

**UNIVERSIDADE FEDERAL DE SANTA CATARINA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA
COMPUTAÇÃO**

JULIANO TONIZETTI BRIGNOLI

**MODELO HÍBRIDO DIFUSO-PROBABILÍSTICO: UMA
ALTERNATIVA PARA SISTEMAS ESPECIALISTAS**

Dissertação submetida à Universidade Federal de Santa Catarina como parte dos requisitos para a obtenção do grau de Mestre em Ciência da Computação.

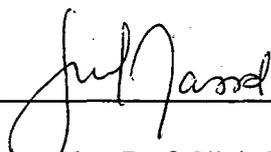
SILVIA MODESTO NASSAR
Prof. Orientador

Florianópolis, fevereiro de 2001

MODELO HÍBRIDO DIFUSO-PROBABILÍSTICO: UMA ALTERNATIVA PARA SISTEMAS ESPECIALISTAS

JULIANO TONIZETTI BRIGNOLI

Esta Dissertação foi julgada adequada para a obtenção do título de Mestre em Ciência da Computação Área de Concentração (Sistemas de Conhecimento) e aprovada em sua forma final pelo Programa de Pós-Graduação em Ciência da Computação.

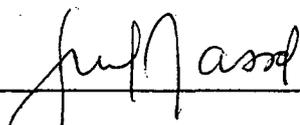


Orientador: Prof. Silvia Modesto Nassar, Dra

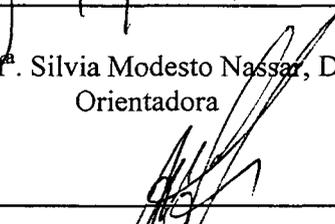


Coordenador do Curso: Prof. Fernando A. O. Gauthier, Dr

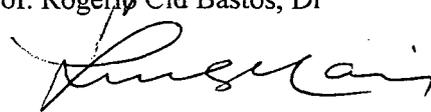
Banca Examinadora:



Prof.^a Silvia Modesto Nassar, Dra
Orientadora



Prof. Rogério Cid Bastos, Dr



Prof. Luiz Fernando J. Maia, Dr

*À minha esposa Ivonete, pelos momentos
sem minha presença e pelas mensagens de
motivação!*

*Ao meu filho Fernando, que por muitas
vezes brincando à só, desejava a
companhia do pai!*

AGRADECIMENTOS

Agradeço em primeiro lugar à Deus, pela força, motivação e coragem na realização desta etapa de minha vida.

À profª Silvia Modesto Nassar, pelo conhecimento, prestatividade, humildade e, como companheira fundamental na conclusão deste projeto.

À UNIDAVI e em especial ao Profº Francisco Fronza pela mobilização junto ao CAPES, os quais financiaram este projeto.

Aos meus pais, pelas suas preocupações em minha ausência e por todo incentivo que me foi atribuído.

SUMÁRIO

	Página
LISTA DE FIGURAS.....	vii
LISTA DE TABELAS.....	ix
LISTA DE ABREVIATURAS.....	x
RESUMO.....	xi
ABSTRACT.....	xii
1. INTRODUÇÃO.....	13
1.1. RELEVÂNCIA DA PESQUISA.....	14
1.2. OBJETIVOS.....	15
1.3. SEGMENTAÇÃO DO TRABALHO.....	16
2. ABORDAGEM SOBRE O DOMÍNIO DA APLICAÇÃO.....	18
2.1. INTRODUÇÃO.....	18
2.2. ASPECTOS RELATIVOS A SAÚDE PÚBLICA.....	19
2.3. MÉTODO DE ANÁLISE LABORATORIAL.....	21
3. INTELIGÊNCIA COMPUTACIONAL versus INTELIGÊNCIA ARTIFICIAL.....	22
3.1. INTRODUÇÃO.....	22
3.2. CONSIDERAÇÕES INICIAIS DA INTELIGÊNCIA ARTIFICIAL.....	22
3.3. MÁQUINA versus HUMANOS.....	23
3.4. ALGUMAS DEFINIÇÕES DE INTELIGÊNCIA ARTIFICIAL.....	24
3.5. INTELIGÊNCIA COMO UMA CARACTERÍSTICA INTERIOR.....	26
3.6. O SINTÁTICO E O SEMÂNTICO.....	26
3.7. O ARGUMENTO DO QUARTO CHINÊS.....	27
3.8. CONSIDERAÇÕES SOBRE AS ABORDAGENS DE IA.....	28
4. SISTEMAS ESPECIALISTAS.....	29
4.1. INTRODUÇÃO.....	29
4.2. NOTAS SOBRE SISTEMAS ESPECIALISTAS.....	34

4.3. BASE DE CONHECIMENTO	36
4.4. INTELIGÊNCIA ARTIFICIAL DISTRIBUÍDA.....	38
4.4.1. SISTEMAS BASEADOS EM CONHECIMENTO.....	40
4.4.2. SISTEMAS DE PRODUÇÃO	43
5. SISTEMAS ESPECIALISTAS PROBABILÍSTICOS.....	47
5.1. INTRODUÇÃO.....	47
5.2. TEOREMA DE BAYES.....	49
5.3. REDES BAYESIANAS	52
6. SISTEMAS ESPECIALISTAS DIFUSOS	55
6.1. INTRODUÇÃO	55
6.2. CONJUNTOS DIFUSOS	57
6.3. SISTEMAS DIFUSOS	58
6.4. NOTAS SOBRE ALGUMAS APLICAÇÕES DE CONJUNTOS DIFUSOS	64
7. SHELL PARA SISTEMAS ESPECIALISTAS	66
7.1. PROBABILÍSTICOS	66
7.2. DIFUSOS.....	69
8. PESQUISAS SIMILARES AO MODELO APRESENTADO.....	74
8.1. ASPECTOS GERAIS SOBRE MEDIDAS DE INCERTEZA	74
8.2. NOTAS SOBRE OUTRAS INVESTIGAÇÕES.....	76
9. MODELO HÍBRIDO	80
9.1. INTRODUÇÃO.....	80
9.2. ESQUEMATIZAÇÃO E FERRAMENTAS AUXILIARES	81
9.3. O QUALIFICADOR PARA O MODELO HÍBRIDO.....	92
9.4. A INTERFACE E DETALHES DE IMPLEMENTAÇÃO DO PROTÓTIPO	97
9.5. SIMULAÇÕES E ILUSTRAÇÕES DE CÁLCULOS	100
10. CONCLUSÕES	113
11. REFERÊNCIAS BIBLIOGRÁFICAS.....	115

LISTA DE FIGURAS

	Página
Fig. 1: Componentes de um sistema especialista probabilístico	49
Fig. 2: Um sistema difuso com duas entradas e uma saída	58
Fig. 3: Processo de inferência para um conjunto de três regras	60
Fig. 4: Esquema do modelo de um sistema difuso	62
Fig. 5: Processo de defuzzificação pelo método do Centro de Massa	63
Fig. 6: Uma rede bayesiana para diagnóstico de reações adversas causadas por antibióticos no leite (Fonte: DIAG, 199).....	67
Fig. 7: Interface de entrada para modelagem difusa no Matlab Toolbox Fuzzy	69
Fig. 8: Construção de um sistema difuso para diagnosticar credibilidade	70
Fig. 9: Interface para evolução do sistema no FUDGE	71
Fig. 10: Simulador construído no software FuzzyTech	72
Fig. 11: O Pêndulo Invertido	72
Fig. 12: Software UNFUZZY para sistema especialista difuso	73
Fig. 13: Esquema representativo do modelo híbrido proposto	82
Fig. 14: Uma Rede Bayesiana para protótipo	84
Fig. 15: Interface principal indicando o número de variáveis de entrada	88
Fig. 16: Especificação das variáveis de entrada e seus conjuntos difusos	88
Fig. 17: Especificação da variável de saída e seus conjuntos difusos	89
Fig. 18: Especificação das regras	89
Fig. 19: Simulação passo a passo de uma entrada ao sistema	90
Fig. 20: Híbridismo do probabilístico-possibilístico em função da qualificação das entradas.....	91
Fig. 21: Formulário de entrada do protótipo	97
Fig. 22: Formulário para entrada das quantidades de antibióticos	98
Fig. 23: Formulário para demonstração dos procedimentos de qualificação da Rede Bayesiana	98
Fig. 24: Formulário para escolha do formato das funções de pertinência	99
Fig. 25: Distribuição inicial das probabilidades da Rede.....	100
Fig. 26: Probabilidades condicionadas pela Presença de Penicilina e Estreptomicina.....	102
Fig. 27: Probabilidades condicionadas pela Presença de Penicilina (sem qualificação).....	104
Fig. 28: Probabilidades condicionadas pela Presença Difusa de Penicilina (com qualificação).....	105
Fig. 29: Probabilidades condicionadas pela Presença de todas as variáveis de entrada (sem qualificação).....	107
Fig. 30: Probabilidades condicionadas pela Presença de todas as variáveis de entrada (sem qualificação).....	107
Fig. 31: Probabilidades condicionadas pela Ausência de Estreptomicina (sem qualificação).....	109
Fig. 32: Probabilidades condicionadas pela Ausência de Estreptomicina (com qualificação).....	109
Fig. 33: Probabilidades condicionadas pela Presença de Tetraciclina (sem qualificação).....	110
Fig. 34: Probabilidades condicionadas pela Presença de Tetraciclina (com qualificação).....	110

Fig. 35: Probabilidades condicionadas pela Ausência de Tetraciclina (sem qualificação).....	111
Fig. 36: Probabilidades condicionadas pela Ausência de Tetraciclina (com qualificação).....	111
Fig. 37: Probabilidades condicionadas pela Presença de Penicilina e Ausência de Estreptomicina (sem qualificação).....	112
Fig. 38: Probabilidades condicionadas pela Presença de Penicilina e Ausência de Estreptomicina (sem qualificação).....	112

LISTA DE TABELAS

	Página
Tabela 1: Comparação entre o conhecimento especializado humano e artificial.....	33
Tabela 2: Probabilidades de Penicilina condicionadas às hipóteses diagnósticas.....	85
Tabela 3: Probabilidades de Estreptomicina condicionadas às hipóteses diagnósticas.....	85
Tabela 4: Probabilidades de Tetraciclina condicionadas às hipóteses diagnósticas.....	86
Tabela 5: Antibióticos comuns no estudo de caso e suas concentrações.....	87
Tabela 6: As funções de pertinência das variáveis de entrada sob a forma triangular.....	96
Tabela 7: Probabilidades $P(e/H_i)$ para Penicilina.....	101
Tabela 8: Probabilidades $P(e/H_i)$ para Estreptomicina.....	101
Tabela 9: Probabilidades $P(e/H_i)$ para Tetraciclina.....	102

LISTA DE ABREVIATURAS

1. SE - Sistema Especialista
2. IA - Inteligência Artificial
3. IAD - Inteligência Artificial Distribuída
4. SP - Sistema de Produção
5. SEP - Sistema Especialista Probabilístico

RESUMO

O trabalho apresenta uma modelagem híbrida para a concepção de sistemas computacionais inteligentes, mais precisamente no que refere-se aos sistemas especialistas. Para tal modelagem, foram utilizadas duas técnicas de Inteligência Computacional voltadas para o tratamento de incertezas, os Conjuntos Difusos e as Redes Bayesianas. Conjuntos Difusos processam incertezas por imprecisão e as Redes Bayesianas processam incertezas por aleatoriedade. Desta forma, a idéia da modelagem híbrida é aplicar Conjuntos Difusos para qualificar ou ajustar uma saída probabilística quando as evidências que constituem parte do conhecimento da Rede Bayesiana estão baseadas em informações imprecisas. Para tal, efetuou-se uma modificação na solução dada pelo Teorema de Bayes. Com objetivo de experimentar o modelo híbrido, propor-se um estudo de caso para o diagnóstico de reações adversas causadas pelo consumo de leite contaminado por quantidades de antibióticos, as quais, são imprecisas pelo fato da nebulosidade existente nas mesmas. Os especialistas estimam percentuais sobre as chances de se ter reação adversa caso haja a presença de antibióticos no leite, porém, estas estimativas não consideram a imprecisão na presença. Exatamente neste ponto é que a modelagem híbrida proposta atua no sentido de corrigir ou qualificar as saídas Rede Bayesiana. Testes e análise de resultados foram obtidos por um protótipo desenvolvido especificamente para simulações. Estes resultados mostraram que as probabilidades acima do ponto de máxima entropia diminuem quando vinculadas à eventos difusos e, as que ficam abaixo deste ponto, aumentam. Os resultados obtidos mantiveram a soma unitária do vetor de saídas da rede.

Palavras-Chave: Sistemas Especialistas, Conjuntos Difusos, Redes Bayesianas.

ABSTRACT

This work presents a hybrid modeling to the conception of intelligent computers systems, more precisely about the specialist systems. To such modeling, two techniques of computational intelligence have been used to the uncertainty approach, the fuzzy sets and the bayesian networks. As the fuzzy sets process the uncertainty by imprecision, the bayesian network process the uncertainty aleatorically. By this way, the idea of hybrid modeling is to apply fuzzy sets to qualify of adjust a probabilistic output when the evidences take part of the bayesian network knowledge based on inaccurate informations. Thus, purpose a change on solution given by Bayes formula. With the purpose of trying the hybrid modeling, it's presented a case study to the adverse reactions diagnostic in people who consume the contaminated milk for such antibiotics. This diagnostic may have an inaccurate result because of the nebulous about the quantity of antibiotics. The specialists range percentages about the chances of having an adverse reaction in case of the evidence being a kind of antibiotic, these estimate numbers do not consider the inaccurate origin of the fact, though. In this moment, specifically, the hybrid modeling bring into action the correction or qualification of the given probability as the bayesian network diagnostics. The simulations to the results analyzing are produced by softwares build specifically for this study. This results show that decrease the probabilities greater that maximum uncertainty point when associated from fuzzy events and, the lower from this point make up. This results gain keeping the unity sum from Bayesian Network output array.

Key-Words: Expert Systems, Fuzzy Sets, Bayesians Network.

1. INTRODUÇÃO

Na busca pela perfeição e melhoria da qualidade de vida, o homem tenta desde tempos mais remotos superar sua própria capacidade física e intelectual, procurando ultrapassar seus limites com o desenvolvimento de máquinas e computadores que possam substituir em atividades de risco ou simplesmente pela satisfação de materializar o pensamento ou idéia. O homem observou os pássaros voando, mas da sua impossibilidade física de imitá-lo, sua mente criativa se inspirou no modelo natural e inventou o avião. Desde a revolução industrial vimos o homem desenvolver inúmeras máquinas com funções variadas sempre objetivando amenizá-lo de tarefas árduas. Sem sombra de dúvidas, a invenção do computador marcou o início de uma era de transformações sociais. Tecnologia, indústria, comunicação, globalização, desemprego, informação, estes temas são intrínsecos de uma era em que computadores presenciam fortemente nosso cotidiano.

Paralelo ao desenvolvimento e melhoria de arquiteturas e sistemas de computação, os estudos da mente humana e da inteligência também alcançaram resultados satisfatórios. Hoje estas duas linhas de pesquisa se unem e resultam num poderoso recurso tecnológico do qual inúmeros benefícios sociais são realizados. Hoje, computadores inteligentes estão presentes em diversos setores: aviação, navegação, controle de satélites, sistemas bélicos, medicina, automação em geral.

A Inteligência Computacional é entendida como o emprego de diversas técnicas de programação fundamentadas por modelos matemáticos e na sua maioria inspirados em funções e atividades de seres do mundo real.

Tendo em vista que técnicas matemáticas são amplamente utilizadas na concepção de modelos computacionais, a idéia central deste trabalho consiste em contribuir com os Sistemas Especialistas, oportunizando os processos de decisão e de diagnóstico partindo da combinação entre Conjuntos Difusos e Redes de Probabilidades para apresentar um novo modelo caracterizado pelo hibridismo das técnicas anteriores. Tal hibridismo denotado por um processo de *qualificação*, tem a característica de redistribuir as probabilidades de uma Rede Bayesiana para diagnóstico, considerando que nas entradas ou eventos desta rede, existe imprecisão, ou seja, os eventos são difusos.

É fundamental que esta nova abordagem para aplicações de Sistemas Especialistas não interfere nas propriedades axiomáticas individuais das anteriores, mas, aproveita as características lógicas dos processos de raciocínio das mesmas.

1.1 RELEVÂNCIA DA PESQUISA

Em termos de contribuição para os modelos empregados em Sistemas Especialistas, acredita-se que o hibridismo apresentado nesta pesquisa permite uma melhor modelagem, no sentido de estar mais próxima à solução natural obtida no raciocínio lógico humano, em processos de diagnose utilizando técnicas únicas de raciocínio aproximado.

Para uma modelagem computacional, a grande importância está no fato de que se pretende conciliar no mesmo estudo a função de uma Rede Bayesiana, fundamentada pela Teoria da Probabilidade e, utilizar Conjuntos Difusos para ajustar as *probabilidades* quando na detecção de imprecisão existente nos eventos.

Entende-se que esta investigação contribua para a Ciência da Computação, mas especificamente, para o conhecimento pertinente às linhas de trabalho da Inteligência Computacional, no sentido de permitir a reflexão sobre a agregação de técnicas matemáticas fortemente calcadas em propriedades axiomáticas, para a construção de sistemas computacionais que possam aproveitar em um único processo as características essenciais destas técnicas. Sabe-se que o raciocínio de um especialista humano é apoiado por vários princípios lógico-matemáticos para alcançar um resultado em processos decisórios, mas, tal raciocínio é caracterizado pela idéia de aproximação, baseado muitas vezes em deduções a partir da observação de fenômenos.

Então, se o objetivo principal da Inteligência Computacional e das técnicas empregadas pelos Sistemas Especialistas é justamente imitar um especialista humano naquilo que ele faz de melhor, pode-se pensar na possível associação ou combinação de todos os métodos matemáticos e inspirações biológicas que já contribuíram para a melhoria do conhecimento na Ciência da Computação, pois, poder-se-á acreditar que o potencial destas combinações produzirão resultados mais próximos dos reais.

De acordo com o domínio desta aplicação, destaca-se o fato principal na inspiração para a realização desta pesquisa, como sendo a imprecisão existente na presença de quantidades de antibióticos no leite de consumo, as quais, podem influenciar resultados de diagnósticos fundamentados em valores probabilísticos. Estes diagnósticos tem como resultado uma probabilidade de uma pessoa vir a ter uma reação adversa com base em evidências da existência de certos antibióticos no leite. Justifica-se a modelagem híbrida pelo fato de verificar-se a imprecisão existente nas evidências ou eventos e assim, processar esta imprecisão em termos de *qualificar* ou *ajustar* uma distribuição de probabilidades.

1.2 OBJETIVOS

O objetivo geral desta pesquisa é desenvolver uma proposta de modelagem híbrida difuso-probabilística. Para alcançar o objetivo geral é necessário que sejam desenvolvidas etapas de trabalho, e assim, destacam-se os objetivos específicos:

- Escolher um domínio de aplicação adequado à validação da proposta híbrida;
- Construir uma Rede Bayesiana para a obtenção dos resultados probabilísticos de diagnose;
- Propôr um modelo difuso considerando as entradas utilizadas na Rede Bayesiana de forma a *ajustar* a saída probabilística do diagnóstico;
- Avaliar o modelo híbrido difuso-probabilístico no diagnóstico.

1.3 SEGMENTAÇÃO DO TRABALHO

O relatório está segmentado em capítulos na seguinte forma:

O capítulo 1 apresenta uma introdução procurando contextualizar o estado da Inteligência Computacional, bem como citar o tema que pretende-se pesquisar e as técnicas de Inteligência Computacional a utilizar. A motivação para o tema proposto é justificada e são apresentados os objetivos geral e específicos.

A fundamentação teórica, bem como, o que pode-se chamar de “Estado da Arte”, será apresentada em vários capítulos contextualizando técnicas desenvolvidas pela Inteligência Computacional, suas propostas e modelos matemáticos, enfatizando principalmente os Sistemas Especialistas e as investigações já realizadas com relação ao tema abordado. Tais capítulos são assim segmentados:

- Capítulo 2: Uma abordagem que visa conhecer o domínio da aplicação, os problemas sobre antibióticos no leite, bem como, doenças e atividades de diagnose laboratoriais oriundas de tal setor;
- Capítulo 3: A Inteligência Computacional como uma nova maneira de referenciar a Inteligência Artificial, com suas inspirações em modelos biológicos, técnicas, vantagens e aplicações, bem como salientando a necessidade de compreender que, *máquina inteligente* não substitui por completo a inteligência humana;
- Capítulo 4: Uma abordagem sobre os Sistemas Especialistas e os principais modelos desenvolvidos, bem como, uma breve discussão sobre Inteligência Artificial Distribuída;
- Capítulo 5: Apresenta-se também, uma revisão de Estatística aplicada em Modelos Inteligentes, no caso a Teoria da Probabilidade e o Teorema de Bayes. Aqui mostra-se o que vem a ser uma Rede Bayesiana e sua grande contribuição aos Sistemas Especialistas de diversas aplicações. Dá-se ênfase aos Sistemas Especialistas na área médica;
- Capítulo 6: Uma revisão sobre os Conjuntos Difusos, sua formalidade matemática e suas áreas de aplicação, bem como, a apresentação de exemplos de implementações com esta técnica.
- Capítulo 7: Uma visão superficial sobre algumas ferramentas para a concepção de Sistemas Especialistas Difusos e Probabilísticos;
- Capítulo 8: Um resumo sobre a investigação realizada por outros pesquisadores que contribuíram com soluções para esta abordagem.

No capítulo 9 apresenta-se o modelo construído e os detalhes de implementação e simulação de resultados.

No capítulo 10, são apresentadas as conclusões sobre a investigação, a análise dos resultados obtidos e, as sugestões para pesquisa com similaridades ao tema abordado. No capítulo 11 é apresentada uma relação de referências bibliográficas.

2. ABORDAGEM SOBRE A IDENTIFICAÇÃO DE ANTIBIÓTICOS NO LEITE

2.1. INTRODUÇÃO

Com o desenvolvimento da antibioticoterapia, notadamente após a 2ª guerra mundial, os antibióticos passaram a ser cada vez mais utilizados na pecuária para o tratamento das doenças infecciosas, como fator de crescimento nas dietas e também como conservantes dos alimentos.

Na pecuária leiteira, a doença mais importante é a mastite e para seu controle nos animais infectados, diversos antibióticos e quimioterápicos são administrados, geralmente em soluções ou suspensões difundidos diretamente no úbere infectado, por via sistêmica ou com as duas formas combinadas. A ordenha dos animais assim tratados, num período de 72 horas após a última aplicação, revela nitidamente a presença dos antibióticos em uso. (SOUZA, 1998)

Após o tratamento de mastite (via intra-mamária e/ou sistêmica), estas substâncias podem ser encontradas em concentrações suficientes para inibir culturas lácteas e causar perdas econômicas nas indústrias de queijos e leite fermentado. Essas perdas são decorrentes de alterações como coagulação inadequada do leite, maturação imprópria do creme durante a produção de manteiga e produtos similares e diminuição do crescimento das culturas lácteas em leite em pó desnatado reconstituído.

Além da mastite, outras doenças acometem o gado leiteiro, havendo a necessidade do emprego de diversos antibióticos administrados por diversas vias e com períodos de eliminação variáveis.

Segundo a portaria 005 de 24 de abril de 1980, publicada no diário oficial da união, ano CXVIII, baixada pela Divisão de Inspeção de Leite e Derivados (DILEI) da Secretaria da Inspeção de Produtos de Origem Animal (SIPA) do Ministério da Agricultura, “a prova da redutase do leite deve ser efetuada com o uso do corante azul metileno, no mínimo duas vezes por semana, por produtor, não devendo o tempo de descoloração ser inferior a duas horas e trinta minutos”.

A redução dos corantes constituem uma prova específica do estado de frescor do leite cru e a sua capacidade de conservação frente a acidificação, sendo uma forma de se estimar o conteúdo geral de germes presentes.

O teste baseia-se no fato da maioria das bactérias que se desenvolvem no leite, abaixarem o potencial de oxidação-redução, provocando mudança na cor de certos corantes, estimando indiretamente o número de bactérias nas amostras de leite, as quais determinam

mudanças no substrato causadas pela atividade metabólica, diminuindo a quantidade de oxigênio dissolvido e pela produção de metabólitos e enzimas específicas.

De acordo com o Standard Methods for the Examination Dairy Products (American Public Health Association, A.P.H.A., 1978) “o tempo de redução deve ser inversamente proporcional à população bacteriana, o que foi comprovado comparando-se a variação do tempo de redução com as populações bacterianas determinadas por outros métodos”.

Torna-se portanto uma questão importante na inspeção de produtos de origem animal, o fato da atuação bacteriostática das substâncias inibidoras criar condições adversas à multiplicação da flora de contaminação, podendo vir a mascarar os testes de redução de corantes, dissimulando assim, a má qualidade de certos produtos.

A estratégia para a utilização dos testes para determinação de resíduos de antibióticos e sulfonamidas (inibidores) no leite, compreendem dois diferentes aspectos: (SOUZA, 1998)

1. **Segurança toxicológica:** critérios governamentais para a saúde pública, por meio de leis para a produção e comercialização de alimentos;
2. **Segurança tecnológica:** pagamento do leite com base na qualidade.

Referindo-se ao preço de mercado do leite, a presença de inibidores freqüentemente resulta em preços mais baixos ou outros arranjos desvantajosos para os produtores, enquanto que sob a ótica das leis que regem a produção e a comercialização de alimentos, são possíveis penalidades judiciais. Quanto aos riscos para a saúde pública que devem ser levados em conta, são:

- Riscos farmacológicos e toxicológicos;
- Riscos microbiológicos (favorecimento de resistência de microrganismos patógenos na flora intestinal);
- Riscos imunopatológicos (alergias).

2.2. ASPECTOS RELATIVOS A SAÚDE PÚBLICA

Em (TAVARES, 1990) é discutido que antibiótico é todo composto químico produzido por um microrganismo, possuindo a propriedade, em solução diluída, de inibir o crescimento de outros microrganismos, ou mesmo de os destruir.

Os antibióticos são substâncias estranhas ao organismo humano, de constituição química diversa e capazes de causar reações locais devidas a natureza cáustica de sua

composição. Estas reações podem resultar da constituição química própria da droga ou de sais sob a forma dos quais é administrado, ou ainda de formulações farmacêuticas em que são apresentadas.

Todos os antibióticos, a par dos extraordinários efeitos curativos, podem causar efeitos indesejáveis, os quais dependem do indivíduo e da droga utilizada. Estes efeitos, por vezes, colocam a vida do usuário em risco ou determinam seqüelas orgânicas. Estima-se que 20% dos pacientes que procuram cuidados médicos, apresentam em sua história progressa algum tipo de efeito adverso causado por substâncias antimicrobianas. Estes efeitos colaterais podem ser: (SOUZA, 1998)

1. De natureza irritativa:

- 1.1. Efeitos irritativos gastrointestinais
- 1.2. Efeitos irritativos para os músculos
- 1.3. Efeitos irritativos para as veias

2. De natureza alérgica:

- 2.1. **Benignos:** urticária, prurido, eosinofilia, eritema polimorfo, eritema nodoso, dermatite de contato, edema de Quincke, rinite alérgica, icterícia colestática, fotossensibilização.
- 2.2. **Graves:** há risco de vida (choque anafilático, dermatite esfoliativa, hemólise, síndrome de Stevens-Johnson, discrasias sangüíneas, edema de glote, broncoespasmo, vasculites, nefrite intersticial). Podem ainda ocorrer manifestações de hiperssensibilidade de natureza grave sem risco para o paciente, porém, com danos sérios a determinados órgãos, como é o caso da surdez e cegueira observadas raramente como manifestação de hiperssensibilidade à ampicilina.

O choque anafilático é mais comum com a penicilina G, sendo esta droga responsável por 76% dos acidentes anafiláticos aos antibióticos, seguindo-se as tetraciclina e o cloranfenicol; a aplasia medular é mais freqüente com o cloranfenicol; a dermatite de contato ocorre em 10% dos pacientes que usam a penicilina G localmente, seguindo-se de estreptomicina em 1% dos casos; a reação semelhante a doença do soro é mais habitual com as penicilinas e estreptomicina; a colestase intra-hepática ocorre com o estolado de eritromicina e a triacetiloleandomicina.

As reações alérgicas aos antibióticos são fundamentalmente dependentes do indivíduo e em geral, aparecem após o seu uso anterior, mas é preciso lembrar que o indivíduo já pode ter

sido sensibilizado por fungos produtores das drogas, existentes no meio ambiente (inalação de *Penicillium*, por exemplo) ou a sensibilização anterior pela introdução do antibiótico no organismo junto a alimentos ou outros produtos (por exemplo, a sensibilização à penicilina pela ingestão de leite procedente de vacas tratadas para mastite mostrando para efeito alérgico mesmo sem a prévia utilização do antibiótico).

2.3. MÉTODO DE ANÁLISE LABORATORIAL

Existem diversos métodos para a obtenção de informações sobre a presença ou não de resíduos de antibióticos no leite. Referencia-se aqui, o método utilizado pela pesquisa a qual serviu como fonte de dados para este trabalho.

De acordo com (NELSON, 1998) a presença de resíduos de antibióticos nas amostras de leite, foi verificada utilizando-se o ADM (Antimicrobial Diffusion Method), comercializado no Brasil pela CHr. Hansen. O Kit consiste de 100 tubos, cada qual com o meio de cultura contendo esporos do Bacilo *Sterothermophilus var. calidolactis*.

Este método foi introduzido em 1975 pela Gist-Brocades Laboratories na Holanda. É um teste de difusão em ágar.

Tabletes contendo nutrientes e o indicador bromocresol púrpura, são adicionados em ampolas contendo ágar simples com esporos de *Bacillus Sterothermophilus var. calidolactis*. Após adicionar 0,1 ml da amostra de leite, a ampola é incubada por 2 horas e meia a $64 \pm 1^\circ$ C. Na ausência de substâncias antimicrobianas, o meio torna-se amarelo (resultado negativo), entretanto, permanecerá na cor púrpura (resultado positivo) na presença de concentrações suficientemente altas de antibióticos, capazes de inibir o crescimento do microorganismo teste.

Frente a uma concentração muito baixa de antibióticos, o meio tomará uma tonalidade intermediária (resultado duvidoso). Este método é altamente sensível para penicilina. Concentrações de 0,003 UI/ml ou menores, dão resultado negativo, de 0,004 UI/ml dão resultado duvidoso e 0,005 UI/ml ou maiores, dão resultado positivo.

Este teste também é suficientemente sensível para muitos outros antibióticos, usualmente administrados em vacas lactantes.

Antes de proceder a análise no ADM, todas as amostras são aquecidas a 80° C por 10 minutos, com o objetivo de eliminar o efeito dos inibidores naturais do leite.

3. INTELIGÊNCIA COMPUTACIONAL versus INTELIGÊNCIA ARTIFICIAL

3.1. INTRODUÇÃO

Com base em análises de estudos e projetos realizados na área de Inteligência Artificial, vêm-se observando a algum tempo que existe uma mudança na forma de como se referenciar aos termos Inteligência Artificial e Inteligência Computacional. O que parece é uma migração para a segunda denotação.

Sendo assim, pretende-se nesta unidade relatar discussões sobre os dois termos. Esta discussão procura ressaltar conceitos oriundos das áreas que concebem sistemas computacionais inteligentes (Ciência da Computação, Engenharias) e, conceitos ligados às Ciências Cognitivas, bem como aos estudos do funcionamento da mente humana.

3.2. CONSIDERAÇÕES INICIAIS DA INTELIGÊNCIA ARTIFICIAL

Um dos principais objetivos dos cientistas da computação é o desenvolvimento de máquinas que interajam com os seus ambientes à maneira sensorial que tradicionalmente caracteriza os humanos, e que desempenhe suas funções, de forma inteligente, sem a necessidade da intervenção humana. A concretização deste objetivo exige que a máquina “entenda” ou perceba o estímulo recebido e seja capaz de tirar conclusões mediante alguma forma de processo de raciocínio. Tanto a percepção quanto a dedução se enquadram na categoria das atividades corriqueiras que, embora naturais para a mente humana, representam intensas dificuldades para as máquinas. O resultado disto tudo é que a área de pesquisa relacionada com esse objetivo, a *Inteligência Artificial*, esteja numa fase muito primitiva em comparação com seus alvos e expectativas. O tema da Inteligência Artificial deve ser discutido em dois contextos: O primeiro é de ordem filosófica, no qual são consideradas questões relativas à *inteligência propriamente dita*, e se as máquinas são capazes de processar inteligência real, ou somente simular a sua presença; o outro, de ordem mais científica, no qual se questiona de que maneira a tecnologia pode ser empregada na produção de máquinas que se *comportem* de forma inteligente. (BROOKSHEAR, 2000)

3.3. MÁQUINAS versus HUMANOS

Embora o computador seja seguidamente personificado, há uma importante distinção entre suas características e as da mente humana. Máquinas algorítmicas são projetadas para executarem tarefas precisamente definidas, com velocidade e exatidão, e fazem isso extremamente bem. Todavia, máquinas são desprovidas do *bom senso*. Quando diante de uma situação não prevista pelo programador, é provável que seu desempenho caia sensivelmente. A mente humana, embora em geral, tenha dificuldades com cálculos complexos, é capaz de compreender e de deduzir. Por conseguinte, embora a máquina consiga ultrapassar o ser humano no cálculo das soluções de problemas de física nuclear, por exemplo, o ser humano, provavelmente, está muito mais apto a compreender os resultados e a determinar qual deve ser o próximo cálculo a ser realizado. (BROOKSHEAR, 2000)

Um exemplo interessante a enfatizar aqui é quanto ao ato de “estar dirigindo e realizar uma ultrapassagem”. Para um ser humano esta atividade pode ser extremamente simples e, as vezes, até *consientemente independente*. Como o cérebro humano realiza o que pode-se chamar de *processamento paralelo*, um motorista realiza a ultrapassagem, talvez até ouvindo música, conversando com o caroneiro, ou até mesmo, lembrando de outras tarefas que deverá fazer, porém, tudo acontece com a maior naturalidade e harmonia, o motorista não está fazendo nenhum cálculo complexo para saber o quanto deverá acelerar e/ou frear, o quanto deverá girar o volante para à esquerda e/ou direita e até mesmo analisar a distância do veículo à sua frente em relação à sua velocidade atual. Agora imaginemos um computador conectado com diversos sensores adaptados ao veículo e um programa “inteligente” que dará as diretivas para conduzir o tal veículo. Este programa muito provavelmente teria que, com base nas informações transmitidas pelos sensores, calcular distância, aceleração, velocidade, considerar o fenômeno da inércia dos corpos, e aplicar diversos cálculos relacionados às leis da física para, ainda assim, aproximar-se de modo inferior às atitudes tomadas pelo motorista humano que não está preocupado com uma precisão matemática.

Em (PEREIRA, 2000), relata-se que a reprodução de características inteligentes em máquinas construídas pelo homem é um dos objetivos mais perseguidos pela comunidade científica e tecnológica. Muitos paradigmas simbólicos de aprendizagem surgiram, e muitos se desenvolveram como métodos computacionais poderosos, incluindo aquisição indutiva de conceitos, sistemas classificadores e aprendizagem baseada em explicadores. Um conjunto de ferramentas analíticas, conhecidas coletivamente como *Inteligência Computacional*, tem sido utilizadas para solucionar problemas relacionados à Aprendizagem de Máquinas, permitindo

assim a resolução de problemas que eram previamente difíceis ou impossíveis de serem tratados. A característica de destaque da Inteligência Computacional é que seu objetivo primário é alcançar tratabilidade, robustez, baixo custo e alto QIM (Quoeficiente de Inteligência de Máquina) com uma exploração da tolerância à imprecisão e incerteza. O que se busca nesse contexto é uma solução aproximada para um problema formulado imprecisamente. Um simples exemplo é o de se estacionar um carro. Estaciona-se um carro com relativa facilidade porque a posição final e a orientação do carro não são especificados precisamente. Se o fossem, o problema seria praticamente insolúvel para os seres humanos. Em contraste, quando o problema de se estacionar é formulado imprecisamente, torna-se difícil de resolvê-lo pelos métodos computacionais porque tais métodos não tratam do problema de se explorar a tolerância por imprecisão. O que este exemplo simples quer ilustrar é o fato que, em geral, uma precisão alta acarreta um alto custo. O desafio, portanto, é explorar a tolerância à imprecisão pela busca de novos métodos de computação que levem a uma solução aceitável a um baixo custo. Pela sua própria natureza, a Inteligência Computacional é muito mais próxima ao raciocínio humano do que os modos tradicionais de computação.

