento médico> is a nondecreasing function of <Resultado Teste>. (For higher values of <Resultado Teste> higher values of <Tratamento médico> should be selected).

Figura 7.8 Rede probabilística qualitativa reduzida.

7.5 Considerações finais

Neste capítulo foi descrita a implementação de uma ferramenta de apoio à decisão que utiliza diagramas de influências como formalismo de representação. A ferramenta descrita foi implementada segundo as diretrizes da análise e programação orientada a objetos e em nível de protótipo. Muitas outras facilidades poderiam ainda ser adicionadas (facilidades de entrada/saída, comunicação com tabelas de cálculos do ambiente utilizado, compilação do aplicativo, etc.), mas elas não condizem com o objetivo deste trabalho.

O objetivo deste trabalho é explorar e estender possíveis mecanismos de explicação para esclarecer conclusões atingidas através de DIs. Com tal finalidade foram implementadas, e descritas neste capítulo, três facilidades de explicação baseadas na análise de sensibilidade e redes probabilísticas qualitativas. A visualização do modelo quando as variáveis mudam seus valores, assim como dos relacionamentos em nível qualitativo, auxilia na compreensão das variáveis envolvidas no problema. Em particular a apresentação gráfica dos resultados, bem como a facilidade de poder modificar o modelo e observar as mudanças, incentiva à experimentação e percepção do problema.

CAPÍTULO 8

CONCLUSÕES

8.1 Considerações Gerais

Neste trabalho foi discutida a necessidade de representação e manipulação de incertezas por sistemas baseados em conhecimento, e como isto pode ser realizado utilizando sistemas normativos –sistemas baseados nos princípios da teoria da decisão.

Pelo exposto ficou demonstrado que diagramas de influências e redes de Bayes fornecem uma representação do conhecimento que permite expressar conhecimento incerto, crenças e preferências qualitativa e quantitativamente de uma forma flexível. Entre as vantagens mais destacadas destes formalismos as seguintes podem ser mencionadas:

- O conhecimento pode ser representado em qualquer direção, não necessariamente na direção causal. Os algoritmos probabilísticos utilizam estas representações (redes de crenças) para raciocinar na direção requerida (causal, de diagnóstico, ou intercausal). Ao contrário das representações baseadas em regras, não existe uma ligação entre a forma de codificar o conhecimento e a forma de raciocínio.
- Os DIs compreendem um modelo explícito para as decisões e as utilidades dos possíveis resultados. Esta modularidade mantém separados os juízos sobre as crenças (probabilidades) e os juízos sobre as preferências (utilidades).

Uma das questões simples embora de difícil resposta neste paradigma é: "O que acontece se as probabilidades não são conhecidas?". Obter as probabilidades para problemas reais pode ser uma tarefa trabalhosa e um limitante que torna esta abordagem adequada só para certo tipo de problemas. Às vezes é possível obter dados experimentais a partir dos quais as probabilidades podem ser calculadas. Em muitos outros casos estas devem ser obtidas a partir da opinião do especialista. Na análise da decisão existem diversas técnicas para facilitar esta tarefa, embora na nossa opinião seja esta uma das principais limitações a serem superada. Com esta finalidade está sendo desenvolvida pesquisa em aprendizado em redes [Spiegelhalter-92] [Gebhardt-95], aprendizado de parâmetros e monitoração ou crítica do modelo (os parâmetros podem ser ajustados quando comparados com dados obtidos) [Spiegelhalter-93]. Estas disciplinas utilizam extensamente métodos estatísticos e assemelham os esforços realizados em aprendizado de máquina com o objetivo de auxiliar o processo de aquisição do conhecimento em SEs.

A síntese de idéias entre a teoria da decisão e IA já tem produzido algumas contribuições importantes, tanto para a teoria do raciocínio e representação do conhecimento quanto no desenvolvimento prático de sistemas de diagnóstico. Entretanto, uma extensiva pesquisa ainda deve ser realizada no desenvolvimento de ferramentas para facilitar a aquisição do conhecimento para este novo paradigma, na busca de algoritmos de inferência mais eficientes e na exploração de técnicas para a explicação deste tipo de raciocínio, entre outros.

8.2 Conclusões

Dentro do contexto discutido anteriormente, neste trabalho foram pesquisadas técnicas de explicação para sistemas baseados em DIs. Duas facilidades de explicação, baseadas em diferentes teorias: análise de sensibilidade e redes probabilísticas qualitativas, foram propostas e implementadas numa ferramenta gráfica. Estes mecanismos exploram e evidenciam a estrutura qualitativa do modelo: para as pessoas é geralmente mais fácil entender as propriedades e relações do problema e a sua resolução do que resultados quantitativos. A ferramenta desenvolvida também permite compreender mais facilmente o modelo já que as variáveis mais significativas podem ser destacadas. Em particular a apresentação gráfica dos resultados bem como a facilidade de poder modificar o modelo e observar as mudanças incentiva à experimentação e uma melhor percepção do problema.

