

UNIVERSIDADE FEDERAL DE SANTA CATARINA

Centro Tecnológico

Curso de Pós-Graduação em Ciência da Computação

***UMA ABORDAGEM PROBABILÍSTICA
PARA SISTEMAS ESPECIALISTAS***

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação, para a obtenção do grau de Mestre em Ciência da Computação.

Cristiane Koehler

Florianópolis, Dezembro de 1998

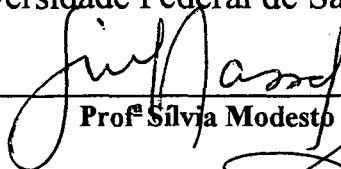
UMA ABORDAGEM PROBABILÍSTICA PARA SISTEMAS ESPECIALISTAS

Cristiane Koehler

Esta Dissertação foi julgada adequada para a obtenção do título de

MESTRE EM CIÊNCIA DA COMPUTAÇÃO

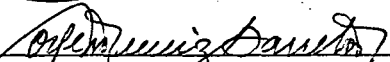
na área de concentração Sistemas de Conhecimento, sub-área Inteligência Computacional e aprovada em sua forma final pelo Curso de Pós-Graduação em Ciência da Computação da Universidade Federal de Santa Catarina.



Prof.^a Silvia Modesto Nassar, Dr. Eng. – Orientadora



Prof.^a Maria Marlene de Souza Pires, Dr. Med. – Co-Orientadora

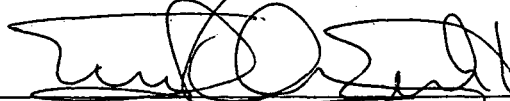


Prof. Jerge Muniz Barreto, Dr. Eng. – Coordenador do Curso

Banca Examinadora:



Prof.^a Aurora Trinidad Ramirez Pozo, Dr. Eng.



Prof. Eugênio Epprecht, Dr. Eng.



Prof. Luiz Fernando Jacinto Maia, Dr. Eng.

*Ao meu querido esposo Nederson.
À minha adorável mãe Solange.
Aos meus irmãos Fernanda, Telmar e Jéssica.*

ARTIGOS PUBLICADOS

“Uma Abordagem Bayesiana para Sistemas Especialistas: Uma Aplicação na Área Pediátrica” em Anais do: QUALISUL – I Encontro do Mercosul Sobre Qualidade na Saúde, realizado no período de 20 a 23 de Agosto de 1998, em Gravatal – SC (Brasil).

“Uma Abordagem Probabilística para Sistemas Especialistas: Avaliação do Estado Nutricional em Crianças de 0 a 2 anos de Idade” em Anais do: III Simpósio Nacional de Informática, pg. 56-59, v.1, no período de 22 a 25 setembro de 1998, em Santa Maria – RS (Brasil).

“Uma Abordagem Bayesiana para Sistemas Especialistas: Uma Aplicação na Área Pediátrica” em Anais do: IV FNCTS – IV Fórum Nacional de Ciência e Tecnologia em Saúde, pg. 265-266, 18 a 23 de outubro de 1998, em Curitiba – PR (Brasil).

“Sistema Especialista Probabilístico de Apoio à Avaliação Nutricional em Pediatria” em Anais do: XIII Congresso Latino Americano y IV Congresso Ibero Americano de Gastroenterología Pediátrica y Nutrición, no período de 27 a 30 de outubro de 1998, (México).

“Sistema Especialista Probabilístico de Apoio à Avaliação Nutricional em Pediatria” em: Revista CIS Ciência e Investigación en Salud, vol. III, nr. especial, outubro, p. 36, 1998, (México).

AGRADECIMENTOS

A Deus, fonte de tudo.

À prof^ª **Sílvia Modesto Nassar** e à prof^ª **Maria Marlene de Souza Pires** pela maravilhosa orientação, dedicação, estímulo, carinho e amizade demonstrados durante o desenvolvimento desta pesquisa.

Ao meu esposo **Nederson da Silva Koehler** pelo grande incentivo, carinho, compreensão, colaboração, apoio e paciência que sempre teve comigo nas horas mais difíceis.

À minha querida mãe **Maria Solange da Costa Folchini** pelo grande exemplo de mulher batalhadora e conquistadora dos seus ideais.

Aos meus queridos avós **Dioclides da Costa** e **Maria Lina da Costa** e à minha tia **Léa Rosange Favaretto** pelas palavras amigas e incentivadoras que muito me ajudaram.

À prof^ª **Lucimar Fossatti de Carvalho** pelo incentivo.

Às minhas queridas amigas do “Clube da Luluzinha”: **Adja Ferreira de Andrade**, **Alessandra Monteiro de Castro Porto**, **Ana Cláudia Garcia Barbosa**, **Fabiane Barreto Vavassori**, **Letícia de Carvalho Pivetta** e **Maria Laura Schneider** pela nossa amizade, companheirismo e coleguismo durante toda a caminhada.

À **Lisiane Albuquerque Reis**, **Moema Luz Cargnin** e **Gilvan Tessari** pela amizade e apoio técnico durante o desenvolvimento desta pesquisa.

À **Verinha** e à **Valdete** pela atenção, carinho e apoio.

SUMÁRIO

| | |
|--|-----------|
| LISTA DE FIGURAS | V |
| LISTA DE TABELAS..... | VI |
| LISTA DE SIGLAS..... | VII |
| RESUMO..... | VIII |
| ABSTRACT..... | IX |
| CAPÍTULO 1 - INTRODUÇÃO..... | 1 |
| 1.1 – OBJETIVOS GERAIS | 2 |
| 1.2 - OBJETIVOS ESPECÍFICOS | 2 |
| 1.3 - JUSTIFICATIVAS..... | 2 |
| 1.4 – ESTRUTURA DO TRABALHO..... | 3 |
| CAPÍTULO 2 – ESTADO DA ARTE EM SISTEMAS ESPECIALISTAS PROBABILÍSTICOS | 5 |
| 2.1 - POR QUE SISTEMAS ESPECIALISTAS PROBABILÍSTICOS ? | 5 |
| 2.2 - SISTEMAS DESENVOLVIDOS | 6 |
| 2.3 - SHELLS DISPONÍVEIS | 9 |
| 2.4 - CONSIDERAÇÕES FINAIS | 16 |
| CAPÍTULO 3 – FUNDAMENTAÇÃO TEÓRICA..... | 17 |
| 3.1 - INTRODUÇÃO | 17 |
| 3.2 - INTELIGÊNCIA ARTIFICIAL (IA) | 20 |
| 3.2.1 - <i>Inteligência Artificial Simbólica (IAS)</i> | 21 |
| 3.2.2 - <i>Inteligência Artificial Conexionista (IAC)</i> | 22 |
| 3.3 - PROBABILIDADE BAYESIANA | 22 |
| 3.3.1 - <i>Espaço de Probabilidade</i> | 24 |
| 3.3.2 - <i>Eventos Compostos</i> | 25 |
| 3.3.3 - <i>Terminologia para os conectivos E e OU</i> | 25 |
| 3.3.4 - <i>Eventos Disjuntos</i> | 26 |
| 3.3.5 - <i>Axiomas e Propriedades de Probabilidade</i> | 26 |
| 3.3.6 - <i>Probabilidade Condicional</i> | 26 |

| | |
|---|-----------|
| 3.3.7 - Propriedades da Probabilidade Condicional | 27 |
| 3.3.8 - Teorema de Bayes | 27 |
| 3.3.9 - Independência de Eventos | 28 |
| 3.3.10 - Probabilidade Condicional e Implicação Lógica | 29 |
| 3.3.11 - Razões de Probabilidade | 29 |
| 3.3.12 - Atualização Bayesiana para uma Nova Evidência | 31 |
| 3.3.13 - Hipóteses Multivaloradas | 32 |
| 3.4 - FORMAS DE RACIOCÍNIO | 33 |
| 3.4.2 - Raciocínio Difuso | 35 |
| 3.4.3 - Raciocínio Qualitativo | 35 |
| 3.4.4 - Raciocínio Probabilístico | 35 |
| 3.5 - SISTEMAS ESPECIALISTAS | 35 |
| 3.5.1 - Conceitos Básicos | 36 |
| 3.5.2 - Componentes de um Sistema Especialista | 37 |
| 3.5.3 - Sistema Especialista Probabilístico | 38 |
| 3.6 - AQUISIÇÃO DE CONHECIMENTO | 39 |
| 3.6.1 - O que é Aquisição de Conhecimento ? | 39 |
| 3.6.2 - Técnicas de Aquisição de Conhecimento | 40 |
| 3.6.3 - Ferramentas computacionais para Aquisição de Conhecimento | 41 |
| 3.7 - REPRESENTAÇÃO DO CONHECIMENTO | 42 |
| 3.7.1 - Representação por Lógica Matemática | 43 |
| 3.7.2 - Regras de Produção | 43 |
| 3.7.3 - Redes Semânticas | 43 |
| 3.7.4 - Redes Neurais Artificiais | 43 |
| 3.7.5 - Redes Bayesianas (RB) | 44 |
| 3.8 - EXPLICAÇÃO DO RACIOCÍNIO DE UM SISTEMA ESPECIALISTA | 49 |
| 3.9 - CONSIDERAÇÕES FINAIS | 50 |
| CAPÍTULO 4 – DOMÍNIO DE APLICAÇÃO | 51 |
| 4.1 - A INCERTEZA E A PRÁTICA CLÍNICA - DR. HÉLIO KOMAGATA | 51 |
| 4.2 – USO CLÍNICO DE MODELOS ESTATÍSTICOS | 53 |
| 4.3 - DESNUTRIÇÃO ENERGÉTICO PROTEICA | 53 |
| 4.3.1 - Conceitos | 53 |
| 4.3.2 - Etiologia da Desnutrição Energético Proteica | 55 |
| 4.3.3 - Gravidade da Desnutrição Energético Proteica | 55 |
| 4.4 - CRESCIMENTO NORMAL DE CRIANÇAS | 57 |
| 4.4.1 - Síndrome da Privação Materna | 58 |

| | |
|--|-----------|
| CAPÍTULO 8 – CONCLUSÕES | 88 |
| 8.1 – TRABALHOS FUTUROS | 90 |
| CAPÍTULO 9 - REFERÊNCIAS BIBLIOGRÁFICAS | 91 |

Lista de Figuras

| | |
|---|----|
| FIGURA 1. COMPONENTES DE UM SISTEMA ESPECIALISTA | 38 |
| FIGURA 2. UM EXEMPLO DE RB [TESSARI 98] | 48 |
| FIGURA 3. SÍNDROME DA PRIVAÇÃO MATERNA [REIS 97] | 59 |
| FIGURA 4. REPRESENTAÇÃO GRÁFICA DA PARTE QUANTITATIVA DA VARIÁVEL EDEMA | 72 |
| FIGURA 5. PRIMEIRA TELA DO <i>SISPAN</i> | 73 |
| FIGURA 6. SEGUNDA TELA DO <i>SISPAN</i> | 74 |
| FIGURA 7. MENU AJÚDA DO SISTEMA | 75 |
| FIGURA 8. AJUDA ON-LINE DO SISTEMA EM FORMA DE TÓPICOS | 75 |
| FIGURA 9. REPRESENTAÇÃO GRÁFICA DA BASE DE CONHECIMENTO (RB) PARA AVALIAÇÃO DO ESTADO NUTRICIONAL | 76 |
| FIGURA 10. TABELA NA <i>SHELL</i> NÉTICA QUE REPRESENTA O VETOR DAS PROBABILIDADES <i>A PRIORI</i> DA VARIÁVEL DE_TIPO | 77 |
| FIGURA 11. REPRESENTAÇÃO GRÁFICA DA PARTE QUANTITATIVA DA VARIÁVEL PESO | 78 |
| FIGURA 12. VETOR DAS PROBABILIDADES <i>A POSTERIORI</i> APÓS UMA CONSULTA À REDE BAYESIANA | 79 |
| FIGURA 13. ANÁLISE DA SENSIBILIDADE DO SISTEMA | 84 |
| FIGURA 14. BASE DE CONHECIMENTO COM O ACRÉSCIMO DE TRÊS VARIÁVEIS | 85 |

Lista de Tabelas

| | |
|--|----|
| TABELA 1. EXEMPLOS DE SEP..... | 6 |
| TABELA 2. RELAÇÃO DE SHELLS PARA DESENVOLVIMENTO DE SEP..... | 9 |
| TABELA 3. DIAGNÓSTICO DO ESTADO NUTRICIONAL..... | 54 |
| TABELA 4. CLASSIFICAÇÃO DA DEP..... | 55 |
| TABELA 5. PRINCIPAIS CARACTERÍSTICAS DO MARASMO E DO KWASHIOKOR..... | 57 |
| TABELA 6. CLASSIFICAÇÃO DE GOMEZ, 1956..... | 62 |
| TABELA 7. CLASSIFICAÇÃO DE KANAWATI, 1970..... | 62 |
| TABELA 8. CLASSIFICAÇÃO DE MCLAREN, 1967..... | 63 |
| TABELA 9. CLASSIFICAÇÃO DE WELLCOME, 1970..... | 63 |
| TABELA 10. RELAÇÃO DE CASOS DE TESTE SUBMETIDOS AO SISTEMA..... | 83 |

Lista de Siglas

API - *Application Program Interface*

DDE - *Dynamic Data Exchange*

DEP – Desnutrição Energético Proteica

DLL - *Dynamic Link Library*

IA – Inteligência Artificial

IAC – Inteligência Artificial Conexionista

IAS – Inteligência Artificial Simbólica

RB – Rede Bayesiana

RBC – Raciocínio Baseado em Casos

SE – Sistema Especialista

SEP – Sistema Especialista Probabilístico

RESUMO

Um dos grandes problemas em Inteligência Artificial, mais precisamente, em Sistemas Especialistas, é o tratamento dos dados incertos, isto é, como tomar uma decisão sem ter as informações necessárias.

A necessidade de tratar a incerteza em Sistemas Especialistas levou a construção de Sistemas Especialistas Probabilísticos. A presente pesquisa realizou um estudo sobre o raciocínio probabilístico em Sistemas Especialistas, tendo como domínio de aplicação a área médica. Foi desenvolvida uma aplicação para auxiliar na avaliação do estado nutricional em crianças de 0 a 2 anos de idade com base nos sinais e sintomas e dados antropométricos. Foi testada a sensibilidade do sistema às mudanças nos valores de probabilidades e foi analisado o comportamento do sistema em função do tamanho da base de conhecimento.

Uma das conclusões desta pesquisa foi que na análise de sensibilidade, o sistema se manteve estável até o ruído de 30% e na análise do comportamento percebeu-se que as variáveis acrescentadas não trouxeram informações relevantes no processo de diagnóstico.

Palavras-Chave: Inteligência Artificial, Raciocínio Probabilístico, Sistemas Especialistas, Incerteza.

ABSTRACT

One of the great problems in Artificial Intelligence, in fact, in Experts Systems, it is the treatment of the uncertain data, that is, as to take a decision without having the necessary information.

The need to treat the uncertainty in Experts Systems took the construction of Probabilistic Experts Systems. To present research it accomplished a study on the probabilistic reasoning in Experts Systems, tends as application domain the medical area. An application was developed to aid in the evaluation of the state nutritional in children of 0 to 2 years of age with base in the signs and symptoms. The sensibility was tested of the system to the changes in the values of probabilities and the behavior of the system was analyzed in function of the size of the knowledge base.

One of the conclusions of this research went that in to sensibility analysis, the system stayed stable until the noise of 30% and in the analysis of the behavior it was noticed that the increased variables didn't bring important information in the diagnosis process.

Key: Artificial Intelligence, Probabilistic Reasoning, Experts Systems, Uncertainty.

Capítulo 1 - Introdução

Um SE é um programa computacional capaz de realizar julgamentos em uma área complexa. A intenção é que o SE tenha capacidade de realizar uma tarefa tão bem quanto um especialista. Em muitos domínios de aplicação, o especialista tem que tomar decisões a partir de informações repletas de incerteza. Estudos sobre SE com capacidade de tratar o conhecimento incerto no domínio de aplicação estão sendo um tópico a mais em IA [NEAPOLITAN 90].

Existem basicamente quatro linhas de estudos na representação do conhecimento incerto [LAURITZEN 88]. O modelo lógico, se utiliza apenas de processamento simbólico [COHEN 85]. O modelo lingüístico baseia-se no raciocínio “fuzzy” para interpretar sentenças imprecisas da linguagem natural [ZADEH 83]. A teoria de Dempster-Shafer utiliza funções de crença [SHAFER 76] e o modelo estatístico está baseado no cálculo de probabilidades, o que lhe dá consistência e confiabilidade [LINDLEY 82]. Possuindo um forte apelo pragmático, já que este modelo possui flexibilidade e meios operacionais de avaliação, crítica e aprendizado de dados [CHEESEMAN 85], além de prover uma metodologia muito adequada à compreensão humana. Nesta pesquisa preocupou-se em estudar uma destas técnicas de representação de incerteza em SE: a abordagem terá a visão estatística.

O primeiro trabalho a sugerir a representação do conhecimento incerto foi baseado na probabilidade condicional de Thomas Bayes, datada de 1763. Daí surgiu a Estatística Bayesiana.

A Estatística Bayesiana passou a ser aplicada em SE sendo uma teoria consistente e que permite a representação de conhecimentos certos e incertos [PEARL 88].

1.1 – Objetivos Gerais

O objetivo geral desta pesquisa é realizar um estudo sobre a abordagem probabilística, no tratamento da incerteza em sistemas especialistas.

1.2 - Objetivos Específicos

- Analisar a sensibilidade do sistema às mudanças nos valores das probabilidades das hipóteses diagnósticas e das probabilidades condicionais,
- Testar o comportamento do sistema em função do tamanho da base de conhecimento,
- Desenvolver uma aplicação na área médica: um protótipo de sistema especialista, o **SISPAN – Sistema Pediátrico para Avaliação Nutricional**.

1.3 - Justificativas

Inicialmente, o fator que motivou o desenvolvimento desta pesquisa foi que o tema escolhido, a aplicação da Teoria da Probabilidade em Sistemas Especialistas para tratar a incerteza, se concentra em uma área emergente de pesquisa aqui na UFSC – Universidade Federal de Santa Catarina, onde o domínio de aplicação foi a área da saúde, mais especificamente, desnutrição infantil de crianças entre zero e dois anos de idade, onde ainda não foi constatado nenhum estudo nesta área.

Em um sistema especialista que utiliza a Teoria da Probabilidade para tratar a incerteza, a base de conhecimento é composta por variáveis (sinais e sintomas), seus respectivos atributos e pela relação entre estas variáveis que foi estabelecida através da definição de fatos e/ou regras na forma de probabilidades.

A estratégia de aquisição do conhecimento foi adquirir estas informações, variáveis e probabilidades, através de entrevistas com o especialista. Por isso a necessidade em analisar a sensibilidade do sistema em função das pequenas alterações nos valores das probabilidades, porque estas probabilidades adquiridas do especialista são probabilidades subjetivas. Com a

análise de sensibilidade do sistema a estas mudanças, será possível observar se as pequenas alterações nos valores das probabilidades alteram, ou não, o diagnóstico. Por outro lado, esta análise será realizada para que o sistema possa ser adotado por outro especialista, desde que este não discorde da base de conhecimentos, permitindo que a base de conhecimento do sistema seja validada externamente.

Além disso, foi constatada a necessidade de testar o comportamento do sistema em função do tamanho da base de conhecimento, devido ao fato de que não se sabe sobre como o sistema responderia caso fossem acrescentadas mais variáveis (sintomas) à base de conhecimento. Com isso poderá ser evidenciado que os sistemas que utilizam a Teoria da Probabilidade para tratar a incerteza, se utilizam de poucas variáveis (sintomas) para decidirem a respeito de um diagnóstico.

1.4 – Estrutura do Trabalho

Este documento está estruturado em oito capítulos. Cada capítulo constitui uma parte essencial sobre o estudo que foi realizado para a realização desta pesquisa.

O capítulo 2 apresenta o Estado da Arte. Uma relação dos SEP desenvolvidos nos últimos quatro anos e uma relação das *shells* disponíveis para o desenvolvimento de SEP.

O capítulo 3 relata a fundamentação teórica desta pesquisa, dando ênfase aos conceitos de IA, Probabilidade Bayesiana, Formas de Raciocínio, Sistemas Especialistas, Aquisição e Representação do Conhecimento.

No capítulo 4, foi detalhado o domínio de aplicação proposto para esta pesquisa.

O sistema proposto para o desenvolvimento desta pesquisa está apresentado no capítulo 5.

No capítulo 6 é apresentada a implementação do sistema proposto: SISPAN - Sistema Pediátrico para Avaliação Nutricional.

Os resultados são apresentados no capítulo 7.

As conclusões obtidas com a elaboração desta pesquisa, bem como sugestões para trabalhos futuros estão comentadas no capítulo 8.

E, posteriormente, está a lista de Referências Bibliográficas.

Capítulo 2 – Estado da Arte Em Sistemas Especialistas Probabilísticos

Este capítulo possui algumas questões sobre a abordagem probabilística em SE, apresenta algumas ferramentas para o desenvolvimento destes sistemas e comenta sobre os SE com abordagem probabilística.

2.1 - Por que Sistemas Especialistas Probabilísticos ?

Os SE são desenvolvidos para resolver uma grande variedade de problemas, sendo que o conhecimento necessário para solucionar estes problemas, geralmente, não é precisamente definido, e além disso, é de natureza incerta, isto é, faltam informações necessárias para a tomada de decisão. Pode-se constatar que muitos problemas na vida real são repletos de incerteza. Especialistas humanos são capazes de formar julgamentos e tomar decisões baseados em informação incerta, incompleta e algumas vezes, até mesmo contraditória. Para que um SE seja confiável, o mesmo deve lidar com este tipo de informação com a mesma facilidade que o ser humano. O desenvolvimento de formalismos para tratar incerteza e os algoritmos, para manipular informação com incerteza tornaram-se um tópico maior de pesquisa em IA, chamado Raciocínio com Incerteza, que será comentado no capítulo 3.

Depois da metade da década de 80, a pesquisa sobre raciocínio probabilístico em SE resultou na introdução de Redes Bayesianas, também chamadas de Redes Causais. Estas redes têm sua origem na teoria da probabilidade e são caracterizadas por um poderoso formalismo que representa o conhecimento no domínio e pelas incertezas associadas a este domínio.

Mais especificamente, o formalismo proporciona uma representação concisa de uma distribuição conjunta de probabilidades em um grupo de variáveis. Associados a este formalismo estão os algoritmos para calcular eficientemente as probabilidades relevantes e para processar as evidências.

Desde sua introdução, a estrutura de RB vem rapidamente ganhando popularidade e agora está começando a mostrar o seu valor em domínios complexos. Aplicações práticas estão sendo desenvolvidas, por exemplo, para diagnóstico e prognóstico médico e para recuperação de informação probabilística [LINDA 96].

A seguir, estão relacionados os principais SEP que foram desenvolvidos no período de 1989 até 1996, sendo que, as áreas de aplicação foram a medicina e a engenharia. Logo após, estão relacionadas, as principais *shells* para desenvolvimento de SEP.

2.2 - Sistemas Desenvolvidos

Tabela 1. Exemplos de SEP

| Data | Nome | Local |
|------|--|---|
| 1996 | MammoNet | Laboratório de IA, Depto de Engenharia Elétrica e Ciência da Computação da Universidade de Wisconsin-Milwaukee, Milwaukee |
| 1994 | DIAVAL | Depto de Informática e Automática - U.N.E.D. - Madrid - Espanha |
| 1989 | <i>General Electric Aircraft Engines Services</i> | <i>Knowledge Industries (KI) , ENTER Software</i> |
| 1989 | <i>Microsoft Pregnancy and Child Care</i> | <i>Microsoft e Knowledge Industries (KI)</i> |
| 1988 | <i>Lógica Autônoma de Controle - Lockheed Martin</i> | <i>Lockheed Martin Marine Systems e Artificial Intelligence Center - Sunnyvale, Califórnia</i> |
| 1972 | DOMBAL | |

1 -DIAVAL : é um SE destinado a ajudar médicos no diagnóstico de enfermidades cardíacas a partir, principalmente, de dados ecocardiográficos. Este software faz a introdução dos dados, apresentação dos resultados do diagnóstico e explicações sobre o raciocínio, destacando as características que diferenciam o DIAVAL de outros SE e de outras RB. Este software é o resultado da tese de doutorado do Prof. Fernandez J. Díez. Este sistema consiste em calcular a probabilidade *a posteriori*, selecionar a doença mais provável e a mais relevante, gerando um relatório [DÍEZ 97].

2 - DXplain : é um sistema de apoio a decisão o qual utiliza um conjunto de dados clínicos (sinais, sintomas e dados laboratoriais) para produzir uma lista de diagnósticos, os quais podem expor (ou serem associados com) as manifestações clínicas. Este software possui uma grande base de dados de probabilidades sobre 5.000 manifestações clínicas associadas com mais de 2.000 diferentes doenças. O sistema utiliza um formato interativo para coletar informações clínicas e faz uso de uma forma modificada do raciocínio Bayesiano para oferecer interpretações clínicas e tem sido utilizado por milhares de médicos e estudantes de medicina. O banco de dados e o sistema estão sendo continuamente aperfeiçoados e atualizados. *DXplain* tem as características de livro texto eletrônico e um manual de referência médico [DXPLAIN 97].

3 - General Electric Aircraft Engines Services : o diagnóstico de falhas em turbinas de aviões em grandes empresas aéreas implica em grandes custos, pois os dados retirados da análise possuem muito ruído e não podem atuar para um defeito específico. As solicitações mecânicas para efetuar os testes e inspeções, assim como a excessiva verificação do motor podem duplicar os custos de manutenção. Para um avião 757 típico, este custo pode chegar a 500.000 dólares. Este sistema possui modelos de RB que guiam um mecânico de motores através de uma seqüência otimizada de testes e passos de inspeção. O sistema de recomendação de testes utiliza os resultados de um modelo do desempenho termodinâmico do motor da aeronave, que processa os dados vindos de sensores montados sobre o motor no estande de teste [KIK 97].

