

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
CURSO DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

**Um Modelo para Sistemas
Especialistas Conexionistas Híbridos**

ELISEO BERNI REÁTEGUI

Dissertação de Mestrado

Prof. Dr. José Mauro Volkmer de Castilho
Orientador

Prof.^a Dra. Beatriz de Faria Leão
Co-orientadora

Porto Alegre, junho de 1993



CIP - CATALOGAÇÃO NA PUBLICAÇÃO

Reátegui, Eliseo Berni

Um Modelo para sistemas especialistas conexionistas híbridos. -
Porto Alegre: CPGCC da UFRGS, 1993.

126p.: il.

- Universidade Federal do Rio Grande do Sul, Curso de Pós-
Graduação em Ciência da Computação, Porto Alegre, 1993.
Orientador: Castilho, José Mauro Volkmer de. Co-orientadora: Leão,
Beatriz de Faria.

Dissertação de Mestrado: Inteligência Artificial, Sistemas
Especialistas, Aprendizado de Máquina, Redes Neurais, Sistemas
Híbridos, Informática aplicada à Medicina.

À minha mãe, pelo carinho a mim sempre dedicado, facilitando minha passagem por difíceis obstáculos com que freqüentemente me deparei.

À Dra. Beatriz de Faria Leão, pelo exemplo humano e profissional em que devo me espelhar por toda a minha vida.

AGRADECIMENTOS

À Dra. Beatriz de Faria Leão, pela competente orientação, incentivo e participação que tornaram possível a realização e conclusão deste trabalho.

Ao Prof. José Mauro Volkmer de Castilho pelo incentivo, ensinamentos e irrestrito apoio.

À amiga Eneida pela amizade e inestimável colaboração durante a coleta de dados de prontuários médicos realizada na etapa de validação deste trabalho.

Ao amigo Alex Guazzelli pelo carinho e incentivo. Também agradeço pela revisão do texto e empréstimo de seu computador que contribuíram imensamente para a finalização do texto.

Ao amigo Altino pela inesquecível amizade, sensibilidade e ajuda, tanto na realização quanto na formatação final desta dissertação.

Aos amigos da Unidade de Pesquisa: Carmem, Lúcia, Rose, Vera e Zezinho pela adorável convivência durante o período em que trabalhei no Instituto de Cardiologia - RS.

Aos funcionários do SAME, Marco Bittencourt, Lorenzo Baddini e Emerson Schmitt pela colaboração na seleção dos prontuários médicos utilizados na etapa de validação desta dissertação de mestrado.

À FAPERGS pelo apoio à pesquisa que tornou possível o desenvolvimento deste trabalho.

LISTA DE ABREVIATURAS

AC	Aquisição de Conhecimento
BC	Base de Conhecimento
BCH	Base de Conhecimento Híbrida
CIA	Comunicação Interatrial
CIV	Comunicação Interventricular
DSAV	Defeito Septal Atrioventricular
EC	Engenheiro de Conhecimento
ESN	Extensional Semantic Network
GF	Grafo de Conhecimento
IA	Inteligência Artificial
ISN	Intensional Semantic Network
MNC	Modelo Neural Combinatório
POO	Programação Orientada a Objetos
RN	Rede Neural
SAD	Sistemas de Apoio à Decisão
SE	Sistema Especialista

LISTA DE FIGURAS

Figura 1.1	Exemplo de representação de conhecimento em <i>frames</i>	23
Figura 1.2	Arquitetura básica dos sistemas especialistas	25
Figura 1.3	Grafo de conhecimento	32
Figura 1.4	Neurônio	34
Figura 1.5	Caracteres ambíguos	35
Figura 1.6	Estrutura geral de um modelo neural	37
Figura 1.7	Topologia do Modelo Neural Combinatório (MNC)	41
Figura 1.8	Modelo <i>E-nebuloso</i> e <i>OU-nebuloso</i>	43
Figura 1.9	Modelos de integração de SEs e RNs.....	47
Figura 3.1	Arquitetura básica de um sistema empregando NEXTOOL	53
Figura 3.2	Rede semântica ISN e ESN	54
Figura 3.3	Arquitetura básica do modelo conexionista híbrido HYCONES.....	59
Figura 3.4	Modelagem de problemas classificatórios através dos conceitos de abstração	61
Figura 3.5	<i>Frames</i> de <i>achado</i>	62
Figura 3.6	<i>Frames</i> de <i>diagnóstico</i>	64
Figura 3.7	Hierarquia de <i>frames</i> <i>achado</i> e <i>diagnóstico</i>	65
Figura 3.8	Mapeamento de um GC para a rede neural e para o <i>frame</i> <i>diagnóstico</i>	66
Figura 3.9	Funcionamento do motor de inferência de HYCONES.....	68
Figura 3.10	Criação das redes neurais por casos de treinamento	72
Figura 3.11	Operador genético de recombinação	80
Figura 3.12	Inserção de grafos de conhecimento.....	83
Figura 4.1	Hierarquia de <i>frames</i> da base de conhecimento híbrida	85

Figura 4.2 Módulos do sistema HYCONES	89
Figura 4.3 Desenvolvimento e operação do sistema HYCONES	90
Figura 4.4 Hierarquia de classes de <i>achados</i> para cardiopatias congênitas	91
Figura 4.5 Selecionando a classe de <i>achados</i>	92
Figura 4.6 Entrada dos fatores de confiança de um <i>achado</i>	93
Figura 4.7 Falta de evidência para concluir um diagnóstico	94
Figura 4.8 Conclusão parcial, com solicitação de informações adicionais.....	95
Figura 4.9 Conclusão diagnóstica por HYCONES	96
Figura 4.10 Mecanismo explanatório de HYCONES	97
Figura 4.11 Inserção do caso diagnosticado na base de casos	98

LISTA DE TABELAS

Tabela 1.1	Características básicas dos sistemas especialistas.....	26
Tabela 1.2	Principais modelos de redes neurais	38
Tabela 1.3	Modelos de redes neurais aplicados a sistemas especialistas	40
Tabela 1.4	Comparação entre os paradigmas de desenvolvimento de SEs	45
Tabela 3.1	Fatores de relevância, conforme definidos no sistema INTERNIST	79
Tabela 5.1	Desempenho das cinco versões da BCH	100
Tabela 5.2	Quadro de contingência comparativo de B1, B2, B4 e B5.....	103
Tabela 5.3	Quadro de contingência comparativo de B1 e B5	104
Tabela 5.4	Quadro de contingência comparativo de B2 e B5	104
Tabela 5.5	Quadro de contingência comparativo de B4 e B5	105
Tabela 5.6	Valores de χ^2 na comparação entre B1 e B5, B2 e B5, B4 e B5	106
Tabela 5.7	Comparação entre os achados essenciais de B1 e B2	106
Tabela 5.8	Comparação entre os achados essenciais de B3 e B4	107
Tabela 5.9	Comparação entre os achados essenciais de B1 e B5	108
Tabela 5.10	Avaliação das funções do protótipo HYCONES	110
Tabela 6.1	Comparação das ferramentas NEXTTOOL e HYCONES	113

SUMÁRIO

LISTA DE ABREVIATURAS	5
LISTA DE FIGURAS	7
LISTA DE TABELAS.....	9
RESUMO	13
ABSTRACT	15
1 INTRODUÇÃO	17
1.1 Introdução aos Sistemas Especialistas	17
1.2 Classificação dos Sistemas Especialistas	17
1.3 Representação de Conhecimento e Arquitetura dos SEs	21
1.4 Aplicações de Sistemas Especialistas	25
1.5 Dificuldades não Equacionadas nos Sistemas Especialistas.....	27
1.6 Engenharia do Conhecimento.....	30
1.7 Introdução à Abordagem Conexcionista.....	33
1.7.1 Modelos conexionistas.....	35
1.8 Sistemas Especialistas Conexionistas	39
1.9 Arquiteturas Conexionistas Híbridas	44
1.9.1 Comparação entre os Paradigmas Simbólico, Conexcionista e Híbrido ..	44
1.9.2 Classificação dos Sistemas de Arquitetura Híbrida.....	47
1.9.3 Aplicações de Sistemas de Arquitetura Híbrida	49
1.9.3.1 Avaliação de Projetos Arquitetônicos.....	50
1.9.3.2 Tratamento do Melanoma.....	50
1.9.3.4 Tratamento de Efluentes Cloacais e Industriais.....	51
1.10 NEXTOOL: Ferramenta para SEs de Arquitetura Híbrida.....	52
2 OBJETIVOS	57

3 METODOLOGIA	59
3.1 Introdução	59
3.2 A Base de Conhecimento Híbrida	59
3.2.1 Integração das Redes Neurais MNC ao Paradigma de <i>Frames</i>	65
3.3 O Mortor de Inferência	67
3.4 O Mecanismo de Explicação	69
3.5 A Máquina de Aprendizado	70
3.5.1 Aprendizado indutivo.....	71
3.5.2 Aprendizado dedutivo.....	79
3.6 Interfaces	82
4 IMPLEMENTAÇÃO	85
4.1 Ambiente	85
4.2 Implementação da Base de Conhecimento Híbrida	85
4.3 Funções do Sistema	88
4.4 Consulta ao Sistema Especialista	90
5 VALIDAÇÃO	99
6 DISCUSSÃO	111
7 CONCLUSÃO	117
BIBLIOGRAFIA	119
APÊNDICE I - Aquisição de Conhecimento por Grafos de Conhecimento	125

RESUMO

O presente trabalho descreve um modelo híbrido para sistemas especialistas classificatórios que integra os mecanismos de representação de conhecimento de *frames* e de redes neurais, assimilando qualidades inerentes de cada um dos dois paradigmas. O mecanismo de *frames* fornece tipos construtores flexíveis para a modelagem do conhecimento do domínio, enquanto as redes neurais possibilitam tanto a automação da aquisição de conhecimento, a partir de uma base de casos, quanto a implementação de aprendizado indutivo e dedutivo.

Empregam-se dois tipos de *frames* para representar o conhecimento relativo aos problemas de classificação: *frames de achado* e *frames de diagnóstico*. Os primeiros descrevem os objetos elementares da aplicação que podem influenciar na detecção de certos diagnósticos. Os *frames de diagnóstico*, por sua vez, são descritos em termos de achados e correspondem às descrições dos problemas de classificação tratados pelo sistema especialista.

É utilizado o Modelo Neural Combinatório (MNC), caracterizado por sua capacidade de aprendizado heurístico através do reconhecimento de regularidades observadas no mundo exterior. A integração do mecanismo de representação de *frames* e do MNC se dá ao nível dos *frames de diagnóstico*, armazenando representações simbólicas do conhecimento das redes neurais que explicitam o conhecimento armazenado nas camadas intermediárias das últimas.

O motor de inferência do sistema opera de maneira integrada sobre ambos os paradigmas, o conexionista e o simbólico. A máquina de aprendizado fornece meios para a automação da aquisição de conhecimento a partir de um banco de casos, bem como para o refinamento do conhecimento do sistema. A máquina de aprendizado é dividida em dois componentes principais: o componente de aprendizado indutivo e o componente de aprendizado dedutivo. O primeiro se volta para o componente indutivo e utiliza os métodos de aprendizado por casos de treinamento, descritos para o MNC, para refinar o conhecimento das redes neurais. O componente dedutivo faz uso de uma estratégia de reorganização das redes neurais, para gerar novos modelos de resolução de problemas.

O sistema implementado, denominado HYCONES (the **HY**brid **CON**nectionist **Expert System**) é apresentado com detalhamento da integração entre os paradigmas simbólico e conexionista. Para validar HYCONES, foram construídas e testadas 5 versões de uma base de conhecimento híbrida para o diagnóstico das 3 cardiopatias congênitas mais freqüentes no Instituto de Cardiologia - RS. A primeira delas (B1) baseou-se no conhecimento extraído de 22 especialistas em cardiologia pediátrica, de 4 instituições brasileiras. Os grafos médios de conhecimento destes especialistas foram diretamente traduzidos para redes neurais MNC, dando origem à primeira versão do sistema. A segunda versão (B2) baseou-se no refinamento da base de conhecimento da versão B1, através da reorganização e treinamento das redes neurais dos especialistas, com aplicação dos métodos de aprendizado dedutivo. A terceira versão (B3) baseou-se no conhecimento de 23 não especialistas, cardiologistas gerais, na sua maioria oriundos do Instituto de Cardiologia - RS. A mesma técnica aplicada para a construção da versão B1 foi utilizada para construir esta versão da base de conhecimento híbrida. A quarta versão (B4) corresponde ao refinamento e reorganização da base de conhecimento da versão B3, com as mesmas técnicas de aprendizado dedutivo, aplicadas na construção da versão B2. Finalmente, na quinta versão (B5), a base de conhecimento foi construída, automaticamente, a partir de base de casos, valendo-se do método de explicitação automática de conhecimento do MNC.

Os resultados obtidos na validação do sistema comprovaram a eficiência dos métodos de explicitação automática de conhecimento, a partir do banco de casos e de aprendizado dedutivo, por reorganização e treinamento das redes neurais.

As principais contribuições deste trabalho foram: a definição do mecanismo para integração de *frames* a redes neurais; a especificação dos métodos de aquisição automática do conhecimento, a partir de um banco de casos e seu posterior treinamento e aprendizado, por métodos indutivos e dedutivos; finalmente, a incorporação das medidas de especificidade e sensibilidade no processo de explicitação automática do conhecimento do banco de casos.

PALAVRAS-CHAVE: Inteligência Artificial, Sistemas Especialistas, Aprendizado de Máquina, Redes Neurais, Sistemas Híbridos, Informática aplicada à Medicina.

ABSTRACT

This dissertation describes a hybrid model for classification expert systems - HYCONES (HYbrid CoNnectionist Expert System), which integrates the knowledge representation mechanism of frames and neural networks, incorporating qualities inherent to each of the knowledge representation paradigms. The mechanism of frames provides flexible constructs for modeling the domain knowledge. The neural networks provide the means for automatic knowledge acquisition and enable the implementation of deductive and inductive learning.

The frames mechanism makes use of semantic modeling concepts to make the new model more generic and to provide flexible constructs for knowledge representation. It employs two basic types of frames to represent the knowledge involved in classification problems: *finding* frames and *diagnosis* frames. The *finding* frames describe the elementary objects of the application world that can influence in the detection of certain diagnoses. *Diagnosis* frames, on the other hand, are described in terms of findings and correspond to the descriptions of the classification problems addressed by the expert system.

The Combinatorial Neural Model (CNM) was the architecture chosen for the neural network environment. This model is characterized by its capacity of learning through the recognition of regularities observed in the outside world. The integration of the knowledge representation mechanism of frames with the CNM happens at the *diagnosis* frames level. The *diagnosis* frames store symbolic representations of knowledge contained in the neural networks.

The system's inference machine operates under an integrated manner with both knowledge representation paradigms. The learning machine provides the means for automatic knowledge acquisition and for knowledge refinement. It is composed by two modules: the inductive and the deductive learning component. The first one uses CNM methods to refine the neural network knowledge. The latter follows a strategy that reorganizes the neural networks, generating new problem solving models.

The implemented HYCONES system is presented, detailing the integration between the symbolic and the connectionist approach. To validate HYCONES, five

versions of a hybrid knowledge base (HKB) for the diagnosis of the three most common congenital heart diseases at the Instituto de Cardiologia - RS were constructed and tested. The first version of the HKB, named B1 for validation purposes, is based on the mean knowledge graphs of 22 experts in congenital heart diseases, from four important Brazilian cardiological institutions. These KG were directly translated to CNM neural networks. The second version of the HKB, named B3 for validation purposes, contains the mean knowledge graphs of 23 non-experts, i.e., general cardiologists, from the Institute of Cardiology of RS. The same method used to built B1 was invoked here: direct translation from the knowledge graphs to CNM neural networks. The second and fourth versions of HKB, named B2 and B4 respectively, consist of trained versions of B1 and B3. To create these trained versions, the CNM neural networks of B1 and B3 were submitted to deductive learning procedures. Finally, the fifth version of HKB, B5, was automatically created, based on the training of the case database. The results obtained and presented in this document point out to the efficiency of the automatic knowledge acquisition methods and the deductive learning techniques implemented, through the reorganization and training of the neural networks.

The main contributions of this research are the design of a mechanism for the integration of frames and neural networks, the specification of the learning methods for the hybrid model and the analysis of factors such as specificity and sensibility in the process of automatic knowledge acquisition.

KEYWORDS: Artificial Intelligence, Expert Systems, Machine Learning, Neural Networks, Hybrid Systems, Computer Science applied to Medicine

1 INTRODUÇÃO

Este capítulo introduz os diversos tópicos envolvidos no desenvolvimento de sistemas especialistas, tais como inteligência artificial, sistemas especialistas, aprendizado de máquina e redes neurais.

1.1 Introdução aos Sistemas Especialistas

As primeiras pesquisas da área de inteligência artificial, na década de 60, eram bastante ambiciosas. Visavam construir "resolvedores de problemas" genéricos, com *interfaces* em linguagem natural. O exemplo mais significativo destes programas foi o GPS (*General Problem Solver*), que podia provar alguns teoremas simples e resolver quebra-cabeças. A gama de problemas tratáveis pelo GPS, todavia, não era muito abrangente.

As pesquisas em inteligência artificial começaram a obter melhores resultados quando passaram a focalizar problemas mais restritos, envolvendo áreas específicas do conhecimento. Os programas capazes de mimetizar o desempenho de um especialista humano na realização de determinadas tarefas foram denominados Sistemas Especialistas (SEs) [HAY 83].

1.2 Classificação dos Sistemas Especialistas

As metodologias de resolução de problemas utilizadas por seres humanos são extremamente dependentes dos tipos de problemas que as pessoas querem resolver. Conseqüentemente, os processos utilizados exibem diferentes peculiaridades de acordo com as classes de problemas.

As características de SEs também são determinadas pelas propriedades da classe de problema para a qual o sistema será empregado [HAY 83]. Por isso, é importante analisar as classes mais significativas.

- **Interpretação:** consiste na análise de dados para determinar seu significado. Um sistema para interpretação elabora explicações para uma situação, com base na análise de dados e atribuição de significado simbólico aos dados observados. O principal problema desta classe é que os dados a serem analisados apresentam erros e ruídos. Por isso, o sistema deve ser capaz de:
 - manipular informação parcial;
 - identificar quando uma informação é incerta, incompleta, ou quando a mesma corresponde a uma inferência do sistema, para evitar interpretações não confiáveis;
 - explicar como a interpretação se apóia nas evidências.

- **Classificação e Diagnóstico:** a classificação consiste na determinação da classe de um elemento específico, baseada no conjunto de suas características. A tarefa de diagnóstico é análoga à classificação e corresponde à detecção de falhas ou defeitos em sistemas, dado um conjunto de sintomas. A maioria dos SEs aplicados à medicina são classificados como sistemas de diagnóstico: uma doença é identificada a partir de um conjunto de achados médicos e sintomas. As principais dificuldades existentes para esta categoria são:
 - alguns sintomas podem ser mascarados por outros sintomas;
 - sintomas podem ser intermitentes;
 - algumas informações podem ser ou muito caras, inacessíveis ou, ainda, de obtenção muito perigosa;
 - o funcionamento e fisiologia dos organismos vivos não são completamente conhecidos.

- **Monitoração:** consiste no processo de observar continuamente o comportamento de um sistema, para tomar as medidas cabíveis quando uma determinada situação ocorre. As dificuldades dos sistemas de monitoração são muito similares àquelas dos sistemas de diagnóstico, já que as situações a reconhecer são descritas por falhas ou mau funcionamento de componentes.

- **Planejamento:** consiste no processo de determinar um conjunto de ações apropriadas para que um objetivo específico seja atingido. Os problemas de planejamento são, geralmente, extensos e complexos, cabendo enfrentá-los com recurso a subtarefas. Como as mesmas exibem interdependências, um sistema de planejamento deve ser capaz de controlar todas estas interações. O sistema deve ser, ainda, capaz de manipular graus de incerteza, já que os requisitos e as situações de um trabalho de planejamento são apenas parcialmente conhecidos, além de sujeitos à mudança.

- **Projeto:** consiste no desenvolvimento de especificações de objetos, de forma a satisfazer um conjunto de requisitos. A maior dificuldade encontrada em sistemas de projeto, tal como em sistemas de planejamento, é que as soluções dos problemas tratados pelos mesmos não podem ser selecionadas de um grupo de soluções possíveis, mas devem ser construídas. Os principais problemas advindos desta dificuldade são:
 - geralmente as restrições de um projeto provêm de diferentes fontes e não há teoria que as integre;
 - aqui os problemas também são muito complexos e devem ser particionados. No entanto, deve-se poder controlar as interdependências dos sub-problemas;
 - um sistema deve ser capaz de manter um protocolo de justificativas de decisões de projeto, para que possa reconsiderar

estas decisões durante um processo de modificação de parte do projeto.

- **Instrução:** objetivam ensinar alunos buscando eliminar deficiências em seus conhecimentos. Um sistema de instrução, tipicamente, tem início no diagnóstico das falhas no conhecimento do aluno, a partir da análise do seu comportamento. Por isto, um sistema de instrução deve manter uma base de modelos de alunos. A partir de então, o sistema planeja uma seqüência de tarefas que podem auxiliar o aluno em suas deficiências. Para determinar a próxima tarefa a apresentar ao aluno, depois de um tempo de interação, o sistema deve ser capaz de identificar o que o estudante já sabia e o que aprendeu. Dadas estas características, verifica-se que os requisitos para um sistema de instrução envolvem aqueles dos sistemas de diagnóstico, interpretação e planejamento.
- **Simulação:** consiste na tradução do processo de predição de alguma situação futura, a partir de uma situação inicial. Sistemas de simulação devem geralmente ser capazes de manipular a variável tempo, já que a maior parte das situações se modificam com o transcorrer deste último. Um sistema de simulação também deve ser capaz de manipular dados imprecisos e incompletos; para problemas que envolvem dados precisos e completos, predições e simulações não são necessárias.

A área de diagnóstico em geral, especificamente o diagnóstico médico, tem sido alvo de desenvolvimento de vários SEs [MIL 86] [FIE 90]. No entanto, os SEs não se restringem a este tipo de aplicação. Também são utilizados na prospecção geológica, na análise de situações militares, no controle do tráfego aéreo, no processamento de imagens entre inúmeras outras aplicações [PEA 88].

1.3 Representação de Conhecimento e Arquitetura dos SEs

A tarefa de resolver problemas em um domínio específico é, na realidade, o objetivo de qualquer programa de computador. O modelo usado pela computação convencional manipula dois tipos de conhecimentos organizados em dados e algoritmos. Os algoritmos determinam as formas com que problemas específicos são solucionados a partir dos dados disponíveis.

Já os SEs, ao invés de dados e algoritmos, utilizam conhecimento e mecanismos de inferência [HAR 85].

O conceito de *dado* é associado aos valores não interpretados que um sistema possui para processamento. O conceito de *informação* aparece quando dados são selecionados e organizados para um determinado fim. Já, o conceito de *conhecimento* é ainda mais abrangente do que o conceito de informação. No *conhecimento*, é introduzida a noção de estrutura, sendo permitida a exploração das relações entre as porções das informações. Assim, o *conhecimento* compõe-se de uma coleção de fatos, agrupados em uma estrutura que inclui as relações entre tipos de objetos, tipos de classes e de indivíduos.

CONHECIMENTO

Representar conhecimento, por sua vez, implica em encontrar as estruturas adequadas para expressar o tipo de conhecimento particular do domínio da aplicação [NIL 80].

A Base de Conhecimento (BC) é o repositório que armazena o conhecimento factual e heurístico do SE. O primeiro consiste em toda a informação disponível, compartilhada e geralmente aceita por especialistas em uma determinada área. O segundo é formado por regras não muito bem discutidas ou especificadas, através das quais os especialistas tomam decisões em suas áreas.

São várias as formas de representação de conhecimento existentes. Algumas se mostram mais adequadas para esta ou aquela tarefa; entretanto, não há uma métrica formal para medir a adequação de um esquema de representação para um problema.

O cálculo de predicados é um mecanismo de representação de conhecimento que enfatiza aspectos declarativos do conhecimento do domínio [HAR 85] [BAR 86] e, por isto, é apropriado para representar porções passivas do conhecimento do domínio. Um predicado é uma construção que expressa algo sobre um objeto, podendo assumir o valor falso ou verdadeiro. Por exemplo, o predicado *Homem (Paulo)* associa a característica *é um homem* para o objeto *Paulo*, podendo esta associação ser verdadeira ou falsa. Os predicados podem ainda ter vários argumentos e usar os conetivos *e*, *ou*, *se..então*, *não* e *somente se*, para formar sentenças compostas, como por exemplo:

- Se voa(X) e tem(X,asas)
Então, avião(x)
- Se voa(X)
Então, tem(X,asas)
- Voa(X)
Logo, avião(X).

No exemplo, chega-se à conclusão de que a variável X é um avião, a partir da ocorrência de voa(X).

As regras de produção constituem uma das mais populares formas de codificação da experiência de especialistas para resolver problemas [CLA 83] [HAR 85] [BAR 86]. Elas enfatizam os aspectos procedimentais do conhecimento e por isto são mais apropriadas para representar características comportamentais da aplicação. Neste esquema, o conhecimento é codificado como um conjunto de regras do tipo *se condição, então, ações*. A condição coloca um estado que deve ser encontrado para que a regra seja aplicável. A ação corresponde a algum procedimento que acarreta uma conclusão ou mudança no estado corrente.

As redes semânticas constituem outro mecanismo de representação de conhecimento que enfatiza aspectos declarativos do conhecimento. As redes semânticas utilizam nodos ligados por arcos para representar objetos, conceitos ou situações e seus relacionamentos [HAR 85] [BAR 86]. Uma ligação pode ser vista como uma afirmação feita para um elemento em relação a outro. Por exemplo, a

ligação *é um*, que conecta o nodo *Paulo* ao nodo *homem*, expressa o relacionamento *é um* entre os dois nodos.

Também existem os esquemas de representação que dão ênfase aos aspectos estruturais do conhecimento do domínio, como por exemplo o mecanismo de *frames*. *Frames* foram originalmente propostos por Minsky, em 1975, como base para compreensão da percepção visual, diálogos em linguagem natural e outras tarefas complexas [MIN 75]. Basicamente, um *frame* é uma estrutura que descreve um objeto ou uma situação estereotipada. Por exemplo, um *frame* pode conter a descrição de uma determinada classe de animais, de uma certa doença ou, mesmo, a descrição de uma pessoa. Os *frames* são formados por *slots* que armazenam propriedades do objeto representado, o que é análogo à função de atributos em uma estrutura do tipo registro. Os *slots* podem armazenar não só valores como também ponteiros para outros *frames*, procedimentos através dos quais um valor é obtido ou, ainda, valores *default* [FIK 85] [HAR 85]. A Figura 1.1 contém um *frame* chamado Pessoa, descrito pelos slots *Nome*, *Idade* e *Endereço*. Este frame está preenchido na instância de nome Paulo, descrita pelas propriedades: Nome: Paulo Flores, Idade: 21, Endereço: POA.