3.4. ALGUMAS DEFINIÇÕES DE INTELIGÊNCIA ARTIFICIAL

“IA é a parte da ciência da computação voltada para o desenvolvimento de sistemas de computadores inteligentes, isto é, sistemas que exibem características, as quais nós associamos com a inteligência no comportamento humano – compreensão da linguagem, aprendizado, raciocínio, resolução de problemas.” (FEIGENBAUM, 1981)

“O campo de estudo que tenta explicar e emular comportamento inteligente em termos de processo computacional.” (SCHALKOFF, 1990)

“Inteligência Artificial é o estudo das idéias que permitem aos computadores serem inteligentes.” (WINSTON, 1984)

“Inteligência Artificial é o estudo das faculdades mentais através do uso de modelos computacionais.” (CHARNIAK, 1985)

“Inteligência Artificial é o estudo de como fazer os computadores realizarem coisas que, no momento, as pessoas fazem melhor.” (KNIGHT, 1993)

“É a atividade que nós associamos com o pensar humano, atividades tais como: tomada de decisão, resolução de problemas, aprendizado” (BELLMAN, 1978)

“A inteligência Artificial (IA) é simplesmente uma maneira de fazer o computador pensar inteligentemente.” (LEVINE, 1988)

Em (LIMA, 1999) são ressaltadas as duas principais abordagens dos princípios teóricos da Inteligência Artificial:

- **Abordagem Cognitiva:** também denominada de descendente, dá ênfase aos processos cognitivos, ou seja, a forma como o ser humano raciocina. Objetiva encontrar explicação para comportamentos inteligentes baseado em aspectos psicológicos e processos algorítmicos. Os pioneiros dessa corrente foram John McCarthy, Marvin Minsky, Newell e Simon. As primeiras modelagens da inteligência surgiram na década de 50 e tiveram como base as regras de produção e a lógica dos predicados. A formalização da lógica facilitou o processo de formalização e representação dos conhecimentos a serem utilizados pelos programas de computador. Inicialmente esses conhecimentos se restringiram a esquemas de raciocínios para jogos, aplicações matemáticas e simuladores.
- **Abordagem Conexionista:** também denominada de biológica ou ascendente, dá ênfase no modelo de funcionamento do cérebro, dos neurônios e das conexões neurais. Os pioneiros dessa corrente foram McCulloch, Pitts, Hebb, Rosenblatt e Widrow. Em 1943 surgiu a representação e formalização matemática dos neurônios artificiais, que fez surgir os primeiros modelos de redes neurais artificiais. A corrente conexionista sofreu grande impacto quando os cientistas Marvin Minsky e Seymour Papert publicaram em 1969 o livro *Perceptrons*, no qual criticaram e sustentaram que os modelos das redes neurais não tinham sustentação matemática suficiente que lhes fosse possível atribuir alguma confiabilidade. Apesar das pesquisas nesta área não terem parado, foi apenas na década de 80 que o físico e biólogo do Instituto de Tecnologia da Califórnia, John Hopfield conseguiu recuperar a credibilidade da utilização das redes neurais.

3.5. INTELIGÊNCIA COMO UMA CARACTERÍSTICA INTERIOR

A dificuldade para determinar se um programa possui ou não inteligência reside na dificuldade de distinguir entre a mera aparência de inteligência e sua existência de fato. Na análise final, inteligência é uma característica interior, cuja presença é externamente detectada apenas de modo indireto, com base no contexto do diálogo estímulo/reação. Essa natureza ilusória da inteligência foi conhecida por Alan Turing em 1950, quando propôs um teste (agora conhecido como o Teste de Turing) para detectar inteligência em uma máquina. A proposta de Turing era a de permitir que um humano, a quem se chama de interrogador, se comunique, por meio de respostas interativas, com um teste, por meio de uma máquina de escrever, sem ser avisado se o teste foi originado por um ser humano ou por uma máquina. Nesse ambiente, a máquina seria declarada inteligente naquele evento se o interrogador não pudesse distingui-la de um ser humano. Até o momento, nenhuma máquina foi capaz de passar no Teste de Turing, embora resultados surpreendentes tenham sido alcançados. (BROOKSHEAR, 2000)

3.6. O SINTÁTICO E O SEMÂNTICO

Conforme em (RICH, 1993), numa das extremidades do *espectro sintático-semântico* estão os sistemas puramente sintáticos, onde não existe preocupação com o significado do conhecimento que está sendo representado. Tais sistemas tem regras simples e uniformes para manipular a representação. Eles não se importam com o conteúdo da representação. No outro extremo, estão os sistemas puramente semânticos, onde não há formato uniforme. Cada aspecto da representação corresponde a um tipo particular de informação, e as regras de inferência são correspondentemente complicadas. As representações lógicas (lógica de predicados e sistemas não-monotônicos) e as estatísticas são as mais puramente sintáticas. Suas regras de inferência são procedimentos estritamente sintáticos que operam sobre fórmulas bem formadas independente do que essas fórmulas representam. Os sistemas de regras de produção também são basicamente sintáticos. Os interpretadores desses sistemas normalmente usam apenas informações sintáticas (como o formato do padrão do lado esquerdo, a posição da regra na base de conhecimentos, ou a posição do objeto casada na memória de curto prazo) para decidir que regras devem ser disparadas. Aqui percebe-se a semelhança entre lógica e regras de produção como o meio de representar e usar o conhecimento. Mas é possível criar sistemas de regras de produção que tenham mais

semânticas embutidas nas regras. As estruturas de escaninhos-e-preenchimento têm orientação tipicamente mais semântica, embora estejam a uma boa distância neste espectro. As redes semânticas, como o próprio nome indica, destinam-se a captar os relacionamentos semânticos entre entidades e normalmente são empregadas com um conjunto de regras de inferência que foi especialmente projetado para controlar corretamente os tipos específicos de arcos presentes na rede.

Existem outras estruturas como os Frames, Sistemas de Raciocínio Estatístico, Dependência Conceitual.

3.7. O ARGUMENTO DO QUARTO CHINÊS

O Argumento do Quarto Chinês proposto por John Searle possui o intuito de demonstrar a falibilidade da Inteligência Artificial Forte, bem como, apresentar uma alternativa de elaboração teórica que sugere a tradicional articulação da solução para o problema mente-corpo. O Quarto Chinês funciona da seguinte forma: em um quarto equipado com uma variedade de caixas contendo cartas em que estão escritos caracteres da língua chinesa, mais um manual com instruções, há uma pessoa (chamada operador), que deve receber papéis com instruções em chinês (que lhe são passados de fora do quarto), os quais devem ser colocados em seqüência. As instruções do manual fornecem ao operador um modo de organizar em seqüência os caracteres chineses que lhe são passados, indicando quais caixas usar e como juntar as cartas com símbolos numa seqüência certa, que, são devolvidas para fora do quarto. É importante ressaltar que o operador do Quarto Chinês não entende chinês, portanto está produzindo *output* sem nenhuma idéia do seu significado, e não pode querer dizer nenhuma frase com ele. No entanto, seguindo as instruções o operador efetivamente produz uma frase, o *output* do Quarto Chinês, perfeitamente inteligível para alguém que entenda chinês, sem que ele próprio saiba a frase que produziu. E é isso que o Quarto Chinês vem ressaltar, que o *output* retornado pelo operador, sendo uma frase em chinês, não significa que este operador (que está no papel de computador) saiba alguma coisa em chinês. O que pode-se afirmar é que quem formulou os manuais seguidos pelo operador, esse sim, tem conhecimento da língua chinesa. (MACHADO, 1999)

3.8. CONSIDERAÇÕES SOBRE AS ABORDAGENS EM IA

Tendo em vista as discussões relatadas sobre a Inteligência Artificial e a confrontação com a denominação Inteligência Computacional, bem como a mensagem deixada por John Searle na sua argumentação sobre o Quarto Chinês, conclui-se que a reprodução da mente humana considerando a sua capacidade de entendimento semântico e não apenas simbólico é praticamente impossível, pois com todas as técnicas e modelos matemáticos existentes a Ciência da Computação e áreas de conhecimento afins conseguiram apenas “imitar” formas de raciocínio humano, porém, tudo está muito distante ao fato de que máquinas podem deduzir, aumentar seu conhecimento com base em suas deduções e possuir “bom senso” nas decisões como os humanos possuem.

4. SISTEMAS ESPECIALISTAS

4.1. INTRODUÇÃO

Conforme (RICH, 1993) os sistemas especialistas são projetados para solucionar problemas que normalmente são solucionados por especialistas humanos. Para solucionar tais problemas, os sistemas especialistas precisam acessar uma substancial base de conhecimentos do domínio da aplicação, que precisa ser criada do modo mais eficiente possível. Eles também precisam explorar um ou mais mecanismos de raciocínio, para aplicar seu conhecimento aos problemas que tem diante de si. Depois eles precisam de um mecanismo para explicar aos seus usuários o que fizeram. Uma maneira de ver os sistemas especialistas é considerar que eles representam a inteligência artificial aplicada em um sentido bem amplo. Os problemas com os quais lidam os sistemas especialistas são altamente diversificados. Há questões gerais que surgem em vários domínios. Mas há também técnicas poderosas que podem ser definidas para classes específicas de problemas. Uma extensão importante do conceito de banco de dados inteligentes é o desenvolvimento de sistemas especialistas – pacotes de softwares projetados para ajudar os seres humanos em situações nas quais é imprescindível a presença de um especialista. Tais sistemas são projetados para simular um raciocínio do tipo causa-e-efeito, que os especialistas realizaram diante das mesmas situações. Assim, um sistema especialista médico busca propor o mesmo procedimento que um médico especialista o faria: sabendo que, sendo notado o aparecimento de uma anormalidade cuja radiografia acuse a presença de massa em tal região, uma biopsia deverá ser efetuada. Uma das principais tarefas na construção de um sistema especialista é a obtenção do conhecimento necessário, fornecido por um especialista. A maneira de resolver este problema se tornou uma importante área de pesquisa. Esse problema é, na verdade, composto de duas tarefas. Uma delas consiste em obter e manter a cooperação do especialista – algo nada fácil de se contornar. Uma vez superados tais problemas, o conhecimento obtido do especialista deve ser organizado em um formato compatível com um sistema computacional. Esta organização é, em geral, realizada expressando tal conhecimento por meio de um conjunto de regras na forma de cláusulas do tipo *if-then*. Por exemplo, a regra para que uma anormalidade, confirmada por radiografias, venha a solicitar uma biopsia pode ser expressa como:

```
If presença_de_anormalidade_e  
    radiografia_mostra_presença_de_massa  
  
then  
    execute_biopsia
```

Observa-se agora a semelhança entre as regras de um sistema especialista e as regras de produção de um sistema de produção (sistemas de produção serão abordados mais adiante). A parte **if** do comando determina, essencialmente, as precondições para executar ou concluir a parte **then** do comando. De fato, muitos sistemas especialistas são essencialmente, sistemas de produção, cujas regras, obtidas a partir do especialista humano, são as produções, e o raciocínio subjacente é baseado na simulação de tais regras pelo sistema de controle. Nesse contexto, o conjunto de produções é denominado base de conhecimento do sistema, e o sistema de controle, máquina de inferência. Alguns sistemas especialistas são organizados como conjuntos de sistemas de produção, que combinam os seus esforços para a resolução de problemas. Como exemplos, pode-se citar os sistemas especialistas baseados no modelo de “quadro-negro”. Nele, vários sistemas de resolução de problema, denominados fontes de conhecimento, compartilham uma área comum de armazenamento, chamada quadro-negro. Este quadro contém o estado corrente do problema a ser resolvido e, sendo compartilhado por todas as fontes de conhecimento, proporciona um meio com o qual as fontes de conhecimento podem contribuir para a solução do problema. Para coordenar as atividades das fontes de conhecimento, há um módulo de controle, encarregado da tarefa de ativar a fonte de conhecimento apropriada no momento adequado. Na terminologia do modelo do quadro-negro, diz-se que este modo de controle determina o “foco de atenção” do sistema. Outra diferença entre um sistema especialista e um sistema simples de produção é que o primeiro não está necessariamente encarregado de alcançar uma meta predeterminada mas, mais provavelmente, encarregado de fornecer respostas bem fundamentadas. Isto significa que as heurísticas empregadas não são medidas de proximidade de uma meta, pois nenhuma meta necessita realmente estar presente. As heurísticas utilizadas em sistemas especialistas tendem a ser os “palpites” usados pelo especialista humano. Veja agora, com mais detalhes, a afirmação de que não há necessidade de uma meta no ambiente de um sistema especialista. Suponha que um especialista ou um sistema especialista esteja encarregado da tarefa de diagnosticar doenças. Ideal seria se ambos pudessem concluir com uma proposição definitiva

da forma: “A doença é X”, em que X é o nome da doença, fornecida pela proposição. Tal precisão pode ser inviável. Em lugar disto, a melhor resposta poderia ser do tipo: “É provável que a doença seja X”, ou talvez “A doença é ou X ou Y. Por favor execute o seguinte teste para determinar qual delas é a mais provável”. Por causa dessa ambigüidade, o sistema de controle contido em um sistema especialista pode seguir várias trajetórias do grafo de estados do sistema, e apresentar os resultados obtidos em cada uma delas. Realmente, se a produção aplicada em algum desses estados for:

```

If presença_do_fator_reumático e
      o_paciente_apresenta_dor_nas_articulações
then
      . 80%_de_probabilidade_de_ser_artrite_reumática

```

então qualquer outro raciocínio baseado no fato de a doença ser artrite reumática tem algum potencial de ser inválido. A exemplo do que ocorre em outras áreas de pesquisa, as primeiras aplicações de sistemas especialistas ficaram restritas a algumas áreas somente. Atualmente são inúmeras as áreas em que os sistemas especialistas encontraram aplicação. Um catalisador para esta expansão foi a conscientização de que um sistema especialista pode ser separado em duas componentes: a de raciocínio e a de conhecimento. Removendo a base de conhecimento de um sistema especialista existente, tem-se um sistema de rotinas de raciocínio que pode ser, provavelmente, muito bem aplicado para outras situações. Sistemas especialistas novos em outras áreas podem ser, portanto, construídos simplesmente anexando uma nova base de conhecimento a este sistema de raciocínio já existente. Nos anos 50, os pesquisadores já haviam estabelecido os fundamentos da Inteligência Artificial, incluindo lógica matemática e Teoria das Funções Recursivas, guiando a formulação de processamento de listas e da própria linguagem *LISP* que fornece um interpretador para desenvolver expressões simbólicas recursivas. Tais capacidades suportaram o surgimento de sistemas práticos de computação simbólica. Ao mesmo tempo, emergiram computadores interativos tornando possíveis ambientes computacionais para desenvolvimento e depuração de programas incrementais. Aproximadamente nesta mesma ocasião, psicólogos cognitivos - estudantes da forma de pensar humana - criaram caminhos padrão do processo de investigação do raciocínio, modelando o aparente processo de tomada de decisão em termos de regras de produção

condicionais. Nos anos 60, os pesquisadores de Inteligência Artificial tentaram simular o complexo processo do pensamento procurando métodos gerais para resolver uma ampla classe de problemas. No entanto, a despeito de alguns progressos interessantes, as dificuldades eram enormes e não frutificaram. Então durante a década de 70 concentraram esforços em técnicas como *Representação* - isto é, modo de formular o problema de maneira a tornar sua solução mais fácil como controlá-la inteligentemente dentro da capacidade de memória do computador. Esta estratégia produziu algum sucesso mas ainda não foi decisiva. Somente no final da década fizeram a descoberta mais importante: o poder do programa em resolver problemas depende mais do conhecimento que possui do que do formalismo ou esquema de inferência empregado. Esta realização levou ao desenvolvimento de programas de computador de propósito particular, sistemas que são peritos em alguma área limitada. Estes programas são chamados Sistemas Especialistas e um novo campo se iniciou. No campo de IA aplicada, após uma década de trabalho, emergiram três subcampos principais: Sistemas Especialistas, Linguagem Natural e *Robótica* - incluindo *visão, fala* e locomoção. Os Sistemas Especialistas começaram a surgir comercialmente entre 1980 e 1981. A primeira companhia formada exclusivamente para produzir sistemas especialistas foi a *Intelli Genetics*, no campo de engenharia genética e com técnicos oriundos do Projeto de Programação Heurística da Universidade de Stanford. Assim os Sistemas Especialistas após terem se desenvolvido por 15 anos como mera curiosidade de IA aplicada em laboratórios de pesquisa, tornaram-se alvo de significativos esforços de desenvolvimento, tanto técnicos como comerciais. Estes sistemas empregam computador de maneira diversa que o processamento de dados convencional abrindo novas e importantes oportunidades. Recentemente, muitas organizações têm explorado esta tecnologia, ampliando suas pesquisas e começando a adaptar suas atividades para tanto. Embora Sistemas Especialistas e Peritos Reais possam em alguns casos desempenhar tarefas idênticas, as características de ambos são criticamente diversas. Mesmo havendo algumas vantagens evidentes dos Sistemas Especialistas eles não poderão substituir os peritos em todas as situações devido a algumas limitações inerentes. A tabela seguinte resume esta comparação: (BROOKSHEAR, 2000)

Tabela 1: Comparação entre conhecimento especializado humano e artificial.

Conhecimento Humano	Conhecimento Artificial
Perecível	Permanente
Difícil de transferir	Fácil de transferir
Difícil de documentar	Fácil de documentar
Imprevisível	Consistente
Caro	Razoável
Discriminatório	Imparcial
Social	Individualizado
Criativo	Sem inspiração
Adaptável	Inflexível
Enfoque amplo	Enfoque restrito
Baseado em senso comum	Técnico

Um Sistema Especialista pode ser comparado com um programa gravado em videocassete enquanto que neste contexto um especialista se compararia com um programa ao vivo. Isto é, a utilização e reprodução, que compensa de sobra seu alto custo de implantação, pode levar vários anos. Os especialistas humanos, especialmente os melhores, tem altos salários, enquanto que Sistemas Especialistas tem o custo nominal do computador rodar o programa. Por outro lado, para demonstrar que os especialistas não podem ser simplesmente substituídos, o conhecimento artificial tem atualmente algumas limitações. Uma delas é falta de criatividade. Um perito pode reorganizar informações e usá-las para sintetizar novos conhecimentos; pode manusear eventos inesperados usando imaginação ou novas abordagens, inclusive raciocínio por analogia de um outro domínio completamente diferente. Sistemas Especialistas trabalham de maneira sem inspiração, rotineiramente. Finalmente, os seres humanos - peritos ou não - possuem o conhecimento advindo do senso comum, que se constitui num largo espectro de conhecimento sobre o mundo, acumulado durante toda a sua vida e que permeia todas as suas decisões. Devido a enorme quantidade de conhecimento de senso comum, fatos de domínio público que todos sabem, torna-se difícil construir um programa inteligente, particularmente um Sistema Especialista. Como um exemplo deste tipo de conhecimento, suponha que você examina um registro médico de um paciente que pese 7 quilos e tenha idade de 25 anos. Imediatamente você suspeitará de erro nos dados, não porque isoladamente sejam impossíveis, mas sua combinação virtualmente o será; desconfiará que os dados foram acidentalmente invertidos. Um Sistema Especialista não detectaria este tipo de erro, a menos que disponha de um tabela idade/peso (e seus limites de tolerância) para

conferi-lo. Senso comum inclui conhecimento sobre o que não se sabe assim como o que se conhece. Por exemplo, se você for perguntado como a televisão noticiou um fato ocorrido no ano passado, puxará pela memória para recuperar a informação. Se a pergunta fosse "como a TV soviética noticiou outro fato", imediatamente responderá que não sabe. Da mesma forma, se lhe perguntarem "como a TV brasileira informou a Proclamação da República", você dirá que tal não ocorreu pois não havia televisão naquela época. Quando um Sistema Especialista é questionado sobre informações que não disponha ou não existam, ele não detecta esta situação por não possuir senso comum. Então, iniciará exaustiva pesquisa nos seus fatos e regras para encontrar a solução. Daí, quando esta não for encontrada, pode julgar que é porque seu conhecimento está incompleto e solicita informação adicional para completar sua base de conhecimento. Por estas razões, Sistemas Especialistas são freqüentemente utilizados no aconselhamento, como consultor ou ajuda para outro perito ou iniciante usar em algum problema. Em contraste com sistemas avançados de processamento de dados, que automatizam algorítmicamente extensos volumes de dados, os Sistemas Especialistas ordinariamente trabalham com pequenas tarefas tipicamente desempenhadas por profissionais: *interpretando, diagnosticando, planejando, escalonando* e assim por diante. Para acompanhamento destas tarefas, os Sistemas Especialistas utilizam judiciosamente os dados e raciocínios com eles. Ao contrário do método usual em processamento de dados, os Sistemas Especialistas geralmente examinam uma extensa gama de possibilidades ou constróem soluções dinamicamente. (SEM, 1999)

4.2. NOTAS SOBRE SISTEMAS ESPECIALISTAS

Basicamente, de acordo com (CRIPPA, 2000) um Sistema Especialista (SE) "é um programa inteligente de computador que usa conhecimentos e procedimentos inferenciais, para resolver problemas que são bastante difíceis, de forma a requererem para sua solução, muita perícia humana. O conhecimento necessário para atuar a esse nível, mais os procedimentos inferenciais empregados, pode considerar-se um modelo da perícia aos melhores profissionais do ramo. O conhecimento de um sistema especialista consiste em fatos e heurísticas. Os fatos constituem um corpo de informação que é largamente compartilhado, publicamente disponível e geralmente aceito pelos especialistas em um campo. As heurísticas são, em sua maioria privadas, regras poucos discutidas, de bom discernimento (regras de raciocínio plausível, regras de boa conjectura), que caracterizam a tomada de decisão a nível de especialista na área. O nível de desempenho de um sistema especialista é função

principalmente do tamanho e da qualidade da base de conhecimento que possui." Para que um sistema seja considerado especialista, alguns componentes são essenciais à sua caracterização, são eles:

- uma *Linguagem* de expressão dos conhecimentos fornecidos pelos especialistas
- uma *Base de Conhecimentos*, para armazenar o conhecimento específico de determinada aplicação, que pode ser diretamente fornecido por um especialista, ou acumulado pelo sistema ao fim dos experimentos
- um *Motor de Inferência*, programa relativamente geral que explora o conhecimento da base precedente, considerando-a como fonte de informações (Assim suscetível a mudanças).

Os SE são programas de Inteligência Artificial (IA) que capacitam um computador a auxiliar pessoas num processo de tomada de decisão. O know-how do perito humano é utilizado para instruir o computador a resolver um problema ou a tomar uma decisão. A máquina pode, então, auxiliar ou aconselhar outros usuários na resolução do mesmo problema por meio:

- do conhecimento especializado do usuário humano;
- da habilidade do computador para armazenar grandes quantidade de textos e dados e considerar todas as possibilidades a alta velocidade.

Assim, (CRIPPA, 2000) encerra: "o computador interage com o usuário num esquema de perguntas e respostas do usuário e chega a uma conclusão baseado nas respostas. O usuário pode perguntar ao computador por que ele quer certa informação e o SE explicará sua necessidade dos dados e como eles serão utilizados. O mais importante é que o SE deveria dizer como chegou a suas conclusões; não dará somente conselhos, mas também justificará a opinião que oferece." Os SE são também conhecidos por sistemas cognitivos, e a tecnologia utilizada para seu desenvolvimento chamada de engenharia do conhecimento. Ainda, alguns autores costumam chamar de sistemas cognitivos apenas os sistemas de pequeno porte.

Em (CRIPPA, 2000) discute-se a definição da engenharia do conhecimento pela definição da própria palavra engenharia: "A engenharia comporta os métodos para estudo de um projeto e realização de um trabalho complexo por equipes diversas, de especialidades complementares." A partir daí ele continua que "desenvolver um SE é aplicar uma "engenharia" do conhecimento: a construção de uma informática lógica, refere-se à fatos e idéias e a capacidade de tratá-los, os quais receberá do exterior, para deduzir um diagnóstico, uma precognição, ou, de fato, uma decisão."

Segundo (SIS, 1999), um Sistema Especialista é aquele que é projetado e desenvolvido para atender a uma aplicação determinada e limitada do conhecimento humano. É capaz de emitir uma decisão, apoiado em conhecimento justificado, a partir de uma base de

informações, tal qual um especialista de determinada área do conhecimento humano. Para tomar uma decisão sobre um determinado assunto, um especialista o faz a partir de fatos que encontra e de hipóteses que formula, buscando em sua memória um conhecimento prévio armazenado durante anos, no período de sua formação e no decorrer de sua vida profissional, sobre esses fatos e hipóteses. E o faz de acordo com a sua experiência, isto é, com o seu conhecimento acumulado sobre o assunto e, com esses fatos e hipóteses, emite a decisão. Durante o processo de raciocínio, vai verificando qual a importância dos fatos que encontra comparando-os com as informações já contidas no seu conhecimento acumulado sobre esses fatos e hipóteses. Neste processo, vai formulando novas hipóteses e verificando novos fatos; e esses novos fatos vão influenciar no processo de raciocínio. Este raciocínio é sempre baseado no conhecimento prévio acumulado. Um especialista com esse processo de raciocínio pode não chegar a uma decisão se os fatos de que dispõe para aplicar o seu conhecimento prévio não forem suficientes. Pode, por este motivo, inclusive chegar a uma conclusão errada; mas este erro é justificado em função dos fatos que encontrou e do seu conhecimento acumulado previamente. Um Sistema Especialista deve, além de inferir conclusões, ter capacidade de aprender novos conhecimentos e, desse modo, melhorar o seu desempenho de raciocínio, e a qualidade de suas decisões.

Conforme (SIS, 1999), de um modo geral, sempre que um problema não pode ser algoritmizado, ou sua solução conduza a um processamento muito demorado, os Sistemas Especialistas podem ser uma saída, pois possuem o seu mecanismo apoiado em processos heurísticos. Um Sistema Especialista deve preservar e transmitir o conhecimento de um especialista humano em uma determinada área. Não pode ser influenciado por elementos externos a ele, como ocorre com o especialista humano, para as mesmas condições deverá fornecer sempre o mesmo conjunto de decisões.

4.3. BASE DE CONHECIMENTOS

De acordo com (SISTEMAS, 2000) um elemento chave em qualquer aplicação de inteligência artificial é o conhecimento. Conhecimento é também informação que foi organizada e analisada de modo a torná-la compreensível e aplicável à solução de problemas e tomada de decisões. O sistema de IA não é baseado em processo algorítmico, mas sim em representação e manipulação simbólica. Usando símbolos é possível criar uma base de conhecimentos que estabelece fatos, conceitos e os relaciona entre eles. O processo é qualitativo e não quantitativo como num algoritmo computacional convencional típico. A base de conhecimento contém todos os fatos, idéias, relacionamentos e interações de um

domínio limitado. A máquina de inferência analisa o conhecimento e deriva conclusões. O interface com o usuário permite que novo conhecimento seja apropriado e implementa a comunicação com o usuário. Existem muitos métodos para descrever o conhecimento e espelham os diferentes modos de abordagem:

- Redes Semânticas descrevem relacionamentos entre objetos;
- Regras de Produção descrevem o conhecimento em termos de regras que reagem aos eventos detectados;
- *Frames* são úteis para o tipo de problema que requer conjuntos de informações a serem coletadas;
- Taxonomias podem descrever conjuntos de informação como nos *Frames* e podem descrever como a informação é interrelacionada, como nas Redes Semânticas.

Uma vez que a arquitetura da rede contém uma hierarquia de componentes, a melhor maneira de representar a descrição lógica e física de rede é com uma hierarquia de *Frames*. Dessa maneira, a Taxonomia também pode ser empregada, pois ela descreve as informações como os *Frames*. Para cada situação haverá uma forma mais ou menos adequada para representar o conhecimento. O analista deve estruturar cuidadosamente como o conhecimento do sistema será representado.

Segundo (SEM, 1999), para solucionar problemas os Sistemas Especialistas precisam acessar uma grande base de conhecimento do domínio da aplicação, portanto o sucesso de um Sistema Especialista depende enormemente da forma de como o conhecimento é representado e dos mecanismos para a exploração deste conhecimento. Um Sistema Especialista é basicamente formado por três componentes:

- "**Base de dados**": *base de conhecimento* - descreve a situação corrente e a meta;
- "**Conjunto de operadores**": *mecanismos para exploração* - unidades que operam sobre a base de dados;
- "**Estratégia de controle**": *Raciocínio Para Frente (bottom-up)* - aplicação de operadores sobre as estruturas na base de dados que descreve a situação do domínio de uma tarefa para produzir uma situação modificada, e *Raciocínio Para Trás (top-down)* - aplicação de operadores sobre as metas para reduzi-las às submetas. A combinação dos dois raciocínios também é possível.

Uma importante técnica de IA envolvendo botton-up e top-down é chamada de Análise Significado-Final (means-end), que faz a comparação da situação de domínio da tarefa corrente com a meta corrente para extrair a diferença entre elas. A diferença é então usada

para indexar o operador mais relevante a fim de reduzir a diferença. Caso estes especialmente relevantes operadores não puderem ser aplicados para o presente estado do problema, submetas são aplicadas para alterar o estado até que um operador relevante possa ser aplicado. Depois que estas submetas são solucionadas, o operador relevante é aplicado produzindo resultado, a situação modificada vem a ser um novo ponto inicial para solucionar a meta original.

4.4. INTELIGÊNCIA ARTIFICIAL DISTRIBUÍDA

(RICH, 1993) define um *sistema de raciocínio distribuído* como aquele que é composto por um conjunto de módulos separados (em geral chamados de agentes, já que cada módulo assume o papel de uma identidade de solução de problemas) e por um conjunto de caminhos de comunicação entre eles. Esta definição é intencionalmente muito vaga. Ela admite sistemas em qualquer ponto do espectro, que vai dos sistemas intimamente ligados, em que há um mecanismo de controle completamente centralizado e uma base de conhecimento compartilhada, até aqueles em que tanto o controle quanto o conhecimento são totalmente distribuídos. Na verdade, a maioria dos sistemas de raciocínio distribuído encontra-se no meio desse espectro. Sistemas de Inteligência Artificial Distribuída são a classe de sistemas que permite a vários processos autônomos, chamados agentes, realizarem atos de inteligência global, somente com processamento local e comunicação interprocessos. Enfim, a meta de sistemas especialistas baseados em conhecimento e distribuídos é poder coordenar atividades de grupo resolvendo problemas que os sistemas compartilham - totalmente ou em parte. Percebe-se então que a IAD não pode ser considerada um subcampo da IA, mas sim um componente essencial na existência de um sistema de IA. Os benefícios da IAD são similares aos benefícios derivados de ter um grupo de pessoas trabalhando juntas para resolver problemas. Problemas que são muito extensos para um só especialista resolver, pode ser resolvido por um grupo. Os problemas podem ser resolvidos em partes por pessoas trabalhando em paralelo. Existem muitas razões para que seja empregado o conceito de Inteligência Artificial Distribuída. Entre elas pode-se ressaltar as que segue:

- Mais Poder de Computação ou *Hardware* mais barato. Esta é uma das principais razões para se aplicar as funcionalidades distribuídas em um sistema: o alto custo do *hardware* necessário para resolver o problema de forma centralizada, que nem sempre é rápido o suficiente para tal tarefa.
- Maior Segurança e Tolerância a Falhas. Se você não está confiando ou não pode confiar a resolução de problemas a um único processo, nada mais justo

que encaminhar o problema a um grupo de *especialistas*. Tendo vários processos especialistas, quando um problema surge, o problema pode ser colocado em evidência, sendo analisado por vários sistemas, paralelamente. Como consequência disso, têm-se várias opiniões - e não apenas uma, sobre qual atitude deve-se tomar. Como exemplo de uma situação real do emprego da IAD, têm-se o sistema de defesa dos E.U.A.. O sistema é composto de mais de um sistema especialista, todos com a mesma função, contudo com formas diferenciadas de armazenar e decidir sobre os problemas. Dessa maneira, para uma resolução ser feita, mais de um sistema deve chegar na mesma solução, de formas diferentes e por métodos diferentes. O sistema é composto por um número ímpar de processos distribuídos para que não ocorra empate de decisões.

- Aproveitamento da Tecnologia existente. Com a IAD a base de conhecimentos pode ser subdividida em áreas, assim como o problema. A IAD estimula decomposição do problema total de gerenciamento, tornando a solução o somatório de módulos (sistemas).

Com a distribuição dos processos, surge o problema da padronização. A linguagem de comunicação entre os processos deve ser padronizada, como também a forma de representar o conhecimento - para que o mesmo seja trocado e tratado por processos especialistas distintos sem a perda de informação. Levando em conta que a informação deve ser padronizada, porque não pensar em uma base de dados comum aos processos. Compartilhando a base de conhecimento, elimina-se o problema da padronização das informações. Tomada a decisão de compartilhamento da base de conhecimentos, pode-se diminuir sensivelmente a comunicação interprocessos, fazendo com que cada processo busque na base de conhecimentos um padrão de dados que ele possa manipular. Dessa maneira, cada processo procura um problema, resolve a sua parte e devolve os dados trabalhados à base, para que um outro processo identifique - então, o seu padrão e possa seguir manipulando os mesmos dados, adiantando mais a resolução do problema em questão. Uma das arquiteturas mais predominantes na área de Inteligência Artificial Distribuída é a *quadro-negro*. Seus principais componentes são os seguintes: (RICH, 1993)

- *quadro-negro*, é um repositório onde está armazenado o conhecimento do problema e o estado de sua solução;

- *fontes de conhecimento*, são processos simples que procuram no *quadro-negro* por um padrão de dados e tomar alguma atitude sobre eles, adiantando a solução do problema;
- a *agenda*, que avalia a lista de *fontes de conhecimento* pronta para rodar e seleciona um para executar.

Desde que a arquitetura *quadro-negro* possui um raciocínio inerentemente distribuído, ela trabalha bem em ambiente onde os dados são passados entre agentes independentes, coordenando a atividade e a resolução parcial de problemas compartilhados. Ao se projetar um sistema especialista distribuído, devem ser feitas algumas perguntas básicas como O que um agente irá dizer ao outro?, Que linguagem eles usarão?, O que eles devem descrever?. É muito importante, como em qualquer projeto, que seja feita uma excelente análise do problema para que o custo de uma reestruturação seja poupado. (SISTEMAS, 2000)

Luis Alberto Gómez e João Bosco Alves ressaltam em (GÓMEZ, 2000) a motivação para o uso de *Agentes Inteligentes* num projeto de Controladores Inteligentes para Edifícios. Em sua obra, os autores falam da evolução do funcionamento e integração dos objetos ou componentes que incorporam uma casa, desde a década de 70 onde as pessoas buscavam acomodação trazendo diversos eletrodomésticos para suas casas, tal período é definido como “*Casa Consumo*”. Nos anos 80 com a chegada de TV a cabo e antenas por satélites define-se “*Casa Comunicação*”. Nos anos 90 a idéia é a integração de sistemas de segurança, condicionamento ambiental caracterizando “*Casa Sistema*” e, (GÓMEZ, 2000) coloca que “*no futuro podemos prever a Casa Humana, onde o sistema casa consegue entender a voz e a linguagem corporal dos habitantes e reage com os mesmos, mudando a iluminação por exemplo*”.

Estas características podem ser implementadas com o uso de IAD e controladores (agentes) compartilhando informações.