4 - Lógica Autônoma de Controle - Lockheed Martin: a *Lockheed Martin Marine Systems*, em Sunnyvale, Califórnia, em conjunto com o *Artificial Intelligence Center* em Palo Alto, vem desenvolvendo um sistema de controle para o UUV (*Unmanned Underwater Vehicle*). Este veículo submarino poderá permanecer semanas submerso, sem necessidade de intervenção humana. O objetivo da pesquisa é desenvolver software que seja capaz de controlar eventos planejados e eventos não-antecipados, de modo a diminuir o risco de perda do veículo e maximizar a probabilidade de sucesso na missão. O sistema possui maior capacidade de manusear a grande gama de anomalias e alarmes falsos que ocorrem em missões de longa duração, o que o torna bastante superior aos sistemas da época. Quando completamente

desenvolvida, a Lógica Autônoma de Controle permitirá ao UUV monitorar o progresso de sua missão, analisar o estado de seu equipamento, detectar e analisar eventos que causam impacto nos objetivos da missão, tomar decisões e efetuar ações para compensar os eventos, além de modificar o plano da missão quando o atual estiver completamente fora de alcance.

5 - *Microsoft Pregnancy and Child Care*: a Microsoft e a Knowledge Industries (KI) uniram-se para criar um serviço *on-line* de informação sobre saúde, com capacidade de multimídia, que emprega raciocínio bayesiano. O produto é chamado de *Microsoft Pregnancy and Child Care*. Uma RB é um componente central da opção *Find by Symptom* (Busca por Sintoma). Sintomas que ocorrem freqüentemente são conectados a módulos especialistas que guiam os pais da criança através de uma série de questões. O sistema repetidamente determina qual a melhor pergunta a se fazer aos pais, mostrando as apresentações em multimídia que correspondem às exatas condições de saúde da criança. Os pais podem escolher entre artigos ilustrados para obter informação detalhada sobre os sintomas e a doença da criança. Os artigos também contém *links* para outras fontes de informação relacionadas, incluindo outros *Web sites*. O *Microsoft Pregnancy and Child Care* foi criado pela Divisão de Consumo da Microsoft. A KI desenvolveu e testou as bases de conhecimentos em RB, trabalhando com os especialistas Dr. Richard Behrman, Professor de Pediatria na *Stanford University* e na *University of California, San Francisco*, e Dr. David Kerns, Chefe de Pediatria no *Santa Clara Valley Medical Center*.

6 - *MammoNet*: é um sistema de suporte à decisão médica para auxiliar no diagnóstico do câncer de mama dado fatores de risco demográficos, informações radiológicas e sintomas físicos. As probabilidades condicionais foram obtidas através de livros e especialistas médicos. *MammoNet* foi implementado como uma base de conhecimento de regras. A rede que especifica o problema foi construída utilizando um algoritmo de construção de RB. A performance do sistema foi avaliada com 77 casos de um livro texto [MAMMO 98].

7 - *DOMBAL*: um dos primeiros, e talvez um dos melhores, sistemas desenvolvidos na área bayesiana foi o de *DOMBAL* et al. para diagnóstico de dor abdominal. Este sistema utilizava os dados de sensibilidade, especificidade e prevalências para vários sinais, sintomas e testes

laboratoriais para calcular a probabilidade de sete possíveis diagnósticos de dor abdominal (apendicite, diverticulite, úlcera perfurada, colecistite, obstrução intestinal, pancreatite e dor não específica). O sistema, implementado em computador, passou a funcionar como um tipo de exame subsidiário, onde o médico interessado preenchia uma requisição padronizada informando os achados clínicos e a enviava, em seguida, ao serviço de computação. Após alguns minutos o médico recebia um impresso, dando as probabilidades para os diagnósticos diferenciais [DOMBAL 72].

2.3 - Shells disponíveis

Tabela 2. Relação de shells para desenvolvimento de SEP

| Data | Nome | Local |
|------|---------------------------|---|
| 1996 | BANTER | Laboratório de IA, Depto de Engenharia Elétrica e Ciência da Computação da Universidade de Wisconsin-Milwaukee, Milwaukee |
| 1996 | Microsoft Belief Networks | Microsoft |
| 1995 | BNG | Depto de Engenharia Elétrica e Ciência da Computação - Universidade de Wisconsin-Milwaukee |
| 1995 | Graphical-Belief | Depto de Estatística - Universidade de Washington |
| 1995 | Hugin | Hugin Expert A/S e Universidade de Aalborg - Dinamarca |
| 1994 | DxPress | Rockwell Palo Alto Laboratory - Universidade de Stanford |
| 1994 | Strategist | Prevision Incorporatec - Davis - Canadá |
| 1993 | Bayes On-Line | Knowledge Based Engineering (KBE) - África do Sul |
| 1990 | IDEAL | Rockwell Palo Alto Laboratory |
| 1990 | Netica | Norsys Software Corp. - Vancouver, BC, Canadá |
| 1990 | SPIRIT | Universidade de Hagen (FernUniversität Hagen) |

A seguir estão relacionadas as principais vantagens e desvantagens de cada ferramenta:

1 - Bayes On-Line (BOL): é uma ferramenta que ajuda a determinar as causas do problema e prognostica eventos futuros em muitas aplicações, principalmente industriais. Ele utiliza um rede para modelar uma aproximação baseada na regra de Bayes da teoria da probabilidade. Automatiza a identificação de causas prováveis de um problema detectado e facilita a resolução do problema. BOL é valioso para sensores de validação, diagnósticos, prognósticos, análise de decisão, análise de risco, análise probabilística de segurança e monitoramento de alarmes. É aplicável a problemas que requerem respostas em tempo real, conectividade, ou ferramentas sofisticadas para decisões estratégicas.

BOL trabalha, preferencialmente, com distribuições de probabilidades complexas, elaboração de árvores de decisão, desenvolvimento de aplicações que possam trabalhar com o nível de componente com BOL, especificando somente as causas imediatas e os efeitos de um evento. Estes relacionamentos são usados para desenvolver gráficos, os quais podem ser compreendidos por usuários não-especialistas. Ele permite uma fácil representação de fenômenos dinâmicos, tais como: persistência, efeitos atrasados e *feedback* [BOL 97].

2 - Bayesian Network Generator (BNG): é um sistema para construção de base de conhecimentos com RB. Uma classe de RB é especificada com uma base de conhecimento de regras. Regras podem conter variáveis temporais e não-temporais, bem como restrições do contexto. BNG constrói uma RB estruturalmente mínima para calcular a probabilidade *a posteriori* de uma consulta. Este sistema oferece duas vantagens sobre o uso das RB tradicionais: representacional e computacional. A incorporação das variáveis quantificadas na base de conhecimento permite a representação da informação não expressível em RB tradicionais. Pois somente uma pequena porção de um grande modelo probabilístico pode ser relevante para a computação de uma particular probabilidade posterior, usando BNG pode resultar em economias computacionais significantes. Estas economias podem ser particularmente grandes com RB temporais. BNG está escrito em CommonLisp e possui interface para a ferramenta IDEAL que será apresentada mais adiante [BNG 97].

3 - Graphical-Belief: é um programa que oferece um ambiente interativo para explorar modelos baseados em RB. Uma RB é um modelo de rede utilizado para descrever os relacionamentos entre variáveis. Este software é dividido em dois módulos principais, a saber: Análise do Modelo e Construção do Modelo. É uma ferramenta para explorar os aspectos preditivos dos modelos. É baseado na tecnologia de modelos gráficos (também conhecidos como diagramas de influência, redes de crença ou RB) as quais já têm sido muito utilizadas na análise de decisão e IA. Estes modelos têm sido utilizados com sucesso em diversas áreas, tais como: segurança de sistemas, tomada de decisões médicas, planejamento financeiro e gerenciamento operacional. Este programa é um ambiente completo para construir e explorar modelos de risco. Este software oferece várias ferramentas para modelar tarefas, tais como:

- 1) Organização: O modelo somente especifica a interação direta entre as variáveis no modelo;

calcula e mostra todas as dependências envolvidas entre variáveis distantes. 2) Engenharia do Conhecimento: Uma grande quantidade de conhecimento fica armazenado dentro de um modelo gráfico tradicional. *Graphical-Belief* oferece ferramentas para manter este conhecimento. O modelo pode ser projetado através de uma biblioteca anteriormente construída e de estruturas de conhecimento genérico. Uma outra vantagem em se utilizar o *Graphical-Belief* é que ele utiliza um esquema orientado a objetos para armazenar o conhecimento; uma alteração de uma simples regra do protótipo ou de uma variável é prontamente propagada para todo o modelo. 3) Flexibilidade: o usuário pode selecionar a forma de representação dos relacionamentos entre as variáveis: probabilidade (para relacionamentos com incerteza) e lógica (para relacionamentos com certeza). *Graphical-Belief* completa esta flexibilidade com uma máquina de inferência genérica a qual pode ser simplesmente expandida para incluir outras representações para os relacionamentos. 4) Visualização Dinâmica: o usuário manipula diretamente o modelo, explorando as implicações entre os relacionamentos [GRAPHICAL 97].

4 - Microsoft Belief Network Tools : este software permite a criação e avaliação de RB. É composto pelo executável MSBN32.EXE que utiliza uma versão proposta pelo *Bayes Net Interchange Format*, isto é, uma representação de um arquivo texto de RB. A aplicação está escrita em Visual Basic 4.0 e possui uma .DLL para Windows (MSBN32.DLL). Este módulo é o componente ativo do conjunto, e suporta as seguintes operações: a) procura e armazena RB em texto para criação e modificação da rede através da adição de nós e arcos da avaliação das probabilidades discretas das RB utilizando um método chamado "*exact clique-tree propagation methods decision-theoretic troubleshooting*"; b) diagramas assimétricos de influência de simples decisão. É uma aplicação padrão com interface multi-documentos para Windows. Ele tem uma barra de ferramentas e um *help* sensível ao contexto.

5 - Netica: este software utiliza RB para realizar vários tipos de inferência usando algoritmos modernos e rápidos. Dado um novo caso, pelo qual o usuário tem conhecimento limitado, Netica encontrará os valores ou probabilidades apropriadas para todas as variáveis desconhecidas. O caso pode ser, convenientemente, salvo em um arquivo e depois incluído na rede (ou em uma rede diferente) para incrementar a consulta, ou para trazer uma nova

informação sobre o caso. Utiliza diagramas de influência para encontrar as decisões ótimas, as quais maximizam os valores esperados das variáveis especificadas. Variáveis que não são de interesse por muito tempo podem ser removidas sem uma mudança geral entre os relacionamentos entre as variáveis remanescentes.

Vantagens:

- gera uma apresentação gráfica de qualidade, a qual pode ser incorporada dentro de outros documentos,
- pode encontrar decisões ótimas para problemas de decisão sequencial,
- soluciona diagramas de influência eficientemente pelo uso de clique na árvore,
- pode aprender relações probabilísticas através de dados,
- permite atualização fácil da RB e dos diagramas de influência, incluindo: excluir, colar e duplicar nós da RB e dos diagramas de influência,
- mantém diagramas complexos ordenados,
- permite comentários,
- permite desfazer e refazer comando ilimitadamente,
- permite a entrada de relações probabilísticas através de equações, com uma extensa biblioteca de funções probabilísticas e funções matemáticas.
- tem facilidade para realizar a discretização de variáveis contínuas,

Netica está disponível em versões para Windows 95, Windows NT 4.0 e Macintosh. A versão para UNIX foi disponibilizada em 1997 [NETICA 97].

6 – Strategist: oferece uma ferramenta para construir e analisar modelos graficamente. O usuário poderá analisar graficamente o seu modelo de decisão. Os modelos incluem diagramas de influência para mostrar a estrutura dos relacionamentos entre as variáveis. *Strategist* é baseado na probabilidade Bayesiana e na teoria de decisão estatística. As tecnologias existentes para operacionalização da análise de decisão, tais como: árvores de decisão e diagramas de influência, ainda deixam muito a desejar por parte do usuário. As principais características do *Strategist* são [STRATEGIST 97]:

- trabalha com RB.
- trabalha com diagramas de influência hierárquico.

- uma rica linguagem para representar o relacionamento entre as variáveis (contingências, relacionamentos *noisy-or* e *noisy-and*, divisões).
- trabalha com variáveis contínuas.
- realiza testes com os modelos utilizando situações do mundo real.
- possui análise de sensibilidade.

7 – Dxpress: é uma ferramenta para aquisição de conhecimento para desenvolvimento rápido de modelos probabilísticos para diagnósticos. Escrito em C++, este programa utiliza várias técnicas para acelerar a aquisição do conhecimento, para reduzir a quantidade de tempo necessário para desenvolver um SE. Por exemplo, foi desenvolvido um protótipo de sistema de diagnóstico para uma turbina à jato usando energia elétrica e pneumática no DC-10. Este sistema diagnostica 60 diferentes falhas utilizando aproximadamente 50 testes ou observações diferentes [KIK 97].

Vantagens:

- avaliação rápida das distribuições de probabilidades utilizando funções de independência causal.
- localiza o progresso durante as diferentes fases da aquisição do conhecimento.
- permite ao usuário dar entrada em notas e definições pertinentes para o processo de aquisição do conhecimento.

8 – IDEAL (Influence Diagram Evaluation and Analysis in Lisp): é uma base de testes para trabalhar com RB e diagramas de influência, ambos estão descritos no capítulo 3. Contém diversos algoritmos de inferência para RB e algoritmos de avaliação para diagramas de influência. Permite criar e editar diagramas de influência e RB [SRINIVAS 90]. IDEAL foi escrito em Common Lisp. Este programa possui uma clareza do código e oferece uma programação de alto nível. Desta forma, o IDEAL é adequado para implementações experimentais que necessitem tecnologia de RB. Em aplicações mais avançadas, IDEAL pode ser usado como uma biblioteca de sub-rotinas as quais oferecem inferência em RB e avaliação de diagramas de influência. IDEAL vem com uma interface gráfica opcional que depende da versão do Common Lisp que o usuário utiliza. Uma versão para pesquisa do

IDEAL pode ser obtida através de *download* na forma de um arquivo .TAR para Unix. O manual vem na forma de um arquivo *Postscript* [IDEAL 97].

9 – SPIRIT : é uma *shell* para SE criada na Universidade de Hagen (FernUniversität Hagen), na Alemanha, desenvolvida pela equipe do Prof. Willhem Rödden. É uma ferramenta que possui interface gráfica de desenvolvimento. Permite a criação de variáveis de diversos tipos, tais como: booleana, ordinal, cardinal e nominal. A parte qualitativa da RB é implementada a partir da inserção de regras de produção. Em seguida, valores de probabilidades são associados a estas regras e às variáveis, caracterizando a parte quantitativa da RB [SPIRIT 97]. Após a caracterização da RB, faz-se necessário inicializá-la, preparando-a para a compilação, ou seja, aprendizagem das regras. Feita a compilação pode-se efetuar inferência sobre a base, inserindo evidências de um caso específico, chegando a um diagnóstico com um determinado valor de probabilidade.

Vantagens:

- facilidade de acesso ao seu criador,
- permite maior interação com o usuário, por não ser um software comercial, e
- roda em várias plataformas, inclusive *Windows*.

10 – Hugin: é uma *shell* de origem dinamarquesa para construção de RB. Foi desenvolvida por Frank Jensen e Lars Bo Nielsen. É uma ferramenta capaz de construir modelos baseados em SE em domínios caracterizados pela incerteza. Os modelos suportados são RB e diagramas de influência. Permite ao usuário definir nós discretos e alguns nós contínuos em seus modelos. A origem desta *shell* deu-se durante o projeto do programa ESPRIT sobre diagnóstico de doenças neuromuscular. Um grupo de pesquisa da Universidade de Aalborg, na Dinamarca, trabalhou no desenvolvimento correto e eficiente dos métodos de cálculo para o problema do diagnóstico. Naquela época, os pesquisadores obtiveram alguns resultados, mas ainda restava um problema que impedia a utilização de RB na construção de SE. O problema foi chamado de *rumour problem*: “Posso ouvir a mesma estória através de diversos canais; mas a estória ainda pode se originar da mesma fonte. Sem saber se meus canais são, ou não, independentes, não posso combinar as declarações corretamente”.

Em RB o *rumour problem* aparece quando uma causa pode influenciar o mesmo evento através de diferentes caminhos na rede. O problema foi resolvido e métodos gerais foram disponibilizados para serem utilizados em vários domínios, os quais podem ser modelados por uma RB. Os métodos têm sido programados dentro de um sistema de desenvolvimento geral, de fácil operação por qualquer usuário que deseja construir um SE baseado em uma RB. Este sistema foi chamado HUGIN. Desde então o sistema tem sido aprimorado com a facilidade dos diagramas de influência.

HUGIN apresenta a parte qualitativa da rede de forma clara, incluindo a direção dos grafos e a parte quantitativa é apresentada em forma de tabelas. A primeira fase na implementação de uma base de conhecimentos no HUGIN é a construção da parte qualitativa. Com estas informações é possível definir todas as variáveis e seus respectivos atributos. A parte quantitativa consiste em preencher as tabelas que são apresentadas quando cada variável é selecionada. O tamanho da tabela varia conforme o número de *links* que o nó observado possui com *patern nodes* e outros nós. *Patern Nodes* são os nós onde chegam grafos de dependências. Finalizando este processo, está terminada a etapa de edição da RB. A etapa seguinte é a compilação e o programa passa para o modo de execução.

O usuário tem a oportunidade de utilizar o HUGIN através HUGIN Runtime - um ambiente gráfico de fácil acesso para o usuário. O usuário pode também utilizar o HUGIN API, que é uma biblioteca de funções para C (ou C++) [HUGIN 97].

11 – BANTER: é uma ferramenta educacional para trazer a informação contida em uma RB para o usuário final de uma forma intelegível. A *shell* BANTER foi projetada para usuários tutores na avaliação de hipóteses e seleção dos melhores processos diagnósticos. BANTER poder ser usado com qualquer RB que contém nós que podem ser classificados dentro de hipóteses, observações e processos diagnósticos. Oferece capacidade para calcular a probabilidade *a posteriori* de hipóteses, determinar o melhor processo de diagnóstico para afirmar ("*rulle in*") ou excluir ("*rulle out*") uma hipótese, interroga o usuário na seleção do processo diagnóstico ótimo e explica o raciocínio do sistema. Oferece também, uma interface para modelar RB e cria problemas para uso instrucional. Foi escrito em C e necessita da *shell*

HUGIN, descrita anteriormente, para executar [BANTER 97].

2.4 - Considerações Finais

Este capítulo teve como objetivo, fazer um relato sobre os SEP desenvolvidos durante os últimos anos, podendo assim perceber que a abordagem probabilística tem sido utilizada como auxílio em tomada de decisões, tanto na área médica quanto na engenharia. Este relato também descreveu as *shells* para desenvolvimento de SEP com o objetivo de auxiliar no desenvolvimento de SEP.

Durante esta pesquisa, foram utilizadas basicamente três *shells* com o objetivo de testá-las com exemplos de RB extraídas de bibliografias. Estas *shells* são as seguintes:

- Spirit: desenvolvida pela Fernuniversität Hagen, da Alemanha,
- Hugin: desenvolvida pela Hugin Expert A/S, da Dinamarca e
- Netica: desenvolvida pela Norsys Software Corp, do Canadá.

Com a relação de todas estas *shells* para SEP e das suas vantagens e desvantagens, escolheu-se como *shell* para desenvolver a aplicação proposta nesta pesquisa, a *shell* Netica. As justificativas referentes à esta escolha estão comentadas no capítulo 6.

Capítulo 3 – Fundamentação Teórica

3.1 - Introdução

Sistemas Especialistas e tratamento de incerteza em Inteligência Artificial têm tido um grande surgimento de atividade de pesquisa durante a última década.

Uma das dificuldades no desenvolvimento de SE é o tratamento da incerteza nas informações e a imprecisão nos dados, as quais são as variáveis fundamentais no processo do raciocínio. Infelizmente, determinar a “melhor” conclusão em SE pode não ser tão fácil nem simples [GIARRATANO 89].

É necessário ressaltar que incerteza difere de imprecisão, isto é, diz-se que existe incerteza em uma informação quando não se tem certeza absoluta de ela ser verdadeira ou não. Geralmente é representado por um valor numérico que indica o **grau de certeza**, este valor pode encontrar-se no intervalo $[0, 1]$, onde 1 indica que se está certo de que o fato é verdadeiro e 0 indica que se está certo de que o fato não é verdadeiro. Por outro lado, imprecisão (nebulosidade) ocorre quando a informação não tem seus limites definidos. Por exemplo, a idade, para alguns, 25 anos é jovem, e para outros 35 anos é jovem. Isto quer dizer que não existem valores quantitativos únicos que definam o termo “jovem”. Ao contrário da incerteza, não depende do quanto se acredita se a informação é verdadeira ou não, mas é inerente às expressões lingüísticas utilizadas [COLLAZOS 97].

Uma alternativa para solucionar o problema do tratamento da incerteza em SE é a utilização de Redes Bayesianas [PEARL 88] que será detalhada posteriormente.

Dois conceitos muito importantes na representação do conhecimento incerto: Crença e Evidência.

Crença é quando uma pessoa acredita na informação e tenta justificá-la. Por exemplo, se uma criança está com diarreia, o médico acredita que esta criança ingeriu uma comida estragada. Esta crença é medida em números estatísticos. Assim, o médico acredita com 90% de certeza ter sido a ingestão de comida estragada que causou a diarreia.

Evidência é quando uma pessoa tem como justificar uma crença. Por exemplo, se uma criança ingeriu uma comida estragada, isto é uma evidência. E, a justificativa do por que a criança está com diarreia.

O problema é como associar um grau de crença ou descrença às afirmações baseadas em evidências.

As RB tratam a incerteza de forma mais simples e os resultados são mais seguros. Por exemplo, geralmente, nos deparamos com as seguintes afirmações:

- João tem febre.
- João tem gripe.
- Se alguém tem gripe, então será que podemos afirmar que esta pessoa terá febre ?

Não se tem certeza de que estas afirmações sejam verdadeiras. Acredita-se e não se acredita ao mesmo tempo.

Sendo assim, “crença” e “evidência” não têm a mesma definição.

As estimativas das crenças citadas acima são obtidas através:

- do especialista (probabilidade subjetiva) e
- da observação da frequência de uma dada informação (no caso de se dispor de uma base de dados grande), isto é, quanto maior a ocorrência de uma informação, mais evidente ela se torna dentro do sistema.

Difícilmente, as crenças têm boas estimativas. Sendo assim, devemos considerar a questão: *O que fazer se a estimativa da crença é ruim ?*

Pode-se dizer que a imprecisão e incerteza são conceitos diferentes, mas os dois pertencem ao tipo de informações inexatas. Para solucionar esse problema, diversos métodos foram desenvolvidos para levar em consideração esses tipos de informação. Esses métodos procuram meios para representar a imprecisão e/ou a incerteza e assim proporcionar uma inferência mais aproximada à realidade. Existem várias formas numéricas para tratar com informações inexatas, pode-se citar a Probabilidade Bayesiana, os Fatores de Certeza, a Teoria de Dempster-Shafer e os Métodos Difusos. A seguir, são conceituadas cada uma destas propostas.

Probabilidade e Teorema de Bayes

A probabilidade sempre foi relacionada com a medida de eventos repetíveis de sistemas ideais. No entanto, hoje em dia se fala de probabilidade subjetiva, trata com eventos que não têm uma base histórica (base de dados) sob a qual se possa extrapolar. A probabilidade subjetiva é atualmente uma crença ou opinião expressa como uma probabilidade mais do que uma probabilidade baseada em medidas empíricas. As crenças e opiniões de um especialista têm um papel importante no desenvolvimento de SEP.

Os SEP utilizam o teorema de Bayes como base para o raciocínio sob a incerteza.

Fatores de Certeza

Esta abordagem teve como pioneiro o sistema MYCIN [SHORTLIFFE 76], o qual tenta recomendar terapias apropriadas para pacientes com infecções bacteriológicas.

No MYCIN, o grau de confirmação foi originalmente definido como um fator de certeza (FC), que é a diferença entre crença e descrença.

O fator de certeza pode ser utilizado para um conjunto de hipóteses em ordem de importância. Ou indicar a crença em uma hipótese sob alguma evidência.

Teoria de Dempster-Shafer

Está baseado no trabalho feito originalmente por Dempster, que tentou modelar a incerteza por uma faixa de probabilidades mais do que um simples número probabilístico

[GIARRATANO 89]. Esta abordagem considera grupos de proposições e atribui a cada um deles um intervalo para o grau de crença. Considera que crença mede a força da evidência em favor de um grupo de proposições. Ela varia de 0 (que indica falta de evidências) a 1 (que denota certeza) [RICH 88].

Métodos Difusos

Nos métodos difusos os quantificadores e o raciocínio utilizam linguagem natural, onde encontra-se palavras com significados imprecisos. Este método trabalha com funções de pertinência e variáveis lingüísticas.

Nesta pesquisa, foi explorada a abordagem Probabilística, destacando as formas de Aquisição do Conhecimento e como este foi representado.

3.2 - Inteligência Artificial (IA)

O objetivo da IA é representar o comportamento inteligente através de modelos computacionais. Este comportamento está fundamentado em dois temas: o conhecimento e a aprendizagem. O conhecimento é sua principal matéria prima, podendo ser representado por muitos formalismos. A aprendizagem também tem despertado muito o interesse da comunidade de IA.

Como apresentar uma definição convincente de IA ? Muita discussão existe em busca da definição ideal, mas nada como comentar algumas das definições correntes [BARRETO 97]:

- Para Patrick Winston “IA é o estudo de conceitos que permitem aos computadores serem inteligentes”;
- Para Elaine Rich e Kevin Knight “IA é o estudo de como fazer os computadores realizarem coisas que, no momento, as pessoas fazem melhor”;
- Para Eugene Chermiak e Drew McDermott “IA é o estudo das faculdades mentais com o uso de modelos computacionais”.

Analisando as definições citadas, observamos que a primeira define IA utilizando o termo

inteligente, o que causa recursividade por não esclarecer inteligência. Já a segunda definição condena a IA a nunca conseguir um sucesso, pois cada vez que for resolvido um problema ele deixa de ser do domínio de IA. A definição mais convincente dentre as citadas é a terceira, pois não apresenta recursividade, se relaciona com outras ciências (por exemplo: psicologia, lógica e fisiologia) no que diz respeito às faculdades mentais e além disso se mostra independente de arquiteturas computacionais, podendo ser adotada tanto por computadores digitais como por neurocomputadores¹ [BARRETO 97].

Em IA existem duas correntes de pensamento chamadas de paradigmas: o Paradigma Simbólico e o Paradigma Conexionista. A seguir, tem-se uma descrição destes dois paradigmas.

3.2.1 - Inteligência Artificial Simbólica (IAS)

Na IAS o comportamento inteligente é simulado, baseado nos princípios da psicologia cognitiva.