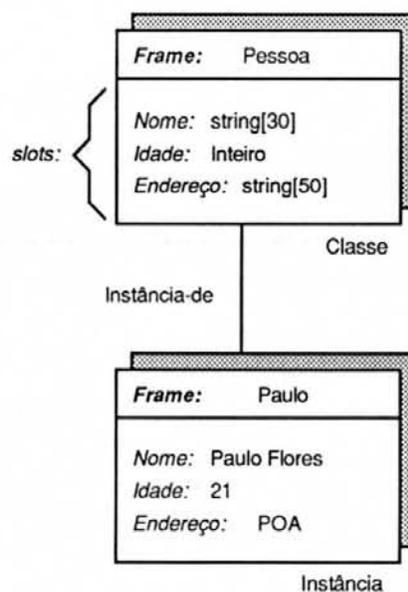


Figura 1.1 Exemplo de representação de conhecimento em *frames*

Independentemente do mecanismo de representação de conhecimento utilizado, todo o SE possui um motor de inferência, que é o componente responsável pela manipulação e busca de fatos na BC, com o objetivo de encontrar soluções para os problemas apresentados.

Tradicionalmente, os sistemas de computador funcionam como "caixas-pretas", ou seja, o usuário coloca um problema para o computador que simplesmente apresenta uma resposta. Assume-se que o usuário não tem interesse em compreender a forma de resolução do problema e que, na maioria dos casos, sequer tem a capacidade para tanto. Estas suposições não devem prevalecer nos SEs, por dois motivos principais [ROL 88]:

Geralmente, os SEs trabalham com dados incompletos. Por isso, é essencial que estes sistemas apresentem conclusões que assegurem os resultados obtidos. Os SEs são geralmente utilizados como sistemas de apoio à decisão ou como tutores. Em qualquer uma das aplicações, a explanação do raciocínio desenvolvido pelo sistema é requisito essencial.

Portanto, além de resolver o problema, o SE também deve ser capaz de descrever seu processo de raciocínio, explicá-lo e justificar as conclusões a que chegou. O mecanismo de explanação é o componente do sistema responsável por estas tarefas.

Os SEs também possuem uma interface com o usuário, encarregada da comunicação entre o SE e o mundo exterior. Esta interface deve ser capaz de operar nos seguintes modos [ROL 88]:

- teste: usuário busca validar o comportamento do sistema;
- aquisição de conhecimento: usuário provê conhecimento adicional ao sistema ou modifica o conhecimento já existente;
- pupilo: usuário procura aprender com o sistema, consultando o conhecimento organizado e destilado da BC.

- consulta: usuário utiliza o sistema para resolver um problema real.

A Figura 1.2 apresenta a arquitetura básica de um sistema especialista.

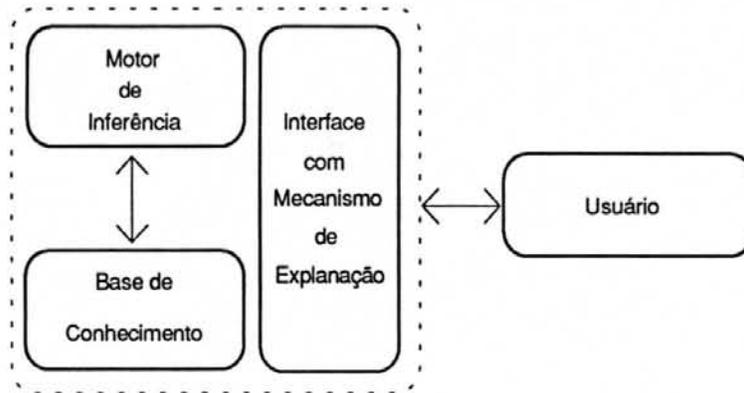


Figura 1.2 Arquitetura básica dos sistemas especialistas

1.4 Aplicações de Sistemas Especialistas

A partir da década de 70, vários sistemas especialistas começaram a emergir. Foram desenvolvidos em diversas áreas de aplicação para tratar diferentes classes de problemas. A Tabela 1.1 apresenta um quadro onde são apresentadas as características básicas dos principais sistemas especialistas desenvolvidos neste período .

A maior parte destes sistemas trouxe resultados bastante significativos para as pesquisas de IA. Apesar de todos os resultados positivos, devido a complexidade envolvida em comportamentos inteligentes como os mimetizados pelos SEs, certas dificuldades continuam a desafiar os pesquisadores de inteligência artificial.

TABELA 1.1 Características básicas dos sistemas especialistas

Nome	Domínio	Classe de problema	Instituição/ano
ABSTRIPS	Planejamento das ações de um robô	Planejamento	SRI Int. 1974
CASNET	Diagnóstico de glaucoma Medicina Interna	Diagnóstico	Rutgers 1977
DENDRAL	Espectropia de massa	Interpretação	Stanford 1969
EL	Análise de circuitos elétricos	Interpretação	MIT 1976
EXPERT	Diagnóstico de glaucoma	Diagnóstico	Rutgers 1976
GA1	Espectropia de massa	Interpretação	Stanford 1976
INTERNINST /QMR	Medicina Interna	Diagnóstico	Pittsburgh 1975
HEARSAY-II	Reconhecimento de linguagem falada	Interpretação	Carnegie-Mellon 1976
MACSYMA	Matemática simbólica	Interpretação	MIT 1971
MDX	Doenças do fígado	Diagnóstico	Colombus/Ohio 1979
MOLGEN	Experimentos em genética molecular	Planejamento	Stanford 1979
MYCIN	Doenças infecciosas do sangue	Diagnóstico	Stanford 1976
NOAH	Planejamento de ações de robôs	Planejamento	SRI Int. 1975
PDS	Diagnóstico de falhas em máquinas	Diagnóstico	Carnegie Melon 1983
PIP	Problemas renais	Simulação	MIT 1976
PUFF	Análise de exames de pulmões	Interpretação	Stanford 1981
PROSPECTOR	Exploração mineral	Interpretação	SRI Int. 1976
R1	Configuração de computadores	Projeto	Carnegie-Mellon 1980

1.5 Dificuldades não Equacionadas nos Sistemas Especialistas

Vários dos SEs desenvolvidos a partir da década de 70 apresentaram resultados muitas vezes superiores aos resultados geralmente alcançados por especialistas humanos [ROL 88]. Todavia, algumas limitações já podiam ser diagnosticadas. Em primeiro lugar, verificava-se uma grande dificuldade na explicitação de conhecimento de especialistas humanos e na formalização deste conhecimento para uma estrutura tratável pelo computador. Este processo foi considerado como uma das etapas mais demoradas e caras na construção de sistemas especialistas e inspirou o desenvolvimento de muitas metodologias e ferramentas para aquisição de conhecimento [BOO 88].

As metodologias de aquisição de conhecimento geralmente usadas consistem em entrevistas, estruturadas ou não estruturadas, a partir das quais o engenheiro de conhecimento vai organizando as informações. No entanto, até agora não existe uma técnica que seja amplamente aceita, marcada por melhor desempenho na explicitação de conhecimento de especialistas.

A criação de grandes bases de conhecimento para sistemas especialistas é, na prática, uma tarefa muito difícil de ser efetuada sem o auxílio de ferramentas para aquisição de conhecimento. São diversos os fatores que dificultam o processo de aquisição de conhecimento, podendo-se destacar os seguintes problemas:

- o conhecimento necessário para construção da base de conhecimento de um sistema especialista não se encontra isolado, nem organizado de forma a permitir um mapeamento direto para um dos formalismos clássicos de representação de conhecimento, por exemplo *frames*, redes semânticas e regras;
- as regras utilizadas pelos especialistas para resolver problemas são geralmente formulações heurísticas sobre as quais nem mesmo os próprios especialistas têm consciência ou controle. Os problemas são resolvidos pela "experiência";

- o engenheiro de conhecimento geralmente não possui conhecimento na área do especialista, o que dificulta o diálogo entre ambos e a compreensão das regras para resolução dos problemas utilizadas pelo especialista;
- a tarefa de explicitação de conhecimento corporativo (ou conhecimento compartilhado por um grupo de especialistas) torna-se ainda mais difícil, já que, além dos problemas existentes para a aquisição de conhecimento de um único especialista, surgem as dificuldades para cálculo do "conhecimento médio" do grupo de especialistas.

Outro problema diagnosticado é que, ao contrário dos especialistas humanos, os SEs não conseguiam aprender com a experiência. Ou seja, os sistemas existentes, na maioria das vezes, só podiam atualizar seu conhecimento com a ajuda de um ser humano, no caso, um engenheiro de conhecimento. Diversas técnicas foram desenvolvidas para incorporar aos sistemas de computador os métodos de aprendizado estudados nas áreas de biologia, psicologia e medicina [MIC 83].

Carbonell divide os paradigmas de aprendizado em 4 classes distintas [CAR 89]:

- aprendizado baseado no paradigma indutivo: consiste na formulação de conceitos (leis) com base na observação de várias ocorrências particulares de um fenômeno. Desta forma, a partir de um conjunto positivo e outro negativo de instâncias de conceitos, deve-se gerar alguma regra de descrição que identifique estes e todos os futuros exemplos como instâncias ou "não instâncias" do conceito;
- aprendizado baseado no paradigma dedutivo: baseia o processo de aprendizagem em uma rica teoria do domínio e em conjuntos de exemplos já existentes. Então, a partir da análise dos modelos gerados pela aprendizagem indutiva, geram-se novos modelos da realidade. Por exemplo; dada a ocorrência de

uma doença desconhecida por parte da população de uma região, pesquisadores passam a analisar os casos, verificando os sintomas apresentados pelas pessoas afetadas. Assim, chegam a conclusões de que, se uma pessoa apresenta determinados sintomas, então é porque está acometida pela nova doença, caso contrário, não;

- aprendizado baseado no paradigma conexionista: esta abordagem consiste na utilização de redes neurais para armazenar conhecimento e conduzir o aprendizado. Os sistemas conexionistas são treinados a reconhecer objetos de um conjunto de classes através da amostra repetitiva de conjuntos semelhantes de padrões. Por exemplo, pode-se treinar uma rede neural para reconhecer um caractere gráfico através da leitura repetida de várias amostras do mesmo caractere;
- aprendizado baseado no paradigma genético: este paradigma é inteiramente inspirado nas teorias de mutação, na reprodução biológica e de seleção natural. Através de algoritmos genéticos, buscam-se indivíduos com maior capacidade de adaptação ao meio em uma população em constante mutação. Estes indivíduos, por sua vez, terão maior capacidade de sobrevivência e reprodução.

Do ponto de vista biológico, contudo, são apenas 2 os paradigmas de aprendizado, um baseado no processamento de sinais sensoriais e outro baseado em processos de pensamento lógico [KUR 90]:

- processamento de sinais sensoriais: utilizado em tarefas como reconhecimento de padrões visuais, auditivos e de tato. Consiste em uma forma de inteligência observada na maioria das espécies animais;

- pensamento lógico: utilizado em tarefas de resolução de problemas. Consiste em uma forma de inteligência particular aos seres humanos.

Os processos lógicos do cérebro são simulados de maneira bem mais fácil pela inteligência artificial. Mesmo os computadores mais antigos já eram capazes de solucionar problemas simples ou simular jogos, tarefas basicamente seqüenciais.

Problemas como o reconhecimento de padrões visuais, que envolvem uma grande quantidade de processamento paralelo, ainda não têm soluções por computador, comparáveis à habilidade dos animais em realizar a mesma tarefa.

O paradigma indutivo baseia-se nos mecanismos de reconhecimento de padrões, os quais envolvem processos maciçamente paralelos, capazes de construir conceitos a partir de repetidas observações de um determinado padrão.

O paradigma dedutivo baseia-se nos processos lógicos do cérebro para gerar novos conceitos a partir de conhecimento já existente.

Assim, comparando-se as definições de aprendizado indutivo e dedutivo às de aprendizado conexionista e genético, é possível observar que os dois últimos paradigmas são casos particulares dos dois primeiros: Nas redes neurais, a repetição constante de um mesmo padrão reforça os pesos das sinapses para a identificação daquele padrão (aprendizado indutivo). Já os algoritmos genéticos, partindo de uma população inicial, geram novas populações de indivíduos que melhor se adaptam ao meio (aprendizado dedutivo).

1.6 Engenharia do Conhecimento

Uma das maneiras de solucionar os problemas da aquisição de conhecimento consiste na automação do processo de construção da BC a partir da análise de casos de treinamento.

No entanto, muitas vezes não é possível automatizar o processo de criação da base de conhecimento, pela inexistência de registros de casos reais para treinamento. Logo, a utilização de uma metodologia de aquisição de conhecimento simples, aplicável a vários domínios e capaz de fornecer informações facilmente transferíveis para um dos formalismos de representação de conhecimento, pode ser de extrema importância.

Uma metodologia particularmente relevante para este trabalho é a de construção de grafos de conhecimento médio (GCM), extraídos de múltiplos especialistas [LEA 90].

Um GCM é definido como um grafo acíclico mínimo que representa o conhecimento de especialistas para uma determinada hipótese de diagnóstico [MAC 90]. O método para obtenção dos grafos inicia por uma fase de entrevistas com um especialista, onde é solicitado que sejam escolhidos fatores importantes na resolução do problema em questão, por exemplo, sintomas importantes para o diagnóstico de uma doença.

A seguir, os especialistas devem ordenar os itens selecionados de acordo com seu grau de importância para a resolução do problema e colocá-los na base do grafo.

Os itens devem, então, ser ligados em nodos intermediários que representam a relevância de um conjunto de fatores para a resolução do problema que, por sua vez, sempre é representado pelo nodo mais ao topo do grafo.

A última etapa na construção dos grafos consiste na atribuição de valores de importância, numa escala de 0 a 10, para os itens da base e para os grupamentos de itens (nodos intermediários).

A Figura 1.3 representa um grafo de conhecimento de um especialista para uma hipótese de diagnóstico.

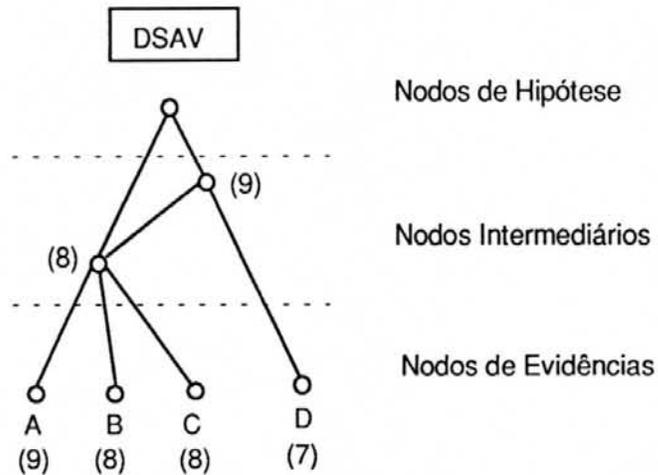


Figura 1.3 Grafo de conhecimento

Três tipos de nodos podem ser identificados na Figura 1.3:

- **nodo de hipótese:** é o nodo mais ao topo do grafo e representa a hipótese de diagnóstico nele considerada;
- **nodos de evidência:** são os nodos situados mais abaixo no grafo. Eles representam os sintomas, sinais e resultados de testes (para diagnósticos médicos) ou simplesmente dados que dão suporte à hipótese de diagnóstico. Estes nodos são organizados da esquerda para a direita em ordem decrescente de importância;
- **nodos intermediários:** representam diferentes agregações de evidências usadas pelo especialista no seu raciocínio. São comumente identificados como problemas, síndromes, etc, que têm relevância para o objetivo visado.

Os nodos intermediários geralmente representam um E-lógico entre as evidências. Já, o nodo de hipótese representa um OU-lógico. Por isso, separando-se todos os sub-grafos conectados por um OU-lógico, um grafo de conhecimento pode ser facilmente transformado em outros tantos.

O método completo para obtenção dos grafos de conhecimento, utilizado em [LEA 88] na explicitação de conhecimento para diagnóstico de cardiopatias congênitas, está descrito no apêndice I.

1.7 Introdução à Abordagem Conexionista

Nos últimos anos, foi tentada a solução de alguns dos problemas de aprendizado em SEs através do uso de redes neurais como mecanismo de representação de conhecimento. Os modelos de redes neurais, na maioria das vezes, visam não apenas desempenhar funções do cérebro como também basear suas operações em conhecimentos da ciência sobre seu funcionamento e arquitetura.

O cérebro humano, pesando não mais que 2 Kg, é muitas vezes citado como o objeto mais complexo do universo. Ele contém cerca de 100 bilhões de neurônios ligados em redes que dão origem à inteligência, criatividade, emoções, consciência e memória. No entanto, apesar do enorme número de células, a capacidade do cérebro em armazenar e buscar informações, experimentar emoções, controlar movimentos, não pode ser explicada pelo simples agrupamento das mesmas.

Em parte, a complexidade do cérebro reside na diversidade de células nervosas e no grande número de ligações entre elas. Santiago Ramón y Cajal - o pai da moderna ciência da mente - descreveu os neurônios como "as misteriosas borboletas da alma cuja batida das asas pode um dia - quem sabe ? - esclarecer o segredo da vida mental" [FIS 92].

Os neurônios são definidos como células polarizadas capazes de receber sinais em suas extensões altamente ramificadas (chamadas dendritos) e transmitir informações por extensões não ramificadas (chamadas axônios). Ao ser excitado, um neurônio transmite informações para outros neurônios através da geração de impulsos conhecidos como potenciais de ação. Estes sinais são propagados como ondas pelo único axônio da célula e são convertidos para sinais químicos nas sinapses, ou pontos de conexão entre os neurônios. A Figura 1.4 apresenta a estrutura básica de um neurônio.

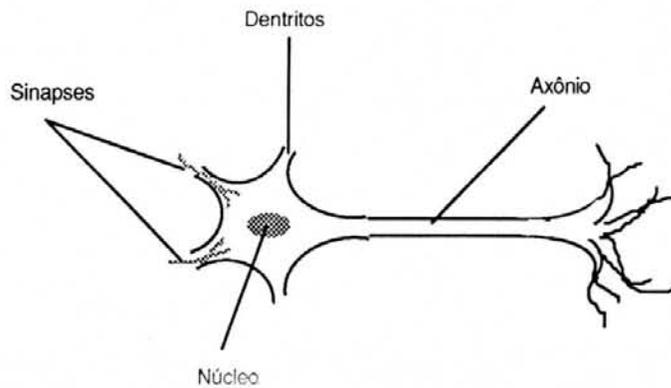


Figura 1.4 Neurônio

A forma de funcionamento de um neurônio pode ser esquematizada do seguinte modo: um neurônio recebe sinais de entrada que são computados e que, se forem fortes o suficiente (ultrapassarem um limiar biológico), são propagados para outras células.

O modelo funcional de um neurônio teria basicamente 3 partes:

- sinapses que ligam os neurônios;
- axônio, através do qual sinais são propagados para outros neurônios;
- um elemento processador (núcleo) que combina os impulsos de entrada e que os propaga, se suficientemente fortes.

A forma básica de funcionamento de um neurônio é muito simples. A complexidade das redes neurais compensa, todavia, a simplicidade de funcionamento de cada célula. Logo, a capacidade de uma rede de aprender, associar e reagir não está associada a cada neurônio em particular, mas sim ao grande número de conexões entre eles [KAN 92].

Os processos de reconhecimento de padrões existentes em todos os animais superiores não são implementados de forma natural nas arquiteturas convencionais de computadores. Estes processos são notadamente paralelos e envolvem a análise simultânea de múltiplos aspectos. A Figura 1.5 apresenta um exemplo onde alguns

caracteres ambíguos são diferenciados através de sua associação com outros caracteres. Isto indica que o sistema humano de percepção é capaz de rapidamente explorar todas as possíveis formas de classificação para um padrão e só decidir por uma delas após a análise de todas as possibilidades [McC 88].



Figura 1.5 Caracteres ambíguos

Análoga ao reconhecimento de padrões visuais é a capacidade de analisar múltiplas situações. Quando em uma festa infantil, em um restaurante, por exemplo, vêm à memória informações sobre festas infantis e sobre restaurantes e compõe-se uma representação média entre as duas situações. Por exemplo, sabe-se que esta festa não será uma junção completa entre festas infantis e restaurantes. Analisando múltiplas situações, verifica-se que, na festa, não haverá jogos do tipo esconde-esconde. Os modelos conexionistas buscam capturar esta capacidade do cérebro, de relacionar elementos, tais que uns influenciem os outros.

1.7.1 Modelos conexionistas

Um dos principais modelos conexionistas, introduzido no final da década de 50, foi o *Perceptron*. A capacidade de aprendizado e adaptação que exibia, despertou grande interesse e trouxe muitas expectativas. Em 1969, no entanto, Minsky e Papert [MIN 69] apontaram uma série de limitações para o modelo, o que fez com que, por vários anos, as redes neurais fossem praticamente abandonadas.

Por volta de 1974, Paul J. Werbos [HIN 92] inventou um procedimento eficiente para o ajuste dos pesos de uma rede. O procedimento, agora conhecido como *backpropagation*, foi ignorado por muitos anos e, apenas na década de 80, foi redescoberto e popularizado por D. E. Rumelhart, D. Parker e Ronald J. Williams [RUM 86].

Atualmente, observa-se crescente interesse pela pesquisa em redes neurais. Psicólogos se interessam pelos modelos conexionistas por suas semelhanças estruturais e comportamentais em relação à mente humana. Os pesquisadores de IA, por sua vez, buscam nos mesmos soluções eficientes para um dos problemas-chave de IA, o aprendizado de máquina [KNI 89]. Os modelos conexionistas podem ser descritos através dos seguintes critérios:

- **propriedades da rede:** um modelo conexionista consiste em uma rede de células, relativamente autônomas, dotadas de simples capacidade de processamento. As células são ligadas por arcos ou conexões, cada uma armazenando um peso p_{ij} , correspondente à influência da célula c_j no processamento do sinal de saída da célula c_i (Figura 1.6). Pesos positivos correspondem a fatores de reforço do sinal de entrada, e pesos negativos correspondem a fatores de inibição do sinal;
- os modelos geralmente apresentam um conjunto de células de entrada, através das quais informações são passadas para a rede. Estas células não possuem arcos de entrada e são representadas na Figura 1.6 pelos círculos mais abaixo do desenho. Também observam-se nos modelos células de saída, ou unidades que apresentam os sinais de saída da rede como um todo (círculos mais acima - Figura 1.6). As outras células da rede são chamadas células intermediárias. Os modelos de redes que não permitem a existência de ciclos entre as células são classificados como *feedforward*. Modelos que permitem estes ciclos classificam-se como *feedback*;

- propriedades das células:** cada unidade c_i é capaz de processar um sinal de entrada e transformá-lo em um sinal de saída, ou ativação. Tipicamente, todas as células da rede empregam a mesma função para cálculo da ativação. Este cálculo é efetuado com base nos sinais conduzidos por todos os arcos que chegam até a célula e seus pesos de conexão. Assim, a ativação da célula c_{m+1} é determinada pela ativação das células c_1 e c_2 , e pelos pesos da conexão $p_{1,m+1}$ e $p_{2,m+1}$, como ilustra a figura 1.6. As entradas e ativações de cada célula podem, ainda, ser discretos, assumindo valores do conjunto $\{0,1\}$ ou $\{0, 1, -1\}$, ou contínuos, assumindo valores no intervalo $[0,1]$ ou $[-1,1]$;
- propriedades dinâmicas:** o método estabelecido para o cálculo das ativações das células pode ser diferente para cada modelo conexionista. Em alguns modelos, as células são percorridas em uma ordem fixa, sendo a ordem do cálculo da ativação de cada célula dependente da ordem com que são analisadas. Em outros, as células calculam suas ativações e apresentam suas saídas de forma simultânea. Outros modelos selecionam uma célula aleatoriamente, computam sua ativação e, então, alteram seu sinal de ativação antes que qualquer outra célula compute sua ativação.

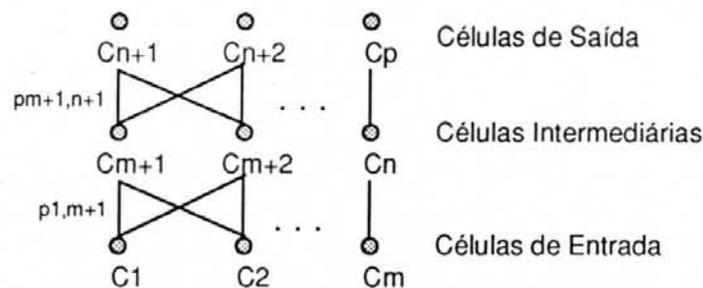


Figura 1.6 Estrutura geral de um modelo neural

As representações dos modelos conexionistas podem ainda ser classificadas como locais ou distribuídas. Nas representações distribuídas, cada conceito ou item de conhecimento é representado como um padrão de ativação em um conjunto de unidades. Numa representação local, cada unidade expressa exatamente um conceito. As vantagens de cada uma destas formas de representação são extensamente discutidas em [HIN 86].

Tabela 1.2 Principais modelos de redes neurais

Modelo/Ano	Exemplo de Aplicação	Propriedades das Células	Propriedades das Redes
Perceptron 1957	Reconhecimento de caracteres impressos	Células recebem valores de entrada binários e apresentam como saída os valores -1 ou +1.	Rede acíclica de 2 camadas
Perceptron Multi-camadas 1974	Controle adaptativo de braços de robôs e processamento de fala	Células semelhantes às do Perceptron, operando porém com valores contínuos	Rede acíclica de no mínimo 3 camadas
Sistema ART 1978	Reconhecimento de sinais de radar e processamento de imagens	Células possuem entradas binárias, podendo assumir valores contínuos	Rede cíclica de 3 camadas
Redes de Kohonen 1980	Reconhecimento de fala e aprendizado da distribuição de probabilidades dos dados	Células recebem valores de entrada contínuos	Rede acíclica de duas camadas
Redes de Hopfield 1982	Reconhecimento de imagens a partir de fragmentos e memória associativa	Células recebem valores de entrada binários e apresentam como saída os valores -1 ou +1.	Rede cíclica de 1 camada

Durante os últimos anos, diversos modelos conexionistas foram criados e aplicados nas mais variadas tarefas, como, por exemplo, no processamento de

sinais, no reconhecimento de padrões, no processamento de imagens, no processamento de conhecimento, em estudos de conexões biológicas, etc [SIMP 90].

A tabela 1.2, modificada de [MAC 91a], apresenta um quadro comparativo entre 5 modelos mais representativos.

