

CENTRO UNIVERSITÁRIO DA FEI

ADRIANE PAULIELI COLOSSETTI

**CONSTRUÇÃO DE UMA REDE BAYESIANA PARA DIAGNÓSTICO
DA DOENÇA DE ALZHEIMER A PARTIR DE NEUROIMAGEM,
HISTÓRICO E SINTOMAS**

**São Bernardo do Campo
2009**

ADRIANE PAULIELI COLOSSETTI

**CONSTRUÇÃO DE UMA REDE BAYESIANA PARA DIAGNÓSTICO
DA DOENÇA DE ALZHEIMER A PARTIR DE NEUROIMAGEM,
HISTÓRICO E SINTOMAS**

Dissertação de Mestrado apresentada ao Centro
Universitário da FEI como parte dos requisitos
necessários para a obtenção do título de
Mestre em Engenharia Elétrica.
Orientador: Prof. Dr. Paulo Eduardo Santos

São Bernardo do Campo
2009

Colossetti, Adriane Paulieli

Construção de uma rede Bayesiana para diagnóstico da doença de Alzheimer a partir de neuroimagem, histórico e sintomas. / Adriane Paulieli Colossetti – São Bernardo do Campo, 2008.

115f: il.

Dissertação de Mestrado em Engenharia Elétrica – Centro Universitário da FEI.

1. Pesquisa Científica. 2. Método Científico. I. Título.

CDU 001.8

ADRIANE PAULIELI COLOSSETTI

**CONSTRUÇÃO DE UMA REDE BAYESIANA PARA DIAGNÓSTICO
DA DOENÇA DE ALZHEIMER A PARTIR DE NEUROIMAGEM,
HISTÓRICO E SINTOMAS**

Dissertação de Mestrado

Centro Universitário da FEI

Departamento de Engenharia Elétrica

Banca Examinadora

Prof. Dr. Paulo Eduardo Santos – FEI

Prof. Dra. Neli R. Ortega – DIM - FMUSP

Prof^o. Dr^o. Orlando Del Bianco Filho – FEI

São Bernardo do Campo

Dedico este trabalho aos meus queridos pais Odete Paulieli Colosetti e Henrique Colosetti, ao meu noivo Emerson Rodolfo Abraham e à minha irmã Priscila Paulieli Colosetti que me apoiaram desde o início dessa jornada, sendo principais fontes inspiradoras para que eu pudesse concluir esse trabalho.

AGRADECIMENTOS

Se fosse descrever toda minha gratidão a todos que me apoiaram nessa jornada, poderia escrever um livro, pois daria muitas páginas com meu eterno carinho e respeito. Então, de forma sucinta deixo registrado nesta página meu eterno agradecimento aos meus pais **Odete Paulieli Colossetti** e **Henrique Colossetti**, meu noivo **Emerson Rofolfo Abraham** e minha irmã **Priscila Paulieli Colossetti** por todo apoio, amparo, carinho, dedicação e compreensão que dispensaram a mim, para que eu pudesse ter ânimo e determinação de seguir sempre adiante.

Ao meu grande mestre e orientador **Prof. Dr. Paulo Eduardo Santos**, agradeço por todos os ensinamentos, conselhos, dicas, correções, direções que contribuíram muito para realização deste trabalho. Sou grata também por acreditar em mim e em meu trabalho.

Aos professores **Carlos Eduardo Thomaz**, **Flávio Tonidandel** e **Reinaldo Bianch** pelos valiosos ensinamentos transmitidos ao longo dos créditos e tempo de mestrado. Aos meus amigos **Cléber William Gomes** por toda ajuda, troca de informações, conhecimentos compartilhados, companheiro de estudos nos finais de semana; à sua Esposa **Carla Renata Delbue** pela ajuda, paciência, amizade e carinho que me deu ao longo desse trabalho, ao amigo **Alexandre Imperatore** pelo apoio, dedicação e ajuda sempre que necessário, ao amigo **Nelson Augusto Oliveira de Aguiar** pela grande ajuda nas análises univariada e multivariada, que com seu conhecimento, boa vontade e bom coração me ajudou nessa fase do meu trabalho também.

Agradeço ao Professor Orlando Del Bianco Filho pelas dicas, apoio e por acreditar em mim desde o principio do mestrado. Estendo ainda meus agradecimentos à Professora Neli R. Ortega que agregou muito na finalização deste trabalho.

Enfim agradeço a todos os amigos e familiares que acreditaram em mim e me apoiaram sempre, entendendo minhas ausências e falta de tempo total.

A **Adriana** e **Ricardo** da secretaria do mestrado e a **todos** que colaboraram direta e indiretamente para que este trabalho fosse concluído.

A todos vocês os meus sinceros e eternos agradecimentos!

“Um pouco de ciência nos afasta de Deus.
Muito, nos aproxima”.

Louis Pasteur.

RESUMO

A perda progressiva da compreensão e da memória, desorientação, ações imprevisíveis, alteração no comportamento e problemas com a linguagem são alguns dos principais sintomas da doença de Alzheimer (DA), considerada a mais comum das demências, cujo diagnóstico só pode ser confirmado após a morte, dificultando de certa forma o diagnóstico mais preciso ao paciente. Visando uma análise probabilística na análise diagnóstica referente a doenças psiquiátricas, mais especificamente a DA, este trabalho tem por objetivo lidar com as incertezas através de técnicas probabilísticas no contexto de redes bayesianas (RB), investigando o uso de neuroimagem e também dados clínicos, como: histórico e sintomas apresentados pelo paciente. Esta dissertação apresenta uma comparação entre dois grupos, sendo 14 casos considerados controles (pessoas que não apresentam sinais de DA) e 14 diagnosticada com DA. Devido a flexibilidade da RB pôde-se construir a rede com base em dados captados na literatura e por meio de consulta com especialista, além dos 28 casos concedidos pela faculdade de medicina de São Paulo (USP). Foi apresentada uma comparação entre as imagens por meio de análises estatísticas univariada e multivariada sendo encontrado diferenças entre imagens cerebrais das pessoas previamente diagnosticadas com DA e pessoas controles. A partir desse resultado utilizamos os dados relevantes como filhos da variável neuroimagem (nó). Os demais nós da rede: histórico e sintomas complementaram a rede com informações da literatura e também do especialista. Foi observado que os resultados apresentaram valores coerentes e assim pode-se concluir que a rede teve o funcionamento satisfatório, embora deve-se melhorá-la com valores e pesquisas mais detalhadas sobre DA; discute-se ainda a possibilidade de usá-la em ambiente real, considerando a obtenção de valores probabilísticos mais fidedignos, baseados por exemplo em meta-análises.

Palavras-chave: Rede Bayesiana, doença de Alzheimer, Noisy-OR, Neuroimagem, Histórico, Sintomas.

ABSTRACT

The progressive loss of memory and understanding, disorientation, unpredictable actions, changes in behavior and problems with language are some of the major symptoms of Alzheimer's disease (AD), considered the most common kind of dementia, whose diagnosis can only be confirmed after the death, making it difficult in some ways the most accurate diagnosis for the patient. Aiming a probabilistic analysis in diagnostic analysis relating to psychiatric disorders, specifically the DA, this study aims to deal with the uncertainties through probabilistic techniques in the context of Bayesian networks (BN), investigating the use of neuroimaging and clinical data, as: history and symptoms presented by the patient. This thesis presents a comparison between two groups, with 14 cases considered controls (people who do not show signs of AD) and 14 diagnosed with AD. Because of the flexibility of the RB we were able to build a network based on data collected in the literature and through consultation with specialists, in addition to the 28 cases granted by the faculty of medicine of São Paulo (USP). It was presented a comparison of the images by means of univariate and multivariate statistical analysis and found differences between brain images of people previously diagnosed with AD and controls people. From this result it was used the relevant data as children of the variable neuroimaging (node). The other network nodes: history and symptoms complemented the network with information from literature and also of Specialist. It was observed that the results shown consistent values and so it can be concluded that the network had the satisfactory functioning, although it should be improving it with values and more detailed research on AD; It also discusses the possibility of using it in real environment, considering the acquisition of probabilistic values more reliable, for example based on meta-analysis.

Keywords: Bayesian Network, Alzheimer's disease, Noisy-OR, Neuroimaging, history, symptoms.

SUMÁRIO

1	INTRODUÇÃO	17
2	MAL DE ALZHEIMER	19
2.1	CAUSAS DA DOENÇA DE ALZHEIMER.....	20
2.2	SINTOMAS DA DOENÇA DE ALZHEIMER.....	21
2.3	ESTÁGIOS DA DOENÇA DE ALZHEIMER	21
2.4	DIAGNÓSTICO EM DOENÇA DE ALZHEIMER	22
2.5	CAMDEX – ENTREVISTA ESTRUTURADA.....	24
3	INTELIGÊNCIA ARTIFICIAL NA MEDICINA.....	25
3.1	REDES BAYESIANAS	27
3.2	SISTEMA JAVABAYES.....	32
3.3	MODELO NOISY-OR	35
3.4	SISTEMAS BASEADOS EM REDES BAYESIANAS	37
3.4.1	SISTEMAS ESPECIALISTAS BASEADOS EM REDES BAYESIANAS	39
4	DIAGNÓSTICO POR NEUROIMAGEM.....	50
4.1	BANCO DE IMAGENS.....	55
4.2	EXPERIMENTO UTILIZANDO STATISTICAL PARAMETRIC MAPPING (SPM2) PARA ANÁLISE UNIVARIADA E MULTIVARIADA	55
4.2.1	ANÁLISE UNIVARIADA	55
4.2.2	ANÁLISE MULTIVARIADA – MAXIMUM UNCERTAINTY LINEAR DISCRIMINANT ANALYSIS (MLDA)	65
5	CONSTRUÇÃO E ANÁLISE DA RB	75
5.1	REDE BAYESIANA COM NOVAS DISTRIBUIÇÕES DE PROBABILIDADE....	85
5.2	AVALIAÇÃO DE DESEMPENHO	92
5.2.1	NOVOS TESTES E COMPARAÇÕES COM DIAGNÓSTICOS DADOS POR ESPECIALISTA	94
6	CONCLUSÕES E TRABALHOS FUTUROS.....	98
7	REFERÊNCIAS BIBLIOGRÁFICAS	102
8	APÊNDICE 1 – PROGRAMA JAVABAYES DA REDE BAYESIANA PARA DOENÇA DE ALZHEIMER.....	110

9	APÊNDICE 2 - TRABALHO SOBRE VISÃO COMPUTACIONAL RELACIONADO À DOENÇA DE ALZHEIMER REALIZADO DURANTE OS CRÉDITOS DESTE MESTRADO	115
---	--	-----

LISTA DE FIGURAS

Figura 3.1: Modelo de grafo acíclico.	28
Figura 3.2: Exemplo de diagnóstico médico	30
Figura 3.3: Representação gráfica de uma RB do problema Doença A.....	32
Figura 3.4: Resultado da inferência realizada representada pela figura 3.3.	33
Figura 3.5: Resultado da inferência realizada representada pela figura 3.3.	34
Figura 3.6: Representação gráfica de uma RB do problema Doença A.....	34
Figura 3.7: MSBNx - Bayesian Network Editor and Toolkits, adaptado de [MSBNx 2008].	38
Figura 3.8: Topologia da RB como modelo do problema das doenças de sono. Adaptado de [MILHO 2000].....	39
Figura 3.9: Páginas da internet do sistema DiagSD, arquivo para ação do diagnóstico e página produzida com os resultados de diagnóstico. Adaptado de [MILHO 2000].	41
Figura 3.10: Tabela de ilustração do sistema SEAMED. Adaptado de [PEROTTO 2001].	42
Figura 3.11: Exemplo de entrada de dados no programa SEAMED. Adaptada de [PEROTTO 2001].	43
Figura 3.12: Topologia da RB de insuficiência cardíaca. Adaptado de [SAHEKI 2003].	44
Figura 3.13: Software iBNetz. Adaptado de [SAHEKI 2003].	45
Figura 3.14: Topologia da RB do Sistema de Avaliação Nutricional. Adaptado de [TIBIRIÇÁ 2006].	46
Figura 3.15: Variáveis de entrada utilizadas no módulo de simulação. Adaptado de [TIBIRIÇÁ 2006].	47
Figura 3.16: Resultado do módulo de simulação. Adaptado de [TIBIRIÇÁ 2006].	47
Figura 3.17: Valores de entrada de uma avaliação completa. Adaptado de [TIBIRIÇÁ 2006].	48
Figura 3.18: Resultado de uma avaliação completa. Adaptado de [TIBIRIÇÁ 2006].	48
Figura 4.1: Efeito relativo ao cérebro dos pacientes que tem a DA. (Adaptado de [AGUIAR 2006]).	51
Figura 4.2: Cérebro humano normal (Adaptado de [AGUIAR 2006]).	53

Figura 4.3: Cérebro humano com DA (Adaptado de [AGUIAR 2006]).	53
Figura 4.4: Cérebro humano normal (Adaptado de [AGUIAR 2006]).	54
Figura 4.5: Cérebro humano com DA (Adaptado de [AGUIAR 2006]).	54
Figura 4.6: Segmentação, adaptado de [AGUIAR 2006].	57
Figura 4.7: Resultado SPM2 do Liquor mostrando as coordenadas das regiões cerebrais, na proporção controle x paciente (1, -1)	59
Figura 4.8: Resultado SPM2 do Liquor mostrando as coordenadas das regiões cerebrais, na proporção paciente x controle (-1, 1)	60
Figura 4.9: Resultado SPM2 da substância cinza mostrando as coordenadas das regiões cerebrais, na proporção controle x paciente (1, -1)	61
Figura 4.10: Resultado SPM2 da substância cinza mostrando as coordenadas das regiões cerebrais, na proporção paciente x controle (-1, 1)	62
Figura 4.11: Resultado SPM2 da substância branca mostrando as coordenadas das regiões cerebrais, na proporção controle x paciente (1, -1)	63
Figura 4.12: Resultado SPM2 da substância branca mostrando as coordenadas das regiões cerebrais, na proporção paciente x controle (-1, 1)	64
Figura 4.13: Ilustração dos três primeiros autovetores adaptado de [AGUIAR 2006].	66
Figura 4.14: Autovetor PCA + MLDA. Adaptado de [AGUIAR 2006].	67
Figura 4.15: Ventrículo. Adaptado de [AGUIAR 2006].	68
Figura 4.16: Ventrículo contornado para melhor visualização. Adaptado de [AGUIAR 2006].	68
Figura 4.17: Hipocampo. Adaptado de [AGUIAR 2006].	69
Figura 4.18: Hipocampo contornado para melhor visualização. Adaptado de [AGUIAR 2006].	69
Figura 4.19: Giro do cíngulo contornado para melhor visualização. Adaptado de [AGUIAR 2006].	70
Figura 4.20: Córtex parietal. Adaptado de [AGUIAR 2006].	71
Figura 4.21: Hipocampo visualizado na análise univariada e multivariada de paciente e controle. Adaptado de [AGUIAR 2006].	72
Figura 4.22: Córtex visualizado na análise univariada e multivariada de paciente e controle. Adaptado de [AGUIAR 2006].	73

Figura 4.23: Giro do cingulo visualizado na análise univariada e multivariada de paciente e controle. Adaptado de [AGUIAR 2006].	74
Figura 5.1: Representação gráfica da rede – Alzheimer.	79
Figura 5.2: Representação gráfica da rede com observações descritas na tabela 5.9.	80
Figura 5.3: Resultado das inferências realizadas representada pela figura 5.2.	81
Figura 5.4: Representação gráfica da rede com observações descritas na tabela 5.10.	82
Figura 5.5: Resultado das inferências realizadas representada pela figura 5.4.	83
Figura 5.6: Representação gráfica da rede com observações descritas na tabela 5.11.	83
Figura 5.7: Resultado das inferências realizadas representada pela figura 5.6.	84
Figura 5.8: Representação gráfica da rede com observações descritas na tabela 5.10.	87
Figura 5.9: Resultado das inferências realizadas na RB representada na figura 5.11.	88
Figura 5.10: Representação gráfica da rede com observações descritas na tabela 5.18.	88
Figura 5.11: Resultado das inferências realizadas na RB representada pela figura 5.10.	89
Figura 5.12: Representação gráfica da rede com observações descritas na tabela 5.12.	90
Figura 5.13: Resultado das inferências realizadas representada pela figura 5.12.	91
Figura 5.14: Comparação dos resultados entre as variáveis “compree”, “recente”, “remota”, “praxia”, “escolar” e “idade”.	93
Figura A2.1: Imagens iniciais usadas como base p/ preenchimento e contagem de pixels vermelhos.	116
Figura A2.2: Imagens com a área ventricular preenchida com pixels vermelhos.	116

LISTA DE TABELAS

Tabela 3.1: Probabilidade Conjunta	30
Tabela 3.2: Tabela de probabilidade conjunta $P(A, B, C)$	31
Tabela 3.3: Tabela de distribuição de probabilidades	36
Tabela 4.1: Tabela de comparações das regiões cerebrais entre as análises univariada e multivariada. Adaptado de [AGUIAR 2006].	67
Tabela 5.1: Distribuição de relevância para o nó Histórico	76
Tabela 5.2: Noisy-or aplicado aos nós: Histórico, Sintomas e Neuroimagem (nós filhos do nó Alzheimer).	77
Tabela 5.3: Noisy-or aplicado aos nós: Cíngulo Posterior, Hipocampo, Ventrículo e Córtex (nós filhos do nó Neuroimagem).	77
Tabela 5.4: Noisy-or aplicado aos nós: Afasia, Perda_de_memória e Apraxia (nós filhos do nó Sintomas).	78
Tabela 5.5: Noisy-or aplicado aos nós: Hereditariedade, Escolaridade e Idade (nós filhos do nó Histórico).	78
Tabela 5.6 : Distribuição das observações feitas nas variáveis destacadas em azul, conforme figura 5.2.	80
Tabela 5.7: Distribuição das observações feitas nas variáveis destacadas em azul, conforme figura 5.4.	82
Tabela 5.8: Distribuição das observações feitas nas variáveis destacadas em azul, conforme figura 5.6.	84
Tabela 5.9: Distribuição de relevância para os nós Neuroimagem, Sintomas, Histórico, Cíngulo Posterior, Hipocampo, Ventrículo, Córtex, Apraxia, Afasia, Perda de memória, Hereditariedade, Idade e Escolaridade.	86
Tabela 5.10: Distribuição das observações feitas nas variáveis destacadas em azul, conforme figura 5.8.	87
Tabela 5.11: Distribuição das observações feitas nas variáveis destacadas em azul, conforme figura 5.13.	89
Tabela 5.12: Distribuição das observações feitas nas variáveis destacadas em azul, conforme figura 5.12.	91

Tabela 5.13: Distribuição das probabilidades para $P(\text{neuroimagem} \mid \text{cíngulo_posterior}, \text{hipocampo}, \text{ventrículo e córtex})$	95
Tabela 5.14: Distribuição das probabilidades para $P(\text{Alzheimer} \mid \text{Histórico}, \text{Sintomas}, \text{Neuroimagem})$	96
Tabela A2.1: Tabela com resultado da figura A2.2.	117

1 INTRODUÇÃO

Em se tratando da ciência de diagnóstico é natural que quanto mais preciso e objetivo for o resultado de uma avaliação sobre determinada doença, mais eficaz será sua prevenção ou tratamento. Entretanto, devido à alta complexidade em função do número de variáveis envolvidas nas doenças, torna-se essencial o uso de ferramentas computacionais para o diagnóstico médico. Ao longo dos anos a tecnologia tem apresentado avanços significativos em todos os setores, inclusive na área médica, gerando dessa forma investigações diagnósticas altamente dinâmicas.

Uma das formas de auxiliar a ciência diagnóstica é a aplicação da estatística, que visa analisar e compreender as informações com o fito de desenvolver métodos cada vez mais eficientes sob dados imprecisos [KLARMANN 1996].

Nossa proposta é valer-se dos axiomas de probabilidade, em um modelo de grafos acíclicos direcionados (Redes Bayesianas - RB), para desenvolver um sistema de diagnóstico automatizado para doença de Alzheimer (DA). RBs representam relações de independências condicionais para consultas eficientes a respeito da probabilidade de um determinado evento [ZHANG 1996]. A aplicabilidade de um sistema computacional de redes Bayesianas é eficaz por tratar-se de uma tecnologia fácil de extrair o conhecimento e que possibilita a inserção de novas variáveis como nós na rede, e flexível no trato de informações por permitir uma continuidade acíclica ao abranger várias variáveis conectadas e independentes ao mesmo tempo, tornando-se uma opção adequada para equacionar o problema de diagnóstico médico automático. Deste modo resolvemos desenvolver um estudo aplicado à doença de Alzheimer tendo como fatos as características da doença dado um conjunto de sintomas (dados clínicos), histórico pessoal e familiar, além de resultados de exames realizados com o paciente (neuroimagem), visando construir um sistema de auxílio ao diagnóstico da doença.

Portanto, o objetivo deste trabalho é investigar o uso de neuroimagem juntamente com os dados clínicos, tais como: histórico do paciente e também os sintomas em um sistema de diagnóstico automático em redes Bayesianas.

Este trabalho é organizado da seguinte forma: o capítulo 2 apresenta as principais características da doença de Alzheimer, suas possíveis causas, sintomas e estágios da doença. O Capítulo 3 apresenta a importância da inteligência artificial na medicina e descreve um de seus métodos chamado redes Bayesianas, mostrando alguns exemplos de aplicabilidade da rede. O capítulo 4 refere-se aos diagnósticos feitos através de exames de neuroimagem e como esses dados integram-se em nosso trabalho. O capítulo 5 apresenta os resultados obtidos através dos experimentos realizados em uma rede Bayesiana construída com base em histórico, sintomas e dados extraídos de exames de ressonância magnética da região cerebral. Por fim, o capítulo 6 apresenta a conclusão desta dissertação e discute trabalhos futuros.

2 MAL DE ALZHEIMER

O termo “Mal de Alzheimer” teve origem em 1901 quando uma senhora chamada August D. foi internada no hospital de Frankfurt, Alemanha, sob os cuidados do Dr. Alois Alzheimer, médico neurologista [MAURES 1997]. Ela apresentava um conjunto de sintomas intrigantes, tais como: redução da compreensão e da memória, desorientação, ações imprevisíveis, paranóia e alucinações auditivas que fez com que Alzheimer investigasse mais detalhadamente o caso. Essa paciente teve óbito em 08 de abril de 1906 e sua aparência era de posição de decúbito, apresentava também ulceração, fraqueza, febre alta e pneumonia em ambos os pulmões inferiores. A *causa mortis* foi dada como septicemia e no diagnóstico da anatomia patológica encontraram hidrocefalia moderada, atrofia cerebral, arteriosclerose das veias cerebrais, pneumonia e nefrite. Em novembro desse mesmo ano, Alzheimer relatou esse caso em um congresso de psiquiatria apoiando-se nos achados dos exames histopatológicos da doença. Entre os achados podemos citar alterações peculiares nas fibras neurais, retratando mudanças que representavam formas mais sérias de demência senil (em função de placas senis excessivamente numerosas) e um terço das células corticais mortas, deixando em seus lugares manchas profundas.

Essa doença é conhecida também como demência senil tipo Alzheimer e é a mais comum das demências, seu diagnóstico só pode ser confirmado após a morte, dificultando dessa forma um diagnóstico mais preciso ao paciente. Atinge homens e mulheres sem distinção de raças e classes sociais.

Alzheimer faleceu em 19 de dezembro de 1915, porém, seu nome é conhecido no mundo inteiro devido à doença que acomete cerca de 10 e 15% das pessoas da 3ª idade¹ [MAURES 1997]. Em outro estudo constata-se que uma em cada dez pessoas maior de 80 anos, uma em cada cem pessoas maior de 70 anos e uma em cada mil pessoas maior de 60 anos são acometidas pela doença em questão [MAURES 1997].

2.1 CAUSAS DA DOENÇA DE ALZHEIMER

Através de estudos realizados, acredita-se que a doença de Alzheimer (DA) surge por várias causas, entre elas: baixo grau de escolaridade² (quanto menos escolaridade maior é a chance de apresentar a DA); idade avançada, geralmente acima de sessenta e cinco anos (a probabilidade de desenvolver DA dobra a cada cinco anos após essa idade); hereditariedade; arteriosclerose; e traumatismo craniano. Há também evidências físicas observadas através da necropsia, tais como: danos cerebrais, atrofias em muitas regiões do cérebro envolvendo estruturas neurais perceptivas, motoras e associativas (córtex, hipocampo lateral, estrutura temporal inferior, anterior e a faixa posterior), características confusas das placas amilóides e da massa neuro-fibrilar, diminuição da concentração de N-acetilaspártato (NAA) e aumento de mio-inositol (MI) na formação do hipocampo, hipoperfusão amigdaló-hipocampal, redução bilateral e frequentemente assimétrica do fluxo sanguíneo e do metabolismo em regiões temporais, deficiência do neurotransmissor acetilcolina e alargamento dos ventrículos [DAVATZIKOS 2006; CHARCHAT 2001; NITRINI 2005; GREENFIELD 2000; AGUIAR 2006; PASCALE 2007].

¹.Varella, Dráuzio. **Alzheimer.** Disponível em: http://drauziovarella.ig.com.br/arquivo/arquivo.asp?doe_id=42. Acessado em 13/10/2008.

².Segundo especialista, a pessoa que tem um maior nível de estudo também pode desenvolver esta doença, porém como tem uma grande rede de sinapses desenvolvida ao longo dos anos de estudo, esta pessoa não percebe precocemente esta doença, pois demora a apresentar os sintomas característicos de DA.

2.2 SINTOMAS DA DOENÇA DE ALZHEIMER

Um dos principais sintomas da DA é a perda de memória. Existem alguns tipos de memórias mais afetadas como, por exemplo, a memória episódica que responde pelo armazenamento de informações relativas aos episódios da vida; a memória de curto prazo, de tal forma que as pessoas que desenvolvem a DA começam a esquecer as lembranças mais recentes; a memória semântica que se refere basicamente à informação relativa ao significado das palavras; e, por fim, a memória de procedimento que responde pela condução dos atos, seja mental ou físico. A perda da memória de procedimento acarreta na dificuldade de cumprir tarefas diárias³ [DAVATZIKOS 2006; CHARCHAT 2001; DAMASCENO 1999; MCCRONE 2002; PASCALE 2007; KANDEL 2000].

São características comuns da DA: a apraxia, que é a perda da capacidade de efetuar movimentos; a afasia, ou seja, a perda da compreensão da linguagem, tanto falada, como escrita; e a agnosia, quando o paciente não consegue mais reconhecer os objetos e nem saber quais suas utilidades. De forma geral, o paciente de Alzheimer também apresenta mudanças de personalidade e de comportamento.

2.3 ESTÁGIOS DA DOENÇA DE ALZHEIMER

A DA é dividida basicamente em três estágios que foram descritos de acordo com os sintomas e o tempo de evolução que os pacientes apresentavam. Abaixo citamos alguns dos sintomas clínicos de cada estágio ou fase⁴ [CONVIVENDO 2007; CAOVIALLA 2002; DAMASCENO 1999; OLIVEIRA 2005].

³. **Convivendo com Alzheimer** – www.alzheimer.med.br/alz1.htm - acessado em 11/04/07.

⁴. **Alzheimer's Association. Alzheimer's disease: 2006**. Disponível em: <http://www.alz.org/>, acessado em 05/06/2007.

Em sua fase inicial (dois a quatro anos precedendo o diagnóstico), a pessoa geralmente apresenta certa distração, dificuldade de lembrar nomes de pessoas, palavras e números de telefones. Apresenta ainda, dificuldade de aprender informações novas e certa redução das atividades sociais. Tais ocorrências são tão sutis que os familiares e amigos quase nunca percebem.

Na fase intermediária (dois a dez anos após o diagnóstico) inicia-se o agravamento dos quadros anteriores afetando mais a memória, sendo que os acontecimentos mais recentes passam a ser esquecidos pelo paciente. Ocorre perda da atividade cognitiva, passando o paciente a não reconhecer a própria família, perdendo-se com facilidade e apresentando tendência a fugir ou perambular pela casa demonstrando também alteração no comportamento (impaciência, inquietação, alucinações, delírios, entre outros).

Na fase final a pessoa perde completamente as funções cognitivas; acentua-se o problema com a linguagem, e a comunicação torna-se monossilábica até cessar por completo. O doente passa a ser totalmente dependente de terceiros, pois este não controla a bexiga ou o intestino e as articulações tornam-se enrijecidas subtraindo a capacidade de se sentar, andar e até engolir, muitas vezes precisando fazer uso de sonda para alimentar-se [ALZHEIMER'S 2007].