4.4.1. SISTEMAS BASEADOS EM CONHECIMENTO

(RICH, 1993) coloca que quando é construído um sistema voltado para uma área muito específica - absorvendo a experiência de um especialista humano, ele é denominado Sistema Especialista. Em sistemas algorítmicos, o conhecimento é armazenado na forma de código de programa. Se precisar aumentar o conhecimento do sistema, mais código deverá ser criado. Os comandos de decisão formam o conhecimento do sistema. Existe um outro método para armazenar esse conhecimento que o sistema possui - essa experiência. Quando o método

utilizado para armazenar o conhecimento é uma base de conhecimentos, o sistema é chamado de Sistema Especialista Baseado em Conhecimento. Dessa forma, o conhecimento é manipulado como uma base de dados - pois é isto na verdade, podendo ser alterada sem alterar o código do sistema. O sistema pode adquirir conhecimentos e esquecer-se se assim for necessário sem maiores complicações. Um Sistema Baseado em Conhecimento difere de Sistemas Algorítmicos comuns pelo fato de que a solução do problema esta na forma de tratar o conhecimento, que em sistemas comuns, o problema é como implementar tal conhecimento. Para ser mais claro, em um Sistema Baseado em Conhecimento, os conhecimentos são armazenados de forma tal que o sistema define o que significa aquele conhecimento, como ele deve ser interpretado. Em Sistemas Algorítmicos, o programador deve construir rotinas para manipular o conhecimento em seus mínimos detalhes, que são dispensadas em sistemas de IA. Em sistemas de IA, o programador preocupa-se em construir rotinas de inferência da base de conhecimentos, rotinas de questionamentos. É como usar a linguagem SQL em banco de dados, o programador preocupa-se em especificar o que deve ser buscado, não interessando como o banco de dados fará a busca na base de dados. Nos últimos anos tem surgido vários Sistemas Especialistas Baseados em Conhecimento dirigidos a várias áreas. No que tange ao gerenciamento de redes de computadores, algumas categorias de sistemas capturam e interpretam, diagnosticam e predizem os dados, monitorando, planejando e controlando com perícia os dispositivos. Sistemas de interpretação e diagnóstico envolve sistemas de inferência e sensores de dados. Estes sistemas são usados para avaliar o desempenho da rede, isolamento de falhas e recuperação de uma queda pelas informações dos sensores de observação e alarmes. Sistemas de monitoração envolvem comparações observáveis para prever e levantar hipóteses sobre o estado interno da rede usando sensores de dados e fazendo a devida interpretação dos mesmos. A monitoração *on-line* é usada para medir o desempenho da rede e planejar as ações corretiva, se necessárias (por exemplo: re-roteamento de tráfego por queda de algum roteador ou congestionamento do mesmo). Planejamento requer engenhosas ações para completar os resultados desejados. O planejamento de reações rápidas pode envolver uma nova estratégia de roteamento ou geração de novas tabelas de roteamento para ajustar o tráfego. Um sistema especialista de planejamento pode também prover auxílio para planejamento de capacidade, expansão e integração de redes de computadores. A atividade de *design* envolve a configuração de elementos da rede. Ela pode ser iniciada por um pedido de serviço de usuário para projetar ou expandir uma rede. O *design* pode requerer passos físicos, lógicos e ações. Predizer é antecipar uma dada situação. Uma situação pode ser prevista com a história dos dados anteriores e com a direção tendenciosa da performance da rede. O controle envolve as ações necessárias para levar o sistema para o estado desejado. A ação de controle

pode ser iniciada com uma ação de interpretação de sensores de dados e serve para levar a rede a um estado aceitável. Ações de controle incluem redirecionamentos de rota e outras reconfigurações diversas. Existem sistemas especialistas voltados a cada uma das funções descritas acima. Existem também sistemas mais complexos que englobam mais de uma das áreas de atuação citada, mesmo porque algumas delas são interdependentes. Também é possível, ao invés de termos sistemas especialistas enormes, termos sistemas especialistas que se comuniquem uns com os outros, cada um com a sua função. Cada sistema é responsável por parte do gerenciamento da rede, comunicando-se com os outros para atingir tal meta. Esse é o conceito de Inteligência Artificial Distribuída. É muito importante o processo de escolha da metodologia correta para abordar o problema. Não basta apenas escolher o que parece ser mais eficaz. Deve-se pesar as medidas com os devidos valores. A decisão entre um sistema especialista baseado em conhecimento distribuído e não-distribuído deve levar em consideração o problema - se exige realmente um sistema distribuído, o *hardware* - veloz e caro no caso de um sistema não-distribuído e complexo, e muitos outros fatores particulares da situação que devem ser analisados com cautela. Sistemas Especialistas fornecem benefícios inegáveis na automação de gerência. Entre eles pode-se destacar os seguintes:

- Provê alta qualidade de Serviço.
 - Substancial aumento de desempenho. Sistemas especialistas são usados para fazer roteamento - de pacotes, e manutenção em redes de computadores.
 - Captura conhecimentos que serão utilizados no futuro. A experiência dos especialistas são perdidas com o tempo, no momento em que eles se especializam em outros problemas ou trocam de trabalho. Um sistema especialista nunca esquece um procedimento, nem mesmo em seus mínimos detalhes.
 - Possui a mesma velocidade e dinâmica de um especialista vinte e quatro (24) horas por dia, durante todos dias.
- Provê baixo custo do serviço pela produtividade adquirida. Isto melhora o desempenho da rede por inteiro.
 - Significativo aumento da produtividade. Um sistema especialista pode levar uma pessoa de pouca experiência para o nível de um especialista, executando tarefas mais rapidamente e mais facilmente.
 - Reduz os custos com funcionários especializados. Com sistemas especialistas uma tarefa de grande complexidade pode ser realizada por funcionários que não sejam *experts*, mas funcionários especializados apenas. Com isso, os

especialistas podem ficar dedicados a tarefas que realmente sejam de altíssima complexidade.

- Sistemas especialistas podem, as vezes, fazer melhor que um especialista. Sistemas especialistas podem combinar a capacidade de um especialista com a velocidade e precisão de uma máquina. Eles podem fazer uma análise detalhada e completa de uma situação que dificilmente um humano seria capaz de fazer. Seguindo, se um sistema foi bem estruturado, ele não cometerá erros que eventualmente um humano pode cometer. Finalmente, um sistema especialista seria capaz de fazer tarefas detalhadas e repetitivas - que um especialista talvez as desprezasse, em uma situação em que as chances de resolver o problema são muito pequenas .
- Podem oferecer ajuda no treinamento de pessoal. Os sistemas especialistas podem fornecer informações detalhadas - de diversos níveis, sobre como foi procedido para solucionar um determinado problema. Esta pode ser a maior ajuda no treinamento de pessoas inexperientes. Em adição, existem sistemas especialistas construídos especialmente para prover treinamento. Neste caso, até didática eles possuem.

Os sistemas especialistas também podem gerar problemas. Com a introdução da automação, os trabalhadores podem se sentir inseguros - eles podem achar que irão perder o emprego (e isto realmente acontece). Esta situação pode ser evitada com uma migração lenta e responsável do sistema humano para o sistema máquina. (SIS, 1999)

Um outro problema pode surgir quanto a política. Sistemas especialistas não possuem senso de política, podendo produzir resultados embaraçosos a alguém . Portanto, nessas situações, é necessário subordinar o sistema a um analista humano, que interage com ele e desvia a situação se necessário. (SEM, 1999)

4.4.2. SISTEMAS DE PRODUÇÃO

Um sistema de produções consiste de três componentes principais:

1. **Um conjunto de estados:** Cada estado é uma situação que pode ocorrer no ambiente da aplicação. O estado a partir do qual o ambiente é iniciado é chamado de *estado inicial*. O(s) estado(s) é (são) chamado(s) *estado(s)-meta*, ou seja, o estado inicial é a configuração de um quebra-cabeças a ser fornecido à máquina, e o estado-meta é a configuração do quebra-cabeças resolvido.

2. **Um conjunto de produções (ou regras):** Uma produção é uma operação que pode ser executada no ambiente da aplicação para migrar de um estado para outro. Cada produção pode estar associada a precondições. Em outras palavras, as precondições são aquelas condições que devem obrigatoriamente estar presentes no ambiente para que a produção possa ser aplicada.
3. **Um sistema de controle:** O sistema de controle consiste da lógica necessária para conduzir um quebra-cabeças do seu estado inicial para o estado-meta. A cada passo do processo, o sistema de controle deve decidir qual das produções, cujas precondições estejam satisfeitas, deve ser aplicada a seguir.

Do ponto de vista de sistemas de produções, a tarefa de desenvolver uma máquina inteligente consiste em implementar o sistema de controle na forma de um programa armazenado na máquina. Este programa inspeciona o estado corrente do sistema desejado, identifica uma seqüência de produções que conduz ao estado-meta e executa tal seqüência. Para essa finalidade, o sistema de controle constrói um algoritmo para resolver o problema inicial, utilizando as produções como blocos de construção. A principal dificuldade para projetar uma máquina para a resolução de um “quebra-cabeças” é o desenvolvimento desses programas de controle. (BROOKSHEAR, 2000)

Conforme (SEM, 1999), os sistemas de produção são uma boa maneira de descrever as operações que podem ser realizadas quando se busca uma solução para um problema. Uma vez que a busca forma o núcleo de muitos processos inteligentes, é útil estruturar os programas de IA de uma forma que facilite a descrição e execução dos processos de busca. Os Sistemas de Produção (SP) proporcionam tais estruturas. Um conjunto de regras, cada uma delas consistindo em um lado esquerdo (um padrão) que determina a aplicabilidade de uma regra e um lado direito que descreve a operação a ser efetuada se a regra for aplicada. Uma ou mais bases de conhecimentos/bases de dados que contenham quaisquer informações apropriadas a uma determinada tarefa. Certas partes da base de dados podem ser permanentes, enquanto que outras, podem pertencer apenas à solução do problema corrente. Uma estratégia de controle que especifique a ordem em que as regras serão comparadas com a base de dados e uma maneira de solucionar conflitos que surgirem quando várias regras puderem ser aplicadas ao mesmo tempo. Um aplicador de regras. Esta definição é bastante genérica, incluindo entre outros os programas para jogar xadrez, resolver o problema das jarras d'água, do fazendeiro, isto é, o processo de solução destes problemas foram modelados como um

sistema de produção. Além disso, inclui também uma família de interpretadores genéricos de sistemas de produção, tais como:

- Linguagens de sistemas de produção básicos, como o OPS5 e o ACT.
- Sistemas mais complexos, em geral híbridos, chamados de shells de sistemas especialistas, que proporcionam ambientes mais complexos para construção de sistemas especialistas baseados em conhecimento.
- Arquiteturas para a solução de problemas gerais como o SOAR, um sistema baseado em um conjunto específico de hipóteses motivadas cognitivamente sobre a natureza da solução dos problemas.

Todos estes sistemas oferecem a arquitetura geral de um sistema de produção e permitem que o programador escreva regras que definam determinados problemas a serem solucionados. Considerando as características de um Sistema Especialista, pode-se formular duas perguntas básicas sobre os Sistemas de Produção:

- Eles podem ser descritos por um grupo de características que coloquem alguma luz sobre: *Como ou quão facilmente eles podem ser implementados?*
- Considere as classes de SP:
 - **SP MONOTÔNICOS:** a aplicação de uma regra NUNCA impede a aplicação posterior de uma outra regra que também pudesse ter sido aplicada quando a primeira regra foi escolhida;
 - **SP NÃO-MONOTÔNICOS:** a aplicação de uma regra PODE impedir a aplicação posterior de uma regra que também, pudesse ter sido aplicada quando a primeira regra foi escolhida;
 - **SP PARCIALMENTE COMUTATIVO:** se a aplicação de uma determinada seqüência de regras transforma um estado x em um estado y, então qualquer permutação permitida dessas regras também transforma o estado x em estado y;
 - **SP COMUTATIVO:** é monotônico e parcialmente comutativo simultaneamente.

Pode-se dizer que: para qualquer problema solucionável, existe um número infinito de SP que descrevem maneiras de encontrar soluções. Alguns serão mais NATURAIS ou mais EFICIENTES que outros. Qualquer problema que possa ser solucionado por qualquer SP, pode ser solucionado por um SISTEMA COMUTATIVO (classe mais restrita), mas o sistema comutativo pode ter o manejo tão complicado que se torna praticamente inútil. Em um

sentido mais formal: "Não existe relação entre tipos de problemas e tipos de SP, já que todos os problemas podem ser solucionados por todos os tipos de sistemas." Em termos práticos: "Existe definitivamente uma relação entre tipos de problemas e tipos de SP que se prestam naturalmente à descrição desses problemas." Os *SP PARCIALMENTE COMUTATIVOS MONOTÔNICOS* são importantes, do ponto de vista da implementação, porque podem ser implementados sem que haja necessidade de retrocesso a estágios anteriores, quando se descobre que um caminho errado foi seguido. Embora normalmente seja útil implementar tais sistemas com o recurso do retrocesso, a fim de garantir uma busca sistemática, a base de dados, que na verdade representa o estado do problema, não precisa ser restaurada, o que se traduz em eficiência. Os SP COMUTATIVOS são bons para problemas nos quais as coisas não mudam e coisas novas são criadas. Isto é, ao invés de alterar coisas antigas envolvem a criação de coisas novas. Os *SP PARCIALMENTE COMUTATIVOS NÃO-MONOTÔNICOS* são úteis para problemas nos quais aconteçam mudanças que possam ser revertidas e nos quais a ordem das operações não seja crucial. A navegação por robôs em terreno liso e sem obstáculos, um problema de manipulação física, é um exemplo. Navegação significa mover-se pelo mundo: Planejar rotas, chegar aos destinos desejados sem trombar com outras coisas no meio do caminho, e assim por diante. Como a visão e o reconhecimento da fala, esta é uma tarefa que os seres humanos executam com facilidade. Suponha que um robô tenha os seguintes operadores: ir para o norte (N), ir para o leste (L), ir para o oeste (O), ir para o sul (S). Para alcançar seu objetivo, não importa se o robô executa N-N-L ou N-L-N. Os *SP NÃO-PARCIALMENTE COMUTATIVOS* têm menos chances de produzir o mesmo nó muitas vezes durante o processo de busca. Quando se lida com aqueles que descrevem processos irreversíveis, é particularmente importante tomar decisões corretas logo na primeira vez, apesar de que, se o universo for previsível, o planejamento poderá ser usado para fazer com a tomada de decisão tenha papel menos importante.

5. SISTEMAS ESPECIALISTAS PROBABILÍSTICOS

5.1. INTRODUÇÃO

Os sistemas especialistas são desenvolvidos para resolver uma grande variedade de problemas, sendo que o conhecimento necessário para solucionar estes problemas, nem sempre é precisamente definido. Pode-se constatar que muitos problemas na vida real são repletos de incerteza. Especialistas humanos nestes domínios são capazes de formar julgamentos e tomar decisões baseados em informação incerta e incompleta. Para que um sistema especialista 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 de grande interesse em Inteligência Artificial e muitas pesquisas vem sendo desenvolvidas nesta área. Desde a metade da década 80, a pesquisa sobre raciocínio probabilístico em sistemas especialistas resultou na introdução de Redes Bayesianas de Crenças. Estas redes têm sua origem na probabilidade e são caracterizadas por um poderoso formalismo que representa o conhecimento no domínio e, pela incerteza associada a este domínio. Mais especificamente, o formalismo proporciona uma representação concisa de uma distribuição conjunta de probabilidades em um grupo de variáveis aleatórias. Associados a este formalismo estão os algoritmos para computar eficientemente as probabilidades relevantes e para processar as evidências. Desde sua introdução, a estrutura de redes bayesianas vem rapidamente ganhando popularidade e agora começa a mostrar o seu valor em domínios complexos. Várias aplicações práticas estão sendo desenvolvidas, por exemplo, para diagnóstico médico, para recuperação de informação probabilística e para reconhecimento de imagens por programas computacionais. Como a teoria da probabilidade é uma teoria matematicamente bem fundamentada, possuindo uma longa e notável tradição de pesquisa, não é surpreendente que esta teoria conquiste lugar proeminente na pesquisa de raciocínio com incerteza em sistemas especialistas. Infelizmente, aplicar a teoria da probabilidade no contexto de um sistema especialista não é tão fácil quanto possa parecer à primeira vista. A aplicação direta dos conceitos básicos da teoria da probabilidade leva a problemas insuperáveis de complexidade computacional, isto é, a representação explícita de uma distribuição conjunta de probabilidade requer espaço exponencialmente proporcional ao número de variáveis envolvidas. Os Sistemas Especialistas Probabilísticos (SEP), também chamados de Sistemas Especialistas Bayesianos, têm em sua base de conhecimentos fatos e regras que representam o conhecimento do especialista num domínio de aplicação. Aos fatos e às regras é associada a

incerteza presente no domínio, e são explicitados a crença em sua ocorrência por meio de valores de probabilidade. O raciocínio realizado pelo sistema deve considerar estas probabilidades para a partir dos dados de entrada (input) associar um vetor de probabilidades ao conjunto de hipóteses diagnosticas (output). A hipótese com maior probabilidade de ocorrência pode ser considerada a conclusão do sistema, note que a esta conclusão está associado o grau de certeza da resposta do sistema. (NASSAR, 1998)

Segundo (RODDER, 1997), as vantagens dos Sistemas Especialistas Probabilísticos estão ligadas ao fato de poderem ser facilmente reproduzidos e serem objetivos. Dispondo-se de um destes sistemas pode-se, tirar tantas cópias quanto se queira e, enviarem para as localidades desejadas. O especialista artificial é mais disponível que o seu colega humano. O tempo de resposta destes sistemas está, de acordo com a complexidade do conhecimento, na casa dos segundos, ideal para a maioria das aplicações. Suas desvantagens estão ligadas aos domínios de conhecimento relativamente pequenos por eles cobertos, à sua ignorância de relações triviais e, na sua linguagem limitada de comunicação com o usuário. Um sistema especialista probabilístico para o diagnóstico de meningite só sabe algo sobre esta doença. Conhecimentos evidentes, relacionados com o assunto, que nós humanos aprendemos do nosso ambiente, o sistema não tem, a não ser que ele tenha sido programado para isto. A linguagem na qual o usuário e o sistema se comunicam está predefinida pelo programador do sistema. Desvios desta linguagem levam à irritações não compreensíveis pelo especialista artificial.

A fig. 1 exibe um esquema de representação dos componentes de um sistema especialista.

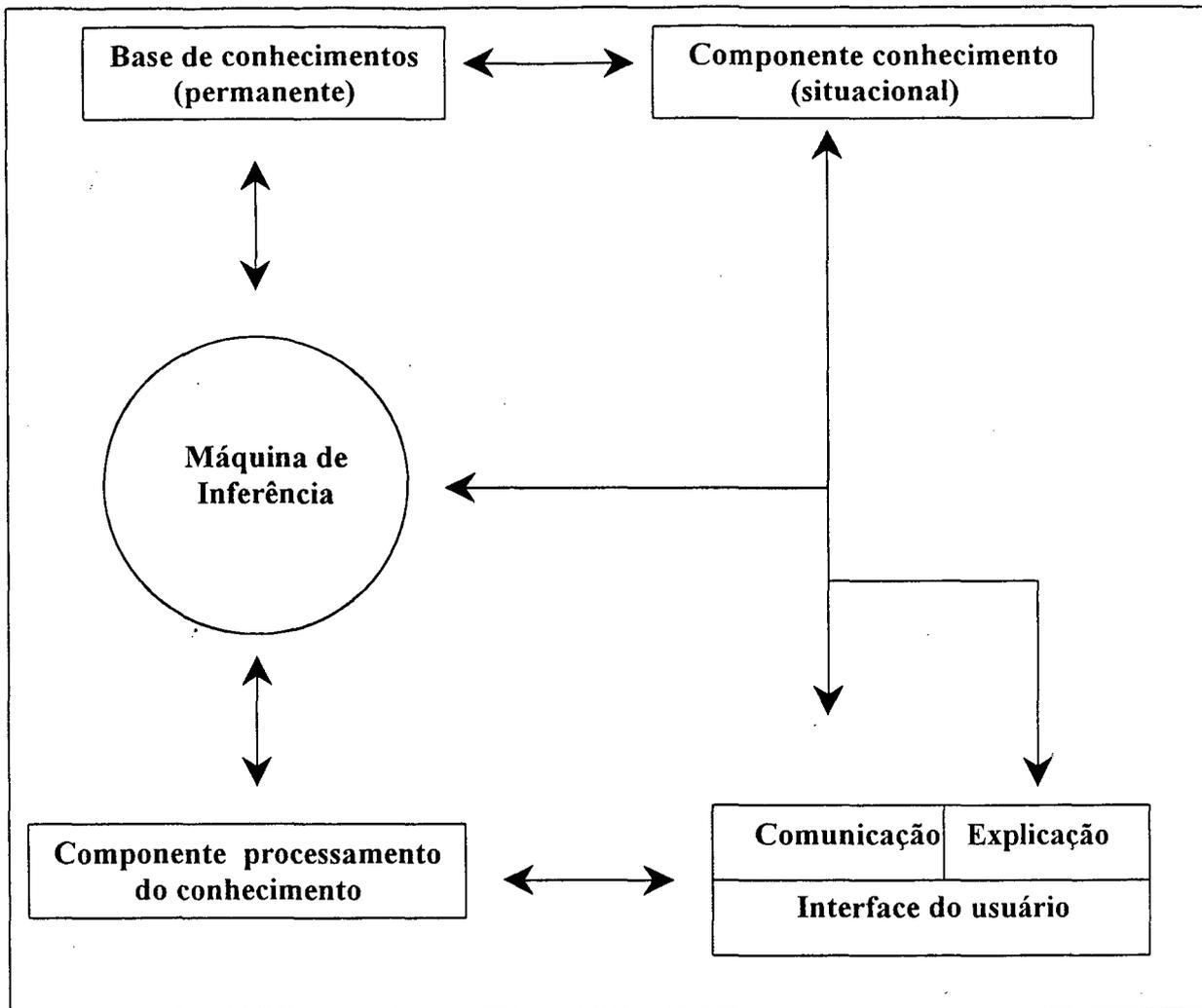


Figura 1: Componentes de um Sistema Especialista Probabilístico.

5.2. TEOREMA DE BAYES

A Teoria da Probabilidade

Segundo (NASSAR, 1998), a Teoria da Probabilidade refere-se a uma estrutura rigorosa de representação de eventos aleatórios. A probabilidade de um evento assume valor de 0 a 1. Aqui é considerado também a probabilidade de que ocorra um evento B (consequência) condicionada a ocorrência de um evento A (causa), esta probabilidade é suportada pelo teorema de Bayes.

Espaço de Probabilidade

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

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

O par (Π, P) é chamado de espaço de probabilidade. Os elementos de Π 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 contido no intervalo $[0, 1]$. Intuitivamente, Π 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.

Axiomas e propriedades de Probabilidade

As propriedades seguintes caracterizam a robustez da Teoria da Probabilidade:

- i) $0 \leq P(A) \leq 1$
- ii) $P(\Pi) = 1$
- iii) $P(A \vee B) = P(A) + P(B)$ se A e B são eventos disjuntos
- iv) Se $A \subseteq B \subseteq \Pi$ então $0 \leq P(A) \leq P(B) \leq 1$
- v) Se $A, B \subseteq \Pi$ então $P(A \wedge B) + P(A \wedge \bar{B}) = P(A)$ e, $P(A \vee B) = P(A) + P(B) - P(A \wedge B)$
- vi) $P(A) + P(\bar{A}) = 1$
- vii) Se $A_i \subseteq \Pi$ para $1 \leq i \leq n$ e $A_i \cap A_j = \emptyset$ então para todo $i \neq j$

$$\Rightarrow P(A_1 \vee A_2 \vee \dots \vee A_n) = P(A_1) + P(A_2) + \dots + P(A_n)$$

Probabilidade Condicional

Considere que duas moedas são lançadas. Adicionalmente, sabe-se que pelo menos uma cara ocorreu. Como construir um novo espaço de probabilidade (Π, P') que seja consistente com o espaço de probabilidade (Π, P) anterior e com nova informação ?

A construção de um novo espaço de probabilidade (Π, P') onde o anterior (Π, P) é alterado devido a uma nova informação é o problema central no desenvolvimento de Sistemas Especialistas Probabilísticos. 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 (Π, 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 (Π, P) um espaço de probabilidade e seja $A \subseteq \Pi$ tal que $P(A) \neq 0$. Definido um novo espaço de probabilidade (Π, f) da seguinte forma:

$$f(e) = \begin{cases} P(e) / P(A), & \text{se } e \in A \\ 0, & \text{se } e \notin A \end{cases}$$

para qualquer $B \subseteq \Pi$ 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.

A Formulação do 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:

$$P(H_i / e) = \frac{P(e / H_i) \cdot P(H_i)}{P(e)} \quad (1)$$

Se $P(H_i \wedge e) \neq 0 \forall i$ então

$$\frac{P(H_i / e)}{P(H_j / e)} = \frac{P(H_i)}{P(H_j)} \cdot \frac{P(e / H_i)}{P(e / H_j)} \quad (2)$$

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

$$P(e) = P(H_1).P(e / H_1) + P(H_2).P(e / H_2) + \dots + P(H_k).P(e / H_k) \quad (3)$$

resultando:

$$P(H_i / e) = \frac{P(e / H_i) \cdot P(H_i)}{\sum_{j=1}^k (P(H_j) \cdot P(e / H_j))} \quad (4)$$

Nas aplicações dos sistemas especialistas probabilísticos os H 'is 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 . (NASSAR, 2000)

Em (RODDER, 1997) descreve-se que na teoria da probabilidade e na estatística procura-se discutir os fenômenos sujeitos à aleatoriedade. Os modelos probabilísticos devem ser concebidos tendo em vista esta realidade. Informações sobre relações entre fenômenos reais são representadas sob forma de parâmetros de distribuições de probabilidade.

5.3. REDES BAYESIANAS

São redes de conhecimento ou estruturas gráficas que representam sistemas baseados em conhecimento calcados em vários tipos de teorias. Quando são formalizados com base na teoria da probabilidade, são chamadas redes bayesianas, causais ou diagramas de influência. Quando sua formalização tem base na teoria de Dempster-Shafer, fala-se então de Galerias, redes quantitativas de Markov. A teoria da probabilidade provê consistência às redes de

conhecimento bayesianas. Por si, a teoria da probabilidade 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 se passa a ter dificuldade em tomar decisões. Com uma representação gráfica pode-se trabalhar localmente o conceito de dependência, assim estando bem mais próximo do raciocínio humano. Apesar de não permitirem a representação de todas as características desejáveis, os grafos são estruturas que melhor representam as relações ou modelos de dependência. Além das vantagens que a estrutura oferece, a rede representa uma distribuição de probabilidade, e com o uso de uma semântica clara, permite processar diagnóstico, aprendizado, inferência, além de outras necessárias em um sistema inteligente. 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. (JUNIOR, 1997)

Para se verificar se um grafo direcionado e acíclico é uma rede bayesiana, existe uma condição necessária e suficiente: cada variável X do grafo deve ser condicionalmente independente de todos os nós que não são seus descendentes exceto seus pais (Π_i). Baseados nesta condição, pode-se formular um procedimento recursivo para a construção de uma rede bayesiana. Dada uma distribuição de probabilidade conjunta $P(x_1, \dots, x_n)$ e uma determinada ordem destas variáveis dada por d , inicia-se a construção do grafo escolhendo o nó raiz (X_1) e atribuindo a ele a probabilidade marginal $P(x_1)$. Em seguida, acrescenta-se mais um nó (X_2) no grafo. Caso X_2 seja dependente de X_1 , X_2 tem X_1 como pai, então traça-se um arco direcionado que vai de X_1 a X_2 e quantifica-se este arco com $P(x_2/x_1)$, caso contrário, mantém-se X_1 e X_2 desconectados e atribui-se uma probabilidade prévia (*à priori*) a X_2 . Ao se atingir o i -ésimo nó (X_i), traça-se os arcos de X_i a todos os seus pais (Π_i) e quantifica-se este grupo de arcos com $P(x_i / \Pi_i)$ e assim obtém-se a rede bayesiana que representa todas as dependências que são conseqüências de definição dos pais das variáveis. (JUNIOR, 1997)

Redes Bayesianas são esquemas de representação de conhecimentos utilizadas para desenvolver a base de conhecimento de um sistema especialista probabilístico. São compostas por uma parte estrutural que reflete as relações causais entre as variáveis, e pelos valores de probabilidade que refletem a força da relação entre as variáveis. De uma forma geral, a estrutura de uma rede consiste de duas partes:

- **Parte Qualitativa:** Representa o modelo gráfico (grafo acíclico direcionado) onde as variáveis são os nós e as regras, que são as relações de dependência condicional entre essas variáveis, são os arcos direcionados;
- **Parte Quantitativa:** É o conjunto de probabilidades condicionais associadas aos arcos existentes no modelo gráfico e as probabilidades estimadas à priori das hipóteses diagnosticas.

Uma vez que a topologia da rede é definida, basta especificar as probabilidades condicionais para os nodos que possuem dependências diretas e usar essas para obter qualquer outro valor de probabilidade. A rede bayesiana pode ser considerada uma base de conhecimento abstrata que contém uma grande variedade de composições diferentes, pois representa a estrutura geral dos processos causais do domínio em lugar de qualquer detalhe da população de indivíduos. A vantagem dessa representação é que permite ao especialista expressar diretamente a relação qualitativa fundamental de "dependência direta" entre as variáveis. No desenvolvimento da rede bayesiana, o direcionamento das setas é essencial para exibir as dependências de não-transitividade. Dois eventos não chegam a ser relevantes entre si somente pelo fato de predizerem uma consequência comum, mas tornam-se relevantes quando a consequência é realmente observada. Se os arcos forem despojados das setas, algumas dessas relações não serão representadas. Exibindo as irrelevâncias do domínio, o raciocínio causal minimiza o número de relações que precisam ser consideradas na construção do modelo e em muitas das conclusões futuras. A codificação do conhecimento por meio de regras e evidências, prática prevaiente dos sistemas especialistas baseados em regras, falha neste ponto. Ao tentar representar todas as exceções em um domínio proposto, o número de regras fica intratável. A inferência bayesiana é feita com a utilização da teoria da probabilidade, mais precisamente do teorema de Bayes. Há dois tipos de cálculos realizados por uma rede bayesiana: a *atualização* de crenças e a *revisão* de crenças. A atualização é o cálculo de probabilidades das variáveis aleatórias e a revisão refere-se à obtenção das probabilidades das hipóteses diagnosticas e a identificação da hipótese diagnostica com maior valor de probabilidade. (PEARL, 1988)

Em capítulos posteriores descreve-se mais sobre redes bayesianas enfocando sobre as ferramentas (shell) para implementação de sistemas especialistas probabilísticos.

6. SISTEMAS ESPECIALISTAS DIFUSOS

6.1. INTRODUÇÃO

O conceito de Conjuntos Difusos foi introduzido, em 1965, por Lotfi A. Zadeh (Universidade da Califórnia, Berkeley). A ele é atribuído o reconhecimento como grande colaborador do Controle Moderno. Em meados da década de 60, Zadeh observou que os recursos tecnológicos disponíveis eram incapazes de automatizar as atividades relacionadas a problemas de natureza industrial, biológica ou química, que compreendessem situações ambíguas, não passíveis de processamento com uso da lógica computacional fundamentada na lógica booleana. Procurando solucionar esses problemas o Prof. Zadeh publicou em 1965 um artigo resumindo os conceitos dos Conjuntos Difusos, revolucionando o assunto com a criação de Sistemas Difusos. Em 1974, o Prof. Mamdani, do Queen Mary College, Universidade de Londres, após inúmeras tentativas frustradas em controlar uma máquina a vapor com tipos distintos de controladores, somente conseguiu fazê-lo por meio da aplicação do raciocínio difuso. Esse sucesso serviu de alavanca para muitas outras aplicações, como em 1980, no controle difuso de operação de um forno de cimento. Vieram em seguida, várias outras aplicações, destacando-se, por exemplo, os controladores difusos de plantas nucleares, refinarias, processos biológicos e químicos, trocador de calor, máquina diesel, tratamento de água e sistema de operação automática de trens. Estimulados pelo desenvolvimento e pelas enormes possibilidades práticas de aplicações que se apresentaram, os estudos sobre Sistemas Difusos e controle de processos avançam rapidamente, culminando com a criação em 1984, da Sociedade Internacional de Sistemas Difusos, constituída, principalmente, por pesquisadores dos países mais avançados tecnologicamente. A propósito disto, e apenas a título de ilustração, mais de 30% dos artigos até hoje publicados são de origem japonesa. O desenvolvimento de técnicas de inteligência artificial (IA) nos últimos anos, ocupa cada vez mais posição de destaque em pesquisas na área de controle de processos industriais e, aos poucos, começam a ser implantadas em plantas industriais com enorme sucesso. Bens de consumo japoneses incorporam extensamente Sistemas Difusos, aspiradores de pó Matsushita usam controladores de 4 bits rodando algoritmos difusos que interrogam sensores de pó e ajustam o poder de sucção; máquinas de lavar Hitachi usam controladores difusos para controle de peso, verificação de tipo de tecido, e sensores de sujeira e automaticamente setam os ciclos de lavagem para o uso otimizado de potência, água, e detergente. Como um exemplo mais específico, a Canon desenvolveu uma câmera com autofoco que usa um dispositivo "charge-coupled" (CCD) para medir a claridade de imagens em seis regiões do

campo de visão e usa a informação obtida para determinar se a imagem está no foco; ela também rastreia a taxa de mudança no movimento da lente durante a procura do foco e controla a velocidade para evitar overshoot. O sistema de controle da câmera usa 12 entradas: 6 para obter a claridade atual fornecendo dados para o CCD e 6 para medir a taxa de mudança no movimento da lente. A saída é a posição do curso da lente. O sistema difuso usa 13 regras e requer 1.1KB de memória. Um outro exemplo de um sistema prático, um ar condicionado industrial projetado pela Mitsubishi usa 25 regras de aquecimento e 25 regras de refrescamento. Um sensor de temperatura fornece a entrada, com o controle da saída alimentado por um inversor, uma válvula compressora, e um ventilador. Comparado com o projeto anterior, o controlador difuso aquece e esfria 5 vezes mais rápido, reduz o consumo de potência em 24%, incrementa a estabilidade da temperatura por um fator de 2, e usa menos sensores. O entusiasmo dos japoneses por Conjuntos Difusos é refletido na ampla faixa de outras aplicações que eles tem investigado ou implementado: reconhecimento de caracteres, sistemas difusos óticos, robôs, helicópteros comandados por voz por meio de um robô, sistemas de elevadores, entre outras. Trabalhos em Sistemas Difusos é também um procedimento nos EUA e Europa, entretanto não com o mesmo entusiasmo visto no Japão. A Agência de Proteção Ambiental nos EUA tem investigado controle difuso para motores de energia eficaz, e a NASA tem estudado controle difuso para ancorar suas naves automaticamente no espaço; simulações mostram que um sistema de controle difuso pode reduzir em muito o consumo de combustível. Firmas como a Boeing, General Motors, Allen-Bradley, Chrysler, Eaton e Whirlpool têm trabalhado com Conjuntos Difusos para uso em refrigeradores de baixa potência, melhorando a transmissão automotiva, e motores elétricos de energia eficaz. Em 1995 Maytag introduziu um máquina de lavar pratos "inteligente" baseado em um controlador difuso e um "one-stop sensing module" que combina um termistor (para medida da temperatura), um sensor condutivo (para medir o nível de detergente por meio dos íons presentes na água), um sensor de turvação que difundi a medida e transmite luz para medir a sujeira na lavagem, e um sensor magnético para ler a taxa de giro. O sistema determina uma otimização no ciclo de lavagem para qualquer carga obter os melhores resultados com o mínimo de energia, detergente, e água. Pesquisa e desenvolvimento está em andamento em aplicações difusas em projeto de software, incluindo Sistemas Difusos expert e integração de Conjuntos Difusos com redes neurais os denominados algoritmos genéticos adaptativos, com o objetivo de construção de um Sistema Difuso capaz de aprender. (KLIR, 1995) e (SAFFIOTI, 1997)

6.2. CONJUNTOS DIFUSOS

A ciência da lógica e, mais especificamente a *Lógica Clássica* proveio de séculos de reflexões sobre o pensamento filosófico grego. As contribuições de diversos filósofos tais como, “Sócrates”, “Platão”, “Aristóteles”, que sistematizou a lógica clássica, bem como os sofistas, estabeleceram um conjunto de regras rígidas para que conclusões pudessem ser aceitas logicamente válidas. O emprego da lógica aristotélica levava a uma linha de raciocínio lógico baseado em premissas e conclusões. Como por exemplo: se é observado que *“todo ser vivo é mortal”* (premissa 1), a seguir é constatado que *“Sarah é um ser vivo”* (premissa 2), como conclusão têm-se que *“Sarah é mortal”*. Desde então, a lógica Ocidental, assim chamada, tem sido binária, isto é, uma declaração é falsa ou verdadeira, não podendo ser ao mesmo tempo parcialmente verdadeira e parcialmente falsa. Esta suposição e a lei da não contradição, que coloca que *“U e não U”* cobrem todas as possibilidades, formam a base do pensamento lógico Ocidental. Os Conjuntos Difusos violam estas suposições. O conceito de dualidade, estabelecendo que algo pode e deve coexistir com o seu oposto, faz os Conjuntos Difusos parecerem naturais, até mesmo inevitáveis. A lógica aristotélica trata com valores *“verdade”* das afirmações, classificando-as como verdadeiras ou falsas. Não obstante, muitas das experiências humanas não podem ser classificadas simplesmente como verdadeiras ou falsas, sim ou não, branco ou preto. Por exemplo, é aquele homem alto ou baixo? A taxa de risco para aquele empreendimento é grande ou pequena? Um sim ou um não como resposta a estas questões é, na maioria das vezes, incompleta. Na verdade, entre a certeza de ser e a certeza de não ser, existem infinitos graus de incerteza. Esta imperfeição intrínseca à informação representada numa linguagem natural, tem sido tratada matematicamente no passado com o uso da teoria das probabilidades. Contudo, Conjuntos Difusos, com base na teoria dos Conjuntos Nebulosos (Fuzzy Set), tem-se mostrado mais adequada para tratar imperfeições da informação do que a teoria das probabilidades. De forma mais objetiva e preliminar, podemos definir Conjuntos Difusos como sendo *uma ferramenta capaz de capturar informações vagas, em geral descritas em uma linguagem natural e convertê-las para um formato numérico, de fácil manipulação pelos computadores de hoje em dia.* (KLIR, 1995) e (ARBEX, 1994)

6.3. SISTEMAS DIFUSOS

Em (ZIMMERMANN, 1999) descreve-se que um sistema difuso é formado pela agregação de um conjunto de regras de inferência que mapeia regiões no espaço das entradas $U_1 \times U_2 \times \dots \times U_r$ para regiões no espaço das saídas $V_1 \times V_2 \times \dots \times V_m$. Este mapeamento se faz com o uso de conjuntos difusos para conjuntos difusos. A figura-2 mostra um sistema difuso de duas entradas e um saída. Assim, este sistema mapeia regiões do espaço R^2 para o espaço R^1 . Este sistema é formado por "n" regras que codificam o conhecimento de um especialista descrito como "Se X_1 é A_i e X_2 é B_i Então Y é C_i ", como exemplo de tal conhecimento, poderíamos ter: *Se a renda do cliente é alta e sua dívida é pequena, então seu crédito é muito alto*. Desta forma, a variável lingüística X_1 estaria associada com a renda do cliente (Cliente.renda); a variável lingüística X_2 estaria associada com a dívida do cliente (Cliente.dívida); a variável lingüística Y estaria associada com o crédito do cliente (Cliente.crédito); o conceito de renda alta estaria associado a um conjunto difuso A_i , digamos A_3 (alta.renda); o conceito de dívida pequena estaria associado a um conjunto difuso B_i , (pequena.dívida) e finalmente, o conceito de crédito muito alto estaria associado a um conjunto difuso C_i , (muito_alto.crédito). Do ponto de vista da lógica difusa e neste exemplo, as variáveis lingüísticas e os conjuntos difusos: X_1 e A_i , X_2 e B_i , e Y e C_i pertencem aos universos de discurso U , V e W , respectivamente.