A primeira facilidade de explicação implementada permite avaliar a importância de cada variável do modelo e a eliminação daquelas irrelevantes. Isto se baseia na observação de que nenhum modelo é completo: geralmente é necessária uma simplificação do conhecimento, das crenças e preferências do especialista, os quais já representam uma simplificação da realidade. Desta forma, existe um compromisso entre simplicidade para facilitar a engenharia do conhecimento e inferência, e a completeza necessária para assegurar uma precisão máxima. Durante a sua construção o modelo pode ser simplificado enquanto este compromisso entre precisão e simplicidade é explorado. Na área da análise de decisão diversos métodos de análise de sensibilidade são utilizados para examinar a importância de cada parâmetro do modelo. Isto é realizado a fim de responder à pergunta: "Pode esta variável afetar a utilidade esperada de forma a mudar a decisão recomendada?". Sob esta perspectiva obtém-se uma medida de relevância que permite podar aqueles parâmetros que são não significativos e fixar-se na compreensão das principais variáveis.

O segundo mecanismo de explicação, que faz uso das redes probabilísticas qualitativas, sofre das limitações impostas por este tipo de representação, tais como:

- somente associações monótonas podem ser exprimíveis,
- modelos qualitativos não especificam completamente a distribuição conjunta das variáveis, por este motivo pode existir perda de informação durante o processo de inferência, como ainda será discutido,
- algumas situações não podem ser resolvidas e ficam indeterminadas, como por exemplo influências paralelas de sinal contrário quando combinadas resultam em uma influência indeterminada do tipo $N^?$.

Ainda com essas limitações esta facilidade de explicação permite visualizar os relacionamentos em nível qualitativo, auxiliando assim na compreensão das ações e

eventos que fazem parte do problema. Este modelo de decisão genérico é inadequado para justificar uma particular escolha, porém o modelo é forte o suficiente para fornecer restrições úteis na estratégia ótima como uma função da informação disponível.

A ferramenta descrita foi implementada em nível de protótipo e diversas melhoras podem ser realizadas tanto em nível geral bem como aquelas destinadas ao usuário entender melhor o problema representado. Por exemplo, a análise de sensibilidade implementada pode ser estendida para variáveis com mais de dois valores, ou outras medidas de sensibilidade podem ser utilizadas, como será discutido a seguir.

8.3 Sugestões para futuros trabalhos

São diversas as linhas de pesquisa que dão continuidade e complementam este trabalho. Algumas delas, porém, enfrentam problemas de difícil resolução, como por exemplo a perda de informação que pode ocorrer no processo de inferência em RPQs, o qual constitui uma das questões em aberto no desenvolvimento destas redes. Dependendo da ordem na qual as transformações são realizadas no processo de inferência, a conclusão obtida pode não ser a conclusão mais restrita possível de ser obtida [Wellman-90]. Ou seja, pode-se obter uma conclusão mais geral (devido à perda de informação pela sequência de operações escolhidas) que leva a restrições (e conseqüentemente explicações) mais fracas sobre a rede analisada. A solução deste problema está relacionada com a construção de novos algoritmos de inferência em RPQs que minimizem esta ambigüidade.

Outras linhas de pesquisa que dão continuidade a este trabalho são:

- A análise de sensibilidade em uma via fornece uma visão limitada, pois examina as mudanças numa única variável, as outras variáveis são mantidas constantes. Às vezes é interessante explorar a melhor estratégia para uma combinação de fatores. Tais questões podem ser abordadas realizando análise de sensibilidade em duas vias – variando os valores de duas variáveis independentemente sobre todos os seus possíveis valores e determinando a melhor estratégia para cada uma destas combinações [Pauker-87].
- Outras medidas de sensibilidade podem ser exploradas, como por exemplo: o valor esperado da informação, do cómputo e de incluir incertezas [Henrion-91].
- Outras abordagens para gerar explicações da estratégia recomendada podem ser estudadas, como por exemplo através dos pesos de evidência.
- RPQs também podem ser usadas para examinar estratégias e separar aquelas que embora sintaticamente corretas são incoerentes ou não têm sentido prático [Wellman-88] (um exemplo de uma estratégia sem sentido é realizar um teste de diagnóstico e a seguir escolher o tratamento médico sem considerar o resultado do teste).
- A utilização de cor como uma facilidade de explicação pode ser analisada: nós podem ser coloridos ou sombreados segundo a sua influência na variável de valor.

- Testes podem ser conduzidos para tentar comparar o grau de compreensão do usuário, quando apresentado com diversos exemplos, utilizando a ferramenta com e sem as facilidades de explicação.
- Uma comparação sistemática com outros sistemas, como por exemplo alguns dos mencionados no capítulo 5, poderia ser realizada.