2.4 DIAGNÓSTICO EM DOENÇA DE ALZHEIMER

Atualmente os métodos comumente utilizados para o diagnóstico de DA envolvem exames de Ressonância Magnética (RM), PET (Positron Emission Tomography) e SPECT (Single – Photon Emission Computed Tomography), EletroEncefalograma (EEG), EletroEncefalograma Quantitativo (EEGQ), além de entrevistas estruturadas [VENTURA 2007].

No ano de 2003 houve uma movimentação intensa envolvendo cientistas procurando levantar as principais normas, condutas ou sugestões para o diagnóstico clínico de DA no Brasil [NITRINI 2005]. Para tanto, foram utilizados diversos trabalhos

científicos, incluindo artigos originais, teses e dissertações a fim de unificar as diversas informações referentes à DA.

Os cientistas designados na IV Reunião de pesquisadores em DA e desordens relacionadas [NITRINI 2005] pesquisaram em bancos de dados importantes e respeitados na área médica, como: o banco de dados da PUBMED (site desenvolvido pela NLM (*National Library of Medicine*), de grande importância para o profissional da saúde e auxilia a pesquisa bibliográfica que tem como base a bibliografia da MedLine⁵); Critérios diagnósticos do NINCDS – ADRDA, APA [MARQUES 2000; PSYCHIATRIC 1987; PSYCHIATRIC 1994] (Associação Psiquiátrica Americana em suas duas versões – DSM-III-R e DSM-IV); Critérios da organização mundial de saúde (CID-10 (Classificação Internacional de Doença)) [PSYCHIATRIC 1994]; CAMDEX (entrevista estruturada válida para auxiliar o diagnóstico diferencial de demência e depressão em pessoas idosas) [WHO 1993; HOLLAND 1998] e Diagnóstico Clínico. Esses critérios e normas são classificações internacionais usadas pelos especialistas para identificar e diagnosticar as doenças. Geralmente dos consensos e as normas citadas acima recomendam mais frequentemente os critérios da APA, na versão IV (DSM-IV) [NITRINI 2005], o qual é utilizado na prática para o diagnóstico de demências e os da NINCDS – ADRDA mais especificamente para o diagnóstico de DA.

De acordo com os critérios diagnósticos da NINDS-ADRDA (*Diagnostic criteria for Alzheimer's Disease (AD)*) a doença de Alzheimer é caracterizada pela perda da memória episódica (por relato da pessoa ou ainda por meio de informações superiores a seis meses), atrofia do lóbulo temporal (perda do volume hipocampal, córtice entorhinal, amígdala) evidenciados em ressonância magnética. Há ainda a anormalidade do líquido cérebro-espinhal, baixa concentração de amiloide β 1-42, aumento das proteínas tau ou phospho-tau, deficiência cognitivas como afasia e apraxia [NITRINI 2005].

⁵. Medline. **Medical Literature Analysis and Retrieval System Online**. Disponível em: <http://www.nlm.nih.gov/pubs/factsheets/medline.html>. Acessado em 13/10/2008.

2.5 CAMDEX – ENTREVISTA ESTRUTURADA

CAMDEX é um instrumento utilizado em todo o mundo com o objetivo de auxiliar o diagnóstico de transtornos mentais em indivíduos idosos [HOLLAND 1998]. Através dele podemos destacar algumas evidências que colocaremos em nossa rede para auxiliar no diagnóstico mais preciso da DA. É importante destacar que não tivemos acesso ao questionário CAMDEX, mas somente aos dados tabulados para trabalharmos como variável do problema escolhido. CAMDEX é vendido para profissionais da área da medicina. Conforme veremos na seção 5.2, o CAMDEX não foi muito informativo para este trabalho.

Esse questionário refere-se aos indivíduos dos quais foram obtidas as 28 imagens estudadas neste trabalho e foi cedido pelo Instituto de Psiquiatria da Faculdade de Medicina da USP.

Deste questionário ressaltamos as variáveis: idade, apraxia, afasia e escolaridade citadas no capítulo 2 que as destaca como as principais características das pessoas que desenvolve a DA.

Percebemos uma média de 70 anos de idade, sendo que os casos estudados foram a partir dos 62 anos na pessoa com menor idade e 82 anos na pessoa com maior idade.

Para este trabalho utilizamos os resultados dos exames físicos, como RM, o CAMDEX e diagnósticos clínicos que servem de embasamento para criação de novas tecnologias valendo-se de diversos métodos computacionais e, em particular, métodos de Inteligência Artificial (IA), conforme veremos no próximo capítulo.

3 INTELIGÊNCIA ARTIFICIAL NA MEDICINA

Os primeiros programas construídos em Inteligência Artificial (IA) para a medicina recebiam como entrada de dados uma grande lista de sintomas e geravam como saída uma lista com diagnósticos possíveis [RAZZOUK 1995].

Hoje em dia, com os métodos de IA, podemos construir programas (agentes inteligentes) que são capazes de selecionar a melhor solução e executar a ação que achar mais adequada para as situações reais, levando em consideração incertezas nas variáveis do domínio [RAZZOUK 1995].

IA pode auxiliar no desenvolvimento de ferramentas mais precisas devido à capacidade que máquinas (computadores) possuem para armazenar todos os dados e gerar informações rapidamente e com isso ganhar uma eficácia na identificação de um diagnóstico ou decisão terapêutica.

Encontramos muitas técnicas de IA em diversas áreas de estudo, inclusive em psiquiatria, no desenvolvimento de sistemas especialistas (*expert systems*) que reproduzem o raciocínio de um especialista no que diz respeito à solução de alguns problemas complexos. Basílio de Bragança Pereira [PEREIRA 2001] citou o uso dos sistemas especialistas para o auxílio ao diagnóstico em que psiquiatras desenvolveram padrões de avaliação, algoritmos e escalas de pontuação para um diagnóstico mais preciso de certas doenças psiquiátricas.

Sistemas especialistas são considerados de grande flexibilidade, pois podem ser alterados a qualquer momento, acrescentando ou eliminando informações sem perder a eficácia. Eles também apresentam uma ilustração de todas as variáveis relevantes de um domínio aos especialistas, o que os tornam úteis para auxiliar na decisão do diagnóstico e na terapêutica a ser aplicada.

Os sistemas inteligentes são desenvolvidos a partir de uma base de conhecimento constituída por um banco de dados onde está armazenado o conhecimento baseado em experiências e teorias. Faz parte destes sistemas uma máquina de inferência que, a partir das informações da base de conhecimento, executa

passos de prova com a finalidade de se obter resposta a alguma pergunta (*query*) [RAZZOUK 1995].

A idéia básica para desenvolver um sistema inteligente é a identificação de regras ou padrões aplicados por especialistas, partindo desse princípio citamos os trabalhos desenvolvidos por Razzouk [RAZZOUK 1995, 2000, 2003, 2006 e 2007].

Razzouk [RAZZOUK 1995] elaborou uma pesquisa em que foi solicitado a três especialistas em psiquiatria (de dois centros Universitários de São Paulo), reconhecer sinais e sintomas típicos presentes e ausentes da esquizofrenia em três casos clínicos apresentadas. A partir disso foi solicitado aos especialistas que elaborassem grafos representando como eles chegaram à decisão de diagnóstico. Depois foram atribuídas probabilidades para cada sintoma de acordo com a especialidade e importância para decisão do diagnóstico.

Tomando como exemplo os trabalhos [RAZZOUK 1990, 1995, 2000, 2003, 2006, 2007 e LEÃO 1990] por se tratar de construção de um sistema especialista em IA e também por ser aplicado a uma patologia psiquiátrica, podemos considerar as etapas da construção do sistema que se resumem em adquirir o conhecimento, organizá-lo, modelá-lo e avaliá-lo em sua performance, para segui-las como diretrizes em nosso trabalho.

Os sistemas desenvolvidos para área da psiquiatria, de um modo geral, podem auxiliar o especialista no processo de tomada de decisão do diagnóstico, pois levam em consideração as inúmeras variáveis que envolvem a doença de maneira rápida e eficaz pela capacidade de processamento computacional.

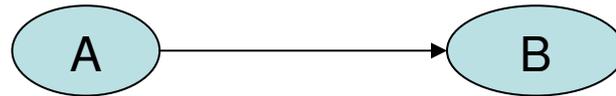
Existem diversas formas para desenvolver um sistema especialista, porém neste trabalho escolhemos rede Bayesiana (RB), pela facilidade de trabalhar com incertezas, por gerar explicações sobre a seleção indicada, e outras funcionalidades que mostraremos na próxima seção.

3.1 REDES BAYESIANAS

Segundo [RUSSELL E NORVIG 1995; MARQUES 2000; IDE 2001] redes Bayesianas (RB) são grafos acíclicos, dirigidos, que representam dependências entre variáveis representadas por nós em que podemos atribuir valores conforme o grau de relevância em um modelo probabilístico. Esta abordagem representa uma boa estratégia para lidar com problemas que tratam incertezas, onde conclusões não podem ser construídas apenas do conhecimento prévio a respeito do problema. Podemos afirmar que a principal vantagem de usar o raciocínio probabilístico sobre raciocínio lógico é que os agentes podem tomar decisões racionais mesmo quando há pouca informação para provar alguma ação [CHARNIAK 1991].

Matematicamente RBs são representações compactas de tabelas de conjunção de probabilidades do universo do problema, equacionadas por uma fórmula pela qual se pode calcular a probabilidade de um evento dadas algumas informações iniciais (probabilidade a priori ou incondicional) e algumas informações adicionais (probabilidade posterior ou condicional).

Deste modo, podemos dizer que uma RB é a representação compacta das variáveis do universo de um problema, através de um modelo probabilístico. Tal modelo pode estar associado a ferramentas computacionais para modelar informações em um contexto probabilístico. Tais modelos consideram a dependência entre as variáveis do problema e produzem uma dedução perante as incertezas existentes a partir do confronto de duas ou mais explicações possíveis para a ocorrência de um evento. RB é representada por grafos dirigidos sem ciclos (DAG) conforme é ilustrado na figura 3.1; as variáveis do problema são nós da rede e as dependências probabilísticas são os arcos ligando nós.



A representa causa de B

Figura 3.1: Modelo de grafo acíclico.

Redes Bayesianas permitem que quaisquer variáveis possam ser consideradas como pergunta (*query*) ou evidência (*evidence*) permitindo quatro tipos de inferências: diagnóstico (de efeito para causa); causal (de causas para evidências); inter-causal (entre causas de um efeito comum) e mista (combina duas ou mais das anteriores) [MILHO 2000].

Antes de adentrarmos nas explicações sobre as representações de interface gráfica [JAVABAYES 2007] com modelos explicativos sobre RB, é importante explanarmos um pouco sobre o fundamento matemático que envolve nosso objeto de estudo.

Na fórmula seguinte mostra-se como calcular a probabilidade de ocorrer os eventos A e B simultaneamente e a partir dela deduz-se a regra de Bayes.

$$P(A \wedge B) = P(A|B)P(B) = P(B|A)P(A)$$

A probabilidade de ocorrer tanto o evento A como o evento B é igual a probabilidade do evento A dado B (ou B dado A) vezes a probabilidade a priori de B (ou A).

A equação geral da regra de Bayes que os sistemas modernos de IA usam como base para inferência probabilística é a seguinte [RUSSELL E NORVIG 1995]:

$$P(A|B) = \frac{P(B|A) P(A)}{P(B)}$$

Para a aplicação da regra acima são exigidos três termos: uma probabilidade condicional $P(B|A) = Z$, onde dado o evento “A” a probabilidade do evento “B” é “Z” e

duas probabilidades incondicionais (a priori): $P(A)$ e $P(B)$. Considerando que probabilidade a priori é o nome que se dá às outras variáveis do mesmo domínio.

Já a regra da soma mostrada a seguir apresenta a probabilidade de ocorrer o evento A ou o evento B , sendo esta a probabilidade calculada pela soma das probabilidades a priori deles mesmos menos a probabilidade dos dois eventos ocorrerem simultaneamente:

$$P(A \vee B) = P(A) + P(B) - P(A \wedge B)$$

A propriedade básica de uma RB é que toda variável X_i tem que ser independente de todos os antecessores de X_i , dado os filhos e os pais dos filhos X_i . Esta é a hipótese da independência entre variáveis, que é base para construção de RB. Portanto, na RB para diagnóstico de Alzheimer a independência é assumida (por construção). Observamos a seguinte expressão [IDE 2001]:

$$P(X) = \prod_i p(X_i | pa\{X_i\}), \text{ onde } pa\{X_i\} \text{ representa os pais de } X_i.$$

Para calcular a distribuição condicional de uma variável qualquer X_q , dada as evidências (observações, representadas em E), temos:

$$P(X_q | E) = p(X_q \wedge E) / p(E).$$

E a expressão geral para computar RBs é:

$$P(X_q | E) = \sum_{\{X/X_q, E\}} [\prod_i p(X_i / pa \{X_i\})] / \sum_{\{X/E\}} [\prod_i p(X_i / pa\{X_i\})].$$

Para demonstrar a aplicação de uma Rede Bayesiana em diagnóstico médico, construímos um exemplo com três variáveis para melhor entendimento e explanação. Na figura 3.2 ilustramos uma rede onde queremos saber a probabilidade de uma pessoa ter uma determinada doença ou o exame de Liquor dar positivo ou negativo dado que ela apresenta ou não dor de cabeça.

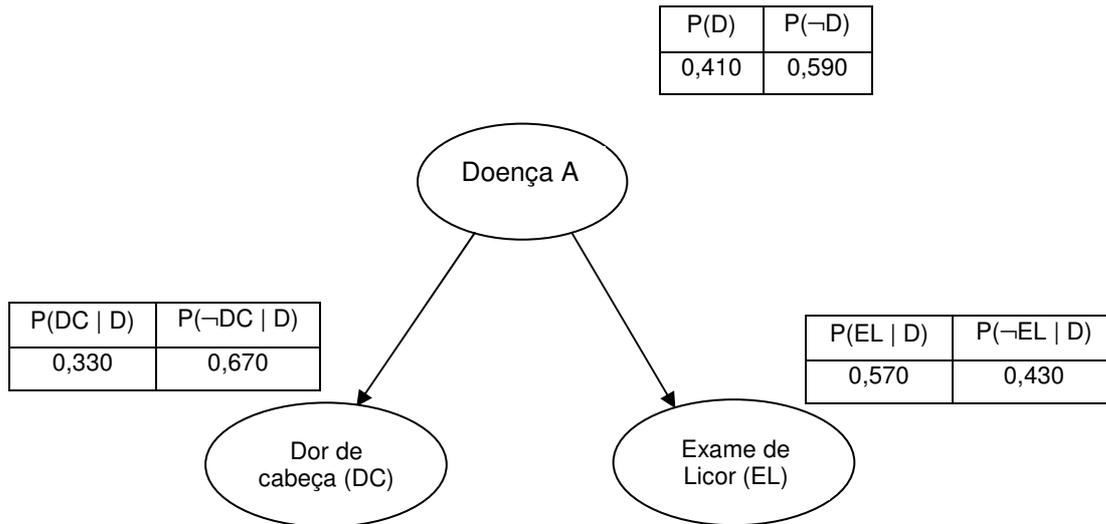


Figura 3.2: Exemplo de diagnóstico médico .

Através dos dados anteriores, construímos na tabela 3.1 a relação de probabilidade conjunta $P(A, B, C)$ para o efeito de cálculo na RB.

Tabela 3.1: Probabilidade Conjunta

P(Doença A, Dor de Cabeça, Líquor positivo)	$P(a_1, b_1, c_1)$	$P(a_1) * P(b_1 a_1) * P(c_1 a_1)$	0,077
P(Doença A, Não dor de cabeça, Líquor positivo)	$P(a_1, b_2, c_1)$	$P(a_1) * P(b_2 a_1) * P(c_1 a_1)$	0,156
P(¬Doença A, Dor de Cabeça, Líquor positivo)	$P(a_2, b_1, c_1)$	$P(a_2) * P(b_1 a_2) * P(c_1 a_2)$	0,110
P(¬Doença A, ¬Dor de cabeça, Líquor positivo)	$P(a_2, b_2, c_1)$	$P(a_2) * P(b_2 a_2) * P(c_1 a_2)$	0,225
P(Doença A, Dor de Cabeça, Líquor negativo)	$P(a_1, b_1, c_2)$	$P(a_1) * P(b_1 a_1) * P(c_2 a_1)$	0,058
P(Doença A, ¬dor de cabeça, Líquor negativo)	$P(a_1, b_2, c_2)$	$P(a_1) * P(b_2 a_1) * P(c_2 a_1)$	0,118
P(¬Doença A, Dor de Cabeça, Líquor negativo)	$P(a_2, b_1, c_2)$	$P(a_2) * P(b_1 a_2) * P(c_2 a_2)$	0,083
P(¬Doença A, ¬dor de cabeça, Líquor negativo)	$P(a_2, b_2, c_2)$	$P(a_2) * P(b_2 a_2) * P(c_2 a_2)$	0,169

Legenda: a_1 : Doença A; a_2 : ¬Doença A; b_1 : Dor de Cabeça; b_2 : ¬Dor de cabeça; c_1 : Líquor positivo; c_2 : Líquor negativo.

¬ Significa Não, ou seja, nega determinada variável.

Na tabela 3.2 apresentamos algumas perguntas feitas à RB, tais como qual é a probabilidade de ter Doença A dado que a pessoa não apresenta dor de cabeça. Pelos cálculos de probabilidade assumindo as independências, podemos observar que a resposta para esta questão é que a probabilidade de apresentar Doença A dado a evidência é de 20%, e assim são mostradas as probabilidades até o final da tabela.

Tabela 3.2: Tabela de probabilidade conjunta P(A, B, C)

P(Doença A ¬dor de cabeça)	P(a1 b2)	$\frac{P(a1,b2,c1) + P(a1,b2,c2)}{P(a1,b2,c1) + P(a1,b2,c2) + P(a2,b2,c1) + P(a2,b2,c2)}$	0,410
P(¬Doença A ¬dor de cabeça)	P(a2 b2)	$\frac{P(a2,b2,c1) + P(a2,b2,c2)}{P(a1,b2,c1) + P(a1,b2,c2) + P(a2,b2,c1) + P(a2,b2,c2)}$	0,589
P(Liquor positivo ¬dor de cabeça)	P(c1 b2)	$\frac{P(a1,b2,c1) + P(a2,b2,c1)}{P(a1,b2,c1) + P(a1,b2,c2) + P(a2,b2,c1) + P(a2,b2,c2)}$	0,570
P(Liquor negativo ¬dor de cabeça)	P(c2 b2)	$\frac{P(a1,b2,c2) + P(a2,b2,c2)}{P(a1,b2,c1) + P(a1,b2,c2) + P(a2,b2,c1) + P(a2,b2,c2)}$	0,429
P(¬Doença A dor de cabeça)	P(a2 b1)	$\frac{P(a1,b1,c1) + P(a1,b1,c2)}{P(a1,b1,c1) + P(a1,b1,c2) + P(a2,b1,c1) + P(a2,b1,c2)}$	0,202
P(Doença A dor de cabeça)	P(a1 b1)	$\frac{P(a2,b1,c1) + P(a2,b1,c2)}{P(a1,b1,c1) + P(a1,b1,c2) + P(a2,b1,c1) + P(a2,b1,c2)}$	0,289

Legenda: a1: Doença A; a2: ¬Doença A; b1: Dor de Cabeça; b2: ¬Dor de cabeça; c1: Liquor positivo; c2: Liquor negativo.

Há vários sistemas computacionais para modelar e raciocinar sobre domínios utilizando redes Bayesianas. Dentre eles podemos citar o Microsoftw MSBNx⁶, VIBES [VIBES 2008], PNL⁷, JavaBayes [COZMAN 2001], entre outros⁸.

Na próxima sessão apresentamos o sistema JavaBayes que realiza esses mesmos cálculos de forma gráfica facilitando a explicação sobre RB. Além disso, a RB para diagnóstico de DA (desenvolvida nesse trabalho) foi construída neste sistema conforme veremos no capítulo 5.

⁶. **Microsoft Belief Networks in Java**. Disponível em: <http://research.microsoft.com/adapt/MSBNx/> acessado em 19 jun. 07

⁷. Murph, Kevin. **Software Packages for Graphical Models / Bayesian Networks**. Disponível em: <http://www.cs.ubc.ca/~murphyk/Bayes/bnsoft.html>. Acessado em: 13/06/2007.

⁸. <http://www.cs.ubc.ca/~murphyk/Bayes/bnsoft.html>.

3.2 SISTEMA JAVABAYES

Escolhemos este sistema para usar em nosso trabalho pela facilidade em lidar com a interface gráfica de modo intuitivo e também por ser um sistema que funciona em diversas plataformas.

JavaBayes é um sistema de distribuição gratuita, desenvolvido por Fábio Cozman [COZMAN 2001; IDE 2001], que utiliza linguagem de programação Java permitindo sua usabilidade em diversas plataformas (Windows, Unix, entre outras). Esse sistema [COZMAN 2001] é um conjunto de ferramentas para criação, manipulação e execução de inferências em RBs. Composto por um editor gráfico que permite de forma intuitiva, construir e modificar redes Bayesianas (conforme ilustra a figura 3.3), JavaBayes possui uma máquina de inferência que é responsável pela manipulação das estruturas de dados representadas pela RB, e um conjunto de interpretadores onde podemos importar RBs em uma variedade de formatos.

Construímos a rede representada na figura 3.3 no JavaBayes para exemplificar o uso deste sistema. Foi construída a rede com os três nós como mostra a figura 3.2 da seção anterior.

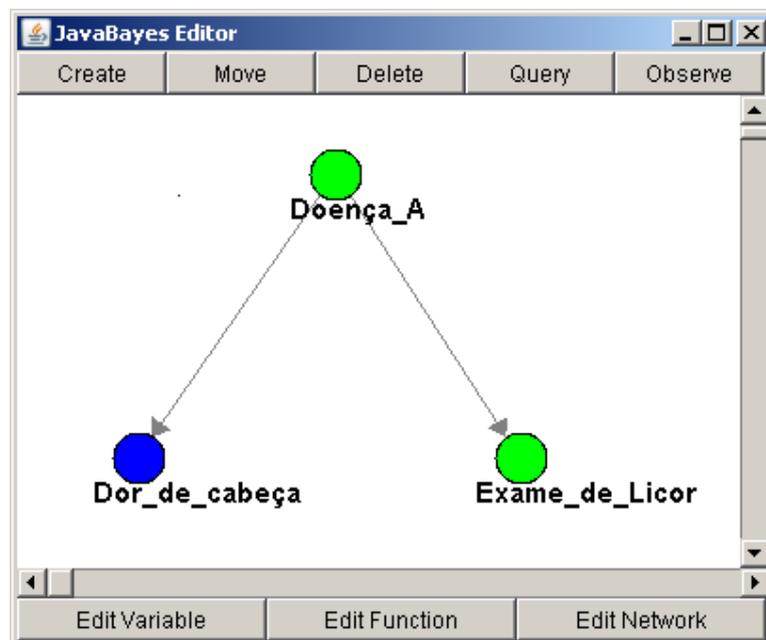
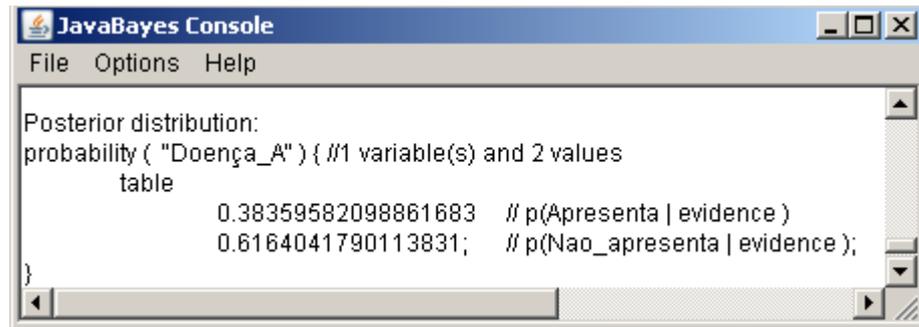


Figura 3.3: Representação gráfica de uma RB do problema Doença A.

Fizemos as mesmas perguntas ao sistema JavaBayes e demonstramos na figura 3.4 os resultados para a probabilidade de apresentar ou não a Doença A, dado a pessoa apresentar dor de cabeça $P(\text{Doença A} \mid \text{dor de cabeça})^9$. Podemos verificar que o resultado é de modo aproximado igual ao da tabela 3.3 que foi calculada manualmente.



```

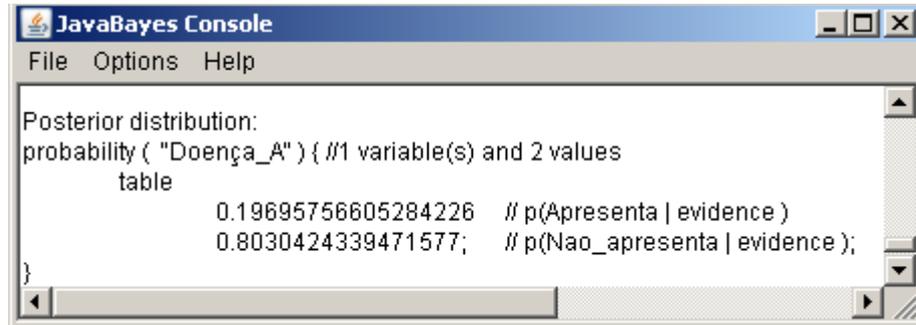
JavaBayes Console
File Options Help

Posterior distribution:
probability ( "Doença_A" ) { //1 variable(s) and 2 values
  table
    0.38359582098861683 // p(Apresenta | evidence )
    0.6164041790113831; // p(Nao_apresenta | evidence );
}
  
```

Figura 3.4: Resultado da inferência realizada representada pela figura 3.3.

⁹. Nas figuras que apresentam os resultados da rede mostra $P(\text{Apresenta} \mid \text{evidence})$ e $P(\text{não_Apresenta} \mid \text{evidence})$, como esse resultado é do próprio software JavaBayes não conseguimos alterar essa informação para português (então, entende-se $P(\text{Apresenta} \mid \text{evidências})$ e $P(\text{Não_Apresenta} \mid \text{evidências}))$).

No resultado representado abaixo (figura 3.5), a pergunta foi para saber qual a probabilidade de apresentar ou não Doença A, dado a pessoa não apresentar dor de cabeça. Nota-se que o resultado mostrou-se, como no exemplo acima, aproximadamente igual ao da tabela calculada manualmente.



```

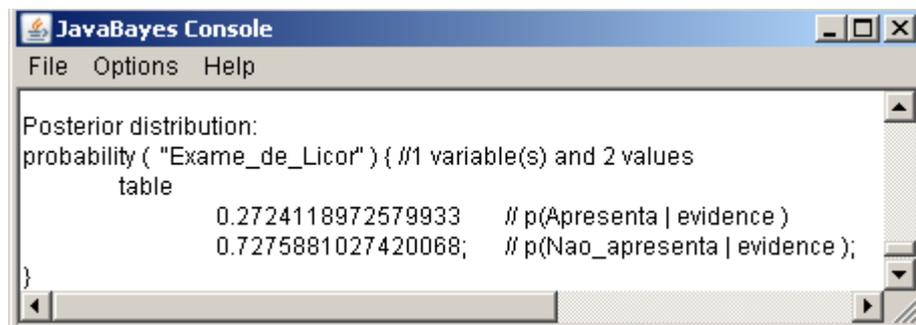
JavaBayes Console
File Options Help

Posterior distribution:
probability ( "Doença_A" ) { //1 variable(s) and 2 values
  table
    0.19695756605284226 // p(Apresenta | evidence )
    0.8030424339471577; // p(Nao_apresenta | evidence );
}

```

Figura 3.5: Resultado da inferência realizada representada pela figura 3.3.

Abaixo (figura 3.6) ilustramos ainda outra pergunta feita a rede para sabermos qual a probabilidade do resultado do exame de Liquor dar positivo ou negativo dado a pessoa não apresentar dor de cabeça $P(\text{Doença A} \mid \neg \text{dor de cabeça})$. Podemos notar que também o resultado se mostrou condizente ao cálculo efetuado manualmente.



```

JavaBayes Console
File Options Help

Posterior distribution:
probability ( "Exame_de_Liquor" ) { //1 variable(s) and 2 values
  table
    0.2724118972579933 // p(Apresenta | evidence )
    0.7275881027420068; // p(Nao_apresenta | evidence );
}

```

Figura 3.6: Representação gráfica de uma RB do problema Doença A.

Na próxima seção apresentamos o método para reduzir a complexidade na definição de tabelas de probabilidades condicionais quando muitas variáveis compõem uma RB.