Na IAS se espera que um conhecimento sobre o problema particular a resolver e das técnicas úteis para o caso, possam levar a manipulação dos conhecimentos básicos e imitando, ao menos até certo ponto, o modo de raciocínio usado por humanos.

Para isto, IAS deve ser utilizada quando o domínio do problema é bem definido, que se tenha idéia de como ele será resolvido e que seja explícito o modo de como achar uma solução.

Alguns resultados da IAS são bem conhecidos, tais como: programas de jogar xadrez, monitoração, previsões financeiras, diagnósticos, etc. Usam especialmente ferramentas como: lógica, regras de produção e técnicas de extração de conhecimento de especialistas. Além disso, muitos aspectos dos mecanismos de cognição são melhor compreendidos hoje graças aos avanços da IAS [BARRETO 97].

¹ Computadores em que os circuitos tentam imitar o sistema nervoso.

3.2.2 - Inteligência Artificial Conexionista (IAC)

A IAC acredita na construção de neurocomputadores e se estes forem bastante parecidos ao cérebro humano ele apresentará um comportamento inteligente.

Segundo Barreto [BARRETO 97] “Se for construído um modelo suficientemente preciso do cérebro, este modelo apresentará um comportamento inteligente. Se apenas uma pequena parte do cérebro for reproduzida, a função exercida por esta parte emergirá do modelo”.

Na IAC, acredita-se que construindo uma máquina que imite a estrutura do cérebro, ela apresentará inteligência como comportamento emergente [BARRETO 97].

A IAC vem atraindo a atenção de pesquisadores, pois apresenta uma melhor performance quando o problema não é bem definido (falta de conhecimento explícito de como realizar a tarefa).

O estudo abordado nesta pesquisa foi desenvolvido considerando os conceitos da IAS.

3.3 - Probabilidade Bayesiana

A interpretação da probabilidade como o grau de certeza de um determinado evento, é chamada Bayesiana, e aquela que associa a probabilidade com a frequência de um evento é chamada frequentista ou objetiva [HECKERMAN 95].

Os métodos Bayesianos possibilitam representar numericamente o grau de certeza sobre condições de incerteza, e manipulá-lo de acordo com as regras definidas na teoria da probabilidade, pois a teoria Bayesiana está fundamentada nesta teoria [HECKERMAN 95].

A estatística Bayesiana tem origem no nome de Thomas Bayes, matemático inglês que viveu no século dezoito, cujos trabalhos na área de probabilidade permitiram o surgimento de estudos nesta área. Na realidade, foi o grande matemático francês, Pierre Simon de LaPlace,

que desenvolveu o teorema de Bayes na forma que este é hoje conhecido e utilizado [HECKERMAN 95].

O teorema de Bayes é de grande importância para o cálculo de probabilidades. Quando melhor compreendido, o teorema se mostra como a lei fundamental que governa o processo de inferência lógica. O teorema é a base para analisar um conjunto de informações disponíveis e chegar a uma conclusão objetiva, expressa numericamente.

Mais especificamente, na área de interesse desta pesquisa, o teorema de Bayes é o mecanismo fundamental que permite relacionar diversas probabilidades conhecidas de um prognóstico clínico e deduzir uma probabilidade que com melhor eficácia estime o resultado.

O teorema de Bayes é um método quantitativo para a revisão de probabilidades conhecidas, com base em nova informação amostral.

Se perguntarmos à você qual a probabilidade de ocorrer cara ao lançar uma moeda de um real? É quase certo que sua resposta seria que a probabilidade é igual a 0,50 . Aqui se deve ressaltar dois aspectos: que você não tinha certeza de que face ocorreria ao lançar a moeda, isto é, você estava diante da aleatoriedade do resultado; e que ao expressar o valor 0,50 você estava explicitando uma medida de crença na ocorrência de cara com base em seu suposto conhecimento da situação.

Nos SEP os valores de probabilidade refletem a crença do especialista sobre o que ele espera que ocorra em situações similares àquelas que tem tido experiência e aprendizado. Isto é, o especialista está tentando extrapolar com base em sua experiência e aprendizado no domínio de aplicação.

Supondo agora a seguinte situação:

Se uma pessoa lança uma moeda de um real, observa o resultado e não diz às outras pessoas. Qual a probabilidade de ter ocorrido cara?

Sua resposta continuaria sendo 0,50 . Porém, a pessoa que viu o resultado poderia declarar um ou zero. Qual a resposta correta, a sua ou a da pessoa que viu o resultado ? As duas respostas estão corretas! Aqui vale ressaltar que a probabilidade é baseada em informação. Eu conheço como a moeda caiu e a pessoa que não viu o resultado somente conhece que a moeda foi lançada. Assim, nossas respostas diferentes refletem nosso estado de conhecimento sobre a situação.

A idéia de que as probabilidades se alteram com a mudança de conhecimento é crucial para os SEP, também chamados de Sistemas Especialistas Bayesianos que serão descritos posteriormente.

Estes sistemas têm em sua base de conhecimentos fatos e regras que representam o conhecimento do especialista num domínio de aplicação. Aos fatos e às regras são associadas as incertezas presentes no domínio, e é explicitada a crença em sua ocorrência através de valores de probabilidade. O raciocínio realizado pelo sistema deve considerar estas probabilidades para a partir dos dados de entrada associar um vetor de probabilidades ao conjunto de hipóteses diagnósticas. A hipótese com maior probabilidade de ocorrência pode ser considerada a conclusão do sistema, note que a esta conclusão está associado ao grau de certeza da resposta do sistema.

Neste momento cabem duas questões iniciais para o desenvolvimento de SEP:

- 1) O que são os valores de probabilidade associados aos fatos e regras do SEP ?
- 2) Como o sistema raciocina com estes conhecimentos ?

Para responder a estas perguntas é feita, a seguir, uma apresentação formal dos conceitos de probabilidade [NASSAR 98].

3.3.1 - Espaço de Probabilidade

Seja ϵ um conjunto finito e P uma função de ϵ para os números reais não negativos, tal que:

$$\sum P(e) = 1 \quad \forall e: e \in \epsilon$$

O par (ϵ, P) é chamado de espaço de probabilidade. Os elementos de ϵ são chamados de eventos simples ou elementares. P é chamado de distribuição de probabilidade ou função de probabilidade. Desta definição resulta que a probabilidade de qualquer evento e denotada por $P(e)$, é medida por um número no intervalo $[0;1]$.

Intuitivamente, ϵ é a coleção de resultados que se pode esperar em um domínio de aplicação. O valor $P(e)$ é uma estimativa da crença de que o resultado e ocorra.

3.3.2 - Eventos Compostos

Seja um espaço de probabilidade (ϵ, P) .

Um subconjunto A do conjunto ϵ é denominado de evento composto.

Para um evento composto A , $A \subseteq \epsilon$ define-se a probabilidade do evento A da seguinte forma:

$$P(A) = \sum_{e \in A} P(e)$$

Se e ocorre e $e \in A$ então diz-se que A ocorre.

Ao evento composto consistindo do conjunto de elementos de ϵ que não estão em A é chamado de negação ou complemento de A e é denotado por $\neg A$ ou \bar{A} .

3.3.3 - Terminologia para os conectivos E e OU

Considerando dois eventos A e B e adotando a terminologia \wedge e \vee para os conectivos e e ou, respectivamente; então:

$P(A \wedge B)$ significa a probabilidade de que ambos os eventos ocorram; isto é a probabilidade do evento A ocorrer e do evento B ocorrer.

$P(A \vee B)$ significa a probabilidade de que pelo menos um dos eventos ocorre; isto é a probabilidade do evento A ocorrer ou do evento B ocorrer.

3.3.4 - Eventos Disjuntos

Denotando por \cup e \cap a união e a intersecção de conjuntos então:

$$A \cup B = \{e: e \in A \text{ OU } e \in B\}$$

$$A \cap B = \{e: e \in A \text{ E } e \in B\}$$

Se $A \cap B = \emptyset$ então A e B são eventos disjuntos ou mutuamente excludentes, e $P(A \cap B) = 0$.

3.3.5 - Axiomas e Propriedades de Probabilidade

a) $0 \leq P(A) \leq 1$

b) $P(\epsilon) = 1$

c) $P(A \cup B) = P(A) + P(B)$ se A e B são eventos disjuntos.

d) Se $A \subseteq B \subseteq \epsilon$ então $0 \leq P(A) \leq P(B) \leq 1$

e) Se $A, B \subseteq \epsilon$ então $P(A \cap B) + P(A \cap \bar{B}) = P(A)$ e $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

f) $P(A) + P(\bar{A}) = 1$

g) Se $A_i \subseteq \epsilon$ para $1 \leq i \leq n$ e $A_i \cap A_j = \emptyset$ então para todo $i \neq j$

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = P(A_1) + P(A_2) + \dots + P(A_n)$$

3.3.6 - Probabilidade Condicional

Considerando agora que duas moedas de um real são lançadas. Adicionalmente, sabe-se que pelo menos uma cara ocorreu. Como construir um novo espaço de probabilidade (ϵ, \mathbf{P}') que seja consistente com o espaço de probabilidade (ϵ, \mathbf{P}) anterior e com a nova informação?

A construção de um novo espaço de probabilidade (ϵ, \mathbf{P}') onde o anterior (ϵ, \mathbf{P}) é alterado devido a uma nova informação, é o problema central no desenvolvimento de SEP. O conceito de probabilidade condicional permite considerar as novas informações de forma a obter as novas probabilidades.

Sejam A e B eventos compostos de um espaço de probabilidades (ϵ, \mathbf{P}) . Suponha que um evento simples e ocorra. A probabilidade $P(B)$ é a probabilidade de que $e \in B$ dado nosso conhecimento inicial refletido por P. Intuitivamente, $P(B|A)$ é a probabilidade que $e \in B$

quando se tem a informação adicional de que $e \in A$.

Seja (ϵ, P) um espaço de probabilidade e seja $A \subseteq \epsilon$ tal que $P(A) \neq 0$. Definindo um novo espaço de probabilidade (ϵ, f) da seguinte forma:

$$P(e) / P(A) \quad \text{se } e \in A$$

$$f(e) = 0 \quad \text{se } e \notin A$$

para qualquer $B \subseteq \epsilon$ a probabilidade condicional de B dado a ocorrência de A é igual a $f(B)$. Observe que neste caso A é o novo espaço de probabilidade, onde B deve ser analisado.

Se $A = \epsilon$ então $P(B|A) = P(B)$.

3.3.7 - Propriedades da Probabilidade Condicional

Seja um espaço de probabilidade (ϵ, P) . Se $C \subseteq \epsilon$ então $P(C) \neq 0$. Segue-se que:

1. $P(A|C) = P(A \cap C) / P(C)$

2. Se $A \subseteq B \subseteq \epsilon$ então $0 \leq P(A|C) \leq P(B|C) \leq 1$

3. Se $A, B \subseteq \epsilon$ então $P(A|C) = P(A \cap B | C) + P(A \cap \bar{B} | C)$ e

$$P(A \cup B | C) = P(A|C) + P(B|C) - P(A \cap B | C)$$

4. Se $A_i \subseteq \epsilon$ para $1 \leq i \leq n$ e $A_i \cap A_j = \emptyset$ então para todo $i \neq j$

$$P(A_1 \cup A_2 \cup \dots \cup A_n | C) = P(A_1 | C) + P(A_2 | C) + \dots + P(A_n | C)$$

5. Se $A \subseteq \epsilon$, $B_1 \cup B_2 \cup \dots \cup B_n$ para $1 \leq i \leq n$ e $B_i \cap B_j = \emptyset$ para todo $i \neq j$ então

$$P(A) = P(A|B_1) \cdot P(B_1) + P(A|B_2) \cdot P(B_2) + \dots + P(A|B_n) \cdot P(B_n)$$

3.3.8 - Teorema de Bayes

Seja o espaço de probabilidade (ϵ, P) e os eventos compostos e $H_1, H_2, \dots, H_K \subseteq \epsilon$, desde que nenhum desses eventos tenha probabilidade nula, então:

$$P(H_i|e) = \frac{P(e|H_i) \cdot P(H_i)}{P(e)}$$

Se $P(H_i \cap e) \neq 0$ para todo i então

$$\frac{P(H_i|e)}{P(H_j|e)} = \frac{P(H_i)}{P(H_j)} \cdot \frac{P(e|H_i)}{P(e|H_j)}$$

Se os eventos $H_1 \cup H_2 \cup \dots \cup H_k = \varepsilon$ e $H_i \cap H_j = \emptyset$ para todo $i \neq j$ então

$$P(e) = P(H_1) \cdot P(e|H_1) + P(H_2) \cdot P(e|H_2) + \dots + P(H_k) \cdot P(e|H_k)$$

resultando

$$P(H_i|e) = \frac{P(e|H_i) \cdot P(H_i)}{\sum_{j=1}^k (P(H_j) \cdot P(e|H_j))}$$

Nas aplicações dos SEP os H_i 's são as hipóteses concorrentes. O evento e pode ser pensado como uma evidência. O conhecimento da ocorrência desta evidência leva a mudanças na probabilidade *a priori* $P(H_i)$ para a probabilidade condicional $P(H_i|e)$, que por sua vez considera a evidência e .

3.3.9 - Independência de Eventos

Seja um espaço de probabilidade (ε, P) .

E, sejam os eventos $e_1, e_2 \subseteq \varepsilon$. Segue-se que:

- Se $P(e_1 \wedge e_2) = P(e_1) \cdot P(e_2)$ então os eventos e_1 e e_2 são independentes.
- Genericamente, para qualquer subconjunto $E = \{e_{i1}, e_{i2}, \dots, e_{iik}\}$ de $\{e_1, e_2, \dots, e_n\}$ se $P(e_{i1} \wedge e_{i2} \wedge \dots \wedge e_{iik} | H) = P(e_{i1} | H) \cdot P(e_{i2} | H) \dots P(e_{iik} | H)$ então

pode-se dizer que os eventos e_i 's são eventos mutuamente independentes dado a hipótese H .

A idéia básica subjacente ao conceito probabilístico de independência é que o conhecimento de certa informação não traz informação adicional sobre outra coisa. Isto é, se e somente se, saber que o evento e_1 ocorreu não trazer informação sobre o evento e_2 e saber que o evento

e_2 ocorreu não trazer informação sobre o evento e_1 , então diz-se que ocorre a independência entre estes eventos.

Por exemplo, dois testes diagnósticos são independentes quando não há associação entre os resultados de ambos os testes.

3.3.10 - Probabilidade Condicional e Implicação Lógica

Na teoria da probabilidade foi possível observar que $P(A \vee B)$, $P(A \wedge B)$, $P(\neg B)$ eram adequadas aos conectivos lógicos OU, E e NÃO, respectivamente. Será que o conceito de probabilidade condicional $P(A|B)$ é adequado à implicação lógica ($B \rightarrow A$), isto é “se B então A”?

Sabe-se que $(B \rightarrow A) = (A \vee \neg B)$ então:

$$\begin{aligned} P(B \rightarrow A) &= P(A \vee \neg B) = P(\neg B) + P(A \wedge B) = P(\neg B) + P(B \wedge A) = \\ &= P(\neg B) + P(B) \cdot P(A|B) \end{aligned}$$

como os valores de probabilidade são não negativos então

$$P(B \rightarrow A) > P(A|B)$$

Dessa forma, pode-se concluir que a probabilidade condicional não é adequada à implicação lógica.

A probabilidade condicional utiliza a ocorrência de evidência(s) para chegar à uma conclusão sobre uma hipótese diagnóstica, assim pode ser considerada uma forma de raciocínio.

3.3.11 - Razões de Probabilidade

Razões de probabilidade são chamadas *ODDS* ou *LIKELIHOOD ratio*. No contexto de SEP as seguintes razões de probabilidade podem ser utilizadas:

a) *ODDS a priori* – é a razão entre as probabilidades de duas hipóteses diagnósticas.

$$O(H) = P(H_i) / P(H_j)$$

b) *ODDS a posteriori* – é a razão entre as probabilidades condicionais de duas hipóteses diagnósticas dado que se conhece a evidência e .

$$O(H|e) = P(H_i|e) / P(H_j|e)$$

c) **LIKELIHOOD ratio** – é a razão entre as probabilidades condicionais de uma evidência e considerando a ocorrência de uma ou outra hipótese diagnóstica.

$$L(e | H) = P(e | H_i) / P(e | H_j)$$

Observe que a *ODDS a priori* mede quão mais provável é a hipótese H_i em relação à hipótese H_j . A *ODDS a posteriori* mostra quão mais provável é a hipótese H_i em relação à hipótese H_j , porém considerando a ocorrência da evidência e . Já a *LIKELIHOOD ratio* mede a sensibilidade (confiabilidade) da evidência e considerando duas hipóteses diagnósticas.

Geralmente estes termos *ODDS* e *LIKELIHOOD ratio* são empregados quando se tem somente duas hipóteses diagnósticas H_1 e H_2 . Estas hipóteses, neste caso, são complementares, isto é, $H_1 = \bar{H}_2$.

Pode-se demonstrar que as seguintes expressões são verdadeiras:

$$O(H | e) = O(H) \cdot L(e | H)$$

$$P(H | e) = O(H | e) / (1 + O(H | e)) = 1 / (1 + O(H | e))$$

Agora será expandido o conceito acima considerando a ocorrência de uma outra evidência e_2 . Isto é, deseja-se saber qual o efeito de uma outra evidência e_2 na probabilidade de uma hipótese H_i .

$$\begin{aligned} O(H | e_1 \wedge e_2) &= P(H_i | e_1 \wedge e_2) / P(H_j | e_1 \wedge e_2) = \\ &= (P(H_i | e_1) / P(H_j | e_1)) \cdot (P(e_2 | H_i \wedge e_1) / P(e_2 | H_j \wedge e_1)) = \\ &= (P(H_i) / P(H_j)) \cdot (P(H_i | e_1) / P(H_j | e_1)) \cdot (P(e_2 | H_i \wedge e_1) / P(e_2 | H_j \wedge e_1)) \end{aligned}$$

supondo que as evidências e_1 e e_2 são independentes então

$$O(H | e_1 \wedge e_2) = (P(H_i) / P(H_j)) \cdot (P(H_i | e_1) / P(H_j | e_1)) \cdot (P(e_2 | H_i) / P(e_2 | H_j))$$

Note que a primeira parcela da multiplicação acima é a *ODDS a priori* que reflete a falta de informação de evidências. A segunda parcela reflete a informação sobre a ocorrência da evidência e_1 . A terceira parcela reflete a informação sobre a ocorrência da evidência e_2 . Estas duas últimas parcelas são as *LIKELIHOOD ratios* para as evidências e_1 e e_2 , respectivamente.

Supondo agora a ocorrência de k hipóteses diagnósticas obtém-se a probabilidade condicional da hipótese H_i dado as evidências e_1 e e_2 :

$$P(H_i | e_1 \wedge e_2) = \left\{ \sum_{j=1}^{k-1} [(P(H_j) / P(H_i)) \cdot (P(H_j | e_1) / P(H_i | e_1)) \cdot (P(e_2 | H_j) / P(e_2 | H_i))] \right\}$$

resultando que

$$P(H_i | e_1 \wedge e_2) = [O(H_j | e_1) \cdot L(e_2 | H_j \wedge e_1)]^{-1}$$

3.3.12 - Atualização Bayesiana para uma Nova Evidência

Uma característica interessante da regra de atualização de Bayes é seu esquema computacional recursivo e incremental.

Seja H uma hipótese e $e^n = (e_1, e_2, \dots, e_n)$ uma seqüência de dados independentes observados no passado e seja e um novo fato. A probabilidade condicional para a nova evidência é:

$$\begin{aligned} P(H | e^n \wedge e) &= P(H \wedge e^n \wedge e) / P(e^n \wedge e) = \\ &= (P(e^n) \cdot P(H | e^n) \cdot P(e | e^n \wedge H)) / ((P(e^n) \cdot P(e | e^n))) \end{aligned}$$

resultando em

$$P(H | e^n \wedge e) = P(H | e^n) \cdot (P(e | e^n \wedge H) / P(e | e^n))$$

O resultado acima mostra que uma vez calculada a probabilidade condicional da hipótese H dado o conjunto e^n de evidências, isto é o valor $P(H | e^n)$, os dados passados e^n podem ser desprezados e assim pode ser obtido o impacto da nova evidência e . A crença velha ($H | e^n$) assume o papel de crença *a priori* no cálculo do impacto da nova informação e ; a probabilidade $P(H | e^n)$ sumariza completamente a experiência passada e para sua atualização necessita somente ser multiplicada pela *LIKELIHOOD ratio* $P(e | e^n \wedge H)$. Esta razão mede a probabilidade do novo dado e considerando a hipótese H e os dados passados e^n .

Geralmente, adota-se que a *LIKELIHOOD ratio* é independente dos dados passados e considera somente a nova evidência e . Por exemplo, é razoável supor que um paciente desenvolverá um certo sintoma e considerando que tem a patologia H independentemente de que sintomas e^n teve no passado.

A natureza incremental do processo de atualização para a nova evidência e pode ser explorado utilizando a razão ODDS :

$$O(H | e^n \wedge e) = O(H | e^n) \cdot L(e|H)$$

$$\log O(H | e^n \wedge e) = \log O(H | e^n) + \log L(e|H)$$

assim o logaritmo da *LIKELIHOOD ratio* da evidência e pode ser visto como um peso da própria evidência e . Caso a evidência e suporte a hipótese H então terá um peso positivo, se for oposta a H então terá um peso negativo.

Vale aqui ressaltar que atualizar recursivamente as medidas de crenças está fortemente relacionada ao conceito de independência condicional visto anteriormente.

3.3.13 - Hipóteses Multivaloradas

Nos domínios de aplicação de SEP é comum a existência de várias hipóteses concorrentes cada uma com um conjunto distinto de evidências, dizemos então que o sistema tem um conjunto multivalorado de hipóteses. Assim, o SEP deve realizar o diagnóstico, isto é, selecionar a hipótese (*output*), mais provável, a partir de um conjunto de evidências conhecidas (*input*) e que representam os dados de uma nova consulta ao sistema.

Supondo que se tenha m hipóteses e n evidências, a seguir, será obtido o vetor de probabilidades condicionais das várias hipóteses concorrentes.

Dado um conjunto de evidências $e_1, e_2, e_3, \dots, e_k, \dots, e_n$ a crença na i -ésima hipótese H_i é dada

$$P(H_i | e_1, e_2, e_3, \dots, e_k, \dots, e_n) = \frac{P(e_1, e_2, e_3, \dots, e_k, \dots, e_n | H_i) \cdot P(H_i)}{P(e_1, e_2, e_3, \dots, e_k, \dots, e_n)}$$

Chamando a probabilidade $[P(e_1, e_2, e_3, \dots, e_k, \dots, e_n)]^{-1}$ de constante de normalização e denotando por α . E, supondo a independência condicional das evidências com relação a cada hipótese H_i , então

$$P(H_i | e_1, e_2, e_3, \dots, e_k, \dots, e_n) = \alpha \cdot P(H_i) \cdot \prod_{k=1}^n P(e_k | H_i)$$

Note que:

- a) $P(H_i)$ é o vetor de probabilidades *a priori* das diferentes hipóteses em consideração;
- b) A constante de normalização α deve ser calculada pela restrição de soma unitária das probabilidades $P(H_i | e_1, e_2, e_3, \dots, e_k, \dots, e_n)$.

Assumindo ainda, que para evidência e_k define-se o vetor λ^k de probabilidades condicionais das m hipóteses, isto é:

$$\lambda^k = (\lambda_1, \lambda_2, \dots, \lambda_m) \quad \text{onde} \quad \lambda_i = P(e_k | H_i)$$

Finalmente, pode-se reescrever a equação original na seguinte forma:

$$P(H_i | e_1, e_2, e_3, \dots, e_k, \dots, e_n) = \alpha \cdot P(H_i) \cdot \Delta$$

onde $\Delta = \lambda^1 \cdot \lambda^2 \cdot \lambda^3 \cdot \dots \cdot \lambda^n$

Assim fica computacionalmente mais fácil obter o vetor de probabilidades condicionais das hipóteses H_i dado um conjunto de evidências. Por outro lado, com a expressão acima, fica também fácil verificar que a regra de atualização Bayesiana apresenta um esquema computacional recursivo e incremental a cada evidência nova que ocorrer.

No raciocínio probabilístico são utilizadas variáveis aleatórias para representar eventos e/ou objetos do domínio de aplicação. Ao se fazer várias instanciações deste pode-se alterar o estado atual do domínio. Assim envolve a obtenção de probabilidade conjunta de variáveis aleatórias. O que se faz necessário obter informações sobre as relações entre estas variáveis. No pior caso, necessitar-se-ia das probabilidades de todas as combinações possíveis de variáveis aleatórias, isto levaria à explosão combinatória. Por outro lado, associa-se valores de probabilidades condicionais a uma cadeia de regras SE-ENTÃO. Dessa forma, ao invés de se explicitar a distribuição conjunta de probabilidades, são explicitadas tabelas de probabilidades condicionais que por sua vez podem ser utilizadas para obter as distribuições conjuntas. Um exemplo destas tabelas de probabilidades condicionais está demonstrado na Figura 4.

3.4 - Formas de Raciocínio

Para ser inteligente, é necessário ter Conhecimento e Estratégias de raciocínio.

Comportamento inteligente implica ligar estas duas características e ainda por cima poder deduzir fatos que não estão explícitos no conhecimento e produzir reações sensíveis a estes fatos.

Em humanos, existe uma consciência, que nos permite entender conceitos como *o que* e *porque*, que é a intencionalidade. Com esta habilidade, somos capazes de realizar julgamentos racionais e agir de acordo. É claro que a “razão” imbuída em nosso julgamento é quase sempre subjetiva, assim como o é também nosso conceito de comportamento inteligente [RICH 88].

A imprecisão e a incerteza das informações podem ser consideradas como a falta de informação necessária para a tomada de decisão.

O raciocínio é a faculdade humana de chegar a conclusões a partir de conhecimentos ou informações sobre um problema. Um estudo sobre as formas de raciocínio pode ser encontrada em [COLLAZOS 97]. Na continuação apresentam-se algumas formas de raciocínio: Raciocínio Plausível, Raciocínio Difuso, Raciocínio Qualitativo e Raciocínio Probabilístico.

3.4.1 - Raciocínio Plausível

Raciocínio Plausível é a maneira de chegar a uma conclusão a partir do conhecimento de evidências sobre o fato que se está analisando. As principais características deste tipo de raciocínio são: mantém conjecturas, é seguro e sem controvérsias, sempre aprende-se algo novo sobre o mundo, o objeto principal é distinguir uma suposição de outra, isto é, uma suposição mais razoável de uma menos razoável. Desenvolve-se uma habilidade prática e leva à aprendizagem por imitação e exercícios.