1.8 Sistemas Especialistas Conexionistas

Os sistemas especialistas conexionistas foram desenvolvidos com o objetivo de solucionar os dois principais problemas de SEs anteriormente citados: o problema do aprendizado e o da construção da base de conhecimento. Por serem mecanismos adaptativos, as redes neurais são capazes de ajustar seu conhecimento com base em exemplos de treinamento ou, até mesmo, de aprender sem que haja qualquer conhecimento prévio.

A tabela 1.3 apresenta um quadro comparativo entre 3 importantes modelos conexionistas: SEC, RUBICON e MNC.

O SEC foi um dos sistemas classificatórios precursores na utilização de um modelo conexionista para uso em sistemas especialistas [GAL 88]. A arquitetura do sistema especialista é composta por uma base de conhecimento (formada por uma matriz de nomes de variáveis, pesos e perguntas), um motor de inferência e uma interface com o usuário. O sistema provê métodos para construção da rede neural por exemplos de treinamento, mas sugere que os métodos não substituam os engenheiros de conhecimento, mas sim, sejam empregadas por eles como ferramenta de apoio à construção de parte ou de toda a base de conhecimento.

O Rubicon [SAM 92] é um modelo conexionista desenvolvido para uso em SEs que emprega tanto representação local quanto distribuída. A representação distribuída é usada na comunicação do sistema com o mundo exterior, oferecendo uma interface tolerante a erros e ruídos. Todo o processamento interno da rede, no entanto, emprega representação local, permitindo a construção de redes para diferentes aplicações. O Rubicon tem arquitetura híbrida que integra sistemas baseados em regras ao modelo de representação distribuída/local. As regras

armazenadas constituem-se de uma cláusula antecedente e de uma cláusula conseqüente. Ambas as cláusulas contêm um número arbitrário de expressões. Cláusulas conseqüentes especificam asserção e retração de expressões. Sempre que alguma parte da informação "casar" com o antecedente de alguma regra, a expressão conseqüente aparecerá na camada de saída da rede. A utilização do sistema conexionista integrado às regras possibilita que o "casamento" de informações seja apenas parcial, de forma que as regras possam ser disparadas mesmo sem que todos os componentes da expressão antecedente estejam presentes.

Tabela 1.3 Modelos de redes neurais aplicados a sistemas especialistas

Modelo/Ano	Propriedades das Células	Propriedades da Rede	Aplicações
SEC 1988	Ativação das células pode assumir os valores: 1, para verdadeiro; 0, para desconhecido e -1 para falso.	Rede acíclica de 3 camadas. A camada de entrada representa os sintomas, a camada intermediária representa as doenças e a camada de saída, os tratamentos.	Diagnóstico de doenças do sarcó-fago e seleção de tratamentos.
RUBICON 1988	As células possuem entradas binárias e saídas que assumem os valores 0 ou 1.	Rede acíclica de 6 camadas.	Identificação de títulos de filmes.
MNC 1989	A ativação das células pode assumir um valor no intervalo [0,1].	Rede acíclica de 3 ou mais camadas. As camadas de entrada representam evidências, as camadas combinatórias agrupamento de evidências e a camada de saída as hipóteses.	Diagnóstico de doenças renais.

O Modelo Neural Combinatório é utilizado na arquitetura híbrida para sistemas especialistas descrita neste trabalho. Por isto, será apresentado com maior nível de detalhe.

"O Modelo Neural Combinatório (MNC) é um modelo conexionista capaz de executar aprendizado heurístico baseado no reconhecimento de regularidades ambientais através de repetidas observações do mundo exterior" [MAC 89]. A topologia de uma rede de tipo MNC segue o paradigma *feedforward* (não permite ciclos) e possui três camadas. A Figura 1.7 mostra um exemplo de uma rede baseada no MNC.

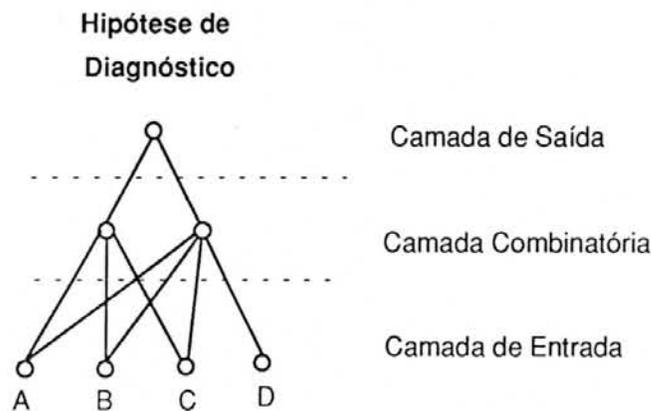


Figura 1.7 Topologia do Modelo Neural Combinatório (MNC)

A camada de entrada é formada por células *fuzzy*, cada uma representando um achado específico. Estas células recebem valores de entrada entre 0 e 1, correspondendo ao grau de confiança que o usuário tem nas informações que ele observa e passa para a rede neural.

Células em camadas diferentes são ligadas por conexões às quais se associam pesos, representando a influência das células da camada inferior na saída daquelas situadas na camada superior.

As conexões da camada de entrada podem ser excitatórias ou inibitórias. Uma conexão excitatória propaga o sinal de entrada utilizando o seu peso de conexão como um fator de atenuação. Uma conexão inibitória primeiramente complementa o

sinal de entrada X , transformando-o em $(1-X)$, para depois propagá-lo através da multiplicação do sinal $(1-X)$ pelo peso da conexão.

A camada combinatória associa diferentes células de entrada em nodos intermediários de conhecimento que representam os padrões (grupos de evidências) que a rede aprende a reconhecer.

A camada de saída armazena as diferentes hipóteses de diagnóstico existentes no domínio do problema. Esta camada implementa um mecanismo de competição entre os diferentes caminhos formados nas camadas anteriores.

O valor do sinal de saída de uma sinapse, ou seu "Fluxo Evidencial", é calculado através da fórmula:

$$FE = w.E.t + w.(1-E).(1-t)$$

onde:

- E é o sinal de entrada;
- w é o peso da sinapse;
- t é o tipo da sinapse, assumindo o valor 1 para sinapses excitatórias e 0 para inibitórias.

Os neurônios podem processar sinais de acordo com uma das seguintes formas:

- *E-nebuloso*: o neurônio seleciona o sinal de entrada mais fraco;
- *Ou-nebuloso*: o neurônio seleciona o sinal de entrada mais forte;

O modelo utiliza o *E-nebuloso* nas camadas combinatórias e o *Ou-nebuloso* na camada de saída.

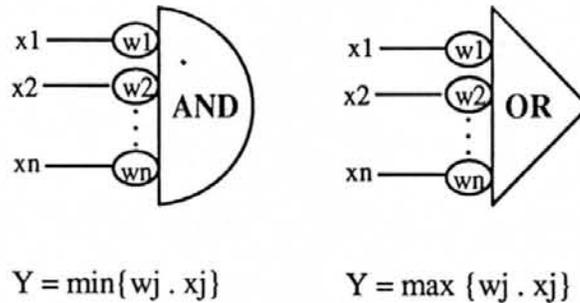


Figura 1.8 Modelo E-nebuloso e Ou-nebuloso

A Figura 1.8, extraída de [MAC 89], ilustra o funcionamento dos dois tipos de neurônios, onde:

- x_j representa os sinais de entrada das células;
- w_j é um fator que representa a força da conexão ou a importância do sinal de entrada x_j para determinar a saída Y ;
- Y representa o estado de ativação da célula ou, ainda, seu sinal de saída, expressando o grau de credibilidade do conceito representado pelo neurônio.

A rede possui duas formas principais de operação: modo de consulta e modo de aprendizado. No modo de consulta, a partir de um conjunto de evidências (entradas), a rede propaga os sinais e obtém as hipóteses (saídas) que obtiverem grau de aceitação maior que um limiar de aceitação pré-estabelecido. No modo de aprendizado, são dois os métodos que podem ser empregados. O primeiro, utilizando um *algoritmo de punição e recompensa - versão de arranque*, é capaz de gerar uma rede operante a partir de casos de treinamento. O segundo, utilizando um *algoritmo de punição e recompensa - versão de refinamento*, ajusta os valores dos pesos de conexão das redes a partir dos casos de treinamento. Ambos os algoritmos são detalhados no capítulo 4.

Além de possibilitar que o aprendizado heurístico seja implementado, O MNC é considerado uma importante conexão entre as áreas de redes neurais e engenharia de conhecimento, já que sua estrutura de funcionamento e sua topologia são

baseadas nos grafos de conhecimento obtidos através de uma metodologia de aquisição de conhecimento.

1.9 Arquiteturas Conexionistas Híbridas

Os sistemas conexionistas apresentaram soluções eficientes para os problemas de aprendizado e de construção da base de conhecimento. Todavia, alguns obstáculos também surgiram pela utilização de redes neurais como única forma de representação de conhecimento. Depois de treinadas para executar uma tarefa específica, as redes passam a manter seu conhecimento na forma de pesos de conexão, o que dificulta a análise e compreensão do conhecimento armazenado. Esta característica também torna complexa a tarefa de obtenção de explicações da resolução de um problema por uma rede neural, o que é essencial para SEs.

Além disto, para a maioria das aplicações, faz-se necessário a utilização de formalismos de representação de conhecimento ricos e flexíveis, que possibilitem o armazenamento de descrições completas e fiéis ao conhecimento factual do domínio. Redes neurais, no entanto, não oferecem tais mecanismos de representação.

Os sistemas especialistas de arquitetura híbrida foram introduzidos como uma forma de solução para estes problemas. Nestes modelos, os paradigmas simbólico e conexionista se complementam: um pela sua natureza lógica, cognitiva e mecânica; o outro pela sua natureza numérica, associativa e de auto-organização. Dados estes atributos, utiliza-se o paradigma simbólico para capturar mais semântica na modelagem do conhecimento do domínio e o paradigma conexionista para prover os recursos necessários tanto para a aquisição quanto para o refinamento do conhecimento.

1.9.1 Comparação entre os paradigmas simbólico, conexionista e híbrido

Historicamente, observa-se que os sistemas especialistas evoluíram para arquiteturas híbridas, com o objetivo de sanar problemas inerentes a um ou outro paradigma de representação de conhecimento.

A Tabela 1.4 apresenta uma comparação entre os paradigmas simbólico, conexionista e híbrido, no desenvolvimento de sistemas especialistas.

Tabela 1.4 Comparação entre os paradigmas de desenvolvimento de SEs

	Sistemas Especialistas Convencionais	Sistemas Especialistas Conexionistas	SEs de Arquitetura Híbrida
Representação de Conhecimento	Regras de Produção Redes Semânticas Frames	Redes Neurais	Frames combinados a Redes Neurais
Reconhecimento padrões	Não	Sim	Sim
Raciocínio lógico	Sim	Não	Sim
Aprendizado Indutivo	Difícil implement.	Fácil implement.	Fácil implement.
Aprendizado Dedutivo	Difícil implement.	Fácil implement.	Fácil implement.
Aquisição de Conhecimento	Não automatizada, toma muito tempo	Totalmente automatizada	Parcialmente automatizada
Explicação do Raciocínio	Fácil obtenção	Difícil obtenção	Fácil obtenção

Quanto à forma de representação de conhecimento, observa-se que a utilização de múltiplos paradigmas pode fornecer aos engenheiros de conhecimento, especialistas e usuários em geral uma visão bem mais ampla do conhecimento do domínio.

As arquiteturas híbridas também possibilitam que o sistema mimetize os 2 tipos de raciocínio presentes nos seres humanos:

- **raciocínio paralelo:** responsável pelas tarefas de reconhecimento de padrões observadas nos cinco sentidos, sobre as quais não há controle consciente. Nos sistemas híbridos, estas tarefas podem ser executadas pelas redes neurais, sem a preocupação de mais tarde apresentar uma

explicação para o método desenvolvido na solução do problema;

- **raciocínio lógico:** processos de raciocínio seqüenciais utilizados em tarefas de resolução de problemas. Estas tarefas são executadas, no cérebro, de forma consciente e controlada, podendo ser efetuadas no sistema híbrido pela unidade simbólica. A análise do processo de raciocínio, que objetiva prover explicações para uma dada consulta, é um procedimento notadamente lógico. Por isso, pode ser executada pela unidade simbólica.

Quanto às formas de aprendizado, os sistemas de arquitetura híbrida possibilitam que aprendizado indutivo e dedutivo sejam implementados. Pode-se implementar aprendizado indutivo através de algoritmos de treinamento das redes neurais por casos e aprendizado dedutivo através da modificação estrutural das redes existentes, utilizando, por exemplo, algoritmos genéticos. Nas arquiteturas convencionais, ambas as formas de aprendizado são de difícil implementação.

O processo de aquisição de conhecimento, *gargalo* no desenvolvimento de sistemas especialistas de arquitetura convencional, pode ser totalmente automatizado em sistemas de arquitetura puramente conexionista, através da criação de redes neurais e seu treinamento por meio de *casos* a tanto destinados. Este mesmo tipo de treinamento também é possível para os sistemas de arquitetura híbrida, permitindo que o processo de aquisição de conhecimento heurístico seja automatizado. Contudo, como parte da base de conhecimento emprega outro mecanismo de representação de conhecimento que não as redes neurais, a construção da porção simbólica da base de conhecimento pode ser apenas guiada pelo sistema, porém não automatizada.

Quanto às facilidades de explanação, os sistemas de arquitetura híbrida, tal como os de arquitetura convencional, possuem mecanismos que facilmente obtêm explicações do processo de raciocínio. Já, nos sistemas de arquitetura puramente conexionista, a obtenção de explicações plausíveis para os problemas solucionados pela rede neural constitui uma tarefa bastante complexa.

1.9.2 Classificação dos sistemas de arquitetura híbrida

A Figura 1.9 apresenta um esquema hierárquico para classificação dos modelos de integração de sistemas especialistas (SE) e de redes neurais (RN) [MED 92].

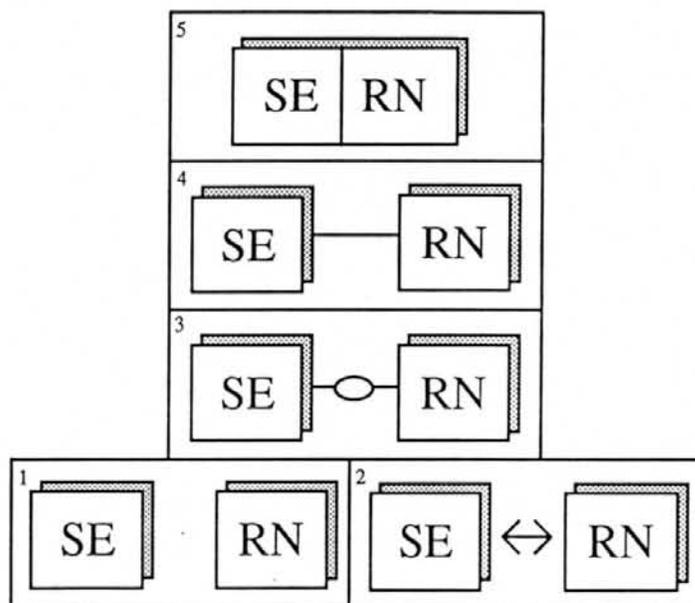


Figura 1.9 Modelos de integração de SEs e RNs.

As estratégias para a integração são as seguintes:

- **arquitetura com componentes independentes:** os modelos baseados em arquiteturas com componentes independentes caracterizam-se por não apresentarem qualquer integração entre os módulos de RNs e de SEs. A principal razão para o emprego de modelos deste tipo é possibilitar a comparação do desempenho na resolução de problemas das duas abordagens para uma aplicação específica. Por exemplo, ao se criar um

sistema híbrido para diagnóstico de falhas, dois módulos distintos (um SE e uma RN) são desenvolvidos para solucionar o mesmo problema de classificação. Quando um componente da máquina apresenta algum problema, os sintomas são apresentados para os 2 sistemas e ambos retornam uma resposta. As soluções são comparadas e o usuário opta por um dos resultados apresentados. Este tipo de modelo não proporciona os benefícios da integração dos dois paradigmas de representação de conhecimento, já que não há interação entre os módulos de SEs e de RNs;

- **modelos transformativos:** nos modelos transformacionais também não existe forma de integração entre os módulos de RNs e de SEs. O que distingue esta abordagem da anterior é que o modelo transformativo começa como um tipo de sistema e termina como outro. As razões para um sistema passar de SE para RN podem incluir a necessidade de assegurar que o sistema apresente características adaptativas e maior velocidade de processamento. Já, o objetivo de um sistema desenvolvido como RN se transformar em SE pode ser o de maior capacidade explanatória, além da representação de conhecimento mais rica. A abordagem dos modelos transformativos tampouco traz os benefícios da integração dos 2 paradigmas de representação de conhecimento;

- **modelos de acoplamento fraco:** este é o primeiro modelo a realmente integrar SEs a RNs. Nele, a aplicação é decomposta em componentes separados de SEs e de RNs, que se comunicam via arquivos de dados. A RN pode assumir tarefas como identificação de padrões regulares, remoção de erros, associação de informações. O SE, por sua vez, pode utilizar as informações vindas da RN para resolver problemas de classificação e apresentar explicações para o raciocínio desenvolvido. Nesta abordagem, como os componentes do modelo híbrido se comunicam para solicitar serviços, obtém-se a maior parte dos benefícios da integração. Algumas limitações que podem ser apontadas nesta abordagem são: redundância na representação em ambos os paradigmas, alto custo de comunicação entre os módulos, difícil manutenção de integridade do conhecimento mantido na RN e no SE devido ao uso de arquivos de dados;

- **modelos fortemente acoplados:** nesta abordagem, a integração entre o SE e a RN se dá via estruturas de dados residentes em memória. Isto aumenta a capacidade de interação dos módulos dos modelos fortemente acoplados em relação aos modelos de acoplamento fraco, além de melhorar seu desempenho. O desenvolvimento e a manutenção de um sistema fortemente acoplado é, no entanto, mais difícil devido à complexidade da interface interna para comunicação. Tal como nos modelos de acoplamento fraco, aqui também ocorre redundância no conhecimento armazenado pelas redes neurais e pelo sistema especialista. Todavia, o modelo oferece flexibilidade e robustez para integração dos dois paradigmas;

- **modelos com integração total:** sua principal característica é a de permitir que seus módulos SE e RN compartilhem estruturas de dados e representação de conhecimento. A comunicação entre ambos é feita por via da natureza dúbia (simbólica/neural) das estruturas. Por exemplo, um nodo da RN pode também representar um conceito simbólico do SE. O raciocínio é, portanto, conduzido de forma cooperativa entre os módulos. Os modelos com integração total apresentam melhor desempenho e menor (ou nenhuma) redundância na representação do conhecimento. Todavia, estes modelos são de desenvolvimento mais complexo e não possibilitam a utilização de pacotes comerciais para a implementação dos módulos de SE e de RN, como é possível em todos os modelos anteriores.

1.9.3 Aplicações de sistemas de arquitetura híbrida

A integração dos paradigmas simbólico e conexionista ainda constitui uma abordagem bastante nova na área de sistemas especialistas. No entanto, alguns sistemas especialistas híbridos já foram implementados com sucesso. As próximas secções apresentam as principais características dos mesmos e suas áreas de aplicação.

1.9.3.1 Avaliação de projetos arquitetônicos

Em [KNA 92] buscou-se representar o conhecimento de especialistas em redes neurais para um problema específico da área de arquitetura. O objetivo inicial era o de criar um SE para dar notas para projetos arquitetônicos. Buscou-se, em uma primeira instância, empregar sistemas baseados em regras. No entanto, constatou-se que, para aquela aplicação, tais sistemas não dispunham dos recursos necessários para representar o conhecimento e conduzir a inferência de forma eficiente. Isto porque os conceitos que deveriam ser manipulados (tais como estética e eficiência) são dificilmente definidos em termos de expressões *booleanas* da lógica proposicional. Então, codificou-se o conhecimento de especialistas em uma rede neural que, guiada por um SE busca solução aproximada para o problema e é capaz de reproduzir o sistema oficial de atribuição de notas para requisitos específicos de projeto. Em outras palavras, o SE encontra uma solução preliminar para o problema e esta é posteriormente ajustada pelas redes neurais.

Neste SE não foram utilizados algoritmos de treinamento para a rede neural, já que os critérios de avaliação representados nesta última são determinados por especialistas e não se podem modificar em função dos trabalhos analisados.

1.9.3.2 Tratamento do melanoma

O melanoma maligno é um tipo de câncer de pele, na maioria dos casos letal, que, a partir de um núcleo escuro, rapidamente forma metástases. Este problema se tem agravado consideravelmente e é cada vez mais notado em áreas do planeta expostas a intensa radiação solar [HUD 92].

O principal objetivo do sistema para tratamento de melanoma não é o de fazer o diagnóstico da doença, mas sim o de estabelecer estratégias de tratamento. Sistemas de apoio à decisão são extremamente úteis para esta aplicação já que há diversos tratamentos para o tumor, porém, apresentam sérios efeitos colaterais.

Um sistema integrando regras a um modelo neural do tipo *feedforward* de 3 camadas foi concebido para o suporte à decisão no tratamento do melanoma.

O objetivo da rede neural é o de usar casos reais (para os quais o tratamento já é conhecido) para determinar os fatores que podem ser utilizados na predição dos tratamentos e a importância relativa de cada fator (fatores de certeza).

A abordagem conexionista é complementada por regras elicítadas de especialistas. A utilização de ambas as abordagens é justificada pelo fato que, para sistemas médicos, é muito importante usar todo o tipo de informação disponível. O modelo híbrido descrito possibilita que uma rede neural, obtida de casos reais, interaja com o conhecimento de especialistas formalizado em regras.

1.9.3.4 Tratamento de efluentes cloacais e industriais

O sistema WATTS - WAsTe Water Treatment System [KRO 92] foi projetado para gerar planos de tratamento de efluentes cloacais.

A maior parte dos efluentes cloacais e industriais contém vários componentes químicos que devem ser removidos antes que a água possa ser liberada para o meio ambiente. Há grande variedade de tratamentos e tecnologias que podem ser invocadas para a redução de um ou mais poluentes. O tratamento dos efluentes depende, então, das concentrações e dos tipos de poluentes existentes no esgoto, na abundância de água e na possibilidade de sua re-utilização.

Uma base de dados para tratamentos foi desenvolvida pelo laboratório Risk Reduction Engineering, em Ohio - EUA, onde são armazenados dados tais como a identificação do composto químico poluente, o tipo de tratamento usado e a concentração do composto poluente na entrada e na saída do processo de tratamento.

O sistema híbrido desenvolvido funciona em 2 fases principais:

- **fase de análise** : consiste em um sistema de aquisição de conhecimento composto por um sistema de aprendizado, que extrai regras da base de dados, e por uma representação de conhecimento baseada em gramática, que integra conhecimento oriundo de diferentes fontes;
- **fase de síntese**: gera planos de tratamento utilizando as regras de tratamento, extraídas na fase de análise. Duas abordagens são possíveis para o problema: na primeira, a fase de síntese é formulada como um problema de busca e uma nova função de busca heurística é desenvolvida para obter o plano de tratamento. Na segunda abordagem, a fase de síntese é formulada como um problema de otimização. Um modelo neural baseado no modelo de Hopfield foi desenvolvido para gerar o plano de tratamento, buscando minimizar o custo total do tratamento que deve fazer com que a concentração do poluente na saída seja menor que um limiar pré-estabelecido.

Qualquer das duas abordagens de síntese pode ser empregada, sendo o enfoque heurístico o mais rápido nos problemas de pequeno porte. Para um número de componentes maior ou concentração de poluentes muito alta, a rede neural produz respostas rápidas com os coeficientes para o tratamento apropriados.

1.10 NEXTOOL: Ferramenta para SEs de Arquitetura Híbrida

NEXTOOL [MAC 91b] é um sistema para desenvolvimento e operação de sistemas especialistas classificatórios que representa conhecimento através de um esquema híbrido, integrando redes semânticas às redes neurais [MAC 92a][MAC 92b]. A Figura 3.1 ilustra a arquitetura básica destes sistemas especialistas.

Nesta arquitetura, a Base de Conhecimento dos Sistemas Especialistas convencionais foi substituída por uma Base de Conhecimento Conexionista, formada por uma rede neural e uma rede semântica interligadas.

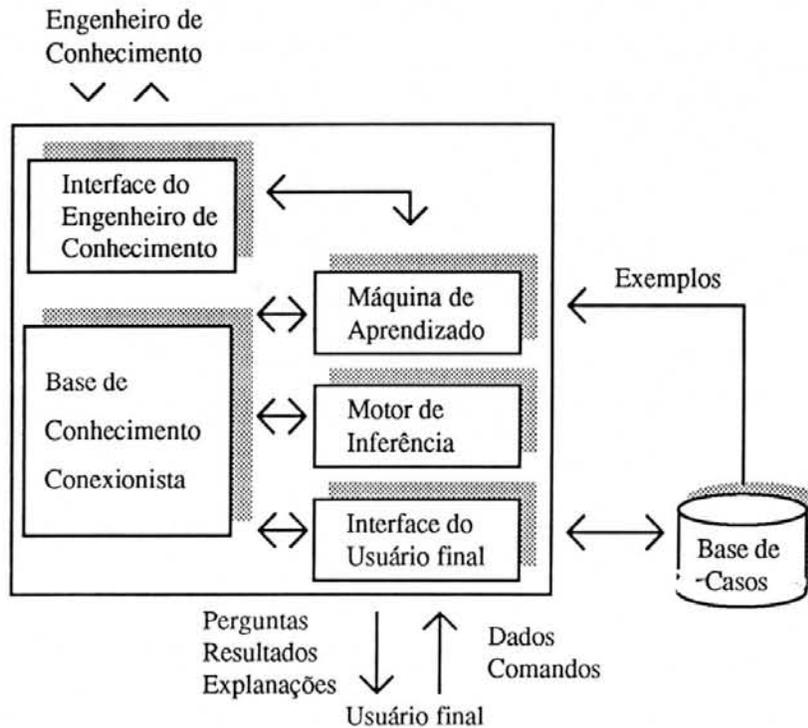


Figura 3.1 Arquitetura básica de um sistema empregando NEXTOOL

A rede semântica é usada para representar os objetos e relacionamentos do domínio do problema. Dois níveis foram definidos para a rede:

- *Intensional Semantic Network (ISN)* : composta por classes de objetos e um conjunto de relacionamentos primitivos entre as classes;
- *Extensional Semantic Network (ESN)*: composta por instâncias das classes e relacionamentos da ISN.

A Figura 3.2, extraída de [MAC 92b], apresenta um fragmento de rede, mostrando os dois níveis acima referidos.