3.3 MODELO NOISY-OR

Tomamos como exemplo o trabalho desenvolvido por André Hideaki Saheki et. al. [SAHEKI 2003] que utiliza a técnica de aproximação noisy-or para lidar com números elevados de variáveis do domínio cardiológico. O sistema baseado em RB para o domínio de diagnóstico de Alzheimer (objeto deste trabalho) pode levar também a um grande número de variáveis. Este método foi utilizado no presente trabalho.

Noisy-or é uma técnica de aproximação utilizada para obter probabilidades condicionais sem precisar de todas as informações. Este modelo é utilizado geralmente quando se tem um número excessivo de variáveis, necessitando de muitos valores de probabilidade para a rede. Esta é uma das maiores dificuldades na aplicação de RBs [PEARL 1988 e SAMPAIO 2007]. Permite o cálculo das probabilidades de ligação que considera que D (doença) pode ser causada pelo fator R (risco) ou ainda por todos os R^s (riscos) que possam levar a D (doença) [SAHEKI 2003].

Para ilustrarmos o funcionamento do método Noisy-or utilizamos um exemplo de probabilidade de apresentar febre ou não apresentar, dadas algumas doenças, como: gripe, resfriado e malária. Abaixo distribuimos os valores para as probabilidades de não apresentar febre dado as evidências de apresentar ou não determinadas características (onde encontramos \neg lê-se “não apresenta”) este exemplo está descrito em [RUSSELL E NORVIG 1995].

- $P(\neg\text{febre} \mid \text{resfriado}, \neg\text{gripe}, \neg\text{malária}) = 0,6$
- $P(\neg\text{febre} \mid \neg\text{resfriado}, \text{gripe}, \neg\text{malária}) = 0,2$
- $P(\neg\text{febre} \mid \neg\text{resfriado}, \neg\text{gripe}, \text{malária}) = 0,1$

A tabela 3.3 apresenta a distribuição de probabilidades condicionais utilizando o método Noisy-or. A letra V (verdadeiro) indica quando apresenta determinada característica e F (falso) quando não apresenta. Na coluna da probabilidade de não apresentar febre ($P(\neg\text{Febre})$) são lançadas as probabilidades a priori citadas acima (destacadas em amarelo para melhor visualização), nas demais colunas onde encontramos mais letras V multiplicamos os valores correspondentes a cada situação

(exemplo linha 8 (destacado em verde): multiplica-se a $P(\neg\text{febre}|\neg\text{resfriado}, \neg\text{gripe}, \text{malária}) = 0,1$, $P(\neg\text{febre}|\neg\text{resfriado}, \text{gripe}, \neg\text{malária}) = 0,2$ e $P(\neg\text{febre} | \text{resfriado}, \neg\text{gripe}, \neg\text{malária}) = 0,6$, logo chega-se no resultado que a $P(\neg\text{febre} | \text{resfriado}, \text{gripe}, \text{malária})$ é de 0,012), portanto a de apresentar febre ($P(\text{febre} | \text{resfriado}, \text{gripe}, \text{malária})$) é de 0,988 e assim é feito sucessivamente.

Tabela 3.3: Tabela de distribuição de probabilidades

Resfriado	Gripe	Malária	P(Febre resfriado, gripe, malária)	P(\neg Febre)
F	F	F	0,0	1,0
F	F	V	0,9	0,1
V	F	F	0,8	0,2
V	F	V	0,98	0,02 = 0,2 x 0,1
F	V	F	0,4	0,6
F	V	V	0,94	0,06 = 0,6 x 0,1
V	V	F	0,88	0,12 = 0,6 x 0,2
V	V	V	0,988	0,012=0,6 x 0,2 x 0,1

Geralmente esses tipos de distribuições são usadas em medicina para modelar relacionamentos entre doenças e sintomas, pois por exemplo: com 448 nós e 906 vínculos, esse modelo exige 8.254 valores, em vez de 133.931.430 para uma rede com tabelas de probabilidade condicionais (TPCs) completas [RUSSELL E NORVIG 1995].

Em qualquer representação do mundo real, existirão alguns fatores que não são modelados. A falta desses fatores pode ser intencional (para evitar um modelo demasiadamente complexo) ou acidental (por mero desconhecimento de quais seriam estes fatores).

A seguir, apresentamos alguns sistemas baseados em RB que utilizam os conhecimentos e metodologias elencados na seção 3.1, entre outros.

3.4 SISTEMAS BASEADOS EM REDES BAYESIANAS

Entre os sistemas que trabalham com inferência em RB, e que são utilizados em diversos domínios, citamos: Hugin Expert¹⁰, MSBNx - Bayesian Network Editor and Toolkits [MSBN 2007; CARVALHO 2004] e JavaBayes [COZMAN 2001; IDE 2001], conforme descrito abaixo.

Em¹⁰ tem-se uma descrição do sistema Hugin Expert desenvolvido pela Hugin A/S, que desde 1989 comercializa produtos com soluções inteligentes para diversas áreas, tais como: financeira, alimentícia, industrial, médica, acadêmica, etc. Ainda em¹⁰ encontramos uma lista desses produtos, sendo todos eles baseados em RBs. Em particular há dois deles que envolvem a área médica: um de uso para orientação do tratamento empírico de infecções bacterianas severas e outro que permite uma solução de suporte à decisão diagnóstica e terapêutica apropriadas para cada caso investigado. Este software possui licença de uso, sendo acessível apenas mediante pagamento.

MSBNx - Bayesian Network Editor and Toolkits [MSBN 2007; CARVALHO 2004] é uma aplicação criada pela Microsoft distribuída gratuitamente pela internet. MSBNx dá suporte à criação, manipulação e avaliação dos modelos de probabilidade bayesiano e possui uma interface visual amigável, conforme mostra a figura 3.7.

¹⁰. Hugin Expert A/S. Gasværksvej 5 . DK-9000 Aalborg . Denmark. Disponível em: <http://www.hugin.com/>. Acessado em 19/06/2007.

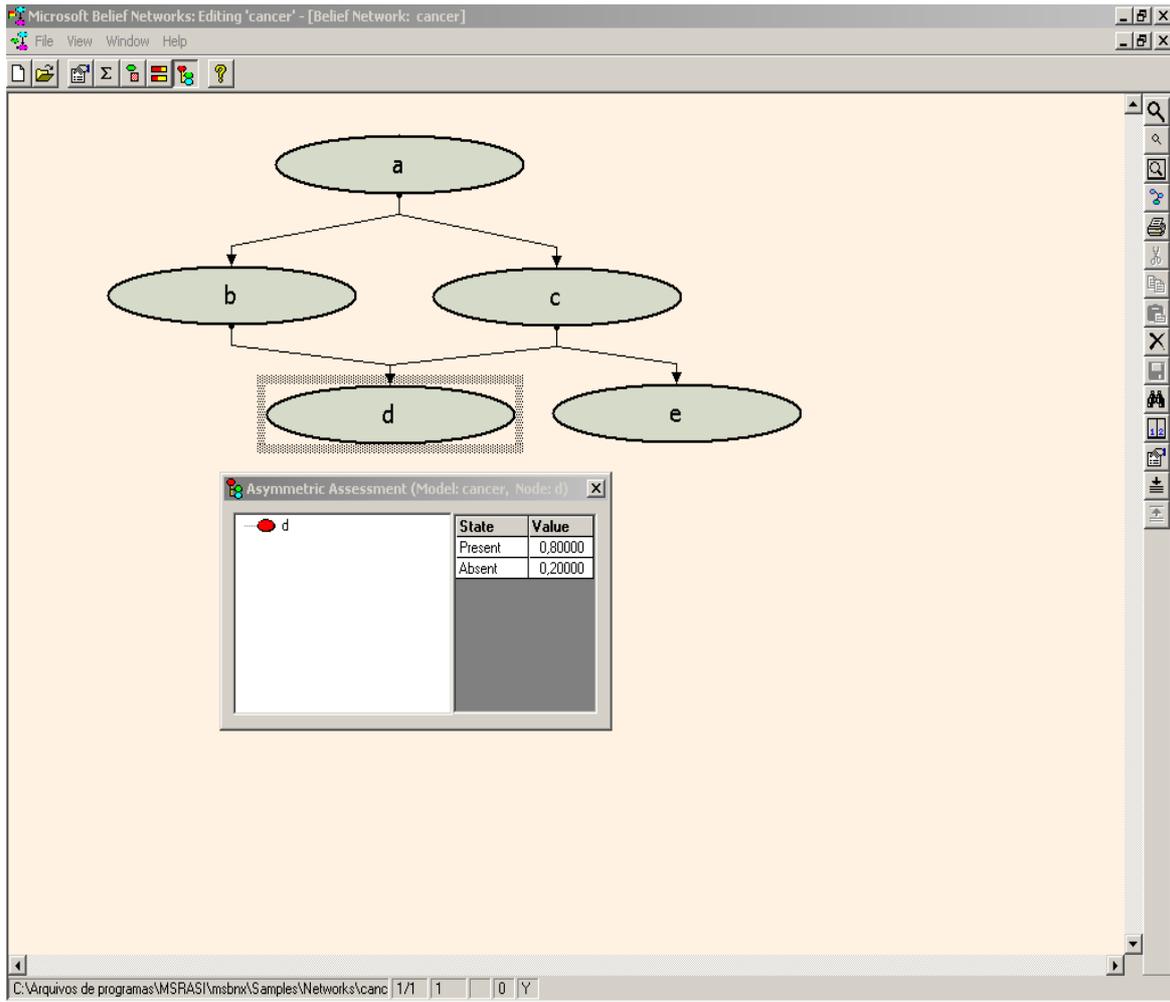


Figura 3.7: MSBNx - Bayesian Network Editor and Toolkits, adaptado de [MSBNx 2008].

Estes sistemas são de aplicação geral, podendo ser utilizados em diversos domínios para diversos fins. A seguir iremos demonstrar alguns sistemas mais específicos para diagnósticos na área médica.

3.4.1 SISTEMAS ESPECIALISTAS BASEADOS EM REDES BAYESIANAS

Nesta sessão iremos descrever alguns sistemas especialistas baseados em RB já existentes [HAIPENG 2002], tais como: DiagSD (*Diagnosing Sleep-Disorders*) [MILHO 2000], SEAMED¹¹, Sistema de Avaliação Nutricional [TIBIRIÇÁ 2006] e iBNetz [SAHEKI 2003].

Isabel Milho e Ana Fred [MILHO 2000] descrevem um sistema especialista (DiagSD (*Diagnosing Sleep-Disorders*)) com a utilização de RB que favorece o diagnóstico médico na doença de desordem do sono (figura 3.8). O sistema DiagSD apresenta uma interface adequada às aplicações dos especialistas do domínio clínico, sem que estes precisem ter conhecimento profundo em RB.

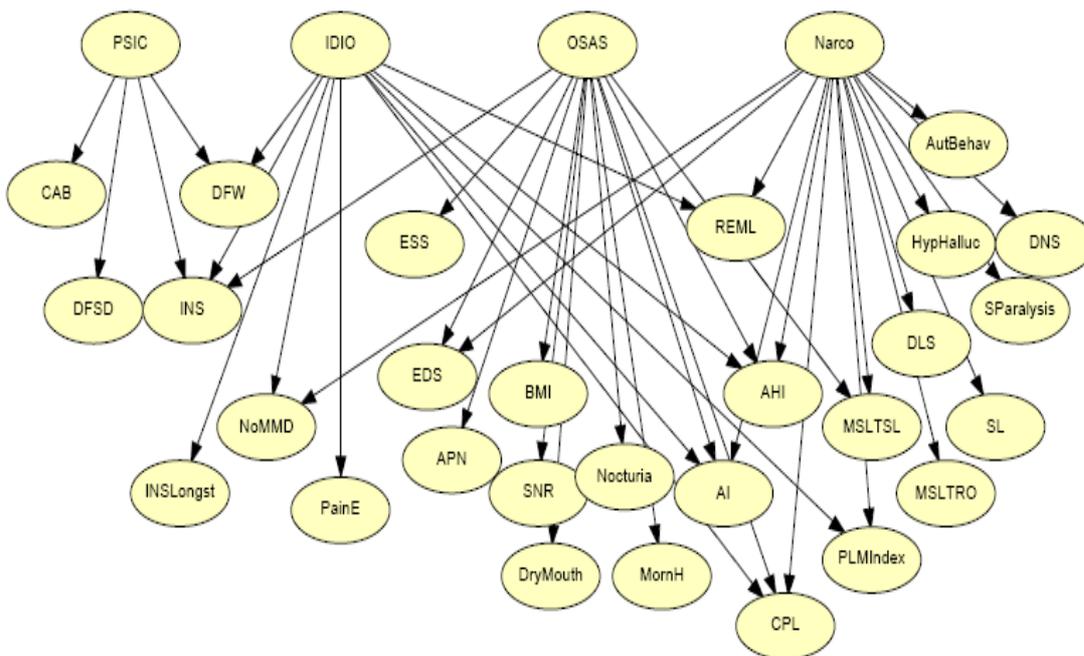


Figura 3.8: Topologia da RB como modelo do problema das doenças de sono. Adaptado de [MILHO 2000]

¹¹. SEAMED – Sistema especialista médico baseado em regras. Disponível em: <http://www.inf.ufrgs.br/~dflores/seamed/default.htm>. Acessado em: 07/07/2007.

Os nós PSIC, IDIO, OSAS e Narco (figura 3.8) representam as doenças psicológicas, insônia, insônia idiopática, síndrome da apnéia do sono obstrutiva e epilepsia por narcóticos relacionados ao sono, sendo a definição da topologia da RB e a das tabelas de probabilidade baseadas na literatura médica e em consultas à especialistas da área [MILHO 2000].

A funcionalidade da RB é representada por um sistema de interface gráfica onde o especialista trabalha via Internet, através das ações de diagnóstico e de relatório conforme figura 3.9. O sistema é composto por três blocos fundamentais: o módulo principal desenvolvido em linguagem Java e arquivos HTML que realiza a troca de informações com o utilizador, e o algoritmo de inferência construído em linguagem C/ C++ que realiza os cálculos probabilísticos. Deste modo, através de uma página na internet, o usuário alimenta os formulários com as evidências sendo estas computadas pelo módulo principal que por sua vez apresenta a resposta via Web para o usuário.

The image shows two overlapping browser windows. The left window displays the main interface of the DiagSD system, featuring the title 'DiagSD Diagnosing Sleep Disorders' and a logo of a person holding a globe. The right window displays the 'DiagSD Diagnostic results' page for patient code 00000. It contains two tables: 'Pathologies' and 'Clinical and Polysomnographic data'.

Pathology	Percentage	Status
PSIC	57.0%	absent
PSIC	63.0%	present
IDIO	93.0%	absent
IDIO	7.0%	present
OSAS	87.0%	absent
OSAS	13.0%	present
Narco	100.0%	absent
Narco	0.0%	present

Condition	Percentage	Status
DFSD	40.0%	absent
DFSD	60.0%	present
INSLongst	89.0%	absent
INSLongst	11.0%	present
NoMMD	89.0%	absent
NoMMD	11.0%	present
APN	81.0%	absent
APN	19.0%	present
EIS	84.0%	absent
EIS	16.0%	present
DryMouth	80.0%	absent
DryMouth	20.0%	present
Mouth	80.0%	absent
Mouth	20.0%	present
SNR	80.0%	absent
SNR	20.0%	present
Nocturia	80.0%	absent
Nocturia	20.0%	present
CEL	90.0%	absent
CEL	1.0%	variant
CEL	9.0%	possible
SPParalysis	90.0%	absent
SPParalysis	10.0%	present
HypHaltuc	80.0%	absent
HypHaltuc	10.0%	present
DHS	80.0%	absent
DHS	10.0%	present

Figura 3.9: Páginas da internet do sistema DiagSD, arquivo para ação do diagnóstico e página produzida com os resultados de diagnóstico. Adaptado de [MILHO 2000].

Pessete [PESSETE 2002] e Perotto [PEROTTO 2001] descrevem o sistema SEAMED, desenvolvido na universidade Federal do Rio Grande do Sul, baseado na aplicação de RBs para o auxílio do diagnóstico médico na área de cardiopatias congênitas.

O sistema SEAMED é formado por um módulo de construção e um módulo de consulta [PESSETE 2002]. O módulo de construção é onde o usuário (especialista) insere os nós que representam as variáveis que ele julga relevantes ao domínio, estes nós são ligados através de arcos (arestas) que representam as influências entre si, depois o especialista informa qual a influência dos nós-pai e os nós-filho e como a

variável em questão se comporta, considerando todas as combinações (hipóteses): $P(\text{sim} \mid \text{sim}, \text{sim})$, $P(\text{sim} \mid \text{sim}, \text{não})$, $P(\text{sim} \mid \text{não}, \text{sim})$ e $P(\text{sim} \mid \text{não}, \text{não})$, nota-se que a somas dos valores em cada coluna é 1, conforme ilustrado na figura 3.10.

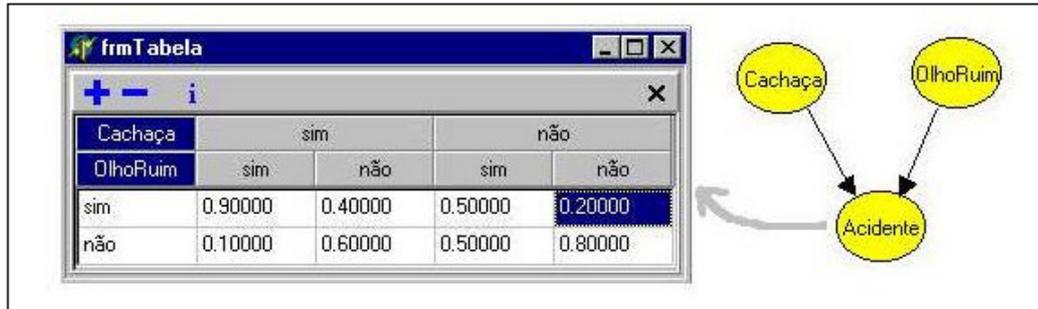


Figura 3.10: Tabela de ilustração do sistema SEAMED. Adaptado de [PEROTTO 2001].

No módulo de consulta, o especialista pode indicar a existência da evidência atribuindo valores de probabilidades. Assim o sistema SEAMED responde quais são os possíveis diagnósticos (considerando os diagnósticos que obtiveram 50% ou mais de ocorrência), mostrando qual o valor de certeza atingida para cada um deles, como mostra a figura 3.11 [PEROTTO 2001; SEAMED 2007; FLORES 2002].

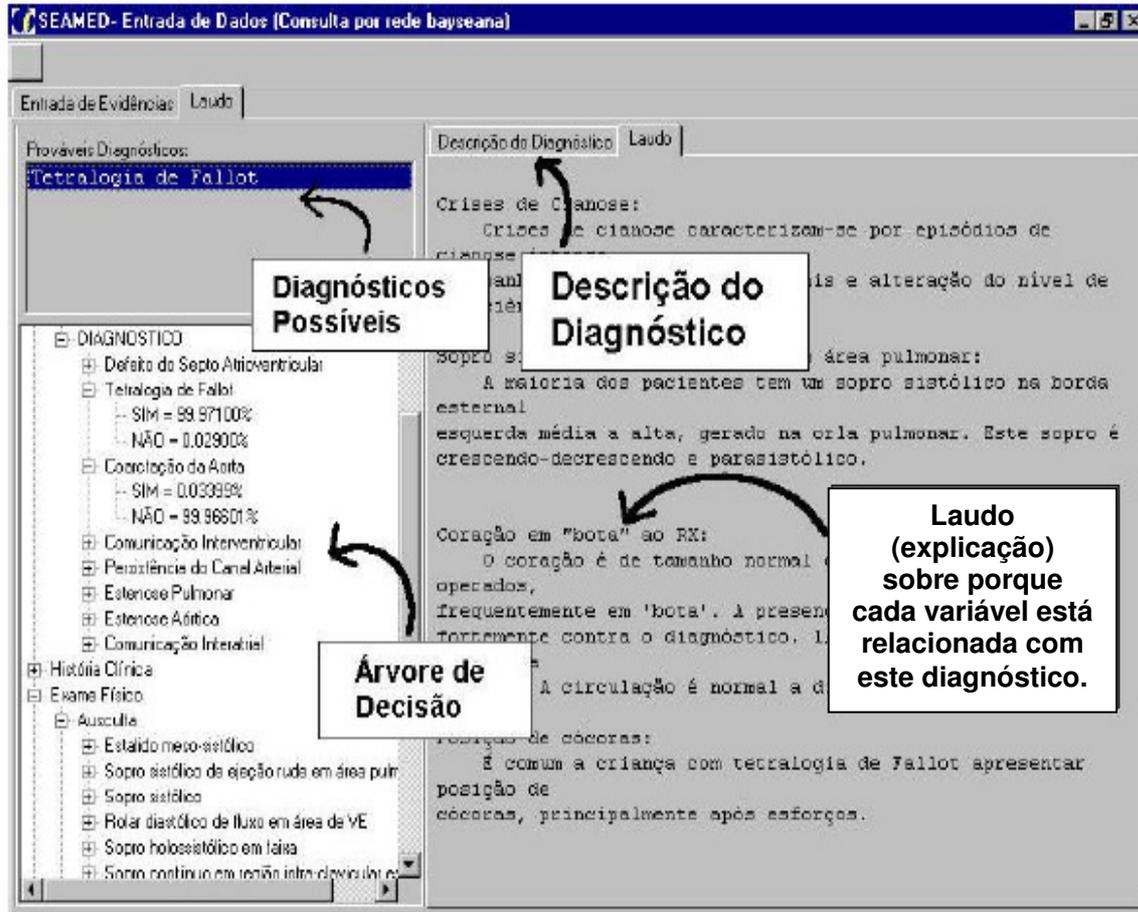


Figura 3.11: Exemplo de entrada de dados no programa SEAMED. Adaptada de [PEROTTO 2001].

É descrito em [SAHEKI 2003] um trabalho desenvolvido entre a Escola Politécnica – USP e o Hospital Universitário – USP cujo objetivo foi desenvolver um sistema utilizando RB e o modelo noisy-or, chamado iBNetz, para o auxílio no diagnóstico de doenças cardíacas, especificamente voltado para pessoas com falta de ar (dispnéia) que dão entrada em postos de saúde.

A rede representada na figura 3.12 realiza uma seleção entre os pacientes que precisam de um encaminhamento ao hospital para o devido tratamento cardíaco, ou se podem permanecer nos postos se o problema apresentado for de menor gravidade. A intenção é diminuir gastos desnecessários com atendimentos hospitalares e melhorar a eficiência do sistema público de saúde [SAHEKI 2003].

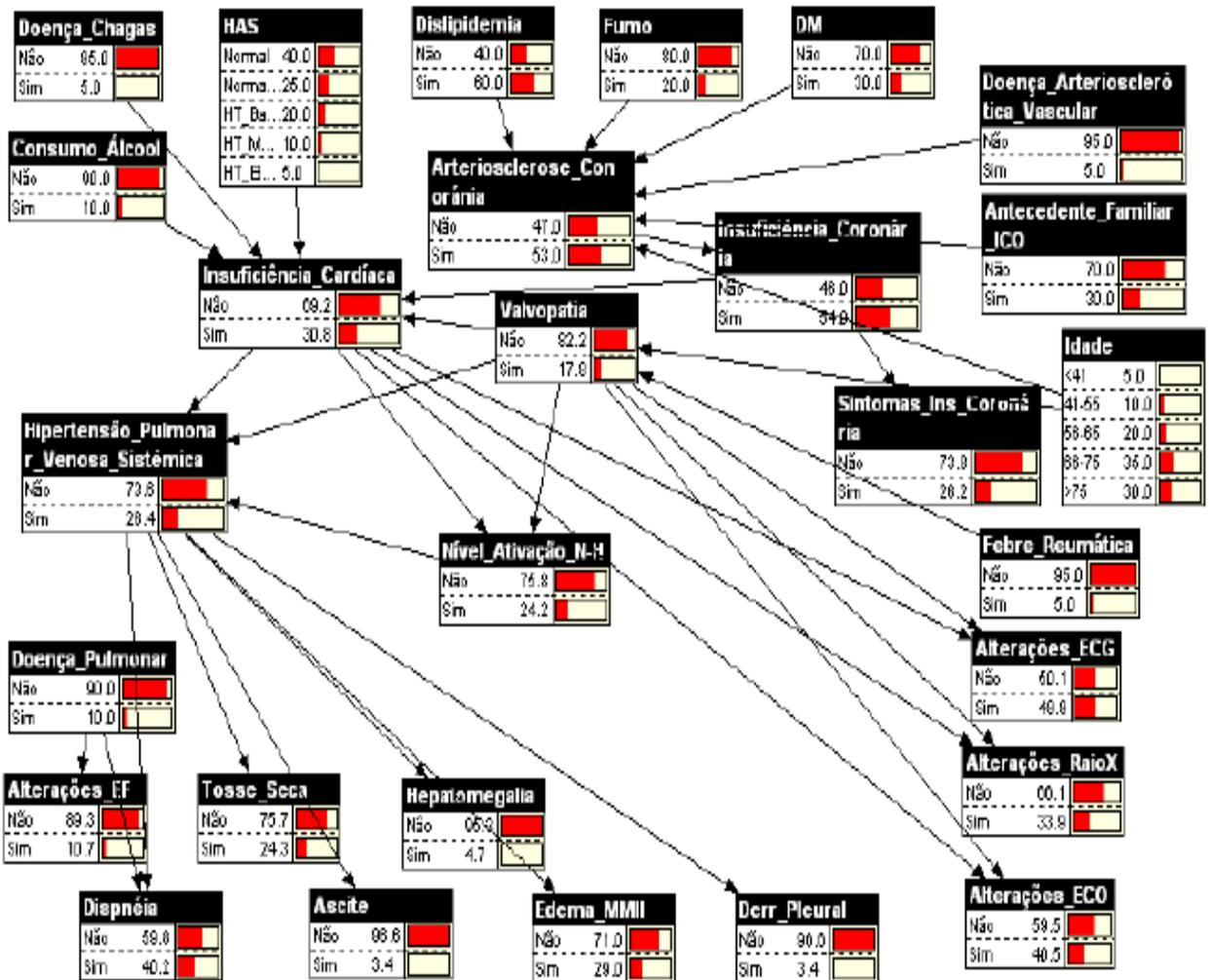


Figura 3.12: Topologia da RB de insuficiência cardíaca. Adaptado de [SAHEKI 2003]

Por ser uma RB com muitos nós, sua utilização depende das observações feitas pelo médico nas variáveis do problema, sendo a variável “insuficiência cardíaca” a de maior impacto na sua decisão [SAHEKI 2003].

Para esta estrutura de RB foi construído o software iBNetz [SAHEKI 2003] conforme figura 3.13 no intuito de executar inferências na rede. O software utiliza a linguagem Java e roda em ambiente Windows, Unix, Linux e similares. Embora o software não permita a criação de uma rede totalmente diferente da rede para diagnóstico de doenças cardíacas, ele permite alterar nós, arcos e probabilidades e sua distribuição é gratuita.