O raciocínio plausível trabalha com suposições e transitoriedade, envolvendo os seguintes processos: Experiência, Indução, Generalização, Especialização e Analogia [COLLAZOS 97].

3.4.2 - Raciocínio Difuso

O Raciocínio Difuso é baseado em dados imprecisos, os quais são representados por graus de pertinência em um conjunto difuso, chegando a uma conclusão. Este raciocínio trabalha com funções de pertinência, grau de pertinência, domínio e universo de discurso [COLLAZOS 97].

3.4.3 - Raciocínio Qualitativo

O Raciocínio Qualitativo baseia-se em informações qualitativas para chegar a uma conclusão. Este raciocínio envolve dois pontos principais: os valores que podem ser adotados pelas variáveis no espaço quantitativo e o tempo qualitativo, isto é, tempo durante o qual os valores qualitativos são válidos [BARRETO 90].

3.4.4 - Raciocínio Probabilístico

O Raciocínio Probabilístico é aquele que, apoiando-se em informações probabilísticas sobre fatos de um domínio, chega a uma conclusão a respeito de um novo caso [COLLAZOS 97].

O raciocínio probabilístico em SE é uma avaliação completa e acessível das fundamentações teóricas e dos métodos computacionais que estão sujeitos ao raciocínio plausível sob incerteza [PEARL 88].

3.5 - Sistemas Especialistas

Sabe-se que, recentemente, houve o surgimento de outros paradigmas de SE em IA, mas nesta pesquisa a aplicação proposta concentrou-se apenas no paradigma da IAS.

Um SE é uma aplicação da área de IA que toma as decisões ou soluciona problemas em um domínio de aplicação, pelo uso do conhecimento e regras analíticas definidas por especialistas neste domínio.

Os SE solucionam problemas que normalmente são solucionados por “especialistas” reais.

Para solucionar tais problemas, os SE precisam acessar uma base de conhecimentos do domínio de aplicação. Estes sistemas também exploram um ou mais mecanismos de raciocínio, para aplicar seu conhecimento aos problemas que têm diante de si. Logo após, o sistema necessita de um outro mecanismo para explicar o que fez ao usuário.

No início dos estudos sobre SE, o primeiro sistema a ser considerado foi o DENDRAL, em 1965, no entanto, sua aplicabilidade ficou restrita ao meio acadêmico. Em 1976 desenvolveu-se o SE mais conhecido, o MYCIN que também ficou restrito aos meios acadêmicos. Somente em 1982, com o desenvolvimento do XCON, os SE saíram das Universidades e ganharam o interesse das indústrias. Atualmente, o setor manufatureiro e empresarial faz uso de SE [LINDA 96].

3.5.1 - Conceitos Básicos

Para compreender o funcionamento de um SE, faz-se necessário a relação dos conceitos básicos para o desenvolvimento do mesmo. São eles:

- **Backward Chaining (Encadeamento regressivo):** é o procedimento para solução de um problema, que começa com uma proposição e um conjunto de regras que conduzem a proposição, e então, trabalha para trás, unificando as regras com informações da base de fatos até a proposição ser verificada ou considerada errada.
- **Forward Chaining (Encadeamento progressivo):** é o procedimento para solução de um problema, que começa com um conjunto de regras e um banco de fatos e trabalha para alcançar uma conclusão baseada em fatos que unifique todo o conjunto de premissas contido nas regras.
- **Shell para SE:** é um *software* que facilita a construção de SE pelo fornecimento de esquemas de representação do conhecimento e de máquinas de inferência. O Engenheiro de Conhecimento apenas adiciona conhecimento.

- **Aquisição do Conhecimento:** é o processo de coletar conhecimento especializado do especialista real para disponibilizar em um SE.
- **Representação do Conhecimento:** é a notação ou formalismo usado para codificar o conhecimento a ser armazenado em um SE.

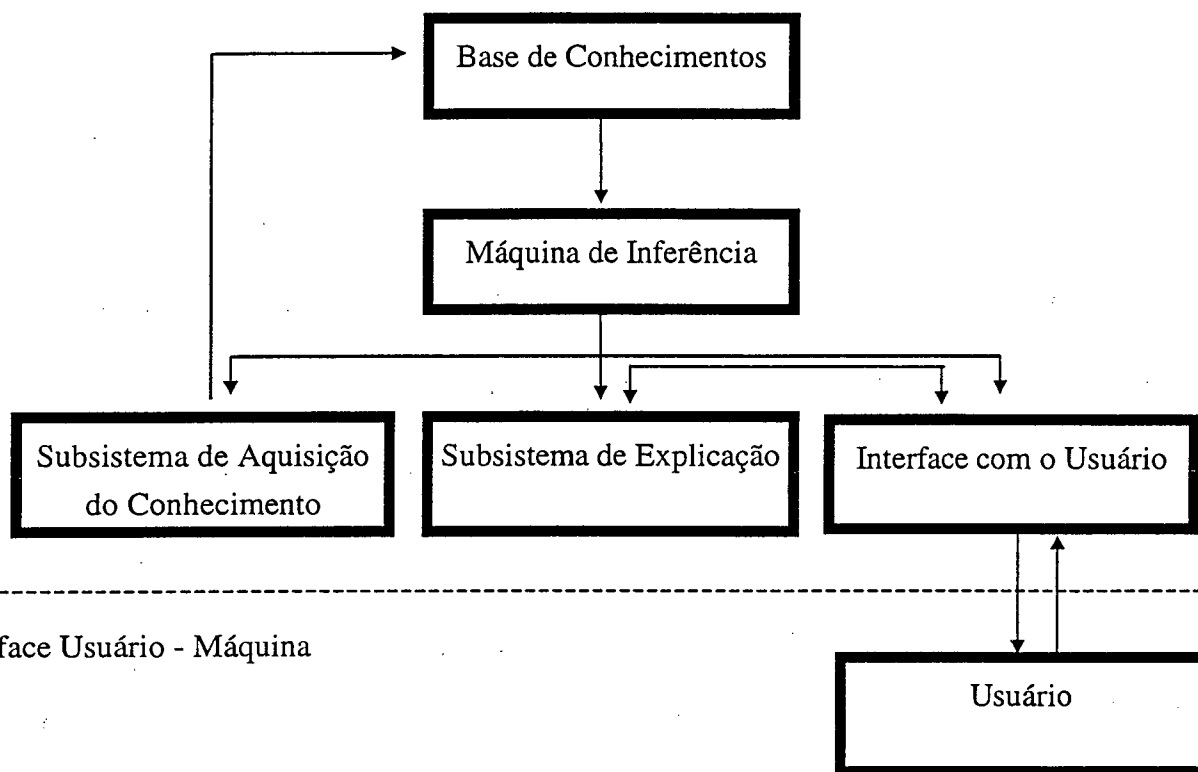
3.5.2 - Componentes de um Sistema Especialista

Os componentes básicos de um SE são:

- **Máquina de Inferência:** indica a operação a ser realizada sobre o conhecimento contido no SE; busca as regras necessárias a serem avaliadas, ordena-as de uma maneira lógica e direciona o processo de inferência; compara *strings* com padrões, e se a comparação for satisfatória, uma outra atitude será tomada, buscando novas regras ou fazendo um outro encaminhamento, até se atingir um determinado objetivo.
- **Base de Conhecimento:** fornece as características de funcionamento do sistema, armazena os fatos e regras a serem usados.
- **Subsistema de Aquisição do Conhecimento:** é usado para alimentar a base de conhecimento. Nele pode-se introduzir, alterar ou excluir regras de acordo com a necessidade; constituído por um editor de texto construído especialmente para ser usado pelo especialista e/ou engenheiro do conhecimento na transferência do conhecimento para o sistema.
- **Subsistema de Explicação:** tem por função, elucidar a linha de raciocínio que o SE usa para chegar a uma conclusão. Interage com o usuário esclarecendo-o de como o sistema chegou a conclusão, se isso for solicitado.
- **Interface:** exibe toda a transação de informações que ocorrem durante a consulta. Pode ser em forma de menus, perguntas ou até mesmos ícones; exibe todas as respostas, perguntas e resultados das consultas e imprime relatórios.

A seguir, na Figura 1, estão ilustrados os componentes básicos de um SE descritos acima.

Sistema Especialista



Interface Usuário - Máquina

Usuário

Figura 1. Componentes de um Sistema Especialista

3.5.3 - Sistema Especialista Probabilístico

Um SEP é assim denominado, quando este atende todas ou uma parte das características citadas anteriormente e que a base de conhecimentos possa ser representada por uma distribuição de probabilidades. A base desta distribuição deverá envolver um conjunto de variáveis discretas com atributos. Entre as variáveis serão necessárias relações de dependência estocásticas, que podem ser especificadas através de fatos e/ou regras. Isto implica na formulação de regras de produção do tipo SE-ENTÃO.

O sistema opera a partir de um tratamento matemático de distribuições marginais sobre o produto cartesiano de todos os atributos de todas as variáveis, e tanto os fatos como as regras, estabelecem as condições iniciais a partir das quais uma distribuição de probabilidade conjunta é processada [RÖDDER 93].

Os SEP têm em sua base de conhecimentos fatos e regras que representam o conhecimento do especialista real. Aos fatos e às regras são associadas as incertezas presentes no domínio e é explicitada a crença em sua ocorrência através dos valores de probabilidades. O raciocínio realizado pelo sistema considera estas probabilidades a partir dos dados de entrada e associa um vetor de probabilidades ao conjunto de hipóteses diagnósticas. A hipótese com maior probabilidade de ocorrência pode ser considerada a conclusão do sistema. Esta conclusão está associada ao grau de certeza da resposta do sistema [NASSAR 98].

3.6 - Aquisição de Conhecimento

Como são construídos os SE ? Tipicamente, um engenheiro do conhecimento interage com um especialista no domínio da aplicação para adquirir o conhecimento do especialista, que é então traduzido para regras. Depois que o sistema inicial estiver pronto, ele precisa ser iterativamente refinado até aproximar-se do nível de desempenho de um especialista. Este processo é caro e lento, portanto vale a pena procurar maneiras mais automatizadas para construir as bases de conhecimentos.

Embora ainda não exista nenhum sistema de aquisição do conhecimento totalmente automatizado, existem vários programas que interagem com os especialistas dos domínios para extrair conhecimento especializado com eficiência. Estes programas estão comentados no item 3.6.3 a seguir e fornecem suporte na inserção de conhecimento, na manutenção da consistência da base de conhecimentos, garantindo sua completude.

3.6.1 - O que é Aquisição de Conhecimento ?

A aquisição de conhecimento é a fase mais complexa da montagem da base de conhecimentos, por tratar-se da interação entre o engenheiro de conhecimentos e o

especialista. Deve-se levar em conta a dificuldade do especialista em deixar clara a sua forma de raciocínio. Algumas vezes pode ser muito difícil a transcrição do conhecimento do especialista para um grupo de regras SE-ENTÃO. Neste ponto, o raciocínio probabilístico demonstra uma vantagem, que é a de representar exceções. Não é viável determinar regras para todas as exceções em um domínio proposto. Podemos, então, fixar uma probabilidade de nosso raciocínio estar correto, deixando uma margem para as exceções, que geralmente são raras, correspondendo a uma porcentagem pequena. Um exemplo destas exceções pode ser observado na Figura 4. Com estas idéias em mente, parte-se para as entrevistas com o especialista. Porém, antes de partir para as entrevistas, o Engenheiro do Conhecimento, deve ter um conhecimento básico sobre o domínio de aplicação e sobre os principais conceitos da Teoria da Probabilidade.

3.6.2 - Técnicas de Aquisição de Conhecimento

O processo de entrevistar um especialista real para extrair conhecimento especializado apresenta várias dificuldades, quer a entrevista seja conduzida por um ser humano quer por uma máquina. Os especialistas são surpreendentemente pouco claros quando se trata de mostrar como eles solucionam problemas, pois o ser humano demonstra uma certa dificuldade em associar números (dados estatísticos) ao quanto ele crê em um determinado fato. É mais fácil observar se um determinado fato depende de outro, ou se um afeta direta ou indiretamente, ou ainda, não afeta um novo fato [PEARL 86]. Daí percebe-se que a representação em termos gráficos, proporciona uma maior aproximação do raciocínio humano, com suas relações de dependências mais explícitas.

Além das entrevistas com o especialista exigirem que haja um certo conhecimento da área de aplicação por parte do engenheiro, este conhecimento pode ser conseguido em conversas com o próprio especialista, ou em pesquisas bibliográficas, o que é bem mais difícil. Ter essa familiaridade com a área de aplicação facilita muito a aquisição de informações.

A primeira conversa com o especialista serve para sentir algumas das dificuldades existentes na transposição de conhecimento do especialista para um sistema. Nessa primeira etapa o engenheiro de conhecimento tem uma idéia de como seguir com as próximas entrevistas.

3.6.3 - Ferramentas computacionais para Aquisição de Conhecimento

Os programas mais eficazes para realizar a aquisição do conhecimento são aqueles que se restringem a um determinado paradigma de solução de problemas. Por exemplo, diagnósticos ou projetos. Um programa de aquisição de conhecimentos deve ser capaz de enumerar os papéis que o conhecimento pode desempenhar no processo de solução de problemas [RICH 88]. Muitas pesquisas na área da aquisição do conhecimento estão sendo realizadas. Na tentativa de suprir algumas destas dificuldades, algumas ferramentas de aquisição do conhecimento têm sido desenvolvidas e utilizadas em diferentes campos de aplicação. Exemplos incluem "MOLE" para diagnósticos, "SALT" utilizado na configuração de tarefas, "KNACK" na avaliação de projetos eletro-mecânicos, "EDE" utilizado para localizar falhas nos equipamentos, entre outros [FOSSATTI 97].

Um exemplo de ferramenta de aquisição do conhecimento em SEP é o *DXpress*. O *DXpress* é uma ferramenta de aquisição de conhecimento para o desenvolvimento rápido de modelos probabilísticos. Esta ferramenta está escrita em C++, e utiliza diversas técnicas de aceleração na aquisição do conhecimento para reduzir a quantidade de tempo necessária para desenvolver um SE. Esta ferramenta facilita uma avaliação rápida das distribuições de probabilidades usando partições e funções de independência causal. Localiza o progresso durante as diferentes fases da aquisição de conhecimento [DXPRESS 97].

Um outro tipo de ferramenta de aquisição de conhecimento é a *M-KAT* [MKAT 97]. Esta ferramenta faz aquisição do conhecimento médico. Ela foi projetada para ajudar os usuários, geralmente um médico especialista, na construção e gerenciamento de um SE capaz de realizar raciocínio médico. Ela foi realizada por um modelo computacional que podia implementar o modelo *Select and Test model (ST-Model)* previamente desenvolvido em colaboração com o *Department of Philosophy of the University of Pavia*. Este modelo ajuda o usuário na realização de uma análise cognitiva, simplificando a construção de qualquer sistema baseado em conhecimento médico. O objetivo de tais análises é para identificar quais são os diferentes tipos/classes de inferência envolvidas em cada tarefa médica, e quais são as entidades em que as inferências atuam. Para que seja útil no mundo real, um SE deve ser capaz de resolver problemas complexos em domínios bem definidos. Para este propósito, os

SE devem possuir uma quantidade suficiente de conhecimento. *M-KAT* permite que o usuário construa um esboço do sistema baseado em conhecimento sobre os resultados obtidos de uma análise cognitiva. *M-KAT* oferece uma coleção de primitivas para a entrada dos vários elementos ontológicos, isto é, o conhecimento relativo para especificar o setor médico em consideração e o modelo de inferência que é o conhecimento estratégico que prescreve como a ontologia deve ser usada para resolver um caso particular. *M-KAT* foi desenvolvida usando a linguagem *Common Lisp Language* e *KEE* que, atualmente, é um dos ambientes de desenvolvimento para sistemas baseados em conhecimento mais poderosos.

3.7 - Representação do Conhecimento

A representação do conhecimento consiste de um sistema que proporciona acesso ao conhecimento, isto é, uma estrutura de dados para representação em memória e um método (o processo computacional) para acessá-lo.

Os métodos são necessários para modelar esses conhecimentos eficientemente e colocá-los prontos para serem acessados, ou seja, a representação do conhecimento é feita através de métodos de IA.

A finalidade da representação do conhecimento é organizar a informação exigida em uma forma tal que o programa em IA possa realmente acessá-lo para tomar decisões.

Uma estrutura de representação de conhecimento deve mostrar de maneira útil e pragmática a informação essencial extraída de uma determinada descrição; deve permitir a busca e a recuperação relativamente eficiente de descrições básicas.

Existem vários métodos para representar o conhecimento, entre eles podemos citar: Representação por Lógica Matemática, Regras de Produção, Redes Semânticas, Redes Neurais Artificiais [RICH 88] e Redes Bayesianas:

3.7.1 - Representação por Lógica Matemática

É um método para representar uma base de conhecimento, empregando lógica de predicados de primeira ordem. Com este método, uma base de conhecimento pode ser vista como uma coleção de fórmulas lógicas, as quais proporcionam uma descrição parcial do mundo. As modificações nestas bases resultam a partir de adições e deleções de fórmulas lógicas.

3.7.2 - Regras de Produção

Consiste de uma base de conhecimentos e uma base de dados globais, os quais representam o *status* do sistema e uma regra de interpretação (estrutura de controle) para escolher as regras a executar. Uma vantagem importante é a facilidade para representar conhecimento heurístico, particularmente informação, em um domínio específico que pode permitir a melhor dedução direta do processo de solução.

3.7.3 - Redes Semânticas

As redes semânticas partem do conceito que a memória humana é associativa, assim o conhecimento é modelado como um conjunto de nós conectados por ligações (arcos), que descrevem relações entre os nós. Os nós representam objetos, conceitos ou eventos, e os arcos representam as relações do tipo “é um” e “é parte de”, dentro de uma hierarquia.

Uma das primeiras finalidades do uso de redes semânticas foi encontrar relações entre objetos ao espalhar a ativação de cada um de dois nós e ver onde esta se propagava. Este processo foi denominado busca de intersecção. Aplicações mais recentes de redes semânticas utilizam procedimentos de busca mais direcionados para responder perguntas específicas. Embora, em princípio, não haja restrições sobre o modo da informação ser representada nas redes, estes procedimentos de busca mais direcionados só podem ser eficazes se houver compatibilidade entre o significado de cada nó e de cada arco.

3.7.4 - Redes Neurais Artificiais

Estas redes são caracterizadas pelos seus neurônios, as conexões entre eles e pelo seu esquema de aprendizado. O neurônio artificial tenta imitar o neurônio biológico. O neurônio biológico é formado por núcleo, axônio e dendritos e se comunica com outros neurônios através de junções entre o axônio de um com um dos dendritos do outro denominadas sinapses. Em um neurônio artificial, ou elemento processador, os axônios e dendritos são modelados através de conexões que interligam dois neurônios, geralmente em um único sentido, com pesos que

representam a sinapse. As redes neurais são consideradas sistemas de processamento distribuído e paralelo porque o processamento é realizado entre os elementos processadores de maneira distribuída, e cada um deles realiza esse processamento isolada e paralelamente. As redes neurais artificiais possuem inspiração biológica, apresentando algumas características semelhantes com o cérebro humano, como [BARRETO 97]:

- **Aprendizagem:** estas redes podem modificar seu comportamento em resposta a eventos ou fatos que ocorreram no meio externo e que fornecem um conjunto de entradas, às quais pode ser associado um conjunto de saídas desejadas. Através de um algoritmo de treinamento, este conjunto provoca um auto-ajuste da rede para produzir um conjunto de respostas adequado e consistente com os padrões de entrada.
- **Generalização:** após o aprendizado de um conceito, a rede é capaz de funcionar com conceitos similares, que ainda não foram utilizados diretamente no treinamento.

Unindo-se vários neurônios artificiais podemos formar uma rede neural artificial que consegue reconhecer uma quantidade maior e mais complexa de padrões. De fato, o poder computacional da rede neural artificial advém das conexões entre os neurônios, ou seja, das sinapses formadas.

3.7.5 - Redes Bayesianas (RB)

São um tipo específico de redes de conhecimento. As redes de conhecimento são estruturas gráficas que representam sistemas baseados em conhecimento embasadas em vários tipos de teorias. Quando são formalizadas com base na teoria da probabilidade, são chamadas RB, Causais ou Diagramas de Influência² [PEARL 88]. Já quando sua formalização tem base na teoria de Dempster-Shafer, então fala-se de Galerias, Redes Qualitativas de Markov ou Redes de Construção [PEARL 88].

O primeiro pesquisador a idealizar a utilização da representação gráfica das informações probabilísticas foi o geneticista Sewall Wright [WRIGHT 21], mas foi muito criticado e seu trabalho só foi reconhecido pela comunidade estatística a partir da década de 50 [PEARL 88].

RB são esquemas de representação de conhecimentos utilizadas para desenvolver a base de

conhecimentos de um SEP. Elas têm uma parte estrutural refletindo relações causais e valores de probabilidade refletindo a força da relação. A RB é extraída de um especialista que transfere seus conhecimentos do domínio de aplicação para o engenheiro de conhecimentos do sistema. Na construção de uma RB, nem sempre pode-se identificar numericamente as dependências entre as variáveis. Na maioria das vezes esta tarefa é feita através do julgamento humano. Para se designar que uma variável B será descendente direto (filho) de uma variável A , deve-se constatar que A é causa direta ou então influencia a variável B diretamente [PEARL 88].

Na tentativa de se representar o conhecimento humano a motivação para a utilização de RB, é que os seres humanos têm uma certa necessidade de moldar os fatos e fenômenos em forma de relacionamentos causais. Sempre que um fato está sendo analisado, busca-se uma explicação que seja a causa do fato em questão, mesmo que esta causa seja imaginária e que não se tenha provas concretas [HECKERMAN 95].

Esta busca por uma relação causa/conseqüência se dá pela facilidade que se tem em representar, questionar e atualizar este tipo de estrutura. Por exemplo, se houvesse uma pesquisa que perguntasse às pessoas que passam pela rua: “que dia é hoje?”. Sem dúvida a resposta da maioria dos entrevistados seria a mesma. Isto porque a maioria das pessoas segue um mesmo calendário e isso causa o fato das respostas serem aproximadamente iguais. Sendo assim, se deseja saber qual será a resposta do próximo entrevistado. Não é necessário observar todas as respostas anteriores, basta observar a causa, ou seja, o fato do calendário utilizado ser referência de uma grande parte da população mundial [HECKERMAN 95].

Num modelo causal não há relações com complexidade muito grande. A estrutura de cálculos terá uma complexidade proporcional ao número de variáveis dependentes e não ao número de variáveis do problema [HECKERMAN 95].

A representação numérica da teoria da probabilidade é que provê consistência às redes de conhecimento Bayesianas. Por si, a teoria da probabilidade é adequada para a prova de

² É uma extensão da representação de uma rede Bayesiana para o tratar problemas de decisão.

teoremas matemáticos, mas não é adequada para representar o raciocínio humano. Isto pelo fato de que os seres humanos, não trabalham com grande quantidade de proposições. Quando o conjunto de proposições se torna grande, então passa-se a ter dificuldade em tomar decisões. Por isso é importante uma representação gráfica onde se possa trabalhar localmente o conceito de dependência, assim estando bem mais próximo do raciocínio humano [HECKERMAN 95].

Apesar de não permitirem a representação de todas as características desejáveis, os grafos são as estruturas que melhor representam as relações de dependência ou modelo de dependências [PEARL 88]. Além das vantagens que a estrutura de uma RB oferece, esta especifica uma distribuição de probabilidade, e através de uma semântica clara, nos permite processar diagnóstico, aprendizado, explicações, além de outras tarefas necessárias em um sistema especialista.

Formalmente, as redes de conhecimento Bayesianas são grafos direcionados acíclicos nos quais os nós representam variáveis aleatórias com medidas de incerteza associadas. Os arcos significam a existência de uma influência causal direta entre as variáveis conectadas, e a força destas influências é quantificada por probabilidades condicionais [PEARL 88].

Ao se utilizar de arcos direcionados, pode-se representar as relações de causalidade, obtendo assim, uma correspondência muito grande entre a topologia do grafo e as relações de dependência (causalidade) nele contidas. Os grafos ainda nos possibilitam utilizar a rede como um mecanismo de inferência [HECKERMAN 95].

Considerando que a causa precede o efeito, é fundamental ter um processo unidirecional para modelar a causalidade. Isto é, se B causa A então B ocorre antes de A . No contexto da lógica clássica, a implicação não capta uma relação causal por problemas de falta de direcionalidade. Nela temos que $(B \rightarrow A)$ é equivalente a $(\neg B \rightarrow \neg A)$ assim não permite que a causalidade seja modelada.

Nos sistemas baseados em regras, as declarações SE-ENTÃO são interpretadas de forma unidirecional. As redes semânticas também incorporam direcionalidade pelo uso de grafos direcionados. O conceito de probabilidade condicional é sugestivo de causalidade, mas sem uma conexão direta. Dessa forma, nas RB são utilizadas estruturas de grafos direcionados no

espaço de eventos ϵ .

O formalismo de RB provê uma representação concisa de uma distribuição conjunta de probabilidades em um grupo de variáveis estatísticas; esta representação é denominada RB de crença ou, simplesmente, RB. A concisão na representação é alcançada através de separação explícita entre a informação sobre as independências entre as variáveis e os valores numéricos envolvidos. Para este fim, uma RB compreende duas partes: **quantitativa** e **qualitativa** [LINDA 96].

A parte **qualitativa** é um modelo gráfico (grafo direcionado acíclico) onde as variáveis são os nós e as regras, que são as relações de dependência condicional entre essas variáveis, são os arcos direcionados. Assim um arco ligando as variáveis A e B , na seguinte forma $A \rightarrow B$, indica que a variável B é a consequência e a variável A é a causa, e estas apresentam uma relação de dependência. Por outro lado, se não houver um arco ligando duas variáveis então supõe-se que essas variáveis são independentes. O conceito de independência de variáveis está explicado no Capítulo 3, item 3.3.9.

A parte **quantitativa** de uma RB é o conjunto de probabilidades condicionais associadas aos arcos existentes no modelo gráfico acima descrito e as probabilidades estimadas *a priori* das hipóteses diagnósticas. Associada à parte qualitativa de uma RB está um grupo de funções representando valores numéricos da distribuição em questão. A cada vértice do grafo é associada uma função de atribuição de probabilidades que é basicamente um conjunto de probabilidades condicionais que descreve a influência dos valores dos vértices predecessores nas probabilidades dos valores deste vértice. Estas funções de atribuição de probabilidade em conjunto constituem a parte **quantitativa** da RB [LINDA 96].