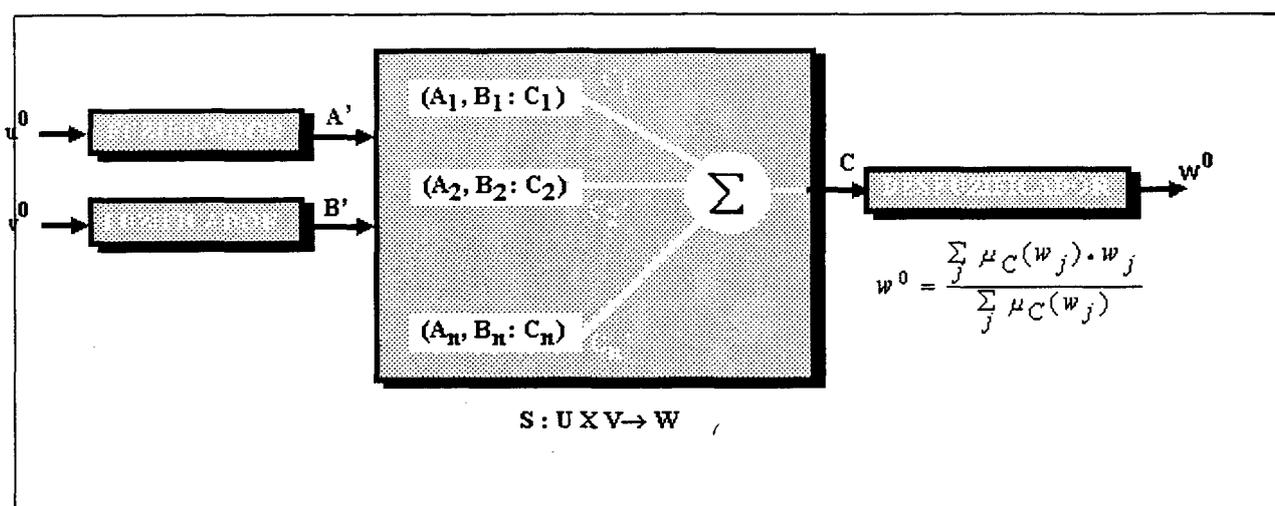


Figura 2: Um Sistema Difuso com duas Entradas e uma Saída.

As duas entradas A' e B' interagem com as "n" regras, aqui representadas na forma compacta por $(A_i, B_i: C_i)$, produzindo "n" conjuntos difusos C_i' que são uma versão dos "n" conjuntos C_i modificados pelo grau de interação das entradas com as respectivas regras. Estes conjuntos difusos são então somados logicamente para produzir o conjunto difuso de saída (C). A partir da regra composicional de inferência, têm-se que:

A solução para:

$$\begin{aligned} C_i' &= (A' \wedge B') \wedge (A_i \wedge B_i) \rightarrow C_i \\ C &= \sum_{i=1}^n (A' \wedge B') \wedge (A_i \wedge B_i) \rightarrow C_i \end{aligned} \quad (5)$$

em termos da sua função de pertinência é:

$$\mu_C(w) = \bigvee_{u,v} [\mu_{A'}(u) \wedge \mu_{B'}(v)] \wedge [(\mu_{A_i}(u) \wedge \mu_{B_i}(v)) \rightarrow \mu_{C_i}(w)] \quad (6)$$

A credibilidade das regras de inferência poderia ser levada em conta, ponderando-se as saídas das regras com um peso θ de forma que a saída seria então:

$$C = \sum_i \theta_i C_i' \quad \text{com} \quad 0 \leq \theta_i \leq 1 \quad (7)$$

Na grande maioria das aplicações dos sistemas difusos em engenharia, a troca de informações entre o ambiente e o sistema deve se dá por meio de variáveis cujos valores são números reais. Exemplificando: a saída de um sensor de temperatura é um valor real (ex.: 65°C) e não um conjunto difuso. Assim, este interfaceamento entre o sistema difuso puro e o ambiente no qual ele está situado, se faz com o uso de dois blocos funcionais chamados de *Fuzzificador* e *Defuzzificador*, respectivamente. O bloco fuzzificador tem a função de converter um valor real em um conjunto difuso. O mais simples dos fuzzificadores é o fuzzificador "singleton" que converte um número real em um conjunto nebuloso singleton. O bloco defuzzificador realiza o processo inverso. Dentre os métodos existentes, o mais popular é o método do centro de gravidade que está exemplificado na fig. 2. A fig. 3 mostra uma visualização gráfica do processo de inferência usando a função de implicação de Mamdani, para um sistema difuso semelhante àquele mostrado na fig. 2, mas com apenas três regras de inferência difusos. Note a partir da fig. 3 que o suporte numérico para os universos de discurso U , V e W podem ser completamente diferentes um dos outros. Note também, que um valor de entrada (u ou v) pode ativar mais de uma regra ao mesmo tempo. Isto é completamente diferente do sistema especialista clássico, onde uma premissa simples só ativa uma única regra.

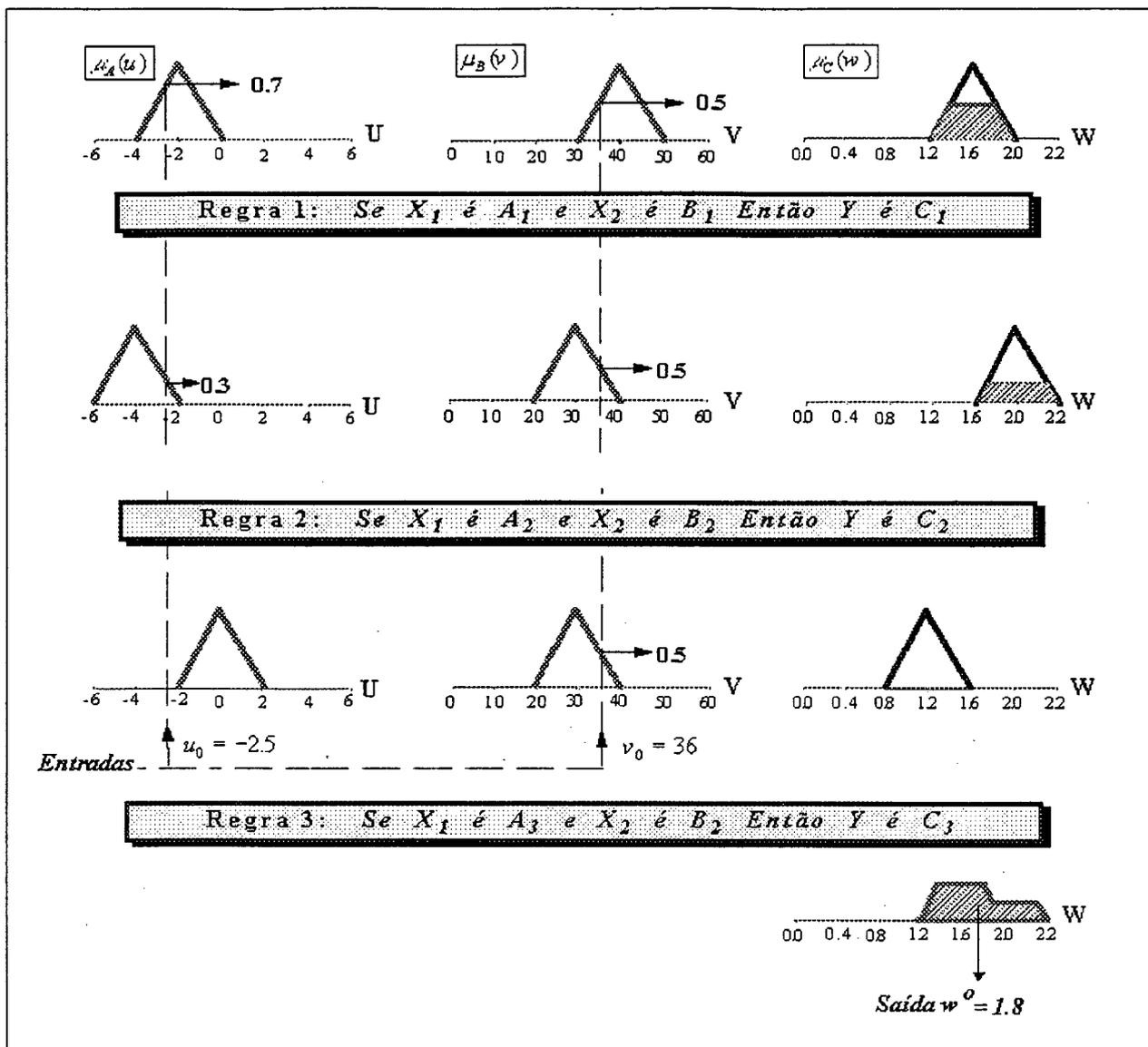


Figura 3: Processo de Inferência para um Conjunto de Três Regras.

Supondo que em um dado instante $X_1 = U_0 = -2.5$ e que $X_2 = V_0 = 36$, o processo de inferência segue como descrito a seguir. Para todas as regras da base de regras, é calculado os valores verdade das premissas de cada regra por meio da função conjunção (\wedge):

$$\text{Regra 1: } \mu_1 = \mu_{A_1}(u_0) \wedge \mu_{B_1}(v_0)$$

$$\text{Regra 2: } \mu_2 = \mu_{A_2}(u_0) \wedge \mu_{B_2}(v_0)$$

$$\text{Regra 3: } \mu_3 = \mu_{A_3}(u_0) \wedge \mu_{B_2}(v_0)$$

Note que cada μ_i é um número real. A saída de cada regra é um conjunto difuso, que é calculada usando a função de implicação de Mamdani:

$$\text{Regra 1: } \mu_{C_1}(w) = \text{Min} [\mu_1, \mu_{C_1}(w)]$$

$$\text{Regra 2: } \mu_{C_2}(w) = \text{Min} [\mu_2, \mu_{C_2}(w)]$$

$$\text{Regra 3: } \mu_{C_3}(w) = \text{Min} [\mu_3, \mu_{C_3}(w)]$$

Estes conjuntos são então somados logicamente para se obter o conjunto difuso de saída:

$$\mu_C(w) = \text{Max} [\mu_{C_1}(w), \mu_{C_2}(w), \mu_{C_3}(w)] \quad (8)$$

Então este conjunto difuso é então convertido em número real pelo processo de defuzzificação utilizando o método, neste caso, do centro de gravidade:

$$w^0 = \frac{\sum_j w_j \mu_C(w)}{\sum_j \mu_C(w)} \quad (9)$$

Conforme abordado em (VIEIRA, 1999), os sistemas difusos são aqueles que tem, em seu sistema de inferência, o emprego de conjuntos difusos. Veja cinco vantagens do uso de sistemas difusos em relação aos sistemas clássicos:

1. Habilidade para modelar problemas extremamente complexos: sistemas convencionais, ao tratarem de problemas reais, geralmente não-lineares, aumentam consideravelmente o custo computacional, enquanto os Conjuntos Difusos tornam possível resolvê-los utilizando o raciocínio, requerendo menos regras;
2. Aumento da Modelagem cognitiva dos sistemas especialistas: a grande vantagem de se trabalhar com sistemas difusos está na forma como é armazenado o conhecimento, possibilitando a associação das regras à forma de pensar dos especialistas;
3. Habilidade para modelar sistemas envolvendo vários especialistas: os sistemas difusos possibilitam relacionar posições e decisões conflitantes, usuais quando são vários os participantes de um projeto de sistema especialista. A forma como as regras interagem em um sistema difuso garantem esta característica;

4. Complexidade do modelo reduzida: os modelos difusos requerem menos regras que os convencionais, traduzindo-se em menor complexidade do modelo. O conhecimento a ser explicitado nas próprias regras do sistema é outra característica que beneficia os sistemas difusos;
5. Melhora na manipulação de incerteza e possibilidades: enquanto boa parte dos sistemas especialistas presentes no mercado são baseados na *probabilidade bayesiana*, os modelos difusos tem como característica intrínseca o tratamento de imprecisão, o que se traduz em uma manipulação mais consistente, comparativamente aos sistemas tradicionais.

A fig. 4 exhibe um esquema resumido de um modelo difuso:

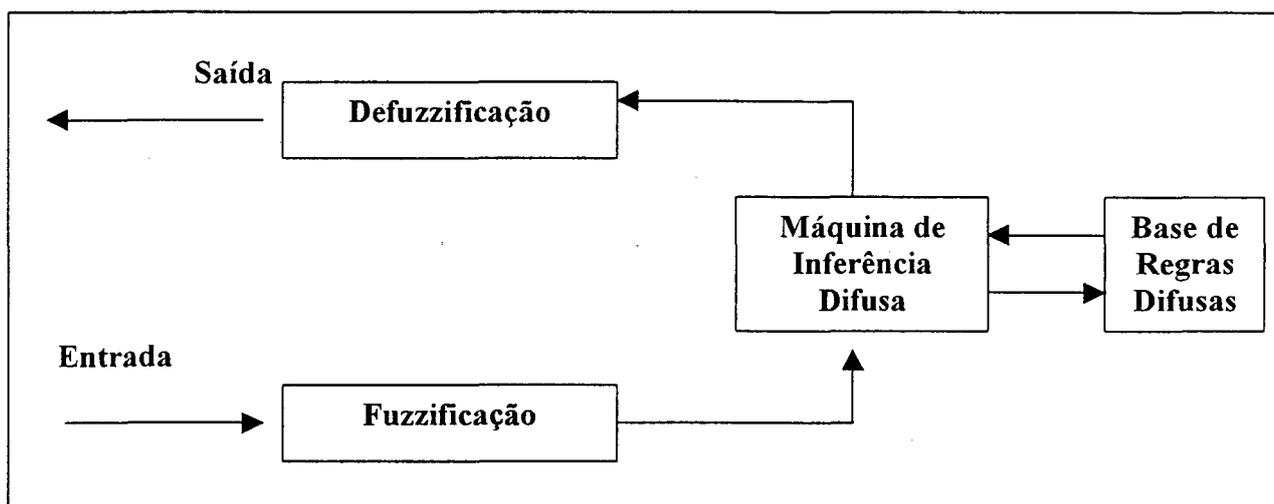


Figura 4: Esquema do modelo de um Sistema Difuso.

Em (ZIMMERMANN, 1999) tem-se as seguintes definições:

- **Fuzzificação:** é o processo que torna qualquer quantidade numérica (CRISP) em quantidade difusa. É, portanto, uma função que garante certo grau de imprecisão a um valor numérico, mapeando o valor físico de uma variável de um em um universo normalizado de discussão. Isto é necessário para que a entrada do processo se torne compatível com a representação difusa adotada na base de regras.
- **Base de Regras:** tem, como objetivo, representar de forma sistemática a maneira como o controlador gerenciará o sistema sob sua supervisão. Adotando valores lingüísticos iguais aos utilizados por nós quando efetuamos um controle sobre determinado processo, as regras envolvidas apresentam a forma sintática: *SE-ENTÃO*, onde *SE* está relacionado ao antecedente ou entrada da regra, enquanto *ENTÃO* está associado ao conseqüente ou a

ação do controle. Relacionadas as variáveis lingüísticas do processo, pode-se claramente estabelecer os antecedentes, ou estados do processo, e associá-los com ações de controle.

- **Máquina de Inferência:** é a responsável pela combinação do dado de entrada, já no formato de número difuso, com as regras difusas existentes, as quais, trabalhando em cima de *regras de produção*, descrevem o processo de tal forma que se obtenha, por meio de inferência, o desejado valor de saída.
- **Defuzzificação:** é definida como a operação inversa da fuzzificação, ou seja, tem como objetivo converter cada conclusão difusa do controlador em uma variável numérica. Existem vários métodos de defuzzificação, tendo por base o *centróide*, o *centro de máxima*, a *média das máximas*, a *média do suporte*, a *soma dos centros*, o *centro da maior área*, o *primeiro dos máximos*, cada qual com suas vantagens e desvantagens, num vasto campo de aplicações.

Segue a representação matemática do processo de defuzzificação pelo método do *Centro de Massa*, de acordo com (ROSS, 1995):

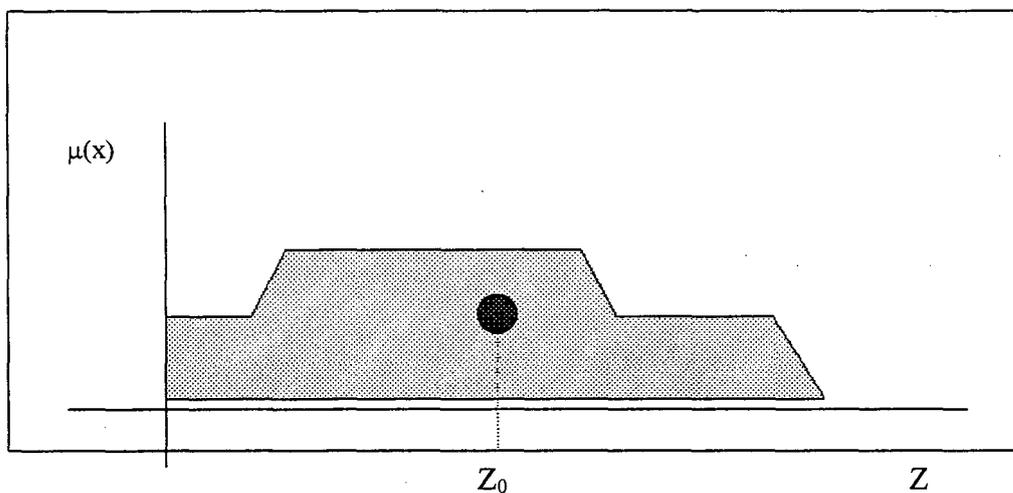


Figura 5: Processo de Defuzzificação pelo Método do *Centro de Massa*.

Assim, o *Centro de Massa* pode ser especificado por:

$$Z_0 = \frac{\int z \mu(x)(Z) \delta z}{\int \mu(x)(Z)} \quad (10)$$

ou

$$Z_0 = \frac{\sum z_i \mu(x)(Z) \delta z}{\sum \mu(x)(Z)} \quad (11)$$

6.4. NOTAS SOBRE ALGUMAS APLICAÇÕES DE CONJUNTOS DIFUSOS

Fábio Duarte de Araújo, em sua obra (ARAÚJO, 2000), fala da Casa do Futuro e seus aspectos digitais, enfatizando a questão da relação *Lar versus Mundo Cibernético*, onde uma casa ou todas as casas pertinentes ao mundo digital proposto seriam conectadas de forma que houvessem interações entre todos os dispositivos domésticos fazendo com que, uma simples falha em apenas um destes dispositivos acionasse um outro para fins de análise e correção. Nesta obra, o autor aborda a contribuição dos *Conjuntos Difusos* para inúmeras transformações e melhorias da qualidade de vida, fala sobre a incorporação desta técnica em aspiradores de pó, tostadeiras, câmeras de vídeo e freios de metrô. Cita o caso da câmera da Panasonic que armazena 120 pontos de referência e, quando o operador balança, por exemplo, ela detecta todos os pontos que se moveram numa mesma direção e é capaz de *entender* que quem se mexeu foi a câmera, e não o objeto, e com base nos primeiros pontos armazenados, corrige a imagem globalmente para compensar a tremulação. (ARAÚJO, 2000) ressalta que “*A relação casa/computador traz então duas possibilidades de discussão: a idéia de lar como interface ao universo digital e a interação com esse universo digital reconfigurando as casas*”.

(ARAÚJO, 2000) quando fala em *Lógica Difusa* “Do Carro à Geladeira” fala sobre os freios do metrô de Tokyo que respondem às variações que ocorrem no cotidiano de suas

operações e, a mesma tecnologia é usada pela Hitachi nas caixas de câmbio dos carros da Nissam para tornar mais suaves as trocas de marcha, de acordo com a velocidade do carro, relevo, mudanças anteriores. São ainda nos carros que vê-se outros usos dos computadores em nosso dia-a-dia, que em alguns casos regulam todo o sistema hidro-pneumático que, dependendo do relevo da pista, ou da vontade do motorista, controla o modo de direção agressiva ou moderada.

7. SHELL PARA SISTEMAS ESPECIALISTAS

Este item objetiva apresentar ao leitor uma breve exploração sobre alguns dos diversos softwares existentes para suportar a implementação de *Redes Bayesianas* ou *Redes de Crenças* e *Sistemas Difusos*. No levantamento realizado sobre ferramentas Shell para a elaboração de *Sistemas Especialistas Probabilísticos e Difusos*, algumas não permitem o uso parcial ou versões DEMO e por uma questão de custo para aquisição de versões FULL (completa), o trabalho ilustra a interface de algumas e apenas relata características operacionais de outras.

7.1. PROBABILÍSTICOS

Inicialmente apresenta-se a *Shell Netica* da *Norsys Software Corp.* em sua versão 1.06 for *Windows 95* e *Windows NT 4.0*. O software permite a utilização com ou sem senha de acesso, caso o usuário não a tenha, mesmo assim poderá utilizar alguns recursos na implementação da Rede. Em sua total capacidade, a shell permite a geração de código na Linguagem C, dando ao usuário (desenvolvedor) maior flexibilidade e poder na elaboração de Interfaces de Sistemas Especialistas. A figura seguinte ilustra uma pequena rede utilizada para diagnóstico de doenças e seus sintomas construída com a ferramenta Netica.

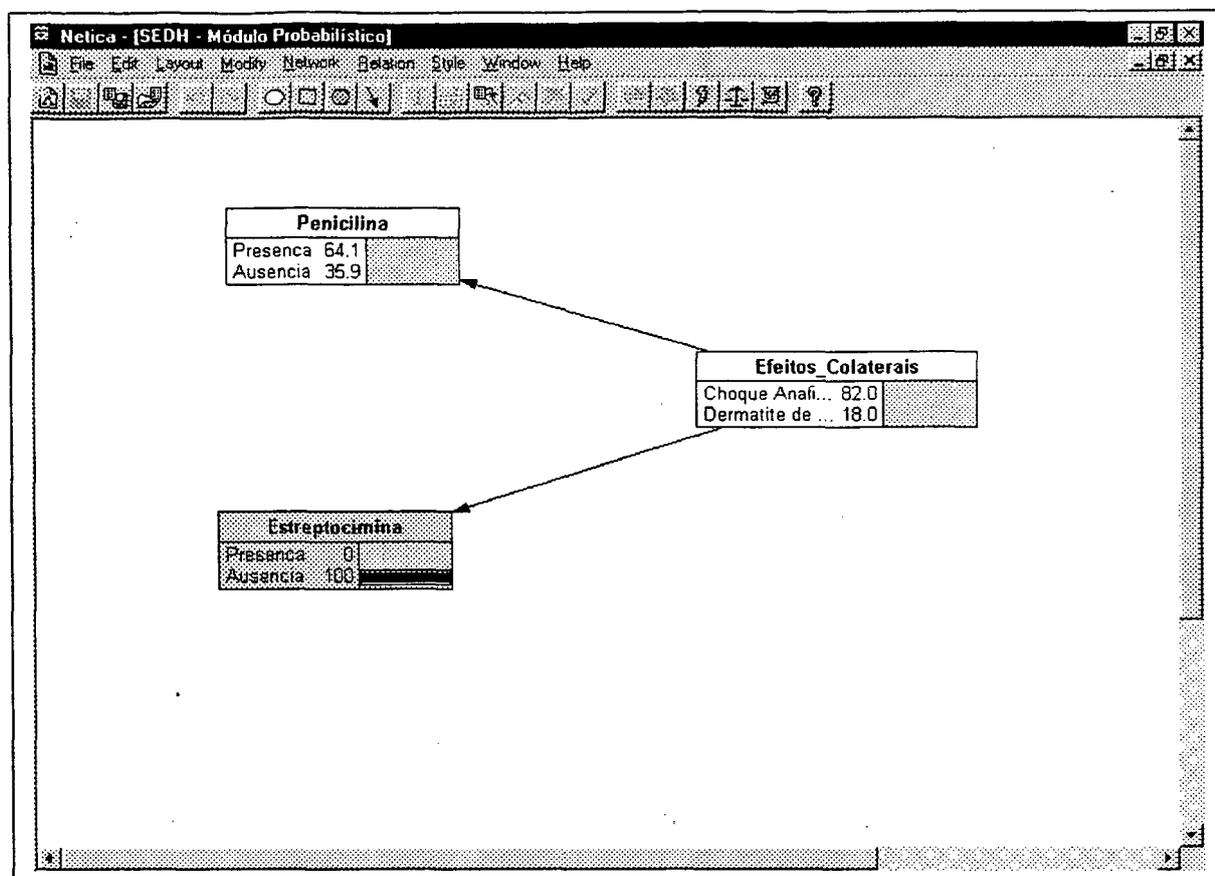


Figura 6: Uma Rede Bayesiana para Diagnóstico de Reações Adversas causadas por Antibióticos.

Em (BAYDA, 2000) tem-se informações sobre o software *BAYDA* em sua versão 1.31. É um software para análise onde o ponto alvo é a construção de um modelo para prever o valor de uma variável discreta (classe, grupo, categoria) usando outras variáveis. No campo de *Aprendizado de Máquina* esta tarefa de discriminação é conhecida como *classificação*. Embora a interface apresentada pelo *BAYDA* tenha sido projetada tendo em sua idéia principal Ciências Sociais, o software deve também ser de interesse para qualquer um que deseja trabalhar com tarefas de classificação em áreas como *Aprendizado de Máquina* e *Data Mining*. *BAYDA* tem sido testado freqüentemente em ambientes Windows 95 e NT, SunOS e plataformas Linux. É 100% implementado em Java e pode ser executado em qualquer plataforma suportando Java Runtime Environment 1.1.3 ou superior. O software é *FREE* para pesquisas e características de ensino.

BayesBuilder está disponível em (BAYES, 2000) como mais uma ferramenta para a implementação de Redes Bayesianas, em especial para a modelagem de Redes voltadas à *diagnóstico médico* e *data mining*. Veja algumas características desta ferramenta:

- Define várias visões (views) nas partes da rede sendo que isto é essencial para redes com muitos nós;
- Importa redes no formato Hugin, Netica, Microsoft Bayesian Network e Bayesian Interchange;
- Exporta o *status* da rede para uma base de dados de casos e importa casos de uma base de dados;
- Layout automático da rede;
- Relatórios.

Outra ferramenta existente para implementação de Redes Bayesianas é o software *Belief NetWork Power Constructor* conforme (BELIEF, 2000). Voltado a plataformas Windows 95 e NT em sua versão 2.1 Beta 1. Possui assistente de Interface passo à passo. Suporta base de dados nos formatos Ms-Access, dBase, Foxpro, Paradox, Excel e formato de texto. Também suporta banco de dados servidores como Oracle, Sql-Server por meio de ODBC. Possui suporte para ActiveX DLL podendo assim integrar-se com outras plataformas de Redes Bayesianas.

Conforme apresentado por (BAYESIAN, 1999), segue outros nomes de ferramentas para modelar Redes Bayesianas tais como: Analytica, Bayes Net Toolbox 2.0 for MatLab 5, Bayesian Knowledge Discoverer, BUGS, CABeN, Cispac, Ergo da Noetic Systems entre diversos outros.

7.2. DIFUSOS

Uma das ferramentas mais tradicionais para a implementação de sistemas especialistas difusos é o módulo fuzzy (Fuzzy Toolbox) embutido no software matemático MATLAB. Permite a elaboração de sistemas difusos e posterior geração em código C++. Veja a interface do ToolBox Fuzzy do MATLAB:

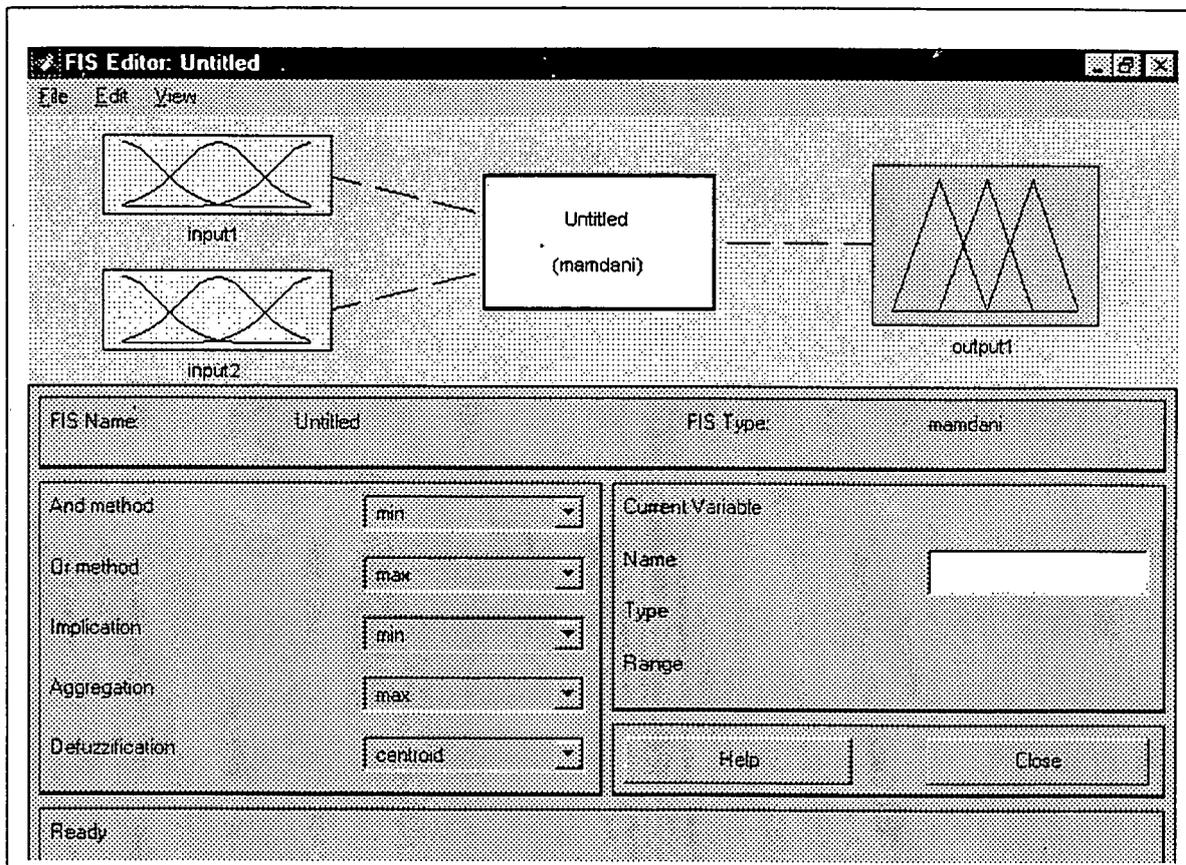


Figura 7: Interface de entrada para modelagem difusa no MATLAB ToolBox Fuzzy.

Já citado anteriormente neste trabalho, o Fuzzy Knowledge Builder é outro software para construção de Sistemas Difusos. Sua interface e exemplos podem ser vistos no item (2.3.11.4.).

A figura-8 exibe a interface do software FUDGE (CASTRO, 1994) para construção de Sistemas Difusos. A versão é DEMO e possui limites quanto às funções de *output*.

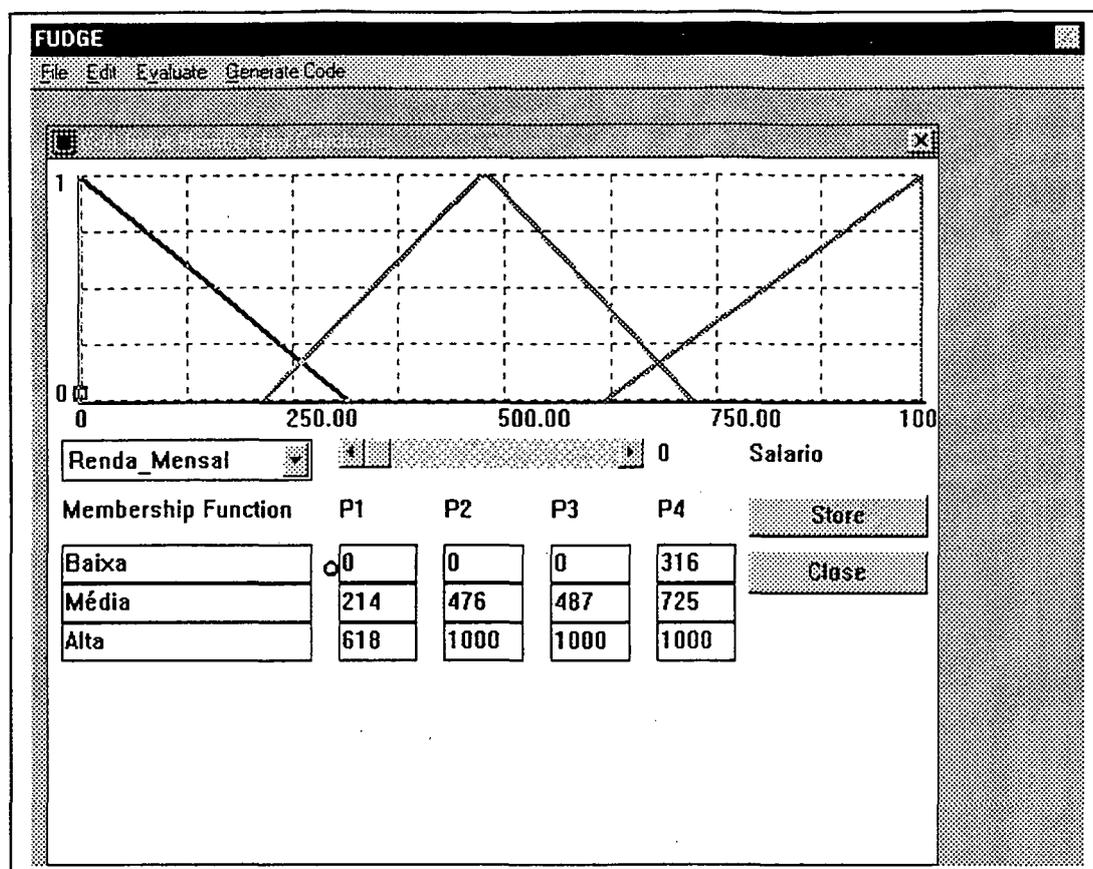


Figura 8: Construção de um sistema difuso para diagnosticar credibilidade.

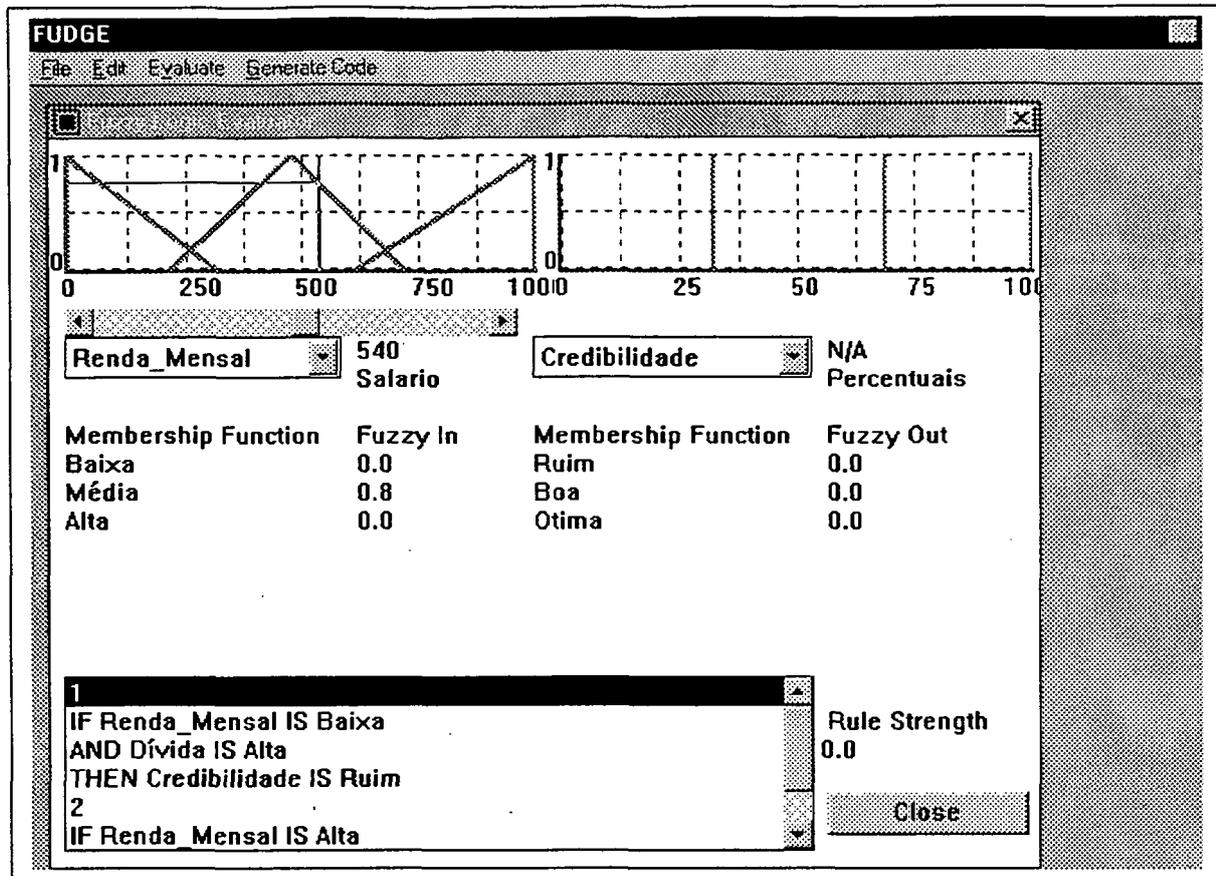


Figura 9: Interface para evolução do sistema no FUDGE.