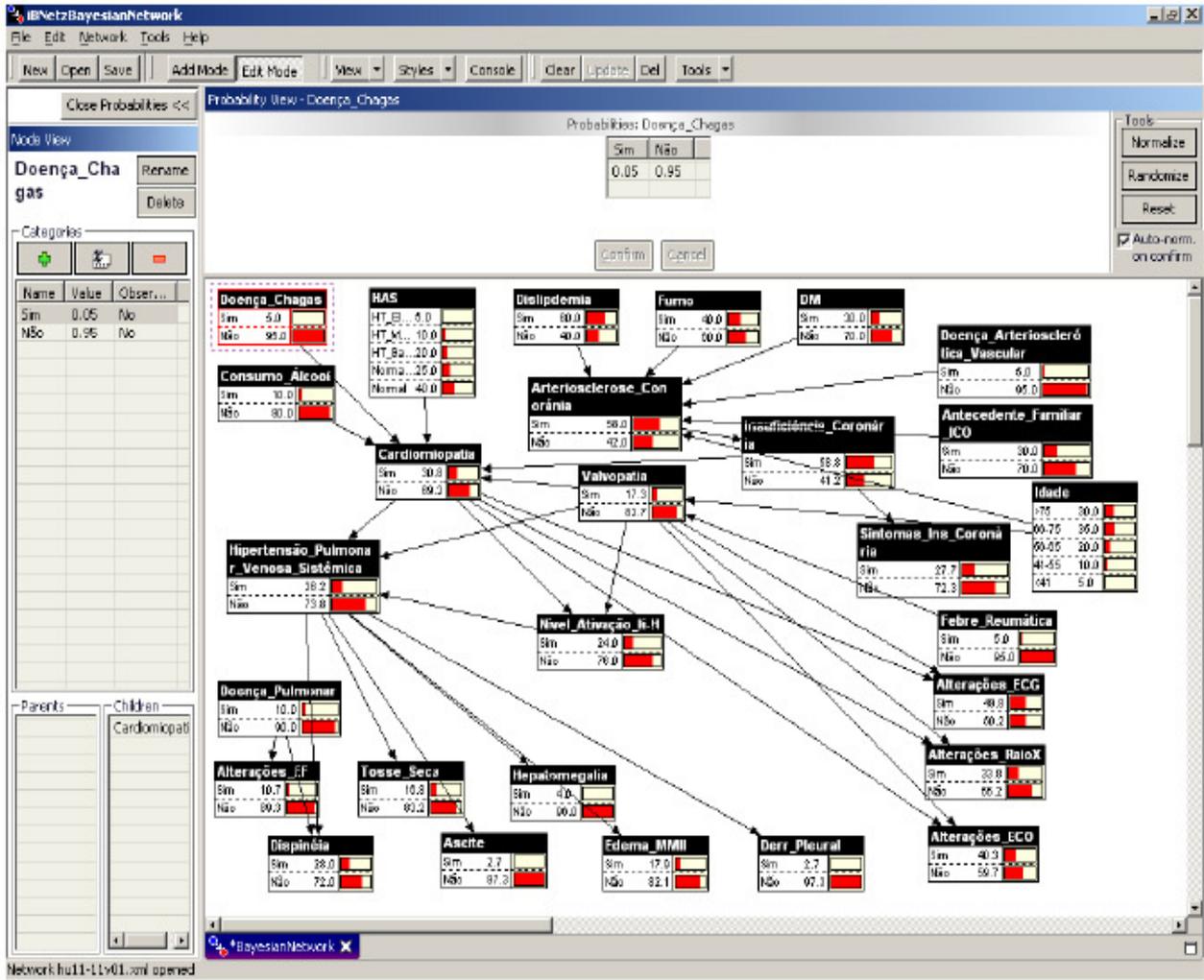


Figura 3.13: Software iBNetz. Adaptado de [SAHEKI 2003].

Um sistema baseado em RB nomeado de Sistema de Avaliação Nutricional é apresentado em [TIBIRIÇÁ 2006] (figura 3.14) e serve de apoio ao diagnóstico de avaliação nutricional. Esse sistema é dividido em três módulos (conhecimento, simulação e especialista) permitindo a interação de leigos e especialistas, com ressalva ao módulo especialista que é liberado apenas para o próprio especialista.

O módulo conhecimento tem por objetivo disponibilizar informações relativas a diversos temas relacionados à desnutrição e à avaliação nutricional; o módulo simulação tem por objetivo permitir pessoas leigas fazerem um pré-diagnóstico de avaliação nutricional; e o módulo especialista é restrito a pessoas da área médica.

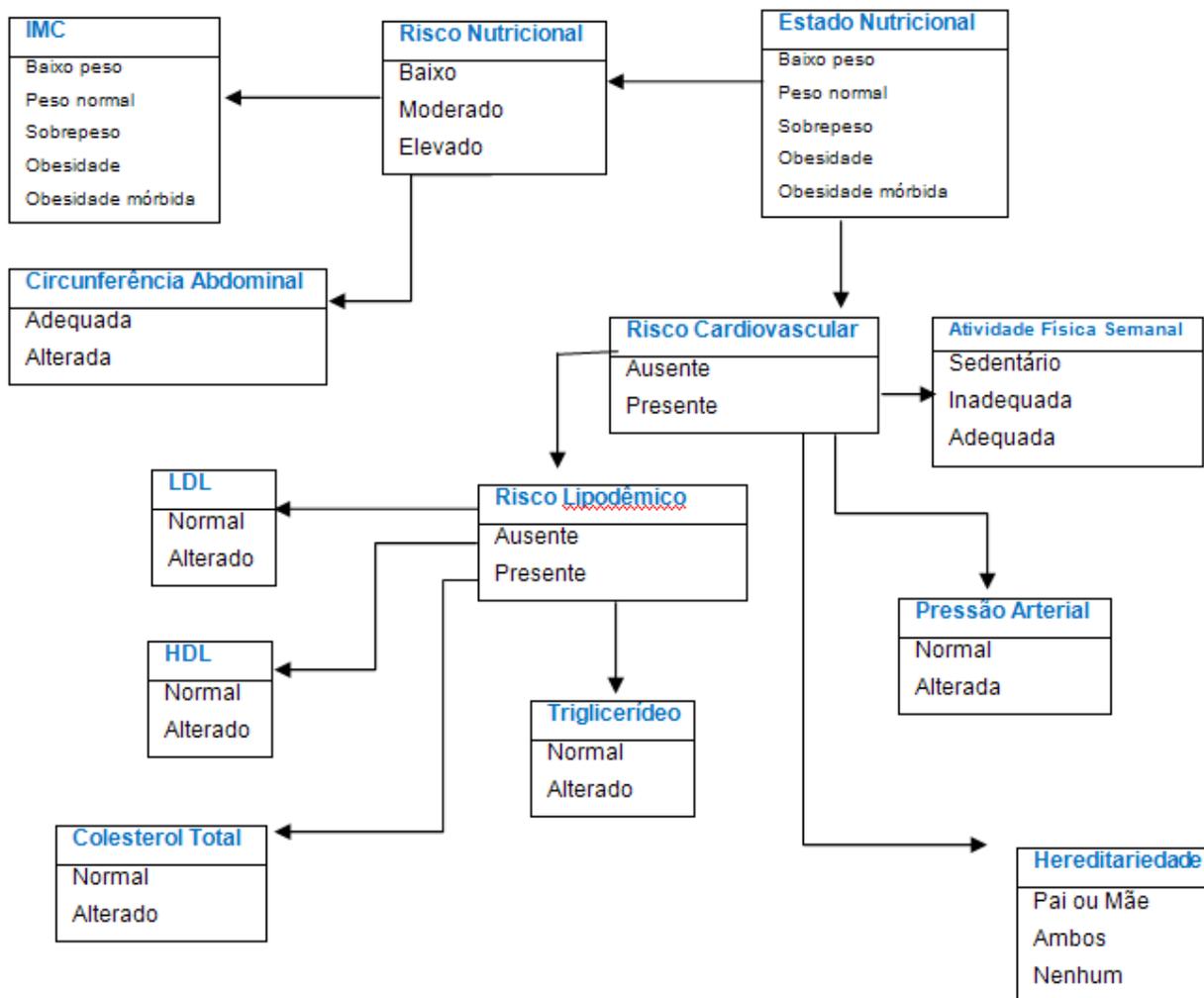


Figura 3.14: Topologia da RB do Sistema de Avaliação Nutricional. Adaptado de [TIBIRIÇÁ 2006].

A RB deste sistema é representada por um sistema que trabalha via Web por meio de formulários captando as informações (variáveis) e apresenta o resultado na forma de gráficos de barras, conforme os testes abaixo (figuras 3.15, 3.16, 3.17 e 3.18).

Sexo do Paciente:

Risco Cardiovascular

Atividade Física Semanal:

Hereditariedade:

Pressão Arterial

Sistólica: mmHg

Diastólica: mmHg

Risco Nutricional

Índice de Massa Corporal

Peso: Kg

Altura: m

Circunferência Abdominal

Medida do Abdômen: cm

Figura 3.15: Variáveis de entrada utilizadas no módulo de simulação. Adaptado de [TIBIRIÇÁ 2006].

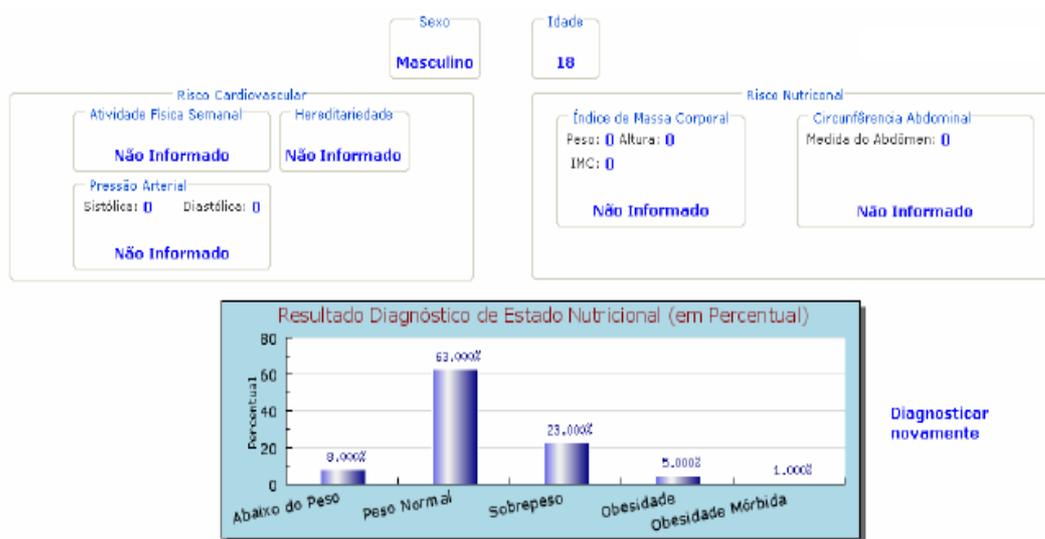


Figura 3.16: Resultado do módulo de simulação. Adaptado de [TIBIRIÇÁ 2006].

<p>• Risco Cardiovascular</p> <p>Atividade Física Semanal: <input type="text" value="Sedentário"/></p> <p>Hereditariedade: <input type="text" value="Ambos"/></p> <p>Pressão Arterial Sistólica: <input type="text" value="9"/> mmHg</p> <p>Pressão Arterial Diastólica: <input type="text" value="6"/> mmHg</p>	<p>• Risco Nutricional</p> <p>Peso: <input type="text" value="75"/> Kg</p> <p>Altura: <input type="text" value="1.5"/> m</p> <p>Medida do Abdômen: <input type="text" value="88"/> cm</p>
---	--

<p>• Risco Lipidêmico</p> <p>Triglicérido: <input type="text" value="90"/> (mg/dl)</p> <p>Colesterol Total: <input type="text" value="280"/> (mg)</p> <p>LDL: <input type="text" value="150"/> (mg)</p> <p>HDL: <input type="text" value="40"/> (mg)</p>	<p>Realizar Diagnóstico >></p>
---	--------------------------------------

Figura 3.17: Valores de entrada de uma avaliação completa. Adaptado de [TIBIRIÇÁ 2006].

<p>• Risco Cardiovascular</p> <p>Atividade Física Semanal: Sedentário</p> <p>Hereditariedade: Ambos</p> <p>Pressão Arterial: Normal</p> <ul style="list-style-type: none"> - Sistólica: 9 mmHg - Diastólica: 6 mmHg 	<p>• Risco Nutricional</p> <p>Índice de Massa Corporal: Obesidade</p> <ul style="list-style-type: none"> - Peso: 75 Kg - Altura: 1.5 m - IMC: 33.33 <p>Circunferência Abdominal: Alterada</p> <ul style="list-style-type: none"> - Medida do Abdômen: 88 cm
---	--

<p>• Risco Lipidêmico</p> <p>Triglicérido: Normal</p> <ul style="list-style-type: none"> - Valor: 90 (mg/dl) <p>HDL: Normal</p> <ul style="list-style-type: none"> - Valor: 40 (mg) 	<p>LDL: Alterado</p> <ul style="list-style-type: none"> - Valor: 150 (mg) <p>Colesterol Total: Alterado</p> <ul style="list-style-type: none"> - Valor: 280 (mg)
--	--

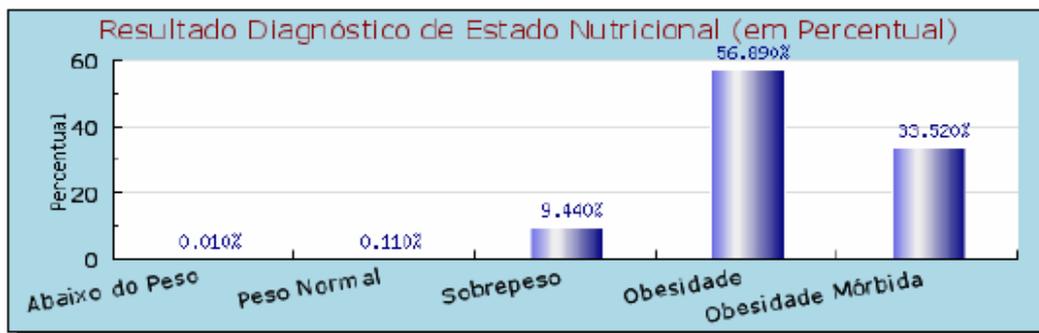


Figura 3.18: Resultado de uma avaliação completa. Adaptado de [TIBIRIÇÁ 2006].

Baseando-se nos sistemas e suas aplicabilidades descritos neste capítulo, desenvolvemos um protótipo de sistema especialista para o diagnóstico da DA utilizando o sistema JavaBayes, tal qual apresentado no capítulo 5.

4 DIAGNÓSTICO POR NEUROIMAGEM

Em função das dificuldades no diagnóstico da DA, dentre o conjunto das demências, um estudo realizado em 2001 [CHARCHAT 2001] investigou os marcadores clínicos dos estágios da DA, onde pode-se observar que as pessoas com essa doença tiveram alterações neurológicas sugestivas de disfunção em estruturas do lobo temporal e frontal, o que difere dos pacientes com amnésia e casos que apresentam exclusivamente déficit de memória episódica.

Charchat e Goldman [CHARCHAT 2001; GOLDMAN 1992] a partir de métodos funcionais de neuroimagens, descreveram um envolvimento do córtex pré-frontal em tarefas usadas como modelo de memória operacional. Maures [MAURES 1997] reforçou esta idéia afirmando que pacientes com lesões no lobo frontal apresentam comprometimento de memória operacional associado a um déficit de funções executivas. Neste artigo [CHARCHAT 2001] podemos encontrar descritas as principais características neuropsicológicas dos estágios iniciais de DA, como por exemplo: lentidão cognitiva em função do aumento do tempo de reação; comprometimento da memória episódica (acometimento bilateral dos lobos temporais, com predomínio do hemisfério dominante esquerdo); comprometimento da memória de curto-prazo (correlacionada com a disfunção no lobo frontal).

É observado em [DAVATZIKOS 2006] que os estágios iniciais da DA podem ser tão sutis que se confundem com o envelhecimento natural, dificultando a diferenciação do diagnóstico. Por este motivo, análises estatísticas em imagens são usadas para capturar as características morfológicas que diferenciam o envelhecimento normal e transtorno cognitivo leve (TCL) da DA.

Muitas análises de imagens de RM de crânio têm evidenciado volumes diferentes em regiões específicas do cérebro, especialmente as regiões do hipocampo e córtex entorhinal, onde há mudanças histopatológicas (estudo dos tecidos doentes – histologia patológica). No estágio inicial de DA, os estudos neuroanatômicos computacionais usados em comparações de indivíduos em processo de envelhecimento natural e DA vem confirmando as atrofia de estruturas lobo temporal médio [DAVATZIKOS 2006].

Podemos utilizar essas informações para construir nossa rede inserindo-as como variáveis do nó neuroimagem apresentado com maior detalhe no capítulo 5.

As figuras a seguir evidenciam algumas diferenças. A figura 4.1 (A) mostra uma comparação do tamanho de um cérebro saudável com um cérebro de uma pessoa com DA. Observou-se que o cérebro do paciente com DA teve seu tamanho reduzido de forma bastante notável. A figura 4.1 (B) mostra também a diferença de um cérebro saudável com um cérebro de uma pessoa com DA, porém num corte coronal para melhor se observar o ventrículo e os sulcos dos cérebros, percebendo claramente o alargamento dessas regiões afetadas pela doença [AGUIAR 2006].

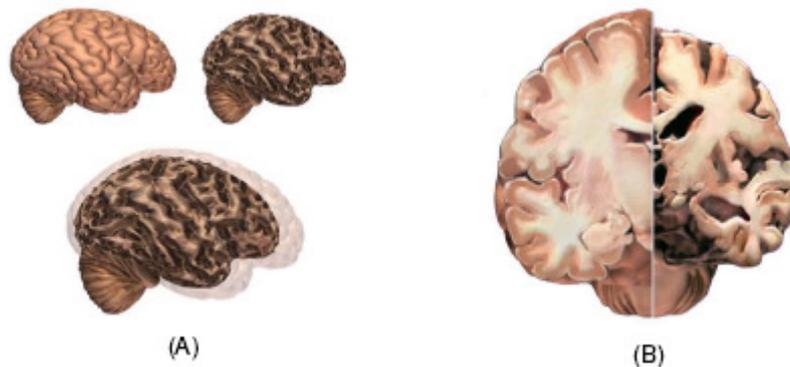


Figura 4.1: Efeito relativo ao cérebro dos pacientes que tem a DA. (Adaptado de [AGUIAR 2006]).

Para se chegar a um consenso do diagnóstico, são realizados exames laboratoriais e também de imagens como tomografia computadorizada (TC) ou, preferencialmente, a ressonância magnética (RM) a fim de afastar outras possíveis doenças. Foi observado que o exame de imagem (TC ou RM) cerebral pode apresentar atrofia da região do córtex entorhinal em possíveis pacientes com DA, porém em sua fase inicial pode apresentar mudanças tão sutis que é necessário uma RM de alta resolução para uma melhor observação, como a espectroscopia por RM e a técnica de subtração de imagens [NITRINI 2005].

Estudos metabólicos do exame de espectroscopia por RM vêm mostrando uma diminuição da concentração de N-acetilaspártato (NAA) e aumento de mio-inositol (MI) na formação do hipocampo. A relação NAA / MI é robusta na diferenciação entre os

casos com possível DA e controles (pessoas que não apresentam características da DA) pareados por idade. Lembrando que outros pacientes com outras demências apresentam níveis significativamente reduzidos de NAA, porém com níveis de MI normais [NITRINI 2005; AZEVEDO 2005].

Em exames como PET (Positron Emission Tomography) e SPECT (Single – Photon Emission Computed Tomography) pode-se observar evidências que pacientes com DA apresentam hipoperfusão amigdaló-hipocampal, uma redução bilateral e frequentemente assimétrica do fluxo sanguíneo e do metabolismo em regiões temporais ou temporo parietais [NITRINI 2005].

Segundo Ricardo Nitrini [NITRINI 2005] “o diagnóstico de DA provável, baseado em critérios clínicos (NINCDS - ADRDA), confere uma probabilidade de 84% para o diagnóstico patológico, podendo subir para 92% se o spect cerebral for positivo (ou cair para 70% se der negativo), e o EEG (Eletroencefalograma) e EEGQ (Eletroencefalograma quantitativo) podem ser úteis para auxiliar na decisão diagnóstica enquanto a mesma estiver em aberto”.

[PASCALE 2002; RAMOS 2001] confirmam que o déficit cognitivo de DA é baseado na deficiência do neurotransmissor acetilcolina, encontraram também o atrofiamento do córtex e o alargamento dos ventrículos.

As seguintes figuras (adaptadas de [AGUIAR 2006]) mostram claramente as diferenças entre cérebros de pessoas controles e pessoas com DA.



Figura 4.2: Cérebro humano normal (Adaptado de [AGUIAR 2006]).

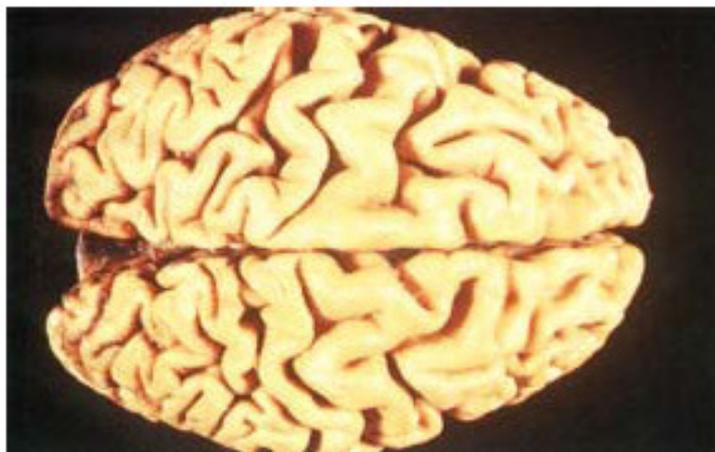


Figura 4.3: Cérebro humano com DA (Adaptado de [AGUIAR 2006]).



Figura 4.4: Cérebro humano normal (Adaptado de [AGUIAR 2006]).



Figura 4.5: Cérebro humano com DA (Adaptado de [AGUIAR 2006]).

A extração do máximo de informações possíveis poderá fazer o diferencial nas análises em RB, para tanto é importante utilizar tanto métodos de diagnóstico por entrevistas quanto por neuroimagens apoiados em padrões de avaliações, isto é, por critérios médicos escalares de pontuação da doença.

O objetivo deste trabalho de mestrado é investigar a aplicação de um modelo em RB capaz de inferir probabilidades a partir de evidências obtidas por estes métodos de análise.

Na próxima seção mostramos o resultado de um teste realizado com algumas imagens com a finalidade de identificar o alargamento dos ventrículos em DA.

4.1 BANCO DE IMAGENS

As imagens utilizadas neste trabalho foram cedidas pelo Instituto de Psiquiatria da Faculdade de Medicina da USP e a permissão ética foi concedida pelo comitê de éticas do Hospital das Clínicas de São Paulo, Universidade de São Paulo (USP). Ao todo são 28 imagens, sendo 14 imagens de controles e 14 de pacientes. Todas foram adquiridas por meio de exames de ressonância magnética [AGUIAR 2006]. Estas imagens não foram apresentadas neste trabalho por questões éticas.

É importante salientar que essas imagens servirão como base para verificar a ocorrência das variáveis escolhidas para a rede neuroimagem de DA. As imagens não foram utilizadas na construção da rede.

4.2 EXPERIMENTO UTILIZANDO STATISTICAL PARAMETRIC MAPPING (SPM2) PARA ANÁLISE UNIVARIADA E MULTIVARIADA

4.2.1 ANÁLISE UNIVARIADA

Statistical Parametric Mapping (SPM2) é um pacote estatístico muito utilizado para análises cerebrais que permite analisar as imagens médicas utilizando técnicas de estatística univariadas. Esse software trabalha com o teste de significância t (modelo linear) para localizar as diferenças que existem em cada voxel (pixel 3D) [FRISTON 1995; AGUIAR 2006], ele faz comparação entre grupos de imagens e destaca as principais diferenças.

A idéia de utilizar esse software foi para conseguirmos destacar as principais diferenças entre as regiões cerebrais e acrescentá-las na Rede Bayesiana que está sendo desenvolvida neste trabalho.

Criamos uma imagem de referência do grupo das 28 imagens baseando-se em uma imagem padrão (TI.img – utilizada mundialmente como uma imagem de referência

para normalização de imagens de ressonância magnética das desordens cerebrais) [TAILARACH 1988], garantindo dessa forma que um determinado voxel de uma imagem esteja na mesma posição das demais. Após a criação da imagem de referência, realizou-se a normalização das imagens para que as imagens estejam alinhadas e assim a análise estatística seja feita de forma correta. O procedimento adotado foi o mesmo utilizado por [AGUIAR 2006].

Após essa etapa fez-se a segmentação das imagens para dividirmos em substância cinzenta, substância branca, Liquor (LCR – Líquido cefalorraquidiano) e tecidos extra-cerebrais, facilitando assim a detecção das regiões afetadas pela doença de Alzheimer, conforme mostra a figura 4.6 [AGUIAR 2006].

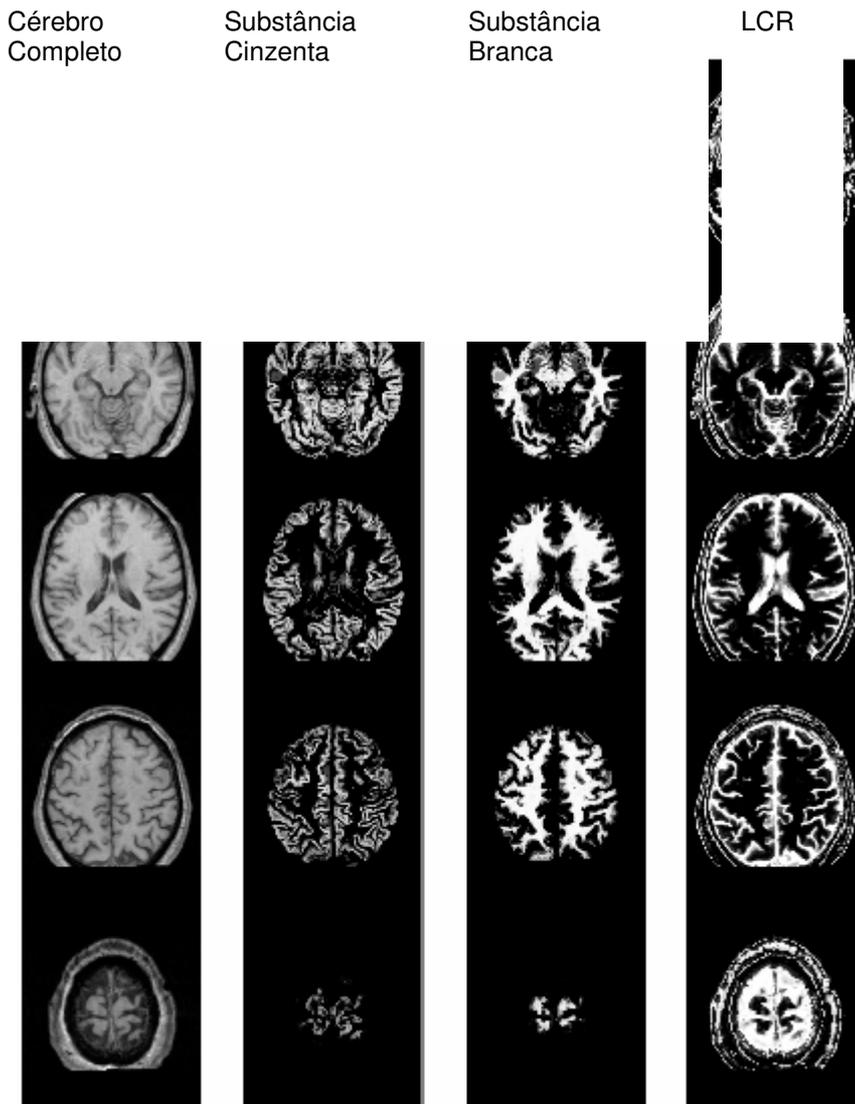


Figura 4.6: Segmentação, adaptado de [AGUIAR 2006].

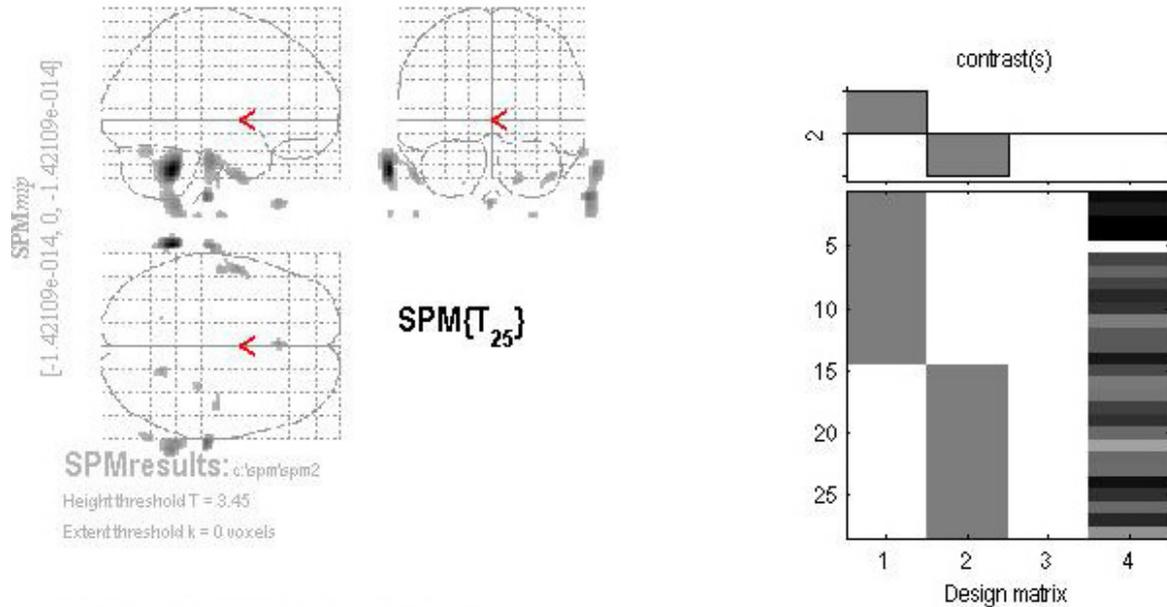
Após a segmentação, realizou-se a suavização para eliminação de eventuais ruídos nas imagens. Após esses três passos, aplica-se a análise estatística univariada, que engloba diferentes tipos de análises, um deles é o teste t de student que é um teste de hipótese para médias. Para este trabalho, foi utilizado o método t de student. E por meio desse cálculo puderam-se encontrar possíveis diferenças entre os cérebros dos controles e pacientes analisados [AGUIAR 2006].