A seguir será demonstrado um exemplo de RB que representa o conhecimento médico relativo ao diagnóstico de edemas subcutâneos [TESSARI 98]. A RB é composta por seis variáveis binárias: a variável *Edemas_Subcutâneos* representa presença de Celulite ou Alergia, a variável *Prurido* representa a *presença* ou *ausência* de prurido, a variável *História_Alergia_Família* representa se o paciente teve casos de edemas na família ou não, a variável *Dor* representa a presença ou ausência de dor, a variável *Cor_Local* representa se a

cor do local é Roxa ou Rósea e a variável *Febre* representa se existe a presença ou não de febre. Neste exemplo, a variável *Edemas Subcutâneos* é modelada como tendo influência direta sobre a presença ou a ausência das outras variáveis.

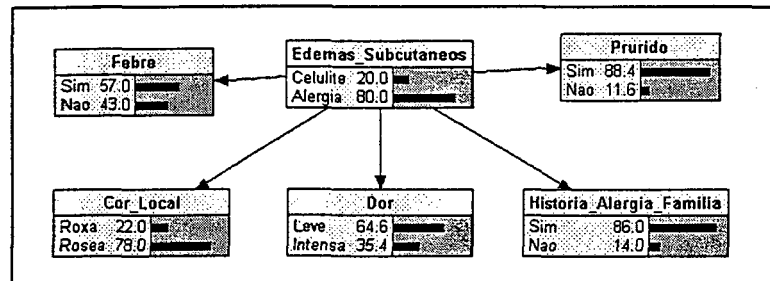


Figura 2. Um exemplo de RB [TESSARI 98]

Neste exemplo, as variáveis estão representadas pelos nós. A relação de dependência condicional entre as variáveis está representada pela seta (arco) (\longrightarrow) que liga um nó ao outro. As setas ligando a variável *Edemas Subcutâneos* às variáveis *Dor*, *Prurido*, *Febre*, *Cor Local* e *História Alergia Família*, significa que as variáveis *Dor*, *Prurido*, *Febre*, *Cor Local* e *História Alergia Família* são consequência da variável *Edemas Subcutâneos*. Assim, pode-se dizer: “*dado que uma pessoa tem Edemas Subcutâneos (causa), a probabilidade dessa pessoa ter Febre (consequência) será de 57%*”. Os nós (variáveis) e as setas (regras e/ou fatos) compreendem a parte **qualitativa** da RB.

O conjunto de probabilidades condicionais associado aos arcos e as probabilidades *a priori* das hipóteses diagnósticas, compreendem a parte **quantitativa** da RB.

Um usuário de um SEP especifica observações quando realiza uma consulta ao sistema, isto é ele instancia o sistema para aquele caso atual de interesse. Cada observação pode ser considerada como uma informação sobre o valor de uma variável aleatória. e geralmente, o usuário está interessado em conhecer valores de probabilidade sobre uma outra variável aleatória e/ou de uma hipótese diagnóstica. Estes valores de probabilidade são estimados pelo SEP através da utilização da teoria da probabilidade e mais precisamente do teorema de Bayes.

Há dois tipos de cálculos realizados por uma RB: a atualização de crenças e a revisão de crenças. A atualização de crenças é o cálculo de probabilidades das variáveis aleatórias. Enquanto que a revisão de crença refere-se à obtenção das probabilidades das hipóteses diagnósticas e a identificação da hipótese diagnóstica com maior valor de probabilidade. Esta hipótese diagnóstica mais provável pode ser vista como a explicação mais satisfatória do conjunto de evidências em mãos [NASSAR 98].

Na realidade, a representação é a formalização do conhecimento de um sistema. Pode ser imaginada como um conjunto de convenções para se descreverem objetos, fatos e situações.

Um mesmo objeto pode ter diversas representações, mas a habilidade de um sistema resolver problemas para os quais foi projetado é altamente influenciada pela escolha da representação.

Nesta pesquisa, o conhecimento do especialista será representado com o uso de uma RB, simplesmente pelo fato de que esta representa e trata a incerteza presente no domínio de aplicação utilizando a probabilidade bayesiana.

3.8 - Explicação do Raciocínio de um Sistema Especialista

Para um SE ser uma ferramenta eficaz, as pessoas têm de ser capazes de interagir com ele facilmente. Para facilitar essa interação, o SE deve ter dois recursos [RICH 88]: **a) Explicar seu raciocínio.** Em muitos domínios nos quais os SE abrangem, as pessoas necessitam de uma explicação de como o SE produziu tais respostas e **b) Adquirir conhecimento novo e modificar conhecimento antigo.** Pode-se chamar esta etapa de “Aprendizagem de regras”, isto é, o SE precisa “pensar” com os novos dados e alterar a base de conhecimentos, se for necessário.

Como já foi citado anteriormente, a exploração do Raciocínio Probabilístico em SE é o principal objetivo desta pesquisa, por isso não será desenvolvido um módulo de explicação do raciocínio da aplicação, e sim, será testado o comportamento do sistema em função do tamanho da sua base de conhecimento e será analisada a sensibilidade do sistema às mudanças nos valores das probabilidades das hipóteses diagnósticas.

3.9 - Considerações Finais

Pode-se constatar através deste estudo bibliográfico que a utilização da abordagem probabilística em SE para realizar o tratamento da incerteza, tem uma grande vantagem sobre as outras técnicas de representação do conhecimento incerto. Esta vantagem é a utilização de bases de conhecimento embasadas na Teoria da Probabilidade, as chamadas Redes Bayesianas.

Outra atração do uso da probabilidade e do Teorema de Bayes é que esta teoria é não-monotônica, isto é, ela suporta conflito de evidências e pode reduzir o valor da crença, ao contrário da lógica tradicional. Por exemplo, quando se tem uma probabilidade *a priori* de um paciente estar com febre, observa-se o seguinte: sabendo-se que a febre pode ser derivada de uma gripe ou de uma inflamação no ouvido, o SEP terá que decidir qual das evidências é mais provável, sem descartar totalmente a menos provável.

Outra vantagem do uso das RB, é a possibilidade de se fazer uma análise de sensibilidade. Esta análise pode ser feita para avaliar quais os números necessitam uma excelente exatidão [CRITCHFIELD 86].

Nesta pesquisa, os conceitos da Teoria da Probabilidade e do Teorema de Bayes foram utilizados para que o sistema pudesse auxiliar o médico na tomada de decisão. O próximo capítulo apresenta o domínio de aplicação desta pesquisa: Diagnóstico de Desnutrição Energético Proteica em crianças de 0 a 2 anos de idade.

Capítulo 4 – Domínio de Aplicação

Nesta pesquisa, o domínio de aplicação é a área médica, com o objetivo de desenvolver um sistema especialista para apoiar no processo da Avaliação do Estado Nutricional em Crianças de 0 a 2 anos de Idade - SISPAN.

A seguir, será explicada a forma de como o médico processa todas as informações necessárias para fornecer um determinado diagnóstico, quais as suas principais dificuldades na tomada da decisão e quais as vantagens do uso de um SEP na prática clínica.

Logo após, será apresentado o domínio de aplicação desta pesquisa: a Desnutrição Energético Proteica (DEP).

4.1 - A Incerteza e a Prática Clínica - Dr. Hélio Komagata

Decidir é uma das atividades essenciais na Medicina. Inicialmente, as informações clínicas são frequentemente imperfeitas, subjetivas ou não específicas [DEGOULET 94]. O médico se defronta, em cada consulta, com cenários ricos em informações incompletas e duvidosas. Tanto as informações que o paciente espontaneamente fornece, ou dele são obtidas por testes, quanto as informações que foram adquiridas ao longo da formação profissional do médico, têm a característica básica de incerteza. A prática do dia-a-dia do médico é processar estas informações incertas para obter uma decisão sobre diagnóstico e conduta.

Sem dúvida, a tomada de decisão é a atividade mais elevada, e mais complexa, do raciocínio clínico. O processo de raciocínio médico é baseado: 1) no raciocínio dedutivo, 2) na aplicação de princípios fisiológicos, 3) no conhecimento obtido por outros profissionais e 4) no conhecimento médico obtido pela própria vivência e processos mentais, até inconscientes, de julgamento. Um processo realizado nestas condições carregará, frequentemente, algum grau de incerteza [SMITH 85].

O raciocínio médico, em grande parte das situações da prática clínica, e em face às citadas situações de incerteza, pode ser caracterizado como essencialmente probabilístico [KOMAGATA 97]. O processo de raciocínio probabilístico tem sido estudado por diversas abordagens. Psicólogos, estudaram o caminho dos julgamentos subjetivos de probabilidades, definido como heurística. Uma das principais características do ato médico, isto é, do processo de analisar sinais e sintomas, identificar um diagnóstico e eleger um tratamento, é a tomada de decisão. Um outro caminho, para o estudo do raciocínio probabilístico, é o da estimação objetiva de probabilidades, como uma disciplina da ciência matemática. Embora o entendimento da heurística seja importante para o processo de decisão médica, o caminho da estimação matemática de probabilidades tem sido o que mais efetivamente tem trazido contribuições ao processo de raciocínio médico probabilístico.

Um exemplo de estimação objetiva de probabilidades muito difundida e utilizada pelos médicos, é o valor da prevalência de uma doença na população ou em sub-grupos de população. Prevalência é a frequência de um evento em uma população. Ela é usada frequentemente, pelos médicos, como estimador da probabilidade de um paciente de uma certa população, possuir uma doença ou ter um prognóstico desfavorável. A literatura médica é usualmente a fornecedora deste tipo de dado. Outro estimador de probabilidade muito utilizado, é a frequência com que um evento ocorre em um grupo com determinada característica clínica. Por exemplo, a evolução para morte perinatal dos conceptos de gestantes com eclâmpsia [KOMAGATA 97].

Este tipo de estimação é provavelmente uma das quantificações mais utilizadas na prática clínica, tendo grande impacto nas decisões diagnósticas e de conduta. Sua maior limitação é de fornecer uma estimação parcial, baseada apenas em uma única variável. Uma segunda limitação aparece, em geral, deixando órfão o médico, quando é necessária uma síntese das estimações obtidas por diversas variáveis. Sistemas de prognósticos baseados em fatores de risco, também, são formas muito difundidas de estimação de probabilidades. Estes sistemas são baseados em estudos prévios, onde se obtêm uma relação de fatores de risco que estão associados a um diagnóstico, ou mais frequentemente, a um prognóstico. Na prática, estes sistemas são utilizados calculando-se a somatória de pontos atribuídos aos fatores de risco

presentes no paciente. O valor total obtido é usado para enquadrar o caso em determinado grupo de risco.

4.2 – Uso Clínico de Modelos Estatísticos

Com o surgimento do computador, e sua propagação pela sociedade, a partir da década de 50, cresceu paralelamente a expectativa de se encontrar aplicações para esta nova tecnologia no processo de tomada de decisões médicas [SHORTLIFFE 90]. Sem dúvida, a importância da aplicação do computador em medicina, é no auxílio ao médico na avaliação diagnóstica, prognóstica e terapêutica do paciente, embora aplicações em outras áreas, como a hospitalar ou processamento de sinais, estejam hoje avançadas e contribuindo importantemente na assistência médica.

Os primeiros sistemas desenvolvidos com a finalidade de assistir ao médico no processo diagnóstico surgiram no final da década de 50.

O papel do microcomputador no desenvolvimento de sistemas baseados em modelos estatísticos foi fundamental, permitindo o processamento de estruturas de dados mais complexas, e a aplicação destes sistemas mais próxima do leito do paciente, por exemplo: MammoNet, Diaval, Dombal.

4.3 - Desnutrição Energético Proteica

A seguir será descrita a área médica de aplicação desta pesquisa [REIS 97].

4.3.1 - Conceitos

O Estado Nutricional é a condição de saúde de um indivíduo, influenciada pelo consumo e qualidade de nutrientes.

Avaliação Nutricional é o mecanismo para identificar o Estado Nutricional de uma pessoa e apresenta um papel decisivo para o diagnóstico da desnutrição energético protéica. É realizada através da adequada interpretação dos dados que descrevem a ingestão e utilização dos nutrientes. A obtenção de informações é realizada através de estudos clínicos,

antropométricos, dietéticos e bioquímicos.

Tabela 3. Diagnóstico do Estado Nutricional

| Eutrofia ³ | | Distrofia ⁴ | |
|-------------------------|------------------|------------------------|-------------|
| Por Deficit | Diagnóstico | Por excesso | Diagnóstico |
| de calorias e proteínas | desnutrição | de calorias | obesidade |
| de ferro | anemia | | |
| de vitamina D | raquitismo | | |
| de vitamina C | escorbuto | | |
| de vitamina A | hipovitaminose A | | |
| de água e eletrólitos | desidratação | | |

De acordo com a OMS/FAO⁵, desnutrição energético protéica é “uma gama de condições patológicas com deficiência simultânea de proteínas e calorias, em variadas proporções que acomete preferencialmente crianças de baixa idade e comumente associadas com infecções”.

A DEP é uma doença que ocasiona um retardo na velocidade de crescimento e desenvolvimento. Este retardo pode ser desaceleração (casos leves), interrupção (casos moderados) ou involução (casos graves).

Segundo a gravidade a DEP é classificada em três graus, de acordo com uma relação entre déficit de peso e peso normal esperado para a idade, estatura e o sexo. Os desnutridos de terceiro grau são sub-divididos clinicamente em tipos: marasmo, marasmo-kwashiokor e kwashiokor.

De acordo com a etiologia do processo, os desnutridos são classificados em primários, secundários e misto.

Em alguns casos de DEP ainda deve ser considerada a presença de efeitos tardios, sendo que os principais são o nanismo e a deficiência mental.

³ Quando o indivíduo (chamado eutrófico) apresenta um estado nutricional normal.

⁴ Quando o indivíduo (chamado distrófico) apresenta um estado nutricional alterado.

⁵ Organização Mundial de Saúde / Organização das Nações Unidas para a Agricultura e Alimentação.

Tabela 4. Classificação da DEP

| Etiologia | Gravidade |
|------------|--|
| Primária | Primeiro Grau (Leve) |
| Secundária | Segundo Grau (Moderada) |
| Mista | Terceiro Grau (Grave) Marasmo Kwashiokor Marasmo-Kwashiokor Presença de efeitos tardios Nanismo Deficiência Mental |

4.3.2 - Etiologia da Desnutrição Energético Proteica

A intensidade da desnutrição é baseada no déficit de peso em relação ao peso normal esperado para a idade cronológica (para crianças menores de 2 anos) e déficit de peso em relação ao peso normal para estatura (crianças maiores de 2 anos).

A DEP primária é quando há deficiência de nutrientes na dieta alimentar da criança, ocasionando insuficiência para abastecer o organismo.

Quando uma criança apresenta desnutrição secundária a fisiologia normal do organismo encontra-se alterada. Pode ser ocasionada por dificuldade de ingestão (vômitos, obstruções no tracto digestivo), digestão alterada (atividade das enzimas digestivas diminuída), má-absorção intestinal (doença celíaca, superfície intestinal diminuída, peristaltismo acelerado), excreção aumentada (hipotireoidismo ou diabetes mellitus), catabolismo exagerado (hipertireoidismo, carcinoma). A desnutrição mista é quando ocorre tanto a falta de uma dieta que satisfaça as necessidades da criança, bem como a alteração da fisiologia normal por algum motivo.

4.3.3 - Gravidade da Desnutrição Energético Proteica

A DEP se divide em três tipos quando considerada do tipo Grave. São eles:

Marasmo: é observada em regiões assoladas por catástrofes e em países subdesenvolvidos. Ocorre em lactentes que recebem uma dieta deficiente (principalmente falta de amamentação no peito - Síndrome da Privação Materna).

Crianças com desnutrição do tipo marasmo apresentam em seu quadro clínico falta de ganho de peso ponderal, seguida de perda de peso e redução do tecido muscular e adiposo. Como conseqüências temos a parada de crescimento e limitação da capacidade física. A criança apresenta aspecto físico senil; caracteriza-se por ser pequena para a idade; com baixa atividade; membros delgados; desaparecimento da bola de Bichat (bochecha); costelas proeminentes (bem visíveis); a pele se mostra solta e enrugada devido ao desaparecimento do tecido subcutâneo frouxo; encontra-se, em geral, irritada; ocorre hipotrofia dos músculos; a temperatura geralmente está abaixo do normal e o pulso pode ser lento; o metabolismo basal tende a diminuir; também pode apresentar um tipo de diarreia, com evacuações freqüentes, em pequena quantidade, contendo muco; o apetite é variável. Proteínas séricas, minerais, enzimas hepáticas são normais.

Kwashiokor: o kwashiokor é uma forma especial de desnutrição, causada pela falta de proteínas de valor biológico com ingestão de carboidratos em quantidade suficiente (arroz, milho), que é encontrada, em sua maioria, nas crianças da África, da América do Sul (principalmente no nordeste brasileiro), de algumas partes da Ásia e nos países subdesenvolvidos.

A palavra, originada da língua Ga de Ghana, quer dizer “doença do primeiro filho quando nasce o segundo”, isto porque a criança foi alimentada com o leite materno, tendo assim apresentado um crescimento satisfatório durante este período e após a parada (para alimentar o novo filho) ela começa a ter uma dieta deficiente.

As características do seu quadro clínico são: a perda de peso é parcialmente disfarçada pela presença de edemas (edemas por déficit protéico). As proteínas plasmáticas estão diminuídas (hipoproteinemias), principalmente a albumina (hipoalbuminemia); há evidência de hepatomegalia (o fígado está aumentado) e apresenta degeneração gordurosa disseminada. Existem ainda sérios distúrbios de desenvolvimento, parada de crescimento e uma resistência diminuída à infecção. Ocorre alterações de pele (lesões hipocrômicas alternadas com hiperocrômicas) dos membros inferiores, alterações de cabelos (facilidade de se soltar do couro cabeludo, coloração e textura), face de lua cheia, anasarca e concentrações séricas de proteínas

e albumina baixas. Déficit importante de estatura e massa muscular seriamente consumida; tecido celular subcutâneo conservado. Aspecto físico de miséria extrema, apatia - dificilmente responde a estímulos.

Marasmo-Kwashiokor: esta categoria intermediária se origina de formas graves de desnutrição seca ou com edema que sofreram deficiências importantes de proteínas e de hidratos de carbono e lipídios. A forma seca pode se transformar na edematosa e vice-versa em certas condições (aumento de perdas protéicas fecais, doenças infecciosas, etc.).

Tabela 5. Principais características do Marasmo e do Kwashiokor

| Achados clínicos laboratoriais | Marasmo | Kwashiokor |
|---------------------------------|-----------|-------------|
| Alterações de crescimento (P,E) | +++ | + |
| Atrofia muscular | +++ | +++ |
| Gordura subcutânea | ausente | presente |
| Edema | ausente | presente |
| Dermatoses | raro | comum |
| Alterações de cabelo | + | +++ |
| Hepatomegalia | rara | freqüente |
| Retardo no DNPM | ++ | ++ |
| Atividade física | diminuída | muito baixa |
| Diarréia | +++ | +++ |
| Albumina sérica | normal | baixa |
| Água corporal | aumentada | muito baixa |
| K corporal | baixo | muito baixo |
| Anemia | comum | muito comum |

+ leve; ++ moderada; +++ grave

Para se fazer o diagnóstico da desnutrição energético protéica deve-se observar, anteriormente, como se dá o crescimento normal da criança e algumas considerações que podem levar a criança a um estado de desnutrição.

4.4 - Crescimento Normal de Crianças

Durante os 5 primeiros meses de vida o lactente tem um ganho ponderal de peso, em média, de 20g por dia e aproximadamente 15g por dia durante o resto do 1º ano. Isso implica, em geral, que uma criança aos 5 meses tenha seu peso de nascimento dobrado e com 1 ano triplicado.

O comprimento normal do bebê, durante o 1º ano aumenta em torno de 25 a 30cm . O tecido subcutâneo também sofre um aumento visível, e aos 9 meses de idade é quando atinge seu auge.

A fontanela anterior do recém-nascido geralmente apresenta diminuição em seu tamanho após os 6 meses, podendo estar fechada por inteiro entre os 9 e 18 meses de vida. Já a fontanela posterior está fechada com 4 meses, em geral.

A circunferência do crânio, que é de 34 a 35cm quando o bebê nasce, aumenta em média para 44cm no 6º mês e para 47 com 1 ano.

4.4.1 - Síndrome da Privação Materna

Desde do nascimento, as crianças se sentem mais confortáveis com os pais e familiares do que com estranhos. Se os pais não estabelecerem relações emocionais e sociais eficazes com seu bebê, não havendo uma relação recíproca entre eles poderá gerar a Síndrome da Privação Materna.

As satisfações da criança são geralmente obtidas, através de atividades orais e contatos corpo-a-corpo da alimentação. Mães que não querem amamentar, podem quebrar o laço afetivo com seu filho, privando-lhe não só de um alimento completo, mas também de carinho, amor e confiança. A incapacidade de atingir as metas de desenvolvimento durante os primeiros meses de vida pode constituir a raiz de desordens emocionais permanentes.

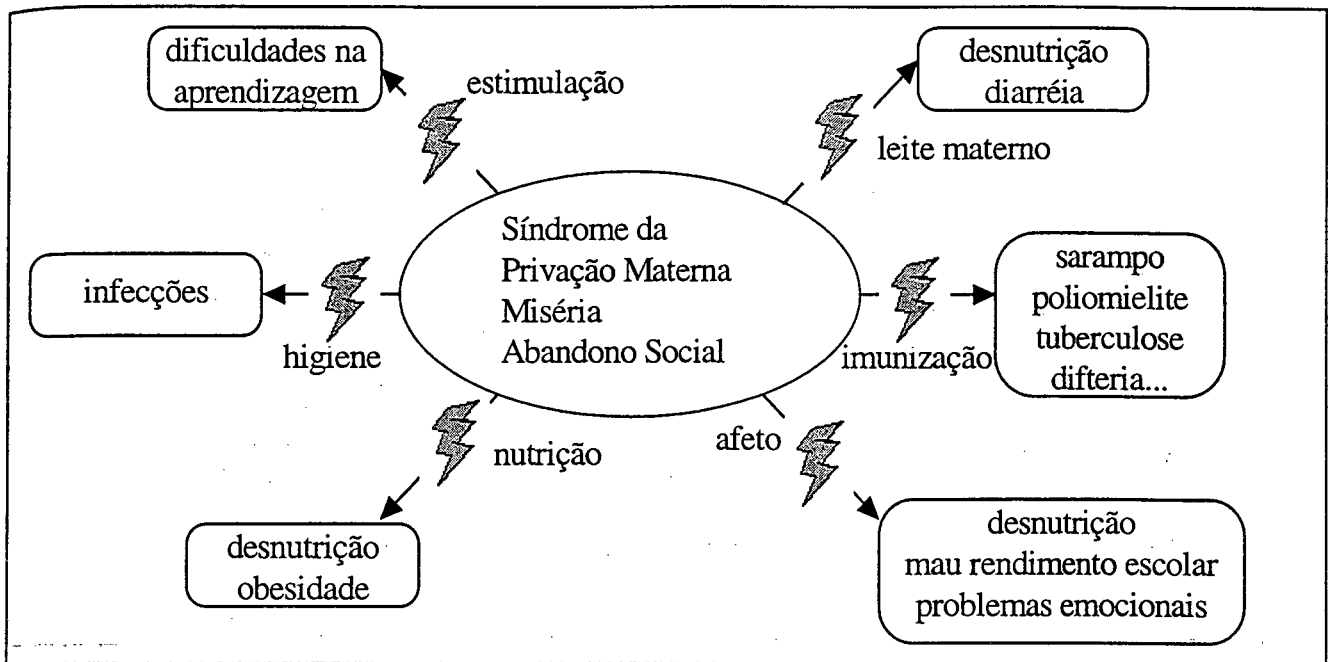


Figura 3. Síndrome da Privação Materna [REIS 97]

4.5 - Diagnóstico

Para se verificar a presença de DEP em uma criança, o especialista (pediatra) deve levar em conta o Estado Nutricional do paciente e a partir daí seguir os seguintes passos:

4.5.1 – Anamnese:

É o processo de observar o histórico clínico do paciente. Deverão ser analisados dados como:

- a idade, o peso e a estatura da criança;
- histórico da doença atual (início, fatores predisponentes, fatores agravantes);
- antecedentes pessoais (gestação, parto e neonatal);
- antecedentes patológicos (doenças crônicas, doenças graves, infecções, DNPM, crescimento comprometido);
- antecedentes alimentares, que através de um “inquérito” verifica-se a introdução de cada alimento, dificuldades de aceitação e reações indesejáveis. É feito através do levantamento do dia anterior (recordatório), registrando-se todas as refeições que o paciente ingeriu durante as 24 horas anteriores ou analisando os hábitos alimentares, a frequência e quantidade de consumo de alimentos;
- condição sócio-econômica-cultural da família, que deve ser avaliada pelo médico

através de informações como renda familiar, educação dos pais e condições de habitação.

- exame clínico, que seria a avaliação do estado geral da criança, observando a vitalidade, estado de hidratação, condições hipotróficas da pele e anexos, aspectos dos cabelos, alterações esqueléticas.

A anamnese na DEP é um instrumento valioso para determinar a etiologia da desnutrição, sendo sobretudo útil na orientação das ações terapêuticas de recuperação nutricional. O recordatório alimentar não fornece um diagnóstico totalmente confiável, pois a quantificação do consumo alimentar e as necessidades nutricionais do indivíduo são procedimentos complexos e imprecisos, no entanto deve ser feito, por ser um instrumento a mais [MONTEIRO 84].

4.5.2 - Medidas Antropométricas

As medidas antropométricas são parâmetros objetivos da Avaliação Nutricional e servem como instrumento importante de seguimento do estado nutricional. As medidas necessárias são: peso, estatura, perímetro cefálico, perímetro torácico, espessura das pregas cutâneas tricipital e subescapular, circunferência do braço e análise no gráfico de crescimento.

4.5.3 - Testes Laboratoriais

É através de testes laboratoriais que se tem medidas específicas de nutrientes e são eles que fornecem uma confirmação objetiva das deficiências nutricionais suspeitadas pela clínica. Devem ser solicitados de acordo com as indicações e suspeitas clínicas e pode detectar doenças associadas (por exemplo, DEP com anemia, ou ainda DEP com carência de vitamina A). Os exames mais comuns são: hemograma, proteínas totais e frações e pré-albumina.