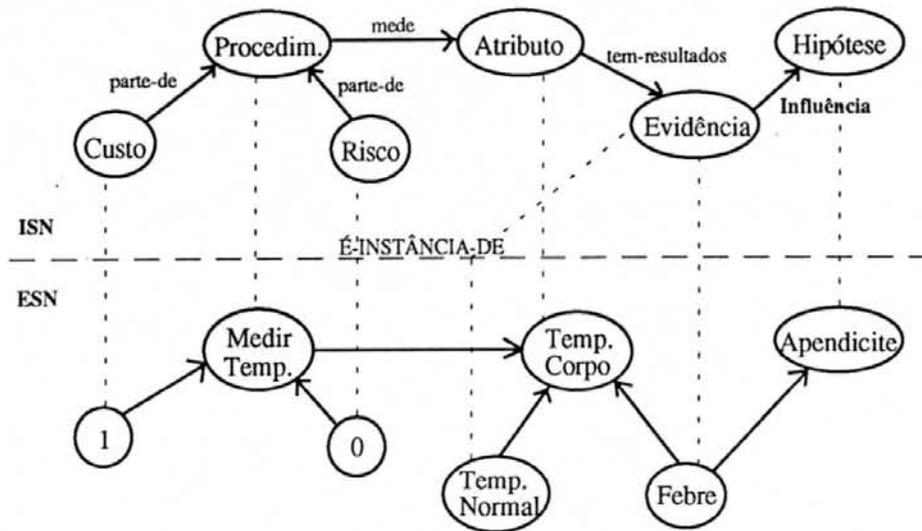


Figura 3.2 Rede Semântica ISN e ESN

Neste exemplo, são evidenciados alguns aspectos importantes, a saber:

- **hipóteses:** representam as categorias dos problemas de classificação;
- **atributos:** armazenam características relevantes do problema em análise;
- **evidências:** apresentam possíveis resultados para atributos;
- **procedimentos:** representam as tarefas através das quais obtêm-se valores para determinados atributos.

Estes conceitos são ligados através dos seguintes arcos que expressam os relacionamentos entre o domínio:

- **mede:** liga um atributo ao seu procedimento de obtenção de valores;
- **tem-resultados:** conecta um atributo aos seus possíveis valores;

- **influência:** permite que seja extraída a inferência a partir de evidências.

As porções de ambas as redes, ligadas por arcos *influência* é chamada de Rede de Inferência. Cada objeto desta rede pode ser visto como uma proposição sobre o problema em análise e terá a ele associado um neurônio da rede neural que representará o grau de possibilidade daquela evidência ou hipótese ser verdadeira. Por exemplo, "a possibilidade do paciente ter apendicite é de 0,7", indica que o neurônio associado ao nodo *Apendicite* foi ativado com força igual a 0,7. Esta é portanto a forma de conectar a porção simbólica do sistema à sua porção neural.

Os arcos *é-instância-de*, *é-um* e *parte-de* têm função de relacionar instâncias da ESN às suas classes da ISN, ligar um conceito à sua generalização e relacionar um objeto às partes que o compõem, respectivamente.

A rede neural utilizada é baseada no **Modelo Neural Combinatório (MNC)**. As redes neurais resolvem os problemas de classificação, buscando o grau de ativação de cada instância de evidência e calculam, a partir destes, a ativação de todas as hipóteses relacionadas àquelas evidências. A hipótese com grau de ativação mais alto é, então, selecionada.

O motor de inferência permite a formulação de consultas ao sistema especialista. O processo de consulta é dividido em uma fase passiva e outra ativa. Na passiva, o usuário introduz no sistema todas as evidências observadas e os graus de possibilidade associados a elas. Na fase ativa, o sistema faz perguntas ao usuário, buscando confirmar uma das hipóteses disparadas pelas evidências apresentadas.

A Máquina de Aprendizado tem o propósito de refinar o conhecimento armazenado na Base de Conhecimento Conexionista, através da análise de conjuntos de exemplos contidos na **base de casos**. O processo de aprendizado pode ocorrer pelo ajuste de pesos das redes neurais usando o algoritmo de punição e recompensa do MNC (aprendizado indutivo) ou, ainda, através da mudança da topologia das redes (aprendizado dedutivo). Esta mudança de topologia pode ocorrer pela aplicação de operadores genéticos para criar novos elementos nas camadas

intermediárias das redes, completando lacunas que havia no conhecimento inicialmente introduzido.

NEXTOOL também fornece mecanismos para automatização do processo de aquisição de conhecimento a partir de uma base de casos. Fundamentado na **base de casos**, a máquina de aprendizado seleciona as combinações de evidências apropriadas e, através dos algoritmos genéticos, gera combinações não existentes. Um poucas iterações são necessárias (talvez uma única) para treinar a rede e torná-la operante com base em um bom conjunto de exemplos.

2 OBJETIVOS

O presente trabalho tem como objetivos:

- 2.1 definir um modelo híbrido para sistemas especialistas, que integre os paradigmas de representação de conhecimento simbólico e conexionista;
- 2.2 projetar um mecanismo de integração que permita ao paradigma simbólico representar o conhecimento armazenado na porção conexionista;
- 2.3 definir métodos de aquisição automática do conhecimento a partir de uma base de casos e incorporar os processos de aprendizado indutivo e dedutivo à base de conhecimento híbrida que possibilitem o refinamento da base de conhecimento;
- 2.4 implementar uma ferramenta para construção e operação de sistemas especialistas conexionistas híbridos com as características acima descritas;
- 2.5 validar esta ferramenta para o domínio de diagnóstico de cardiopatias congênitas.

3 METODOLOGIA

3.1 Introdução

Este capítulo descreve o modelo conexionista híbrido HYCONES desenvolvido neste trabalho, que enseja o projeto de sistemas especialistas classificatórios, integrando o formalismo de representação de conhecimento de *frames* ao modelo conexionista MNC. A Figura 3.3 ilustra a arquitetura básica do modelo conexionista híbrido para SEs aqui proposto.

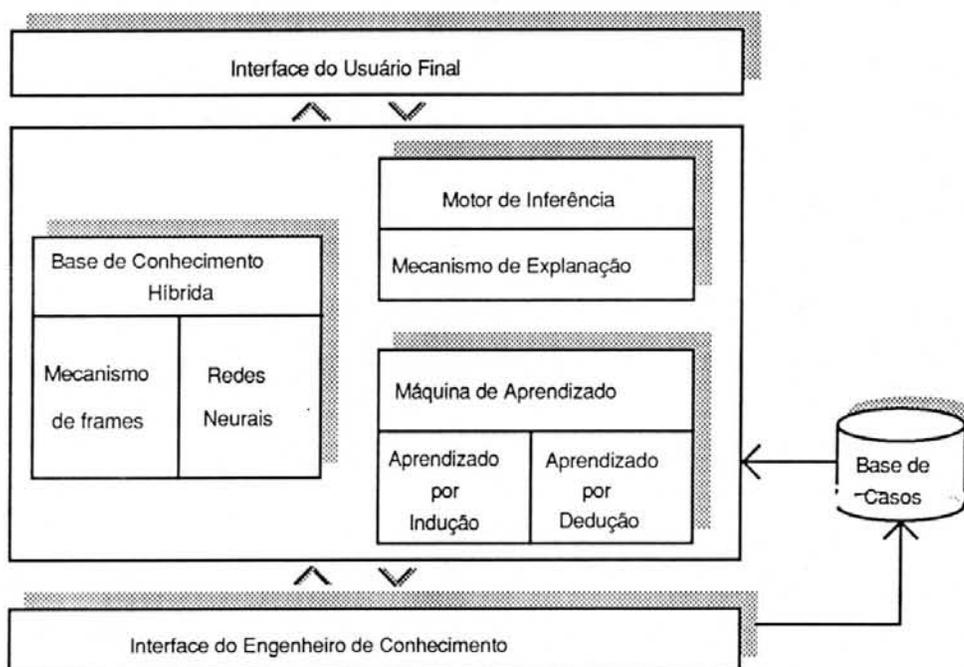


Figura 3.3 Arquitetura básica do modelo conexionista híbrido HYCONES

3.2 A Base de Conhecimento Híbrida

A base de conhecimento híbrida emprega duas formas de representação de conhecimento: *frames* e redes neurais.

A modelagem de conhecimento baseada no mecanismo de representação de *frames* faz-se em dois níveis: nível de **diagnósticos** e nível de **achados**. Ao nível de **achados**, descrevem-se os objetos elementares da aplicação que influenciam a detecção de certos diagnósticos. Ao nível de **diagnósticos**, descrevem-se os problemas de classificação que o sistema objetiva solucionar.

Os objetos armazenados no nível de achados são agrupados em classes (*classificação*). Estas classes são reunidas em grupos mais genéricos, definidos de acordo com os conceitos de abstração de dados de *generalização*, *agregação* e *associação* [HUL 87].

A *classificação* é alcançada a partir do agrupamento de objetos com propriedades comuns em um novo objeto, para os quais todas as condições uniformes são válidas. Em outras palavras, a *classificação* consiste na definição de um objeto composto, chamado classe, com base em um conjunto de objetos simples com as mesmas propriedades, no caso chamados de instâncias.

O conceito de *generalização* é, na realidade, complementar ao de *classificação*. Na *generalização*, as propriedades comuns a várias classes são abstraídas para uma superclasse (hierarquicamente superior). Deste modo, é possível referenciar um conjunto de classes através de um elemento mais genérico. Definindo-se objetos mais complexos (superclasses) em função de objetos menos complexos (subclasses), estabelece-se uma hierarquia de classes do tipo *É-um*.

A *agregação* permite tratar objetos não como entidades atômicas, mas sim como uma composição de outros objetos. O conceito de *agregação*, desta forma, corresponde à composição de objetos, expressando que objetos mais simples são *Parte-de* objetos mais complexos.

A *associação* permite descrever propriedades de um conjunto de objetos como um todo. As propriedades assim definidas são, então, calculadas a partir daquelas propriedades já definidas para os elementos do conjunto.

Cada achado, mapeado em uma classe do nível de achados, aponta para um ou mais diagnósticos do nível de **diagnósticos**. A Figura 3.4 apresenta um exemplo de modelagem de problemas classificatórios, utilizando os conceitos de abstração de dados.

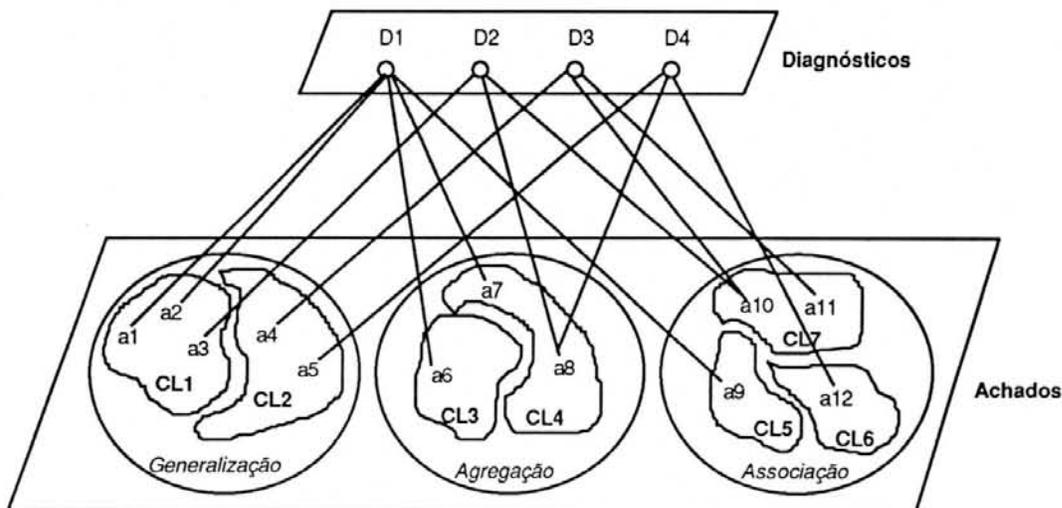


Figura 3.4 Modelagem de problemas classificatórios através dos conceitos de abstração

No exemplo, descrevem-se 12 achados elementares (a_1, \dots, a_{12}), estruturados em 7 classes distintas (CL_1, \dots, CL_7). Estas classes são reunidas em 3 diferentes grupos, um representando uma *generalização*, outro representando uma *agregação* e o terceiro representando uma *associação*. Os achados a_1, a_2, a_6, a_7 e a_9 são ligados ao diagnóstico D1, indicando sua influência na detecção do mesmo. Os achados a_3, a_8 e a_{10} influenciam o diagnóstico D2; a_4, a_{10} e a_{11} influenciam D3; finalmente, a_5, a_8 e a_{12} influenciam D4.

A representação de conhecimento baseada nos conceitos de abstração possibilita retratar diferentes tipos de relacionamentos entre as porções de conhecimento da BCH. Os tipos construtores, oferecidos pelos conceitos de abstração, têm grande potencial de representação semântica, possibilitando representação mais fiel da realidade e permitindo a utilização do modelo em diferentes domínios de aplicação.

Os níveis de **diagnóstico** e de **achados** são representados na BCH por dois tipos distintos de *frames*: *frames de diagnóstico* e *frames de achados*.

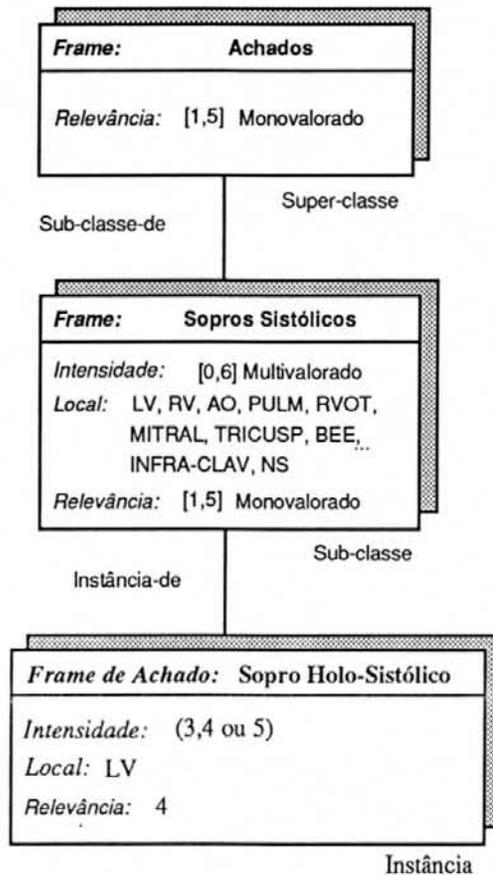


Figura 3.5 Frames de achado

Os *frames de achado* são descritos por um conjunto específico de atributos, definidos de acordo com o conhecimento do domínio. Estes atributos podem ser estabelecidos localmente, ou herdados de *frames* hierarquicamente superiores. Apresenta-se, a seguir, um exemplo de parte de uma hierarquia de *frames de achados*, definida em uma base de conhecimento para o diagnóstico de cardiopatias congênitas (Figura 3.5). No exemplo, define-se o atributo *relevância* na classe dos *frames de achados*, para representar a importância deste último em relação a todos os outros achados da base de conhecimento. Este atributo é herdado para a subclasse de nome *Sopros Sistólicos*, onde são definidos outros dois atributos: *intensidade* e *local*. O *frame de achado Sopro holo-sistólico* é definido como uma

instância do frame *Sopros sistólicos*. Desta forma, *Sopro holo-sistólico* herda as definições dos atributos de *Sopros sistólicos* e recebe os valores (3,4 ou 5) para o atributo multivalorado *intensidade*, LV para o atributo *local* e 4 para o atributo *relevância*.

Os *frames de diagnóstico* são utilizados para armazenar as descrições do conhecimento necessário para resolver o problema diagnóstico da aplicação. Para tanto, um *frame de diagnóstico* armazena referências a conjuntos de achados que, quando presentes, influenciam em sua detecção.

São 4 os tipos de *slots* empregados na descrição de um *frame de diagnóstico*:

- **gatilho:** referencia um *frame de achado* que, quando presente, aponta o *frame de diagnóstico* como solução potencial para o caso apresentado ao SE;
- **achados-essenciais:** contém uma lista de *frames de achados* que, quando presentes, asseguram a identificação do diagnóstico representado pelo *frame de diagnóstico*;
- **achados-complementares:** contém uma lista de *frames de achado* que pode auxiliar na confirmação da hipótese segundo a qual o diagnóstico representado pelo *frame de diagnóstico* é a solução do problema;
- **achados-de-negação:** referencia uma lista de *frames de achado* que, quando presente, pode eliminar o *frame de diagnóstico* do conjunto de possíveis soluções do caso apresentado ao SE.

A Figura 3.6 apresenta um *frame de diagnóstico* usado na descrição do diagnóstico de *Comunicação Interventricular (CIV)*. Neste *frame*, está definido como *gatilho* o *Sopro holo-sistólico*, e como *achados-essenciais*: *hiperfluxo*, *crescimento biventricular* e *cardiomegalia não especificada*. Os achados *Infecção respiratória de repetição* e *B2-hiperfonética* são definidos como *achados-complementares*.

<i>Frame de diagnóstico:</i>		CIV
<i>Gatilho:</i>	S.HOLO-SISTÓLICO	
<i>Achados</i>	S.HOLO-SISTÓLICO	
	HIPERFLUXO	
<i>Essenciais:</i>	CRESCIM.BIVENTRIC.	
	CARDIOMEG.N.ESPEC.	
<i>Achados</i>	INFEC.RESP.REPETIÇÃO	
<i>Complementares:</i>	B2-HIPERFONÉTICA	

Figura 3.6 *Frames de diagnóstico*

A ligação entre a hierarquia de *frames de achados* e os *frames de diagnósticos* faz-se, portanto, através das referências aos *frames de achados* existentes nos *frames de diagnósticos*. A Figura 3.7 apresenta parte de uma hierarquia de achados médicos e sua ligação com o diagnóstico de CIV. No exemplo, pode-se observar como as instâncias *sopro holosistólico*, *hiperfluxo*, *crescimento biventricular*, *cardiomegalia não especificada*, *infecção respiratória de repetição* e *B2-hiperfonética* da hierarquia de frames de achados influenciam a detecção do diagnóstico de CIV (Comunicação Interventricular).

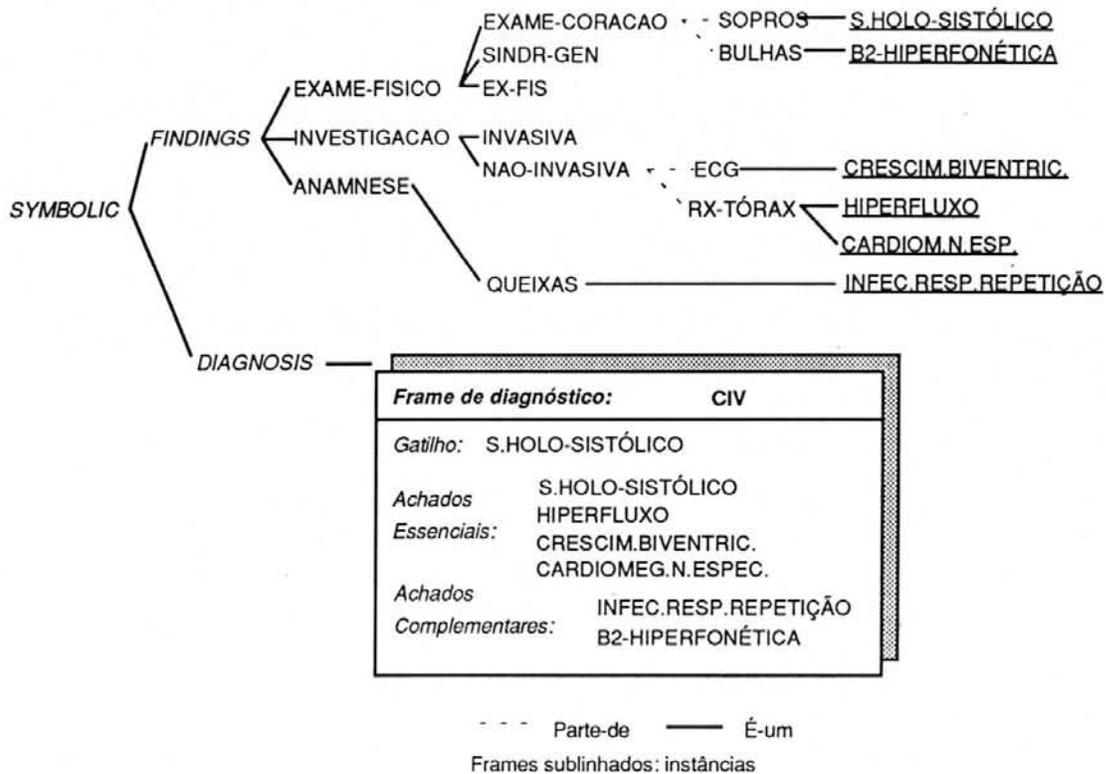


Figura 3.7 Hierarquia de frames achado e diagnósticos

Através da divisão dos *frames* em *frames de diagnóstico* e *frames de achados*, o formalismo de representação de conhecimento simbólico abrange completamente todo o espectro da aplicação. A hierarquia de achados descreve os aspectos declarativos do conhecimento do domínio, enquanto que os *frames de diagnóstico* mantêm as possíveis soluções para os problemas tratados.

3.2.1 A integração das redes neurais MNC ao paradigma de *frames*

O mesmo conhecimento descrito nos *frames de diagnóstico* é também representado nas redes neurais do sistema. O modelo conexionista adotado é o Modelo Neural Combinatório (MNC), descrito no capítulo 1.

A correspondência entre as estruturas do MNC e do mecanismo de *frames* é baseada no mapeamento dos grafos de conhecimento para as redes neurais e para os *frames de diagnóstico*. A 3.8 apresenta um grafo de conhecimento para o

diagnóstico do Defeito Septal Atrioventricular (DSAV) e a sua tradução para uma rede neural do tipo MNC e para um *frame de diagnóstico*.

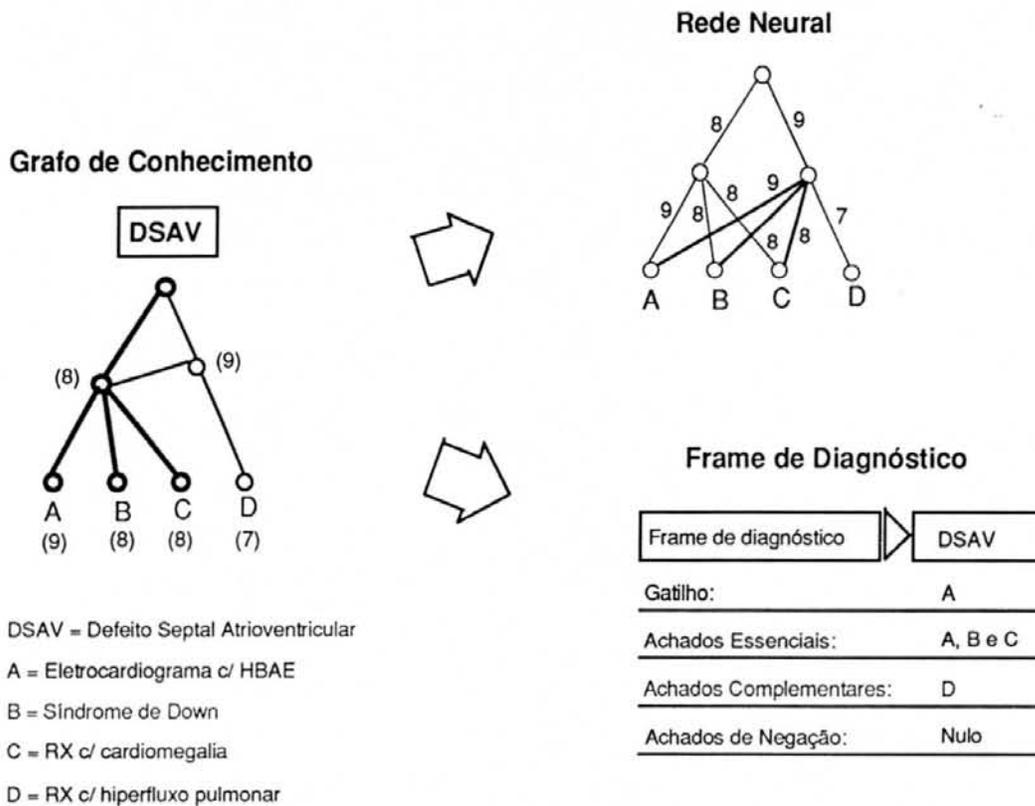


Figura 3.8 Mapeamento de um GC para a rede neural e para o *frame diagnóstico*

A tradução do grafo de conhecimento para a rede neural é direta. Cada caminho do grafo de conhecimento, que leva até a hipótese de diagnóstico, é representado por um caminho semelhante na rede neural. Os pesos correspondentes aos graus de importância de cada um dos nodos do grafo são também diretamente mapeados para os pesos de conexão da rede neural.

Já a tradução do grafo de conhecimento para os frames de diagnóstico é baseada na seguinte estratégia:

- o nodo mais à esquerda do grafo de conhecimento é tomado como o *gatilho* do *frame de diagnóstico* (nodo A da Figura 3.8);

- os nodos que compõem a sub-árvore mais à esquerda do grafo de conhecimento são tomados como os *achados-essenciais* do *frame de diagnóstico* correspondente. Os achados representados por estes nodos estão, na realidade, presentes em todos os caminhos que levam ao diagnóstico (nodos A, B e C da Figura 3.8);
- os outros nodos do grafo de conhecimento que formam sub-árvores que alcançam a hipótese de diagnóstico são considerados *achados -complementares* (nodo D da Figura 3.8).

Os *achados de negação* devem ser explicitamente apontados nos grafos de conhecimento, para que possam ser traduzidos para os *frames de diagnóstico* e para as redes neurais (na forma de ligações com pesos de conexão negativos).

O mapeamento entre redes neurais e *frames de diagnósticos* define um mecanismo de compartilhamento de estruturas de memória, para representar conceitos dos componentes simbólico e conexionista (característica dos modelos de integração total): os *achados* e *diagnósticos* referenciados nos nodos de evidências e de hipóteses da rede neural são, na realidade, os *frames de achados* e de *diagnósticos* da hierarquia de *frames*.

3.3 O Motor de Inferência

A função do motor de inferência é a de resolver problemas através da análise do conhecimento armazenado na base de conhecimento. Como o conhecimento dos sistemas especialistas híbridos aqui tratados está armazenado na forma de *frames* e de redes neurais, o motor de inferência do sistema deve operar sobre ambas as formas de representação de conhecimento.

A inferência para o modelo aqui descrito baseia-se no mecanismo de reconhecimento de padrões, implementado pelo modo de consulta do MNC. Um

conjunto de evidências é apresentado às redes neurais as quais, ativadas, passam a apresentar sinais de saída para cada um dos diagnósticos tratados pelo sistema.