LISTA DE ABREVIACÕES

BC	Base de Conhecimento
DI	Diagrama de Influência
MI	Motor de Inferências
IA	Inteligência Artificial
POO	Programação Orientada a Objetos
RB	Redes de Bayes
SAD	Sistemas de Apoio a Decisão
SBC	Sistema Baseado em Conhecimento
SE	Sistemas Especialistas
SGI	Sistemas de Gerenciamento da Informação
SPOO	Sistemas de Programação Orientada a Objetos
TD	Teoria da Decisão
UE	Utilidade Esperada
VME	Valor Monetário Esperado

REFERÊNCIAS BIBLIOGRÁFICAS

- [Abramson-91] Abramson, B. ARCO1: An Application of Belief Networks to the Oil Market. *Proceedings of the Seventh Conference on Uncertainty in Artificial Intelligence*, 1-8, 1991.
- [Abramson-93] Abramson, B. & Finizza, A.J. Using Belief Networks to Forecast Oil Prices, *International J. Forecasting*, 7, 3, 299-316, 1993.
- [ACM-95] *Communications of the ACM*. Vol. 38, N° 3, 25-57, March 1995.
- [Almond-94] Almond, R.G., Bradshaw J.M & Madigan, D. Reuse and Sharing of Graphical Belief Network Components, in Cheeseman, P. & Oldford, W. (Eds.) *Selecting Models from Data: Artificial Intelligence and Statistics IV*, Springer-Verlag, 113-122, 1994.
- [Almond-95] Almond, R.G. *Graphical Belief Modelling*. Chapman and Hall, 1995.
- [Almond-96] Almond, R.G. *Software for Manipulating Belief Networks*, in <http://bayes.stat.washington.edu/almond/belief.html>.
- [Almond-96a] Almond, R.G. *Software for Learning Belief Networks from Data*, in <http://bayes.stat.washington.edu/almond/belfit.html>.
- [Andersen-89] Andersen, S.K., Olesen, K.J., Jensen, F.V. & Jensen, F. HUGIN - A Shell for Building Bayesian Belief Universes in Expert Systems. *Proc. 11th International Joint Conference on AI*, Detroit, 1080-1085, 1989.
- [Andreassen-87] Andreassen, S., Woldbye, M., Falck, B. & Andersen, S.K. MUNIN: A Causal Probabilistic Network for Interpretation of Electromyographic Findings. In *Proceedings of the Tenth National Conference on Artificial Intelligence*, American Association for AI: Menlo Park, California, 121-123, 1987.
- [Baker-85] Baker, K.R. & Kropp, D.H. *Management Science, An Introduction to the Use of Decision Models*. John Wiley & Sons: New York, 1985.
- [Bonissone-90] Bonissone, P.P., Cyrluk, D.A., Goodwin, J.W. & Stillman, J. Uncertainty and Incompleteness: Breaking the Symmetry of Defeasible Reasoning. *Uncertainty in Artificial Intelligence 5*, Henrion, M., Shachter, R.D., Kannal, L.N. & Lemmer, J.F. (Eds), Elsevier Science Pub.: North Holland, 1990
- [Booch-91] Booch, G. *Object Oriented Design with Applications*. Benjamin/Cummings: Redwood City, 1991.

- [Boose-89] Boose, J.H. A Knowledge Acquisition Techniques and Tools. *Knowledge Acquisition*, 1, 3-37, 1989.
- [Bourne-92] Bourne, J.R. *Building Engineering Systems using Smalltalk-80*, Irwin and Aksen Associates: Illinois, USA, 1992.
- [Breese-94] Breese, J.S., Goldman, R.P. & Wellman, M.P. Introduction to the Special Section on Knowledge-Based Construction of Probabilistic and Decision Models. *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 24, N° 11, 1577-1579, November 1994.
- [Buchanam-84] Buchanam B.G. & Shortliffe, E.H. *Rule-Based Expert Systems*. Addison-Wesley: Reading, Massachusetts, 1984.
- [Buede-94] Buede D. Aiding Insight II (Survey). *OR/MS Today*, 62-74, June 1994.
- [Castiñeira-90] Castiñeira, M.I. *Aprendizado de Máquina por Exemplos Usando Árvores de Decisão*. Dissertação de Mestrado, ICMSC-USP, São Carlos, SP, 1990.
- [Castiñeira-90a] Castiñeira, M.I. & Monard, M.C. *Análise de um Método de Poda para Árvores Indutivas de Decisão*, Anais do VII Simpósio Brasileiro de Inteligência Artificial, Paraíba, 17-27, Novembro de 1990.
- [Castiñeira-91] Castiñeira, M.I., Monard, M.C. & Nicoletti, M.C. *Aprendizado de Máquina: Descrição e Implementação de um Algoritmo Geral para a Construção de Árvores de Decisão*, Notas do ICMSC-USP, N° 98, ICMSC-USP, São Carlos, SP, 1991.
- [Castiñeira-96] Castiñeira, M.I. & Monard, M.C. *Mecanismos de Explicação em Diagramas de Influências*. XIII Simpósio Brasileiro de Inteligência Artificial, Curitiba, 23-25 de Outubro, 1996. Trabalho aceito.
- [Castiñeira-96a] Castiñeira, M.I. & Monard, M.C. *Descrição do Projeto e Implementação do Sistema IDEX: uma Ferramenta para a Representação e Avaliação de Diagramas de Influência*. Notas do ICMSC-USP, Série Computação, a ser publicado.
- [Charniak-91] Charniak, E. Bayesian Networks without Tears. *AI Magazine*, 12(4), 50-63, Winter 1991.
- [Clancey-84] Clancey, W.J., Hasling, D.W. & Rennels, G. Strategic Explanation for a Diagnostic Consultation System. *International Journal of Man-Machine Studies*, 20, 3-19, 1984.
- [Clark-90] Clark, D.A. Numerical and Symbolic Approaches to Uncertainty Management in AI. *Artificial Intelligence Review*, 4, 109-146, 1990.
- [Cooper-84] Cooper G.F. NESTOR: a a Computer based Medical Diagnostic Aid that Integrates Causal and Probabilistic Knowledge. Ph.D. Thesis, Computer Science Department, Stanford University, November 1984: Rep. No. STAN-CS-84-48.