4.5.4 - Considerações

Depois de analisados os sintomas do paciente vários métodos são disponíveis para classificar a desnutrição quanto a duração, a intensidade, e tipo.

Para verificar se a DEP é aguda, subaguda ou crônica utilizamos o critério de Waterlow, tomando-se como parâmetros o peso para altura e a altura para idade. Este teste consiste em

colocar, num plano cartesiano, o peso para altura na abscissa e a altura para idade na ordenada; dividi-se o plano em 4 partes iguais e adota-se como pontos de corte 80% da mediana do peso/altura e 90% da mediana da altura/idade da curva internacional de referência do NCHS. Crianças com valores de peso/altura acima de 80% são consideradas as crianças com desnutrição grave, instalada de modo agudo. Crianças com valores acima de 90% da altura/idade são consideradas normais, abaixo disto são consideradas crianças desnutridas cronicamente, cuja característica clínica é a falta de crescimento. Esta classificação foi adotada pela Organização Mundial de Saúde.

Na DEP Aguda (15 dias) a criança tem uma perda aguda do peso sem perda da estatura. O processo é rápido e o organismo não tem tempo de se adaptar a perda desacelerando a estatura. Por exemplo, um quadro de meningite ou diarréico agudo.

Na DEP Crônica (em média 3 meses) a criança entra em falência do crescimento e inicia com perda de peso. A criança desacelera o peso. Se isso acontece gradativamente por uma doença crônica, por exemplo, anemia ou alimentação diluída, o organismo tenta compensar desacelerando a velocidade do crescimento estatural e por um bom período não se percebe até que não tendo mais como desacelerar a estatura há então o consumo da gordura e posteriormente da musculatura (DEP Grave). Na DEP SubAguda (1 ou 2 meses), por exemplo, se a criança estiver com diarréia crônica.

O critério de Gomez [GOMEZ 56], que está apresentado na Tabela 6, é utilizado para avaliar a intensidade da DEP. Neste critério é feito uma adequação do peso para a idade e deve ser utilizada em crianças menores de 2 anos, pois acima desta idade o peso varia mais em função da altura do que da idade. Este critério é muito utilizado, porém não é capaz de distinguir entre desnutrição aguda e crônica.

Quando a criança que apresente edema clínico estiver sendo classificada seguindo o critério de Gomez e seu déficit de peso for de II grau, deve-se admitir que a DEP é de III grau.

Tabela 6. Classificação de Gomez, 1956

| Peso/Idade % | Graus de desnutrição |
|--------------|--------------------------|
| 91 - 100 | Normal |
| 76 - 90 | Primeiro Grau ou Leve |
| 61 - 75 | Segundo Grau ou Moderada |
| < 60 | Terceiro Grau ou Grave |

Outro critério utilizado para classificar a intensidade da desnutrição é o de Jelliffe, que propõe a variação dos pesos para a idade em intervalos de 10% ao redor dos padrões de referência (curvas de Harvard), admitindo I, II, III e IV graus, quando a variação for, respectivamente, 81-90%, 71-80%, 61-70% e < 60%.

Ainda, para classificar a intensidade há o critério Kanawati [KANAWATI 70], que pode ser observado na Tabela 7, onde existe uma correlação entre índices de ganho ponderal (estado nutricional) e a relação das medidas do ponto médio da circunferência do braço e do perímetro cefálico.

Tabela 7. Classificação de Kanawati, 1970

| Relação CB/PC | Intensidade |
|---------------|-----------------|
| > 0,32 | normal |
| 0,25-0,31 | leve e moderada |
| < 0,25 | grave |

Depois de classificado a intensidade e verificado Desnutrição Energéteico Proteica de III grau é necessário diagnosticar o tipo para se aplicar um tratamento conveniente. Para verificar o tipo existe o critério de McLaren, que pode ser observado na Tabela 8 e o de Wellcome que pode ser visto na Tabela 9.

Tabela 8. Classificação de McLaren, 1967

| Sintomas | Pontuação |
|--------------------------------|---------------|
| Edema | 3 |
| Alteração de pele | 2 |
| Edema + alteração de pele | 6 |
| Alteração de cabelo | 1 |
| Concentração de albumina (g/l) | |
| > 3,5 | 0 |
| 3 - 3,4 | 1 |
| 2,5 - 2,9 | 2 |
| 2,0 - 2,4 | 3 |
| 1,5 - 1,9 | 4 |
| 1,0 - 1,4 | 5 |
| 0,5 - 0,9 | 6 |
| Classificação | |
| Marasmo | 0 a 3 pontos |
| Marasmo- Kwashiokor | 4 a 8 pontos |
| Kwashiokor | 9 a 15 pontos |

Tabela 9. Classificação de Wellcome, 1970

| Déficit de peso/idade % do padrão | Edema | |
|--------------------------------------|---------------------|-------------|
| | Presente | Ausente |
| 60-80 | Kwashiokor | Subnutrição |
| <60 | Marasmo- Kwashiokor | Marasmo |

A classificação a ser utilizada no sistema será através do Escore Z, da avaliação clínica (aspecto da pele, cabelos, trofismo, e antropometria) e da avaliação laboratorial de McLaren. Para compreendê-la melhor ser faz necessário conceituar a expressão *Escore_Z*. Segundo [CARAZZA 91], a desnutrição leve e moderada podem ser avaliadas quanto à intensidade utilizando-se as medidas antropométricas mais comuns e disponíveis, como peso e altura. Estas medidas podem ter a capacidade de discriminar indivíduos portadores de desnutrição. Mas, para isso, é necessário que se selecionem padrões normais de crescimento que sejam parâmetros de referência para a população a ser avaliada. O Departamento de Pediatria da Faculdade de Medicina da Universidade de São Paulo adota a curva de crescimento (curva NCHS) de crianças brasileiras de 0 a 2 anos de idade originada do estudo de Santo André (1971 e 1982).

Por exemplo, o **Escore Z do Peso** é calculado a partir da fórmula:

$$\text{Escore Z} = \frac{\text{Peso} - \text{Média Peso}}{\text{Desvio Padrão Peso}}$$

onde,

Peso = é o peso atual da criança;

Média Peso = é a média de peso para a idade.

Desvio Padrão Peso = é o desvio padrão de peso para aquela idade.

Nesta pesquisa estimou-se a curva de média e desvio padrão de peso e altura em função do sexo e idade da criança.

Quanto aos resultados, os valores são classificados da seguinte maneira:

Escore Z = 0, significa Peso e/ou Altura normal,

Escore Z1 é quando os resultados ocorrem entre 0 e -1, isto significa Peso e/ou Altura com deficiência Leve,

Escore Z2 é quando os resultados ocorrem entre -1 e -2, isto significa Peso e/ou Altura com deficiência Moderada, e

Escore Z3 é quando os resultados ocorrem entre -2 e -3, isto significa Peso e/ou Altura com deficiência Grave.

4.6 - Tratamento

Pode ser realizado, de acordo com o estado da criança, utilizando:

- antibióticos
- dietas especiais com uso de leites hidrolisados de proteínas ou isentos de lactose
- nutrição através de sonda enteral
- suporte metabólico através da nutrição parenteral periférica ou total
- cálculo adequado e monitorado diariamente dos requerimentos protéicos, lipídicos e carboidratos
- cálculo da relação calórico nitrogenada de acordo com a condição clínica de cada paciente
- controle da ingestão de sódio.

4.7 – Considerações Finais

Neste capítulo foram apresentados os seguintes tópicos:

- uma apresentação do domínio de aplicação desta pesquisa: a área Médica,
- um relato sobre a forma de como o médico processa todas as informações necessárias para fornecer um determinado diagnóstico, quais as suas principais dificuldades na tomada da decisão e quais as vantagens do uso de um SEP na prática clínica, e
- uma apresentação dos principais conceitos sobre o domínio de aplicação que é a Avaliação do Estado Nutricional em Crianças de 0 a 2 anos de idade.

O capítulo seguinte propõe o sistema SISPAN – Sistema Pediátrico para Avaliação Nutricional.

Capítulo 5 – Proposta de um Sistema Especialista para Avaliação do Estado Nutricional em Crianças de 0 a 2 anos de Idade

5.1 - Considerações Iniciais

A flexibilidade das RB para representar dependências probabilísticas e a eficiência das técnicas computacionais para realizar inferência fazem das RB uma poderosa ferramenta para resolver problemas envolvendo incerteza e representam uma promissora técnica para suporte de decisão clínica.

Conforme visto no Capítulo 2, diversos SEP já foram desenvolvidos para auxiliar em problemas de tomada de decisão tanto na medicina, quanto na engenharia.

Propõe-se neste Capítulo uma análise sobre a sensibilidade do sistema às mudanças nos valores das probabilidades das hipóteses diagnósticas e das probabilidades condicionais. Será realizado, também, um estudo sobre o comportamento do sistema em função do tamanho da base de conhecimento e será desenvolvida uma aplicação na área médica o SISPAN – Sistema Pediátrico para Avaliação Nutricional.

5.2 - Arquitetura Geral

A arquitetura desta aplicação médica proposta foi composta da seguinte forma:

- a base de conhecimento extraída do médico especialista na área de Nutrição,
- a interface gráfica para fazer a interação com o usuário,
- uma base de casos de consulta e
- o módulo de Help.

A seguir, tem-se uma descrição das partes que compõem esta arquitetura.

5.2.1 - Base de conhecimento

As variáveis de entrada do sistema são os sinais e sintomas do paciente relativos à desnutrição. A variável de saída é um vetor de probabilidades que indica quão provável é cada uma das hipóteses diagnósticas para o paciente em questão. Assim sendo, a base de conhecimento foi composta por variáveis, seus respectivos atributos e pela relação entre estas variáveis que foi estabelecida através da definição de regras e/ou fatos. Esta base de conhecimento foi representada por distribuições de probabilidades.

5.2.2 - Shell para SEP

Foi utilizada a *shell* Netica 1:06 para o desenvolvimento do sistema. Esta *shell* armazena a base de conhecimento probabilístico e permite a visualização gráfica da RB que representa o conhecimento incerto. Após a definição das variáveis e das regras e/ou fatos, a *shell* gera uma distribuição conjunta de probabilidades. Esta distribuição de probabilidades representa o conhecimento do especialista no domínio pesquisado.

5.2.3 - Interface Gráfica

A interface tem como objetivo principal, fazer a comunicação entre o usuário e o sistema SISPAN - Sistema Pediátrico para Avaliação Nutricional. Será através da interface que o usuário escolherá quais os sintomas do paciente e visualizará qual o diagnóstico mais provável considerando os sintomas para o diagnóstico. A interface será projetada utilizando-se a linguagem de programação CBuilder, versão 1.0, da Borland.

5.2.4 – Base de Casos de Consulta

No sistema SISPAN foi implementada uma base de casos com o objetivo de, posteriormente, possibilitar uma avaliação quantitativa do desempenho do sistema, por exemplo, estimando-se a taxa de acerto do sistema. A base de casos armazenará todas as consultas feitas ao sistema. No momento em que o usuário solicita o diagnóstico ao sistema, automaticamente, todas as informações sobre o paciente e o seu respectivo diagnóstico são armazenados no arquivo Diagnostico.DB. Este arquivo com extensão .DB é um arquivo do tipo Paradox 7.0.

5.2.5 – Módulo de Help do Sistema

Este módulo tem como objetivo auxiliar o usuário oferecendo explicações, em forma de tópicos, sobre o domínio pesquisado. Contém textos e figuras sobre Desnutrição Energético Protéica implementados com recursos de hipertexto.

5.3 - Recursos Necessários

Os recursos necessários para o desenvolvimento desta pesquisa foram: o médico especialista em Nutrição, hardware e software. O software inclui a *shell* para desenvolvimento de SEP e a linguagem de programação CBuilder, versão 1.0, da Borland, para projetar a interface.

5.4 - Considerações Finais

Neste capítulo foi apresentada a proposta de implementação do sistema SISPAN, onde foi descrita a sua arquitetura geral e os recursos necessários. Com a definição da arquitetura geral do sistema partiu-se para a aquisição do conhecimento com o especialista.

O próximo capítulo apresenta uma descrição do sistema SISPAN, os métodos que foram utilizados para o seu desenvolvimento e descreve alguns aspectos relativos à abordagem probabilística em SE.

Capítulo 6 – Implementação

Neste capítulo serão apresentadas todas as etapas realizadas para o desenvolvimento do sistema SISPAN. Serão feitos comentários sobre a linguagem de programação e sobre a *shell* utilizadas, será descrito como foi realizada a implementação do sistema, desde a aquisição do conhecimento, representação do conhecimento, até os últimos testes de validação da base de conhecimentos do sistema.

6.1 - Software de Suporte

6.1.1 – *Shell* para SEP

Para a construção do sistema, o primeiro passo foi decidir como seriam implementados os algoritmos para o raciocínio probabilístico. Na primeira etapa da pesquisa, foi realizado um estudo sobre a abordagem bayesiana apenas em *shells*. Como um dos objetivos propostos é o desenvolvimento de uma aplicação médica, foi necessário desenvolver um programa em forma de executável com a finalidade de representar a base de conhecimentos específica, e o mecanismo de inferência referente à base de conhecimentos. As *shells* em potencial para se utilizar são: o Spirit, o Hugin e o Netica. Para desenvolver a interface, poderia ser utilizada uma ferramenta como o Kappa-PC, que possui tradição em SE, ou utilizar linguagens de programação que apresentassem elementos visuais, como Visual Basic, Delphi ou CBuilder. Após feito uma análise, a escolha foi sobre as linguagens com elementos visuais, devido à maior flexibilidade na utilização do sistema que estas permitem .

6.1.1.1 - Método de Comunicação de Dados

Além de desenvolver a interface, foi necessário acoplar os algoritmos de raciocínio probabilístico à interface. Isto levou à pesquisa de métodos de comunicação entre programas.

Basicamente, eram possíveis dois métodos: DDE (*Dynamic Data Exchange*) e DLL (*Dynamic Link Library*). Existem outros métodos, mas estes foram os que mais se encaixavam com os objetivos desta pesquisa. O método DDE, que pode ser observado em [CECHINEL 98], não foi utilizado devido a dois problemas básicos:

- a necessidade de um programa auxiliar, que deve ser instalado na máquina onde o sistema será utilizado e as
- dificuldades de comunicação com o programa auxiliar, que entrega dados ao sistema, e necessidade de execução simultânea, podendo sobrecarregar máquinas com baixa capacidade de processamento.

Desta forma, analisou-se as DLL's. A seguir, são apresentadas as principais vantagens e desvantagens deste método observadas em [TESSARI 98].

Vantagens das DLL's:

- **menor influência no executável:** a DLL é compilada separadamente do arquivo .EXE principal, e *linkada* dinamicamente durante a execução. Por esta razão, não aumenta o tamanho do executável que a chama.
- **Facilidade de manutenção:** o código é encapsulado em seu próprio módulo (a DLL). Atualizar esta parte do código, independente do seu tamanho, é tão simples quanto enviar aos usuários a DLL modificada. Evidentemente, quando se utiliza DLL's, é necessário manter a coerência entre as versões da DLL e os executáveis.
- **Aceleração do tempo de desenvolvimento:** a construção total do executável não inclui a DLL fonte, uma vez que ela é compilada e *linkada* separadamente.
- **código compartilhado implica em revisões globais:** diversas aplicações podem utilizar o mesmo arquivo .DLL.

Desvantagens das DLL's:

- **outro arquivo a ser distribuído:** o processo de chamada deve estar preparado para encontrar o arquivo. Além disso, mais arquivos implicam em mais documentação.
- **maior probabilidade de erros de programação:** mais considerações e garantias são necessários na programação, haja visto que o módulo carregado deve se entender com

as convenções de chamada de DLL e prevenir choque de nomenclatura em C++, problema existente também quando utilizados arquivos .OBJ. Além disso, a locação da DLL precisa ser conhecida.

6.1.1.2 - Shell utilizada

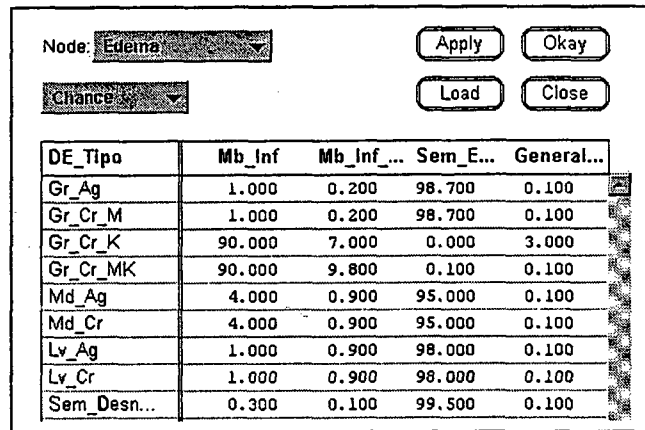
A opção de fornecimento de DLL está presente apenas no Hugin e no Netica, de forma que o Spirit, na versão disponível, mostrou-se inadequado para a aplicação proposta nesta pesquisa, que tem como um dos seus objetivos, o desenvolvimento de uma interface gráfica para o sistema.

Restava decidir entre o Hugin e o Netica, e a DLL adquirida foi a DLL do Netica, devido a fatores econômicos. A DLL oferecida pela Hugin Expert S/A apenas na sua versão profissional, que extrapolou os recursos financeiros do laboratório. Optou-se, então, pela DLL da Norsys Software Corp., que apresentou as características básicas necessárias à aplicação proposta nesta pesquisa e possui uma versão dentro dos recursos financeiros do laboratório.

O Netica 1.06 é composto por dois ambientes de trabalho: **1) Netica Application** e **2) Netica API**. O Netica Application é a interface gráfica padrão *Windows* onde a base de conhecimentos pode ser visualizada na forma de uma RB. O Netica API é a DLL (*Dynamic Link Library*) que contém todas as funções definidas na linguagem C para gerenciar a base de conhecimentos. Estas funções são para criar nós, adicionar *links*, compilar e gravar a rede, realizar as inferências na rede, dentre outras funções que podem ser observadas em [NETICA 97].

Nesta pesquisa, a base de conhecimentos foi construída através da Aquisição do Conhecimento com o médico especialista no domínio, utilizando-se o Netica Application. Nesta fase foi definida a parte qualitativa e a parte quantitativa da rede a partir da facilidade de criação dos nós e *links* da RB com o Netica Application. Após a definição dos nós e *links*, a RB foi salva com o nome de **Diagnostico.DNET**. Para acessar esta rede pré-gravada foi utilizada a linguagem Cbuilder 1.0 juntamente com a DLL do Netica, o Netica API.

Trabalhou-se, então, com os dois ambientes de trabalho da *shell* para SEP Netica: o **Netica Application** e o **Netica API**. O Netica Application possui uma interface gráfica para desenvolvimento, com exibição do grafo direcionado acíclico, que constitui a parte qualitativa de uma RB. A parte quantitativa da RB é apresentada na forma de tabelas. A Figura 4 ilustra a tabela de probabilidades condicionais da variável Edema.



| DE_Tipo | Mb_Inf | Mb_Inf_... | Sem_E... | General... |
|-------------|--------|------------|----------|------------|
| Gr_Ag | 1.000 | 0.200 | 98.700 | 0.100 |
| Gr_Cr_M | 1.000 | 0.200 | 98.700 | 0.100 |
| Gr_Cr_K | 90.000 | 7.000 | 0.000 | 3.000 |
| Gr_Cr_MK | 90.000 | 9.800 | 0.100 | 0.100 |
| Md_Ag | 4.000 | 0.900 | 95.000 | 0.100 |
| Md_Cr | 4.000 | 0.900 | 95.000 | 0.100 |
| Lv_Ag | 1.000 | 0.900 | 98.000 | 0.100 |
| Lv_Cr | 1.000 | 0.900 | 98.000 | 0.100 |
| Sem_Desn... | 0.300 | 0.100 | 99.500 | 0.100 |

Figura 4. Representação gráfica da Parte Quantitativa da variável Edema

A Netica API é uma biblioteca completa de funções em C, para trabalhar com RB e diagramas de influência. Essa DLL contém funções para construir, ensinar, modificar, transformar, salvar e ler redes, assim como uma poderosa máquina de inferência. Pode ser embutida em programas escritos em qualquer linguagem, desde que esta possa executar funções em C. Os programas que a utilizam controlam-na completamente. Por exemplo, as funções do Netica não executam nenhuma ação enquanto não forem chamadas, e não levam um intervalo de tempo imprevisível para retornar um resultado. Também podem ser utilizadas com qualquer outra biblioteca em C ou C++ e não necessitam de nenhuma outra biblioteca a não ser a biblioteca C Standard. Depois de decidido qual *shell* utilizar, foi necessário decidir, também, sobre qual seria a linguagem de programação mais adequada para o projeto da interface, pois esta linguagem de programação deverá se comunicar com os algoritmos que realizam o raciocínio probabilístico.

6.1.2 – Linguagem de Programação

A linguagem de programação utilizada para desenvolver a interface gráfica do sistema foi o

Cbuilder 1.0 da Borland. Utilizou-se esta linguagem pela sua facilidade de lidar com objetos visuais e devido à DLL Netica API ser toda programada em C++. Isto facilitou a comunicação da linguagem Cbuilder 1.0 com as funções do Netica API que possui os algoritmos que executam o raciocínio probabilístico.

6.2 – Implementação do Sistema

6.2.1 – Projeto da Interface Gráfica

A interface gráfica foi projetada em quatro partes principais como descrito abaixo:

- **Visualização das Probabilidades *a priori* da RB:** nesta primeira tela do sistema o usuário visualizará as probabilidades *a priori* da base de conhecimentos, como mostrado na Figura 5.

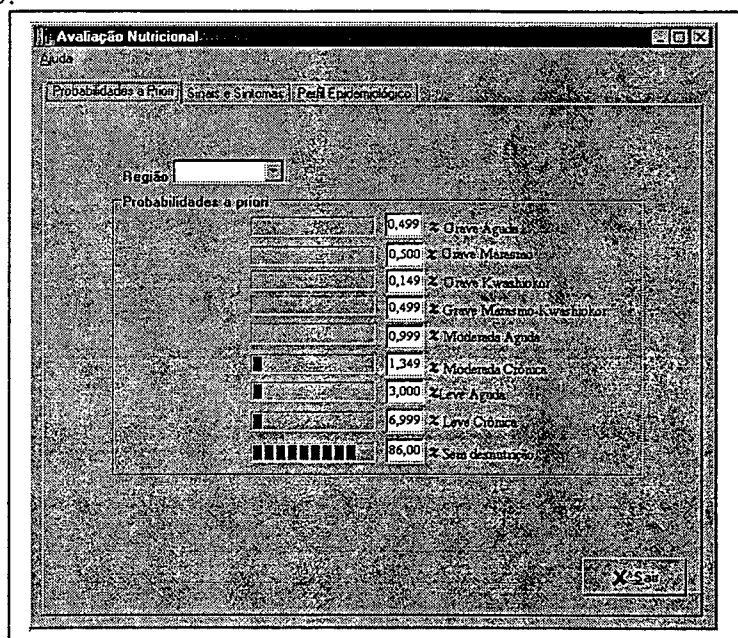


Figura 5. Primeira tela do SISPAN

- **Digitação dos Dados:** na segunda tela do sistema o especialista digitará os dados do paciente, entre eles o peso, a estatura e a idade e os sintomas observados pelo médico. Logo após, o usuário deverá clicar no botão Diagnóstico. Neste momento o sistema calcula as probabilidades a posteriori e exibe um vetor de probabilidades, informando ao usuário qual o diagnóstico mais provável. Esta tela pode ser visualizada na Figura 6.

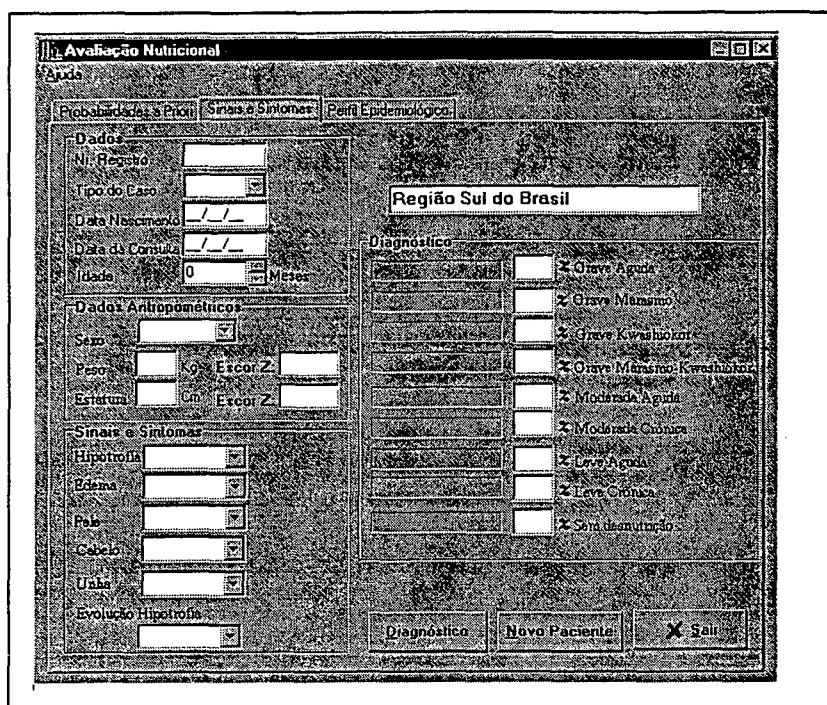


Figura 6. Segunda tela do SISPAN

- **Atualização do Perfil Epidemiológico:** nesta terceira tela, o usuário informará um novo vetor de probabilidades *a priori* para as probabilidades das hipóteses diagnósticas da RB. Pode surgir a seguinte questão: *Por que atualizar apenas as probabilidades das hipóteses diagnósticas e não atualizar as probabilidades condicionais ?*

Porque o conhecimento das relações causais entre diagnóstico e sinais e sintomas podem ser considerados estáticos, isto é, não se alteram. E o perfil epidemiológico da patologia na população não é o mesmo em todas as regiões do país ou do mundo, por isso a necessidade de atualizar a base de conhecimento. Esta atualização permite ao usuário personalizar o sistema para a região geográfica que trabalha. Esta flexibilidade pode levar a uma maior aceitação do sistema por parte do usuário.