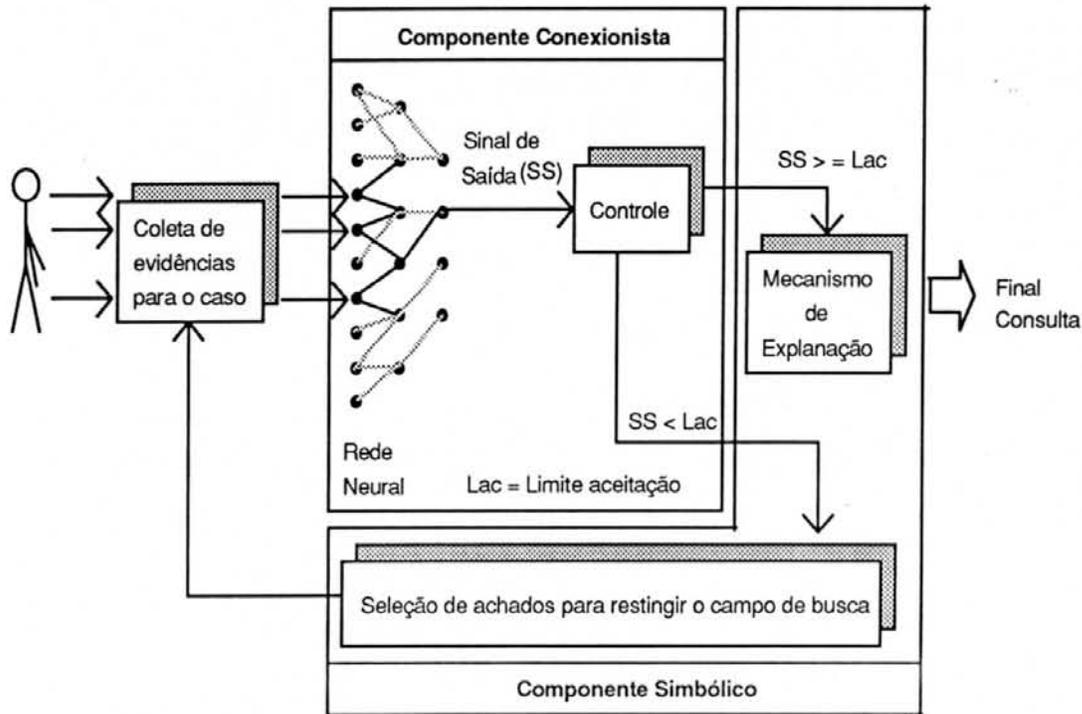


Figura 3.9 Funcionamento do motor de inferência de HYCONES

Se algum caminho da rede apresentar sinal de saída maior que um limiar de aceitação previamente estabelecido, o sistema apresenta o diagnóstico apontado pela rede neural como solução para o problema e uma explicação para o raciocínio. Caso contrário, o sistema passa a trabalhar com o formalismo simbólico de representação, na tentativa de restringir o campo de busca e guiar a subsequente coleta de informações. Para tanto, o sistema identifica, nos *frames de diagnóstico*, todos aqueles cujo gatilho corresponde a alguma evidência apresentada pelo usuário. Identificados os frames, os achados ainda não apontados para o caso e presentes nos *slots achados-essencias* e *achados-complementares* são selecionados. Estes achados são apresentados ao usuário como uma lista de elementos que pode auxiliar o sistema a encontrar a solução do problema. Após esta nova coleta de informações, o sistema volta a ativar as redes neurais. Este processo prossegue até que o sinal de saída apresentado por uma rede neural seja maior que o limiar de aceitação ou até

que o usuário, em uma das etapas de coleta de informações, decida interromper a inferência. A Figura 3.9 apresenta o esquema de funcionamento do motor de inferência.

A execução de tarefas de forma integrada entre os componentes conexionista e simbólico no motor de inferência é uma característica dos *Modelos híbridos de conexão total*.

3.4 O Mecanismo de Explicação

No modelo proposto, as explicações são obtidas pelo componente simbólico do sistema. Como o conhecimento da rede neural é simbolicamente representado pelos *frames de diagnóstico*, a inferência executada pela rede neural pode ser explicada através da análise destes *frames*.

Existem diferentes formas de se conduzir a explicação do raciocínio de um sistema especialista, tais como a explicação baseada em modelos causais, a invocação da racionalização, o recurso à clarificação [ROL 88].

O mecanismo de explicação projetado para este modelo híbrido emprega o modelo de explicação por racionalização, ou seja, o sistema busca convencer o usuário da validade da conclusão apresentada.

O principal meio de obter explicações em sistemas especialistas baseia-se no armazenamento de justificativas para cada passo adotado pelo sistema. Todavia, como o raciocínio é aqui desenvolvido pelo componente conexionista, o armazenamento de justificativas durante o processo de inferência se torna inviável.

Por isso, adotou-se um mecanismo de explicação baseado na classificação dos elementos dos *frames de diagnóstico* em *gatilho*, *achados-essenciais*, *achados-complementares* e *achados de negação*. Estes elementos podem ser vistos como regras compostas da seguinte forma:

Se
o *gatilho* do frame de diagnóstico estiver presente,
Então,
o diagnóstico é a provável solução para o problema.

Se
todos os *achados-essenciais* estiverem presentes,
Então
o diagnóstico é solução.

Se
algum *achado complementar* estiver presente,
Então,
ratifica-se a hipótese de que o diagnóstico é a solução.

Se
algum dos *achados de negação* estiver presente,
Então
o diagnóstico possivelmente não é a solução para o problema.

A classificação das regras em quatro tipos possibilita que seja elaborada uma explicação com base no papel de cada um dos elementos que aparece na parte esquerda da regra.

A conclusão apresentada é, portanto, obtida através da análise dos *frames de diagnóstico* que contém representações simbólicas do conhecimento armazenado nas redes neurais.

3.5 A Máquina de Aprendizado

A Máquina de Aprendizado tem o propósito de adquirir conhecimento ou refinar o conhecimento armazenado na Base de Conhecimento Híbrida, através da análise de conjuntos de exemplos contidos na Base de Casos.

A máquina de aprendizado, na arquitetura aqui proposta, é dividida em 2 componentes principais, o primeiro responsável pelo aprendizado indutivo e o segundo responsável pelo aprendizado dedutivo.

3.5.1 Aprendizado indutivo

A máquina de aprendizado indutivo possibilita dois tipos de aprendizado para as redes neurais: explicitação automática de conhecimento por casos de treinamento e aprendizado incremental.

A explicitação automática de conhecimento visa sanar o problema considerado gargalo no desenvolvimento de SEs, ou seja, a aquisição de conhecimento para a construção da base de conhecimento. Já o aprendizado incremental visa dotar o sistema da capacidade de aprendizado observada nos especialistas humanos, i.e., aprender pela experiência.

A explicitação automática de conhecimento a partir da base de casos é executada em três fases principais:

- criação de uma rede neural sem conhecimento prévio, com topologia baseada no MNC;
- treinamento da rede pelo algoritmo de punição e recompensa;
- normalização e eliminação de células não operantes da rede, através do algoritmo de poda e normalização.

O método para criação da rede neural baseia-se na análise dos casos da base de casos para gerar, já em primeira instância, uma rede com certo conhecimento. Nesta rede, somente são ligados aos nodos de hipótese os achados observados para aquele diagnóstico em algum caso. Estes achados são combinados em grupos de dois e três elementos, de modo a esgotar as possibilidades de combinações entre eles. A Figura 3.10 ilustra a criação das redes neurais "vazias" a partir de casos de treinamento.

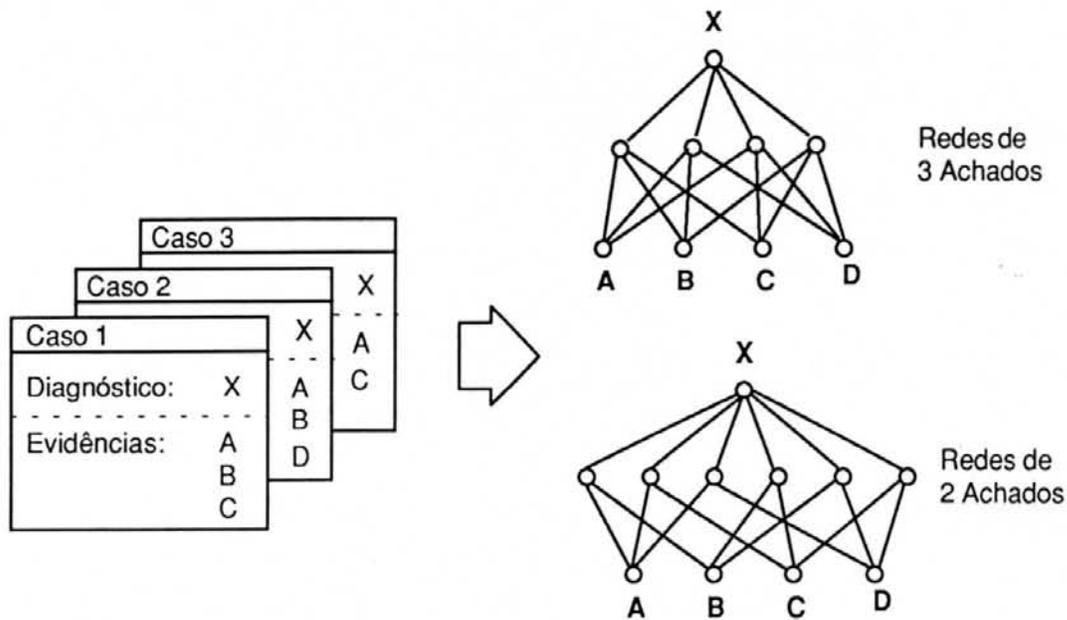


Figura 3.10 Criação das redes neurais por casos de treinamento

No método para construção da rede, originalmente proposto para o MNC [MAC 91a], a geração da rede neural *vazia* se dá através do agrupamento de todos os achados da base de conhecimento em nodos combinatórios de m elementos, $m-1$ elementos, ..., e 1 elemento, para todos os diagnósticos, sendo m um número da ordem de 3 ou 4. Todavia, mesmo gerando nodos intermediários que agrupam um número limitado de elementos, as redes neurais geradas por este método são extremamente grandes e contêm excesso de informação irrelevante, já que reúnem, em nodos intermediários, achados que não são importantes para o diagnóstico. Por isso, adotou-se a seguinte estratégia para geração das redes neurais não treinadas, para cada um dos diagnósticos proceder como segue:

- calcular a frequência com que cada achado contribui para o diagnóstico;
- selecionar os achados mais frequentes e combiná-los em nodos intermediários da rede neural de dois e três elementos, de modo a esgotar as possibilidades de combinações entre eles;
- criar acumuladores de punição e de recompensa para cada conexão da rede, colocando-os em 0;

- colocar todos os pesos de conexão da sub-rede em 1.

Após sua geração, a rede neural *vazia* é treinada a partir dos casos armazenados na *base de casos*. Este treinamento consiste na modificação dos pesos de cada conexão da rede neural, segundo o seguinte critério: *recompensar* os ramos da rede que levam aos diagnósticos corretos e *punir* os ramos da rede que levam aos diagnósticos incorretos. O treinamento da rede neural tem início através da aplicação do *Algoritmo de punição e recompensa - versão de arranque*, definido para o MNC [MAC 91a]. Este algoritmo produz uma rede treinada não-operante, contendo acumuladores com valores entre $-n$ e $+n$, sendo n o número de casos. O algoritmo, que utiliza uma versão simples do mecanismo de *backpropagation* [RUM 86], é apresentado a seguir. Para cada exemplo da base de casos de treinamento, adotar o seguinte procedimento:

- propagar os sinais de entrada até a camada de hipóteses;
- para cada arco que alcance algum nodo de hipótese, fazer:

Se o nodo de hipótese alcançado corresponde a um resultado correto,

Então, propagar retroativamente, a partir deste nó até os nós de entrada, um acréscimo no valor do acumulador de recompensas, de acordo com a fórmula:

$$FE * AD * IE$$

onde:

FE = fluxo evidencial,

AD = a ativação do neurônio de destino,

IE = a importância do exemplo para o treinamento.

Senão, propagar retroativamente um acréscimo no valor do acumulador de punições deste nodo até os nodos de entrada, de acordo com a fórmula:

$$FE * AD * IE$$

onde:

FE = fluxo evidencial,

AD = a ativação do neurônio de destino,

IE = a importância do exemplo para o treinamento.

Para transformar a versão treinada e não-operante da rede neural em versão operante, aplica-se o algoritmo de *poda e normalização* à rede [MAC 91a]. Este algoritmo tem o objetivo de transformar os acumuladores de punição e recompensa da rede neural em pesos de conexão no intervalo [0,1]. O algoritmo de *poda e normalização* é apresentado a seguir. Para cada arco da rede, proceder como segue:

- Calcular o valor líquido dos acumuladores (**ACUMLIQ**) pela fórmula:

$$\text{ACUMLIQ} = \text{acumulador de recompensas} - \text{acumulador de punições}$$

- Se **ACUMLIQ** ≤ 0 ,

Então,

Remover o arco da rede

Senão,

Se acumulador de punições do arco > 0 ,

Então,

Computar o peso do arco como:

$$\text{ACUMLIQ} / \text{MAXLIQ}$$

onde:

MAXLIQ = ao maior **ACUMLIQ** da subrede da hipótese à qual o arco pertence

Senão,

Computar o peso do arco como:

$$\text{SQRT}(\text{Laceit}) + (1 - (\text{SQRT}(\text{Laceit}))) * \text{ACUMLIQ} / \text{MAXLIQ}$$

onde:

Laceit = limiar de aceitação, pré-estabelecido pelo engenheiro de conhecimento.

- **Se** o peso do arco < Limiar de poda,

Então,

Remover o arco da rede.

onde:

Limiar de poda = valor pré-estabelecido pelo engenheiro de conhecimento.

- Remover todos os neurônios das camadas de entrada e combinatória que percam conectividade com os nós de hipótese.

O componente da máquina de aprendizado responsável pelo aprendizado incremental faz uso do algoritmo de *punição e recompensa - versão de refinamento*, para mimetizar a capacidade de aprendizado por experiência dos seres humanos [MAC 91a]. Este algoritmo também produz uma rede treinada não operante contendo acumuladores com valores entre -n e n, para n o número de casos. O princípio dos dois algoritmos é o mesmo: recompensar os ramos da rede que levam ao diagnóstico correto e punir os ramos da rede que levam a diagnósticos incorretos. A versão de refinamento do algoritmo é apresentada a seguir.

Para cada exemplo da base de casos de treinamento fazer:

- **Passo 1** Propagar os sinais de entrada até a camada de hipóteses,

- **Passo 2** Para cada arco que alcance algum nodo de hipótese H, fazer:

Se o nodo de hipótese alcançado corresponde a um resultado correto,

Então,

Calcular $a = 1 - \text{ativação (H)}$,

propagar retroativamente, a partir deste nó, até os nós de entrada, um acréscimo no valor do acumulador de recompensas, de acordo com a fórmula

$$a * FE * AD * IE$$

onde:

FE = fluxo evidencial,

AD = ativação do neurônio de destino,

IE = importância do exemplo para o treinamento;

Senão,

Calcular $a = \text{ativação (H)}$,

propagar retroativamente um acréscimo no valor do acumulador de punições deste nodo, até os nodos de entrada, de acordo com a fórmula

$$a * FE * AD * IE$$

onde:

FE = fluxo evidencial,

AD = ativação do neurônio de destino,

IE = importância do exemplo para o treinamento;

Para transformar a rede treinada pelo algoritmo de *punição e recompensa - versão de refinamento*, em uma rede operante, aplica-se o mesmo algoritmo de *poda e normalização*, anteriormente apresentado, com o objetivo de transformar os

acumuladores de punição e de recompensa em pesos de conexão e de remover conexões de peso muito baixo.

Desta forma, a máquina de aprendizado refina o conhecimento das redes neurais. Como o sistema deve garantir que o conhecimento dos *frames de diagnóstico* representem simbolicamente o conhecimento das redes neurais, qualquer modificação na estrutura de uma rede deve disparar um processo para que o *frame de diagnóstico* também seja modificado (característica dos *modelos fortemente acoplados*). A manutenção de consistência é efetuada do seguinte modo:

- O modelo se inspira no fator de *sensibilidade* para determinar os *achados-essenciais*. A noção de sensibilidade mede a probabilidade de que um determinado achado (f) esteja presente, num paciente com um determinado diagnóstico (D). A sensibilidade de um achado, em relação a um diagnóstico, corresponde, portanto, à frequência com que o achado (f) é identificado para aquele diagnóstico (D). A sensibilidade do achado f, em relação ao diagnóstico D, é obtida através da fórmula [OWE 90]:

$$\text{Sensibilidade} = \frac{\text{Número de casos, onde um determinado achado f está presente no diagnóstico D}}{\text{Número total de casos com o diagnóstico D}}$$

Estabeleceu-se um limiar de aceitação *L-essen* para que um achado fosse tomado como essencial para um diagnóstico. Este limiar está no intervalo [0,1] e indica a frequência com que um achado deve aparecer nos casos de treinamento do mesmo diagnóstico, para que seja considerado essencial.

- São *achados-complementares* aqueles que compõem algum caminho que conduz a hipótese de diagnóstico, mas que apareçam com frequência menor que *L-essen* nos casos de treinamento;

- os achados de negação são aqueles cujo somatório de conexões negativas ultrapasse o somatório de conexões positivas para o diagnóstico.

A identificação do *gatilho* de um *frame de diagnóstico* é feita após a determinação de seus *achados-essenciais*. O *gatilho* do *frame* indica qual destes achados é o mais "importante" no processo de identificação do diagnóstico. O modelo se inspira no fator de *especificidade* para determinar a importância dos achados para os diagnósticos. A *especificidade* de um achado (f) em relação a um diagnóstico (D) indica quão específico o achado (f) é para o referido diagnóstico (D), ou ainda, a possibilidade do achado (f) não estar presente quando da detecção de outros diagnósticos (DO) .

A *especificidade* do achado f, em relação ao diagnóstico D, é obtida através da fórmula:

$$\text{Especificidade} = \frac{\text{Número de casos sem o achado f com diagnósticos diferentes de D}}{\text{Número total de casos com diagnósticos diferentes de D}}$$

Dado que o *gatilho* de um *frame de diagnóstico* é selecionado de sua lista de *achados-essenciais*, ele consistirá no mais específico dos achados de maior sensibilidade para o diagnóstico.

Contudo, é possível que a *especificidade* de vários achados seja a mesma, para um determinado diagnóstico. Neste caso, emprega-se outro fator na determinação do *gatilho* do *frame de diagnóstico*: a **relevância**. Esta propriedade é definida para cada achado da base de conhecimento e determina quão significativo um achado é em relação a **todos os demais achados da base**. Para o médico, o conceito de relevância pode ser compreendido como a morbidade de um achado, ou a sua gravidade. Por exemplo, no diagnóstico de problemas cardíacos, o sintoma de dor no peito tem maior morbidade do que o sintoma *coceira* no pé, ou prurido. Por este motivo, quando vários achados apresentam especificidade semelhante, a propriedade **relevância** é utilizada para determinar o *gatilho* do *frame de diagnóstico*.

O modelo emprega a mesma escala de fatores de relevância utilizada pelo sistema INTERNIST-I [MIL 86]. A Tabela 3.1 mostra a interpretação dos fatores de relevância empregados.

Tabela 3.1 Fatores de relevância, conforme definidos no sistema INTERNIST

Relevância	Interpretação
1	A manifestação não é muito importante, ocorre comumente em pessoas normais e pode ser facilmente desconsiderada
2	A manifestação é de relativa importância, mas pode ser freqüentemente ignorada
3	A manifestação é de importância média, porém pode ser um indicador de vários diagnósticos específicos
4	A manifestação é de grande importância e raramente pode ser desconsiderada como, por exemplo, um resultado falso-positivo.
5	A manifestação necessariamente deve ser explicada por um dos diagnósticos finais

3.5.2 Aprendizado dedutivo

O aprendizado por dedução é visto aqui mais como a criação de novos conceitos a partir do conhecimento já existente que como uma operação lógica formal.

O modelo de aprendizado por dedução, utilizado na ferramenta NEXTTOOL, emprega algoritmos genéticos (AG) para fazer a análise dos modelos gerados pela aprendizagem indutiva e gerar novos modelos da realidade [MAC 91a].

A abordagem de aprendizado através de algoritmos genéticos foi introduzida por Holland como uma primeira tentativa de compreender sistemas adaptativos. Os algoritmos baseiam-se em que, a partir de uma população em constante mutação, podem ser gerados indivíduos que se adaptam melhor ao meio e que, por isto, têm

mais possibilidades de sobreviverem e de se reproduzirem (Seleção Natural) [LAW 91].

Os principais componentes dos algoritmos genéticos são:

- um conjunto de cromossomos;
- uma função de avaliação;
- um conjunto de operadores genéticos;
- um conjunto de parâmetros.

Os cromossomos, geralmente codificados como seqüências de 0's e 1's, formam a população que representa o conjunto de soluções do problema. Esta população é modificada e adaptada através dos seguintes operadores genéticos:

- **Operador de Reprodução:** duplica cromossomos já existentes na população, dando maior prioridade para aqueles que apresentarem maiores aptidões. No futuro, cromossomos com múltiplas cópias têm maiores chances de se reproduzirem;
- **Operador de Recombinação:** segmentos selecionados de forma aleatória dos pais são permutados, gerando cromossomos filhos, como apresentado na Figura 3.11.



Figura 3.11 Operador genético de recombinação

- **Operador de Mutação:** este operador gera um filho a partir da substituição, adição ou eliminação de elementos que formam o cromossomo pai.

- **Operador Gradiente:** este operador gera filhos considerando uma função de avaliação aplicada aos pais. Ele toma um múltiplo do gradiente relativo a esta função e soma às entradas do cromossomo pai para gerar os filhos.

A função de avaliação constitui-se basicamente em um meio através do qual valores quantitativos são obtidos para classificar a aptidão de cada cromossomo.

O principal ciclo dos algoritmos genéticos é apresentado a seguir:

- **passo 1** do conjunto de redes neurais existentes, selecionar as mais fortes (pesos mais altos) para serem os *pais* dos novos conceitos classificatórios;
- **passo 2** aplicar operadores genéticos aos *pais*, criando novas redes neurais;
- **passo 3** avaliar a população de redes neurais produzida;
- **passo 4** substituir os *pais* mais fracos pelos *filhos* mais fortes.

O ciclo formado pelos componentes conexionista e genético da máquina de aprendizado foi denominado *Aprendizado Evolutivo* e provou ser muito eficiente no desenvolvimento de sistemas classificatórios com poderosas capacidades adaptativas [MAC 91a].

Todavia, o aprendizado por dedução aqui desenvolvido não faz uso de algoritmos genéticos para gerar novos modelos da realidade. Concebeu-se um mecanismo de reorganização e treinamento das redes neurais, análogo ao operador de recombinação dos algoritmos genéticos. A estratégia para recombinação dos elementos das redes neurais segue os seguintes passos:

Para cada diagnóstico fazer:

- **passo 1** coletar todos os achados do *frame de diagnóstico* e combiná-los em nodos intermediários de dois e três elementos,

de modo a esgotar as possibilidades de combinações entre eles;

- **passo 2** calcular a freqüência com que cada combinação aparece nos casos de treinamento da base de casos, eliminando as combinações menos freqüentes;
- **passo 3** criar acumuladores de punição e de recompensa para cada conexão da rede, colocando-os em 0;
- **passo 4** colocar todos os pesos de conexão da rede em 1.
- **passo 5** retreinar a rede pelos casos da base de casos através do algoritmo de *Punição e Recompensa* - versão de arranque;
- **passo 6** normalizar e podar a rede, com o algoritmo de *poda e normalização*.

Através deste mecanismo, uma rede neural que conecta certas evidências a uma determinada hipótese pode ser reestruturada, gerando novas possibilidades de solução para os problemas de classificação.

A busca de informações no componente simbólico da arquitetura para a recombinação e treinamento das redes neurais é outra característica incorporada nos modelos híbridos de conexão total.

3.6 Interfaces

A interface do sistema, encarregada da comunicação entre o SE e o mundo exterior, foi dividida em *interface do engenheiro de conhecimento* e *interface do usuário final*.

A *interface do engenheiro de conhecimento* é a responsável pelas funções de teste e de *aquisição manual de conhecimento*.

No modo de teste, a interface possibilita ao engenheiro de conhecimento dirigir consultas ao sistema especialista e modificar o conhecimento, até que o sistema apresente um bom nível de desempenho.

No modo de *aquisição manual de conhecimento*, é possível introduzir na base de conhecimento grafos de conhecimento médio que representam o conhecimento de especialistas para as hipóteses de diagnósticos.

Os grafos de conhecimento são traduzidos para *frames de diagnóstico* e redes neurais, utilizando o método descrito na seção 3.2.1. A Figura 3.12 esquematiza o uso da interface do engenheiro de conhecimento para a inserção de grafos de conhecimento.

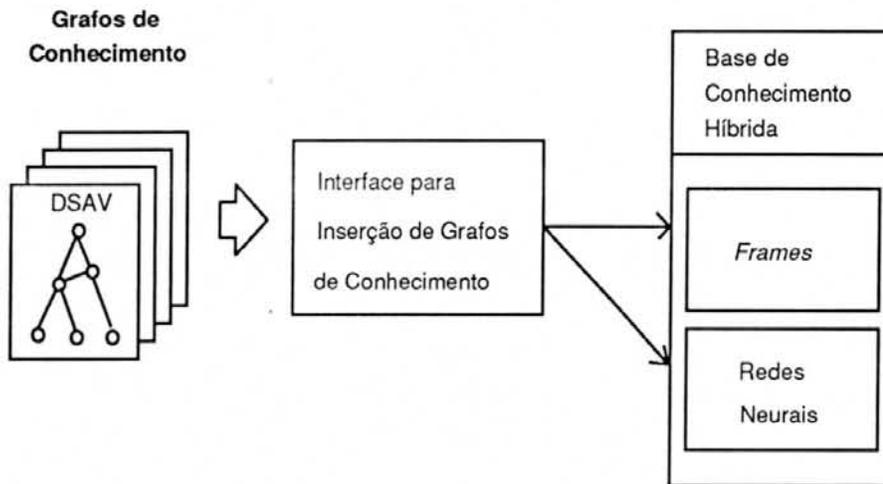


Figura 3.12 Inserção de grafos de conhecimento

A existência da interface para inserção dos grafos de conhecimento possibilita que seja utilizada a metodologia de aquisição de conhecimento, descrita na seção 1.6, onde o conhecimento formalizado em grafos é diretamente inserido no sistema. Isto é particularmente interessante para situações em que não há casos disponíveis a invocar para a construção da base de casos de treinamento.

A interface do usuário final é a responsável pelos modos de operação *pupilo* e *consulta*. O modo *pupilo* é executado pela porção do sistema responsável pelas

consultas à base de conhecimento. Estas consultas podem ser efetuadas percorrendo-se as hierarquias de *frames* através de uma interface do tipo hipertexto. A técnica de hipertexto permite que os dados armazenados sejam relacionados na forma de uma rede de nodos interligados, possibilitando o acesso não-seqüencial por parte do usuário [RAD 91].