Para visualizarmos quais regiões foram encontradas é necessário o uso de um atlas cerebral. Escolhemos para este trabalho o Atlas de Talairach [TALAIRACH 2008] que atinge nosso objetivo relacionando as coordenadas apresentadas pela análise com as regiões correspondentes do cérebro. Atlas de Talairach tem como referência o atlas

desenvolvido por Talairach e Tournaux [TALAIRACH 1988], portanto, para o utilizamos temos que normalizar estas imagens com a imagem de referência do atlas [ABDALA 2006].

A partir deste ponto apresentaremos resultados que obtivemos neste trabalho de mestrado. Nas figuras 4.7 e 4.8 obtivemos as coordenadas dos pontos onde encontraram-se as maiores diferenças nas regiões cerebrais, mais especificamente onde encontramos o liquor, como por exemplo a região do giro do cíngulo, giro parahipocampal, entre outros. Na figura 4.7 encontram-se as coordenadas da relação paciente / controle (1, -1) e na figura 4.8 a relação controle / paciente (-1, 1), dessas coordenadas escolhemos as três primeiras para verificarmos junto ao Atlas de Tailairach [TAILARACH 1988] quais regiões do cérebro correspondem, porém não foi encontrada nenhuma região, pois como podemos perceber nas imagens de Liquor as diferenças foram encontradas em regiões espalhadas não caracterizando pontos no cérebro.

1-1



Statistics: *p-values adjusted for search volume*

set-level		cluster-level			voxel-level					x,y,z (mm)		
p	c	$P_{corrected}$	k_E	$P_{uncorrected}$	$P_{FWE-corr}$	$P_{FDR-corr}$	T	$(Z_{\bar{c}})$	$P_{uncorrected}$			
0.507	15	0.104	329	0.007	0.009	0.022	7.23	5.26	0.000	-74	-48	-36
		0.951	62	0.198	0.502	0.147	5.09	4.18	0.000	70	-24	-56
		0.075	364	0.005	0.841	0.271	4.60	3.88	0.000	76	-50	-36
					0.931	0.322	4.41	3.76	0.000	70	-46	-62
		0.758	110	0.093	0.868	0.277	4.55	3.85	0.000	-72	-20	-30
		0.982	46	0.265	0.968	0.339	4.28	3.67	0.000	-56	-2	-40
		0.996	31	0.360	0.985	0.364	4.17	3.60	0.000	-2	28	-60
		1.000	11	0.599	0.997	0.413	3.99	3.47	0.000	20	-56	-44
		0.999	22	0.444	0.999	0.452	3.87	3.39	0.000	-56	-14	-72
		0.999	18	0.491	1.000	0.460	3.82	3.36	0.000	44	-16	-40
		1.000	11	0.599	1.000	0.482	3.70	3.27	0.001	30	-30	-70
		0.998	27	0.394	1.000	0.485	3.67	3.25	0.001	66	-68	-22
		1.000	10	0.619	1.000	0.500	3.65	3.23	0.001	-74	-46	-68
		1.000	1	0.902	1.000	0.516	3.55	3.16	0.001	20	-54	-38
		1.000	1	0.902	1.000	0.516	3.46	3.10	0.001	-20	-56	-46
		1.000	1	0.902	1.000	0.517	3.46	3.10	0.001	-62	-10	-32

table shows 3 local maxima more than 8.0mm apart

Height threshold: T = 3.45, p = 0.001 (1.000)
Extent threshold: k = 0 voxels, p = 0.965 (1.000)
Expected voxels per cluster, <k> = 39.935
Expected number of clusters, <c> = 14.73
Expected false discovery rate <= 0.52

Degrees of freedom = [1.0, 25.0]
Smoothness FWHM = 14.1 16.2 16.0 (mm) = 7.1 8.1 8.0 (voxels)
Search vol: 4343944 cmm; 542993 voxels; 1143.2 resels
Voxel size: [2.0, 2.0, 2.0] mm (1 resel = 454.92 voxels)

Figura 4.7: Resultado SPM2 do Liquor mostrando as coordenadas das regiões cerebrais, na proporção controle x paciente (1, -1) .

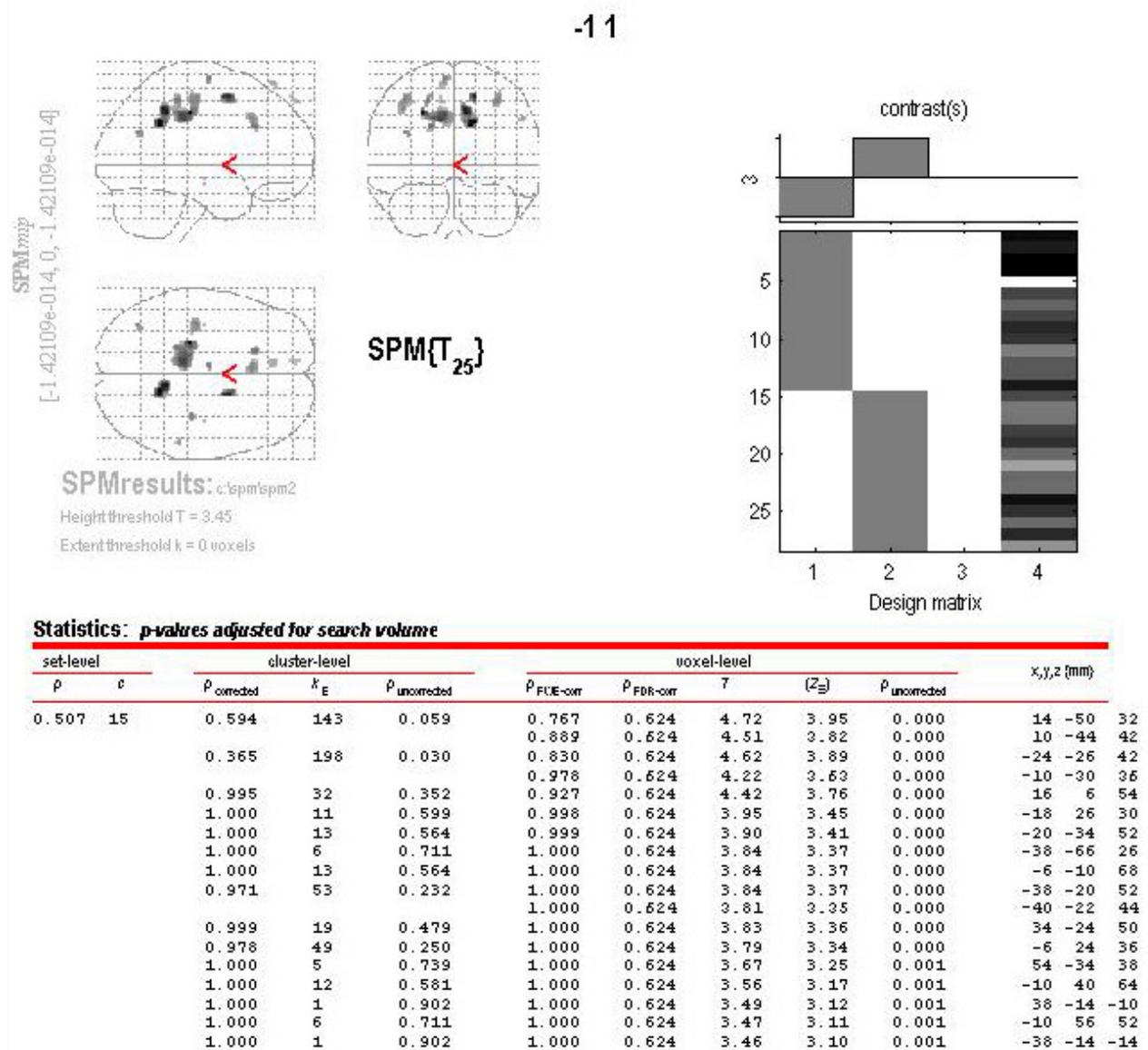
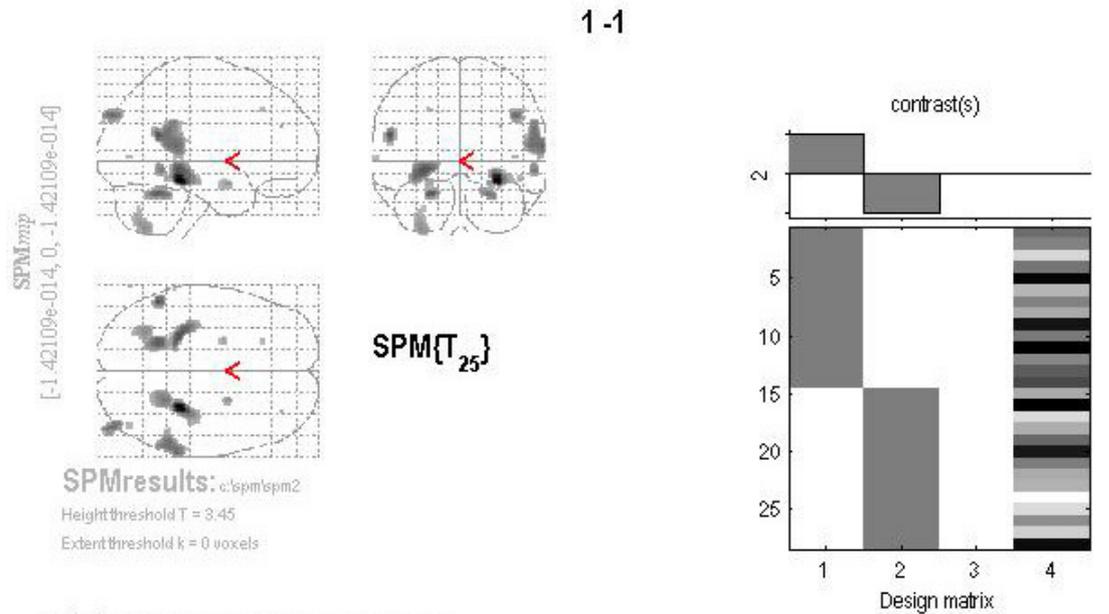


Figura 4.8: Resultado SPM2 do Liquor mostrando as coordenadas das regiões cerebrais, na proporção paciente x controle (-1, 1) .

Nas figuras 4.9 e 4.10 mostram as principais diferenças encontradas nas regiões da substância cinza, onde apresentam as coordenadas da relação paciente / controle (1, -1) e controle / paciente (-1, 1) respectivamente. Quando pesquisamos as três primeiras coordenadas (figura 4.9 e 4.10) no atlas de Talairach encontramos as seguintes regiões do cérebro: giro para-hipocampal, giro fusiforme, lobo temporal, giro temporal superior, giro lateral supramarginal e ventrículo lateral.



Statistics: p-values adjusted for search volume

set-level		cluster-level			voxel-level					x,y,z (mm)
p	c	p _{corrected}	k _E	p _{uncorrected}	p _{FWE-corr}	p _{FDR-corr}	T	(Z _z)	p _{uncorrected}	
0.487	18	0.042	239	0.002	0.134	0.177	5.96	4.66	0.000	28 -36 -14
		0.594	84	0.049	0.790	0.177	4.78	3.99	0.000	-54 -52 20
		0.031	257	0.002	0.793	0.177	4.78	3.99	0.000	-28 -36 -8
		0.360	115	0.024	0.863	0.177	4.66	3.92	0.000	-22 -52 -24
		0.003	407	0.000	0.933	0.177	4.50	3.82	0.000	64 -38 12
					0.959	0.177	4.42	3.76	0.000	56 -44 12
					0.970	0.177	4.37	3.73	0.000	58 -44 24
		0.433	104	0.031	0.951	0.177	4.45	3.78	0.000	44 -88 36
		0.330	120	0.022	0.980	0.177	4.31	3.69	0.000	24 -46 -26
		0.971	34	0.191	0.986	0.177	4.27	3.66	0.000	58 -50 -8
		0.209	146	0.013	0.998	0.177	4.05	3.52	0.000	-30 -64 -52
					1.000	0.212	3.54	3.16	0.001	-30 -68 -38
		0.999	17	0.353	0.999	0.179	3.96	3.45	0.000	24 2 -18
		0.997	20	0.313	1.000	0.191	3.74	3.30	0.000	-24 -2 -18
		1.000	1	0.854	1.000	0.200	3.66	3.24	0.001	32 44 28
		1.000	4	0.672	1.000	0.201	3.64	3.23	0.001	42 -76 4
		1.000	1	0.854	1.000	0.201	3.64	3.23	0.001	-66 -74 -18
		1.000	6	0.594	1.000	0.204	3.62	3.22	0.001	48 -46 -14
1.000	4	0.672	1.000	0.214	3.53	3.15	0.001	-66 -40 2		
			1.000	0.631	1.000	0.214	3.53	3.15	0.001	-24 30 40
			1.000	0.854	1.000	0.214	3.52	3.14	0.001	46 -96 30

Figura 4.9: Resultado SPM2 da substância cinza mostrando as coordenadas das regiões cerebrais, na proporção controle x paciente (1, -1) .

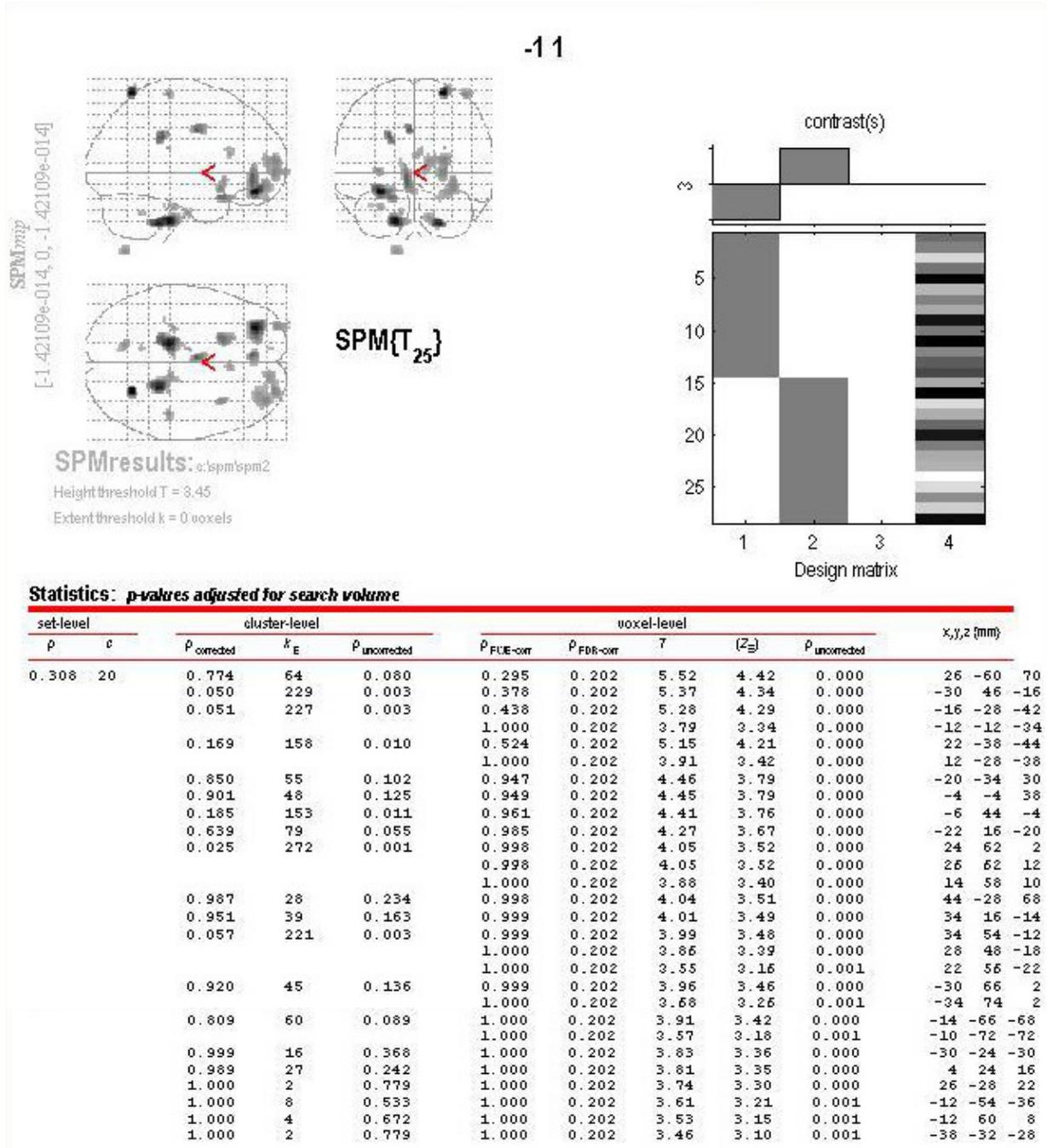


Figura 4.10: Resultado SPM2 da substância cinza mostrando as coordenadas das regiões cerebrais, na proporção paciente x controle (-1, 1).

As figuras 4.11 e 4.12 mostram as principais diferenças encontradas nas regiões da substância branca, e também apresentam as coordenadas da relação paciente / controle (1, -1) e controle / paciente (-1, 1) respectivamente. Igualmente o passo anterior, quando pesquisamos as três primeiras coordenadas (figura 4.11 e 4.12) no

atlas de Talairach encontramos as seguintes regiões do cérebro: lobo parietal, giro poscentral e precentral, lobo frontal, giro médio frontal, lobo temporal e lobo occipital.

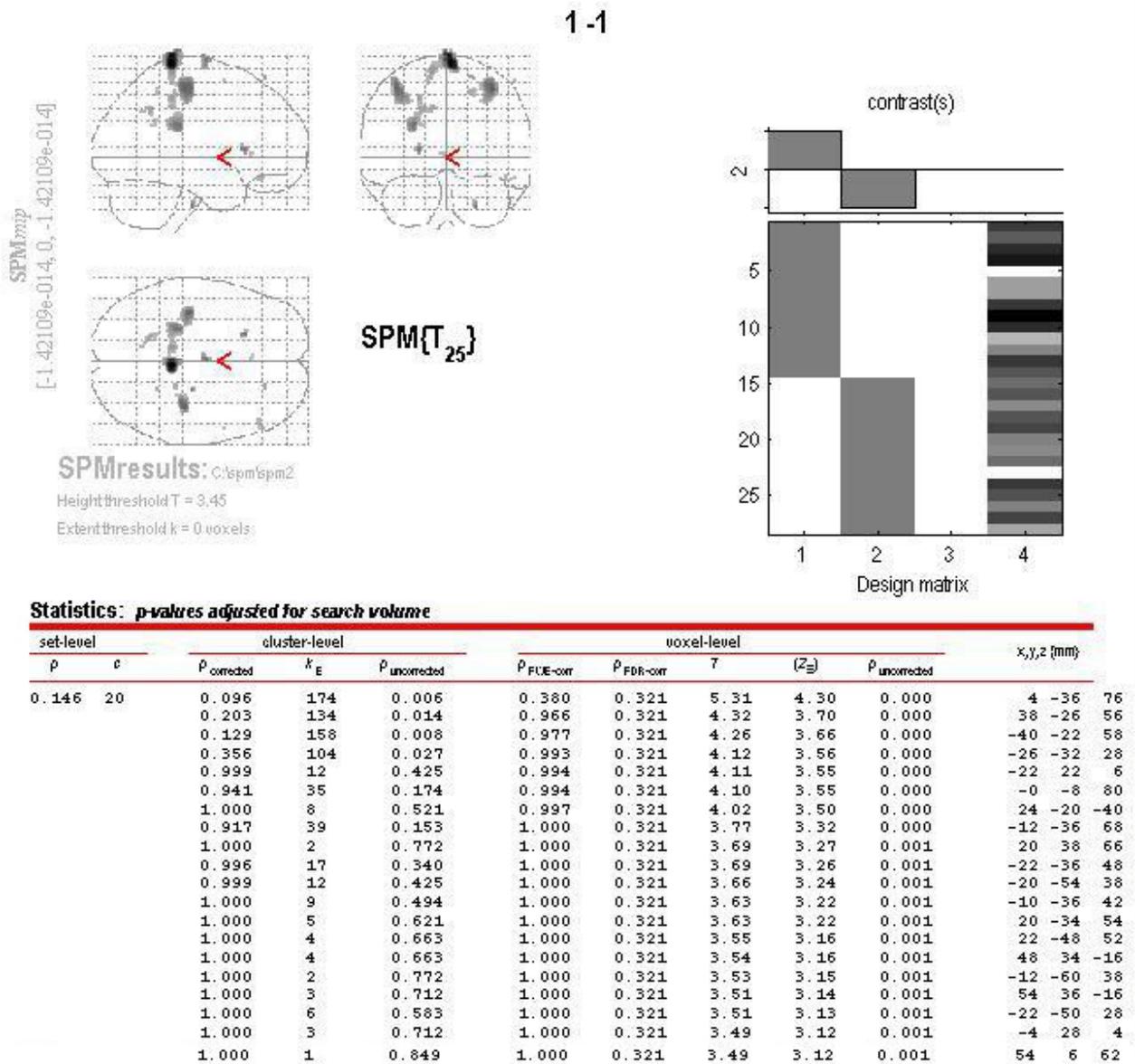


Figura 4.11: Resultado SPM2 da substância branca mostrando as coordenadas das regiões cerebrais, na proporção controle x paciente (1, -1).

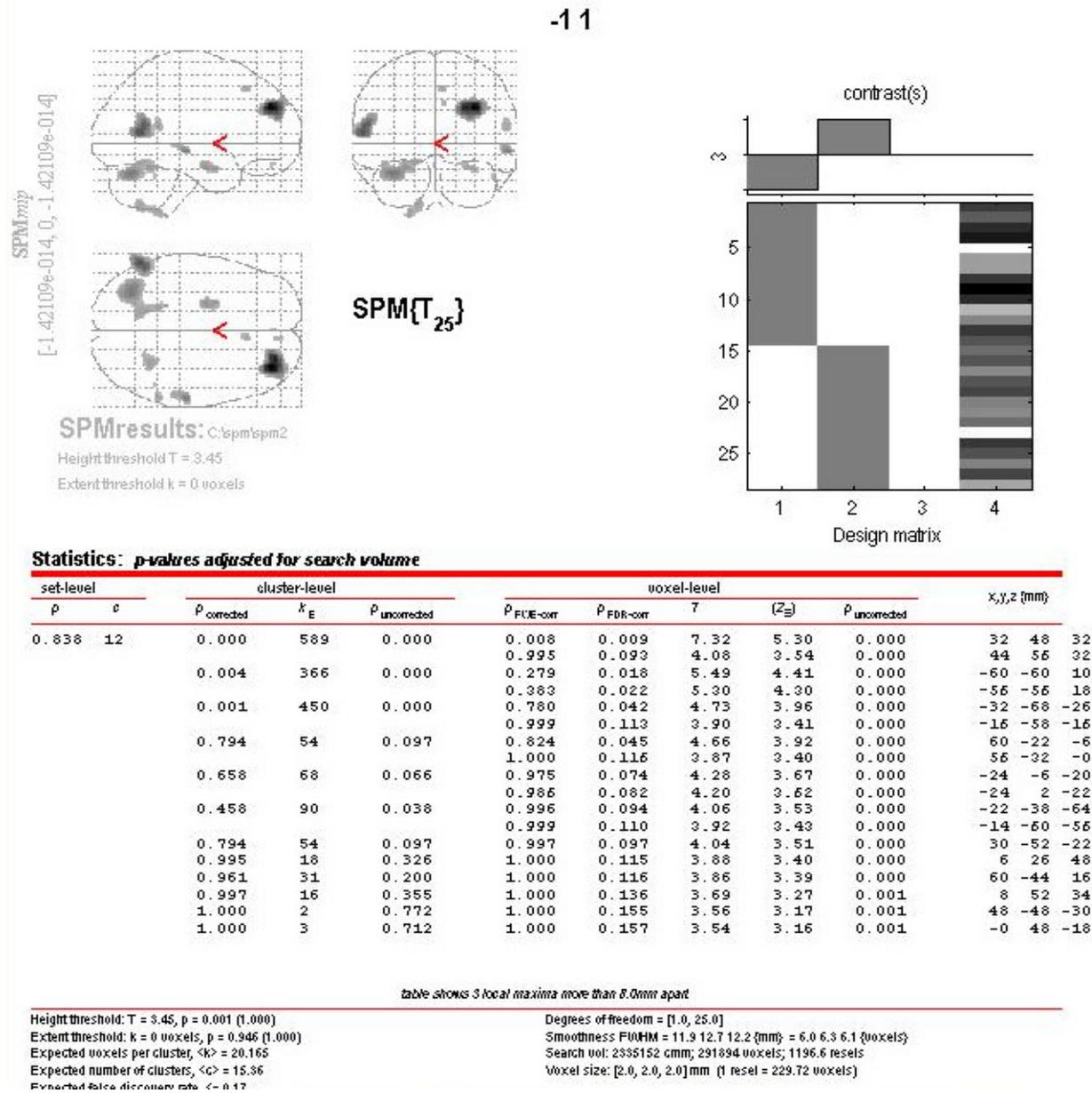


Figura 4.12: Resultado SPM2 da substância branca mostrando as coordenadas das regiões cerebrais, na proporção paciente x controle (-1, 1) .

Das coordenadas encontradas pelo SPM2 e apresentadas da figura 4.7 até a 4.12, escolhemos de todas as imagens as que apresentaram o valor de *t de student* maior que 5,0, pois, quanto maior o valor de t, maior a significância ou relevância da região encontrada [AGUIAR 2006]. A partir daí, como já vimos, utilizamos o Atlas de Talairach [TAILARACH 1988] para sabermos quais regiões são e em resumo podemos descrever as seguintes regiões encontradas: Giro do cíngulo, giro para-hipocampal, giro temporal, giro fusiforme, giro poscentral e precentral, lobo parietal, lobo occipital, lobo

temporal dentre outras. Por meio dessas regiões, juntamente com a bibliografia [CHARCHAT 2001; GOLDMAN 1992; MAURES 1997; DAVATZIKOS 2006; NITRINI 2005; AZEVEDO 2005; PASCALE 2002 e RAMOS 2001], podemos usar para construção dos nós da neuroimagem em nossa rede.

4.2.2 ANÁLISE MULTIVARIADA – MAXIMUM UNCERTAINTY LINEAR DISCRIMINANT ANALYSIS (MLDA)

Na análise multivariada utilizaram-se as mesmas imagens criadas como referência para a análise univariada, porém após a normalização, retirou-se o crânio das imagens utilizando o software MRIcro (Rorden; Brett, 2000 [AGUIAR 2006], deixando na imagem apenas o que é relevante para o estudo.

Diferente da análise univariada, que compara voxel a voxel de cada imagem e depois tira-se a média, a análise multivariada é usada para comparar imagens inteiras entre grupos, em nosso caso, usamos para comparar as imagens dos grupos controle e paciente. Outra diferença, é que na análise univariada como faz-se a comparação voxel a voxel, obtém-se inicialmente resultados separados para cada substância (cinzenta, branca e liquor) e é necessário realizar análise para verificar a expansão de áreas e outra análise para verificar redução de áreas; enquanto na análise multivariada é possível visualizar todas as alterações (em todas as substâncias, redução e expansão) ao mesmo tempo, permitindo que seja encontrada outras áreas de alterações no cérebro sem que seja necessário ajustes de programação no software.

Neste trabalho utiliza-se o método PCA (*Principal Component Analysis* ou Análise de Componentes Principais (ACP)) + MLDA (Maximum Uncertainty Linear Discriminant Analysis) para fazer a separação dos dois grupos. MLDA trabalha somente com uma transformação linear e combina duas técnicas de análise estatística multivariada PCA e LDA (Linear Discriminant Analysis), maiores informações em [AGUIAR 2006]. As imagens são transformadas em vetores multi-dimensionais, formando uma matriz de dados, após esse processo aplicou-se a matriz de transformação do PCA com a finalidade de reduzir dimensão de cada imagem deixando

apenas os componentes principais (dessa forma facilita o cálculo do MLDA, pois reduz o custo computacional). Com isso obtêm-se autovetores que contêm uma ou mais características cerebrais das imagens analisadas como mostra a figura 4.13, sendo que o primeiro autocérebro (autovetor) contém 32,05% de todas as características estudadas entre os dois grupos, o segundo apresenta 4,98% e o terceiro apresenta 4,57% [AGUIAR 2006; THOMAZ 2004].