- **Ajuda On-Line:** a Ajuda On-Line está disponível no Menu Ajuda que pode ser visualizado na Figura 7. Esta ajuda permite o usuário acessar textos explicativos sobre o domínio de aplicação em questão. Segundo o médico especialista, o processo de avaliação do estado nutricional possui algumas particularidades que o usuário pode ficar em dúvida frente a um diagnóstico realizado pelo sistema. Esta Ajuda irá auxiliar no sentido de explicar melhor a classificação da Desnutrição, sua etiologia, seus tipos e o seu quadro clínico e também, o porquê da necessidade da solicitação de determinados testes

laboratoriais como: Albumina e Hemoglobina.

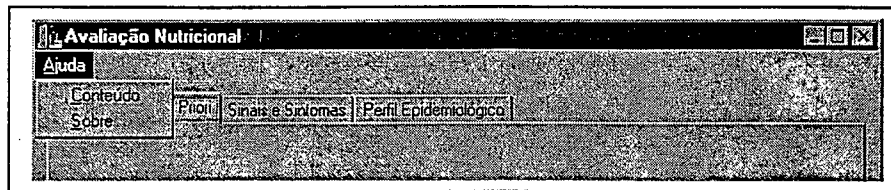


Figura 7. Menu Ajuda do sistema.

A seguir está demonstrada na Figura 8 a primeira tela da Ajuda disponível no sistema. Esta Ajuda é padrão Windows. Ela está toda organizada em tópicos e deve ser utilizada da mesma forma que qualquer Ajuda dos programas para Windows 95.

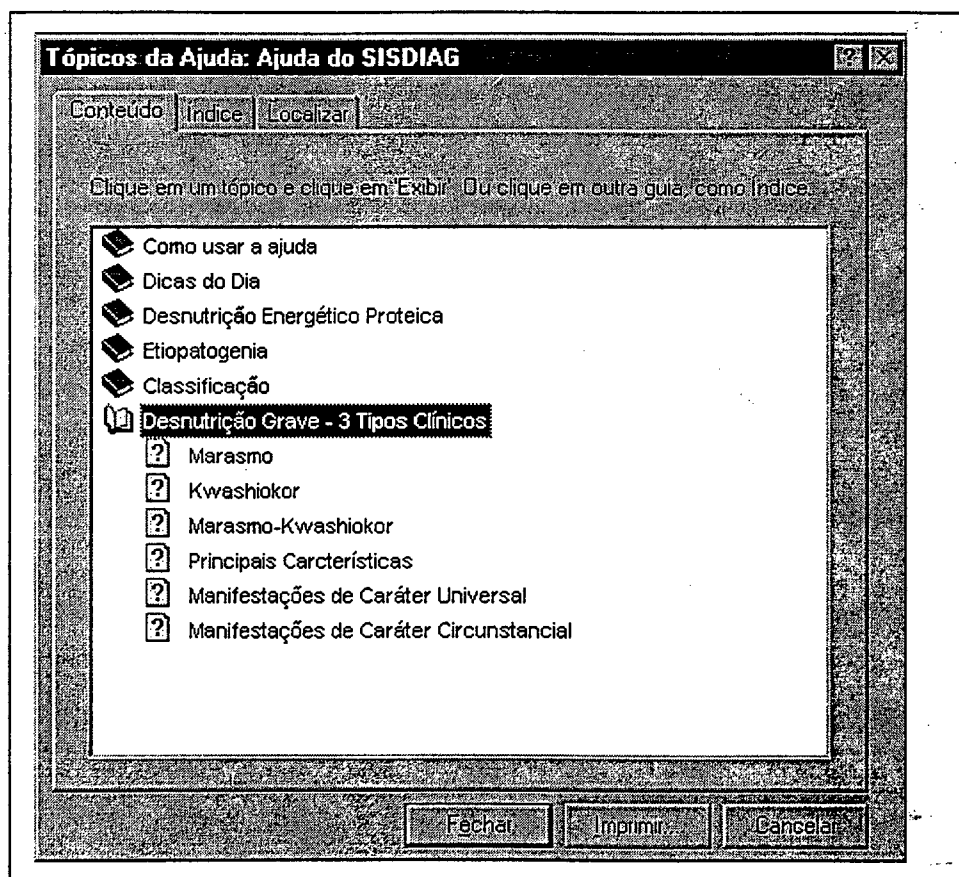


Figura 8. Ajuda On-Line do sistema em forma de tópicos

6.2.2 – Aquisição do Conhecimento

Para o desenvolvimento deste sistema, uma das primeiras etapas a ser realizada foi a Aquisição do Conhecimento com o especialista.

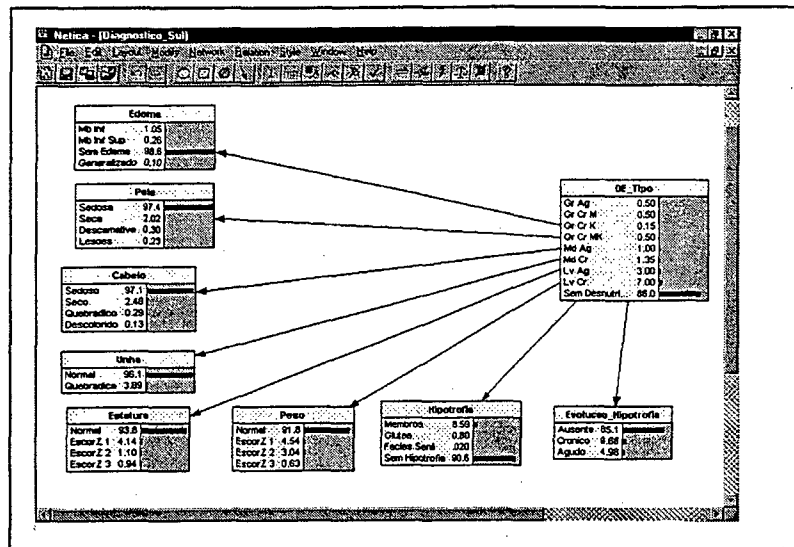


Figura 9. Representação gráfica da base de conhecimento (RB) para Avaliação do Estado Nutricional

Primeiramente, foi necessário definir qual o conhecimento que seria adquirido do especialista. Neste caso, foi constatado que seria necessário saber sobre quais eram as variáveis (sinais e sintomas) mais relevantes para o raciocínio clínico na tomada de decisão e quais os valores de probabilidades destas variáveis ocorrerem em determinadas situações.

Posteriormente, foi observado como estas informações seriam armazenadas na base de conhecimento. A maneira de armazenar estas informações na *shell* Netica foi na forma de tabelas como pode ser observado na Figura 11.

Para obter o conhecimento inerente ao domínio de aplicação foram realizadas inúmeras reuniões baseadas em entrevistas com o médico especialista. Tanto a parte qualitativa (sinais e sintomas) da RB, quanto a parte quantitativa (probabilidades) foram obtidas nestas entrevistas através de perguntas como : “Dado que uma criança tem desnutrição do tipo Grave Aguda, qual a probabilidade desta criança ter o Peso no Escore Z2 ? ”, e assim por diante, com todas as variáveis e seus respectivos atributos.

Para representar o conhecimento do especialista foi utilizada uma RB, como está ilustrado na Figura 9. A *shell* Netica possui uma interface gráfica interativa, onde foi possível criar os nós da RB que representam as variáveis, definir os arcos entre esses nós que são as dependências causais entre as variáveis e atribuir as probabilidades condicionais *a priori* para cada nó

(variável).

Esta RB é composta por nove nós, onde um representa o diagnóstico e os restantes representam os sinais e sintomas considerados. Uma vez que se tem a parte qualitativa da rede representada no modelo, pode-se construir a parte quantitativa da rede determinando as probabilidades condicionais *a priori* para cada variável. Esta etapa de determinação das partes Qualitativa e Quantitativa da rede compõe a fase de Aquisição de Conhecimento de um SE.

A parte Quantitativa da RB foi construída através de perguntas feitas ao médico especialista e digitada em forma de tabelas como mostrado na Figura 4.

6.2.3 - Inicialização da RB

Para a demonstração de como a *shell* Netica inicializou a rede foi necessário considerar o vetor das probabilidades *a priori* das hipóteses diagnósticas, que representa a variável DE_Tipo e os vetores das probabilidades condicionais *a priori* de cada uma das outras variáveis.

A seguir está apresentado o vetor das probabilidades *a priori* da variável DE_Tipo

| Gr_Ag | Gr_Cr_M | Gr_Cr_K | Gr_Cr_MK | Md_Ag | Md_Cr | Lv_Ag | Lv_Cr | Sem_D... |
|-------|---------|---------|----------|-------|-------|-------|-------|----------|
| 0.500 | 0.500 | 0.150 | 0.500 | 1.000 | 1.350 | 3.000 | 7.000 | 86.000 |

Figura 10. Tabela na *shell* Netica que representa o vetor das probabilidades *a priori* da variável DE_Tipo

onde,

| | |
|-----------------|--|
| Gr_Ag | significa Desnutrição do tipo Grave Aguda |
| Gr_Cr_M | significa Desnutrição do tipo Grave Crônica Marasmo |
| Gr_Cr_K | significa Desnutrição do tipo Grave Crônica Kwashiokor |
| Gr_Cr_MK | significa Desnutrição do tipo Grave Crônica Marasmo-Kwashiokor |
| Md_Ag | significa Desnutrição do tipo Moderada Aguda |
| Md_Cr | significa Desnutrição do tipo Moderada Crônica |
| Lv_Ag | significa Desnutrição do tipo Leve Aguda |
| Lv_Cr | significa Desnutrição do tipo Leve Crônica |
| Sem_Desnutrição | significa ausência de Desnutrição |

Os vetores das probabilidades condicionais *a priori* da variável **Peso** está apresentado a seguir:

| DE_Tipo | Normal | EscorZ_1 | EscorZ_2 | EscorZ_3 |
|-------------|--------|----------|----------|----------|
| Gr_Ag | 0.200 | 4.800 | 70.000 | 25.000 |
| Gr_Cr_M | 0.100 | 0.900 | 39.000 | 60.000 |
| Gr_Cr_K | 2.000 | 13.000 | 65.000 | 20.000 |
| Gr_Cr_MK | 0.500 | 1.500 | 63.000 | 35.000 |
| Md_Ag | 0.500 | 7.500 | 91.900 | 0.100 |
| Md_Cr | 1.000 | 20.000 | 78.900 | 0.100 |
| Lv_Ag | 89.900 | 10.000 | 0.100 | 0.000 |
| Lv_Cr | 93.000 | 6.900 | 0.100 | 0.000 |
| Sem_Desn... | 96.000 | 3.900 | 0.100 | 0.000 |

Figura 11. Representação gráfica da Parte Quantitativa da variável **Peso**

Após entrar com todas as probabilidades, a *shell* Netica gerou uma distribuição de probabilidades que pode ser visualizada na Figura 9.

A seguir será demonstrado como a *shell* calculou que 91,8% das crianças tem **Peso Normal** na base de conhecimentos do sistema (Figura 9).

$$\begin{aligned}
 P(\text{Peso Normal}) &= P(\text{Gr_Ag} \wedge \text{Peso Normal}) + \\
 &P(\text{Gr_Cr_M} \wedge \text{Peso Normal}) + \\
 &P(\text{Gr_Cr_K} \wedge \text{Peso Normal}) + \\
 &P(\text{Gr_Cr_MK} \wedge \text{Peso Normal}) + \\
 &P(\text{Md_Ag} \wedge \text{Peso Normal}) + \\
 &P(\text{Md_Cr} \wedge \text{Peso Normal}) + \\
 &P(\text{Lv_Ag} \wedge \text{Peso Normal}) + \\
 &P(\text{Lv_Cr} \wedge \text{Peso Normal}) + \\
 &P(\text{Sem_Desnutricao} \wedge \text{Peso Normal}) \\
 P(\text{Peso Normal}) &= P(\text{Gr_Ag}) \cdot P(\text{Peso Normal} \mid \text{Gr_Ag}) + \\
 &P(\text{Gr_Cr_M}) \cdot P(\text{Peso Normal} \mid \text{Gr_Cr_M}) + \\
 &P(\text{Gr_Cr_K}) \cdot P(\text{Peso Normal} \mid \text{Gr_Cr_K}) + \\
 &P(\text{Gr_Cr_MK}) \cdot P(\text{Peso Normal} \mid \text{Gr_Cr_MK}) + \\
 &P(\text{Md_Ag}) \cdot P(\text{Peso Normal} \mid \text{Md_Ag}) + \\
 &P(\text{Md_Cr}) \cdot P(\text{Peso Normal} \mid \text{Md_Cr}) + \\
 &P(\text{Lv_Ag}) \cdot P(\text{Peso Normal} \mid \text{Lv_Ag}) + \\
 &P(\text{Lv_Cr}) \cdot P(\text{Peso Normal} \mid \text{Lv_Cr}) + \\
 &P(\text{Sem_Desnutricao}) \cdot P(\text{Peso Normal} \mid \text{Sem_Desnutricao})
 \end{aligned}$$

$$P(\text{Peso Normal}) = 0,5 \cdot 0,2 + 0,5 \cdot 0,1 + 0,15 \cdot 2 + 0,5 \cdot 0,5 + 1 \cdot 0,5 + 1,35 \cdot 1 + 3 \cdot 89 + 7 \cdot 93 + 86 \cdot 96$$

$$P(\text{Peso Normal}) = 91,8$$

Pode-se notar que o cálculo das probabilidades condicionais *a priori* foi realizado para todas

as variáveis da rede. Aqui apenas foi demonstrado a probabilidade de crianças estarem com o Peso Normal.

Logo após, a aquisição do conhecimento da parte qualitativa e da parte quantitativa da rede, pode-se realizar as inferências na rede, isto é, informar os sinais e sintomas e obter-se como resposta o vetor de probabilidades *a posteriori* das hipóteses diagnósticas. Mais especificamente, a RB propagará a(s) evidência(s) e atualizará as probabilidades das hipóteses diagnósticas. A seguir está comentada uma destas consultas que podem ser realizadas na rede e como a *shell* realiza os cálculos.

6.2.4 – Consultas ao Sistema

Considera-se a evidência **Peso Escore_Z3**. A pergunta é: *Qual a probabilidade de uma criança estar com Desnutrição do tipo Grave Crônica Marasmo-Kwashiokor, dado que esta criança esta com uma perda de Peso no Escore_Z3 ?*

A seguir, a *shell* Netica mostra o vetor das probabilidades *a posteriori* dado a evidência de que uma criança tem o Peso no Escore Z3.

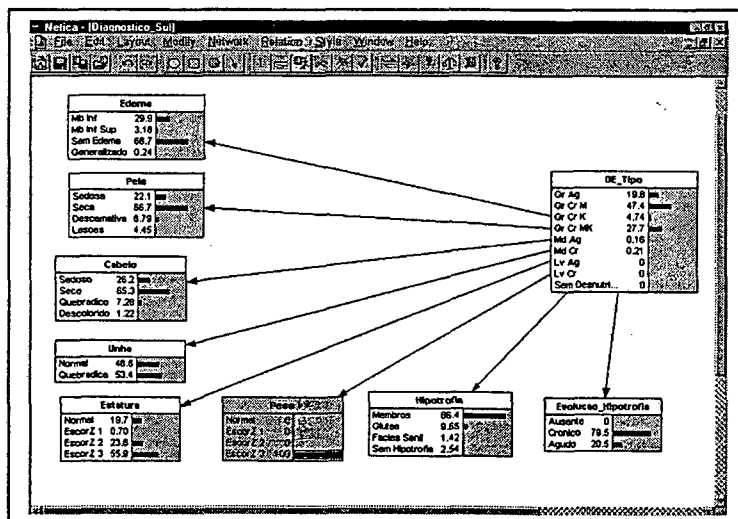


Figura 12. Vetor das probabilidades *a posteriori* após uma consulta à rede bayesiana

A seguir será demonstrado como a *shell* calculou que 27,7% das crianças tem **Desnutrição do tipo Grave Crônica Marasmo-Kwashiokor** dado que o **Peso é Escore_Z3**.

$$P(\text{Gr_Cr_MK} \mid \text{Peso_Escore_Z3}) = \frac{P(\text{Gr_Cr_MK}) \cdot P(\text{Peso_Escore_Z3} \mid \text{Gr_Cr_MK})}{P(\text{Peso_Escore_Z3})}$$
$$P(\text{Gr_Cr_MK} \mid \text{Peso_Escore_Z3}) = \frac{0,50 \cdot 35}{0,63}$$
$$P(\text{Gr_Cr_MK} \mid \text{Peso_Escore_Z3}) = 27,77\%$$

6.3 – Considerações Finais

Na etapa final de implementação do SISPAN verificou-se alguns aspectos no desenvolvimento que serão relatados a seguir.

- a) Observou-se a necessidade de se realizar a aquisição do conhecimento com o médico especialista por etapas, pois a parte quantitativa da RB faz com que o especialista repense todo o seu raciocínio e se realizada em apenas um dia, poderá confundir o especialista no momento de estipular um probabilidade condicional. Isto pode ser chamado de “Intoxicação de Números”, chega um certo momento em que o especialista precisa refazer todo o seu raciocínio para estimar a parte quantitativa da rede.
- b) Foi observado, a necessidade de se fazer um refinamento na base em paralelo com a aquisição do conhecimento, para testar a base de conhecimento.
- c) Em relação ao especialista, o desenvolvimento deste sistema levou-o a repensar o raciocínio clínico e a considerar uma forma melhor para repassar o seu conhecimento aos alunos, conforme relato verbal: *“o especialista, após longo período de atuação adquire um raciocínio clínico rápido baseado na sua experiência clínica diária. Esse raciocínio ainda que se utilize de um mínimo de sinais e sintomas clínicos, conclui normalmente um diagnóstico inicial correto, isso porque o especialista com sua experiência cria modelos de diagnóstico próprios – o chamado olho clínico. Ao trabalhar com sistemas especialistas, tornou-se necessário repassar esse olhar clínico de forma didática, e para isso foi necessário repensar cada passo dado antes do diagnóstico final. Esse processo levou o especialista a repensar quais seriam realmente os sinais e sintomas clínicos importantes na validação de um diagnóstico; quais os sinais e sintomas que reforçariam essa validação e quais que pouco ou nada interfeririam nessa validação. Desse processo*

dois pontos importantes se tornaram evidentes e contribuíram para reformular a postura do especialista como assistente, pesquisador e docente: 1) diante de uma patologia é importante repensar sempre o diagnóstico inicial e valorizar cada sinal e sintoma para evitar automatismos e erros decorrentes desse automatismo; se aproximar do paciente através da valorização desses sinais e sintomas reforçando a relação médico-paciente; 2) repassar o conhecimento com uma melhor metodologia dando um peso adequado a cada sinal e sintoma.

- d) Uma das principais vantagens do uso das RB para a representação do conhecimento incerto, foi a sua forma de transmitir graficamente o conhecimento do especialista, sendo esta representação uma ótima ferramenta de comunicação,
- e) quanto à utilização imediata, o SISPAN será disponibilizado no Hospital Infantil, no serviço de Metabologia e Nutrição como ferramenta de apoio ao ensino a residentes em Pediatria durante seu período de estágio nesse setor. O objetivo de iniciar o uso no próprio serviço é testar a funcionalidade do sistema. Posteriormente, será utilizado como ferramenta de apoio ao ensino e treinamento de Pediatras das unidades de saúde da rede básica municipal.

Capítulo 7 – Resultados

Este capítulo apresenta os resultados da Análise da Sensibilidade do sistema a eventuais alterações nos valores das probabilidades das hipóteses diagnósticas e das probabilidades condicionais, além dos comentários sobre a Análise do Comportamento do sistema em função do tamanho da base de conhecimentos e a Validação do Sistema.

7.1 – Análise de Sensibilidade

Foi realizada a análise de sensibilidade do sistema às mudanças nos valores das probabilidades da base de conhecimentos. Esta análise foi feita utilizando-se do teste *MANOVA – Análise Multivariada de Variância*, ao nível de significância 0,05.

Inicialmente, foram analisadas as seguintes questões:

- quais as variáveis a serem analisadas ?
- quais os valores de probabilidades que serão considerados ? os maiores ou menores valores ?
- qual o ruído a aplicar: um desconto e/ou um acréscimo sobre os valores de probabilidade ?
- quais as taxas de ruído ?

Primeiramente, foram realizadas várias reuniões com o médico especialista para decidir quais as variáveis que seriam analisadas. Foi decidido que seriam as seguintes variáveis: Edema, Hipotrofia, Peso e DE_Tipo. Esta decisão foi tomada considerando a afirmação do médico especialista de que estas variáveis são as mais importantes do sistema, isto é, são as variáveis de maior peso na hora de tomar uma decisão clínica.

Quanto à decisão sobre quais valores de probabilidades considerar, foi decidido que seriam considerados os maiores valores de probabilidades. Isto porque são estes valores que decidem o diagnóstico e nada mais lógico que fazer estas alterações nos valores de maior probabilidade.

Quanto ao ruído: para aplicar um desconto e/ou um acréscimo sobre esses maiores valores, foi decidido apenas descontar um determinado índice, isto é, enfraquecer a força da relação, porque se fosse acrescentado qualquer índice a um determinado valor de probabilidade, o acréscimo conduziria o diagnóstico para a certeza (100%), e assim, não conseguiríamos analisar a sensibilidade do sistema a estas mudanças. Dessa forma, decidiu-se diminuir os

Com estas consultas ao sistema, foi possível se fazer a seguinte análise:

- **Quanto às alterações nos valores das probabilidades das hipóteses diagnósticas:** submeteu-se os casos de teste à base inicial e às bases alteradas comparando-se o vetor de probabilidades *a posteriori* das hipóteses diagnósticas, percebeu-se que, mesmo alterando os valores das probabilidades *a priori* das hipóteses diagnósticas, não ocorreram mudanças nas probabilidades *a posteriori* das hipóteses diagnósticas, isto é, o sistema continuou informando o mesmo vetor de probabilidades. Com isso, pode-se afirmar que o Perfil Epidemiológico da região não influi nos resultados do vetor de probabilidade das hipóteses diagnósticas.
- **Quanto às alterações nos valores das probabilidades condicionais:** percebeu-se que quando consultados alguns casos de teste à base, o sistema se manteve estável, mas teve outros casos em que podem ser observadas alterações no diagnóstico.

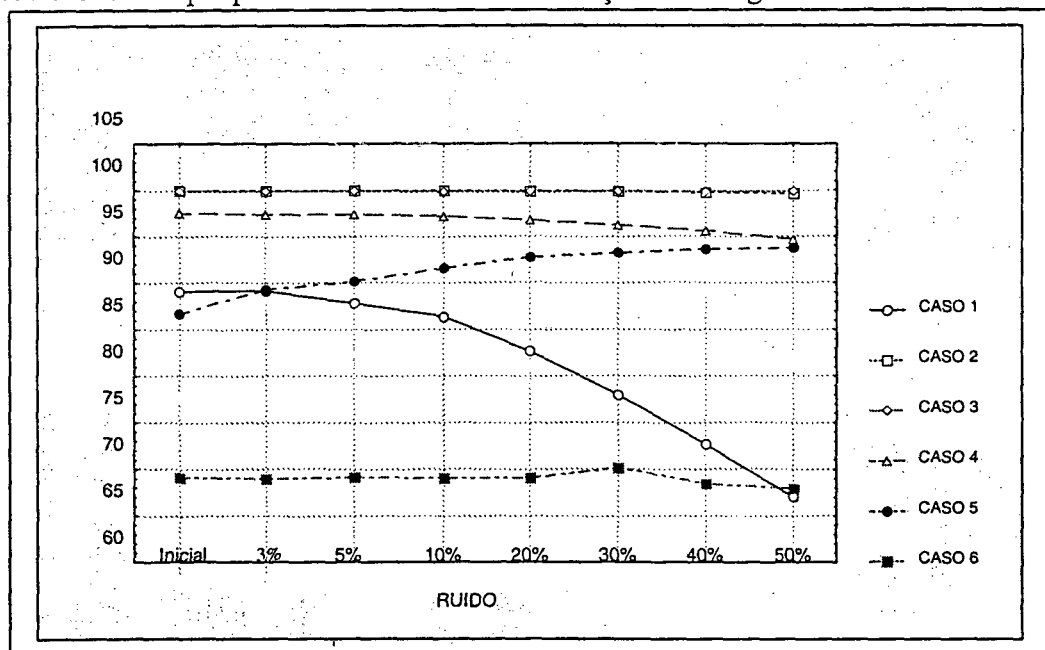


Figura 13. Análise da sensibilidade do sistema

Como pode ser visualizado na Figura 13, o sistema comportou-se da seguinte maneira de acordo com o tipo dos casos: **típico** e **atípicos**. Adotando uma probabilidade mínima de 80% como ponto de corte para selecionar uma hipótese diagnóstica, observou-se o seguinte:

- a) **Caso 1:** quando consultado na base de conhecimento com um ruído de 30%, pode-se observar que o diagnóstico se deteriorou. Segundo o médico especialista, pode-se acreditar

no diagnóstico a partir de 80%, então segundo o especialista, quanto ao caso 1, quando consultado na base com ruídos a partir de 30% não se pode confiar no diagnóstico porque houve uma queda muito significativa nos valores do vetor de probabilidades *a posteriori*.

b) **Caso 2 e Caso 3:** mantiveram-se estáveis.

Quanto aos casos atípicos observou-se o seguinte:

a) **Caso 4:** pode se observar que houve uma estabilidade no vetor de probabilidades *a posteriori*.

b) **Caso 5:** este caso teve uma confirmação dos resultados.

c) **Caso 6:** pode ser observado que este caso não tinha informação suficiente para o diagnóstico, pois já estava com pouca evidência desde a sua primeira consulta ao sistema.

7.2 – Análise do Comportamento do Sistema

Para se analisar o comportamento do sistema em função do tamanho da base de conhecimento, foram acrescentadas três variáveis ao sistema, são elas: **Potássio Sérico**, **Albumina Inferior a 3 g/dl** e **Hemoglobina Inferior a 10**. Estas variáveis podem ser vistas na Figura 14.