A *consulta* ao sistema especialista propriamente dito, inicia através da coleta de informações sobre o caso a ser diagnosticado. O usuário percorre as hierarquias de achados médicos, indicando as evidências que são observadas para o caso e o processo de inferência é disparado automaticamente. Depois de diagnosticado o problema, a interface para consulta informa ao usuário o resultado obtido e apresenta a sua respectiva explicação, a partir do mecanismo de explanação.

4 IMPLEMENTAÇÃO

4.1 Ambiente

O protótipo da ferramenta para desenvolvimento e operação de sistemas especialistas de arquitetura híbrida, chamada HYCONES, foi implementado em equipamento PC compatível 486, configuração EISA. Utilizou-se a ferramenta Goldworks III para criar a estrutura de *frames* do modelo híbrido [HEN 89], e a linguagem de programação LISP para implementar as funções de acesso e manipulação da base de conhecimento híbrida.

4.2 Implementação da Base de Conhecimento Híbrida

A base de conhecimento híbrida é definida a partir da estrutura básica de *frames* apresentada na Figura 4.1.

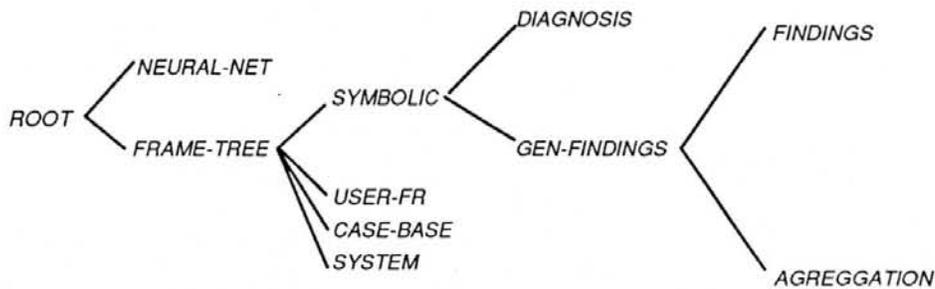


Figura 4.1 Hierarquia de *frames* da Base de Conhecimento Híbrida

O *frame ROOT* corresponde ao elemento situado no topo da hierarquia de *frames*. Associam-se a ele todos os *frames* definidos sem qualquer especificação de posição na hierarquia.

O *frame NEURAL-NET* corresponde à classe sob a qual estão associados todas os nodos da rede neural da base de conhecimento híbrida. Como o modelo neural adotado possui apenas 3 camadas, representam-se apenas os nodos da camada combinatória da rede. Cada um destes nodos é definido como uma instância

da classe *NEURAL-NET*. A classe *NEURAL-NET* é definida com 8 atributos. Os 4 primeiros atributos são relativos à ligação do nodo combinatório com a hipótese de diagnóstico por ele apontada. Estes atributos têm seus identificadores iniciados pela letra D. São eles:

- **D-link** : liga o nodo combinatório a um *frame de diagnóstico* (simbólico)
- **D-weight** : armazena o peso de conexão da ligação do nodo combinatório com o *frame de diagnóstico*
- **D-pun** : é o acumulador de punições da conexão do nodo combinatório com o *frame de diagnóstico*
- **D-rew** : é o acumulador de recompensas da conexão do nodo com o *frame de diagnóstico*

Os 4 atributos restantes têm seus identificadores iniciados pela letra F e possuem as seguintes funções:

- **F-links**: armazena a lista dos achados aos quais o nodo combinatório é ligado
- **F-weights**: mantém uma lista de valores referentes aos pesos de conexão dos achados ao nodo combinatório
- **F-pun**: armazena uma lista de valores correspondentes aos acumuladores de punição de cada uma das conexões dos achados ao nodo combinatório
- **F-rew**: armazena uma lista de valores correspondentes aos acumuladores de recompensa de cada uma das conexões dos achados ao nodo combinatório

O *frame* de nome *FRAME-TREE* tem o objetivo de agrupar toda a informação da base de conhecimento que não diga respeito às redes neurais.

O *frame USER-FRAME* serve para armazenar as informações fornecidas pelo usuário durante uma consulta ao sistema especialista. Ele é definido por dois atributos: *Findings* e *Confidence*, que armazenam, respectivamente, a lista de achados observados pelo usuário e os fatores de confiança que o usuário tem na informação que está fornecendo.

O *frame CASE-BASE* corresponde à classe sob a qual estão associados todos os casos para treinamento introduzidos no sistema. Cada caso é definido por 4 atributos: *Findings*, *Confidence*, *Diagnosis* e *Importance*. O atributo *Findings* armazena os achados existentes para o caso. O atributo *Confidence* armazena os fatores de confiança do usuário em relação à presença dos achados naquele caso. O atributo *Diagnosis* referencia o diagnóstico determinado para o caso, enquanto o atributo *Importance* determina quão importante o caso é para o treinamento do sistema, em uma escala de 0 a 1. Este último fator pode ser empregado quando se quer diferenciar o tratamento que o sistema dará para cada caso, durante o treinamento das redes neurais. Casos com o fator de importância maior serão mais considerados no treinamento, caso contrário, serão menos considerados.

O *frame SYSTEM* possui uma instância associada, *SYSTEM-INFO*, que armazena informações relativas ao sistema em geral. Os atributos definidos para *SYSTEM* são:

- **L-essen**, representando a frequência com que um achado deve estar presente em um certo diagnóstico para ser considerado como *achado essencial*;
- **L-or**, representando o limiar de aceitação utilizado para verificar se a saída apresentada pela rede neural é suficientemente alta para assegurar a correção do diagnóstico apontado como solução;
- **L-prun**, representando o limiar de poda da rede neural. Conexões da rede com peso menor que este limiar são removidas da rede.

O *frame* *SYMBOLIC* foi definido com o objetivo de agrupar as informações referentes ao conhecimento do domínio. A ele estão associados os *frames* *DIAGNOSIS* e *GEN-FINDINGS*. O primeiro representa a classe à qual são associados todos os diagnósticos tratados pelo sistema especialista. Este *frame* é definido pelos atributos *trigger*, *essential*, *complementary* e *negative* (*gatilho*, *achados-essenciais*, *achados-complementares* e *achados de negação*). O *frame* *GEN-FINDINGS* agrupa informações relativas aos achados. Os *frames* *FINDINGS* e *AGGREGATION* são definidos como sub-classes de *GEN-FINDINGS*, visando implementar o conceito de agregação. Para tanto, define-se o atributo *has-parts* (*tem-partes*) para o *frame* *GEN-FINDINGS*, indicando que qualquer classe de achados pode ser composta por outras classes. O *frame* *FINDINGS* é utilizado para descrever a hierarquia de classes de achados. Estas classes podem referenciar outras classes, que as compõem, através do atributo *has-parts*. O *frame* *AGGREGATION* é usado para definir as classes referenciadas nas sub-classes de *FINDINGS* pelo atributo *has-parts*. Os *frames* das folhas da hierarquia de *AGGREGATION* são instâncias de achados da base de conhecimento, que representam os dados de história, exame físico e investigação (Raio X e eletrocardiograma), que são geralmente encontrados no domínio das cardiopatias congênitas.

4.3 Funções do Sistema

A maior parte das funções de acesso à base de conhecimento híbrida e a sua manipulação foi implementada na linguagem de programação *Common LISP*. No entanto, como a ferramenta Goldworks III possui interface gráfica que possibilita a definição, modificação e consulta aos *frames* da base de conhecimento, a mesma pode ser usada para algumas operações de construção e consulta à base de conhecimento. A Figura 4.2 apresenta as três principais funções do protótipo HYCONES e seu esquema de interação com a interface da ferramenta Goldworks III.



Figura 4.2 Módulos do sistema HYCONES

Os módulos de *Construção* são os responsáveis pela estruturação da base de conhecimento do sistema especialista. O conhecimento declarativo do domínio é introduzido na base de conhecimento com o auxílio da interface gráfica da ferramenta GOLDWORKS III. O conhecimento heurístico pode ser inserido na base através da Interface para Inserção de grafos de conhecimento ou através do módulo de explicitação automática de conhecimento do componente de aprendizado indutivo.

Além da automação da explicitação de conhecimento a partir da base de casos, os módulos de *Aprendizado* também são responsáveis pelo refinamento e reorganização da base de conhecimento.

Os módulos de *Consulta* são utilizados tanto para a análise do conhecimento armazenado, através da interface do Goldworks III, quanto para efetuar consultas ao sistema especialista, através do motor de inferência.

O desenvolvimento e operação de um sistema especialista utilizando o protótipo HYCONES e a ferramenta Goldworks III são descritos no diagrama da Figura 4.3.

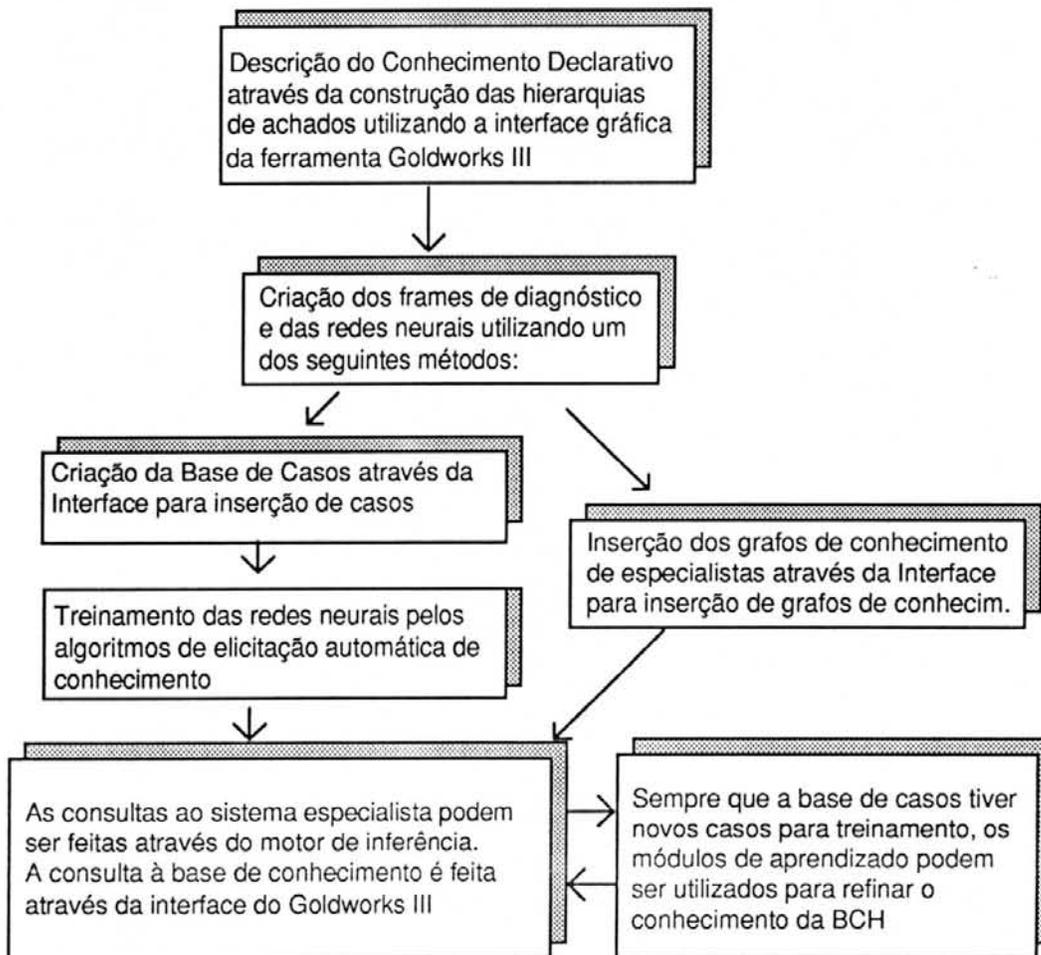


Figura 4.3 Desenvolvimento e operação do sistema HYCONES

4.4 Consulta ao Sistema Especialista

O protótipo HYCONES foi utilizado para construir um sistema especialista na área de cardiologia, especificamente para o diagnóstico de cardiopatias congênitas. As cardiopatias congênitas são causas significativas da mortalidade e morbidade na população pediátrica e afetam 0,8/1000 dos recém-nascidos (EUA). As cardiopatias mais comuns são a Comunicação Intraventricular (CIV), a Comunicação Inter-Atrial (CIA) e o Defeito Septal Atrioventricular (DSAV). O diagnóstico das cardiopatias é feito através da análise da história clínica do paciente (*Anamnese*), de *Exames Físicos* e de *Exames de Investigação*.

A Figura 4.4 apresenta a hierarquia de classes de achados médicos e de diagnósticos para as cardiopatias CIA, CIV e DSAV.

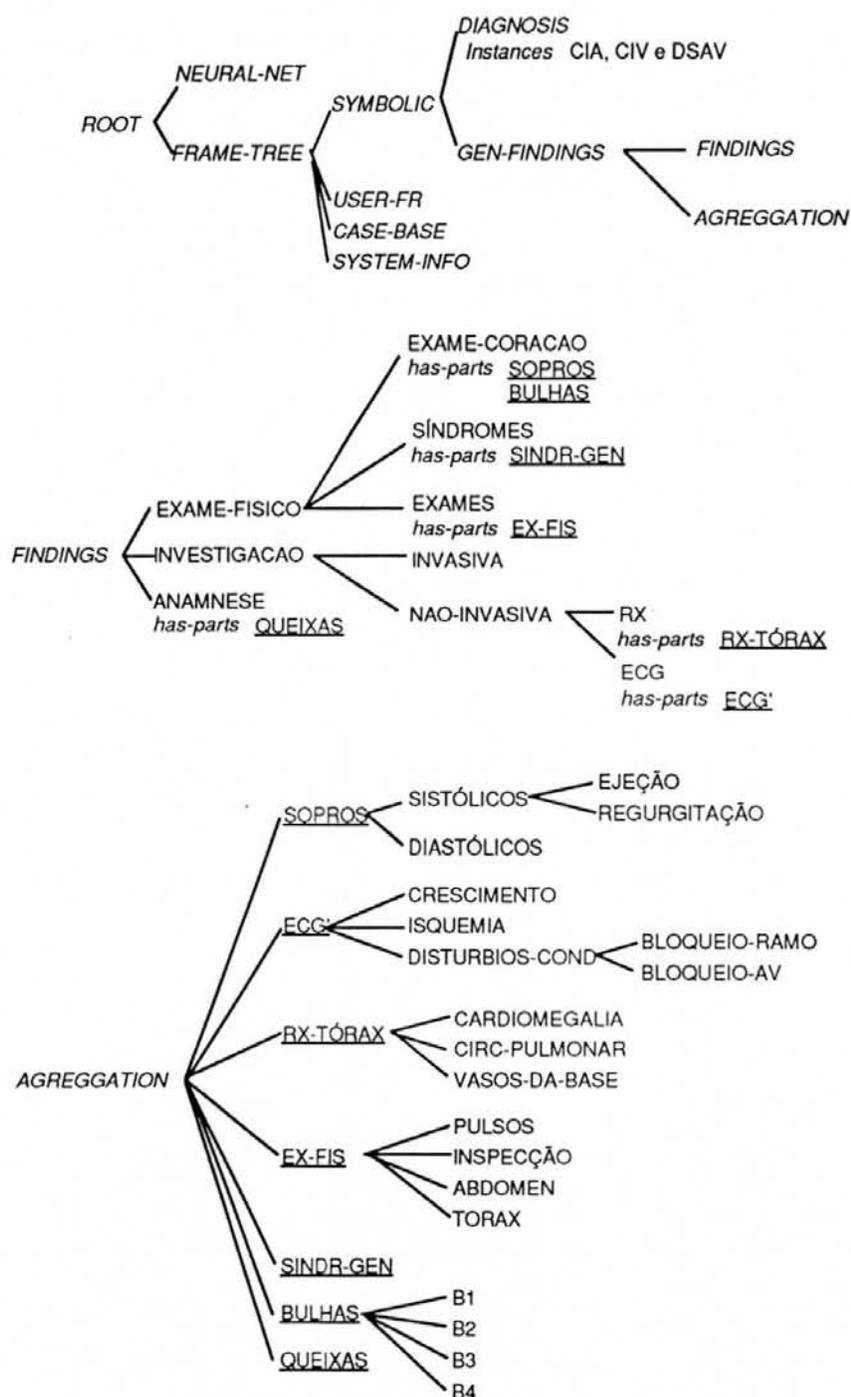


Figura 4.4 Hierarquia de classes de achados para Cardiopatias Congênitas

A seguir, apresenta-se uma consulta feita ao sistema especialista. A consulta tem início com a coleta de informações sobre o caso. O usuário percorre as hierarquias de *achados* médicos, indicando as evidências que são observadas para o caso. A Figura 4.5 mostra a primeira seleção da classe de achados. No exemplo, a classe de achados *Exames Físicos* é selecionada.

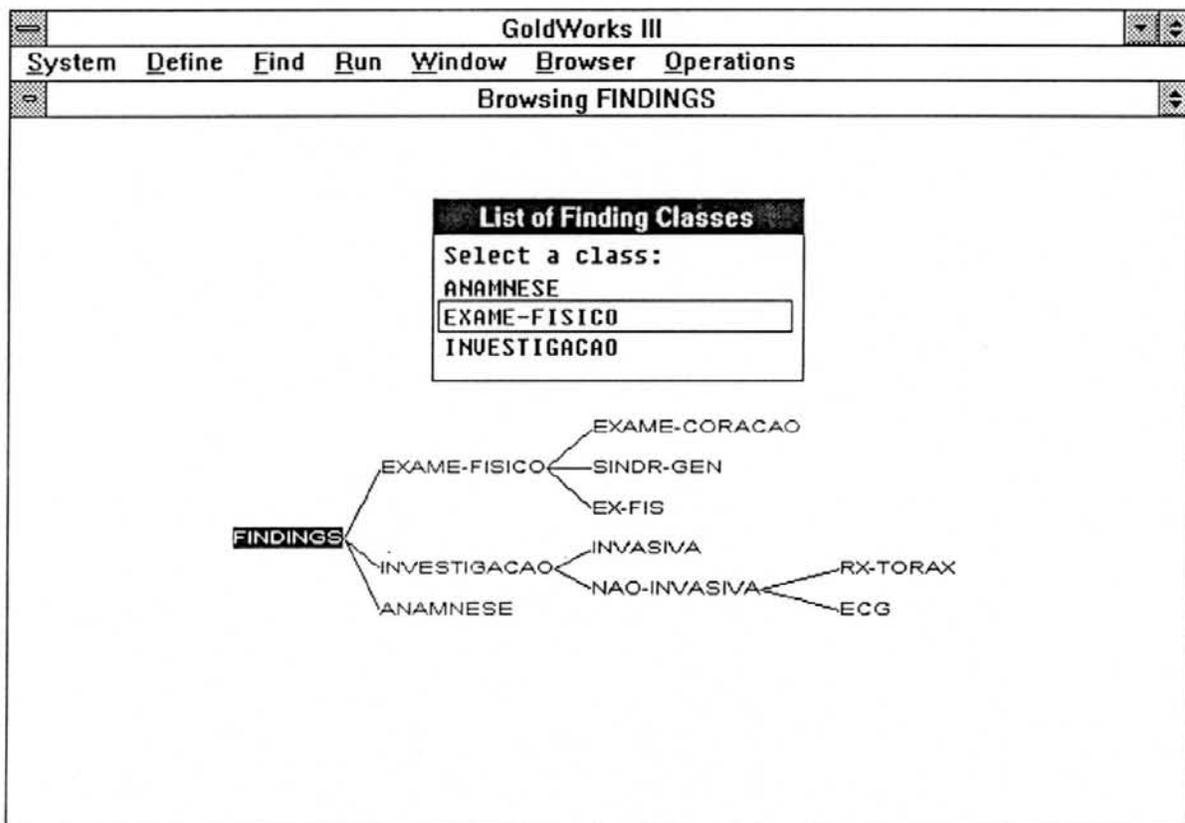


Figura 4.5 Selecionando a classe de achados

Percorre-se a hierarquia de *achados*, até chegar nos nodos folha, que representam os achados médicos propriamente ditos. Neste momento, o usuário seleciona algum achado correspondente a determinada evidência observada para o caso. Depois de selecionar o achado, o usuário é questionado em relação ao grau de

confiança da informação que está sendo fornecida ao sistema, como mostra a Figura 4.6.

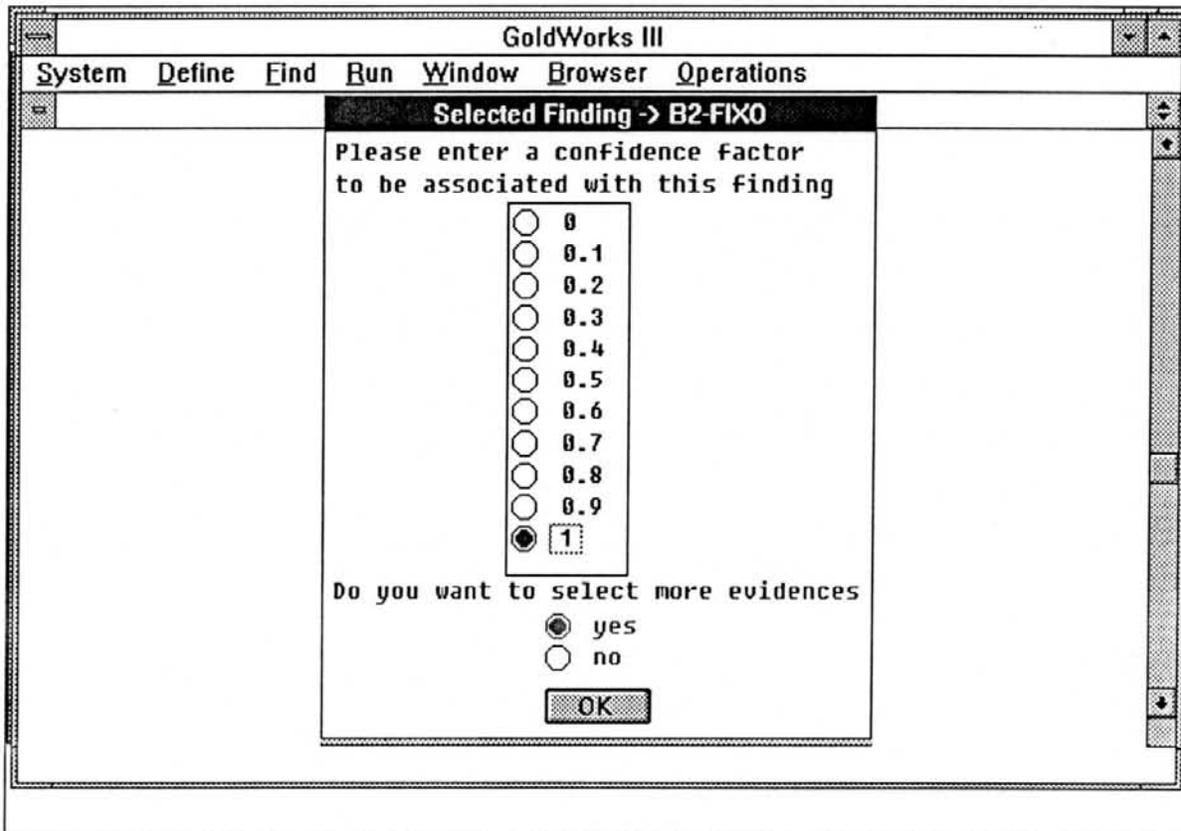


Figura 4.6 Entrada dos fatores de confiança de um achado

O sistema prossegue com a coleta de informações até que o usuário esgote as evidências a apontar. A máquina de inferência é, então, ativada. Calcula-se, a seguir, a saída para cada caminho da rede neural que leve até alguma hipótese diagnóstica. No exemplo ilustrado na Figura 4.7, as informações não foram suficientes para o sistema determinar o diagnóstico. Solicita-se, então, o fornecimento de evidências adicionais.

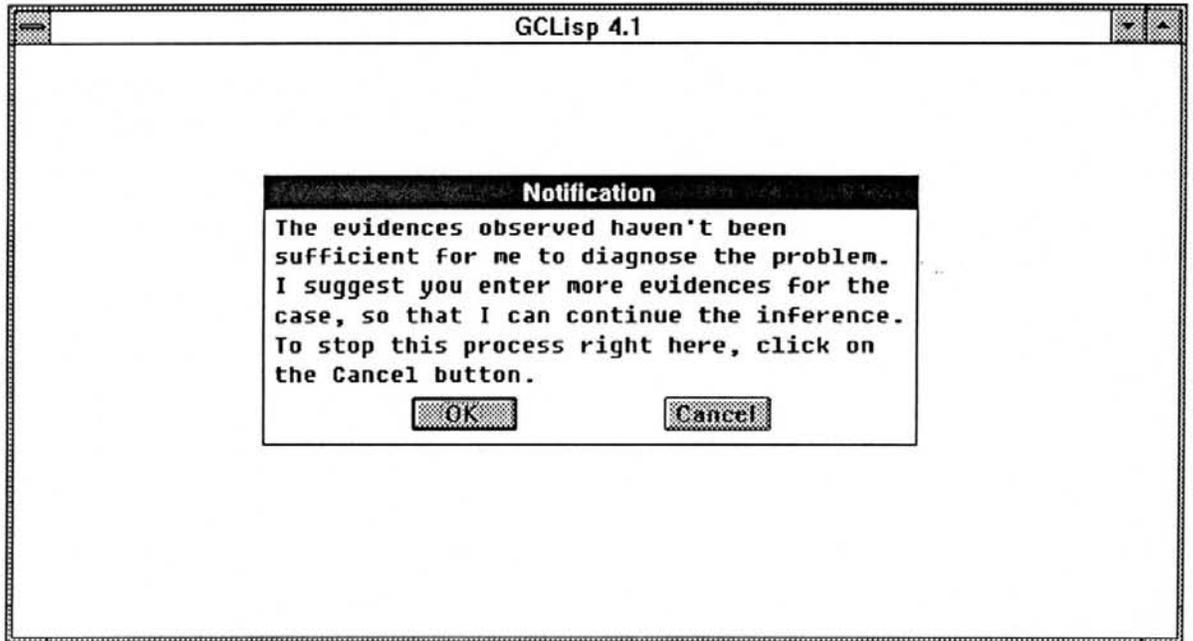


Figura 4.7 Falta de evidência para concluir um diagnóstico

Enquanto as redes neurais do sistema não apresentarem sinal de saída mais alto que o limiar de aceitação (Laceit), o sistema abstém-se de enunciar sua conclusão final. No entanto, quando o achado referenciado no *gatilho* de alguns dos *frames de diagnóstico* estiverem presentes, os correspondentes diagnósticos serão apresentados como conclusões parciais. Após as conclusões parciais, o sistema guia a coleta de novas informações, através da busca dos *achados* referenciados nos *slots de achados-essenciais* e de *achados-complementares* dos *frames de diagnóstico* selecionados como possíveis soluções (Figura 4.8).