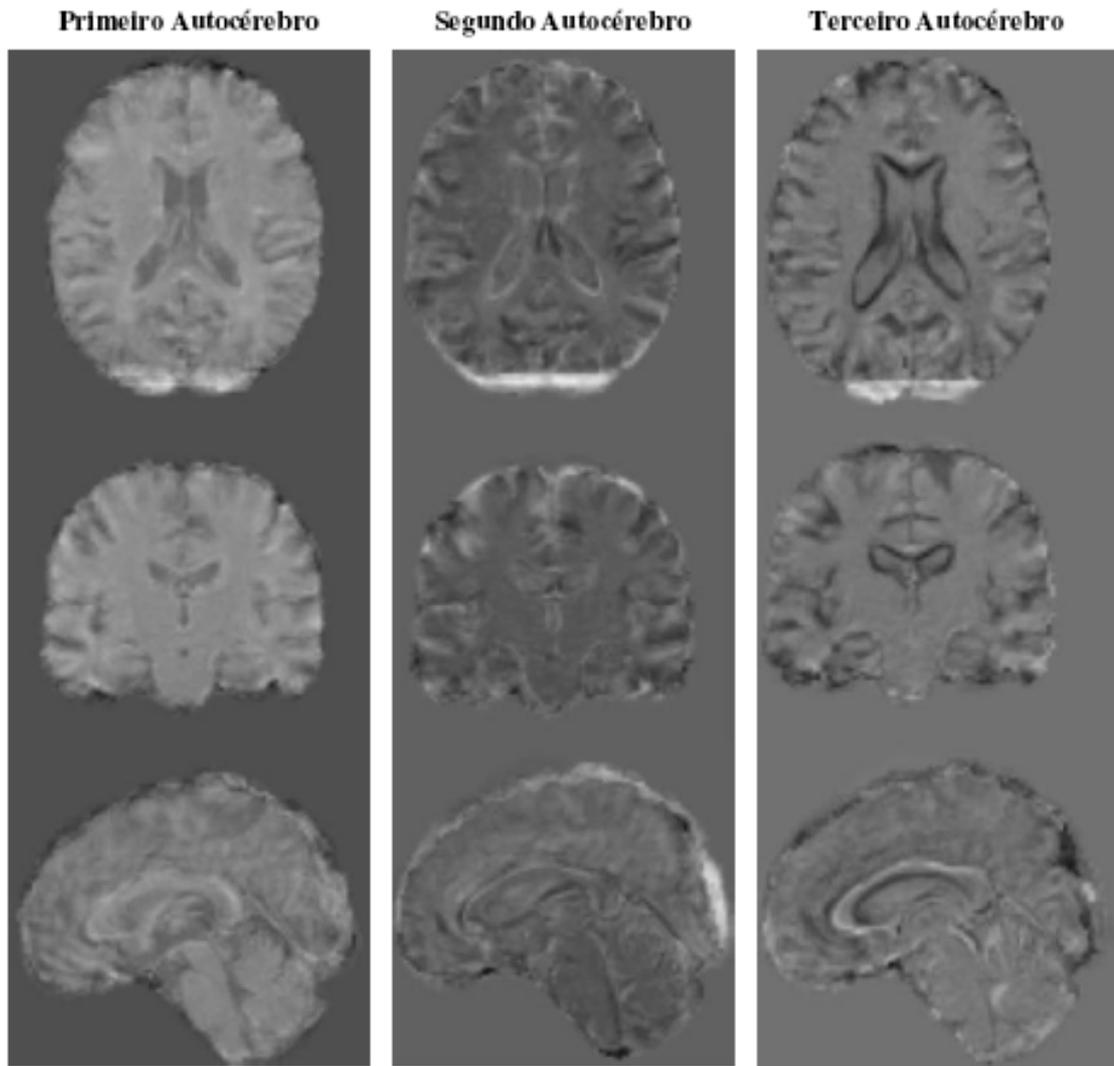


Figura 4.13: Ilustração dos três primeiros autovetores adaptado de [AGUIAR 2006].

Após essa etapa é aplicado o modelo estatístico MLDA para visualizar os desvios padrões das médias dos grupos dos pacientes e dos controles como mostra a figura 4.14, onde os “x” são indicações de pacientes e “O” indicam controles. Essa

técnica separa as imagens (dos 28 casos estudados) e também permite a extração de descrições entre eles. Além de reconstruir as imagens estudadas para possibilitar a navegação no hiperplano MLDA e dessa forma perceber as diferenças entre os grupos [AGUIAR 2006].



Figura 4.14: Autovetor PCA + MLDA. Adaptado de [AGUIAR 2006].

Em [AGUIAR 2006] apresenta-se uma tabela com a comparação entre as principais diferenças encontradas nas análises univariada e multivariada, representada neste trabalho na tabela 4.1.

Tabela 4.1: Tabela de comparações das regiões cerebrais entre as análises univariada e multivariada. Adaptado de [AGUIAR 2006].

Região do cérebro	Análise Univariada	Análise Multivariada
Hipocampo	Redução	Redução
Córtex	Redução	Redução
Ventrículos	Não relevante	Expansão
Giro do Cíngulo	Expansão no LCR	Redução
	Redução	

As imagens com os resultados das diferenças apresentadas na tabela 4.1 são mostrados nas figuras 4.15, 4.16 e 4.17 com as devidas regiões destacadas, onde o amarelo indica os pontos mais significantes e o vermelho indica os pontos menos significantes [AGUIAR 2006].

A figura 4.15 mostra as principais diferenças entre a região do ventrículo cerebral, onde ao lado esquerdo da figura apresenta-se imagem de controle e ao lado direito, paciente. Podemos perceber nitidamente essas diferenças na figura 4.16 que a região do ventrículo é destacada com a segmentação (contorno em amarelo).

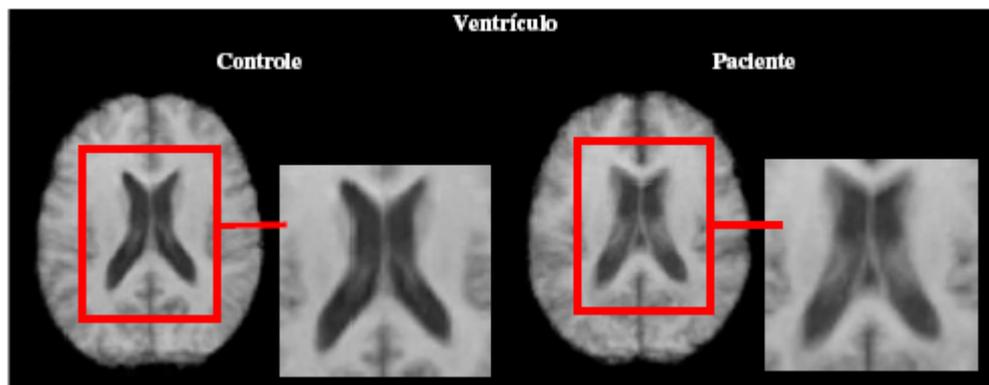


Figura 4.15: Ventrículo. Adaptado de [AGUIAR 2006].

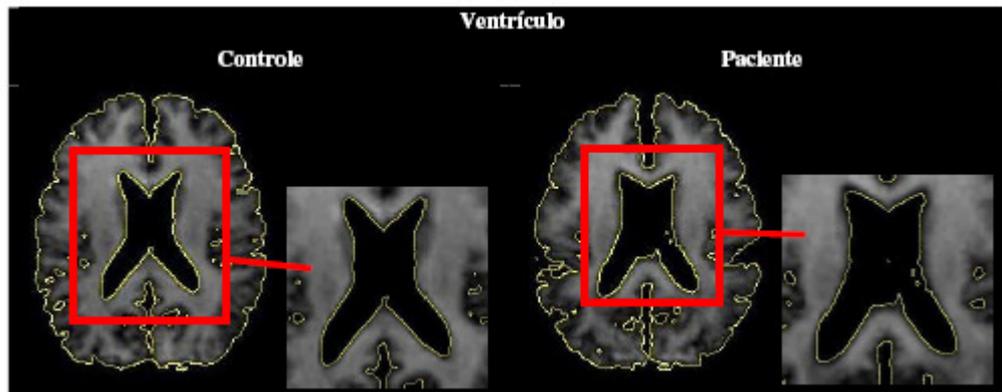


Figura 4.16: Ventrículo contornado para melhor visualização. Adaptado de [AGUIAR 2006].

As figuras 4.17 e 4.18 mostram as principais diferenças entre a região do hipocampo cerebral, onde na parte de cima de cada figura apresenta-se imagem de controle e embaixo o paciente. Podemos perceber nitidamente essas diferenças na figura 4.18 onde a região do ventrículo é destacada com a segmentação (contorno em amarelo).

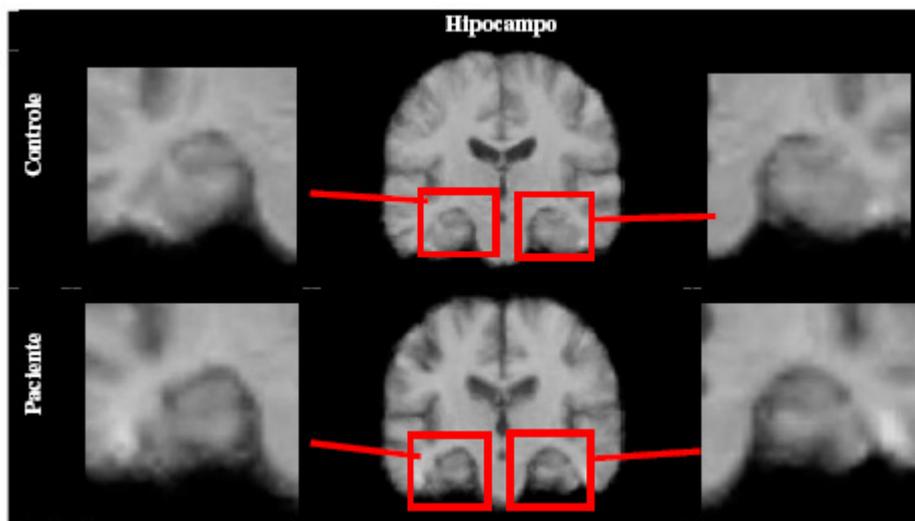


Figura 4.17: Hipocampo. Adaptado de [AGUIAR 2006].

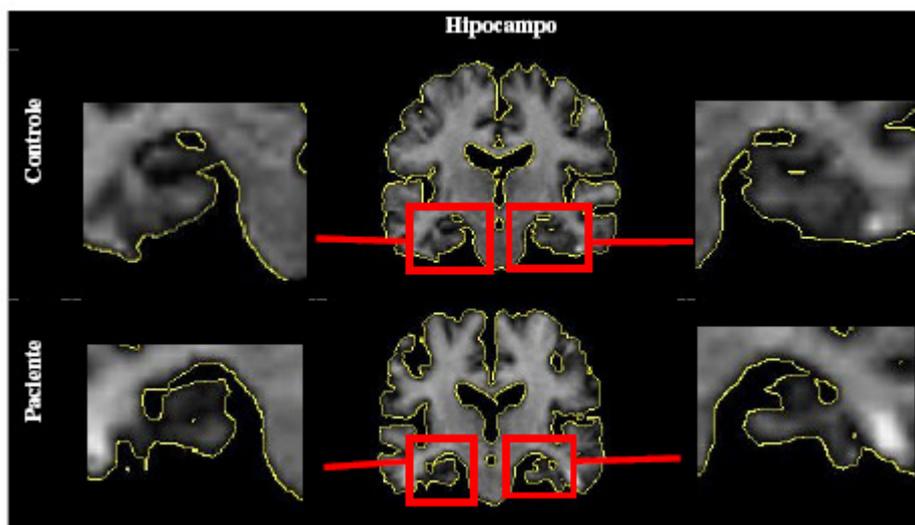


Figura 4.18: Hipocampo contornado para melhor visualização. Adaptado de [AGUIAR 2006].

A figura 4.19 mostra as principais diferenças entre a região do giro do cíngulo, onde na parte de cima da figura ilustra-se a imagem do controle e abaixo a do paciente. Podemos perceber nitidamente essas diferenças, pois já está destacada com a segmentação (contorno em amarelo).

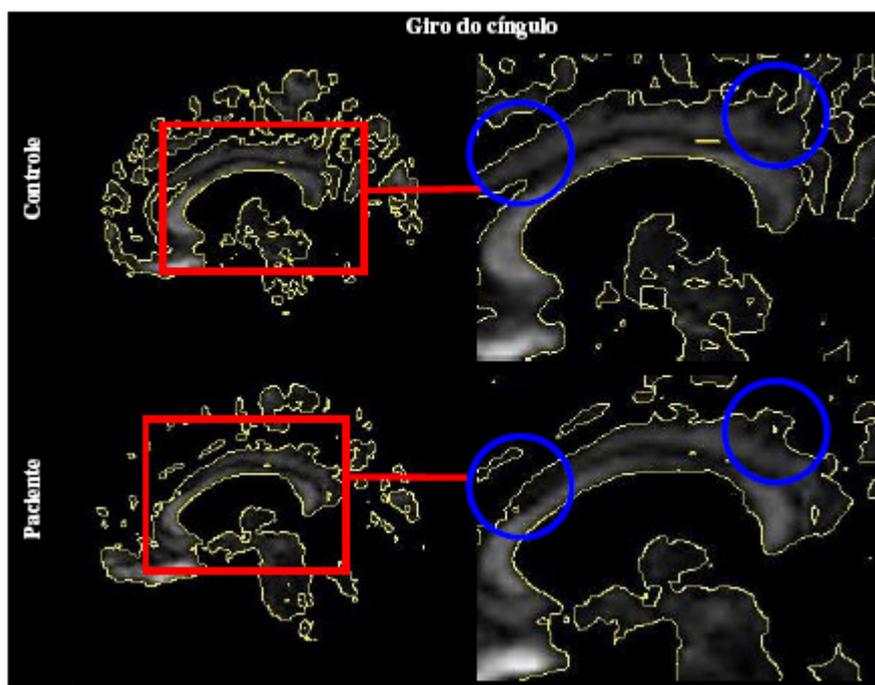


Figura 4.19: Giro do cíngulo contornado para melhor visualização. Adaptado de [AGUIAR 2006].

A figura 4.20 mostra as principais diferenças entre a região do córtex cerebral, sendo ao lado esquerdo, apresentada as imagens do controle e ao lado direito as do paciente. Podemos perceber nitidamente essas diferenças na parte de baixo da figura, pois está destacada com a segmentação (contorno em amarelo).

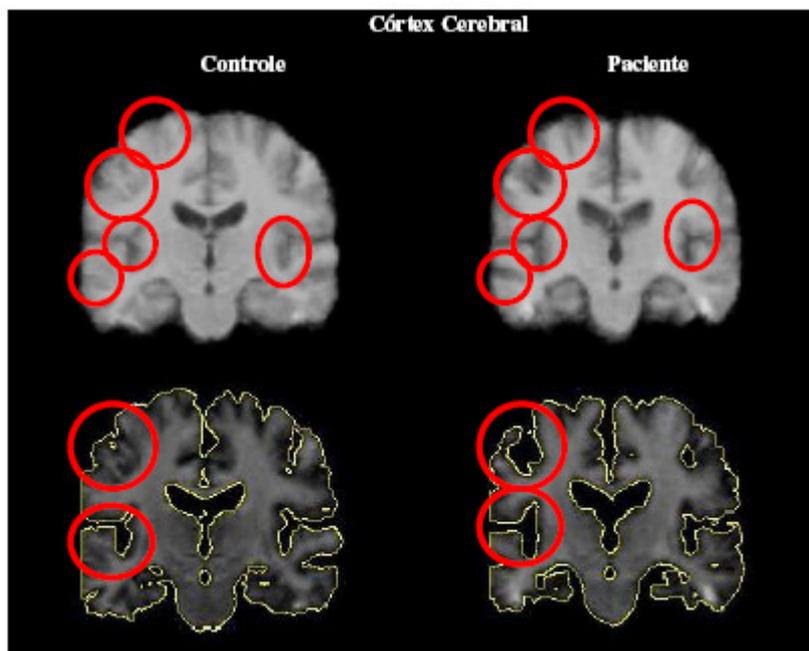


Figura 4.20: Córtex parietal. Adaptado de [AGUIAR 2006].

As figuras 4.21, 4.22 e 4.23 mostram a comparação das análises univariada e multivariada das mesmas regiões citadas acima. Na figura acima está a imagem da análise univariada e mostra-se em cores amarelo para maior significância e vermelho para menor, as regiões encontradas com maiores diferenças. Já na análise multivariada podemos perceber as diferenças tanto para expansão da região apontada quanto para redução [AGUIAR 2006].

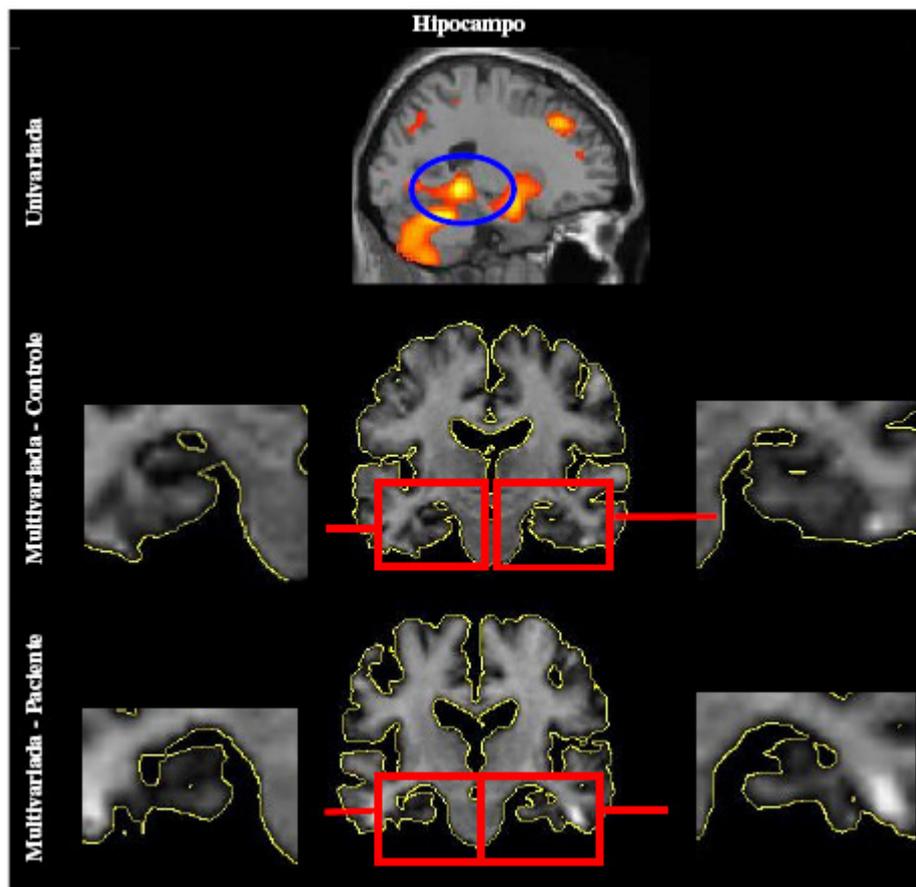


Figura 4.21: Hipocampo visualizado na análise univariada e multivariada de paciente e controle. Adaptado de [AGUIAR 2006].

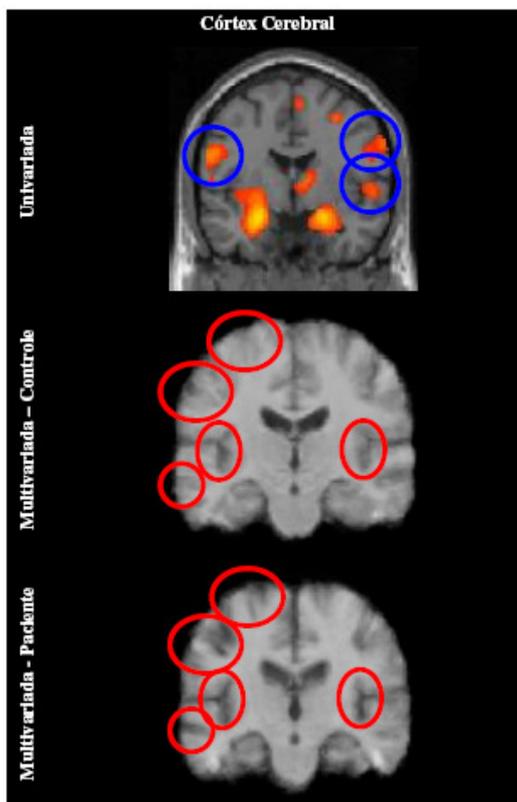


Figura 4.22: Córtex visualizado na análise univariada e multivariada de paciente e controle. Adaptado de [AGUIAR 2006].

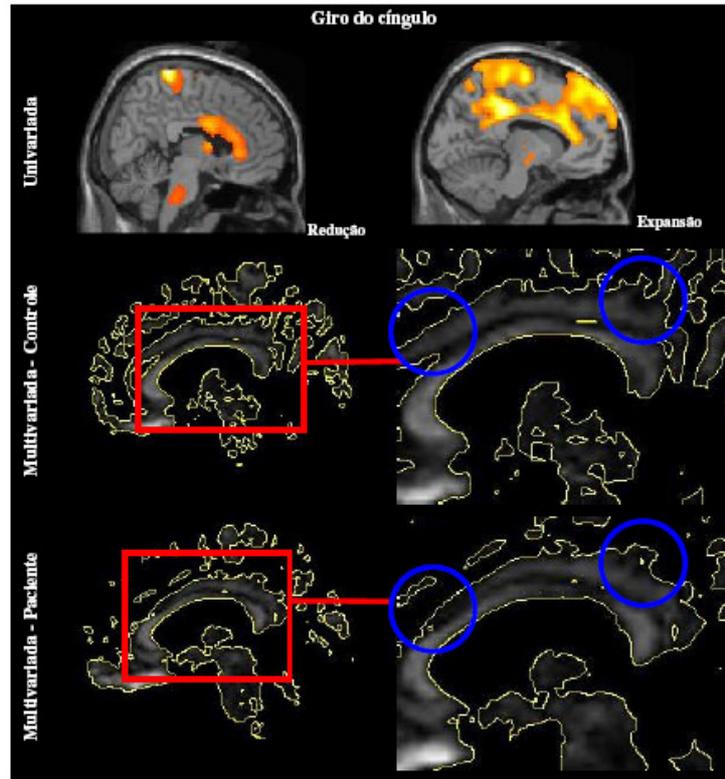


Figura 4.23: Giro do cíngulo visualizado na análise univariada e multivariada de paciente e controle. Adaptado de [AGUIAR 2006].

Após o uso dessas duas análises escolhemos as estruturas cerebrais cujas diferenças entre controle e paciente mostraram-se mais significantes para serem acrescentadas na Rede Bayesiana desenvolvida neste trabalho. Utilizaremos, portanto, dados que foram encontrados na análise multivariada e univariada para construir o nó neuroimagem em nossa rede, com as seguintes variáveis como filhos: cíngulo posterior, hipocampo / giro parahipocampal, ventrículo e córtex parietal.

Em concordância com o especialista consultado, optamos por utilizar os dados destacados nas figuras 4.15 a 4.23 encontradas pela análise multivariada por ser nítida a visualização das regiões do cérebro, enquanto que na análise univariada destacam-se inúmeros pontos, dificultando a clareza da visualização das regiões.

No próximo capítulo será explicado de onde extraímos os nós que completam nossa rede: histórico e sintomas.

5 CONSTRUÇÃO E ANÁLISE DA RB

Neste capítulo será apresentada a construção da rede e os resultados obtidos em experimentos realizados na rede Bayesiana para o diagnóstico da DA desenvolvido neste trabalho.

A princípio a RB foi construída com probabilidades a priori de 0,5 (considerando que a cada cem pessoas que consultam o especialista, cinquenta se queixem de alguns dos sintomas, apresentem algum histórico ou tenham algum dos resultados de neuroimagem positiva, características da DA), deste modo, consideramos como variáveis do problema (nós) o histórico do paciente dadas as evidências: hereditariedade, idade e grau de escolaridade; os sintomas da doença, dada as evidências: apraxia, afasia e perda de memória; e os diagnósticos por neuroimagem, dada as evidências: cíngulo posterior, hipocampo / giroparahipocampal, alargamento do ventrículo e atrofia do córtex parietal. Os resultados de neuroimagem foram tratados como descrito na seção anterior.

Primeiramente, foi agregado à RB o grau de relevância para cada variável, ou seja, o quão importante é cada variável para se chegar ao diagnóstico de DA. Na tabela 5.1 demonstramos a distribuição da relevância para cada nó que compõem a rede inteira. Os valores para relevância foram definidos em entrevistas com um médico psiquiatra, especialista em DA

Tabela 5.1: Distribuição de relevância para o nó Histórico

Variáveis condicionais	Variáveis	Probabilidade
Alzheimer	Neuroimagem	0,8
Alzheimer	Sintomas	0,8
Alzheimer	Histórico	0,8
Neuroimagem	Cíngulo Posterior	0,4
Neuroimagem	Hipocampo	0,7
Neuroimagem	Ventrículo	0,8
Neuroimagem	Córtex	0,9
Sintomas	Apraxia	0,5
Sintomas	Afasia	0,5
Sintomas	Perda	0,9
Histórico	Hereditariedade	0,3
Histórico	Escolaridade	0,3
Histórico	Idade	0,9

A distribuição das probabilidades das tabelas de probabilidades condicionais (TPC) foram baseadas no método Noisy-or descrito na seção 3.4 (demonstrada na tabela 5.2).

Tabela 5.2: Noisy-or aplicado aos nós: Histórico, Sintomas e Neuroimagem (nós filhos do nó Alzheimer).

Histórico (H)	Sintomas (S)	Neuroimagem (N)	Multiplicação das variáveis T	$P(\neg\text{Alzheimer} H, S, N)$
80%	80%	80%		
20%	20%	20%	← Porcentagem complementar de não apresentar	
F	F	F	0	100,00%
F	F	V	0,20	20,00%
F	V	F	0,20	20,00%
F	V	V	$0,20 * 0,20$	4,00%
V	F	F	0,20	20,00%
V	F	V	$0,20 * 0,20$	4,00%
V	V	F	$0,20 * 0,20$	4,00%
V	V	V	$0,20 * 0,20 * 0,20$	0,80%

O mesmo fazemos para os demais nós em nossa rede, a tabela 5.3 mostra a distribuição de probabilidade condicional o nó neuroimagem seguindo a mesma lógica da distribuição do nó Alzheimer, e assim fazemos para todos os demais nós da rede.

Tabela 5.3: Noisy-or aplicado aos nós: Cíngulo Posterior, Hipocampo, Ventrículo e Córtex (nós filhos do nó Neuroimagem).

Cíngulo Posterior (P)	Hipocampo (H)	Ventrículo (V)	Córtex (C)	Multiplicação das variáveis T	$P(\neg\text{Neuroimagem} P, H, V, C)$
40%	50%	60%	70%		
60%	50%	40%	30%	← Porcentagem complementar de não apresentar	
F	F	F	F	0	100,00%
F	F	F	V	0,30	10,00%
F	F	V	F	0,40	20,00%
F	F	V	V	$0,40 * 0,30$	2,00%
F	V	F	F	0,50	30,00%
F	V	F	V	$0,50 * 0,30$	3,00%
F	V	V	F	$0,50 * 0,40$	6,00%
F	V	V	V	$0,50 * 0,40 * 0,30$	0,60%
V	F	F	F	0,60	60,00%
V	F	F	V	$0,60 * 0,30$	6,00%
V	F	V	F	$0,60 * 0,40$	12,00%
V	F	V	V	$0,60 * 0,50 * 0,30$	1,20%
V	V	F	F	$0,60 * 0,50$	18,00%
V	V	F	V	$0,60 * 0,50 * 0,30$	1,80%
V	V	V	F	$0,60 * 0,50 * 0,40$	3,60%
V	V	V	V	$0,60 * 0,50 * 0,40 * 0,30$	0,36%

Na tabela 5.4 demonstramos a distribuição dos nós filhos do nó Sintomas

Tabela 5.4: Noisy-or aplicado aos nós: Afasia, Perda_de_memória e Apraxia (nós filhos do nó Sintomas).