- [Cooper-88] Cooper G.F. A Method for using Belief Networks as Influence Diagrams. in *Proceedings of the 4th Workshop on Uncertainty in Artificial Intelligence*, Univ. of Minnesota: Minneapolis, M.N. August, 1988.
- [Cooper-90] Cooper, G.F. Computational Complexity of Probabilistic Inference using Bayesian Belief Networks. *Artificial Intelligence*, N° 42, 393-405, 1990.
- [Cooper-91] Cooper, G.F. & Herskovits, E. A Bayesian Method for Constructing Bayesian Belief Networks from Databases, in Smets, D'Ambrosio & Bonissone (Eds.), *Proceedings of the Seventh Conference on Uncertainty in AI*, University of California: Los Angeles, CA, 86-94, July 1991.
- [Cowell-93] Cowell R.G., Dawid A.P. & Spiegelhalter D.J. Sequential Model Criticism in Probabilistic Expert Systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 15, N° 3, 209-219, March 1993.
- [D'Ambrosio-91] D'Ambrosio, B. Local Expression Languages for Probabilistic Dependence, in *Proceedings of the 7th Conference on Uncertainty in Artificial Intelligence*. Morgan Kaufmann: Calif., 95-102, 1991.
- [Dagum-93] Dagum, P. & Luby, M. Approximately Probabilistic Reasoning in Bayesian Belief Networks is NP-hard. *Artificial Intelligence*, 141-153, 1993.
- [Dempster-67] Dempster, A.P. Upper and Lower Probabilities Induced by a Multivalued Mapping. *Ann. Math. Statistics*, 38, 325-339, 1967.
- [Digitalk-91] Digitalk Inc. *Smalltalk/V Windows: Object-Oriented Programming System (OOPS), Tutorial and Programming Handbook*. Los Angeles, 1991.
- [Dubois-88] Dubois, D. & Prade, H. *Possibility Theory: an Approach to Computerized Processing of Uncertainty*. Plenum Press: New York, 1988.
- [Duda-76] Duda, R.O., Hart, P.E. & Nilsson, N.J. Subjective Bayesian Methods for Rule-based Inference Systems. *Proc., Natl. Comp. Conf. (AFIPS)* 45, 1075-1082, 1976.
- [Elsaesser-89] Elsaesser, C. Explanation of Probabilistic Inference. *Uncertainty in Artificial Intelligence 3*, Kanal, L.N., Levitt, T.S. & Lemmer, J.F. (Eds.). Elsevier Science Publishers B.V.: North Holland, 387-400, 1989.
- [Elsaesser- 90] Elsaesser, C. & Henrion, M. Verbal Expressions for Probability Updates: How Much more Probable is "Much More Probable"? *Uncertainty in Artificial Intelligence 5*, Henrion, M., Shachter, R.D.,

Kanal, L.N. & Lemmer, J.F. (Eds.). Elsevier Science Publishers B.V.: North Holland, 319-328, 1990.