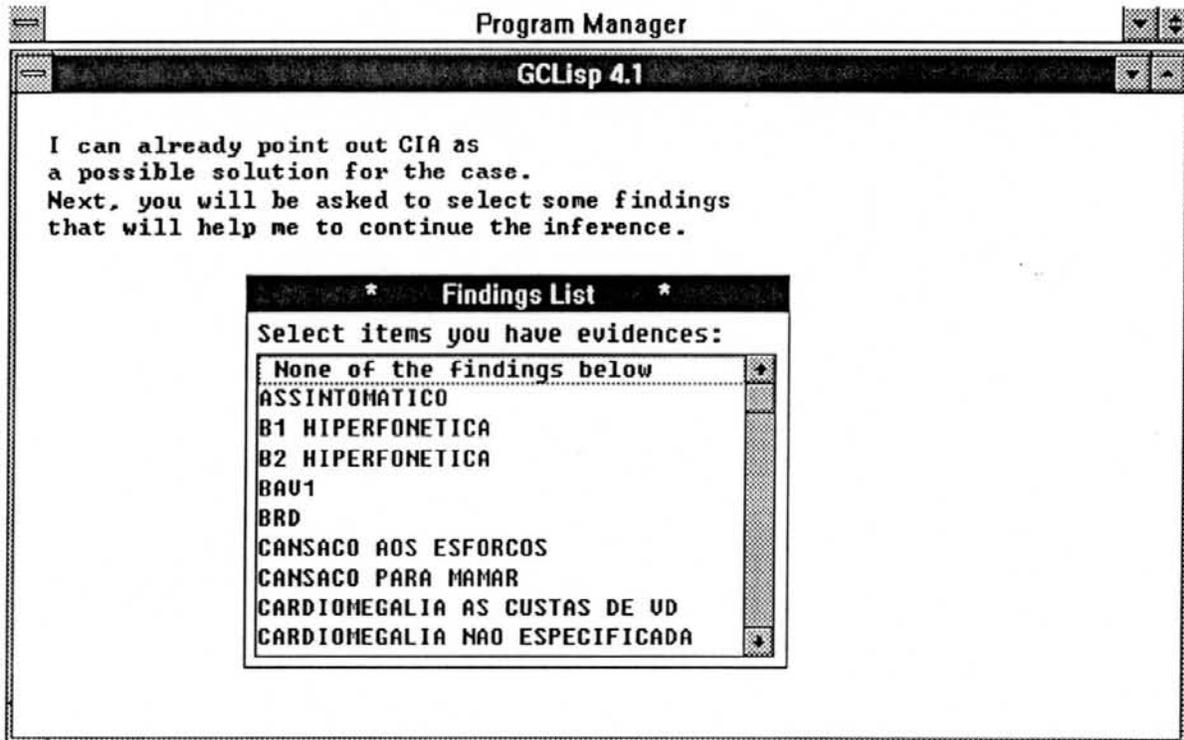
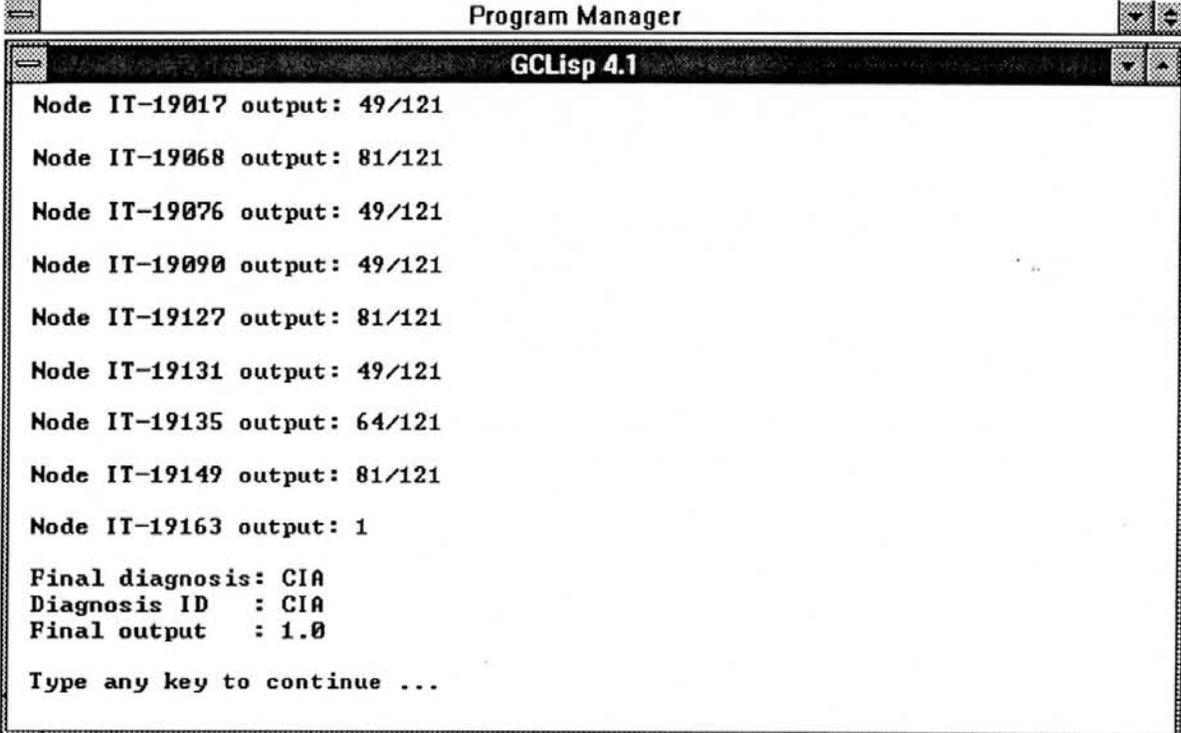


Figura 4.8 Conclusão parcial, com solicitação de informações adicionais

Quando algum caminho da rede neural apresentar sinal de saída maior que o limiar de aceitação, o sistema aponta o diagnóstico referenciado pela rede neural como solução para o problema (Figura 4.9). A seguir, apresenta uma explicação para a seleção deste diagnóstico como solução do problema (Figura 4.10).



The image shows a window titled "Program Manager" containing a sub-window titled "GCLisp 4.1". The sub-window displays the following text:

```
Node IT-19017 output: 49/121
Node IT-19068 output: 81/121
Node IT-19076 output: 49/121
Node IT-19090 output: 49/121
Node IT-19127 output: 81/121
Node IT-19131 output: 49/121
Node IT-19135 output: 64/121
Node IT-19149 output: 81/121
Node IT-19163 output: 1

Final diagnosis: CIA
Diagnosis ID   : CIA
Final output   : 1.0

Type any key to continue ...
```

Figura 4.9 Conclusão diagnóstica por HYCONES

Uma vez concluído o diagnóstico, o sistema apresenta a explicação da inferência realizada (Figura 4.10).

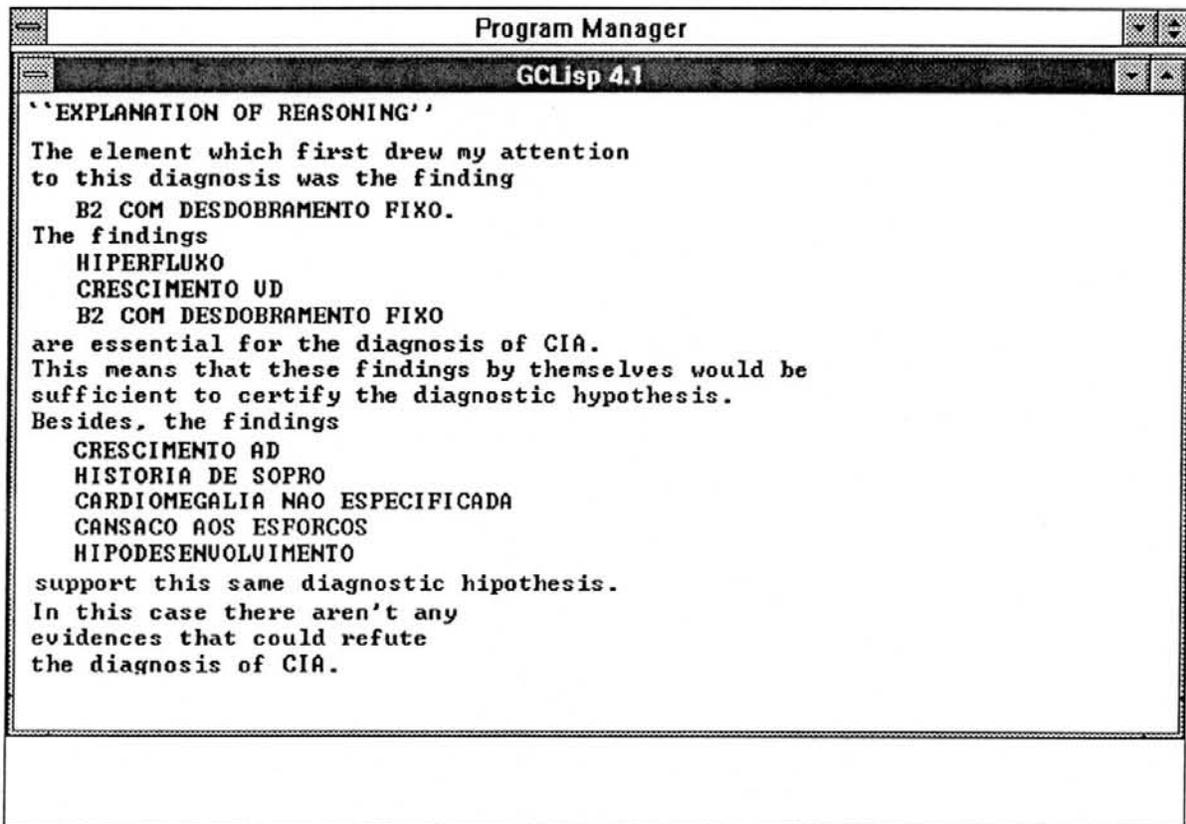


Figura 4.10 Mecanismo Explanatório de HYCONES

Após terminada a consulta, o usuário pode inserir o caso diagnosticado na base de casos, para posterior uso no treinamento das redes neurais (Figura 4.11).

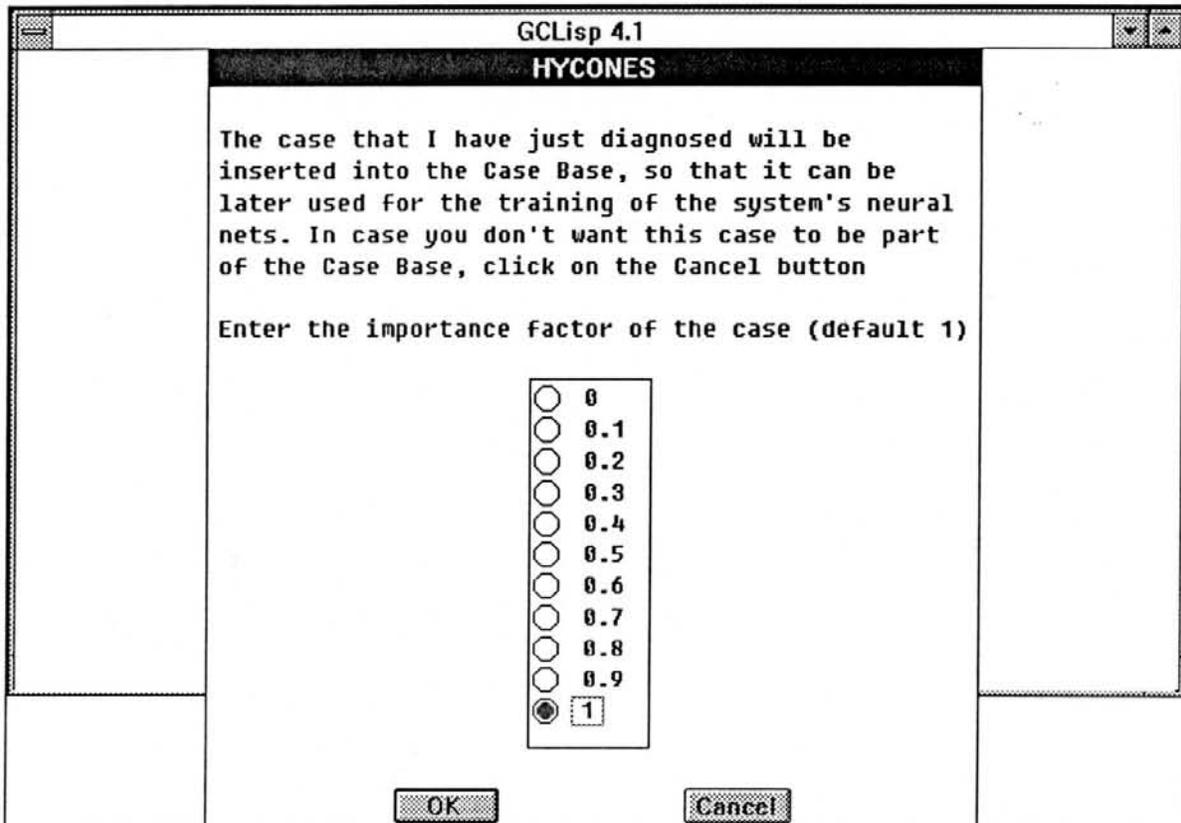


Figura 4.11 Inserção do caso diagnosticado na base de casos

5 VALIDAÇÃO

O desempenho de HYCONES no diagnóstico das três cardiopatias congênitas mais freqüentes foi validado conforme descrição que se segue.

Inicialmente, foi construída a base de casos do sistema. Para a criação desta, foram selecionados, aleatoriamente, 66 casos de pacientes portadores de cardiopatia congênita, submetidos a cirurgia cardíaca no Instituto de Cardiologia - RS, no período de jan/86 a dez/90. Estes 66 casos exibem os três diagnósticos mais freqüentes, ocorridos no período supra-mencionado, na seguinte proporção: 22 com Comunicação Interatrial (CIA), 29 com Comunicação Interventricular (CIV) e 15 com Defeito Septal Atrioventricular (DSAV).

Para validar HYCONES, foram construídas 5 versões da base de conhecimento híbrida, denominadas respectivamente **B1, B2, B3, B4 e B5**, conforme abaixo descrito :

- **versão B1** baseou-se no conhecimento extraído de 22 especialistas em cardiologia pediátrica, de 4 instituições brasileiras, de acordo com a metodologia de aquisição de conhecimento descrita por [LEA 88]. Os grafos médios de conhecimento destes especialistas foram diretamente traduzidos para redes neurais MNC, conforme descrito em 3.2.1;
- **versão B2** baseou-se no refinamento da base de conhecimento da versão **B1**, através da reorganização e treinamento das redes neurais representativas do conhecimento dos especialistas, com aplicação dos métodos de aprendizado dedutivo. Para aplicar este método, o sistema utilizou as evidências da camada de entrada das redes neurais MNC, recombinao-as, com o objetivo de criar várias novas redes, a partir das originais, de acordo com o algoritmo de aprendizado dedutivo descrito na sessão 3.2.5.2.

- **versão B3** baseou-se no conhecimento de 23 não-especialistas, cardiologistas gerais, na sua maioria oriundos do Instituto de Cardiologia - RS (ICFUC). A mesma técnica aplicada para a construção da versão **B1** foi utilizada para construir esta versão.
- **versão B4** corresponde ao refinamento e reorganização da base de conhecimento da versão **B3**, com as mesmas técnicas de aprendizado dedutivo aplicadas à construção da versão **B2**;
- **versão B5** constituída pela construção automática da base de conhecimento híbrida, a partir da base de casos, valendo-se do método de explicitação automática de conhecimento do **MNC**.

Para validar as cinco versões do sistema foram, também, selecionados, aleatoriamente, 33 casos adicionais, do mesmo período acima mencionado: 13 com **CIA**, 10 com **CIV**, e 10 com **DSAV**. Estes 33 casos constituem a base de testes do sistema. Foram excluídos desta seleção aqueles casos que já haviam sido selecionados para a base de casos. A tabela 5.1 abaixo exhibe o desempenho das cinco versões de HYCONES.

Tabela 5.1 Desempenho das cinco versões da BCH

	Versão B1		Versão B2		Versão B3		Versão B4		Versão B5	
	N	%	N	%	N	%	N	%	N	%
Acertos	14	42,5	26	78,7	0	-	28	84,9	31	93,9
Erros	3	9	0	-	0	-	0	-	0	-
Falta de evidências para o diagnóstico	16	48,5	7	21,3	33	100,00	5	15,1	2	6,1
Total	33	100,00	33	100,00	33	100,00	33	100,00	33	100,00

A versão **B1**, baseada nas redes neurais originadas dos grafos médios dos especialistas, concluiu corretamente 14 (42,5%) dos 33 diagnósticos. Faltaram evidências suficientes para concluir o diagnóstico correto em 16 (48,5%) dos 33 casos testados. Em 3 (9%) casos, esta versão concluiu erroneamente pelo diagnóstico.

A versão **B2**, caracterizada pela reorganização e treinamento das redes neurais baseadas nos grafos dos especialistas, concluiu corretamente 26 (78,7%) dos 33 diagnósticos, não tendo ocorrido erros diagnósticos, todavia, para 7 (21,3%) dos 33 casos testados, a versão **B2** não logrou encontrar a solução correta por falta de evidências suficientes.

A versão **B3**, baseada nos grafos de conhecimento dos 23 não-especialistas, não apresentou desempenho satisfatório, não encontrando a solução correta nos 33 casos testados, por falta de evidência suficiente.

A versão **B4**, caracterizada pela reorganização e treinamento das redes neurais baseadas nos grafos dos não especialistas, concluiu corretamente 28(84,9%) dos 33 diagnósticos, sendo que para 5 (15,1%) dos 33 casos, o sistema não alcançou a resposta correta por falta de evidências suficientes.

A versão **B5**, criada pelo processo de automação da aquisição de conhecimento heurístico, apresentou excelente desempenho, concluindo corretamente 31 (93,9%) dos 33 diagnósticos da base de testes. Em 2 (6,1%) dos 33 casos testados, esta versão não concluiu corretamente pelo diagnóstico, por falta de evidência suficiente. Os insucessos se dirigiram a um caso de CIA e um de CIV. Os prontuários médicos relativos a estes casos foram revisados. Constatou-se, então, que, para o primeiro deles, o sistema não foi capaz de identificar o diagnóstico de CIA porque o paciente possuía outro diagnóstico associado, *Estenose Pulmonar Valvar*, que não constava da base de conhecimento. Para o segundo caso, constatou-se que o sistema não foi capaz de diagnosticar CIV porque o paciente associava à lesão congênita cardíaca um quadro que lhe é secundário, a *Hipertensão Pulmonar Severa*, não apresentado, por este motivo, o sopro sistólico, característico dos quadros de CIV.

A validação de HYCONES visava apenas testar pacientes com diagnósticos isolados, sem qualquer outra cardiopatia associada. Para tal, foi realizada uma pesquisa no banco de dados, buscando apenas registros que possuísem diagnósticos isolados. Verificou-se, através desta avaliação, que existem casos erroneamente cadastrados no banco de dados do Instituto de Cardiologia do RS. Ainda que este não fosse um dos objetivos do presente trabalho, o resultado secundário foi útil, pois permitiu corrigir, pelo menos para estes dois casos, os respectivos registros no banco de dados. Portanto, a versão **B5** não concluiu a inferência para estes dois casos mais por um problema de cadastramento de prontuários que por uma falha do mecanismo de inferência ou por deficiência da base de conhecimento.

Concluiu-se, portanto, que a rede neural criada a partir da base de casos foi capaz de formar, de maneira correta, *germes* de conhecimento, i.e., caminhos mínimos que conduzem aos diagnósticos. Também verificou-se a correta criação dos *frames de diagnóstico*, em relação aos *achados-essenciais* e complementares. Foi constatado um problema no cálculo da especificidade de *achados* para determinar o *gatilho* de um diagnóstico. Consiste este num problema semântico, onde alguns achados médicos podem ser compostos por outros achados. Esta relação semântica entre os achados médicos, torna-se fundamental para o cálculo da especificidade de um determinado achado, em relação a um determinado diagnóstico. Por exemplo, o achado ECG com crescimento biventricular, significa que existem ambos *crescimento de VE* e *crescimento de VD* no eletrocardiograma do paciente. Para resolver este problema, obrigou-se o usuário a fornecer a lista de achados a partir de suas componentes. No caso acima citado, para expressar crescimento biventricular, o usuário informou o sistema que ambos os sinais, *crescimento de VE* e *crescimento de VD* estavam presentes. Esta solução resolveu os problemas de forma imediata, porém, futuramente, pretende-se dotar o sistema da capacidade de cálculo de especificidade, mesmo que existam na BCH achados compostos. Isto poderá ser feito através da análise da representação simbólica dos mesmos.

Aplicou-se o teste qui-quadrado (χ^2), para comparar o desempenho das versões **B1**, **B2** e **B4** e **B5**, conforme demonstra a Tabela 5.2. A versão **B3**, por não ter concluído qualquer diagnóstico não foi comparada. A diferença entre os

desempenhos das versões foi altamente significativa. Com o objetivo de cotejar individualmente a versão **B5** com as demais, aplicou-se novamente o teste qui-quadrado, para comparar-se o desempenho das versões **B1**, **B2** e **B4** contra o da versão **B5**. Os quadros de contingência 2 X 2 figuram a seguir.

Tabela 5.2 Quadro de Contigência comparativo de **B1**, **B2**, **B4** e **B5**

Versão	Diagn. Corretos	Diagn. Incorretos ou falta de evidência para a conclusão	TOTAL
B1	14	19	33
B2	26	7	33
B4	28	5	33
B5	31	2	33

$\chi^2 = 31.88$, 3 graus de liberdade , $p = 0,00000055$

Tabela 5.3 Quadro de Contigência comparativo de B1 e B5

Versão	B1	B5	TOTAL
Diagnósticos corretos	14	31	33
Diagnósticos incorretos ou falta de evidência para a conclusão	19	2	33
TOTAL	33	33	33

Tabela 5.4 Quadro de Contigência comparativo de B2 e B5

Versão	B2	B5	TOTAL
Diagnósticos corretos	26	31	33
Diagnósticos incorretos ou falta de evidência para a conclusão	7	2	33
TOTAL	33	33	33

Tabela 5.5 Quadro de Contigência comparativo de B4 e B5

Versão	B4	B5	TOTAL
Diagnósticos corretos	28	31	33
Diagnósticos incorretos ou falta de evidência para a conclusão	5	2	33
TOTAL	33	33	33

Os valores de qui-quadrado obtidos encontram-se na Tabela 5.6. O nível de significância adotado foi o de 5% unilateral. Não foram detectadas diferenças estatisticamente significativas, entre as versão B2 e B4, quando cotejadas com a versão B5. Em outras palavras, quer a BCH haja sido treinada automaticamente, a partir da base de casos (B5), quer este treinamento tenha ocorrido a partir dos grafos dos especialistas (B2) ou daqueles dos não-especialistas (B4), o desempenho foi considerado semelhante. A diferença entre os desempenhos das versões B1 e B5 foi altamente significativa, com forte predominância da precisão diagnóstica da versão automaticamente treinada.

Tabela 5.6 Valores de χ^2 na comparação entre B1 e B5, B2 e B5, B4 e B5

Versão	χ^2	p (unilateral)	Conclusão
B1 X B5	17,88 (*)	0,0000235	altamente significativo
B2 X B5	(**)	0,0743604	não significativo
B4 X B5	(**)	0,2133249	não significativo

(*) Com correção de Yates para as freqüências reais

(**) Pelo teste exato de Fisher

A tabela 5.7 mostra um quadro comparativo entre o conjunto de achados essenciais para B1 e B2.

Tabela 5.7 Comparação entre os achados essenciais de B1 e B2

	B1	B2
CIA	B2-Fixo * Hiperfluxo	B2-Fixo * Hiperfluxo Cresc. VD
CIV	Sopro Holo-sist. * Hiperfluxo Cresc. VD Cresc. VE Cardiom. não esp.	Cresc. VE * Cresc. VD Hiperfluxo
DSAV	HBAE * Síndr. Down Cardiom. não esp.	Síndr. Down * HBAE Hiperfluxo

* Gatilho

Os dados da tabela 5.7 mostram que, tanto para CIA quanto para DSAV, não ocorreram grandes modificações no conhecimento dos achados essenciais. Já para CIV, verifica-se que o sistema retirou desta lista de achados *Sopro Holo-sist.*, sendo

este o achado considerado o mais importante pelos especialistas para o diagnóstico de CIV. Isto pode ser explicado pelo fato de que o relato do exame físico dos pacientes é feito, na maior parte das vezes, por não especialistas, que descrevem o sopro observado não na mesma maneira que o especialista o faz. Portanto, o sistema distribui o grau de importância de sopro para os diversos sopros observados nos casos de treinamento. Este problema pode ser resolvido incorporando-se ao motor de inferência a capacidade de identificação de conexões semânticas entre os achados.

A tabela 5.8 mostra um quadro comparativo entre o conjunto de achados essenciais para B3 e B4.

Tabela 5.8 Comparação entre os achados essenciais de B3 e B4

	B3	B4
CIA	Hiperfluxo * BRD Sopro Sist. ñ. esp. Dilatação TAP	B2-Fixo * Hiperfluxo Cresc. VD
CIV	Hiperfluxo * Sopro Holo-sist. Cresc. VD Cresc. VE Infecção resp.	Cresc. VE * Cresc. VD Hiperfluxo
DSAV	Síndr. Down * Hiperfluxo HBAE Sopro Sist. VE	Síndr. Down * HBAE Hiperfluxo

* Gatilho

Os dados da tabela 5.8 mostram diferenças tanto qualitativas quanto quantitativas em relação ao conjunto de achados essenciais para cada um dos diagnósticos. Como a base de conhecimento B4 apresentou um número de acertos significativamente maior do que B3, pode-se concluir da capacidade do mecanismo de aprendizado dedutivo em reorganizar o conhecimento da BCH e gerar novas formas de resolução para os problemas já existentes.

A tabela 5.9 mostra um quadro comparativo entre o conjunto de achados essenciais para B1 e B5.