A fig. 10 exhibe um simulador construído no FuzzyTECH (FUZZYTECH, 1999) para controlar um carro em seu percurso. Em sua versão completa o FuzzyTECH permite controle de aplicações comerciais e financeiras e suporte para sistemas neuro-fuzzy.

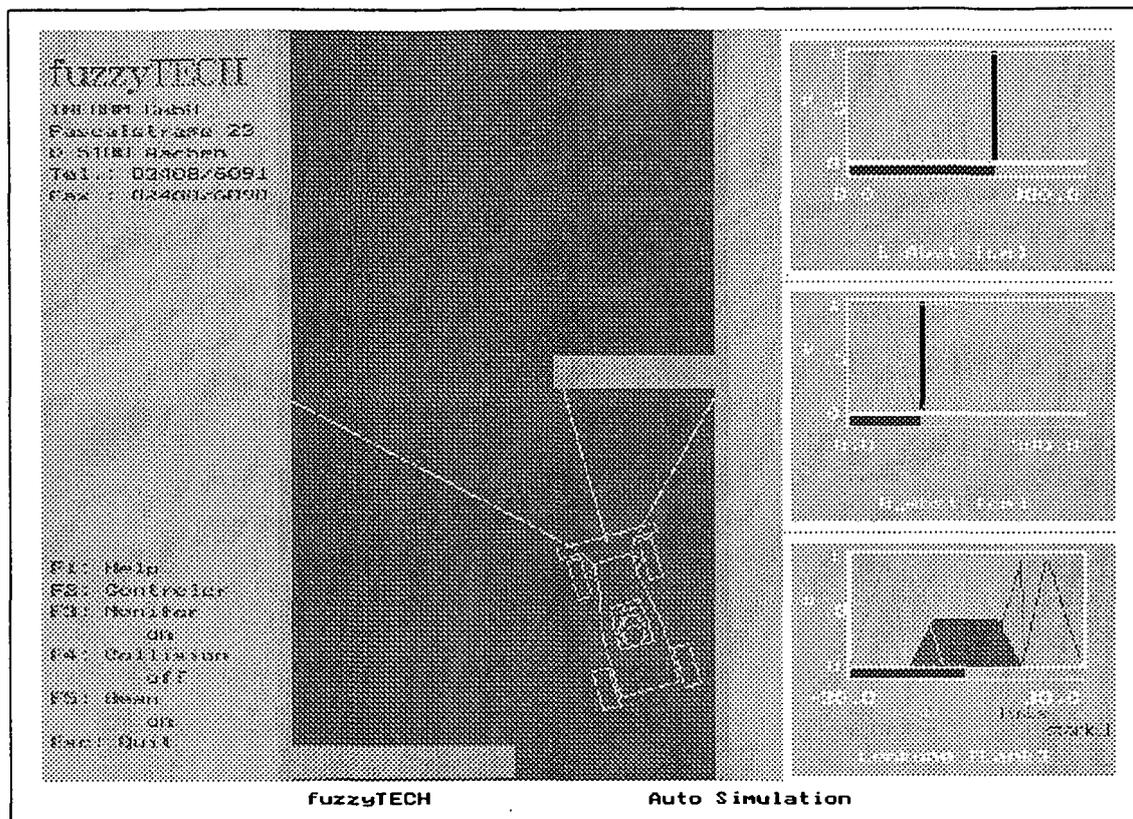


Figura 10: Simulador construído no software FuzzyTECH.

Em (FLDE, 2000) encontra-se o FLDE – Fuzzy Logic Development Environment, uma ferramenta para desenvolver aplicações complexas com Conjuntos Difusos. O software gera código em C. Não disponibiliza versão DEMO. Possui exemplos de aplicação como o problema do *Pêndulo Invertido* conforme fig. 11.

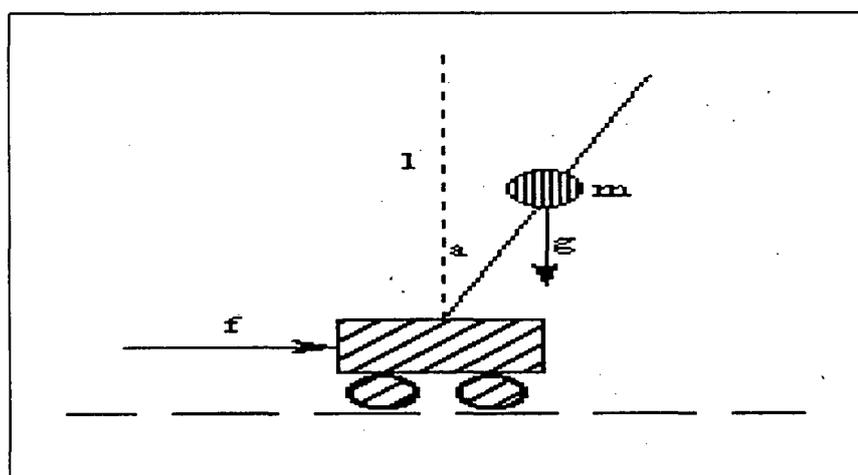


Figura 11: O Pêndulo Invertido¹.

¹ O problema do Pêndulo Invertido consiste num móvel que procura equilibrar um bastão (pêndulo) vertical sofrendo assim, ações de variáveis como gravidade, posição do ângulo do bastão x a velocidade do móvel.

Em (FUZZY, 1999) encontra-se referências de uma Shell (Fl.exe) que implementa um sistema especialista para classificação. No contexto é abordado o problema do reconhecimento da flor “íris”. Pode-se editar os arquivos contidos no software e, assim, implementar qualquer outro sistema. Apesar de ser nada agradável, pois trata-se de interface DOS, ele permite a modelagem de sistemas complexos disponibilizando funções triangulares, trapezoidais, sigmóides e gaussianas. Sua manipulação é totalmente por arquivos tipo texto.

Alguns outros nomes de softwares para Conjuntos Difusos conforme (FUZZY, 2000): CLOUDS C++ Library Organizing Urcentainty in Database Systems, Fuzzy Logic Inferencing Engine (FLIE), FuzzyCLIPS e NEFCLASS.

Cabe ainda ressaltar a utilização do software UNFUZZY desenvolvido na Universidade Nacional de Colômbia, na elaboração de protótipo para simulação de resultados desta pesquisa. A fig. 12 exhibe a janela *about* do software. Maiores detalhes desta ferramenta serão vistos posteriormente.

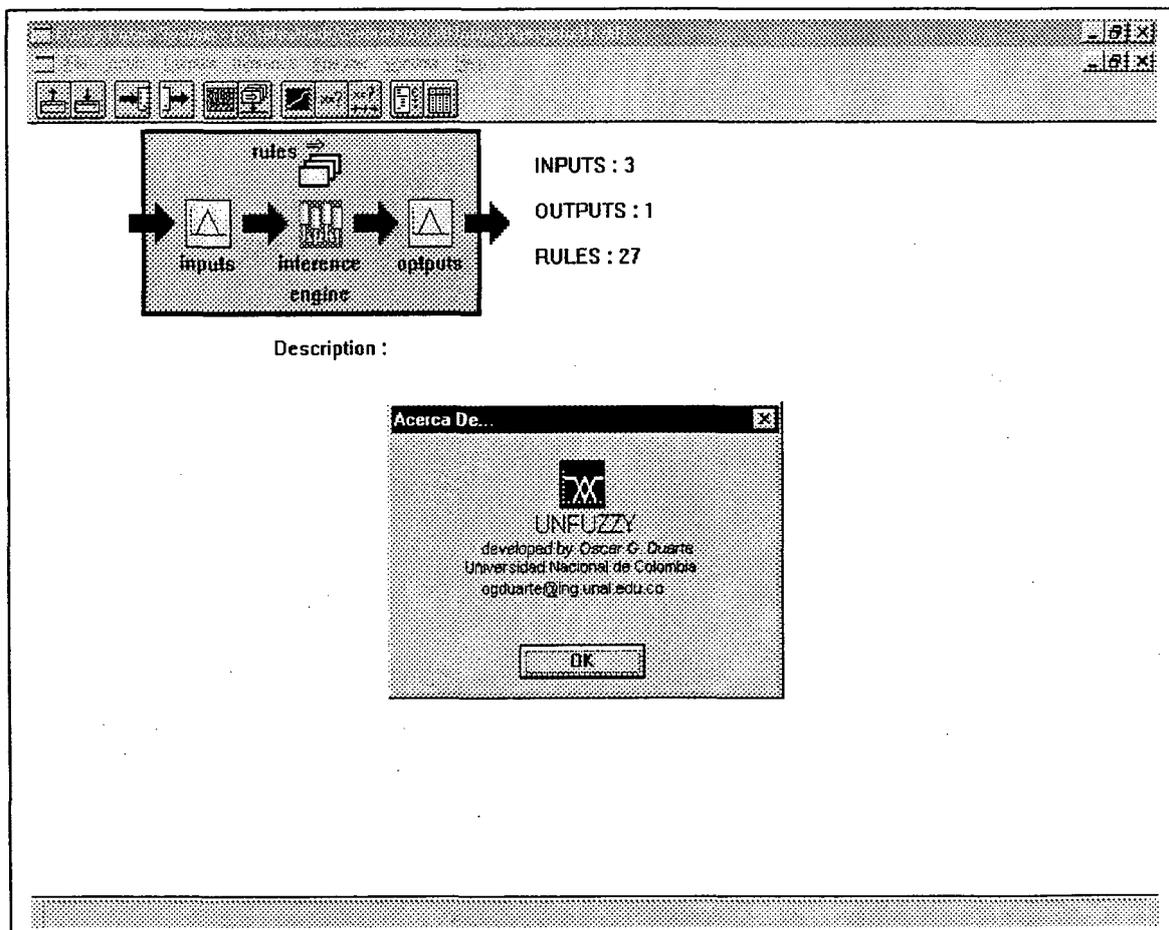


Figura 12: Software UNFUZZY para sistema especialista difuso.

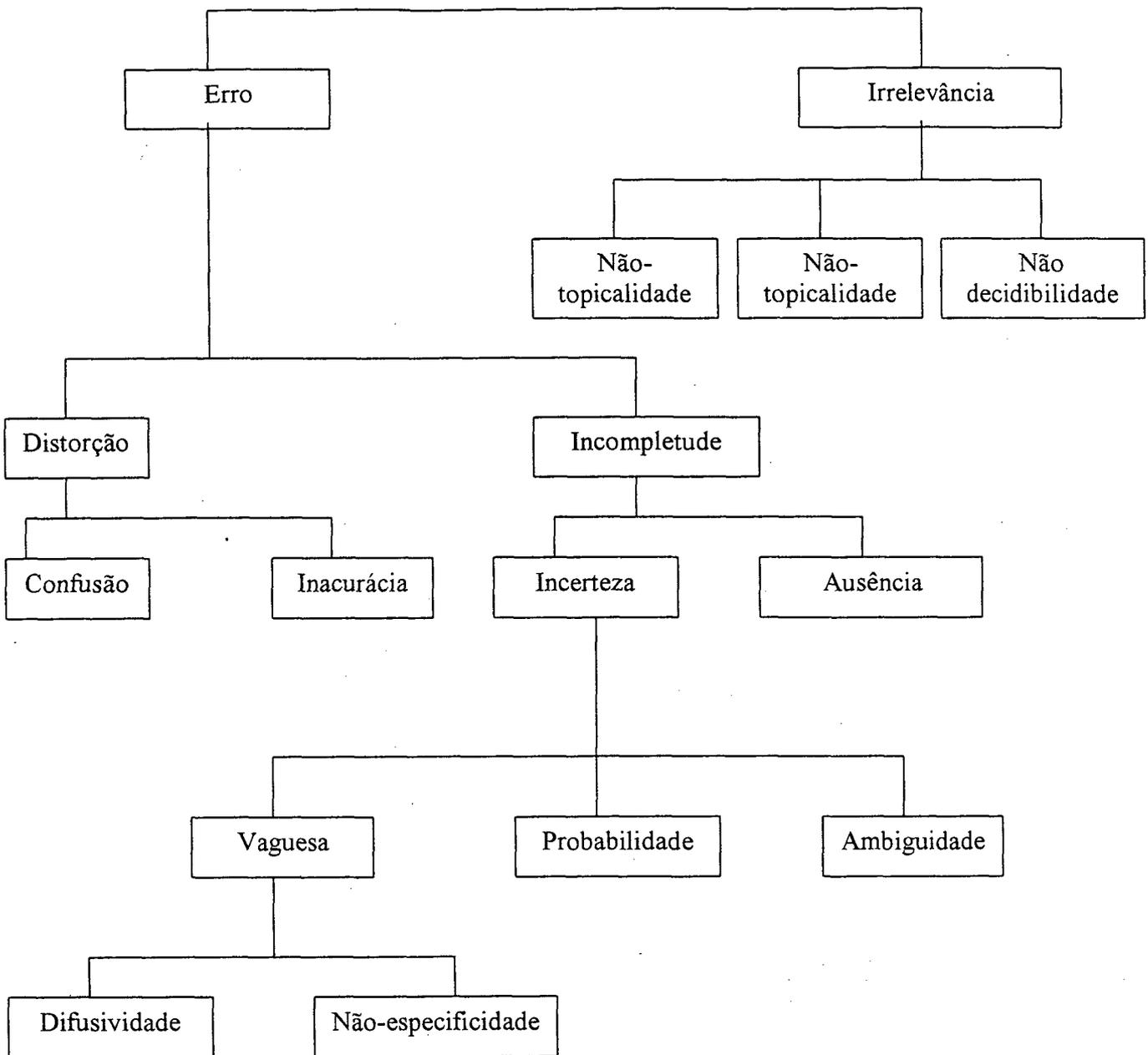
8. PESQUISAS SIMILARES AO MODELO APRESENTADO

8.1. ASPECTOS GERAIS SOBRE MEDIDAS DE INCERTEZA

Na modelagem de sistemas inteligentes, deve-se considerar características essenciais relacionadas aos mesmos, tais como, complexidade, utilidade e incerteza. Existe uma forte relação entre essas três características. Sob o ponto de vista da incerteza, pode-se dizer que, em geral, se for possível trabalhar com maiores níveis de incerteza, o modelo que descreve o sistema poderia ter a maior credibilidade e menor complexidade. Quando se modela um sistema é imposta uma negociação entre essas características. Se o nível de incerteza permitido for grande, o sistema seria possivelmente mais manipulável e útil. Depois de uma pesquisa clínica por exemplo, poder-se-ia encontrar que o tempo esperado de vida de um grupo após uma dada intervenção cirúrgica seria de dez anos mais ou menos dois anos, com 95% de confiança. Essa declaração tem mais informação útil do que quando se diz que com 99.9% de confiança, aquele grupo tem expectativa de vida de dez anos mais ou menos nove anos. Nesse caso, o aumento da incerteza fornece mais credibilidade para o resultado encontrado. Considerando-se um outro exemplo com relação à complexidade. Supondo um modelo de regressão múltipla para determinar as fontes da diabetes gestacional, procura-se um número limitado de variáveis para a construção do modelo, atribuindo-se mais variáveis menos importantes ao erro estatístico o que diminui possíveis problemas de multicolinearidade. Nesse caso o sistema tem a sua complexidade diminuída e como custo o aumento da incerteza. (BRACARENSE, 1999)

De acordo com (BRACARENSE, 1999), o esquema que segue retrata os níveis de incerteza, inserindo incerteza no contexto da ignorância:

Ignorância



8.2. NOTAS SOBRE OUTRAS INVESTIGAÇÕES

Esta unidade tem por objetivo apresentar uma explanação superficial sobre as investigações de (ROSS, 1995) e (KAND, 1986), os quais trabalharam sobre uma abordagem de raciocínio difuso-bayesiano expondo suas idéias por meio de modelos matemáticos.

Inicialmente, mostra-se uma breve visão dos estudos realizados pelo Professor Timothy J. Ross da Universidade do Novo México, o qual cita com clareza em (ROSS, 1995), “*O problema do esquema Bayesiano está nos eventos vagos e ambíguos*”. A idéia partiu da associação de várias alternativas de escolha sobre estados naturais a valores de utilidade para cada uma destas alternativas. Isso mostra que em seus estudos, o Professor Ross abordou a *Teoria da Utilidade*. Ele especifica que a utilidade esperada associada com a *j-ésima* alternativa deve ser:

$$E(u_j) = \sum_{i=1}^n u_{ij} p(s_i)$$

onde, $E(u_j)$ é uma função de utilidade para um valor esperado.

Em seu trabalho, (ROSS, 1995) utiliza de árvores de decisão para a escolha de valores de probabilidades associados às alternativas de um dado problema. Na modelagem proposta pelo Professor Rossi, percebe-se a integração entre várias alternativas matemáticas, como a já citada Teoria da Utilidade e um tratamento de problemas de decisão envolvendo matrizes numéricas que são compostas por valores de utilidades de alternativas de escolha, bem como a especificação de probabilidades condicionais por meio de *informação imperfeita* e de *informação perfeita*. No que refere-se estritamente aos eventos difusos, partiu-se da construção de tabelas com informações perfeitas associadas com estes eventos, onde, estas informações foram classificadas como “boa”, “moderada” e “pobre”. O evento difuso deve ter uma função de pertinência $\mu_M(x_k)$, $k= 1,2, \dots,r$. (ROSS, 1995) define como a “*probabilidade de um evento difuso*” ser:

$$P(\bar{M}) = \sum_{k=1}^r \bar{\mu}_M(x_k) p(x_k)$$

É notável na expressão acima, de acordo com as citações de Ross que, se um evento difuso, for em fato Crisp, $\bar{M} = M$ e, a probabilidade se reduz para:

$$P(M) = \sum_{x_k \in M} p(x_k) \qquad \mu_M = \begin{cases} 1, & \text{se } x_k \in M \\ 0, & \text{se caso contrário} \end{cases}$$

desta forma,

$$P(\bar{M}/S_i) = \sum_{k=1}^r p(x_k/s_i) \bar{\mu}_M(x_k)$$

Para ao final obter-se uma informação difusa, são utilizadas funções de utilidade. Em sua obra, Ross fala de tabelas com funções de pertinência ortogonais para eventos difusos ortogonais.

É notável em seus estudos a preocupação de manter integra as propriedades axiomáticas da Teoria da Probabilidade.

Outro pesquisador desta área de formulação de métodos de raciocínio foi o Professor Abraham Kandel da Universidade do Estado da Flórida, o qual faz em sua obra uma exposição de opiniões sobre o que ele chama de “*Probabilidade Subjetiva*”, com base também na obtenção de probabilidades condicionadas a eventos difusos.

Não se detalhará aqui a modelagem realizada em (KAND, 1986) pelo fato da intensa formulação matemática abordada na pesquisa, que retrata uma necessidade da especificação de idéias por meio de processos de *Diferenciação e Integração de Funções Difusas* e uma resolução que parte da reflexão sobre o conceito de *Limite de Funções Difusas* até modelos de *Equações Diferenciais*.

Kandel faz uma breve análise do uso das teorias da *Possibilidade* e da *Probabilidade* na resolução ou tratamento de incertezas e cita em seu trabalho que “*Ao contrário da crença expressada por alguns pesquisadores, a teoria da probabilidade, por si só, ou em combinação com o princípio da máxima-entropia, não provê de uma ferramenta adequada para a análise de problemas nos quais a informação disponível está incompleta, imprecisa ou inconfiável*”. Ainda em seu trabalho no que refere-se à Estatística Difusa, ele retrata que “*Probabilidade é uma característica objetiva; as conclusões da teoria da probabilidade podem, em geral, serem testadas por experiência*”; “*Um grau de pertinência é subjetivo, embora é natural associar um baixo grau de pertinência a um evento que, considerado sobre o aspecto de probabilidade, pode ter uma baixa probabilidade de ocorrência*”.

De acordo com (KANDEL, 1986), a motivação para o desenvolvimento de Estatística Difusa é filosófica e conceitualmente relacionada com Probabilidade Subjetiva.

CONSIDERAÇÕES FINAIS

É de importância fundamental compreender a maneira que as técnicas enfatizadas anteriormente, Conjuntos Difusos e Redes de Probabilidades, efetuam o processamento do raciocínio artificial, sendo assim, deve-se evidenciar a diferença principal entre tais técnicas, visto que ambas são utilizadas para o tratamento de incertezas. Enquanto Conjuntos Difusos processam incertezas por imprecisão, as Redes de Probabilidades processam incertezas por aleatoriedade.

Tendo conhecimento desta diferença, pode-se pensar em termos de uma interseção existente entre estas duas técnicas, ou seja, o que existe em comum no tratamento de incertezas. O fato que deve ser evidenciado aqui é que pode existir uma probabilidade vinculada a valores incertos por existência de imprecisão, uma espécie de *probabilidade difusa*. De acordo com o modelo de probabilidade condicional $P(H_i/e)$, sabe-se que H_i representa um vetor de probabilidades de ocorrência das hipóteses condicionadas às evidências e , desta forma pode-se partir do princípio que haja imprecisão nas evidências, fazendo com que esta imprecisão torne também imprecisa a probabilidade das hipóteses, ou seja:

 dado um intervalo $I[x_i, x_j]$, sendo que I é subintervalo de probabilidades $[0, 1]$, considera-se uma probabilidade $p\%$ munida de imprecisão como sendo um número $p \in I$, assim p passa a representar uma *probabilidade de evento difuso*.

Esta idéia de *qualificar* ou *ajustar* uma probabilidade condicionada à evidências imprecisas é que justifica o que se chama de *modelagem difuso-probabilística*, nesta pesquisa.

9. MODELO HÍBRIDO

9.1. INTRODUÇÃO

A idéia central desta pesquisa é verificar a possibilidade de integrar duas técnicas para modelagem de sistemas computacionais inteligentes: os Conjuntos Difusos e as Redes de Probabilidades. Como já abordado com maior ênfase em capítulos anteriores, estas técnicas são aplicadas na concepção de sistemas especialistas, sendo que, cada uma possui um enfoque diferente quanto à interpretação do resultado final (*output*). Enquanto um modelo baseado em Conjuntos Difusos verifica a *pertinência* no intervalo $[0, 1]$ de um determinado parâmetro de entrada, uma rede bayesiana verifica a *probabilidade* no mesmo intervalo. Desta forma é que classificam-se Conjuntos Difusos como um *modelo possibilístico* e, Redes Bayesianas como um *modelo probabilístico*. A utilização do termo *híbrido* se dá justamente ao fato da integração ou combinação das duas técnicas citadas.

A motivação para a elaboração e estruturação deste trabalho se deu pelo fato do conhecimento prévio sobre as atividades realizadas no campo da Ciência dos Alimentos, mais precisamente nos estudos sobre a contaminação do leite de consumo causada pelos resíduos de antibióticos administrados em animais, tratamento de lavoura ou industrialização. Conforme citado na introdução, existem métodos de diagnosticar a presença destes resíduos no leite. Sabendo que uma das maneiras de fazer este diagnóstico é por meio da verificação da concentração de quantidades de antibióticos numa determinada amostra de leite em teste, e que estas quantidades variam num intervalo numérico cada qual relativo a um tipo de antibiótico, observa-se uma forma de utilizar Conjuntos Difusos para calcular a pertinência de uma amostra aos resultados dados pelos especialistas do domínio pesquisado. Por outro lado, sabe-se que a contaminação do leite por estas quantidades de antibióticos leva em discussão aspectos ligados à saúde pública em termos toxicológicos e, também, a repercussão sobre o controle da qualidade do leite consumido, sendo assim, pretende-se utilizar as redes bayesianas para determinar probabilidades de doenças com base em evidências sobre presença ou não de tipos de antibióticos no leite analisado.

Tendo em vista que as quantidades de antibióticos presentes numa amostra de leite analisado possuem imprecisões e, que estas quantidades são as evidências usadas pela rede bayesiana ao calcular a probabilidade de ocorrência de uma doença com base nestas evidências, torna-se oportuno encontrar um *qualificador ou ajustador* da saída probabilística, pois esta probabilidade está vinculada a uma imprecisão.

9.2. ESQUEMATIZAÇÃO E FERRAMENTAS AUXILIARES

Esta unidade tem por objetivo detalhar o modelo que se propôs especificar. Foi elaborado um protótipo para permitir um melhor entendimento dos objetivos desejados com a modelagem, bem como a simulação de casos. A fig. 13 mostra um esquema representativo do modelo proposto, envolvendo dois subsistemas: um probabilístico e outro difuso. A idéia do esquema é justamente tornar possível a compreensão da existência de uma fusão entre as duas técnicas, o que caracteriza o *hibridismo* já mencionado.

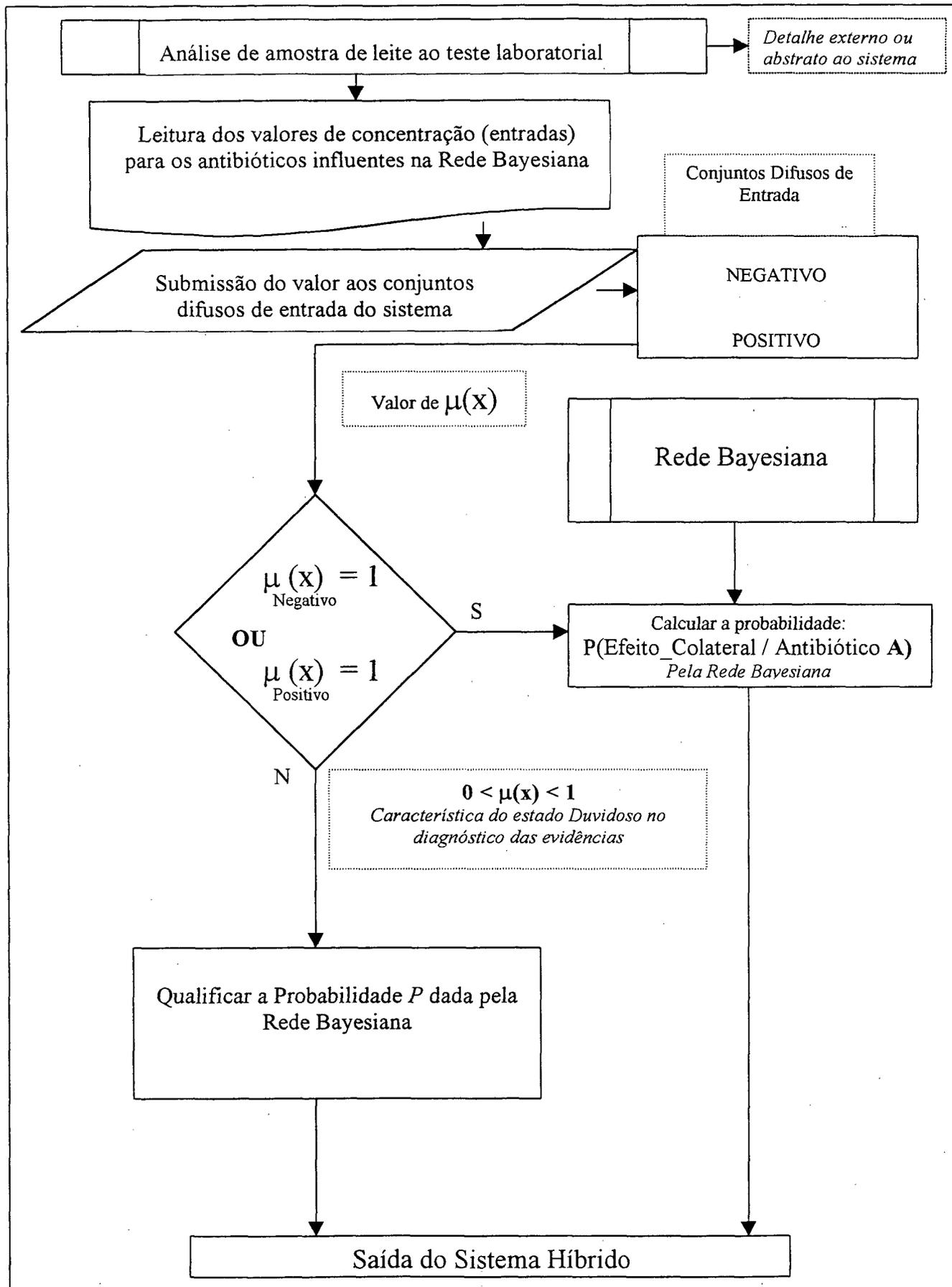


Figura 13: Esquema representativo do modelo híbrido proposto.

O objetivo principal na simulação é obter uma saída que será interpretada como a *probabilidade qualificada* de acontecer um determinado efeito colateral em uma pessoa que consome (sem considerar o período) uma quantidade de leite que esteja contaminado por resíduos de antibióticos. Pelo fato da extensão do problema, bem como dos inúmeros tipos de antibióticos existentes e que fazem parte do contexto deste domínio de aplicação, o modelo irá se concentrar nos poucos antibióticos mais evidentes que foram levantados, da mesma forma, com somente algumas das reações adversas causadas por estes antibióticos.

O objetivo da rede bayesiana construída para a simulação é de gerar *probabilidades* de certos efeitos colaterais segundo tipos de antibióticos. É exatamente aqui que pode-se enfatizar o ponto chave da proposta.

Quando a rede bayesiana calcula uma probabilidade **p** de acontecer a doença **D** caso existam evidências **e**, de antibióticos, este valor de **p** é fundamentado, de acordo com a teoria da probabilidade, de que $P(e)=1$ (existe) ou, $P(e)=0$ (não existe), sendo assim, a rede bayesiana não considera se existe uma gradação para a existência de quantidades de antibióticos no leite, pois, pode-se dizer que no probabilístico, o grau de verdade é zero ou 1. Neste espaço, no qual o modelo probabilístico não pode apresentar uma saída considerando não somente os valores 0 e 1, mas sim, os valores compreendidos no intervalo $[0, 1]$ é que propõem-se o uso de conjuntos difusos que possam calcular um valor o qual pode-se chamar de **Ajustado ou Qualificado** para o probabilístico, assim, se a probabilidade de acontecer a doença **D** for **p**, mas a possibilidade de existência da evidência **e** estiver compreendida no intervalo $]0, 1[$, então o valor de **p** sofre uma perturbação de modo a acontecer uma redistribuição das probabilidades no vetor de hipóteses ou saída da rede. Assim, cada nova probabilidade deste vetor está vinculada à imprecisão existente na evidência do antibiótico, influenciando o diagnóstico de determinada doença.

Para fins de aprimoramento das idéias e, antes de se ter clareza dos objetivos, parâmetros e do fluxo dos procedimentos a serem implementados no protótipo proposto para a demonstração do hibridismo, foram efetuadas operações de pré-simulação em duas plataformas para implementação de sistemas especialistas: uma para construir a rede bayesiana que representa o conhecimento do especialista no domínio da aplicação escolhida e, outra para testar os conjuntos difusos e as pertinências calculadas das variáveis de entrada do sistema a ser abordado.

A fig. 14 mostra uma Rede Bayesiana implementada na Shell Netica da Norsys Software Corp, a qual possui alguns efeitos adversos e seus antibióticos mais influentes, de acordo com especificações técnicas.

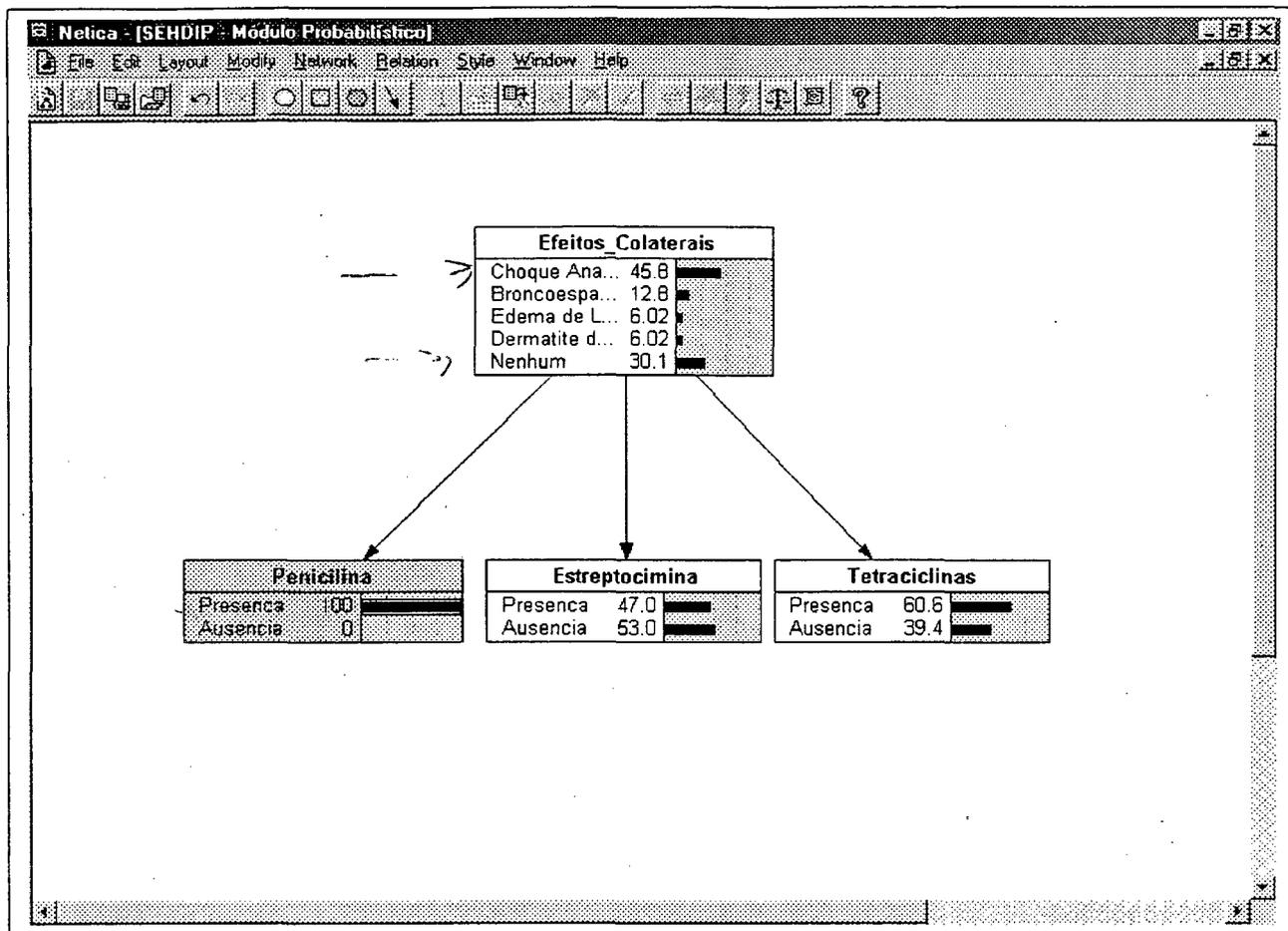


Figura 14: Uma Rede Bayesiana de aplicação.

Com esta rede pode-se ver o vetor de hipóteses diagnósticas ou vetor das variáveis de saída do sistema (nodo superior da rede) e, os nodos que representam as variáveis de entrada do sistema ou, os eventos da rede. Em cada nodo tem-se a distribuição das probabilidades que representam o conhecimento do especialista. Para fins de prototipação, a rede bayesiana está formada por:

1. Um nodo representando o vetor de hipóteses ou saídas da rede, colocados da seguinte forma:
 - a. Choque Anafilático
 - b. Broncoespasmo
 - c. Edema de Laringe
 - d. Dermatite de Contato
 - e. Nenhum

Todos com probabilidades iniciais de 20%. Isso se dá ao fato da dificuldade de obter-se as probabilidades à priori por parte de especialistas. Desta maneira, a rede começa a executar com a máxima entropia na base de conhecimento referente às hipóteses diagnósticas.