Apraxia (AP)	Afasia (AF)	Perda_de_memória (PM)	Multiplicação das variáveis T	P(\neg Sintomas AP, AF, PM)
50%	50%	90%		
50%	50%	10%	← Porcentagem complementar de não apresentar	
F	F	F	0	100,00%
F	F	V	0,10	10,00%
F	V	F	0,50	50,00%
F	V	V	$0,50 * 0,10$	5,00%
V	F	F	0,50	50,00%
V	F	V	$0,50 * 0,10$	5,00%
V	V	F	$0,50 * 0,50$	25,00%
V	V	V	$0,50 * 0,50 * 0,10$	2,50%

A tabela 5.5 mostra a distribuição de probabilidade condicional o nó histórico.

Tabela 5.5: Noisy-or aplicado aos nós: Hereditariedade, Escolaridade e Idade (nós filhos do nó Histórico).

Hereditariedade (H)	Escolaridade (E)	Idade (I)	Multiplicação das variáveis T	P(\neg histórico H, E, I)
30%	30%	90%		
70%	70%	10%	← Porcentagem complementar de não apresentar	
F	F	F	0	100,00%
F	F	V	0,10	10,00%
F	V	F	0,70	70,00%
F	V	V	$0,70 * 0,10$	7,00%
V	F	F	0,70	70,00%
V	F	V	$0,70 * 0,10$	7,00%
V	V	F	$0,70 * 0,70$	49,00%
V	V	V	$0,70 * 0,70 * 0,10$	4,90%

Com as distribuições, realizamos inferências na RB no intuito de conferir se os resultados apresentados se mostraram consistentes, lembrando que esta rede foi construída com valores aleatórios somente para testar os cálculos. Ainda neste capítulo mostraremos a RB construída com valores baseados na literatura e no conhecimento do especialista consultado.

A RB obtida através dessas distribuições está apresentada na figura 5.1.

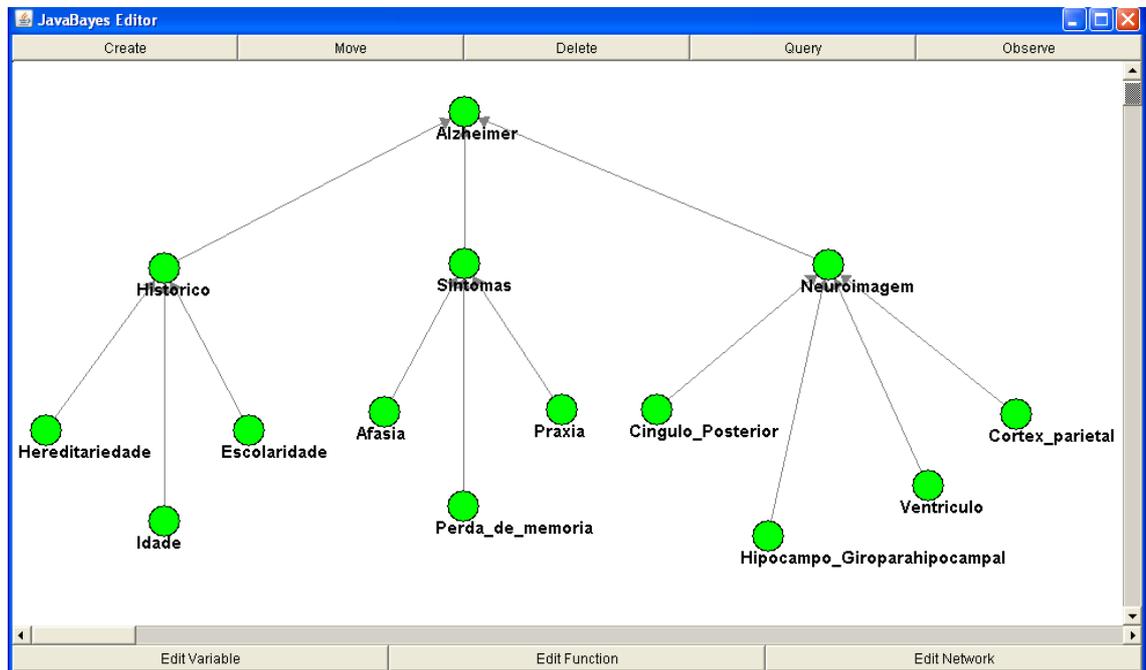


Figura 5.1: Representação gráfica da rede – Alzheimer.

A partir do modelo construído fizemos diferentes observações, verificando-se a probabilidade da incidência de Alzheimer, conforme os resultados abaixo relacionados (as variáveis observadas, tanto se apresentar ou não determinada variável, ficam na cor azul). Para observar as variáveis, clica-se na guia “observe” e marca qual observação deverá ser feita, no caso da variável idade escolhe-se “Media ou Avançada” como mostra a figura 5.2:

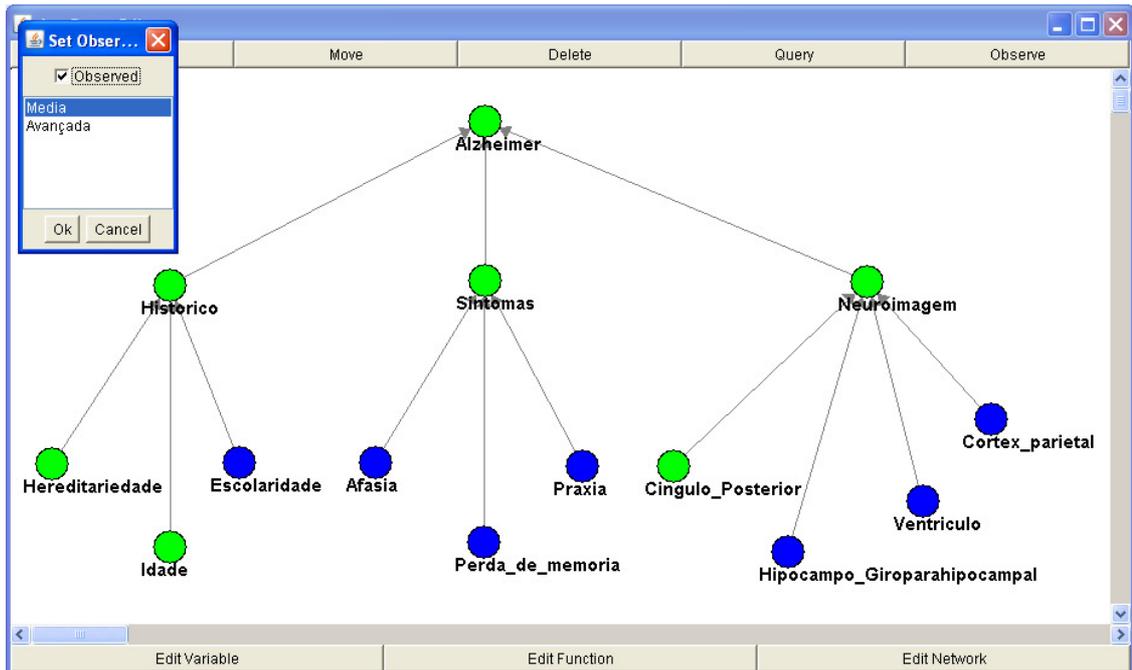


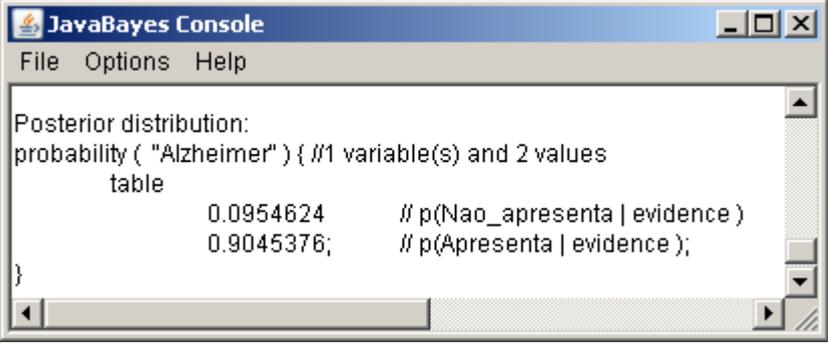
Figura 5.2: Representação gráfica da rede com observações descritas na tabela 5.9.

Neste ensaio fizemos as seguintes suposições (observações) apresentadas na tabela 5.6 (F quando a variável é falsa, V quando é verdadeira e X quando não se sabe seu valor).

Tabela 5.6 : Distribuição das observações feitas nas variáveis destacadas em azul, conforme figura 5.2.

Variáveis	Observações
Hereditariedade	X
Idade	V
Escolaridade	V
Afasia	V
Perda de memória	F
Apraxia	V
Cíngulo posterior	X
Hipocampo_giroparahipocampal	V
Ventrículo	V
Córtex_parietal	V

O resultado do ensaio foi o seguinte (figura 5.3):



```
JavaBayes Console
File Options Help

Posterior distribution:
probability ( "Alzheimer" ) { //1 variable(s) and 2 values
  table
    0.0954624      // p(Nao_apresenta | evidence )
    0.9045376;    // p(Apresenta | evidence );
}
```

Figura 5.3: Resultado das inferências realizadas representada pela figura 5.2.

A figura 5.3 apresenta que no caso estudado na figura 5.2 a probabilidade de ser diagnosticado a DA, dado as evidências observadas é de 0,90. O valor se mostra coerente baseado que as variáveis observadas na rede: idade, escolaridade, afasia, apraxia, hipocampo / giroparahipocampal, ventrículo e córtex parietal são de extrema relevância para se aproximar do diagnóstico de DA conforme consulta com especialista.

Fizemos outras observações na RB para verificarmos se as respostas eram consistentes com a realidade. Na figura 5.4 a maioria das variáveis foram observadas como “não apresenta”, deste modo, percebeu-se que o valor da probabilidade de apresentar a DA tende a diminuir.

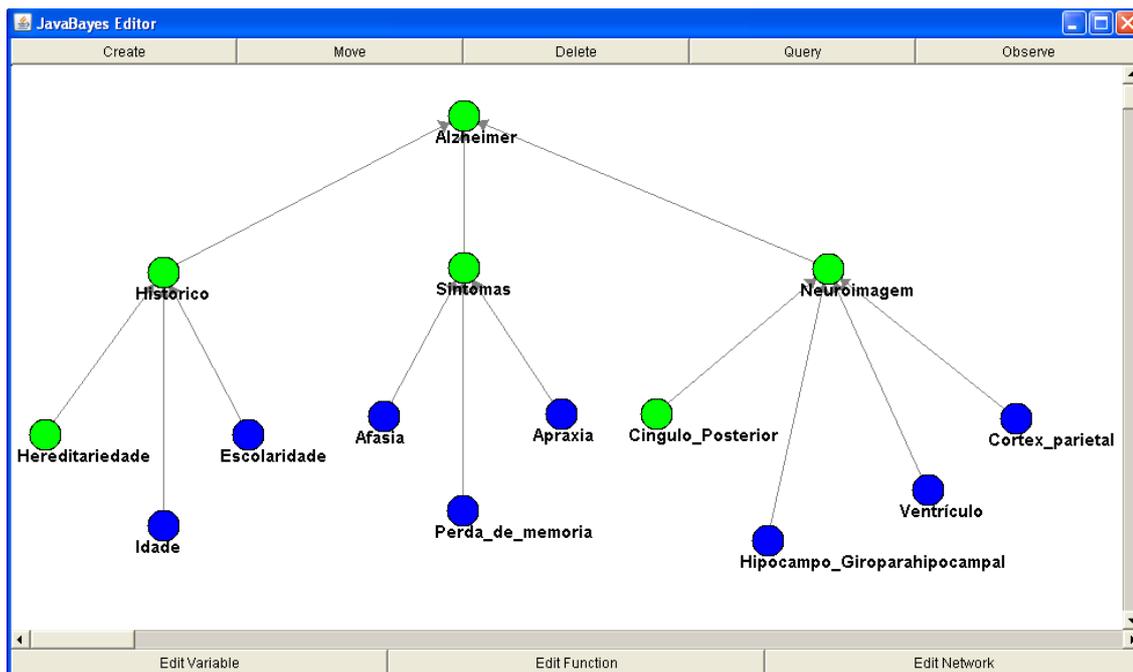


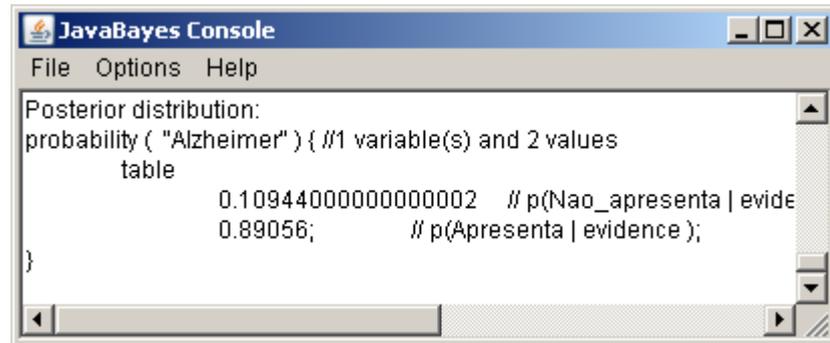
Figura 5.4: Representação gráfica da rede com observações descritas na tabela 5.10.

Neste ensaio fizemos as seguintes suposições (observações) apresentadas na tabela 5.7 (F quando a variável é falsa, V quando é verdadeira e X quando não se sabe seu valor).

Tabela 5.7: Distribuição das observações feitas nas variáveis destacadas em azul, conforme figura 5.4.

Variáveis	Observações
Hereditariedade	F
Idade	F
Escolaridade	V
Afasia	V
Perda de memória	F
Apraxia	F
Cíngulo posterior	F
Hipocampo_giroparahipocampal	F
Ventrículo	X
Córtex_parietal	F

O resultado do ensaio foi o seguinte (figura 5.5):



```

JavaBayes Console
File Options Help
Posterior distribution:
probability ( "Alzheimer" ) { //1 variable(s) and 2 values
  table
    0.10944000000000002 // p(Nao_apresenta | evide
    0.89056;           // p(Apresenta | evidence );
}
  
```

Figura 5.5: Resultado das inferências realizadas representada pela figura 5.4.

A figura 5.5 apresenta que no caso estudado na figura 5.4 a probabilidade de ser diagnosticado a DA, dado as evidências observadas é de 0,89. Isto é, somente com a observação de afasia e baixa escolaridade a probabilidade de 0,89 não se faz coerente, pois são poucas evidências apresentadas para que a rede retornasse um valor tão alto.

Podemos também fazer as observações diretamente nos nós pais (histórico, sintomas e neuroimagem como mostra a figura 5.6) verificando-se a probabilidade da existência de Alzheimer conforme os resultados abaixo relacionados:

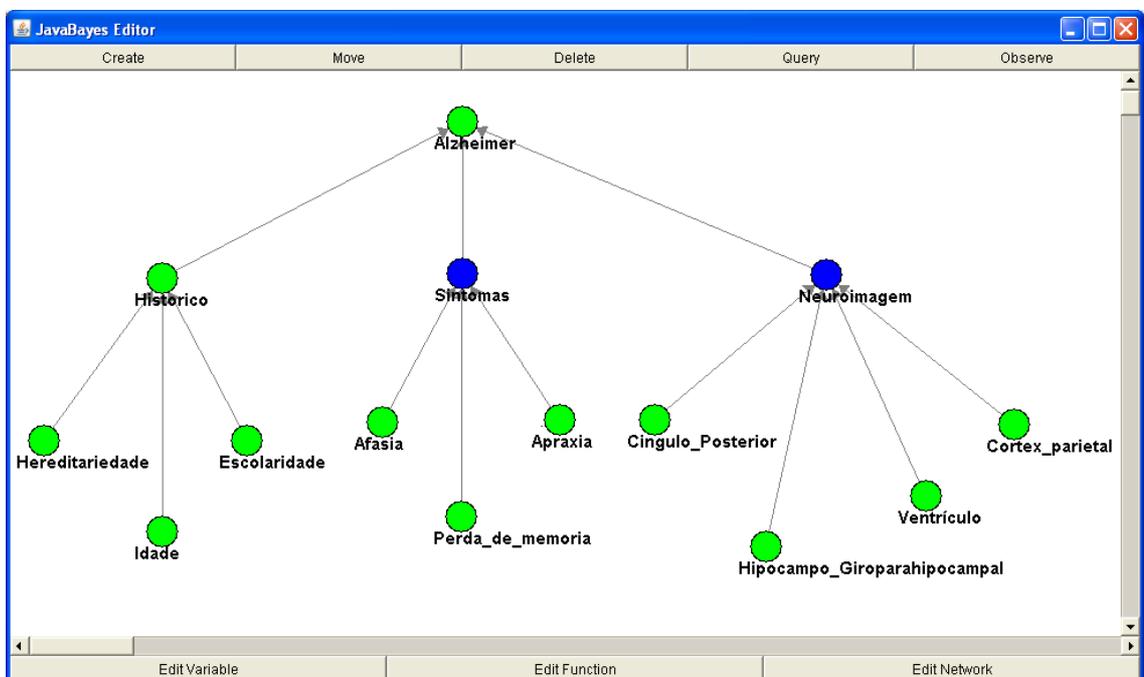


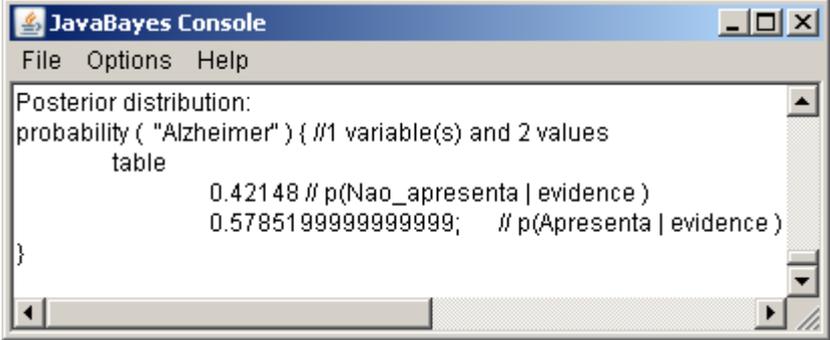
Figura 5.6: Representação gráfica da rede com observações descritas na tabela 5.11.

Neste ensaio, fizemos as seguintes suposições (observações) apresentadas na tabela 5.8 (F quando a variável é falsa, V quando é verdadeira e X quando não se sabe seu valor).

Tabela 5.8: Distribuição das observações feitas nas variáveis destacadas em azul, conforme figura 5.6.

Variáveis	Observações
Histórico	X
Sintomas	F
Neuroimagem	F

O resultado do ensaio foi o seguinte (figura 5.7):



```

JavaBayes Console
File Options Help
Posterior distribution:
probability ( "Alzheimer" ) { //1 variable(s) and 2 values
  table
    0.42148 // p(Nao_apresenta | evidence )
    0.5785199999999999; // p(Apresenta | evidence )
}
  
```

Figura 5.7: Resultado das inferências realizadas representada pela figura 5.6.

A figura 5.7 apresenta que no caso estudado na figura 5.6, a probabilidade de ser diagnosticado DA (dado as evidências observadas) é de 0,57. Prevíamos este comportamento, pois não há evidências suficientes para que a rede convergisse para um valor ainda maior de ocorrer a DA. Porém, a rede apresenta valores incoerentes, sendo possível e desejável o aprimoramento ao longo do uso do sistema.

Na próxima seção apresentaremos estudos mais detalhados para se ajustar as probabilidades, objetivando uma RB mais eficiente no auxílio ao especialista.

5.1 REDE BAYESIANA COM NOVAS DISTRIBUIÇÕES DE PROBABILIDADE

Nesta seção apresentamos novos valores de probabilidade a priori baseando-se na literatura, porém os valores de relevância dos nós permanecem os mesmos citados no início deste capítulo, de acordo com o especialista consultado.

A RB foi construída com probabilidade a priori de 0,10 (considerando que a cada cem pessoas que consultam o especialista, dez se queixam de alguns dos sintomas, apresentam algum histórico ou tenham algum dos resultados de neuroimagem positiva, característicos da DA) baseando-se nos dados obtidos dos estudos de Dráuzio Varella onde relata que 10% das pessoas com idade igual ou superior a 65 anos apresentam alguns dos sintomas característicos da DA [VARELLA 2008]. Mais especificamente esse valor a priori foi atribuído às variáveis: cíngulo posterior, hipocampo, ventrículo, córtex, apraxia, afasia, perda de memória e hereditariedade. A variável idade recebe o valor a priori de 0,086 considerando o estudo realizado pelo Instituto Brasileiro de Geografia e Estatística (IBGE) onde apresenta um índice de 8,6% da população idosa do Brasil [IBGE 2000] e para a variável escolaridade o valor a priori é de 0,091, de acordo com o índice de analfabetismo no Brasil (9,1%) demonstrado pelo IBGE em um estudo realizado em 2004 [IBGE 2004]. Assim, espera-se tornar a rede mais realista.

Consideramos como variáveis do problema (nós) os mesmos apresentados no início deste capítulo. Na tabela 5.9 demonstramos a distribuição da relevância e a probabilidade a priori (citada no parágrafo anterior) para a rede inteira.

Tabela 5.9: Distribuição de relevância para os nós Neuroimagem, Sintomas, Histórico, Cíngulo Posterior, Hipocampo, Ventrículo, Córtex, Apraxia, Afasia, Perda de memória, Hereditariedade, Idade e Escolaridade.

Variáveis condicionais	Variáveis	Probabilidade	A priori (Apresenta)
Alzheimer	Neuroimagem	0,8	0,10
Alzheimer	Sintomas	0,8	0,10
Alzheimer	Histórico	0,8	0,10
Neuroimagem	Cíngulo Posterior	0,4	0,10
Neuroimagem	Hipocampo	0,7	0,10
Neuroimagem	Ventrículo	0,8	0,10
Neuroimagem	Córtex	0,9	0,10
Sintomas	Apraxia	0,5	0,10
Sintomas	Afasia	0,5	0,10
Sintomas	Perda	0,9	0,10
Histórico	Hereditariedade	0,3	0,10
Histórico	Escolaridade	0,3	0,091
Histórico	Idade	0,9	0,086

A distribuição das probabilidades foi baseada no método Noisy-or em conformidade com a rede citada no início deste capítulo, seguindo inclusive o mesmo grau de relevância das tabelas 5.2, 5.3, 5.4 e 5.5.

A seguir demonstramos o primeiro teste feito na rede com os novos valores a priori, onde a figura 5.8 representa a rede e a tabela 5.10 mostra as observações que foram feitas nesta rede.

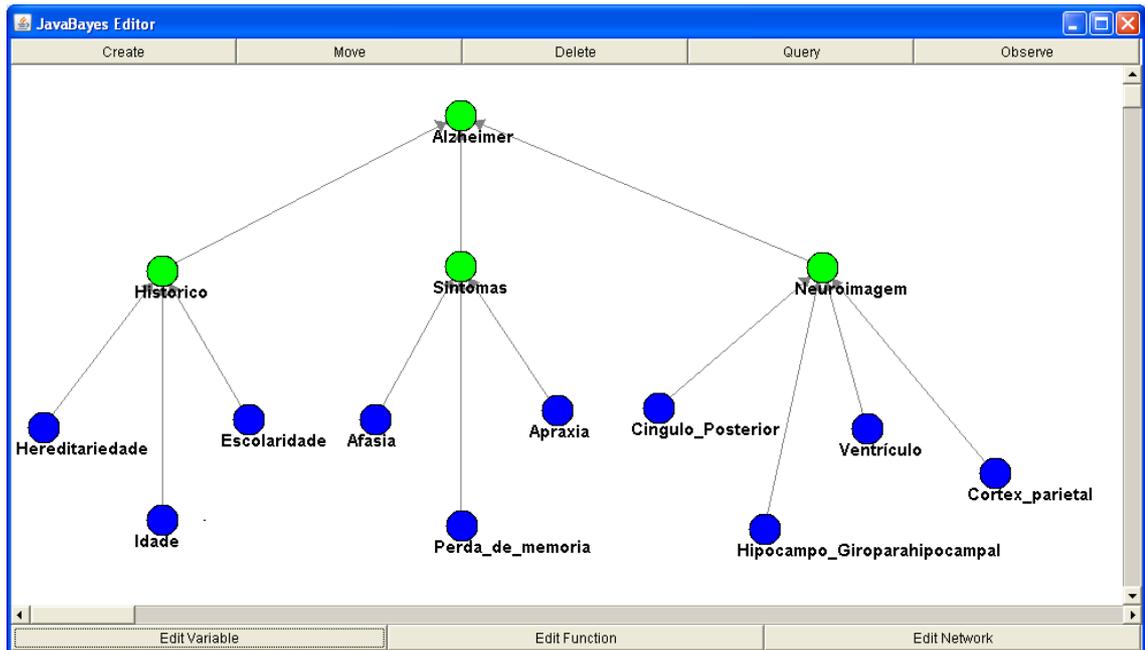


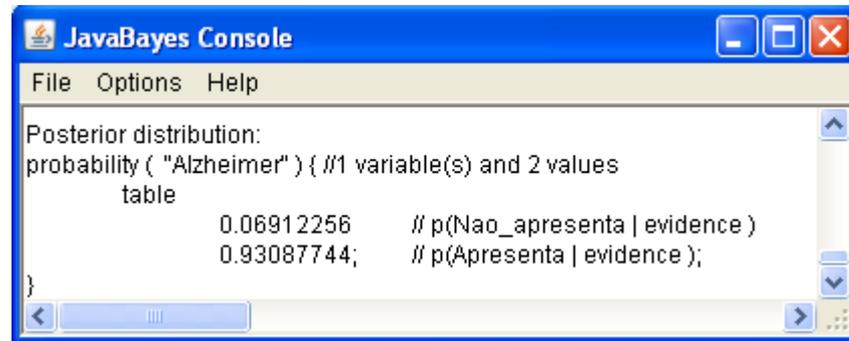
Figura 5.8: Representação gráfica da rede com observações descritas na tabela 5.10.

Neste ensaio fizemos as seguintes suposições (observações) apresentadas na tabela 5.10 (F quando a variável é falsa, V quando é verdadeira e x quando não se sabe seu valor).

Tabela 5.10: Distribuição das observações feitas nas variáveis destacadas em azul, conforme figura 5.8.

Variáveis	Observações
Hereditariedade	F
Idade	V
Escolaridade	V
Afasia	F
Perda de memória	V
Apraxia	F
Cíngulo posterior	F
Hipocampo_giroparahipocampal	V
Ventrículo	V
Córtex_parietal	V

O resultado do ensaio foi o seguinte (figura 5.9):



```

JavaBayes Console
File Options Help
Posterior distribution:
probability ( "Alzheimer" ) { //1 variable(s) and 2 values
  table
    0.06912256 // p(Nao_apresenta | evidence )
    0.93087744; // p(Apresenta | evidence );
}
  
```

Figura 5.9: Resultado das inferências realizadas na RB representada na figura 5.11.

A figura 5.9 apresenta que no caso estudado na figura 5.8 e tabela 5.10 a probabilidade de ser diagnosticado a DA dado as evidências observadas é de 0,93.

Fizemos outros testes com diferentes observações (apresentadas na figura 5.10) na RB para verificarmos se as respostas eram consistentes.