Tabela 5.9 Comparação entre os achados essenciais de B1 e B5

	B1	B5
CIA	B2-Fixo * Hiperfluxo	B2-Fixo * Hiperfluxo Cresc. VD
CIV	Sopro Holo-sist. * Hiperfluxo Cresc. VD Cresc. VE Cardiom. não esp.	Cresc. VE * Cresc. VD Hiperfluxo
DSAV	HBAE * Síndr. Down Cardiom. não esp.	Síndr. Down * HBAE Hiperfluxo

* Gatilho

Os dados da tabela 5.9 também mostram que, tanto para CIA quanto para DSAV, não ocorreram grandes modificações no conhecimento dos achados essenciais. Para CIV, ocorre o mesmo problema relatado na comparação entre B1 e B2. Logo, pode-se concluir da capacidade do mecanismo de aquisição automática de conhecimento em identificar o conjunto de achados importantes para a detecção dos diagnósticos. No entanto, apesar de apresentarem um conjunto de achados essenciais semelhantes, a versão B5 apresenta um número de acertos maior do que a versão B1. Isto pode ser explicado pelo fato de que a lista de achados essenciais de um frame de diagnóstico armazena uma representação média do conhecimento das redes neurais, baseada na frequência em que os achados aparecem para cada diagnóstico. Contudo, outros caminhos para resolução dos problemas também são armazenados nas redes. Por isto, mesmo para casos não descritos pelos seus frames de diagnóstico, B5 é capaz de encontrar o diagnóstico correto com base em casos anteriormente analisados.

O sistema foi posteriormente avaliado em relação aos seguintes componentes:

- Estrutura para representação de conhecimento
- Mecanismo de inferência e explanação

Para validar o modelo híbrido, o conhecimento armazenado na BCH foi avaliado por um especialista, visando diagnosticar qualquer ausência de informações ou restrições impostas pelo modelo. Nenhuma restrição foi identificada. Ao contrário, verificou-se que os mecanismos de abstração são realmente eficientes para descrever o conhecimento do domínio e que a divisão em *frames de achados* e *frames de diagnóstico* constitui-se em uma abordagem potencialmente eficaz para problemas de classificação.

Ainda em relação aos mecanismos de representação de conhecimento, constatou-se que a integração entre *frames* e redes neurais fornece a estas últimas uma eficiente representação simbólica do conhecimento. Comprova-se esta afirmação pelas inúmeras ocasiões em que se recorre aos *frames de diagnóstico* para analisar o conhecimento armazenado na rede neural, pois a análise da rede propriamente dita é muito difícil, pelo grande número de células e conexões.

Os resultados obtidos na tabela 5.9 mostram que o mecanismo de inferência utilizado pode apresentar um ótimo desempenho, com altos índices de acerto, dependendo do conhecimento armazenado. Para os casos que o sistema não conseguiu resolver, por insuficiência de informações, a busca de outras informações nos *frames de diagnóstico* se mostrou muito eficiente ao usuário, auxiliando-o no processo de coleta adicional de informações relevantes para o prosseguimento da inferência.

Quanto ao mecanismo de explanação, as explicações apresentadas ao final de uma consulta são bastante convincentes e ajudam o usuário a compreender o modelo de raciocínio empregado por especialistas (e reproduzido pelo sistema) na resolução dos problemas de diagnóstico.

A Tabela 5.10 mostra a avaliação das funções do protótipo HYCONES, atualmente disponíveis.

Tabela 5.10 Avaliação das funções do protótipo HYCONES

Funções	Estado	Observações
Inferência	Implementado	Desempenho plenamente satisfatório
Aquisição de conhecimento por casos de treinamento	Implementado	Ótima capacidade de explicitação de conhecimento heurístico.
Construção da base de conhecimento por grafos de conhecimento	Implementado	Gera bases de conhecimento híbridas, fiéis ao conhecimento dos especialistas
Aprendizado incremental	Implementado	Desempenho satisfatório
Aprendizado dedutivo por reorganização e treinamento das redes neurais	Implementado	Ótima capacidade de reorganização e refinamento do conhecimento da base
Aprendizado dedutivo por algoritmos genéticos	Não implementado	x
Módulo de inserção de casos de treinamento na base de casos	Implementado	Obriga o usuário a fornecer as evidências, de acordo com jargão utilizado na definição dos <i>frames de diagnósticos</i> e <i>achados</i> presentes na BCH.

6 DISCUSSÃO

Em inteligência artificial, tanto a abordagem conexionista como a simbólica visam:

- encontrar métodos mais eficientes de busca;
- representar o conhecimento de forma mais clara para um especialista, mas também mais computável por uma máquina;
- descobrir novos métodos de aprendizagem de máquina.

Cada abordagem, contudo, soluciona estes problemas de maneira diferente. Por exemplo, a representação de conhecimento em sistemas conexionistas é feita através de redes neurais. Já na abordagem simbólica, o conhecimento é representado através de sistemas de produção, redes semânticas, *frames* e outros.

Tanto a abordagem conexionista como a simbólica, sem dúvida, têm pontos fortes e fracos. Uma das principais vantagens dos sistemas conexionistas está em que o esquema utilizado para a representação do conhecimento permite a introdução de métodos de aprendizado de forma mais natural. Entretanto, o conhecimento armazenado em uma rede neural treinada é geralmente de difícil análise e compreensão.

Os sistemas híbridos apresentam soluções eficientes para estes problemas. Tais sistemas integram diferentes paradigmas, como: o processamento simbólico, a lógica nebulosa, os modelos conexionistas e os algoritmos genéticos. Esta integração visa extrair de cada um destes paradigmas, o que de melhor oferecem, com o objetivo de simular, de forma mais apropriada, a capacidade humana para representar conhecimento e aprender.

O modelo para sistemas especialistas classificatórios, aqui descrito, foi baseado na arquitetura híbrida que integra redes neurais a redes semânticas, desenvolvida por Machado et. al. [MAC 92a] [MAC 92b]. NEXTTOOL é a ferramenta que implementa o modelo híbrido de redes neurais e redes semânticas, empregada

com sucesso no desenvolvimento de um sistema especialista para o diagnóstico de síndromes renais. Embora NEXTOOL apresente poderosas propriedades adaptativas fornecidas pela sua máquina de aprendizado indutivo e dedutivo, o sistema não possui mecanismos para representação de conhecimento suficientemente flexíveis. A ferramenta para sistemas especialistas conexionistas híbridos - HYCONES - foi desenvolvida sobre os mesmos princípios básicos; no entanto, emprega mecanismos de representação de conhecimento baseados nos conceitos de abstrações de dados, visando tornar o modelo mais flexível. Acredita-se que os tipos construtores oferecidos pelas formas de abstração permitem a modelagem de conhecimento em diversas áreas de aplicação. Adicionalmente, possibilitam que a base de conhecimento, implementada como uma hierarquia de objetos, seja mapeada para um banco de dados [MAT 89]. Este é um requisito essencial para aplicações que trabalham com grandes volumes de dados. Para estas, não é possível manter o conhecimento integralmente em memória principal, impondo-se manter parte da base de conhecimento em disco.

Uma das dificuldades do sistema NEXTOOL é que seu mecanismo de representação de conhecimento simbólico (redes semânticas) não representa de maneira eficiente o conhecimento armazenado nas redes neurais. A ligação entre os dois paradigmas de representação de conhecimento é feita pelos arcos *influência* da rede semântica. Estes arcos, no entanto, não exprimem a importância das evidências e sintomas para os diagnósticos. No modelo de HYCONES, a integração das redes neurais aos *frames* é direta e clara. Os atributos *gatilho*, *achados-essenciais*, *achados-complementares* e *achados de negação* dos *frames de diagnóstico* explicam como os achados, que são os nodos de evidências das redes neurais, se comportam em relação aos diagnósticos. Isto possibilita à ferramenta HYCONES construir explicações mais completas para o seu processo de inferência. A classificação dos achados que influenciam o diagnóstico em quatro maneiras diferentes, através dos diferentes atributos do *frame diagnóstico*, permite ao sistema explicar a detecção de um diagnóstico a partir de cada um dos achados e do papel que desempenham na obtenção do mesmo.

Quanto à automação do conhecimento heurístico, NEXTOOL emprega um algoritmo de criação da rede neural que inicialmente gera combinações de todos os

achados do domínio da aplicação para cada diagnóstico, sem verificar se são realmente relevantes para os diagnósticos. Este método de agrupamento de achados em nodos intermediários do MNC traz o problema da explosão combinatória. O protótipo HYCONES melhora este problema, ao agrupar em nodos combinatórios, apenas os achados que efetivamente pertencem a um determinado diagnóstico, eliminando, portanto, as combinações sem significado semântico. A incorporação de novas informações a estas redes fica a cargo do componente de aprendizado dedutivo do sistema.

Tabela 6.1 Comparação das ferramentas NEXTOOL e HYCONES

	NEXTOOL	HYCONES
Representação de conhecimento	Redes Semânticas X Redes Neurais (MNC)	<i>Frames</i> X Redes Neurais (MNC)
Modelo de integração dos paradigmas simbólico e conexionista	Modelo simbólico não representa conhecimento das redes neurais	Modelo simbólico representa o conhecimento armazenado nas redes neurais através da divisão dos achados em <i>gatilho</i> , <i>achados-essenciais</i> , <i>achados-complementares</i> e <i>achados de negação</i>
Construção da base de conhecimento por casos de treinamento	Problema de explosão combinatória	Melhora o problema da explosão combinatória combinando apenas achados relevantes para os diagnósticos e eliminando combinações não freqüentes
Aprendizado incremental	Sistema capaz de aprender com a experiência	Emprega o mesmo modelo de aprendizado incremental implementado no NEXTOOL
Aprendizado dedutivo	Emprega algoritmos genéticos para modificar a estrutura das redes neurais	Emprega estratégia de re-estruturação e treinamento das redes neurais
Mapeamento da base de conhecimento para um banco de dados	Mapeamento complexo	Mapeamento mais simples devido à utilização de <i>frames</i> como forma de representação de conhecimento

A Tabela 6.1, ilustra as principais diferenças entre as ferramentas HYCONES e NEXTOOL.

A validação do modelo comprovou a eficiência dos métodos de explicitação de conhecimento a partir de uma base de casos, bem como do aprendizado dedutivo por reorganização e treinamento das redes neurais. A base de conhecimento gerada pelo método de explicitação automática de conhecimento apresentou bom desempenho, quando utilizada para efetuar o diagnóstico das três cardiopatias congênitas isoladas. Já as bases geradas a partir dos grafos de conhecimento dos especialistas e não-especialistas não apresentaram bom desempenho. Constatou-se, portanto, que o método de explicitação de conhecimento a partir da base de casos é capaz de gerar bases de conhecimento mais eficientes que as geradas através da metodologia de aquisição de conhecimento por grafos de conhecimento.

Tendo em vista que os grafos médios de conhecimento dos especialistas contavam com pequeno número de evidências de entrada, característica esta do raciocínio de especialistas [LEA 88], poucas redes neurais foram criadas, a partir do algoritmo de aprendizado dedutivo. Estas redes, por conseguinte, não foram capazes de demonstrar desempenho semelhante ao da base de conhecimento treinada a partir do banco de casos. Note-se, todavia, que a comparação não se refere a um especialista, mas a um grafo de conhecimento obtido a partir de determinada metodologia. Por outro lado, na construção do banco de casos, foram utilizadas TODAS as informações disponíveis no prontuário médico referentes à história do paciente, o exame físico, o Raio X de tórax e o eletrocardiograma. É evidente pois, que o aprendizado dedutivo, a partir da reorganização e treinamento das redes neurais, foi capaz de gerar um número muito maior de novas redes, quanto maior fosse a população inicial de evidências ou achados. A comparação entre as diferentes versões treinadas do sistema, i.e. **B2**, **B4** e **B5**, a partir, respectivamente, dos grafos dos especialistas, dos grafos dos não-especialistas e, finalmente do banco de casos, evidenciam a eficiência do método de aprendizado dedutivo implementado em HYCONES. A diferença no número de acertos entre os sistemas treinados por este método e aquele criado pelo método de explicitação automática de conhecimento não foi estatisticamente significativa, concluindo-se, deste modo, que

ambos os métodos de aprendizado são eficazes na definição de bases de conhecimento para problemas classificatórios.

As principais contribuições deste trabalho foram:

- a definição do mecanismo para integração de *frames* a redes neurais. Este mecanismo possibilitou a incorporação das características adaptativas das redes neurais ao paradigma simbólico, bem como tornou explícito o conhecimento das camadas intermediárias do componente conexionista;
- a especificação de um método de explicitação automática de conhecimento, a partir de uma base de casos, capaz de extrair, de um conjunto de casos reais, o conhecimento necessário para identificação de diagnósticos. O método especificado emprega os fatores de especificidade e sensibilidade para capturar e interpretar o tipo de influência dos achados em relação aos diagnósticos, reunindo, com sucesso, os algoritmos de aprendizado do modelo neural MNC à análise dos fatores especificidade e sensibilidade, efetuadas pelo sistema INTERNIST.
- a definição de uma estratégia para mimetizar o aprendizado dedutivo, capaz de reorganizar o conhecimento de uma base de conhecimento já existente e incrementar o desempenho do sistema.

Outras contribuições importantes foram a construção do protótipo HYCONES e o desenvolvimento do sistema especialista para o diagnóstico das cardiopatias congênitas CIA, CIV e DSAV. No futuro, este sistema especialista deve ser expandido para diagnosticar um total de 12 cardiopatias congênitas e deve ser implantado, em caráter experimental, no Instituto de Cardiologia do Rio Grande do Sul.

Embora já tenha oferecido resultados satisfatórios, o protótipo HYCONES pode ser aperfeiçoado em vários aspectos. Em relação à representação de conhecimento do domínio, seria interessante definir funções capazes de identificar conexões semânticas entre os achados da BCH, para efeito do cálculo da

especificidade dos achados, em relação aos diagnósticos e do treinamento das redes neurais.

O protótipo HYCONES também pode ser melhorado em relação ao seu mecanismo de aprendizado dedutivo. A implementação de algoritmos genéticos possibilitaria a criação de nodos da camada combinatória das redes neurais envolvendo maior número de achados. Os algoritmos criariam e restaurariam conexões entre nodos, de forma a representar de maneira mais fiel o conhecimento heurístico do domínio, evitando, ao mesmo tempo a explosão combinatória.

7 CONCLUSÃO

Este experimento, pelos resultados alcançados na construção e validação da ferramenta HYCONES, permite a formulação das seguintes conclusões:

- O modelo de representação de conhecimento, integrando *frames* a redes neurais é bastante flexível para representar conhecimento relativo a problemas de diagnóstico. O mecanismo de integração entre os paradigmas simbólico e conexionista fornece meios para representar de forma eficiente o conhecimento de difícil análise e compreensão, armazenado nas redes neurais.
- O mecanismo de inferência baseado no modo de consulta do MNC apresenta bom desempenho. Sua combinação com a busca de informações nos *frames* de diagnóstico para resolução de problemas não solucionados por falta de evidência também se mostra eficiente.
- A validação do protótipo HYCONES, através da construção do sistema especialista para cardiopatias congênitas e dos testes para os 33 casos clínicos, mostrou a eficiência do mecanismo de aprendizado indutivo por explicitação automática de conhecimento. Os resultados obtidos levam à conclusão de que, dada a existência de um banco de casos, a estratégia de aquisição automática de conhecimento por casos reais produz resultados significativos. Se o banco de casos não existir, a base de conhecimento do sistema especialista deve ser necessariamente construída a partir de um trabalho de aquisição de conhecimento com especialistas. A versão do sistema criada a partir do aquisição de conhecimento, feita com o grupo de não-especialistas não apresentou bons resultados e provou serem ineficientes os grafos médios de conhecimento por eles definidos .
- A validação do protótipo também mostrou ser eficiente o mecanismo de aprendizado dedutivo. As bases de conhecimento treinadas por este método de aprendizado apresentaram melhoras significativas no desempenho dos sistemas especialistas.

Em relação à pesquisa em sistemas híbridos, inúmeras possibilidades são existentes. Estudos neste sentido talvez auxiliem a compreender melhor como funciona a mente humana e contribuam para a realização do antigo sonho do homem, de construir máquinas inteligentes.

BIBLIOGRAFIA

- [BAR 86] BARR, A.; FEIGENBAUM, E. A. **The Handbook of artificial intelligence**. Reading: Addison-Wesley, 1986. v.1, p.161-222.
- [BOO 88] BOOSE, J.; GAINES, B. (eds). **Knowledge acquisition tools for expert systems**. London: Academic Press, 1988. v.2, 200 p.
- [BOO 89] BOOKER, L. B.; GOLDBERG, D. E.; HOLLAND, J. H. Classifier systems and genetic algorithms. **Artificial Intelligence**, Amsterdam, v. 40, n. 1-3, p. 235-282, Sept. 1989.
- [CAR 89] CARBONELL, J. G. Introduction: paradigms for machine learning. **Artificial Intelligence**, Amsterdam, v. 40, n. 1-3, p. 1-9, Sept. 1989.
- [CLA 83] CLANCEY, W. J. The Epistemology of a rule-based expert system : a framework for explanation. **Artificial Intelligence**, Amsterdam, v. 20, n. 3, p. 215-251, May 1983.
- [FIE 90] FIESCHI, M. **Artificial intelligence in medicine: expert systems**. Londres: Chapman and Hall, 1990. 166 p.
- [FIK 85] FIKES, R.; KEHLER, T. The Role of frame-based reasoning. **Communications of the ACM**, New York, v. 28, n. 9, p. 904-920, Sept. 1985.
- [GAL 88] GALLANT, S. I. Connectionist expert systems. **Communications of the ACM**, New York, v. 31, n. 2, p. 152-169, Feb. 1988.
- [HAR 85] HARMON, P.; KING, D. **Artificial intelligence in business**. New York: John Wiley, 1985. 283 p.
- [HAY 83] HAYES-ROTH, I. et al. **Building experts systems**. Reading: Addison-Wesley, 1983. 444 p.

- [HEN 89] HENSON, D. G. **Goldworks III** : reference manual. Cambridge: Gold Hill Computers, 1989. 328 p.
- [HIN 86] HINTON, G. E.; McCLELLAND, J. L.; RUMERHART, D. E. Distributed representation. In: RUMELHART, D. E.; McCLELLAND, J. L. (eds). **Parallel distributed processing: explorations in the microstructure of cognition**. Cambridge: MIT Press, 1986. 547 p. p. 77-110.
- [HIN 92] HINTON, G. E. How neural networks learn from experience. **Scientific American** : Special Issue Mind and Brain, New York, v. 267, n. 3, p. 27-45, Sept. 1992.
- [HUD 92] HUDSON, D. L.; BANDA, P. W.; COHEN, M. E.; BLOIS M. S. Medical diagnosis and treatment plans derived from a hybrid expert system. In: KANDEL, A.; LANGHOLZ, G. **Hybrid architectures for intelligent systems**. Boca Raton: CRC Press, 1992. 420 p. p. 329-344.
- [HUL 87] HULL, R.; KING, R. Semantic database modeling: survey, applications, and research issues. **ACM Computing Surveys**, New York, v. 19, n. 3, p. 201-260, Sept. 1987.
- [KAN 92] KANDEL, E. R.; HAWKINS, R. D. The Biological basis of learning and individuality. **Scientific American** : Special Issue Mind and Brain, New York, v. 267, n. 3, p. 48-57, Sept. 1992.
- [KNA 92] KNAUS, R. Representing expert knowledge in neural nets. In: KANDEL, A.; LANGHOLZ, G. **Hybrid architectures for intelligent systems**. Boca Raton: CRC Press, 1992. 420 p. p. 345-356.
- [KNI 89] KNIGHT, K. **A Gentle introduction to subsymbolic computation: connectionism for the AI researcher**. Pittsburgh: Carnegie-Mellon University, 1989. 28 p.

- [KRO 92] KROVVIDY S. ; WEE W. G. An Intelligent Hybrid System for wastewater treatment. In: KANDEL, A.; LANGHOLZ, G. **Hybrid architectures for intelligent systems**. Boca Raton: CRC Press, 1992. 420 p. p. 357-378.
- [KUR 90] KURZWEIL, R. **The Age of intelligent machines**. Cambridge: MIT Press, 1990. 569 p.
- [LAW 91] LAWRENCE, D. **The Handbook of genetic algorithms**. New York: Van Nostrand Reinhold, 1991. 383 p.
- [LEA 88] LEÃO, B. F. **Construção da base de conhecimento de um sistema especialista de apoio ao diagnóstico de cardiopatias congênitas**. São Paulo: Universidade de São Paulo, Escola de Saúde Pública, 1988. 230 p. (Tese de Doutorado)
- [LEA 90] LEÃO, B. F.; ROCHA, A. F. Proposed methodology for knowledge acquisition: a study on congenital heart disease diagnosis. **Methods of Information in Medicine**, Amsterdam, v. 29, n. 1, p. 30-40, Jan. 1990.
- [MAC 89] MACHADO, R. J. ; ROCHA, A. F. **Handling knowledge in high order neural networks: the combinatorial neural model**. Rio de Janeiro: IBM Rio Scientific Center, 1989. 22 p. (Technical Report CCR076).
- [MAC 90] MACHADO, R. J.; ROCHA, A.F.; LEÃO, B. F. Calculating the mean knowledge representation from multiple experts. In: FEDREZZI, M.; KACPRZKYK, J. (eds.). **Multiperson decision making models using fuzzy sets and possibility theory**, Amsterdam: Kluwer Academic Publishers, 1990. p. 113-127.
- [MAC 91a] MACHADO, R. J.; DENIS, F. A. R. M. **O modelo conexionista evolutivo**. Rio de Janeiro: IBM Rio Scientific Center, 1991. 97 p. (Technical Report CCR-128)

- [MAC 91b] MACHADO R. J. et al. **Next: the neural expert tool**. Rio de Janeiro: IBM Rio Scientific Center, 1991. 57 p. (Technical Report CCR-120)
- [MAC 92a] MACHADO, R. J. et al. **Combining semantic networks and neural networks in expert systems**. Rio de Janeiro: IBM Rio Scientific Center, 1992. (Technical Report CCR-140)
- [MAC 92b] MACHADO, R. J.; ROCHA, A. F. A Hybrid Architecture for Fuzzy Connectionist Expert Systems. In: KANDEL, A.; LANGHOLZ, G. **Hybrid architectures for intelligent systems**. Boca Raton: CRC Press, 1992. 420 p. p. 135-152.
- [MAT 89] MATTOS, N. M. **An approach to knowledge base management : requirements, knowledge representation and design issues**. Kaiserslautern: Universitaet Kaiserslautern, 1989. 255 p. (Tese de Doutorado)
- [McC 88] McCLELLAND, J. L.; RUMELHART, D. E.; HINTON, G. E. The Appeal of parallel distributed processing. In: RUMELHART, D.E.; McCLELLAND J.L. **Parallel distributed processing**. Cambridge: MIT Press, 1988. v. 1, p. 4-44.
- [MED 92] MEDSKER, L. R.; BAILEY, D. L. Models and Guidelines for Integrating Expert Systems and Neural Networks. In: KANDEL, A.; LANGHOLZ, G. **Hybrid architectures for intelligent systems**. Boca Raton: CRC Press, 1992. 420 p. p. 329-344.
- [MIC 83] MICHALSKI, R. S.; CARBONELL, J. G.; MITCHELL, T. M. **Machine learning : an artificial intelligence approach**. Palo Alto: Morgan Kaufmann, 1983. 572 p.

- [MIL 86] MILLER, R. A. et al. Internist - I: an experimental computer based diagnostic consultant for general internal medicine. In: REGGIA, J. A.; STANLEY, T. **Computer-assited medical decision making**. New York: Springer-Verlag, 1986. v. 2, p. 139-158.
- [MIN 75] MINSKY M. A Framework for representing knowledge. In: WINSTON, P.H. **The Psychology of computer vision**. New York: McGraw-Hill, 1975.
- [NIL 80] NILSSON, N. J. **Principles of artificial intelligence**. Palo Alto, CA: Tioga, 1980. 476 p.
- [OWE 90] OWENS, D. K. et al. Medical decision making: probabilistic medical reasoning. In: SHORTLIFE, E. H. et al. (eds.). **Medical informatics: computer applications in health care**. Reading: Addison-Wesley, 1990. p. 70-116.
- [PEA 88] PEAT F. D. **Artificial intelligence: how machines think**. New York: Ben Publishing Enterprises, 1988. 356 p.
- [RAD 91] RADA, R. **Hypertext: from text to expertext**. London: McGraw-Hill, 1991. 273 p.
- [ROL 88] ROLSTON, D. W. **Artificial Intelligence and expert systems development**. New York: McGraw-Hill, 1988. 257 p.
- [RUM 86] RUMELHART, D. E.; McLELLAND, J. L. **Parallel distributed processing**. Cambridge: MIT Press, 1986. vol. 1, 547 p.
- [SAM 92] SAMAD, T. Hybrid distributed/ local connectionist architectures. In: KANDEL, A.; LANGHOLZ, G. **Hybrid architectures for intelligent systems**. Boca Raton: CRC Press, 1992. 420 p. p. 199-220.

APÊNDICE I

Metodologia de Aquisição de Conhecimento por Grafos de Conhecimento

Fase 0 - Definição

1. Definir o escopo do problema.
2. Definir as hipóteses de diagnóstico que farão parte do projeto.
3. Definir uma lista de sintomas, sinais e resultados de testes, com a assistência de um especialista.

Fase 1 - Aquisição de Conhecimento

1. Selecionar, sem conhecimento prévio do especialista, uma das hipóteses de diagnóstico, se possível em ordem decrescente de frequência de ocorrência.
2. Pedir para o especialista indicar, na lista de sintomas, os itens necessários para a formulação das hipóteses diagnósticas, definindo, assim, um subconjunto de trabalho.
3. Pedir para o especialista ordenar o subconjunto de trabalho de acordo com a importância dos itens para o diagnóstico.
4. Pedir para o especialista assumir os itens da lista ordenada como nodos de evidência de um grafo de conhecimento e associá-los, de forma a tornar possível a conclusão sobre a hipótese diagnóstica (gerando nodos intermediários que convergem para a hipótese de diagnóstico).
5. Pedir para o especialista determinar um grau de importância entre 0 e 10 para a informação representada em cada nodo do grafo.

6. Pedir para o especialista definir os operadores lógicos (E, OU, NÃO) associados aos nodos do grafo.
7. Repetir o processo para outras hipóteses de diagnóstico, formando uma família de grafos.



Informática
UFRGS

Um modelo para sistemas especialistas conexionistas híbridos.

Dissertação apresentada aos Senhores.:

Prof. Dr. Armando Freitas Rocha (UNICAMP)

Profa. Dra. Beatriz de Faria Leão (Inst. de Cardiologia-RS)

Prof. Dr. Luis Otávio Campos Alvares

Prof. Dr. Paulo Martins Engel

Vista e permitida a impressão.
Porto Alegre, 25/08/93.

Prof. Dr. José Mauro Volkmer de Castilho,
Orientador.

Prof. Dr. Ricardo A. da L. Reis,
Coordenador do Curso de Pós-Graduação
em Ciência da Computação.