2. Três nodos representando as variáveis de entrada (antibióticos):

- a. Penicilina
- b. Estreptomicina
- c. Tetraciclina

Cada uma com dois estados - Presença e Ausência.

As tabelas seguintes mostram as probabilidades de cada variável de entrada condicionada às hipóteses diagnósticas.

Tabela 2: Probabilidades de Penicilina condicionadas às hipóteses diagnósticas.

HIPÓTESES DIAGNÓSTICAS	PRESENÇA %	AUSÊNCIA %
Choque Anafilático	76	24
Broncoespasmo	20	80
Edema de Laringe	10	90
Dermatite de Contato	10	90
Nenhum	50	50

Tabela 3: Probabilidades de Estreptomicina condicionadas às hipóteses diagnósticas.

HIPÓTESES DIAGNÓSTICAS	PRESENÇA %	AUSÊNCIA %
Choque Anafilático	50	50
Broncoespasmo	50	50
Edema de Laringe	50	50
Dermatite de Contato	1	99
Nenhum	50	50

Tabela 4: Probabilidades de Tetraciclina condicionadas às hipóteses diagnósticas.

HIPÓTESES DIAGNÓSTICAS	PRESENÇA %	AUSÊNCIA %
Choque Anafilático	60	40
Broncoespasmo	85	15
Edema de Laringe	80	20
Dermatite de Contato	50	50
Nenhum	50	50

De acordo com a fig. 14 observe que, se a probabilidade da existência do antibiótico penicilina for 100%, então, sem considerar a influência dos demais antibióticos (Presença = 100%), o valor 45.8% seria calculado para a probabilidade de ocorrência de choque anafilático a uma pessoa que consome leite contaminado por tal antibiótico. Leva-se em consideração então que o valor 45.8% está baseado na existência do antibiótico penicilina (fator probabilístico). Assim, ao considerar que existem dúvidas quanto à presença ou não de antibióticos no leite analisado, percebe-se a existência de *nebulosidade* que leva o termo *duvidoso* (usado pelos especialistas) a influenciar o número 45.8% tornando-o um *número difuso*. Desta forma, quando a pertinência dos valores de concentração dos antibióticos em relação aos conjuntos NEGATIVO ou POSITIVO não for igual a 1 (um), então existe uma margem de nebulosidade que faz o valor 45.8% ser reajustado dentro do espaço de probabilidades, influenciado pela imprecisão explícita na pertinência da quantidade deste antibiótico aos conjuntos de entrada.

Sendo assim, o fato da existência do termo DUVIDOSO na interpretação do diagnóstico pelos especialistas no domínio desta aplicação é a principal motivação para verificar se valores probabilísticos podem ser influenciados por valores possibilísticos, ou seja, tenta-se aqui certificar-se de que o probabilístico sofre um determinado *ruido* quando há existência de nebulosidade, no caso, nas evidências.

Na parte que cabe ao módulo difuso e, para efeito de simulação, propõem-se como entrada, 3 (três) variáveis:

1. Concentração de Penicilina (valores em UI/g)
2. Concentração de Estreptomicina (valores em ug/ml)
3. Concentração de Tetraciclina (valores em ug/ml)

Cada uma destas variáveis possui dois conjuntos difusos, não necessariamente complementares:

1. Negativo (no caso da ausência confirmada de antibióticos na amostra analisada)
2. Positivo (no caso da presença confirmada de antibióticos na amostra analisada).

A tabela 5 mostra as concentrações dos antibióticos mais comuns, segundo especificações coletadas em (SOUZA, 1998).

Tabela-5: Antibióticos comuns no estudo de caso e suas concentrações.

CONCENTRAÇÕES DE ANTIBIÓTICOS	PRESENÇA DE ANTIBIÓTICOS					
	NEGATIVO		DUVIDOSO		POSITIVO	
Concentrações	Mínima	Máxima	Mínima	Máxima	Mínima	Máxima
Penicilinas (UI/g)	/	0.002	> 0.002	0.004	0.005	Acima
Estreptomicina (ug/ml)	/	4	> 4	< 13	13	Acima
Tetraciclina (ug/ml)	/	0.1	> 0.1	< 0.4	0.4	Acima

Os valores na tabela 5 foram usados no suporte aos conjuntos difusos respectivamente às suas variáveis de entrada.

O módulo difuso não apresenta saída em termos de um sistema difuso completo (entradas, fuzificação, defuzificação, saída), ou seja, não se justifica como diagnóstico, uma saída difusa para o propósito estabelecido. As pertinências das entradas calculadas por este módulo servirão como parâmetros para o processo de *qualificação*, ou seja, irão compor o módulo híbrido, o qual fará um ajuste na probabilidade dada pela Rede Bayesiana, isto é, a redistribuição das probabilidades à partir dos eventos difusos.

Utilizou-se a ferramenta shell Unfuzzy para obtenção das pertinências das entradas com objetivo de comparar os resultados com aqueles obtidos nas simulações feitas com o protótipo desenvolvido para esta aplicação. As figuras seguintes exibem algumas partes que compõem o módulo difuso construído com o software Unfuzzy, com o único intuito de testar os conjuntos e algumas funções de pertinência para agregar ao protótipo final.

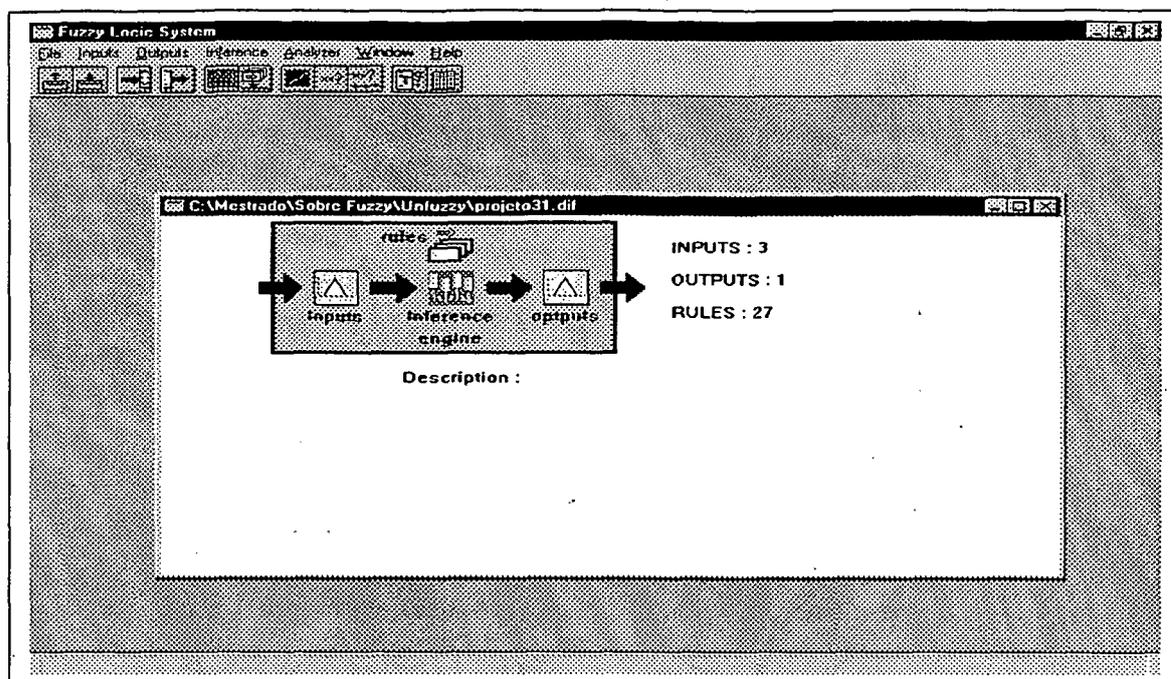


Figura 15: Interface principal indicando o número de variáveis de entrada.

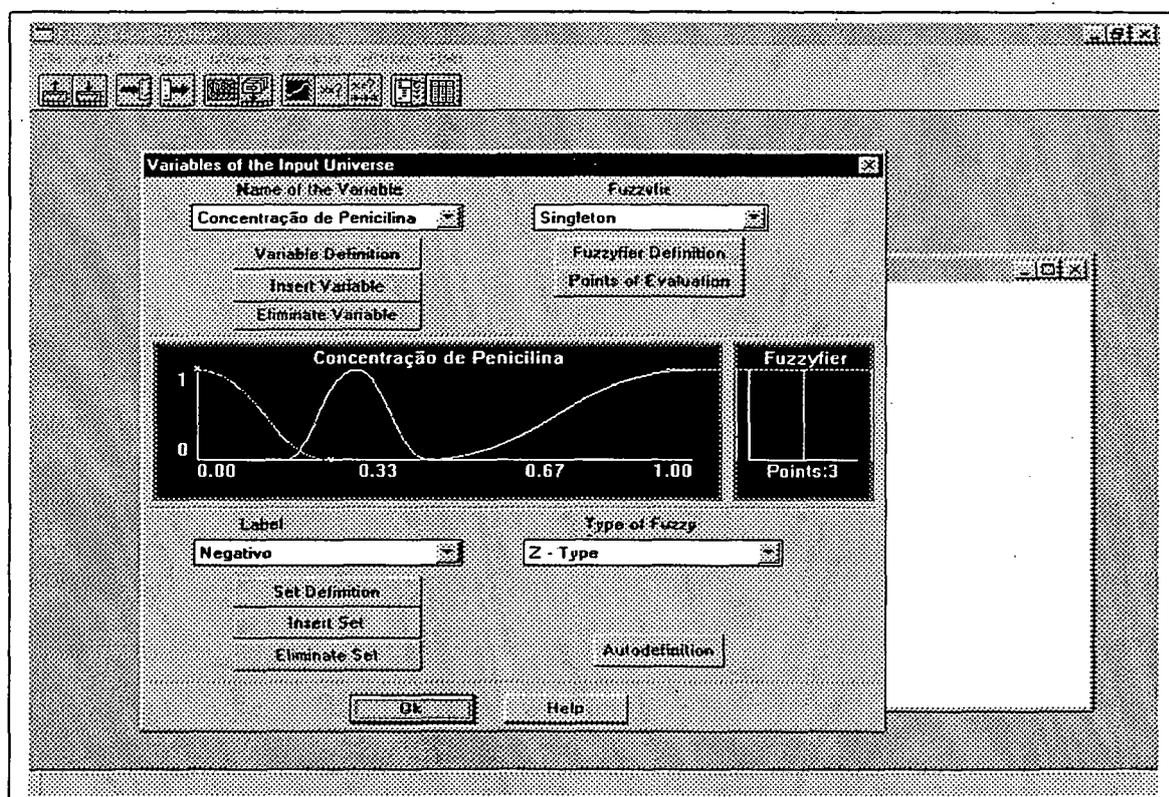


Figura 16: Especificação das variáveis de entrada e seus conjuntos difusos.

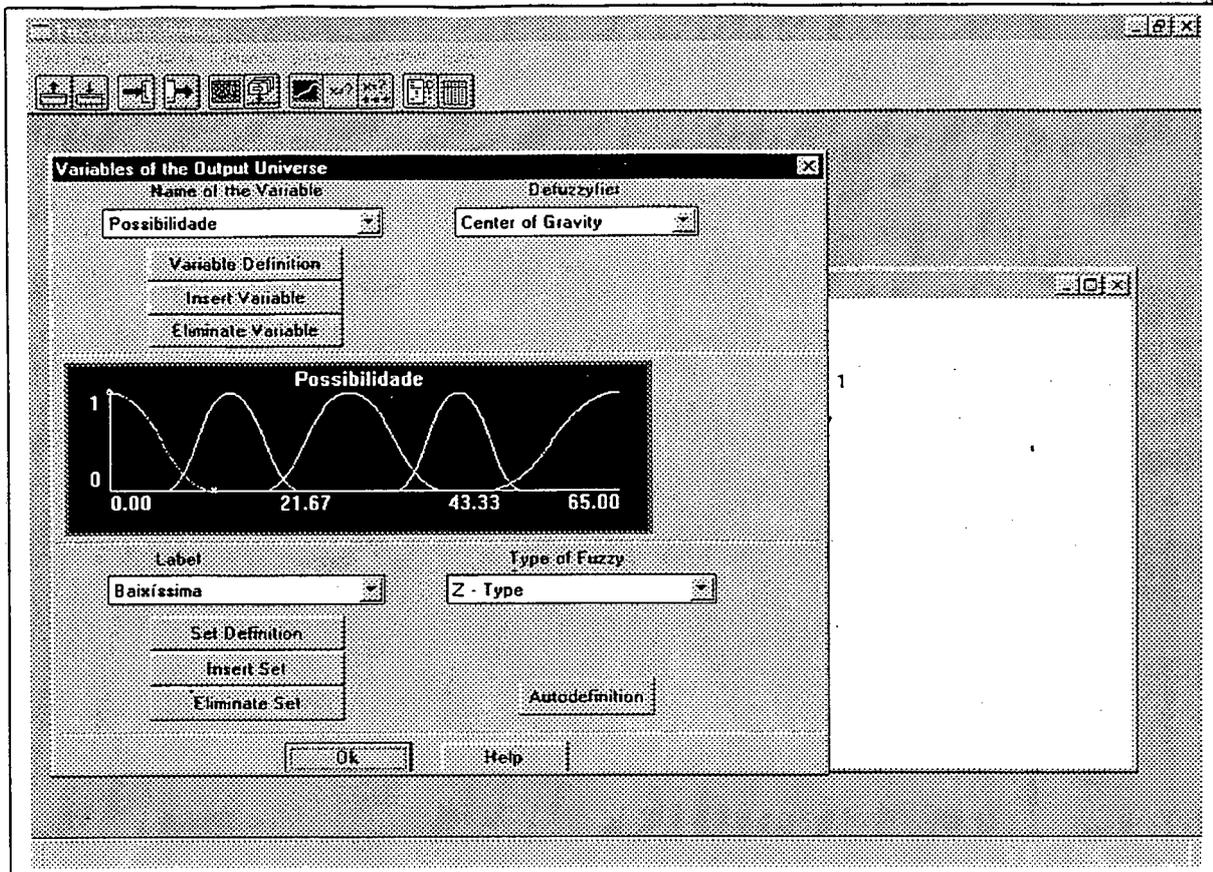


Figura 17: Especificação da variável de saída e seus conjuntos difusos.

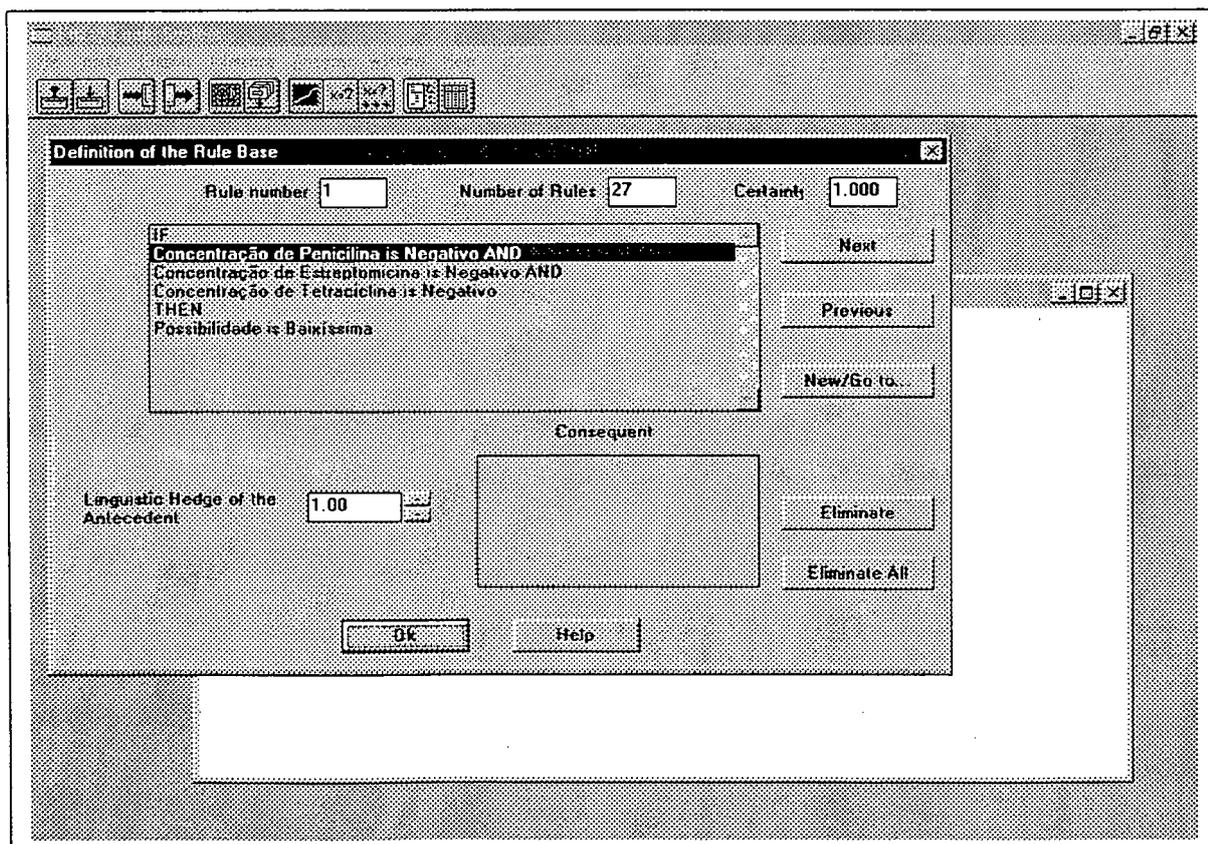


Figura 18: Especificação das regras.

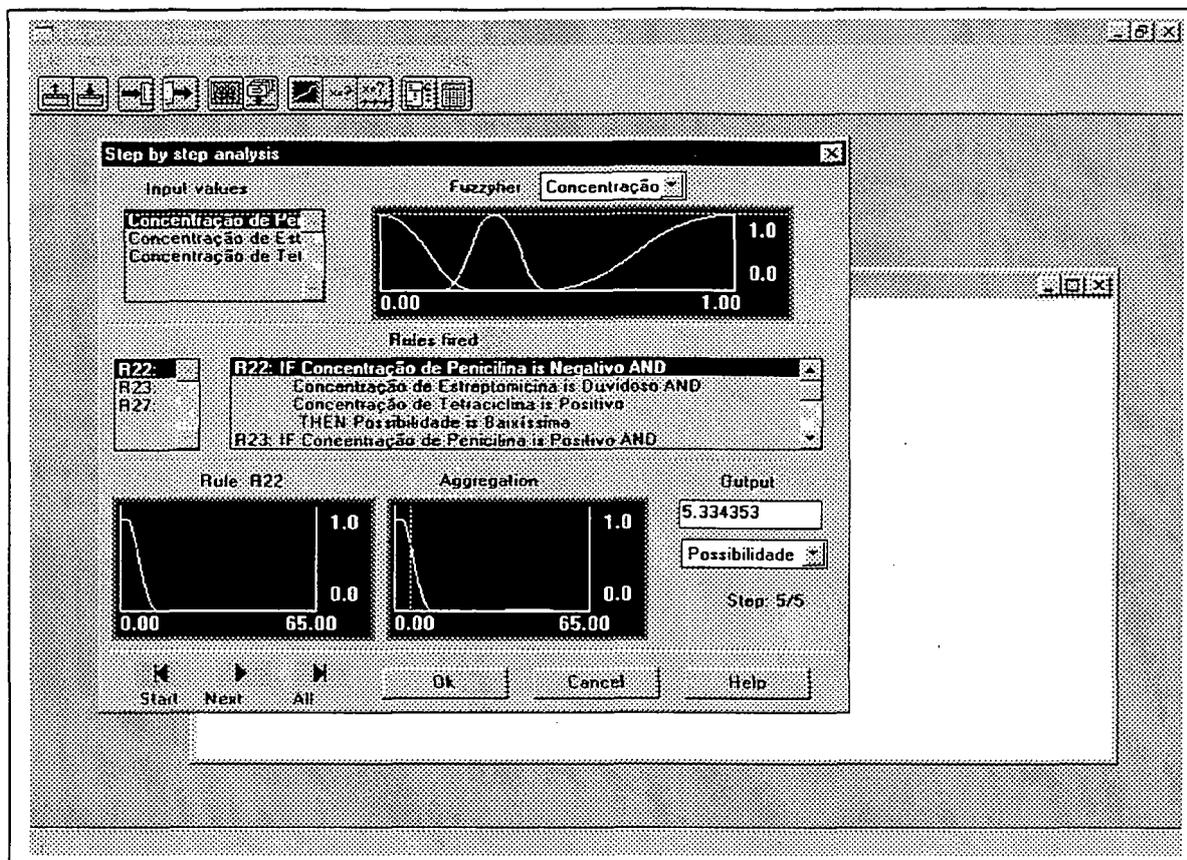


Figura 19: Simulação passo a passo de uma entrada ao sistema.

É conveniente neste ponto da proposta ressaltar a importância do módulo *híbrido*, quando se refere ao *qualificador* ou *ajustador* de um valor probabilístico dado como saída do diagnóstico.

Ao considerar que cada evidência que compõem a rede de probabilidades pode possuir imprecisão, então nota-se que estas mesmas evidências são responsáveis pelo aumento da nebulosidade do diagnóstico probabilístico, ou seja, cada tipo de antibiótico especificado na rede bayesiana construída que deixar de ter sua pertinência $\mu(x) = 1$ aos conjuntos difusos de entrada (NEGATIVO ou POSITIVO), contribuirá para que o nível de qualificação da rede seja cada vez maior, ou seja, quanto mais eventos difusos, maior será a perturbação na rede e, conseqüentemente, maior será a redistribuição das probabilidades no processo de qualificação para as hipóteses diagnósticas.

Para melhor entendimento deste *qualificador*, considera-se hipoteticamente a existência de 3 (três) tipos de antibióticos (A, B, C) e suas pertinências aos conjuntos difusos de entrada (NEGATIVO e POSITIVO), assim define-se três níveis de *qualificação*:

- Se existir apenas um $\mu(x) < 1$ então a saída probabilística será *fracamente qualificada*;
- Se existirem apenas dois $\mu(x) < 1$ então a saída probabilística será *moderadamente qualificada*;
- Se existirem três $\mu(x) < 1$ então a saída será *fortemente qualificada*.

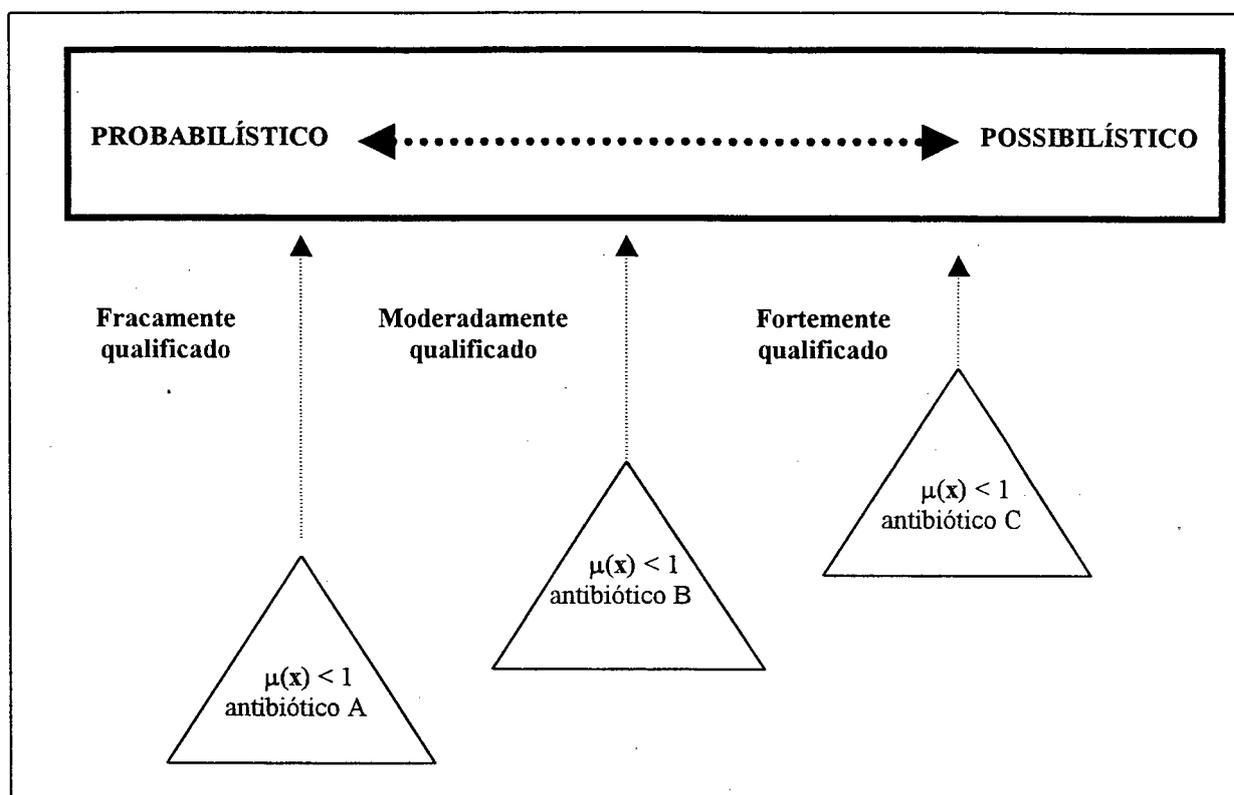


Figura 20: Hibridismo do probabilístico-possibilístico em função da qualificação das entradas.

A fig. 20 mostra a possível transição de estados de incertezas no diagnóstico influenciadas pelo *qualificador* do módulo híbrido na modelagem proposta.

É notável que com tal processo de qualificação os valores das probabilidades se aproximam de *possibilidades*, mas é necessário ressaltar que o *qualificador* deverá ajustar as saídas produzidas pela Rede Bayesiana de modo a mantê-la no intervalo $[0, 1]$.

9.3. O QUALIFICADOR PARA O MODELO HÍBRIDO

Seja n o número de variáveis de saída da Rede Bayesiana e, seja e , um evento não-difuso. Pelo Teorema de Bayes sabe-se que:

$$P(H_i / e) = \frac{P(e / H_i) \cdot P(H_i)}{P(e)} \quad \text{onde,} \quad (12)$$

$$P(e) = P(e / H_i) + P(e / \neg H_i) \quad (13)$$

Esta 4ª é a forma que caracteriza as saídas da Rede Bayesiana como *puramente probabilísticas*, ou seja, não existe *ruido* ou *perturbação* pelo fato da não imprecisão intrínseca nas entradas (eventos e). Pode-se aqui dizer que ao invés de se estar calculando probabilidades sobre eventos difusos, estes eventos são *Crisp*.

Considere agora a existência de eventos difusos compondo a Rede Bayesiana. Assim, δ representa um evento munido de imprecisão, ou seja, a *possibilidade* de existência de δ está no intervalo $]0,1[$. Sendo assim, apresenta-se $\phi = f(\rho, \delta)$ como uma notação matemática que, expressa sob a forma de função, um qualificador para o modelo híbrido.

Entende-se então que ϕ é que causará a redistribuição das probabilidades da Rede Bayesiana tendo ρ como uma probabilidade condicional sem imprecisão e, δ representando um ou mais eventos difusos. Para tal, utilizou-se a fórmula do Teorema de Bayes com algumas modificações em sua forma original, conforme segue:

Se n é o número de hipóteses diagnósticas da Rede Bayesiana, então para variáveis de entrada com dois estados (conforme o modelo proposto: Presença, Ausência) têm-se:

$$P(H_i / \delta) = \frac{P(\delta / H_i)}{P(\delta)} \quad \text{onde,} \quad (14)$$

$$P(\delta/H_i) = P(H_i) \cdot [P(\delta/H_i) \cdot \mu_{\mathcal{E}_1}(\delta) + (1-P(\delta/H_i)) \cdot \mu_{\mathcal{E}_2}(\delta)] \quad (15)$$

e,

$$P(\delta) = \sum_{i=1}^n P(H_i) \cdot [P(\delta/H_i) \cdot \mu_{\mathcal{E}_1}(\delta) + (1-P(\delta/H_i)) \cdot \mu_{\mathcal{E}_2}(\delta)] \quad (16)$$

Os símbolos \mathcal{E}_1 e \mathcal{E}_2 foram usados para representar respectivamente os conjuntos difusos de entrada, Ausência e Presença, para os quais são submetidos os valores de entrada para o modelo híbrido, no caso, as quantidades dos antibióticos especificados. No caso tratado, o conjunto Presença é complementar do conjunto Ausência, ou seja, tal complementaridade é referente ao padrão existente na Teoria da Probabilidade, dentro do intervalo $[0,1]$ e que, na Teoria dos Conjuntos Difusos é apenas um caso específico, pois nesta, a soma do complemento de um conjunto não necessariamente deve ser unitária, mas, pode ser maior ou menor do que a unidade.

É importante ressaltar que após ser aplicada a função ϕ de qualificação, a propriedade $\sum_{i=1}^n P(H_i) = 1$ continua verdadeira, bem como, a soma das pertinências das entradas não necessariamente tenha que somar 1, de acordo com o *complemento* padrão existente na Teoria dos Conjuntos Difusos.

Em termos de generalização da fórmula 15, estende-se o caso para três ou mais estados de cada variável de entrada e , naturalmente para o mesmo número de conjuntos difusos para os quais cada um destes estados seja submetido. Assim, teremos:

$$P(\delta/H_i) = P(H_i) \cdot [P_1(\delta/H_i) \cdot \mu_{E_1}(\delta) + P_2(\delta/H_i) \cdot \mu_{E_2}(\delta) + \dots + P_r(\delta/H_i) \cdot \mu_{E_r}(\delta)] \quad (17)$$

assim, se s representa o número de estados para cada variável de entrada, então:

$$P(\delta/H_i) = P(H_i) \cdot \sum_{r=1}^s P_r(\delta/H_i) \cdot \mu_{E_r}(\delta) \quad (18)$$

e,

$$P(\delta) = \sum_{i=1}^n P(H_i) \cdot \sum_{r=1}^s P_r(\delta/H_i) \cdot \mu_{E_r}(\delta) \quad (19)$$

Para tornar mais ampla a validação do modelo híbrido proposto, mostra-se uma alteração na estrutura das fórmulas 15 e 16, de modo a considerar o cálculo de hipóteses diagnósticas à partir de vários eventos difusos, ou seja, $\delta = \{e_1, e_2, \dots, e_m\}$. No caso desta aplicação, o número máximo de variáveis de entrada (eventos) e , que conseqüentemente poderiam ser eventos difusos é igual a 3 (três) e, considera-se neste modelo que cada variável de entrada tenha 2 (dois) estados.

Desta forma, seja n o número de hipóteses diagnósticas da Rede Bayesiana e, m o número de eventos difusos desta Rede, assim:

$$\begin{aligned}
P(e_1 \cap e_2 \cap \dots \cap e_m / H_i) &= P(H_i) \cdot ([P(e_1 / H_i) \cdot \mu_{\xi_1}(e_1) + (1 - P(e_1 / H_i)) \cdot \mu_{\xi_2}(e_1)] \times \\
&\times [P(e_2 / H_i) \cdot \mu_{\xi_1}(e_2) + (1 - P(e_2 / H_i)) \cdot \mu_{\xi_2}(e_2)] \times \dots \times \\
&\times [P(e_m / H_i) \cdot \mu_{\xi_1}(e_m) + (1 - P(e_m / H_i)) \cdot \mu_{\xi_2}(e_m)])
\end{aligned} \tag{20}$$

então, se $\delta = \{e_1, e_2, \dots, e_m\}$ teremos:

$$P(\delta / H_i) = P(H_i) \cdot \prod_{k=1}^m [P(e_k / H_i) \cdot \mu_{\xi_1}(e_k) + (1 - P(e_k / H_i)) \cdot \mu_{\xi_2}(e_k)] \tag{21}$$

e,

$$P(\delta) = \sum_{i=1}^n P(H_i) \cdot \prod_{k=1}^m [P(e_k / H_i) \cdot \mu_{\xi_1}(e_k) + (1 - P(e_k / H_i)) \cdot \mu_{\xi_2}(e_k)] \tag{22}$$

É evidente que pode-se pensar na aplicação ou extensão das fórmulas 17, 18 e 19 para as fórmulas 20, 21 e 22 respectivamente, tornando possível a realização do hibridismo da Rede Bayesiana nos casos em que existe três ou mais estados para cada evento difuso, bem como para o número de conjuntos difusos relacionados a estes eventos.

No módulo difuso que compõem o protótipo de demonstração desta modelagem híbrida, foram utilizadas, apenas as funções de pertinências para as entradas, conforme já mencionado em unidades anteriores não necessitou-se, pelo menos, nestes estudos, de um sistema difuso completo com formação de regras e processo de defuzificação. Para tal, utilizou-se funções triangulares e sigmóides, assim, pode-se configurar uma destas opções durante a simulação de entradas ao sistema híbrido com objetivo de comparar resultados.

Com as funções triangulares ficaram as variáveis e seus conjuntos na forma:

Tabela 6: As funções de pertinência das variáveis de entrada sob a forma triangular.

Variável	Conjuntos/Funções	
	Ausência	Presença
Penicilina	$y = -333.333x + 1.667$	$y = 333.333x - 0.667$
Estreptomicina	$y = -0.111x + 1.444$	$y = 0.111x - 0.444$
Tetraciclina	$y = -3.333x + 1.333$	$y = 3.333x - 0.3332$

Para uma modelagem mais similar à forma natural do raciocínio aproximado do especialista, é notável a utilização de outras funções de pertinência, no caso, não lineares. Em termos desta aplicação e, para o entendimento e simulação da proposta do hibridismo, utilizou-se apenas a forma triangular.

É notável, que os conjuntos difusos para todas as variáveis de entrada são dicotômicos, mas, isso não limita a idéia de que o qualificador especificado anteriormente não possa ser aplicado nos casos em que as variáveis de entrada possuem mais do que dois estados, apenas para isso, seria necessário considerar vários estados onde nas fórmulas 15, 16, 20, 21 e 22 utilizou-se apenas os estados ε_1 e ε_2 . Seria então o caso de se especificar $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_j$, para j estados de uma variável de entrada.

9.4. A INTERFACE E DETALHES DE IMPLEMENTAÇÃO DO PROTÓTIPO

Para a demonstração das idéias introduzidas neste trabalho e, para simulação e obtenção de resultados, foi desenvolvida uma versão prototipada de um sistema híbrido, ou seja, pretende-se aqui deixar claro que o objetivo principal não é desenvolver uma ferramenta Shell para a implementação de Sistemas Especialistas Híbridos, mas, lançar a idéia de um modelo matemático que possa ser incorporado às técnicas de construção de Sistemas Especialistas.

O protótipo foi desenvolvido no ambiente Borland DELPHI Enterprise 5.0 (Build 5.62) onde para o armazenamento do conhecimento da Rede Bayesiana foi utilizada a plataforma de Banco de Dados PARADOX versão 7.0. No total foram implementadas 6 tabelas, as quais armazenavam as probabilidades iniciais do vetor de hipóteses diagnósticas as probabilidades condicionais e ainda, os cálculos resultantes das operações do módulo difuso.

As figuras seguintes mostrarão os formulários que compõem o protótipo.



Figura.21: Formulário de entrada do protótipo.

Simulador Difuso-Probabilístico >>>> SEHIB v1.0

Entradas

Variáveis de entrada:

Penicilina (UI/g)

Estreptomina (ug/ml)

Tetraciclina (ug/ml)

Figura 22: Formulário para entrada das quantidades de antibióticos.

Simulador Difuso-Probabilístico >>>> SEHIB v1.0 <<<<<

Relações de Pertinências

AUSÊNCIA PRESENCIA

Penicilina (UI/g)

Estreptomina (ug/ml) $\mu(x)$

Tetraciclina (ug/ml)

Evidências (tipo de informação)

CRISP

DUFUSA

Distribuição de probabilidades da Rede Bayesiana

SAIDAS ou HIPÓTESES DIAGNÓSTICAS

Choque Anafilático	20%	<input type="text"/>
Broncoespasmo...	20%	<input type="text"/>
Edema de Laringe	20%	<input type="text"/>
Dermatite de Contato	20%	<input type="text"/>
Nenhuma	20%	<input type="text"/>

ENTRADAS

Penicilina	Estreptomina	Tetraciclina
Presença 33.2% <input type="text"/>	Presença 40.2% <input type="text"/>	Presença 65% <input type="text"/>
Ausência 66.8% <input type="text"/>	Ausência 59.8% <input type="text"/>	Ausência 35% <input type="text"/>

Amostras para simulação

Penicilina_qtd	Estreptomina_qtd	Tetraciclina_qtd	
0,0025	15		C
0,0039	12		C
0,0039	12		C
0,0038	15		C

Hipóteses Diagnósticas (sem qualificação)

Hipóteses Diagnósticas (sem qualificação)

Figura 23: Formulário para demonstração dos procedimentos de qualificação da Rede Bayesiana.