- [Gebhardt-95] Gebhardt, J. & Kruse, R. Reasoning and Learning in Probabilistic and Possibilistic Networks: An Overview. *Lectures Notes in Artificial Intelligence*, 912, Springer-Verlag: Berlin, 1995.
- [Geiger-91] Geiger, D. & Heckerman, D. Advances in Probabilistic Reasoning, in Smets, D'Ambrosio & Bonissone (Eds.), *Proceedings of the Seventh Conference on Uncertainty in AI*, University of California: Los Angeles, CA, 118-126, July 1991.
- [Glymour-87] Glymour, C., Scheines, R., Spirtes, P. & Kelly, K. *Discovering Causal Structures*, Academic Press, San Diego, CA, 1987.
- [Goldberg-89] Goldberg, A. & Robson, D. *SMALLTALK-80: The Language and its Implementation*, Addison Wesley: Reading, 1989.
- [Good-85] Good, I.J. Weight of Evidence: A Brief Survey, in *Bayesian Statistics 2*. Bernardo, DeGroot, Lindley & Smith (Eds), Elsevier Science Publishers: North-Holland, 249-270, 1985.
- [Harnett-82] Harnett, D.L. *Statistical Methods*. Addison-Wesley: U.S.A., 1982.
- [Hayes-Roth-92] Hayes-Roth, F. Expert Systems. *Encyclopedia of Artificial Intelligence*, S.C. Shapiro (Ed.), 2nd. Ed., Wiley & Sons: New York, 477-489, 1992.
- [Heckerman-88] Heckerman D.E. & Horvitz E.J. The Myth of Modularity in Rule-based Systems for Reasoning with Uncertainty. *Uncertainty in Artificial Intelligence 2*, Lemer & Kanal, (Eds.) Elsevier Sc.P.: North-Holland, 1988.
- [Heckerman-90] Heckerman, D.E. A Tractable Algorithm for Diagnosing Multiples Diseases, in *Uncertainty in Artificial Intelligence 5*, Henrion, M., Shachter, R.D., Kanal, L.N. & Lemmer, J.F. (Eds.). Elsevier Science Publishers B.V.: North Holland, 1990.
- [Heckerman-91] Heckerman, D. *Probabilistic Similarity Networks*, MIT Press: Cambridge, MA, 1991.
- [Heckerman-92] Heckerman D.E., Horvitz E.J. & Nathwani, B.N. Toward Normative Expert Systems: Part I. The Pathfinder Project. *Methods of Information in Medicine*, 31, 90-105, 1992.
- [Heckerman-92a] Heckerman D.E. & Nathwani, B.N. Toward Normative Expert Systems: Part II. Probability-based Representations for Efficient Knowledge Acquisition and Inference. *Methods of Information in Medicine*, 31, 106-116. 1992.
- [Heckerman-95] Heckerman, D. & Wellman, M.P. Bayesian Networks. *Communications of the ACM*, Vol.38, N°3, 27-30, March 1995.

- [Henrion-91] Henrion, M., Breese, J.S. & Horwitz, E.J. Decision Analysis and Expert Systems. *AI Magazine*, 12(4), 64-91, Winter 1991.
- [Henrion-91a] Henrion, M. & Druzdzel, M.J. Qualitative Propagation and Scenario-based Schemes for Explaining Probabilistic Reasoning. *Uncertainty in Artificial Intelligence 6*, Bonissone, P.P., Henrion, M., Kanal, L.N. & Lemmer, J.F. (Eds.). Elsevier Science Publishers B.V.: North Holland, 17-32, 1991.
- [Holtzman-89] Holtzman, S. *Intelligent Decision Systems*, Addison-Wesley: Massachusetts, 1989.
- [Horvitz-88] Horvitz E.J., Breese J.S. & Henrion M. Decision Theory in Expert Systems and Artificial Intelligence. *International Journal of Approximate Reasoning*, 2, 247-302, 1988.
- [Howard-84] Howard, R.A. & Matheson, J.E. Influence Diagrams, in Howard, R.A. & Matheson, J.E (Eds.), *Readings on the Principles and Applications of Decision Analysis*. Vol II. Strategic Decisions Group: Menlo Park, CA, 719-762, 1984.
- [Hsia-89] Hsia, Y.T. & Shenoy, P.P. An Evidential Language for Expert Systems, in Ras, Z.W. (Ed.). *Methodologies for Intelligent Systems 4*, North-Holland, 9-16, 1989.
- [IEEE-93] *IEEE Transactions on Pattern Analysis and Machine Intelligence (Special Section on Probabilistic Reasoning)*, Vol. 15, N° 3, March 1993.
- [Jensen-89] Jensen, F. *Bayesian Updating in Recursive Graphical Models by Local Computations*, Technical Report, R-89-15, Dep. of Mathematics and Computer Science, University of Aalborg, 1989.
- [Jimison-90] Jimison, H.B. Generating Explanations of Decision Models Based on an Augmented Representation of Uncertainty, in *Uncertainty in Artificial Intelligence 4*, Shachter, Levitt, Lemer & Kanal (Eds.). Elsevier Sc.P.: North-Holland, 351-365, 1990.
- [Jimison-92] Jimison, H.B., Fagan, L.M., Shachter, R.D. & Shortliffe, E.H. Patient-Specific Explanation in Models of Chronic Disease. *Artificial Intelligence in Medicine*, 4, 191-205, 1992.
- [Keeney-76] Keeney, R. & Raiffa, H. *Decisions with Multiple Objectives: Preferences and Values Trade-Offs*. Wiley: New York, 1976.
- [Keshavan-93] Keshavan H.R., Barnett J., Geiger D. & Verma T. Introduction to the Special Section on Probabilistic Reasoning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 15, N° 3, 193-195, March 1993.
- [Kim-92] Kim, S.H. *Statistics and Decisions*. Van Nostrand Reinhold: New York, 1992.