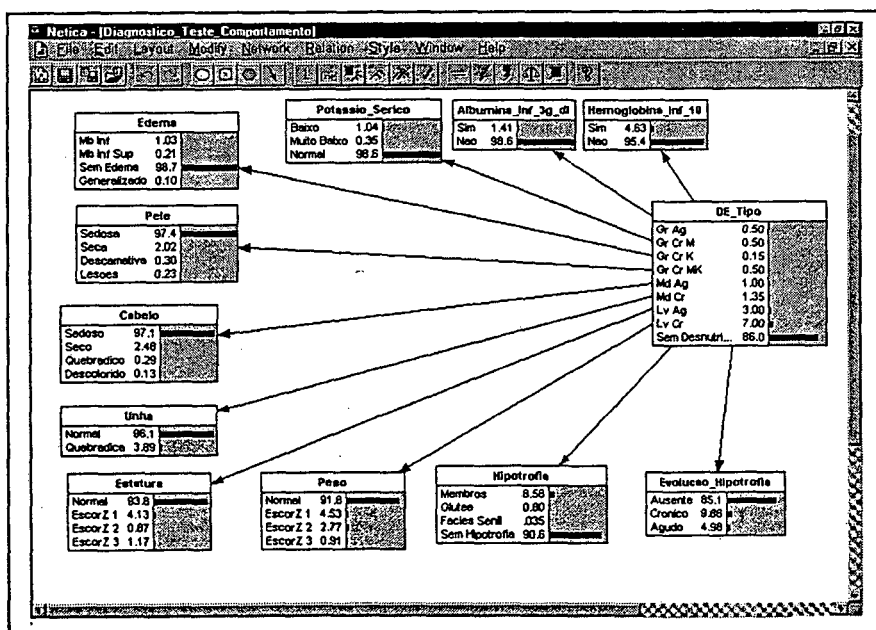


Figura 14. Base de conhecimento com o acréscimo de três variáveis

Após ter acrescentado as variáveis à rede, foram submetidas consultas tanto na rede com mais estas três variáveis quanto na rede sem o acréscimo das variáveis. Após estas consultas ao sistema, percebeu-se que a rede teve um comportamento estável, isto é, ela não alterou as probabilidades das hipóteses diagnósticas com o acréscimo de três variáveis na base de conhecimentos.

Nesta base conhecimento, as variáveis acrescentadas não trouxeram informações relevantes no processo de diagnóstico. Evidenciando que, neste caso, o especialista faz seu raciocínio clínico levando em consideração poucas variáveis, sendo estas, altamente relevantes na tomada da decisão.

7.3 – Comparação com Raciocínio Baseado em Casos

No projeto apresentado à banca como Trabalho Individual (TI), foi proposto realizar uma comparação com outra técnica: o Raciocínio Baseado em Casos. Entretanto, no decorrer do desenvolvimento desta pesquisa, o médico especialista julgou importante aumentar o escopo do sistema, isto é, acrescentar mais variáveis ao sistema, inicialmente desenvolvido por [REIS 97] considerando a técnica de Raciocínio Baseado em Casos. Este fato ocorreu pois a imersão do especialista nas técnicas de IA levou-o a considerar a possibilidade de tornar este sistema, também, uma ferramenta para apoio ao ensino médico. Desta forma, ficou inviável a comparação com RBC, partindo do pressuposto de que, com o acréscimo de mais variáveis no sistema atual, os dois sistemas possuíam conhecimentos bastante distintos, portanto, não sendo adequada sua comparação.

7.4 – Validação do Sistema

A validação qualitativa do sistema foi realizada com a ajuda de três médicos pediatras e experientes em Avaliação Nutricional, do Hospital Joana de Gusmão, de Florianópolis – SC.

A seguir estão relacionadas as observações feitas pelos três médicos pediatras:

- o sistema SISPAN tem conhecimento referente à avaliação nutricional,
- o sistema SISPAN fez diagnósticos válidos para os casos que foram apresentados pelos

médicos avaliadores,

- os textos explicativos são bons e didáticos,
- a interface do sistema é agradável,
- a possibilidade de personalizar o perfil epidemiológico foi julgada muito oportuna dando maior flexibilidade e credibilidade para usar o sistema,
- a base de conhecimentos, quando mostrada aos médicos, foi julgada muito interessante pois explicita o raciocínio clínico.

7.5 – Considerações Finais

Neste capítulo foram apresentados os Resultados desta pesquisa, entre eles:

- a) a Análise de Sensibilidade do sistema às mudanças no valores das probabilidades da base de conhecimento,
- b) a Análise do Comportamento do sistema em função do tamanho da base de conhecimento,
- c) a Validação Qualitativa do sistema.

O próximo capítulo, apresenta as Conclusões desta pesquisa, como também, sugestões de trabalhos futuros.

Capítulo 8 – Conclusões

Decidir é uma das atividades essenciais na Medicina. Inicialmente, as informações clínicas são freqüentemente imperfeitas, subjetivas ou não específicas. A área Médica está repleta de incerteza. Em geral, esta incerteza se refere a informações incompletas e inexatas fornecidas pelo próprio paciente ou por exames e testes laboratoriais.

A partir da década de 50, com o surgimento do computador, e sua propagação pela sociedade cresceu paralelamente a expectativa de se encontrar aplicações para esta nova tecnologia no processo de tomada de decisões médicas. Sem dúvida, a importância da aplicação do computador na Medicina, é no auxílio ao médico na avaliação diagnóstica, prognóstica e terapêutica do paciente.

Nesta pesquisa foi explorado o uso da Teoria da Probabilidade para a representação da incerteza em sistemas de apoio ao diagnóstico médico. Estes sistemas são chamados de Sistemas Especialistas Probabilísticos. Estas foram as conclusões ao longo desta pesquisa:

- a) a teoria da probabilidade para representar a incerteza é matematicamente bem definida, bem fundamentada e tem sido extensivamente estudada, por isso, da credibilidade em utilizá-la para o tratamento da incerteza em sistemas especialistas de apoio à decisão médica,
- b) a base de conhecimentos de um SEP contém valores de probabilidades das hipóteses diagnósticas e valores das probabilidades condicionais e é disponibilizada com uma interface gráfica (RB) para representar o conhecimento do especialista. Isto permite uma melhor comunicação entre o Engenheiro de Conhecimento e o Especialista,
- c) a base de conhecimentos de um SEP utilizando-se a shell Netica, tem que ser completa, isto é, todas as evidências relevantes às hipóteses consideradas devem estar explícitas nas base conhecimentos,

- d) as estimativas necessárias de probabilidade foram feitas pelo especialista a partir de seu conhecimento e experiência no domínio pesquisado. A obtenção das probabilidades das hipóteses diagnósticas H_i e as condicionais $P(H_i|e)$ foi considerada uma tarefa árdua,
- e) a RB ilustra o raciocínio dedutivo e abduutivo do especialista, sendo um grande potencial para apoiar o ensino, neste caso, na avaliação do estado nutricional em crianças de 0 a 2 anos de idade,
- f) dentro de um quadro de normalidade, se for informado ao sistema somente um grupo de evidências que não são tão relevantes para a tomada de decisão, como por exemplo, o aspecto da Pele, do Cabelo e da Unha do paciente, pode-se observar que o sistema não consegue decidir somente com estas evidências, similarmente ao especialista. No entanto, estas evidências juntamente com as outras evidências mais relevantes, reforçam o diagnóstico médico,
- g) o SEP “raciocina” com dados incompletos, isto é, não se faz necessário informar todas as evidências (variáveis de entrada) para se obter o vetor de probabilidades das hipóteses diagnósticas,
- h) um dado (sinal ou sintoma) que tem pouca probabilidade de ocorrer, quando ocorre traz um grande ganho de informação ao sistema, mudando o vetor das probabilidades das hipóteses diagnósticas. Isto é, quando se dá a entrada de um dado que já é esperado pelo sistema, o ganho de informação não é tão significativo para o sistema, quanto a entrada de uma informação que não está sendo esperada,
- i) quando feita a validação qualitativa do sistema, o mesmo foi capaz de fazer diagnósticos válidos para os casos que foram apresentados pelos médicos avaliadores e a sua base de conhecimentos foi julgada muito interessante, pois explicita o raciocínio clínico,
- j) quanto à análise da sensibilidade do sistema constatou-se que o Perfil Epidemiológico da região não influi no diagnóstico,

- k) quanto à análise da sensibilidade do sistema às mudanças nos valores das probabilidades condicionais, observou-se que: quando consultados alguns casos de teste à base, o sistema se manteve estável até o ruído de 30%,
- l) quanto à análise do comportamento do sistema em função do tamanho da base de conhecimentos, percebeu-se que as variáveis acrescentadas não trouxeram informações relevantes no processo de diagnóstico. Evidenciando que, neste caso, o especialista faz seu raciocínio clínico levando em consideração poucas variáveis, sendo estas, altamente relevantes na tomada da decisão, o que deixou o sistema consistente,
- m) O especialista refinou o seu raciocínio clínico, assim como sua forma de transmitir o conhecimento médico,

8.1 – Trabalhos Futuros

Como trabalhos futuros, foram constatadas as seguintes propostas:

- a) Utilizar técnicas de *Data Mining* como um suporte à aquisição do conhecimento.
- b) Buscar estruturas de redes bayesianas mais complexas que representem e tratem as relações de dependência entre as variáveis envolvidas.
- c) Realizar a validação quantitativa do sistema estimando a taxa de acertos do sistema.

Capítulo 9 - Referências Bibliográficas

[BANTER 97] **Banter**. [on line] Documento disponível na internet via WWW: <URL: <http://www.cs.uwm.edu/~haddawy/>>, fev-98.

✓ [BARRETO 97] **BARRETO, J.M. Inteligência Artificial. No limiar do século XXI**, Florianópolis, 1997.

[BARRETO 90] **BARRETO, J.M. A topological model for qualitative time**. Proceedings: IASTED Artificial Intelligence Application and Neural Networks, junho, 25-27, 1990, Zurich(Switzerland), p. 16-21.

[BAYES 91] **BAYES, T. An essay towards solving a problem in the doctrine of chances**. Reprinted in: MD Computing 8: 157 171, 1991.

[BNG 97]. **Bayesian Network Generator**. [on line] Documento disponível na internet via WWW: <URL: <http://www.mcw.edu/midas/bng.html>>, set-97.

[BOL 97] **Bayes-On-Line**. [on line] Documento disponível na internet via WWW: <URL: <http://www.kbe.co.za/products/bol.htm>>, nov-97.

[BUCHANAN 84] **BUCHANAN, B. G. , SHORTLIFFE, E.H. Rule-based Expert Systems: The Mycin Experiments of the Stanford Heuristic Programming Project**. Addison-Westley, Reading, MA, 1984.

[CARAZZA 91] **CARAZZA, Francisco R., MARCONDES, Eduardo. Nutrição Clínica em Pediatria**. Faculdade de Medicina da Universidade de São Paulo, Universidade de São Paulo, São Paulo, pg. 217, 1991.

✓ [CECHINEL 98] **CECHINEL, Cristian., AMORIM, Lidiane., NASSAR, Sílvia Modesto. SESTAT-Sistema Especialista de Apoio ao Ensino de Estatística**. Trabalho de Conclusão de Curso de Graduação em Ciência da Computação, Universidade Federal de Santa Catarina, 1998.

[CHARD 89] **CHARD, T. The effect of dependence on the performance of Bayes' theorem: an evaluation using a computer simulation**. Comput. Meth. Prog. Biomed., v.29, p.15-9, 1989.

[CHEESEMAN 85] **CHEESEMAN, P. In defense of probability**. Proceedings of 9th International Joint Conference on Artificial Intelligence. Los Angeles, pp 1002-1009. 1985.

- [CHEESEMAN 88] CHEESEMAN, P. **An Inquiry into Computer Understanding**. Computational Intelligence, 4, 58-66, 1988.
- [COHEN 85] COHEN, P. R. **Heuristic reasoning about uncertainty: artificial intelligence approach**. Boston: Pitman. 1985.
- [COLLAZOS 97] COLLAZOS, K. S. **Sistema Especialista Nebuloso para Diagnóstico Médico**. Dissertação de Mestrado, Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Santa Catarina, 1997.
- [CRITCHFIELD 86] CRITCHFIELD, G. C., WILLIARD K. E. **Probabilistic analysis of decision trees using Monte Carlo simulation**. Med Decis Making 6:85-92, 1986.
- [DEGOULET 94] DEGOULET, P. FIESCHI, M. **Informatique Médicale**. Paris: MASSON, 1994.
- [DÍEZ 93] DÍEZ, F. **Parameter adjustment in Bayes Networks**. The Generalized noisy-or gate. UAI-93, 99 – 105. 1993.
- [DÍEZ 97] DÍEZ, F. J., J. Mira, E. Iturralde and S. Zubillaga. **DIIVAL, a Bayesian expert system for echocardiography**. Artificial Intelligence in Medicine, 1997. [on line] Documento disponível na internet via WWW: <URL: <http://www.dia.uned.es/~fjdiez/tesis.html>> set-97.
- [DOMBAL 72] DOMBAL, F. T. , LEAPER, D. J. , STANILAND, J. R. , McCANN, A. P. **Human and Computer-aided Diagnosis of Abdominal Pain: Further Report with Emphasis on the Performance of the Clinicians**. British Medical Journal, 4, 376-380, 1972.
- [DXPLAIN 97] **Ruth Lilly Medical Library - Test Connection to DXplain**. [on line] Documento disponível na internet via WWW: <URL: <http://www.medlib.iupui.edu/dxplain.html>>, abr-97.
- [DXPRESS 97] **DXPRESS**. [on line] Documento disponível na internet via WWW: <URL: <http://foxtrot.rahul.net/kic/kic/dxpress.htm>>, jan-98.
- [EELLS 94] EELLS, E., SKYRMS, B. **Probability and Conditionals - Belief Revision and Rational Decision**. Cambridge Studies in Probability, Induction and Decision Theory, 1994.
- [EISENSTEIN 94] EISENSTEIN, E.L.; ALEMI, F. **An evaluation of factors influencing bayesian learning systems**. J. Am. Med. Informatics Assoc., v.1, p.272-284, 1994.

- [FOSSATTI 97] **Um sistema de aquisição de conhecimento na área médica.** [on line] Documento disponível na internet via WWW: <URL: http://www.upf.tche.br/trab_prof/lucimar/projeto.htm>, fev-98.
- [FRYBACK 78] FRYBACK, D. G. **Bayes Theorem and conditional nonindependence of data medical diagnosis.** *Comp Biomed Res* 11:423-434, 1978.
- [GAMMERMAN 92] GAMMERMAN, A.; THATCHER, A.R. **Bayesian diagnostic probabilities without assuming independence of symptoms.** *Meth. Inform. Med.*, v.30, p.15-22, 1992.
- [GIARRATANO 89] GIARRATANO, J. **Experts Systems Principles and Programing.** Boston: PWS-KENT Publishing Company, 1989.
- [GOMEZ 56] GOMEZ, F. et al., **Mortality in Second and Third Degree Melnutrion,** *J. Trop. Pediatr. Environ. Child Hlth*, 2: 77-83, 1956.
- [GRAPHICAL 97] **Graphical-Belief.** [on line] Documento disponível na internet via WWW:<URL:<http://bayes.stat.washington.edu/almond/gb/graphical-belief.html>>, dez-97.
- [HADDAWY 94] HADDAWY P. et al. **Preliminary Investigation of a Bayesian Network for Mammographic Diagnosis of Breast Cancer.** Department of Eletrical Engineering and Computer Science. University of Wisconsin-Milwaukee. Milwaukee, Wisconsin 53201, USA. 1994.
- [HADDAWY 94] HADDAWY P., JACOBSON, J., KAHN CE Jr. **Generating explanations and tutorial problems from Bayesian Networks.** In: Ozbolt JP, ed. *Proceedings of the 18th Annual Symposium on Computer Applications in Medical Care.* Philadelphia: Hanley & Belfus, 770-774. 1994.
- [HADDAWY 94] HADDAWY P., ROBERTS, Linda M., KAHN, Charles E. **Development of a Bayesian Network for Diagnosis of Breast Cancer.** Department of Eletrical Engineering and Computer Science. University of Wisconsin-Milwaukee. Milwaukee, Wisconsin 53201, USA. 1994.
- [HADDAWY 94] HADDAWY, P. **Generating Bayesian Networks from Probability Logic Knowledge Bases.** *Proceedings of the Tenht Conference on Uncertainty in Artificial Intelligence.* Seattle. 262 – 269. 1994.
- [HADDAWY 96] HADDAWY P. et al. **BANTER: A Bayesian Network Tutoring Shell.** Decision Systems and Artificial Intelligence Laboratory. Department of Eletrical Engineering and Computer Science. University of Wisconsin-Milwaukee. Milwaukee, Wisconsin 53201, USA. 1996.

- [HEATGFIELD 90] HEATGFIELD, H. A, et al. **Computer assisted diagnosis of fine needle aspirate.** Journal of Clinical Pathology, 43: 168 – 170. 1990.
- [HECKERMAN 95] HECKERMAN, David. **A Bayesian Approach to Learning Causal.** Technical Report MSR-TR-95-04, Microsoft Research, March, 1995. Documento disponível em: <http://www.research.microsoft.com/research/dtg/heckerma/TR-95-04.htm>
- [HRUSCHKA 95] HRUSCHKA, Estevan Rafael., SILVA, Wagner Teixeira. **Propagação de Crença e Aprendizado em Redes Bayesianas.** Relatório Técnico CIC/UnB. Documento disponível em <ftp://ftp.cic.unb.br/pub/publications/research/tr96-03.ps.gz>. 1995.
- [HUGIN 97] Hugin. [on line] Documento disponível na internet via WWW: <URL: <http://www.hugin.dk>>, set-97.
- [IDEAL 97] **Ideal - Influence Diagram Evaluation and Analysis in Lisp.** [on line] Documento disponível na internet via WWW: <URL: <http://rpal.rockwell.com/ideal.html>>, set-97.
- [IGNIZIO 91] IGNIZIO, J. P. **Introduction to Expert Systems - The Development and Implementation of Rule-Based Expert Systems.** McGraw-Hill, 1991.
- [KANAWATI 70] KANAWATI, AA & McLAREN, DS. **Assessment of marginal malnutrition.** Nature 228:573, 1970.
- [KIK 97] **About Knowledge Industries.** [on line] Documento disponível na internet via WWW: <URL: <http://www.kic.com>>, ago-97.
- [KOMAGATA 97]. **A incerteza e a prática clínica.** [on line] Documento disponível na internet via WWW: <URL <http://www.hcnet.usp.br/cia/forum/bayes.htm> > dez-97.
- [KRIPPENDORFF 86] KRIPPENDORFF, K. **Information Theory: Structural Models for Qualitative Data.** Sage Publications, 1986.
- [LAURITZEN 88] LAURITZEN, S. L. & SPIEGELHALTER, D. J. **Local computations with probabilities on graphical structures and their applications to expert systems.** J. Royal Statist. Soc., B, 50(2):154-227. 1988.
- [LEDLEY 59] LEDLEY, R.S.; LUSTED, L.B. **Reasoning foundations of medical diagnosis.** Science, v.130, p.9, 1959.
- [LINDA 96] LINDA, C. Van Der Gaag. **Bayesian Belief Networks: Odds and Ends.** The Computer Journal, vol.39, nr. 02, p. 97-113, 1996.

- [LINDLEY 82] LINDLEY, D. V. **Scoring rules and the inevitability of probability.** International Statistical Review, (50):1-26. 1982.
- [MACHADO 85] MACHADO, J.M. **Sistemas de apoio à decisão médica - diagnóstico diferencial e geração de planos de investigação.** Rio de Janeiro. Tese de Doutorado - Universidade Federal do Rio de Janeiro, COPPE, 1985.
- [MAMMO 98] **MammoNet - Mammography Decision Support System.** [on line]. Documento disponível na internet via WWW: <URL: <http://www.mcw.edu/midas/mammo.html>>, jan-98.
- [MKAT 97] **M-Kat.** [on line] Documento disponível na internet via WWW: <URL: <http://ipvaimed9.unipv.it/lab/mkat.html>>, fev-98.
- [MONTEIRO 84] MONTEIRO, C. **Recentes Mudanças Propostas na Avaliação Antropométrica do Estado Nutricional Infantil: Uma Avaliação Crítica.** Rev. Saúde Pública, São Paulo, 18: 56-63, 1984.
- [MSBN 97] **Microsoft Belief Networks - Tools for Bayesian Inference.** [on line]. Documento disponível na internet via WWW: <URL: <http://www.research.microsoft.com/research/dtg/msbn/>>, out-97.
- X [NASSAR 95] NASSAR, Sílvia Modesto. **Sistema Estatístico Inteligente para apoio a pesquisas médicas.** Tese de Doutorado, Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Santa Catarina, Florianópolis (Brasil), 1995.
- X [NASSAR 98] NASSAR, Sílvia Modesto. **Informática e Estatística: Uma Interação entre duas Ciências.** Trabalho submetido ao concurso de Professor Titular. Departamento de Informática e Estatística, Universidade Federal de Santa Catarina, Florianópolis (Brasil), 1998.
- [NEAPOLITAN 90] NEAPOLITAN, R. E. **Probabilistic Reasoning in Expert Systems: Theory and Algorithms.** Northeastern Illinois University. Chicago. 1990.
- X [NETICA 97] **Netica.** [on line] Documento disponível na internet via WWW: <URL: <http://www.norsys.com>>, out-97.
- [NI 79] NI, H. AIELLO, N. **AGE: A knowledge-based program for building knowledge-based programs.** Proceedings: 6th. IJCAI, 1979.
- [OHMANN 88] OHMANN, C., QIN, Y; KUNNEKE, M.; STOLTZING, H.; THON, K.; LORENZ, W. **Bayes' theorem and conditional dependence of symptoms: different models applied to data of upper gastrointestinal bleeding.** Meth. Inform. Med., v.27, p.73-83, 1988.

- [OWENS 90] OWENS, D.K.; SOX Jr., H.C. Medical decision making: probabilistic medical reasoning. In: SHORTLIFFE, E.H., 1990.
- [PEARL 86] PEARL, J. **Fusion, Propagation and Structuring in Belief Networks**. Artificial Intelligence, vol 29(3):241-288. 1986.
- [PEARL 88] PEARL, J. **Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference**. San Mateo, Calif.: Morgan Kaufmann, 1988.
- [PETERSON 53] PETERSON, W.; BIRDSALL, T. The theory of signal detectability. Electronic Defense Group, University of Michigan, Ann. Arbor. Technical Report, n.13, 1953.
- [POLYA 90] POLYA, G. **Mathematics and Plausible Reasoning Patterns of Plausible Inference**. New Jersey: Princenton University Press. 1990.
- ✓ [RABUSKE 95] RABUSKE, Renato Antonio. **Inteligência Artificial**. Editora UFSC. 1995.
- [REIS 97] REIS, Lisiane Albuquerque., CARGNIN, Moema Luz. **SDDEP – Uma Aplicação na Área Médica utilizando Raciocínio Baseado em Casos**. Trabalho de Conclusão de Curso de Graduação em Ciência da Computação, Universidade Federal de Santa Catarina, Florianópolis (Brasil), 1997.
- [RICH 88] RICH, E., **Inteligência Artificial**, McGraw-Hill, São Paulo, 1988.
- [RICH 93] RICH, E.; KNIGHT, K. **Inteligência Artificial**, Makron Books do Brasil Editora Ltda., São Paulo, 1993.
- [RÖDDER 93] RÖDDER, W. **Logical Inferencing in the Presence of Uncertainty**. FernUniversität Hagen. Germany, 1993.
- [RUSSEK 83] RUSSEK, E.; KRONMAL, R.A.; FISCHER, L.D. The effect of assuming independence in applying Bayes' theorem to risk estimation and classification in diagnosis. *Comput. Biomed. Res.*, v.16, p.537-52, 1983.
- [SHACHTER 88] SHACHTER, R.D. **Probabilistic inference and influence diagrams**. *Operations Research*, 36:589 – 604. 1988.
- [SHAFER 76] SHAFER, G. **A mathematical theory of evidence**. Princeton, Princeton University Press. 1976.
- [SHORTLIFFE 76] SHORTLIFFE, E.H. **MYCIN: Computer-based medical consultations**. New York: Elsevier Scientific, 1976.

- [SHORTLIFFE 90] SHORTLIFFE, E.H. **Clinical decision-support systems**. In: SHORTLIFFE, E.H.; PERREAULT, L.E., ed. *Medical Informatics: computer application in health care*. Reading (Massachusetts), Addison-Wesley publishing, p.466-502. 1990.
- X[SIMONS 88] SIMONS, G.T., **Introdução a Inteligência Artificial**. Editora Classe. 1988.
- [SMITH 85] SMITH, L. **Medicine as an art**. In: WYNGAARDEN, J.; SMITH, L.; ed. *Cecil Textbook of Medicine*. Philadelphia, Saunders, 1985.
- [SOX 86] SOX Jr., H.C. **Probability theory in the use of diagnostic tests**. *Ann. Internal Med.*, v.104, p. 60, 1986.
- [FRAWLEY 92] W. Frawley, G. Piatetsky-Shapiro, C. Matheus. **Knowledge Discovery in Databases: An Overview**. *AI Magazine*, Fall 1992, pgs 213-228.
- X[SPIRIT 97] **Spirit**. [on line] Documento disponível na internet via WWW: <URL: <http://pawpaw.fernuni-hagen.de/BWLOR/forsch.html>>, ago-97.
- [SRINIVAS 90] Srinivas, S . and Breese, J. **IDEAL: A software package for analysis of influence diagrams**. *Proceedings of the Sixth Uncertainty Conference in AI*, Cambridge, MA, 1990.
- [TESSARI 98] **Raciocínio Probabilístico em Sistemas Especialistas**. Relatório de iniciação científica, Departamento de Informática e Estatística, Universidade Federal de Santa Catarina, Florianópolis (Brasil), 1998.
- [STRATEGIST 97] **Strategist**. [on line] Documento disponível na internet via WWW: <URL: <http://www.prevision.com/strategist.html>>, dez-97.
- [VAN DER GAAG 96] VAN DER GAAG, L. C. **Bayesian Belief Networks: Odds and Ends**. *The Computer Journal*, Vol.39, No. 2, 97 - 113, 1996.
- [VAREJÃO 91] VAREJÃO, F. M., MACHADO, R. J., MILIDIU, R. L. **ALBA – Um ambiente para Classificadores Bayesianos por Aglomeração**. Rio de Janeiro. Technical Report CCR-125. Junho, 1991.
- [WEINTRAUB 84] WEINTRAUB, W.S.; MADEIRA, S.W.; BODENHEIMER, M.M.; et al. **Critical analysis of the application of Bayes' theorem to sequential testing in the noninvasive diagnosis of coronary artery disease**. *Am. J. Cardiol.*, v.54, p.43-9, 1984.
- [WRIGTH 21] WRIGTH, S. **Correlation and causation**. *Journal of Agricultural Research*, 20:557-85. 1921.
- [ZADEH 83] ZADEH, L. A. **The role of fuzzy logic in the management of uncertainty in expert systems**. *Fuzzy Sets and Systems*. (11):199-228. 1983.