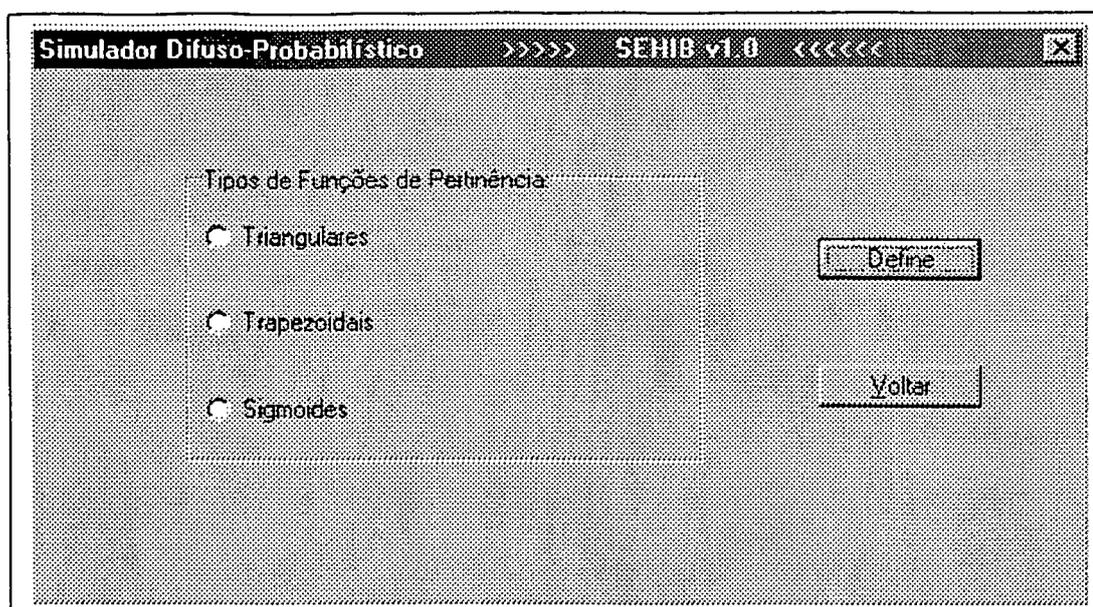


Figura 24: Formulário para escolha do formato das funções de pertinência.

Outros formulários sem influencia direta ao modelo não foram aqui exibidos.

9.5. SIMULAÇÕES E ILUSTRAÇÕES DE CÁLCULOS

Inicialmente serão mostrados alguns resultados calculados pela Rede Bayesiana sem o processo de qualificação. Após, serão simuladas algumas entradas de eventos difusos e ilustrados os cálculos que condizem com o hibridismo proposto.

A figura seguinte mostra a distribuição inicial das probabilidades da Rede Bayesiana.

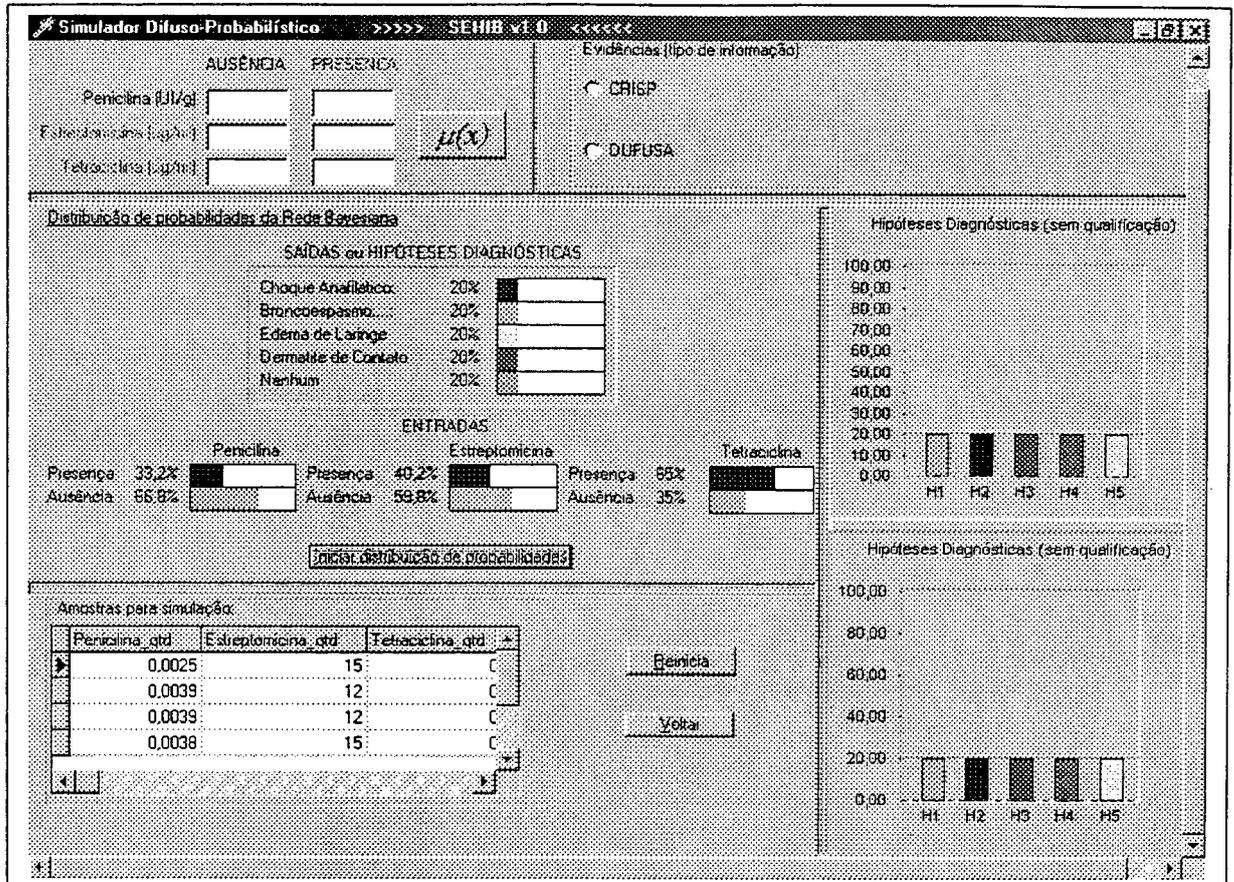


Figura 25: Distribuição inicial das probabilidades da Rede.

Na fig. 25 são mostrados, o vetor de hipóteses diagnósticas composto pelos rótulos: Choque Anafilático, Broncoespasmo, Edema de Laringe, Dermatite de Contato, Nenhum, cada um com a probabilidade inicial de 20%, ou seja, a rede inicia operando sobre o princípio da máxima entropia. Aparecem também as variáveis de entrada: Penicilina, Estreptomycina e Tetraciclina, cada uma com dois estados: Ausência e Presença, e suas probabilidades condicionadas às hipóteses. Para que se entenda como foram obtidos os valores das probabilidades destas entradas, serão ilustrados três exemplos:

1. Qual a probabilidade da Presença de Penicilina ?

A tabela abaixo mostra as probabilidades $P(e/H_i)$ para a variável penicilina:

Tabela 7: Probabilidades $P(e/H_i)$ para Penicilina.

Hipótese	Presença	Ausência
→ Choque Anafilático	76%	24%
Broncoespasmo	20%	80%
Edema de Laringe	10%	90%
Dermatite de Contato	10%	90%
Nenhum	50%	50%

Então, se $P(e) = P(e/H_i) + P(e/\neg H_i)$, têm-se:

$$P(\text{Presença_Penicilina}) = (0.76 \times 0.2) + (0.2 \times 0.2) + (0.1 \times 0.2) + (0.1 \times 0.2) + (0.5 \times 0.2) = 0.332 = 33.2\%.$$

2. Qual a probabilidade da Ausência de Estreptomicina,?

A tabela abaixo mostra as probabilidades $P(e/H_i)$ para a variável Estreptomicina:

Tabela 8: Probabilidades $P(e/H_i)$ para Estreptomicina.

Hipótese	Presença	Ausência
Choque Anafilático	50%	50%
Broncoespasmo	50%	50%
Edema de Laringe	50%	50%
Dermatite de Contato	1%	99%
Nenhum	50%	50%

Assim,

$$P(\text{Ausência_Estreptomicina}) = (0.5 \times 0.2) + (0.5 \times 0.2) + (0.5 \times 0.2) + (0.01 \times 0.2) + (0.5 \times 0.2) = 0.402 = 40.2\%.$$

3. Qual a probabilidade da Presença de Tetraciclina ?

A tabela abaixo mostra as probabilidades $P(e/H_i)$ para a variável Tetraciclina:

Tabela 9: Probabilidades $P(e/H_i)$ para Tetraciclina.

Hipótese	Presença	Ausência
Choque Anafilático	60%	40%
Broncoespasmo	85%	15%
Edema de Laringe	80%	20%
Dermatite de Contato	50%	50%
Nenhum	50%	50%

Assim,

$$P(\text{Presença_Tetraciclina}) = (0.6 \times 0.2) + (0.85 \times 0.2) + (0.8 \times 0.2) + (0.5 \times 0.2) + (0.5 \times 0.2) = 0.65 = 65\%.$$

A figura seguinte mostra os resultados calculados considerando-se que os eventos são: Presença de Penicilina e Presença de Estreptomicina.

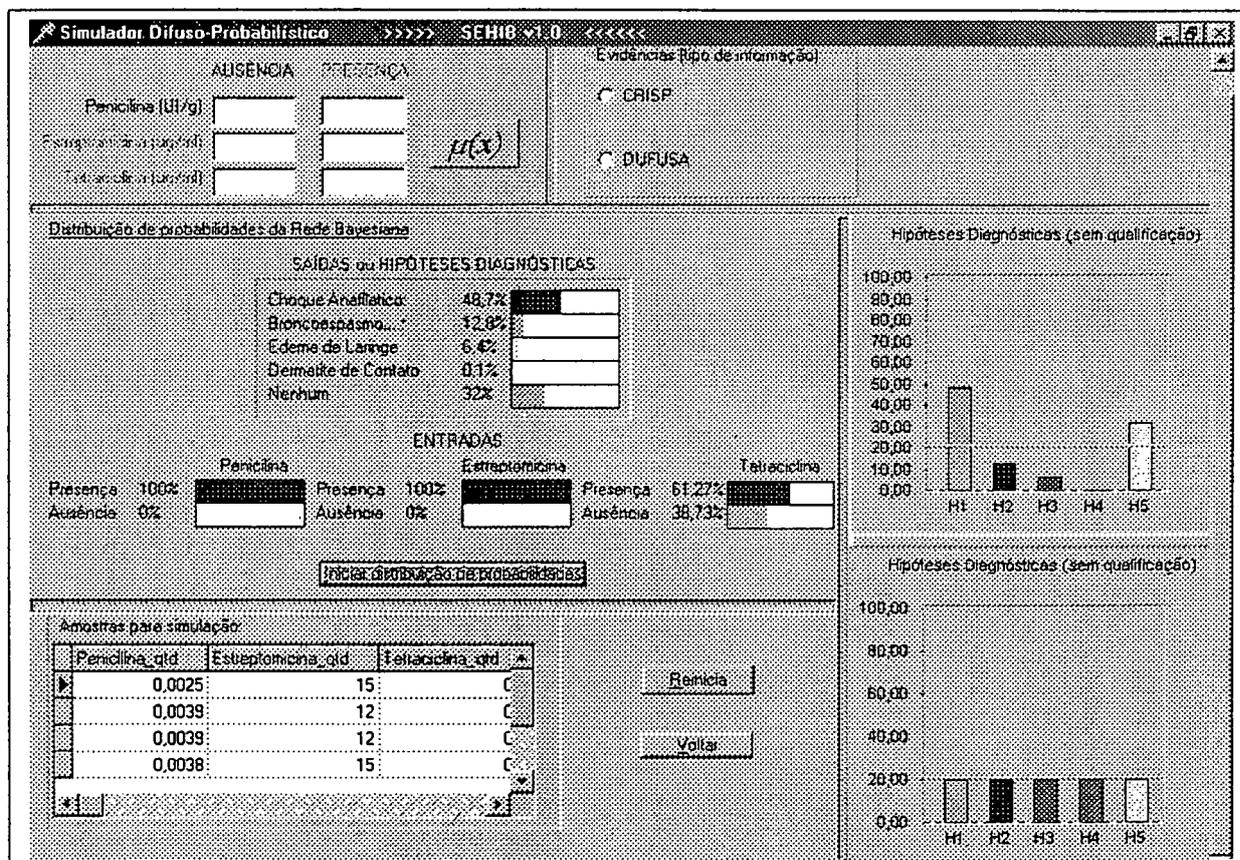


Figura 26: Probabilidades condicionadas pela Presença de Penicilina e Estreptomicina.

Para calcular a probabilidade de cada hipótese dada a condição da Presença de Penicilina e Estreptomicina, a Rede aplicou a fórmula $P(H_i/e_1 \cap e_2)$. Assim, os procedimentos de cálculo foram:

$$P(\text{Presença_Penicilina} \cap \text{Presença_Estreptomicina}) = \\ P(\text{Presença_Penicilina} \cap \text{Presença_Estreptomicina} / H_i) + \\ P(\text{Presença_Penicilina} \cap \text{Presença_Estreptomicina} / \neg H_i)$$

$$P(\text{Presença_Penicilina} \cap \text{Presença_Estreptomicina}) = (0.76 \times 0.5 \times 0.2) + (0.2 \times 0.5 \times 0.2) + \\ (0.1 \times 0.5 \times 0.2) + (0.1 \times 0.01 \times 0.2) + (0.5 \times 0.5 \times 0.2) = 0.1562 = \mathbf{15.62\%}$$

Com este resultado é possível agora obter as probabilidades de cada hipótese H_i :

- ✓ $P(\text{Choque Anafilático} / \text{Presença_Penicilina} \cap \text{Presença_Estreptomicina}) = (0.76 \times 0.5 \times 0.2) / 0.1562 = 0.4865 = \mathbf{48.65\%}$
- ✓ $P(\text{Broncoespasmo} / \text{Presença_Penicilina} \cap \text{Presença_Estreptomicina}) = (0.2 \times 0.5 \times 0.2) / 0.1562 = 0.128 = \mathbf{12.8\%}$
- ✓ $P(\text{Edema de Laringe} / \text{Presença_Penicilina} \cap \text{Presença_Estreptomicina}) = (0.1 \times 0.5 \times 0.2) / 0.1562 = 0.064 = \mathbf{6.4\%}$
- ✓ $P(\text{Dermatite de Conato} / \text{Presença_Penicilina} \cap \text{Presença_Estreptomicina}) = (0.1 \times 0.01 \times 0.2) / 0.1562 = 0.00128 = \mathbf{0,12\%}$
- ✓ $P(\text{Nenhum} / \text{Presença_Penicilina} \cap \text{Presença_Estreptomicina}) = (0.5 \times 0.5 \times 0.2) / 0.1562 = 0.3201 = \mathbf{32.01\%}$.

Os cálculos anteriormente realizados são suficientemente necessários para uma reflexão sobre o funcionamento da aplicação do Teorema de Bayes à resolução de diagnósticos por meio da Rede Bayesiana. É importante saber que o Teorema de Bayes por si só não considera a existência de eventos difusos, pois para isso é que esta investigação propôs uma adequação da Fórmula caracterizando o hibridismo mencionado.

Em seguida serão apresentados alguns resultados obtidos com simulações executadas no protótipo, tendo em vista a entrada de eventos difusos no sistema. Será exibida a resolução de alguns cálculos objetivando o entendimento da função qualificadora. Em cada exemplo

apresentado serão mostradas as probabilidades das hipóteses diagnósticas sem o processo de qualificação e com o processo de qualificação realizado.

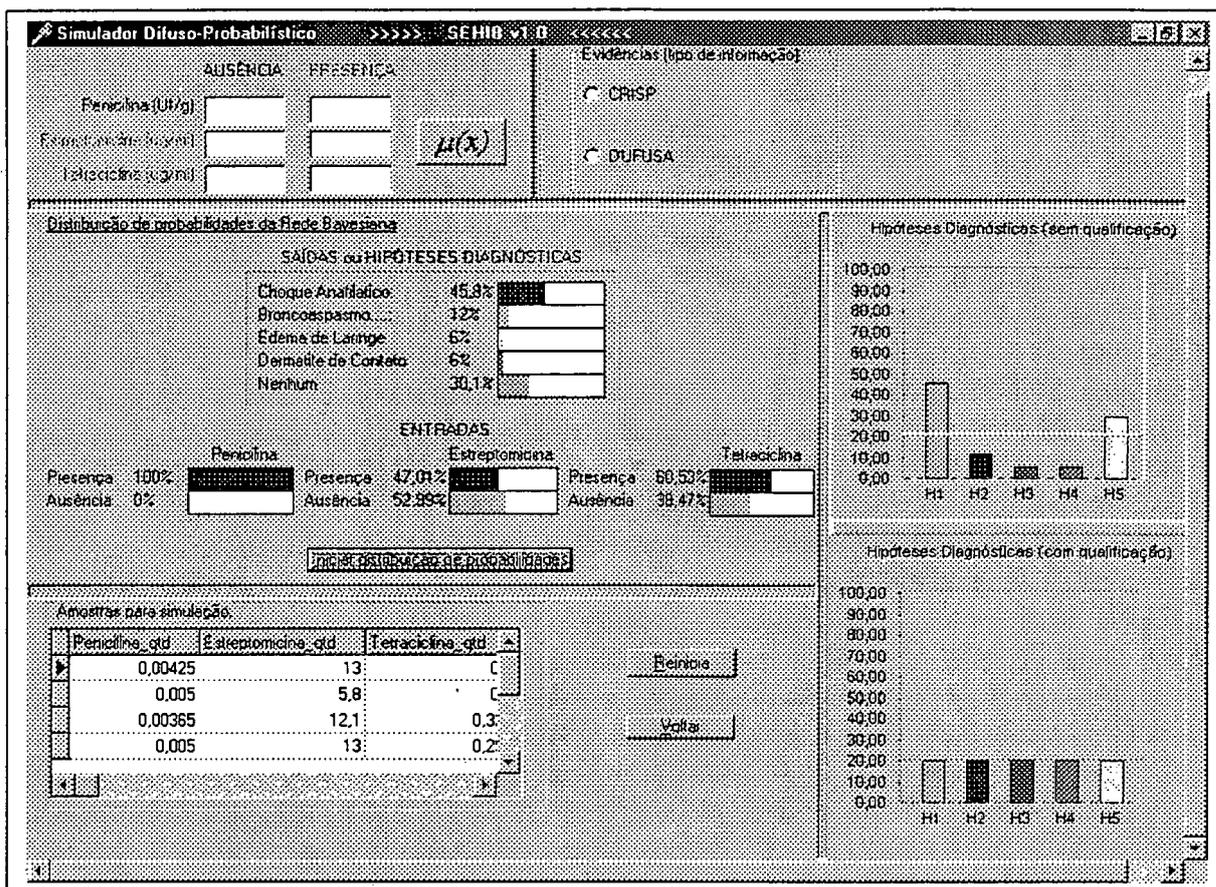


Figura 27: Probabilidades condicionadas pela Presença de Penicilina (sem qualificação).

A fig. 27 exibe os resultados das hipóteses diagnósticas condicionadas à existência de Penicilina, ou seja, $P(H_i/Presença_Penicilina)$. Os resultados são obtidos com a mesma aplicação do caso da fig. 26. A figura seguinte irá exibir os resultados considerando uma entrada para a variável penicilina que a torna um evento difuso para a Rede Bayesiana.

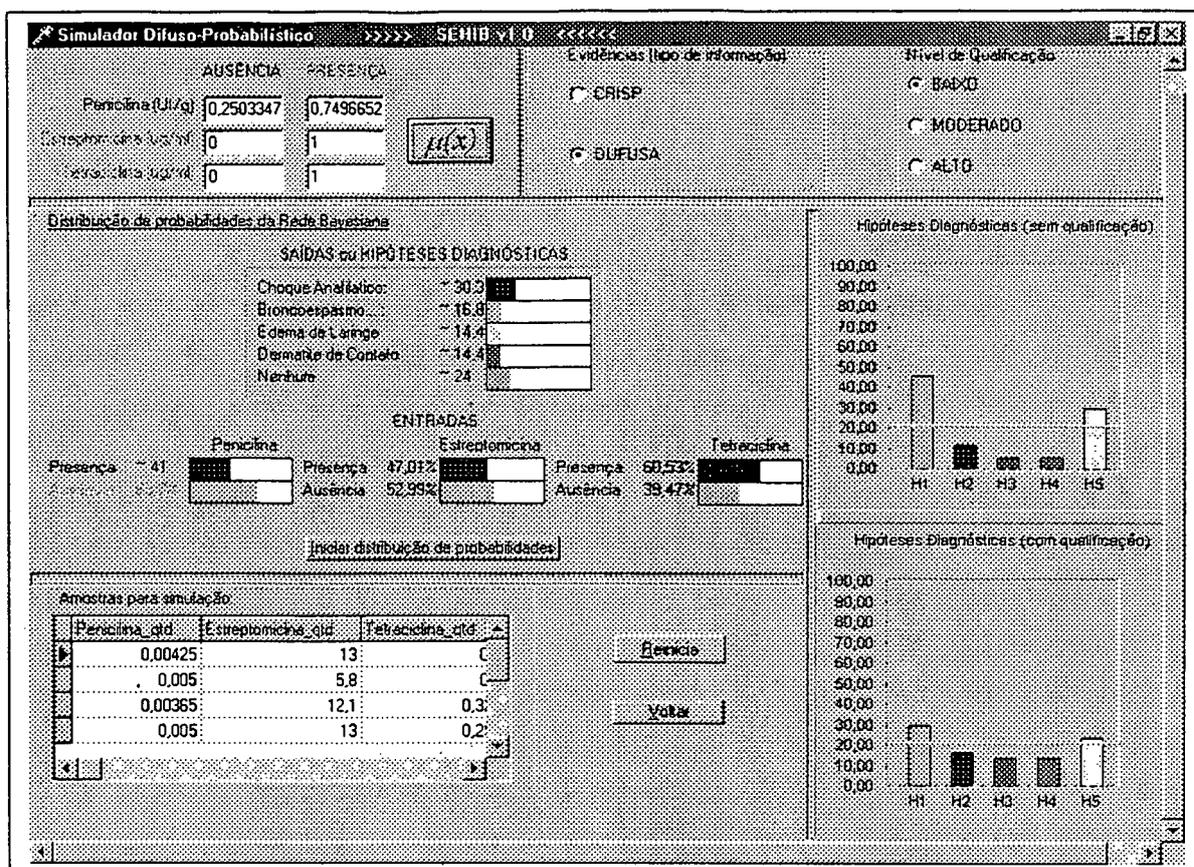


Figura 28: Probabilidades condicionadas pela Presença Difusa de Penicilina (com qualificação).

Para a variável Penicilina foi considerada uma entrada de 0.00425 (em UI/g), fazendo com que sua pertinência aos conjuntos difusos Presença e Ausência seja respectivamente 0.75 e 0.25. Observe que o conjunto Presença é o mais influente neste exemplo e, é por este motivo que deve-se comparar este caso com o caso da fig. 27, sendo que esta apresenta a Presença não difusa de Penicilina. Se o conjunto Ausência tivesse o maior valor de pertinência desta entrada, então a comparação teria que ser com os resultados condicionados à Ausência de Penicilina.

O resultado de $P(H_i/\delta)$, lembrando que δ representa um evento ou conjunto de eventos difusos, foi calculado com base nos valores da tabela 7 aplicando-se a fórmula 16 da seguinte forma:

$$\begin{aligned}
 P(\text{Presença_difusa_Penicilina}) &\approx (0.2) \times [(0.76 \times 0.75) + (0.24 \times 0.25)] + (0.2) \times [(0.2 \times 0.75) \\
 &+ (0.8 \times 0.25)] + (0.2) \times [(0.1 \times 0.75) + (0.1 \times 0.25)] + (0.2) \times [(0.1 \times 0.75) + (0.1 \times 0.25)] + \\
 &(0.2) \times [(0.5 \times 0.75) + (0.5 \times 0.25)] = 0.416 = 41.6\%
 \end{aligned}$$

Então, pode-se agora obter às probabilidades do vetor de hipóteses diagnósticas, aplicando-se fórmula 14:

- ✓ $P(\text{Choque Anafilático/Presença_difusa_Penicilina}) = \{(0.2) \times [(0.76 \times 0.75) + (0.24 \times 0.25)]\} / 0.416 = 0.303 = \mathbf{30.3\%}$
- ✓ $P(\text{Broncoespasmo/Presença_difusa_Penicilina}) = \{(0.2) \times [(0.2 \times 0.75) + (0.8 \times 0.25)]\} / 0.416 = 0.168 = \mathbf{16.8\%}$
- ✓ $P(\text{Edema de Laringe/Presença_difusa_Penicilina}) = \{(0.2) \times [(0.1 \times 0.75) + (0.9 \times 0.25)]\} / 0.416 = 0.144 = \mathbf{14.4\%}$
- ✓ $P(\text{Dermatite de Contato/Presença_difusa_Penicilina}) = \{(0.2) \times [(0.1 \times 0.75) + (0.9 \times 0.25)]\} / 0.416 = 0.144 = \mathbf{14.4\%}$
- ✓ $P(\text{Nenhum/Presença_difusa_Penicilina}) = \{(0.2) \times [(0.5 \times 0.75) + (0.5 \times 0.25)]\} / 0.416 = 0.24 = \mathbf{24\%}$.

Observe na fig. 28 que aparece o nível de qualificação classificado como Baixo, pois, houve a entrada de apenas um evento difuso na Rede. Também, note que houve um achatamento na distribuição das probabilidades, comparando-se os gráficos de barra superior e inferior em relação ao valor 20% do eixo vertical. O ponto 20% ou 0.2 representa o ponto de máxima entropia existente nas informações dadas pelas hipóteses diagnósticas.

Outro caso que será ilustrado causará uma qualificação que pode-se classificar como Alta qualificação, pois às hipóteses estarão condicionadas à 3 (três) eventos difusos. Como no exemplo anterior, mostrar-se-á em primeiro lugar os resultados obtidos sem a qualificação (fig. 29) para que depois possa ser feito uma comparação com os resultados da qualificação (fig. 30).

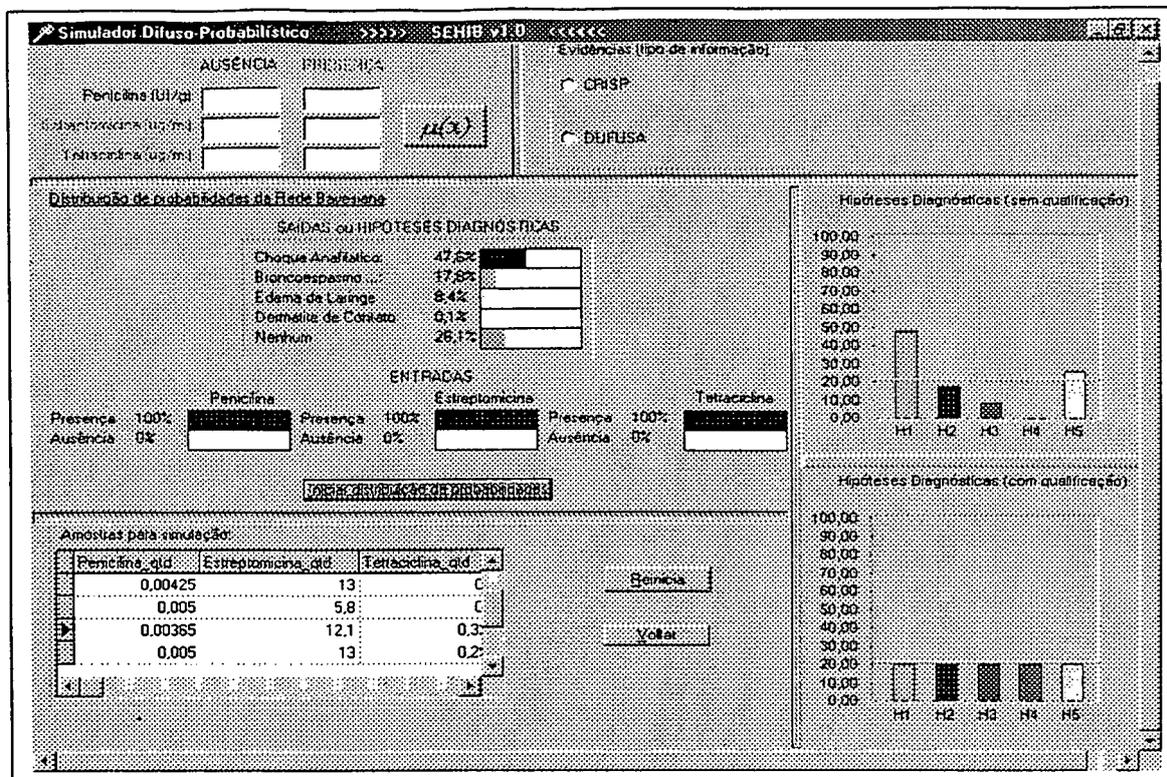


Figura 29: Probabilidades condicionadas pela Presença de todas as variáveis de entrada (sem qualificação).

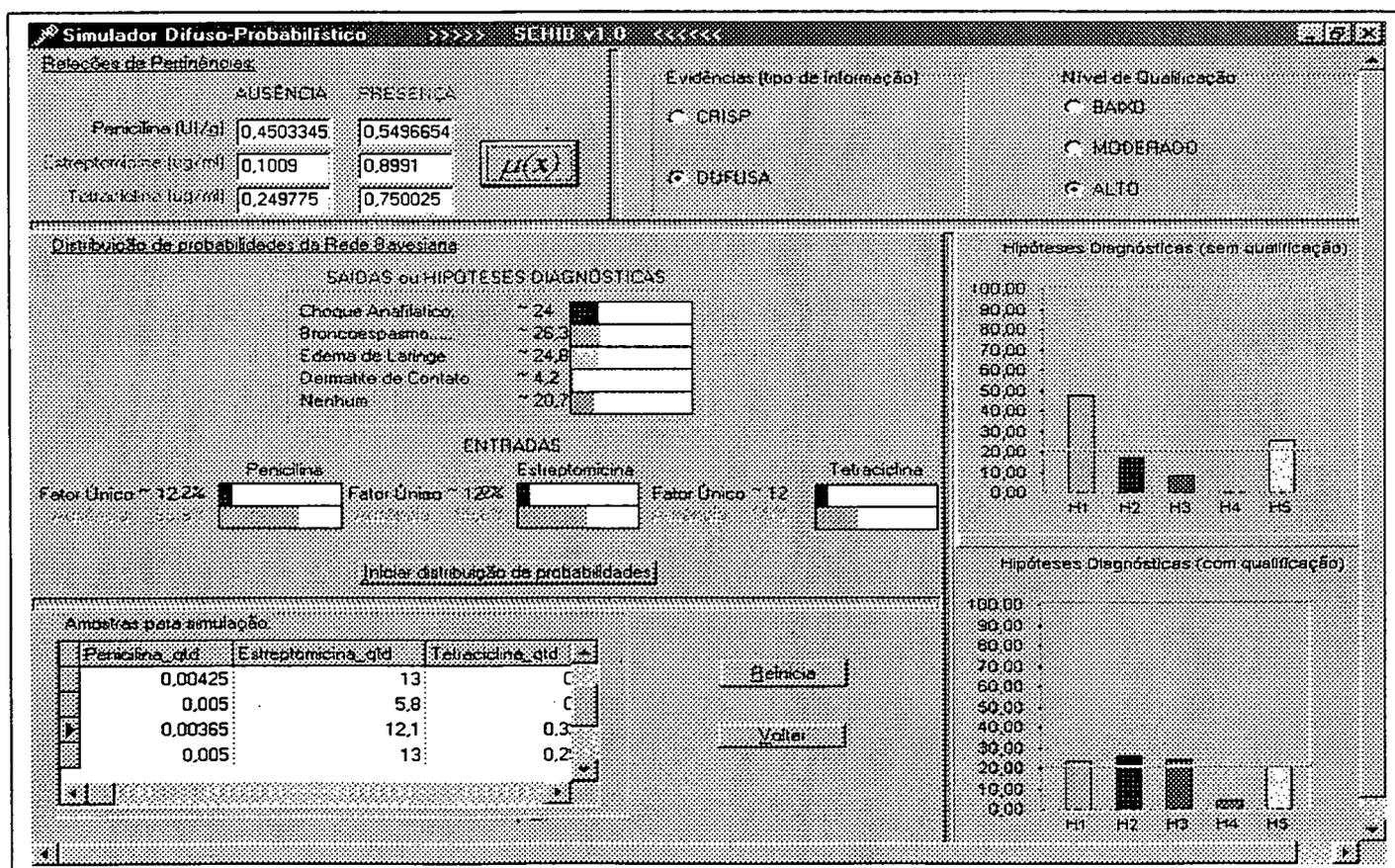


Figura 30: Probabilidades condicionadas pela Presença de todas as variáveis de entrada (sem qualificação).

As entradas para às variáveis Penicilina, Estreptomicina e Tetraciclina foram respectivamente 0.00365 (UI/g), 12.1 (ug/ml) e 0.325 (ug/ml) que, às tornaram todas difusas com às pertinências de 0.55, 0.9 e 0.75, tendo a maior influência do conjunto Presença. Desta forma, os resultados da fig. 30 foram obtidos assim:

Penicilina: Probabilidades condicionadas à cada H_i * às respectivas pertinências para Presença e Ausência

$$\begin{aligned}x_1 &= (0.76 * 0.55) + (0.24 * 0.45) \\x_2 &= (0.2 * 0.55) + (0.8 * 0.45) \\x_3 &= (0.1 * 0.55) + (0.9 * 0.45) \\x_4 &= (0.1 * 0.55) + (0.9 * 0.45) \\x_5 &= (0.5 * 0.55) + (0.9 * 0.45)\end{aligned}$$

Estreptomicina: Probabilidades condicionadas à cada H_i * às respectivas pertinências para Presença e Ausência

$$\begin{aligned}y_1 &= (0.5 * 0.9) + (0.5 * 0.1) \\y_2 &= (0.5 * 0.9) + (0.5 * 0.1) \\y_3 &= (0.5 * 0.9) + (0.5 * 0.1) \\y_4 &= (0.01 * 0.9) + (0.5 * 0.1) \\y_5 &= (0.5 * 0.9) + (0.5 * 0.1)\end{aligned}$$

Tetraciclina: Probabilidades condicionadas à cada H_i * às respectivas pertinências para Presença e Ausência

$$\begin{aligned}z_1 &= (0.6 * 0.75) + (0.4 * 0.25) \\z_2 &= (0.85 * 0.75) + (0.15 * 0.25) \\z_3 &= (0.8 * 0.75) + (0.2 * 0.25) \\z_4 &= (0.5 * 0.75) + (0.5 * 0.25) \\z_5 &= (0.5 * 0.75) + (0.5 * 0.25)\end{aligned}$$

$$w = (x_1 * y_1 * z_1 * 0.2) + (x_2 * y_2 * z_2 * 0.2) + (x_3 * y_3 * z_3 * 0.2) + (x_4 * y_4 * z_4 * 0.2) + (x_5 * y_5 * z_5 * 0.2)$$

$$w = 0.12 = 12\%.$$

Agora, w através da aplicação da fórmula 7 representa $P(H_i/\delta)$, onde δ é o conjunto dos eventos difusos da rede, então é possível obter as probabilidades das hipóteses diagnósticas mostradas na fig. 30.

- ✓ $P(\text{Choque Anafilático} / \delta) = (x_1 * y_1 * z_1 * 0.2) / \delta = 0.24 = 24\%$
- ✓ $P(\text{Broncoespasmo} / \delta) = (x_2 * y_2 * z_2 * 0.2) / \delta = 0.263 = 26.3\%$
- ✓ $P(\text{Edema de Laringe} / \delta) = (x_3 * y_3 * z_3 * 0.2) / \delta = 0.248 = 24.8\%$
- ✓ $P(\text{Dermatite de Contato} / \delta) = (x_4 * y_4 * z_4 * 0.2) / \delta = 0.042 = 4.2\%$
- ✓ $P(\text{Nenhum} / \delta) = (x_5 * y_5 * z_5 * 0.2) / \delta = 0.207 = 20.7\%$.

Observe novamente o achatamento causado na distribuição de probabilidades das saídas (hipóteses diagnósticas) comparando os gráficos das figuras 29 e 30.

As figuras seguintes ilustram outros casos simulados apresentando respectivamente as situações sem qualificação e com o processo de qualificação das saídas da Rede Bayesiana.