- [Klein-94] Klein, D.A. & Shortliffe, E.H. A Framework for Explaining Decision-Theoretic Advice. *Artificial Intelligence*, 67, 201-243, 1994.
- [Langlotz-83] Langlotz, C.P. & Shortliffe, E.H. Adapting a Consultation System to Critique User Plans. *International Journal of Man-Machine Studies*, 19(5), 479-496, 1983.
- [Langlotz-88] Langlotz, C.P., Shortliffe, E.H. & Fagan, L.M. A Methodology for Generating Computer-based Explanations of Decision-theoretic Advice. *Medical Decision Making*, 8, 290-303, 1988.
- [Langlotz-89] Langlotz, C.P. & Shortliffe, E.H. *The Critiquing Approach to Automated Advice and Explanation: Rationale and Examples*. Knowledge Systems Laboratory, Report KSL-89-46, Medical Computer Science, Stanford University, 1989.
- [Lauritzen-88] Lauritzen, S.L. & Spiegelhalter D.J. Local Computations with Probabilities on Graphical Structures and their Application to Expert Systems. *J.R.Statist.Soc. (B)*, 50, N° 2, 157-224, 1988.
- [Lauritzen-89] Lauritzen, S. L. & Wermuth, N. Graphical Models for Associations Between Variables, some of which are Qualitative and some Quantitative. *Annals of Statistics*, (17), 31-57, 1989.
- [Levitt-90] Levitt, T.S., Agosta, J.M. & Binford, T.O. Model-based Influence Diagrams for Machine Vision, in *Uncertainty in Artificial Intelligence 5*, Henrion, M., Shachter, R.D., Kannal, L.N. & Lemmer, J.F. (Eds), Elsevier Science Pub.: North Holland, 371-388, 1990.
- [Lowrance-86] Lowrance, J.D., Garvey, T.D., & Strat, T.M. A Framework for Evidential-reasoning Systems. *Proc., 5th Natl. Conf. on AI (AAAI-86)*, Philadelphia, 896-901, 1986.
- [Madigan-95] Madigan, D., Mosurski, K. & Almond, R.G. *Explanation in Belief Networks*. 1995. Artigo disponivel por anonymous ftp em: <ftp://ftp.stat.washington.edu/pub/tech.reports/explanation.ps>
- [Matzkevich-95] Matzkevich, I. & Abramson, B. Decision Analytic Networks in Artificial Intelligence. *Management Science*, Vol. 41, N° 1, January 1995.
- [Microsoft-96] Microsoft Decision Theory Group, *Software for Bayesian Nets*, <http://www.research.microsoft.com/dtg/msbn>.
- [Miller-76] Miller, A.C., Merkhofer, M.W., Howard, R.A., Matheson, J.E. & Rice, T.R. *Development of Automated Aids for Decision Analysis*, Stanford Research Institute: Menlo Park, Calif, 1976.
- [Miller-82] Miller, P.L. ATTENDING: a System which critiques an anesthetic management Plan. *Proceedings of the First AMIA Congress on Medical Informatics*, American Medical Informatics Association.

- [Miller-83] Miller, P.L. ATTENDING: Critiquing a Physician's Management Plan. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 5, 449-461, 1983.
- [Miller-86] Miller, P.L. Expert Critiquing Systems: Practice-based Medical Consultation by Computer, Computers and Medicine Series, Springer, New York, 1986.
- [Mittendorfer-92] Mittendorfer, J. *Objektorientierte Programmierung mit Smalltalk/V for Windows*. Addison-Wesley: Bonn, 1992.
- [Moore-89] Moore, J.D. & Swartout, W.R. A reactive Approach to Explanation. *Proceedings IJCAI-89*, Detroit, MI, 1989.
- [Ndilikiliksha-94] Ndilikiliksha, P.C., Potential Influence Diagrams. *International Journal of Approximate Reasoning*, Vol. 10, 155(3), 251-285, April 1994.
- [Neapolitan-90] Neapolitan, R.E. *Probabilistic Reasoning in Expert Systems: Theory and Algorithms*. John Wiley & Son: New York, 1990.
- [Nicoletti-92] Nicoletti, M.C., Castiñeira, M.I. & Monard, M.C. *Descrição e Implementação dos Mecanismos de Janela e Poda em Árvores de Decisão no Aprendizado Indutivo de Máquina*. Notas do ICMSC-USP, Nº 110, ICMSC-USP, São Carlos, SP, 1992.
- [Nicoletti-94] Nicoletti, M.C., *Ampliando os Limites do Aprendizado Indutivo de Máquina através das Abordagens Construtiva e Relacional*. Tese de Doutorado, IFSC-USP, São Carlos, SP, 1994.
- [Oliver-90] Oliver, R.M. & Smith J.Q. *Influence Diagrams, Belief Nets and Decision Analysis*. J. Wiley & Sons: Great Britain, 1990.
- [Olmsted-83] Olmsted, S.M. *On Representing and Solving Decision Problems*. Ph.D. thesis, EES Department, Stanford University: Standorf, Calif, 1983.
- [Paris-88] Paris, C.L., Wick, M.R., & Thompson, W.B. The Line of Reasoning Versus the Line of Explanation. *Proceedings AAAI Workshop on Explanation*, 4-7, 1988.
- [Pauker-87] Pauker, S.G. & Kassirer, J.P. Medical Progress: Decision Analysis. *New England Journal of Medicine*, 316, 250-258, 1987.
- [Pearl-86] Pearl, J. Fusion, Propagation, and Structuring in Belief Networks. *Artificial Intelligence*, 29, 241-288, 1986.
- [Pearl-88] Pearl, J. *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann: San Mateo, Calif., 1988.

- [Pearl-91] Pearl, J. & Verma, T. A Theory of Inferred Causation , in *Principles of Knowledge Representation and Reasoning: Proceedings of the Second International Conference*, 441-452, 1991.
- [Polya-54] Polya, G. *Patterns of Plausible Inference*. Princeton University Press: Princeton, 1954.
- [Raiffa-68] Raiffa, H. *Decisions Analysis*. Addison Wesley: Reading, Mass., 1968.
- [Ramoni-95] Ramoni, M., Riva, A., Stefanelli, M. & Patel, V. Medical Decision Making using Ignorant Influence Diagrams. *Lectures Notes in Artificial Intelligence*, N° 934, 139-150, 1995.
- [Rational-92] *Rational Rose for Windows User's Manual*. Rational: California, 1992.
- [Reggia-85] Reggia, J.A. & Perricone, B.T. Answer Justification in Medical Decision Support Systems Based on Bayesian Classification. *Comput. Biol. Med.* Vol. 15, N° 4, 161-167, 1985.
- [Rennels-86] Rennels, G.D. *A Computational Model of Reasoning from the Clinical Literature*. Ph.D.Thesis, Program in Medical Information Science, Stanford University, June, 1986.
- [Russel-90] Russel, S., Srinivas, S. & Agogino, A. Automated Construction of Sparse Bayesian Networks, in Henrion, M., Shachter, R.D., Kanal, L.N. & Lemmer, J.F. (Eds.) *Uncertainty in Artificial Intelligence 5*, Elsevier Science Publishers B.V.: North Holland, 295-308, 1990.
- [Ryan-88] Ryan, J.P. & Bridges, S. Constructing Explanations from conceptual Graphs, in: *Proceedings Third Annual Workshop on Conceptual Graphs*, 4-12, 1988.
- [Saffiotti-91] Saffiotti, A. & Umkehrer, E. *PULCINELLA: User's Manual*, Tech.Rep. TR/IRIDIA/91-5, Univ. Libre de Bruxelles, Belgium, 1991.
- [Saffiotti-94] Saffiotti, A. & Umkehrer, E. Inference Driven Constructions of Valuation Systems from First-order Clauses. *IEEE Trans. on Systems Man and Cibernetics*, 24, 1611-1624, 1994.
- [Shachter-86] Shachter, R.D. Evaluating Influence Diagrams. *Operations Research*, 34, 871-882, 1986.
- [Shachter-87] Shachter, R.D. & Heckerman D.E. Thinking Backward for Knowledge Acquisition. *The AI Magazine*, 55-61, Fall 1987.
- [Shachter-88] Shachter, R.D. Probabilistic Inference and Influence Diagrams. *Operations Research*, 36 (4), 589-604, 1988.
- [Shachter-90] Shachter, R.D. An Ordered Examination of Influence Diagrams. *Networks*, Vol. 20, 535-563, 1990.

- [Shafer-76] Shafer, G. *A Mathematical Theory of Evidence*. Princeton University Press: Princeton, 1976.
- [Sheridan-91] Sheridan, F.K.J. A Survey of Techniques for Inference under Uncertainty. *Artificial Intelligence Review*, 5, 89-119, 1991.
- [Shortliffe-75] Shortliffe, E.H. & Buchanan, B.G. A Model of Inexact Reasoning in Medicine. *Mathematical Biosciences*, 23, 351-379, 1975.
- [Shortliffe-76] Shortliffe, E.H. *Computer-based Medical Consultation: MYCIN*. Elsevier: New York, 1976.
- [Shwe-91] Shwe, M., Middleton, B., Heckerman, D., Henrion, M., Horvitz, E. & Lehmann, H. Probabilistic Diagnosis Using a Reformulation of the Internist-1/QMR Knowledge Base: I. The Probabilistic Model and Inference Algorithms. *Methods of Information in Medicine*, 30, 241-255, 1991.
- [Smith-88] Smith, J.Q. *Decision Analysis, A Bayesian Approach*. Chapman and Hall: London, 1988.
- [Smith-89a] Smith, J.Q. Influence Diagrams for Statistical Modeling. *The Annals of Statistics*, (17) 2, 654-672, 1989.
- [Smith-89b] Smith, J.Q. Influence Diagrams for Bayesian Decision Analysis. *European Journal of Operations Research*. (40) 363-376, 1989.
- [Smith-93] Smith, J.E., Holtzman, S. & Matheson, J.E. Structuring Conditional Relationships in Influence Diagrams. *Operations Research*, 41(2), 280-297, March-April 1993.
- [Spiegelhalter-90] Spiegelhalter, D.J. & Lauritzen, S.L. Sequential Updating of Conditional Probabilities on Directed Graphical Structures. *Networks*, 20, 579-605, 1990.
- [Spiegelhalter-92] Spiegelhalter, D.J. & Cowell R.G. Learning in Probabilistic Expert Systems. *Bayesian Statistics 4*, Bernardo J.M., Berger J.O., Dawid A.P. & Smith A.F.M. (Eds.). Oxford University Press, 447-465, 1992.
- [Spiegelhalter-93] Spiegelhalter, D.J., Dawid A.P., Lauritzen, S.L., & Cowell R.G. Bayesian Analysis in Expert Systems. *Statistical Science*, Vol. 8, No. 3, 219-283, 1993.
- [Spirtes-92] Spirtes, P., Glymour, C., Scheines, R. & Meek, C. *TETRAD II User's Manual*, Technical Report, Carnegie Mellon University, Department of Philosophy, Pittsburgh, PA, 1992.
- [Srinivas-90] Srinivas, S. & Breese, J. IDEAL: A Software Package for Analysis of Influence Diagrams, in *Proceedings of the Sixth Conference on Uncertainty in Artificial Intelligence*, Assoc. for Uncertainty in AI: California, 212-219, 1990.