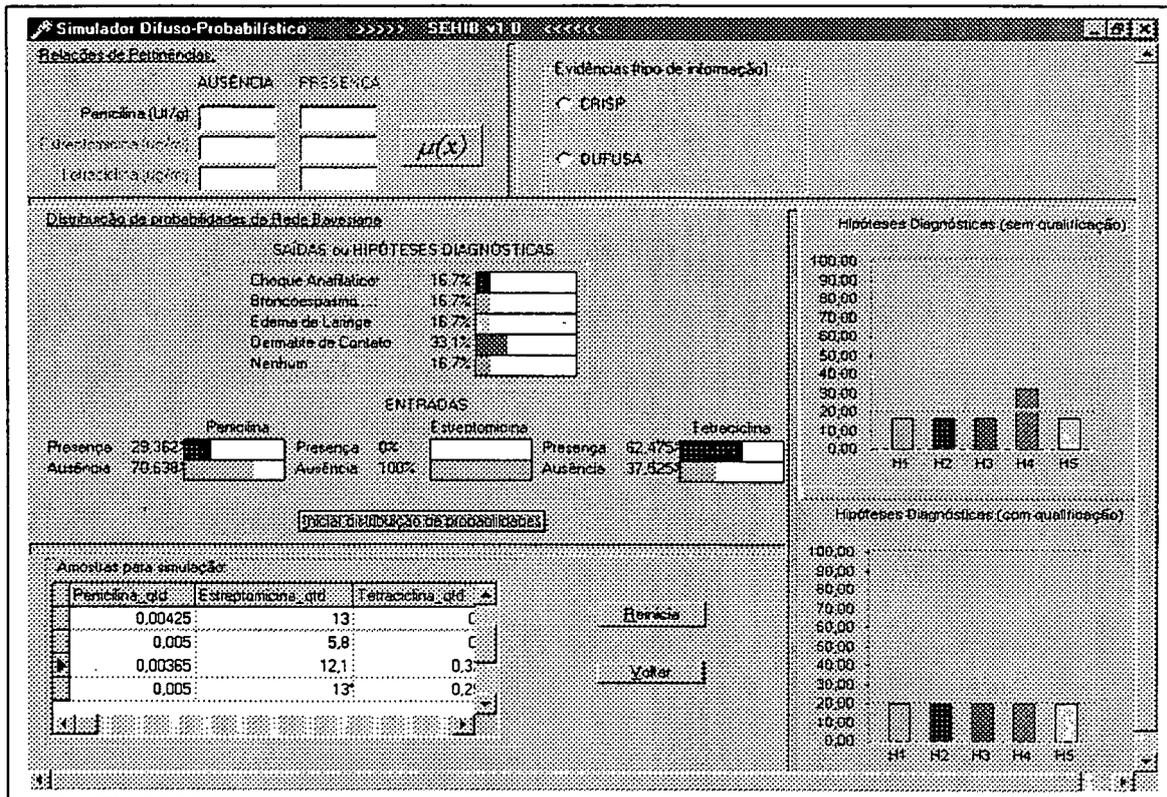


Figura 31: Probabilidades condicionadas pela Ausência de Estreptomomicina (sem qualificação).

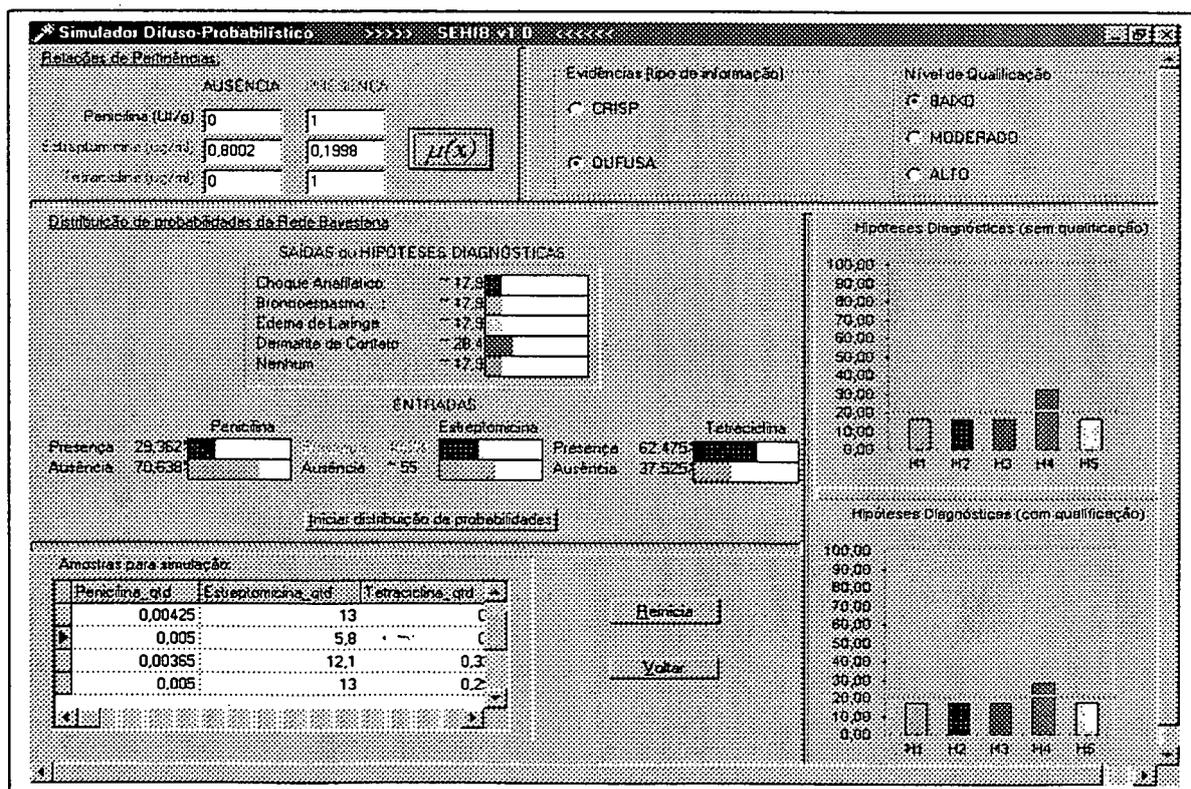


Figura 32: Probabilidades condicionadas pela Ausência de Estreptomomicina (com qualificação).

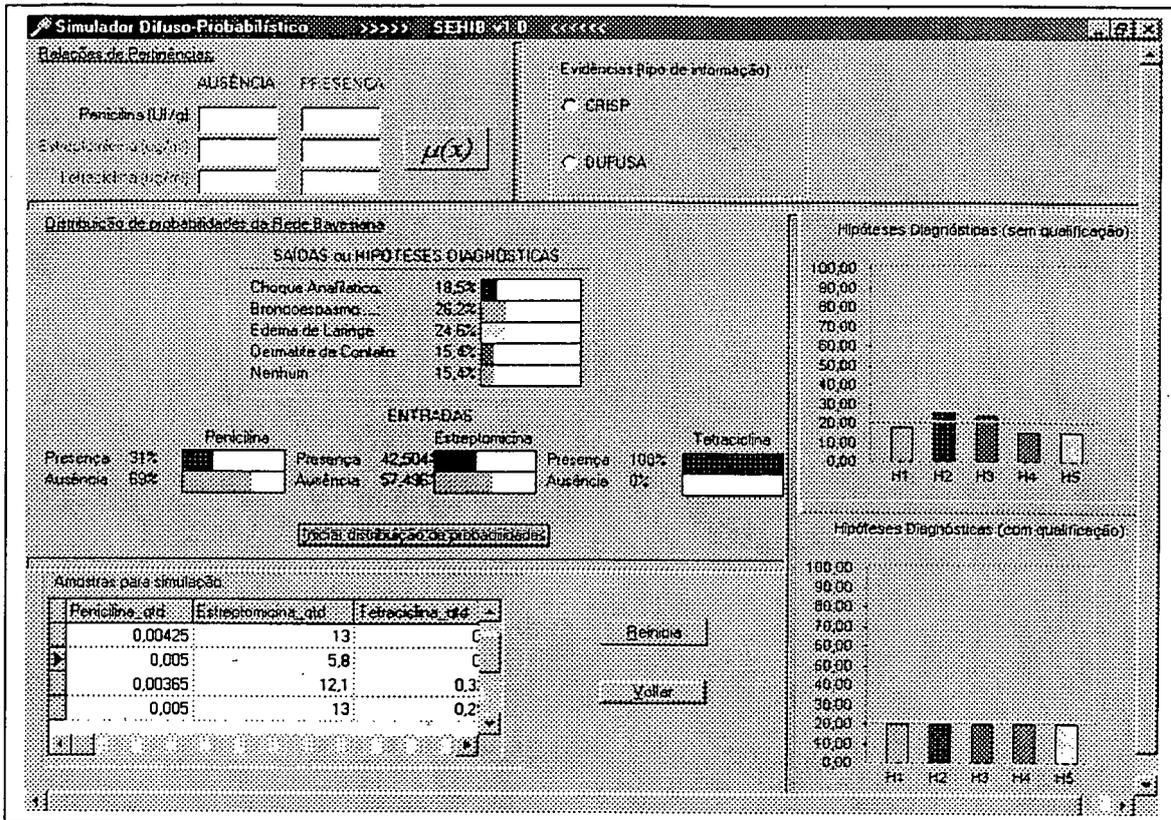


Figura 33: Probabilidades condicionadas pela Presença de Tetraciclina (sem qualificação).

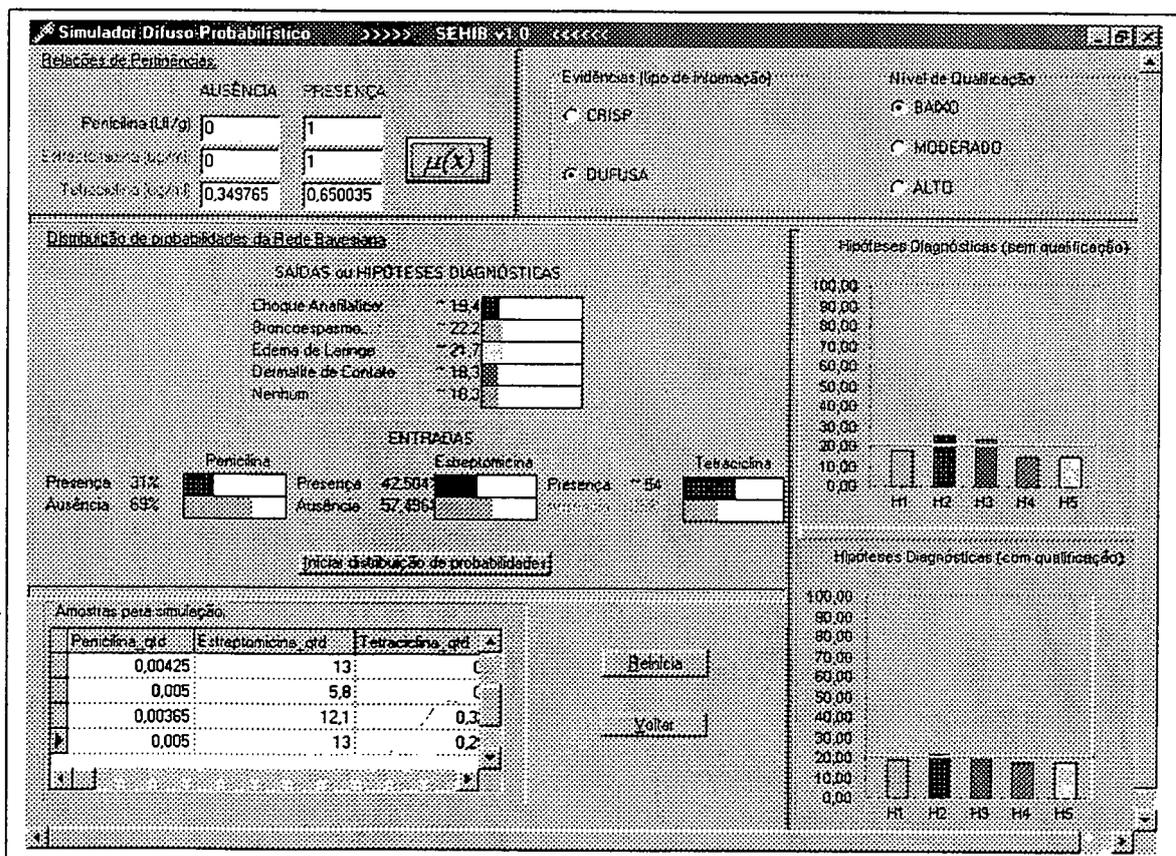


Figura 34: Probabilidades condicionadas pela Presença de Tetraciclina (com qualificação).

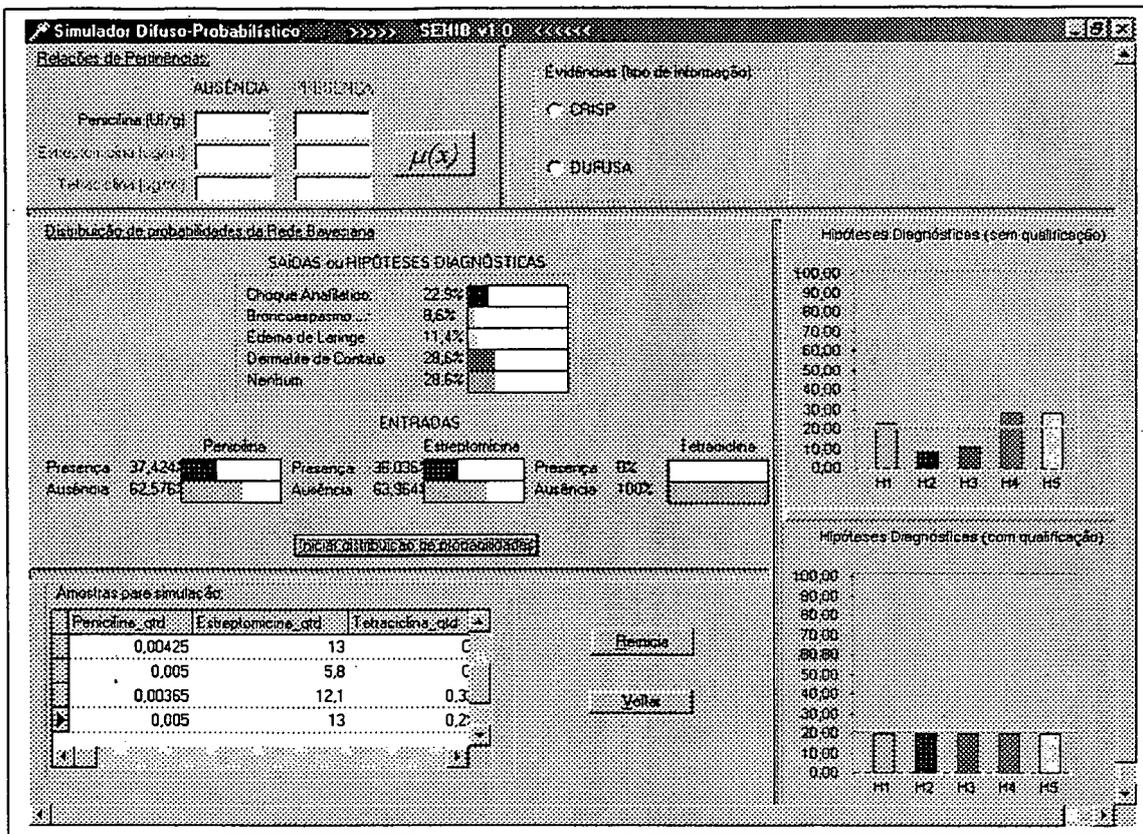


Figura 35: Probabilidades condicionadas pela Ausência de Tetraciclina (sem qualificação).

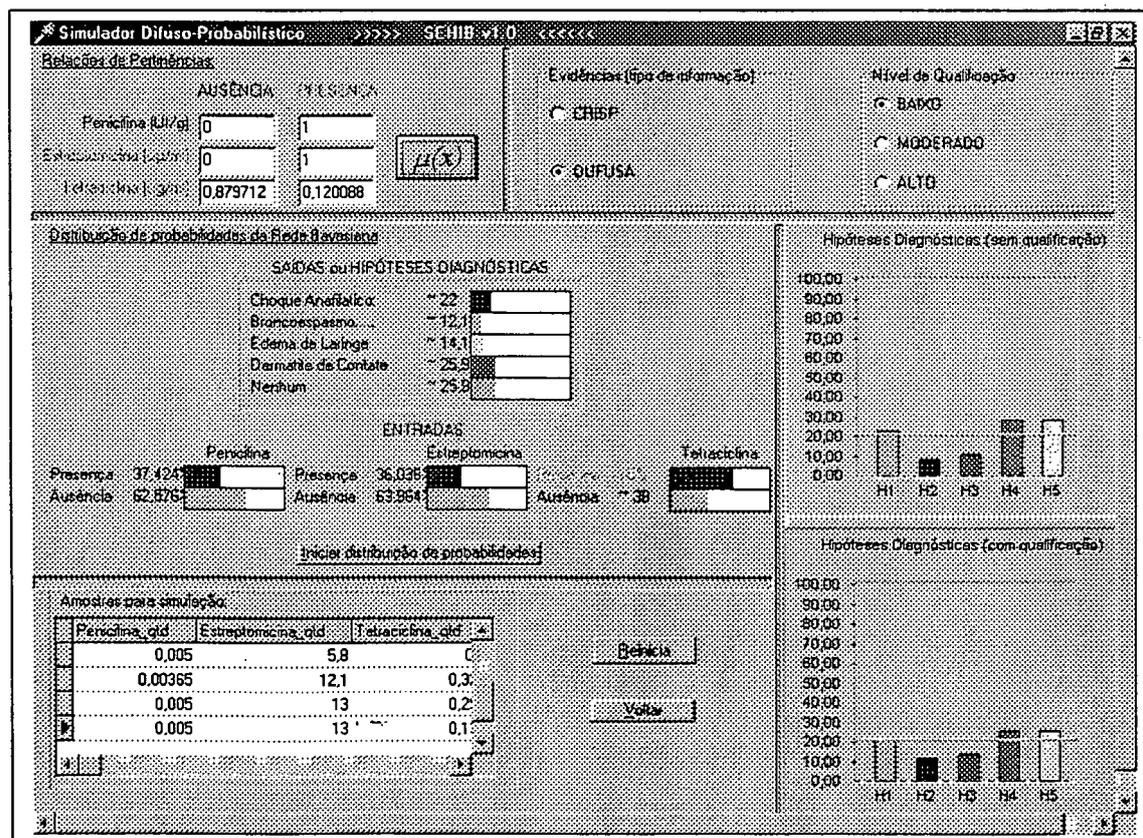


Figura 36: Probabilidades condicionadas pela Ausência de Tetraciclina (com qualificação).

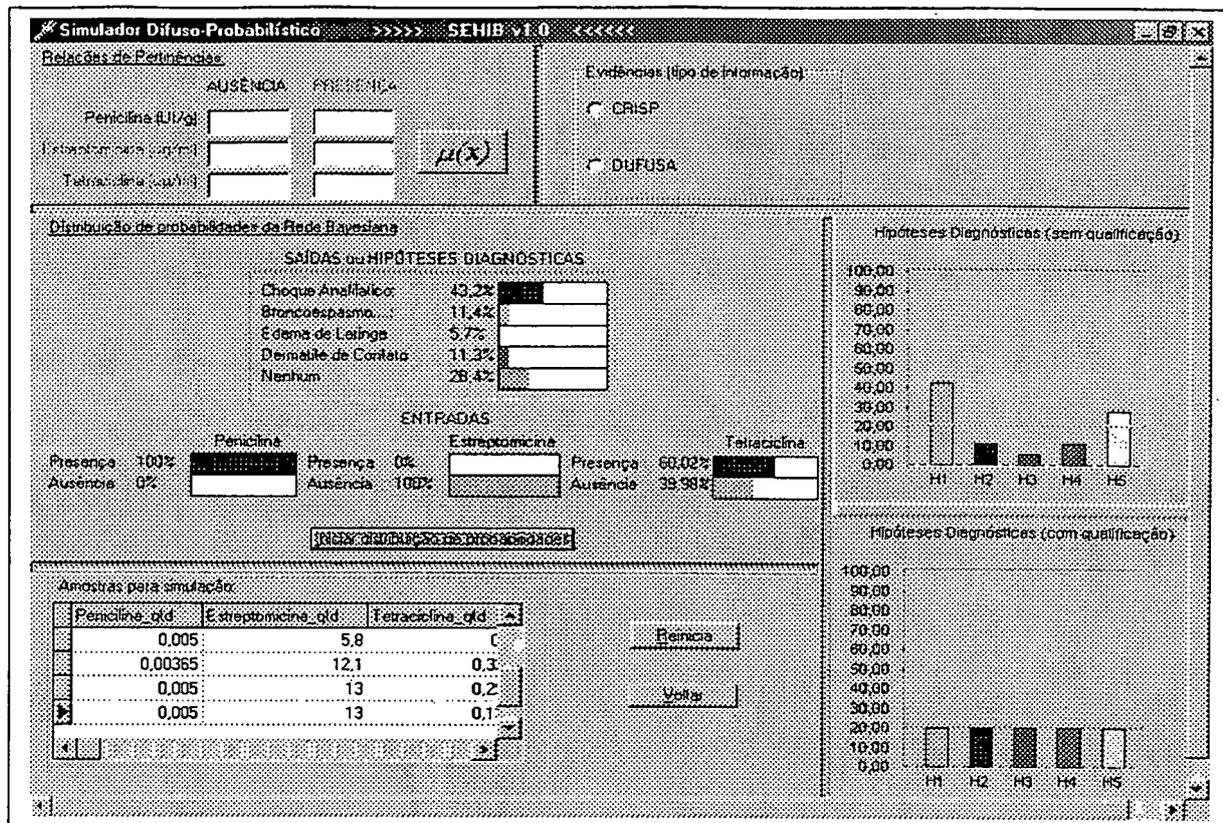


Figura 37: Probabilidades condicionadas pela Presença de Penicilina e Ausência de Estreptomina (sem qualificação).

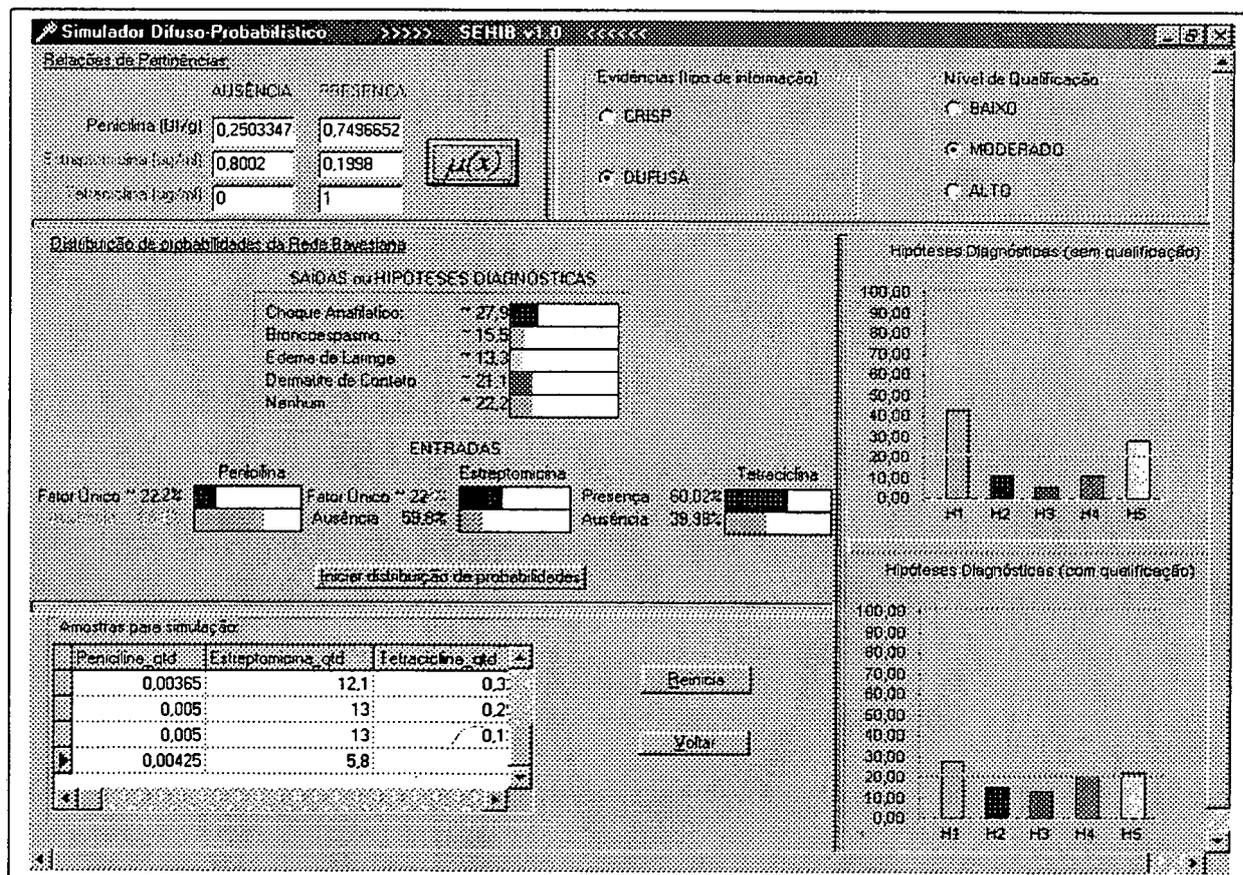


Figura 38: Probabilidades condicionadas pela Presença de Penicilina e Ausência de Estreptomina (sem qualificação).

10. CONCLUSÕES

Os resultados obtidos com base nas simulações mostraram-se válidos no sentido da *qualificação* proposta para que se alcançasse a idéia do hibridismo aplicado às saídas da Rede Bayesiana.

O estudo de caso que comprova a existência de reações adversas em pessoas que consomem leite contaminado por resíduos de antibióticos, mostrou-se adequado quanto a escolha do domínio de aplicação para a validação do modelo proposto.

O modelo de Rede adotado para a representação do conhecimento do especialista, possibilitou o processo de diagnose e orientou os procedimentos para a realização da aplicação do hibridismo.

De acordo com o módulo difuso implementado, foi possível perceber a influência causada nas saídas da Rede Bayesiana, por parte do fator numérico integrante da função de qualificação obtido com base nas pertinências calculadas pela função linear utilizada. Em testes realizados, porém, não implementados nesta aplicação, funções não-lineares apresentaram graus de pertinência que demonstram maior flexibilidade na transição de um estado difuso para outro, sendo assim, mais próximos aos valores reais obtidos da análise de um especialista.

O hibridismo aplicado à Rede Bayesiana qualificou as saídas das hipóteses diagnósticas e caracterizou dois importantes fatos, sendo um deles, a permanência da soma unitária das probabilidades contidas no vetor das hipóteses da Rede, mantendo assim, uma propriedade fundamental da Teoria da Probabilidade. Outro fato notável é o que se pode chamar de um *espalhamento* ou um *achatamento* na distribuição das probabilidades deste vetor. A argumentação para este espalhamento está baseada na observação de que, quando as probabilidades diagnósticas estão condicionadas a eventos difusos, aquelas com percentuais acima do ponto de máxima entropia, diminuem, enquanto as demais aumentam. Desta forma, o processo de qualificação provoca uma perturbação na distribuição das probabilidades diagnósticas e, o referencial para esta mudança de informação é justamente o ponto em que está situada a máxima incerteza no conhecimento da Rede.

Sugere-se para futuras pesquisas relacionadas ao tema aqui apresentado, uma investigação quanto a possibilidade de aplicar o hibridismo sob a esfera de vários eventos difusos, bem como, da existência de variáveis de entrada da Rede compostas por vários estados. É interessante para contribuição às técnicas utilizadas em Sistemas Especialistas que sejam desenvolvidas ferramentas que incorporem a idéia do hibridismo, permitindo a construção de estruturas de conhecimento baseadas no aspecto da imprecisão presente na aleatoriedade.

11. REFERÊNCIAS BIBLIOGRÁFICAS

1. (ARAÚJO, 2000) ARAÚJO, Fábio Duarte de. A Casa Numérica e a Lógica Fuzzy. [On-line]. Available: www.unicap.com.br [2000, Abr. 10].
2. (ARBEX, 1994) ARBEX, Roberto Taiar. Controle Fuzzy: Circuito e aplicações. Revista Instec, junho, 1994.
3. (BAYDA, 2000) BAYDA 1.31 – Bayesian Predictive Discriminant Analysis. [On-line]. Available: <http://www.cs.helsinki.fi/research/cosco> [2000, Mar. 17].
4. (BAYES, 2000) BayesBuilder – A tool for constructing and testing Bayesian networks. [On-line]. Available: <http://www.mbfys.kun.nl/snn/research/bayesbuilder> [2000, Mar. 17].
5. (BAYESIAN, 1999) Bayesian Network Software Packages. [On-line]. Available: <http://nup.es.berkeley.edu/~murphyk/Bayes/bnsoft.html> [1999, Dez. 03].
6. (BELIEF, 2000) Belief Network (BN) PowerConstructor. [On-line]. Available: <http://www.cs.ualberta.ca/~jchenj/bnpc.htm> [2000, Mar. 18].
7. (BELLMAN, 78) BELLMAN. Artificial Intelligence. 1978.
8. (BRACARENSE, 99) BRACARENSE, Paulo Afonso. Um enfoque Segundo a Lógica Difusa para a Meta-Análise. UFSC – Florianópolis, 1999.
9. (BROOKSHEAR, 2000) BROOKSHEAR, J. Glenn. Ciência da Computação: Uma Visão Abrangente. 5ª Ed. Porto Alegre: Bookman, 2000.
10. (CASTRO, 1994) CASTRO, Alex de, Spielman, Jason, Dumas, John. Fuzzy Design Generator (FUDGE) Version 1.02. Copyright Motorola Inc, 1994.
11. (CHARNIAK, 1985) CHARNIAK. Introducion to Artificial Intelligence, 1985.
12. (CRIPPA, 2000) CRIPPA, Mauricio. Sistemas Especialistas. [On-line]. Available: <http://www.n27.udesc.br/demo/trabalhos/alunos/mc/se.htm>. [2000, Jan. 7].
13. (DIAG, 1999) Diagnóstico de efeitos colaterais ocorridos pelo consumo de leite contaminado por resíduos antibióticos. Trabalho para a Disciplina de Tratamentos de Incerteza I – Sistemas Especialistas Probabilísticos. Profª Silvia Modesto Nassar, Dra. Curso de Pós-Graduação em Ciência da Computação – UFSC, 1999.
14. (FEIGENBAUM, 1981) FEIGENBAUM, E., BARR, A. The Handbook of Artificial Intelligence. Vol 1. William Kaufmann, Inc. 1981.
15. (FLDE, 2000) FLDE – Fuzzy Logic Development Environment. [On-line]. Available: www.flde/fldeprd.htm [2000, Abr. 18].
16. (FUZZY, 1999) Fuzzy Systems Implementations. Computational Intelligence PC Tools. 275p. Disciplina de Computação Evolucionária. Profº Paulo Sérgio da

Silva Borges, Dr. Curso de Pós-Graduação em Ciência da Computação – UFSC, 1999.

17. (FUZZY, 2000) Fuzzy Logic – Tools. [On-line]. Available: <http://www.aaai.org/Resources/Education/Repository-Mirror/fuzzy-tools.html> [2000, Jan. 12].
18. (FUZZY, 2000) Fuzzy Systems Engineering. [On-line]. Available: www.fuzzysys.com/index.htm [2000, Abr. 02].
19. (FUZZYTECH, 1999) FUZZYTECH - INFORM GmbH, Gesch., ftsbereich Fuzzy-Technologien, Pascalstr. 23, D-5100 Aachen. [On-line]. Available: http://www.fuzzytech.com/e_apnote.htm [1999, Jun. 05].
20. (GÓMEZ, 2000) GÓMEZ, Luis Alberto, ALVES, João Bosco. Controladores Multiagentes em Edifícios Inteligentes. [On-line]. Available: www.inf.ufsc.br/~jbosco/currr.html [2000, Abr. 14].
21. (JUNIOR, 1997) JUNIOR, Estevam Rafael Hrusehka. Propagação de Evidências em redes Bayesianas – Diagnóstico sobre doenças pulmonares. Dissertação de Mestrado. Universidade de Brasília – Depto. de Ciência da Computação – Brasília, 1997.
22. (KAND, 1986) Kandel, Abraham. Fuzzy Mathematical Techniques with Applications. Florida: Addison-Wesley Publishing Company, 1986.
23. (KLIR, 1995) KLIR, George J., YUAN, Bo. Fuzzy sets and fuzzy logic: theory and applications. New Jersey: Prentice Hall, 1995.
24. (KNIGHT, 1993) KNIGHT, Kevin. Artificial Intelligence. Traduction. São Paulo: Makron Books, 1993.
25. (LEAL, 2000) LEAL, Cassiano Noiman. OBESIDADE – Por que livrar-se deste mal? [On-line]. Available: www.ez.poa.com.br [2000, Abr. 01].
26. (LEVINE, 1988) LEVINE, Robert et al. Inteligência Artificial e Sistemas Especialistas. São Paulo, McGraw-Hill, 1988.
27. (LEWIS, 2000) LEWIS, Harry R., PAPADIMITRIOU, Christos H. Elementos de Teoria da Computação. 2ª Ed. Porto Alegre. Bookman, 2000.
28. (LIMA, 1999) LIMA, Cynthia Moreira. Introdução a Inteligência Artificial. [On-line]. Available: www.elo.com.br/~cynthia/ia.html [1999, Dez. 20].
29. (MACHADO, 1999) MACHADO, Nivaldo. Um andróide em minha vida: Estudo do Argumento do Quarto-Chinês: Um contributo para a análise da problemática da possibilidade ou impossibilidade da tentativa de redução de estados mentais à coisas. Monografia apresentada ao Curso de Pós-Graduação em Psicopedagogia. Unidavi, 1999.
30. (NASSAR, 1998) NASSAR, S. M. A estatística como Apoio à Inteligência Artificial: Sistemas Especialistas Probabilísticos. In: Estatística e Informática: um processo

interativo entre duas ciências. Trabalho apresentado no Concurso para Professor Titular, INE, CTC, UFSC, 1998.

31. (NETICA, 1999) Norsys Software Corp. [On-line]. Available: www.norsys.com [1999, Dez. 02].
32. (NORSYS, 2000) Norsys Software Corp. – Netica em Data Mining. [On-line]. Available; www.norsys.com [2000, Abr. 01].
33. (NOTES, 1999) Notes on Diagnosis Reasoning. Bayesian Systems. Applications for the Creation of Computer-Assisted Tools for Diagnosis. Produced for the Development of Minerva. [On-line]. Available: <http://www.cid.ch/doc/baye.html> [1999, Nov. 15].
34. (OLIVEIRA, 1996) OLIVEIRA, Gladis T. Borges, NISHIDA, Waleska. Sistemas Baseados em Regras Fuzzy. 1996.
35. (PEARL, 1988) PEARL, J. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. California: Morgan Kaufmann, 1988.
36. (PEREIRA, 2000) PEREIRA, Vanderlei de AMO, DOTTO, Fábio Romano Lofrano. Fuzzy Logic Page. [On-line]. Available: <http://www.geocities.com/Athens/crete/2647/indexp.html> [2000, Fev. 10].
37. (RICH, 1993) RICH, Eliane, KNIGHT, Kevin. Inteligência Artificial. 2ª Ed. São Paulo: Makron Books, 1993.
38. (RODDER, 1997) RODDER, Wilhelm, KULMANN, Friedhelm, KOPITTKKE, Bruno Harmut. SPIRIT: Operations Research. FrenUniversitat – Gesamthochschule in Hagen, 1997.
39. (ROSS, 1995) ROSS, Timothy J. Fuzzy Logic With Engineering Applications. Universidade do México: McGraw-Hill, Inc, 1995.
40. (SAFFIOTI, 1997) SAFFIOTI, Alessandro. Using fuzzy logic for autonomous robotics: an on-line workshop. The Knowledge engineering review, Cambridge, v.12, n.1, p.91-94, 1997.
41. (SCHALKOFF, 1990) SCHALKOFF, R. Artificial Intelligence: na Engineering Approach. São Paulo, McGraw-Hill, 1990.
42. (SEM, 1999) Sistemas Especialistas na Medicina. [On-line]. Available: http://www.din.uem.br/~ia/medicina/se_medicina.html [1999, Dez. 02].
43. (SIS, 1999) Sistemas Especialistas. [On-line]. Available: <http://www.din.uem.br/~ia/especialistas> [1999, Dez. 02].
44. (SISTEMAS, 2000) Sistemas Especialistas Baseados em Conhecimento. [On-line]. Available: <http://www.penta.ufgs.br/gr952/tab1/geren2.html> [2000, Jan. 5].

45. (SOUZA, 1998) SOUZA, Nelson Grau. Ocorrência de Resíduos de Antibióticos no Leite de Consumo no Estado de Santa Catarina. Universidade Federal de Santa Catarina – UFSC – Florianópolis, 1998.
46. (TAVARES, 1990) TAVARES, S. W. Manual de antibióticos e quimioterápicos antiinfeciosos. ed. Atheneu, p.3-124, 1990.
47. (VIEIRA, 1999) VIEIRA, Rodrigo. Protótipo de um Sistema de Monitoramento Remoto Inteligente. Dissertação de Mestrado. Curso de Pós-Graduação em Engenharia de Produção. UFSC, 1999.
48. (WISTON, 1984) WISTON. Artificial Intelligence. 2nd Edition, 1984.
49. (ZIMMERMANN, 1999) ZIMMERMANN, H. Logic for Mobile Robot Control. Handbook of fuzzy sets and possibility theory. Boston: Klower Academic, 1999.