- [Steele-84] Steele, G. *Common Lisp: The Language*. 1st. Ed. Digital Pr.: Bedford, Mass, 1984.
- [Steele-90] Steele, G. *Common Lisp: The Language*. 2nd. Ed. Digital Pr.: Bedford, Mass, 1990.
- [Suermondt-93] Suermondt, H.J. & Cooper, G.F. An Evaluation of Explanations of Probabilistic Inference. *Computers and Biomedical Research*. 26, 242-254, 1993.
- [Swartout-83] Swartout W.R. XPLAIN: a System for creating and Explaining Expert Consulting Programs. *Artificial Intelligence*, 21, 285-325, 1983.
- [Teach-81] Teach, R. & Shortliffe, E. An Analysis of Physicians Attitudes Regarding Computer-based Clinical Consultation Systems. *Comp. Biomed. Res.* 14, 542-558, 1981.
- [vonWinterfeldt-86] von Winterfeldt, D. & Edwards, W. *Decision Analysis and Behavioral Research*. Cambridge University Press: New York, 1986.
- [Wallis-82] Wallis, J.W. & Shortliffe, E.H. Explanatory Power for Medical Systems: Studies on the Representation of causal Relationships for Clinical Consultations. *Methods Inf. Med.* 21, 127-136, 1982.
- [Waterman-86] Waterman, D.A. *A Guide to Expert Systems*. Addison-Wesley Publishing Company, 1986.
- [Wellman-88] Wellman, M.P. Qualitative Probabilistic Networks for Planning under Uncertainty, in Lemmer J.F. & Kanal L.N. (Eds.), *Uncertainty in Artificial Intelligence 2*, Elsevier Sc.: Amsterdam, 1988.
- [Wellman-90] Wellman, M.P. Graphical Inference in Qualitative Probabilistic Networks. *Networks*. Band 20(5), 687-701, 1990.
- [Wellman-90a] Wellman, M.P. Fundamental Concepts of Qualitative Probabilistic Networks. *Artificial Intelligence*, 44, 257-303, 1990.
- [White-93] White, I. *The Booch Method: A Case Study for Rational Rose*. Rational: California, 1993.
- [Wick-89] Wick, M.R. & Slagle, J.R. An Explanation Facility for today's Expert Systems. *IEEE Expert*, 4 (1), 26-36, 1989.
- [Wick-92] Wick, M.R. & Thompson, W.B. Reconstructive Expert System Explanation. *Artificial Intelligence*, 54, 33-70, 1992.
- [Wolverton-95] Wolverton, M. Presenting Significant Information in Expert System Explanation, in *Lecture Notes in Artificial Intelligence*, 990. Pinto Ferreira, C. & Mamede, N. J. (Eds.). Progress in AI. 7th Portuguese Conf. on AI, EPIA 95 Proceedings. Springer Verlag: Berlin, 435-439, 1995.

- [Xu-94] Xu, H. & Kennes, R. Steps Towards the Efficient Implementation of Dempster-Shafer Theory, in Fedrizzi, Kacprzyk & Yager (Eds.), *Advances in the Dempster-Shafer Theory of Evidence*, J.Wiley & Sons, 153-174, 1994.
- [Yager-87] Yager, R.R., Ovchinnikov, S., Yong, R.M. & Nguyen, H.T. (Eds.). *Fuzzy Sets and Applications: Selected Papers by L. A. Zadeh*. Wiley: New York, 1987.
- [Zadeh-78] Zadeh L. A. Fuzzy Sets as a Basis for a Theory of Possibility. *Fuzzy Sets and Systems* 1, 1-28, 1978.