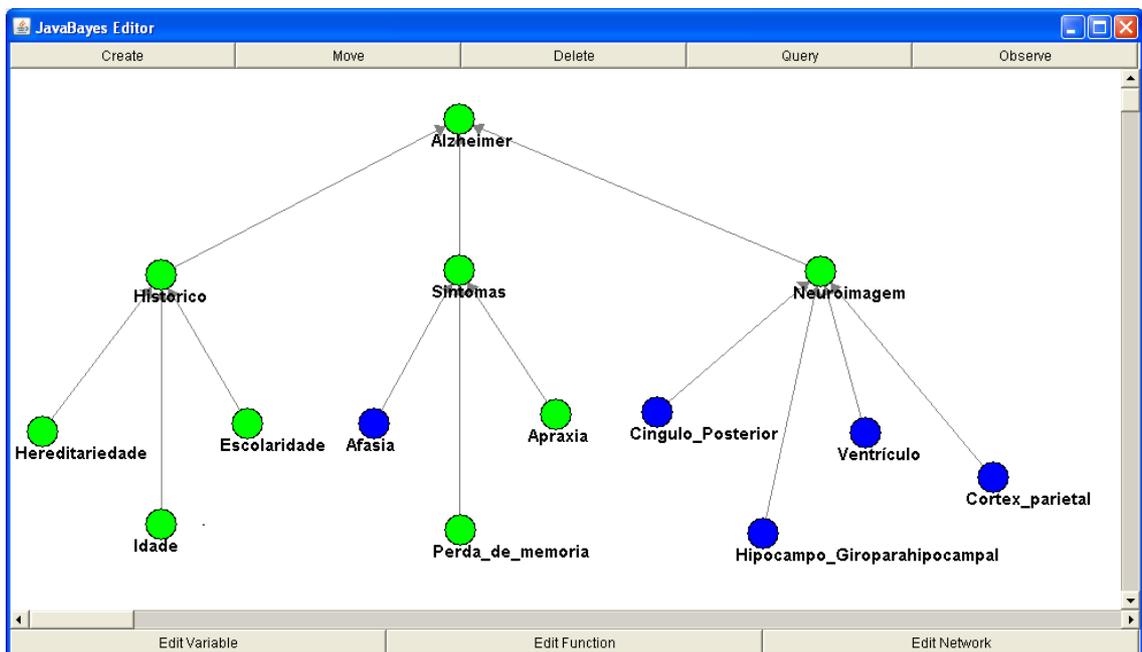


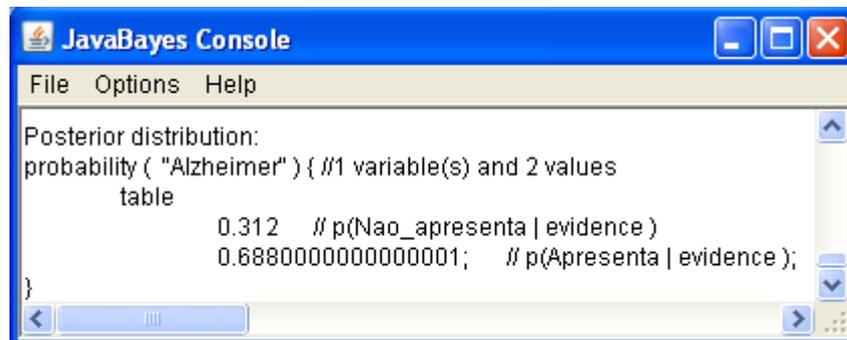
Figura 5.10: Representação gráfica da rede com observações descritas na tabela 5.18.

Neste ensaio fizemos as seguintes suposições (observações) apresentadas na tabela 5.11.

Tabela 5.11: Distribuição das observações feitas nas variáveis destacadas em azul, conforme figura 5.13.

Variáveis	Observações
Hereditariedade	F
Idade	F
Escolaridade	F
Afasia	V
Perda de memória	F
Apraxia	F
Cíngulo posterior	F
Hipocampo_giroparahipocampal	F
Ventrículo	V
Córtex_parietal	F

O resultado do ensaio foi o seguinte (figura 5.11):



```

JavaBayes Console
File Options Help
Posterior distribution:
probability ( "Alzheimer" ) { //1 variable(s) and 2 values
  table
    0.312 // p(Nao_apresenta | evidence )
    0.6880000000000001; // p(Apresenta | evidence);
}
  
```

Figura 5.11: Resultado das inferências realizadas na RB representada pela figura 5.10.

A figura 5.11 apresenta que no caso estudado na figura 5.10 e tabela 5.11 a probabilidade de ser diagnosticado a DA, dado as evidências observadas é de 0,69. Deste modo, percebemos que a rede mostrou-se coerente, visto que temos somente duas variáveis observadas como “apresenta” determinada característica (afasia e ventrículo).

Posteriormente, foram feitos outros ensaios para verificarmos a funcionalidade da rede. A figura 5.13 mostra outro teste com diferentes observações apresentadas na tabela 5.12.

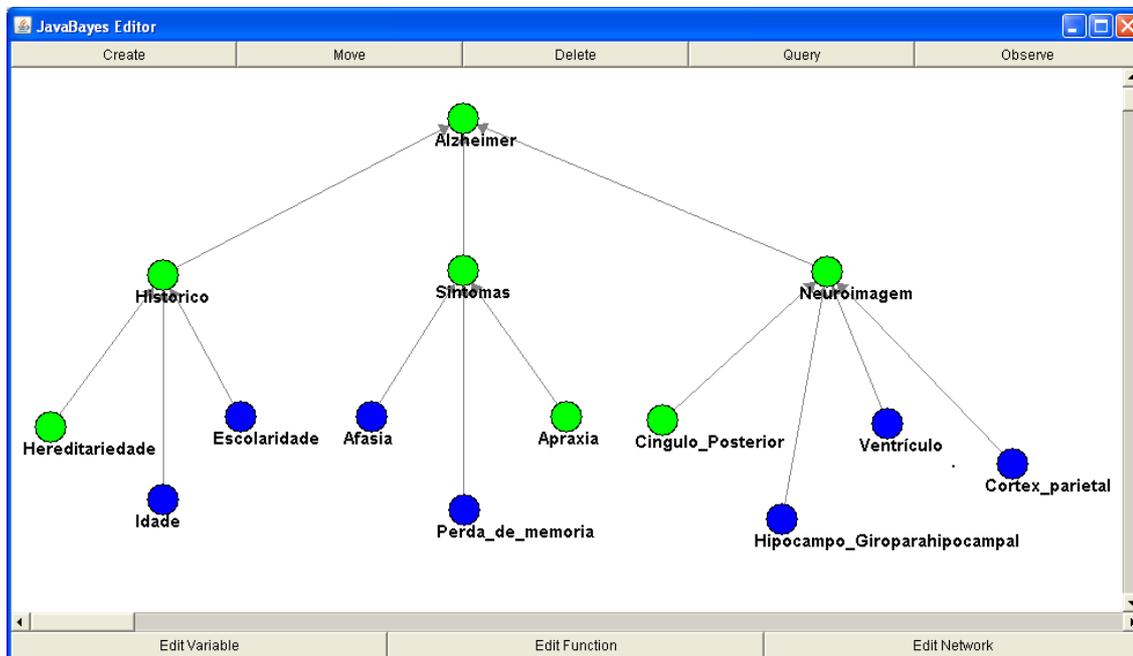
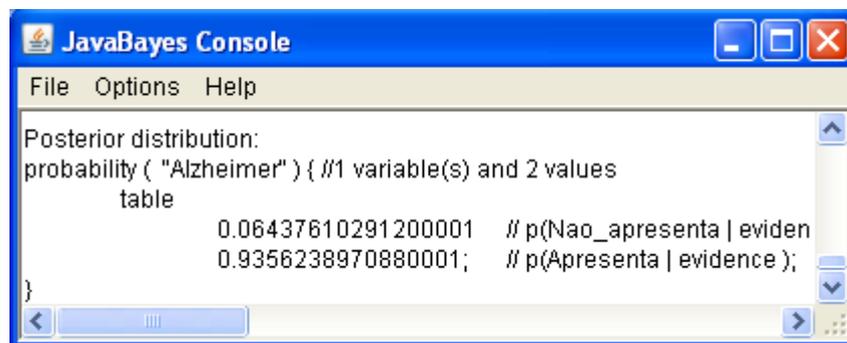


Figura 5.12: Representação gráfica da rede com observações descritas na tabela 5.12.

Tabela 5.12: Distribuição das observações feitas nas variáveis destacadas em azul, conforme figura 5.12.

Variáveis	Observações
Hereditariedade	X
Idade	V
Escolaridade	F
Afasia	V
Perda de memória	V
Apraxia	X
Cíngulo posterior	X
Hipocampo_giroparahipocampal	V
Ventrículo	V
Córtex_parietal	V

O resultado do ensaio foi o seguinte (figura 5.13):



```

JavaBayes Console
File Options Help
Posterior distribution:
probability ( "Alzheimer" ) { //1 variable(s) and 2 values
  table
    0.06437610291200001 // p(Nao_apresenta | eviden
    0.9356238970880001; // p(Apresenta | evidence);
}

```

Figura 5.13: Resultado das inferências realizadas representada pela figura 5.12.

A figura 5.13 apresenta que no caso estudado tabela 5.12 a probabilidade de ser diagnosticado a DA (dado as evidências observadas) é de 0,94, como esperado devido ao maior número de variáveis relevantes para o diagnóstico de DA que foram observadas como “apresenta”.

Com esses novos valores veremos na próxima seção a validação da rede comparando-a com dados reais dos 28 casos estudados neste trabalho.

5.2 AVALIAÇÃO DE DESEMPENHO

A base de dados composta de neuroimagem e entrevista estruturada (CAMDEX) [HOLLAND 1998] utilizada neste estudo foi cedida pelo Instituto de Psiquiatria da Faculdade de Medicina da USP conforme descrito na seção 4.1.

A base de dados contém imagens e a entrevista estruturada (CAMDEX) válida para auxiliar o diagnóstico diferencial de demência e depressão em pessoas idosas, de 28 pessoas, sendo 14 considerados controles e 14 diagnosticados com a DA (paciente).

Foram feitas comparações entre os dois grupos citados no parágrafo anterior (controle / paciente) para identificar as reais diferenças entre eles, para que pudéssemos colocar em nossa rede a informação mais sensata e organizada de acordo com casos reais. Utilizamos os dados: compree (compreensão – indicada na rede pelo nó afasia), recente (memória recente – indicada na rede pelo nó perda de memória), remota (memória remota – indicada na rede pelo nó perda de memória), praxia (antônimo de apraxia – indicada na rede pelo nó apraxia), escolar (grau de instrução escolar – indicada na rede pelo nó escolaridade) e idade (indicada como idade).

Como podemos constatar na figura 5.14, com os dados do CAMDEX não conseguimos demonstrar as reais diferenças entre os grupos (controle versus paciente), em função da redundância entre os dados estudados. Como exemplo, comparando os valores: controle mínimo e paciente mínimo, podemos constatar que não há uma diferença plausível para utilizarmos em nossa rede, principalmente nas variáveis: escolar e idade. O mesmo acontece quando comparamos os valores: controle máximo e paciente máximo, não conseguimos identificar uma diferença considerável entre os grupos, ou seja, se verificarmos o exemplo da variável idade temos os valores mínimo de cada grupo igual a 62 e o máximo varia de 79 a 82, esses valores tiveram uma diferença tão sutil, que constatamos que essa base de dados especificamente não poderíamos utilizar em nosso trabalho.

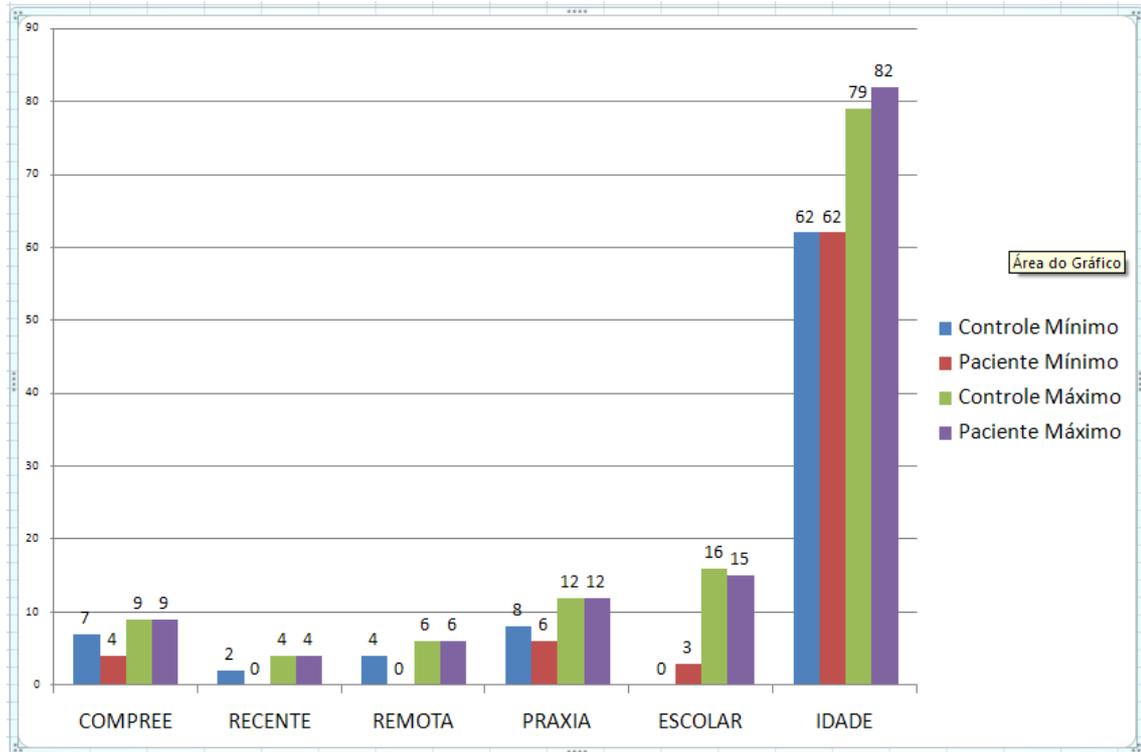


Figura 5.14: Comparação dos resultados entre as variáveis “compree”, “recente”, “remota”, “praxia”, “escolar” e “idade”.

Devido à dificuldade em apresentar os fatos com o gráfico acima (figura 5.14), em nossa rede utilizamos as variáveis que compõem o nó histórico: idade, escolaridade e hereditariedade e o nó sintomas: afasia, perda de memória e apraxia respeitando todo estudo feito com base em referências bibliográficas e opinião do especialista consultado para a realização deste trabalho. As demais variáveis para o nó neuroimagem foram definidas no capítulo 4.

Na próxima seção apresentaremos os resultados da rede com base nos valores de relevância de cada nó (tabela 5.12) e os dados levantados em entrevistas com um especialista e a partir da literatura como mostra a seção 5.1.

5.2.1 NOVOS TESTES E COMPARAÇÕES COM DIAGNÓSTICOS DADOS POR ESPECIALISTA

Por sigilo médico, nos reportaremos aos pacientes e controles como indivíduos i (indivíduo 1, indivíduo 2 e assim por diante). Na tabela 5.13 iremos utilizar as letras F para quando a variável é falsa (não apresenta a característica), T quando é verdadeira (apresenta a característica) e x quando não se sabe seu valor.

Para preenchimento da tabela 5.13 consideramos a base de dados que contém os 28 indivíduos conforme visto na seção 5.2, considerando também os dados da neuroimagem descritos na seção 4.1, lembrando que são dados das mesmas pessoas. Cruzando a informação da entrevista estruturada e neuroimagens com as variáveis que compõem nossa rede, podemos preencher a tabela 5.13 e assim fazer as avaliações da rede para cada indivíduo.

Tabela 5.13: Distribuição das probabilidades para $P(\text{neuroimagem} \mid \text{cíngulo_posterior}, \text{hipocampo}, \text{ventrículo e córtex})$.

Indivíduos	Hereditariedade	idade	escolaridade	afasia	perda	apraxia	cíngulo	hipocampo	ventrículo	córtex
Indivíduo 1	X	V	V	F	F	F	F	F	F	F
Indivíduo 2	X	V	V	F	F	F	F	F	F	F
Indivíduo 3	X	V	F	F	F	F	F	F	F	F
Indivíduo 4	X	V	F	F	F	F	F	F	F	F
Indivíduo 5	X	V	F	F	F	F	F	F	F	F
Indivíduo 6	X	V	F	F	F	F	F	F	F	F
Indivíduo 7	X	V	F	F	F	F	F	F	F	F
Indivíduo 8	X	V	F	F	F	F	F	F	F	F
Indivíduo 9	X	V	F	F	F	F	F	F	F	F
Indivíduo 10	X	V	F	F	F	F	F	F	F	F
Indivíduo 11	X	V	F	F	F	F	F	F	F	F
Indivíduo 12	X	V	F	F	F	F	V	F	F	F
Indivíduo 13	X	V	F	F	F	F	V	F	V	F
Indivíduo 14	X	V	F	F	F	F	V	V	F	F
Indivíduo 15	X	V	F	F	F	V	V	F	F	V
Indivíduo 16	X	V	F	V	V	V	V	V	F	V
Indivíduo 17	X	V	F	F	V	V	V	V	F	V
Indivíduo 18	X	V	F	F	V	V	F	V	V	V
Indivíduo 19	X	V	F	F	V	F	V	V	V	V
Indivíduo 20	X	V	F	F	V	F	V	V	V	V
Indivíduo 21	X	V	F	F	V	F	V	V	V	V
Indivíduo 22	X	V	F	F	V	F	V	V	V	V
Indivíduo 23	X	V	F	F	V	F	V	V	V	V
Indivíduo 24	X	V	F	F	F	F	V	F	V	V
Indivíduo 25	X	V	F	F	V	F	V	V	V	V
Indivíduo 26	X	V	F	F	F	F	V	F	V	V
Indivíduo 27	X	V	F	F	V	F	V	V	V	V
Indivíduo 28	X	V	F	F	V	F	V	V	V	V

Legenda: X = não identificado / V = apresenta / F = não apresenta / cor amarela = diagnosticado pelo especialista como pessoas que não apresentam a DA / cor rosa = diagnosticado pelo especialista como pessoas que apresentam DA.

Antes de apresentarmos os resultados, consideramos como limiar para decisão do diagnóstico de DA. Se o resultado for menor ou igual a 65%, consideramos como “não apresenta a DA” e se o resultado for maior que 65%, consideramos como “apresenta DA”.

Na tabela 5.14 apresentamos os resultados da rede, comparando-a com o diagnóstico do médico para o referido grupo estudado.

Tabela 5.14: Distribuição das probabilidades para P(Alzheimer | Histórico, Sintomas, Neuroimagem).

Indivíduos	Diagnóstico médico	Diagnóstico da RB	Diagnóstico pela rede
Indivíduo 1	Não apresenta DA	47%	Não apresenta DA
Indivíduo 2	Não apresenta DA	47%	Não apresenta DA
Indivíduo 3	Não apresenta DA	26%	Não apresenta DA
Indivíduo 4	Não apresenta DA	26%	Não apresenta DA
Indivíduo 5	Não apresenta DA	26%	Não apresenta DA
Indivíduo 6	Não apresenta DA	26%	Não apresenta DA
Indivíduo 7	Não apresenta DA	26%	Não apresenta DA
Indivíduo 8	Não apresenta DA	26%	Não apresenta DA
Indivíduo 9	Não apresenta DA	26%	Não apresenta DA
Indivíduo 10	Não apresenta DA	26%	Não apresenta DA
Indivíduo 11	Não apresenta DA	26%	Não apresenta DA
Indivíduo 12	Não apresenta DA	49%	Não apresenta DA
Indivíduo 13	Não apresenta DA	70%	Apresenta
Indivíduo 14	Não apresenta DA	69%	Apresenta
Indivíduo 15	Apresenta DA	85%	Apresenta
Indivíduo 16	Apresenta DA	94%	Apresenta
Indivíduo 17	Apresenta DA	92%	Apresenta
Indivíduo 18	Apresenta DA	93%	Apresenta
Indivíduo 19	Apresenta DA	91%	Apresenta
Indivíduo 20	Apresenta DA	91%	Apresenta
Indivíduo 21	Apresenta DA	91%	Apresenta
Indivíduo 22	Apresenta DA	91%	Apresenta
Indivíduo 23	Apresenta DA	91%	Apresenta
Indivíduo 24	Apresenta DA	80%	Apresenta
Indivíduo 25	Apresenta DA	91%	Apresenta
Indivíduo 26	Apresenta DA	80%	Apresenta
Indivíduo 27	Apresenta DA	91%	Apresenta
Indivíduo 28	Apresenta DA	91%	Apresenta

Legenda: cor amarela = diagnosticado pelo especialista como pessoas que não apresentam a DA / cor rosa = diagnosticado pelo especialista como pessoas que apresentam DA.

Destacamos dois valores (indivíduos 13 e 14) que vão de encontro com a opinião do especialista (rede diagnosticou como “apresenta DA” e o especialista como “não apresenta DA”). Portanto, podemos perceber que dos 28 dados que temos como base em nosso trabalho a rede concordou com 26 conferindo uma margem de erro de 7,14%.

Embora a rede apresente uma boa eficiência em relação aos dados observados é necessário aumentar ainda mais a base de dados dessa pesquisa para garantir a confiabilidade dos resultados.

No próximo capítulo apresentamos a conclusão deste trabalho.

6 CONCLUSÕES E TRABALHOS FUTUROS

Este trabalho teve como objetivo a construção de uma rede Bayesiana com a finalidade de auxiliar a decisão médica no diagnóstico da doença de Alzheimer.

Para o desenvolvimento de um sistema utilizando a Inteligência Artificial primeiramente tivemos que escolher um método, neste trabalho escolheu-se trabalhar com rede Bayesiana, depois a definição do problema médico a ser aplicado neste caso escolhemos a doença de Alzheimer. A escolha de utilizar redes Bayesianas foi devido à flexibilidade encontrada neste método para modificá-la ao passo que fosse necessário, isto é, podendo fazer alteração (inserção, remoção de alguma variável do problema) na medida em que encontramos novas informações sobre a doença de Alzheimer. RB trabalha com a hipótese da independência e considerando que ainda não existem certezas quando se trata de DA, assumimos esta hipótese para construção da rede. É de interesse para trabalhos futuros a utilização de lógica Fuzzy para verificar como o sistema reage à pergunta sobre a pessoa apresentar ou não a DA.

Considerando que a medicina, como também em outras diversas áreas, está sempre em evolução e apresentando novas descobertas, faz com que precisamos alterar o sistema para que acompanhe o dinamismo dos estudos publicados sobre a DA, por exemplo. O desenvolvimento da RB pode ser iniciado independente de conhecermos o valor da probabilidade de cada variável ou não, pois ao longo do estudo podemos complementá-las com as informações necessárias até atingir um consenso com a opinião do especialista consultado para realização deste trabalho.

Para a construção da RB consideramos os dados coletados na literatura sobre o problema médico e também a opinião de um especialista. Dessa forma inserimos na rede todas as variáveis mais relevantes para se chegar ao diagnóstico médico da doença em questão. Dividimos as variáveis do problema em três grupos (nós da rede): histórico, sintomas e neuroimagem e a elas adicionamos outras variáveis relacionadas a cada grupo, de acordo com a literatura médica e outros métodos utilizados para verificar as diferenças entre os grupos de pacientes e controles das neuroimagens estudadas neste trabalho. Entretanto, dados epidemiológicos sobre neuroimagem em DA não estão disponíveis, o que dificulta a escolha das probabilidades. Para resolver

esta questão, propõem-se utilizar meta-análises. Porém, não é claro transformar meta-análises em probabilidades bayesianas. Resolver estes problemas será tema de trabalhos futuros.

Começamos pela construção do grupo histórico e a ele adicionamos as variáveis: hereditariedade, idade e escolaridade; ao grupo sintomas adicionamos as variáveis: afasia, perda de memória e apraxia e por fim o grupo neuroimagem recebeu as variáveis: cíngulo posterior, hipocampo / giroparahipocampal, ventrículo e córtex parietal, lembrando que são regiões que conseguimos visualizar por meio de métodos estatísticos univariado e multivariado descritos neste trabalho. Um estudo mais detalhado sobre as regiões afetadas pela DA no cérebro é bem vindo aos trabalhos futuros com a finalidade de enriquecer a rede e torná-la o mais completo possível.

A quantificação das probabilidades (valores das relevâncias (importância que tem cada variável)) das variáveis obtidas se deu por intermédio do especialista. As probabilidades a priori foram adquiridas por meio da literatura, embora de um modo não aprofundado, somente para tornar a rede o mais próximo possível da realidade. Um estudo mais aprofundado, envolvendo a utilização exaustiva da rede em casos reais, seria uma extensão interessante deste trabalho.

Uma das dificuldades encontradas neste trabalho foi a obtenção dos valores probabilísticos para implementar a rede, portanto a pesquisa decorreu da seguinte forma: manteve-se os valores de relevância (o quão significativo é cada variável para chegar ao diagnóstico de DA) de cada variável fornecidos pelo especialista, porém para o valor da probabilidade a priori de cada variável (neuroimagem, sintomas e todos seus filhos, incluindo a variável hereditariedade (filho da variável histórico)) foi baseado no estudo de Dráuzio Varella [Varella 2008], o qual cita que 10% das pessoas que tem idade igual ou superior a 65 anos e 25% das pessoas com idade superior a 85 anos podem apresentar alguns sintomas de DA.

Como consideramos neste trabalho que as pessoas acima de 65 anos estão sujeitas à doença, adotamos o valor a priori de 0,10 (considerando a informação de 10% para esta faixa etária) para a construção da rede nas variáveis citadas neste parágrafo, ou seja, esse valor foi aplicado nas variáveis que compõem o nó neuroimagem, sintomas e apenas uma do nó histórico que é a variável hereditariedade,

pois as outras duas (escolaridade e idade) adquirimos informações do instituto brasileiro de geografia e estatística.

Para a variável escolaridade obtivemos o valor a priori por meio de órgão específico em realização do censo do Brasil. Na pesquisa divulgada no site do IBGE (2004) [IBGE 2004], relata que o nível de analfabetismo no Brasil teve uma queda nos últimos anos, registrando um índice de 9,1%, portanto assumimos para variável em questão o valor a priori de 0,091. O mesmo foi feito para a variável idade que recebe o valor a priori de 0,086 considerando o estudo do IBGE de 2000 que apresenta um índice de 8,6% de idosos no Brasil [IBGE 2000].

A rede apresentada nesta dissertação teve um número pequeno de variáveis, mesmo assim decidimos utilizar um modelo matemático chamado Noisy-or para distribuição das probabilidades das variáveis, pois é capaz de lidar com números elevados de variáveis do sistema utilizando-se de uma técnica de aproximação para obter probabilidades condicionais sem precisar usar todas as informações. Portanto, na medida em que há um aprofundamento do estudo da doença vai se encontrando mais variáveis relevantes para chegar ao diagnóstico de DA, ou seja, por meio de estudo mais detalhado da doença em questão pode-se elevar o número de variáveis e conseqüentemente seria complicado lidar com elas de forma útil e eficaz sem esse modelo.

Na última fase deste trabalho fizemos uma avaliação na rede por meio de comparações entre diagnóstico médico e o diagnóstico obtido pela rede para os 28 casos cedidos pela USP. Os resultados apresentam coerência nos valores retornados e com isso concluímos que a rede teve o funcionamento satisfatório. Tendo em vista alguma discrepância, ainda podemos melhorá-la com valores e pesquisas mais detalhadas sobre a doença estudada neste trabalho.

Para trabalhos futuros propõe-se a realização de uma investigação mais aprofundada sobre a doença com o objetivo de conseguir valores de probabilidade mais fidedignos, baseados, por exemplo, em meta-análises. Além disso, pretendemos estudar o maior número de casos possível, aumentando gradativamente a rede com variáveis importantes para a questão levantada neste trabalho. Após essa etapa, pode-

se ainda realizar testes práticos com a rede em consultórios e ou centros médicos, cuja avaliação pode ser valiosa para determinar a viabilidade do uso desta rede na prática.

Outro ponto importante a ser tratado em trabalhos futuros é contrastar o funcionamento da rede com dados de outras doenças, como esquizofrenia.

7 REFERÊNCIAS BIBLIOGRÁFICAS

[ABDALA 2006] Abdala, D. Daniel; Jr. da Luz, Antônio; Harley M. Wagner; Wangenheim, V. Aldo. **Um Modelo Para o Ajuste Espacial Automático do Atlas Estereotáxico de Talairach**. VI Workshop de Informática Médica – WIM 2006.

[AGUIAR 2007] Aguiar, Nelson A. Oliveira. **Comparação entre as análises estatísticas univariada e mulvariada no diagnóstico da desordem cerebral de Alzheimer**. Centro Universitário da FEI. Dissertação de mestrado. São Bernardo do Campo, 2007.

[AZEVEDO 2005] Azevedo, Dionísio, et al; **Espectroscopia de prótons na doença de Alzheimer e no comprometimento cognitivo sem demência: estudo de uma amostra comunitária**. Arq. Neuro-Psiquiatria v.63 n.4; dezembro,2005.

[CARVALHO 2004] Carvalho, H. S; Castro, L. S. S; Barbosa, Talles M. G. A.; Sene, Iwens G. Jr; Branisso, Henrique J. P.; Figueiredo, Erika C. **Sistema Pessoal móvel de monitoração da saúde: algoritmo para captura inteligente de sintomas (Personal Health Monitoring: na intelligent algorithm for symptoms capture)**. In SBIS 2004, 2004, Ribeirão Preto – SP. Anais CBIS 2004, 2004.

[HAIPENG 2002] Haipeng Guo; Hsu William. A Survey of Algorithms for Real-Time Bayesian Network Inference. In In The Joint AAAI – 02 / KDD – 02 / UAI – 02 Workshop on Real-Time Decision Support and Diagnosis Systems, 2002.

[BELLHOUSE 2004] BELLHOUSE, D. R. **The Reverend Thomas Bayes, FRS:A Biography to Celebrate the Tercentenary of His Birth**. Institute of Mathematical Statistics, Vol. 19, No. 1, 2004.

[HOLLAND 1998] Holland AJ; Hon J; Huppert F. A.; Stevens F.; Watson P. **Population-based Study of the Prevalence and Presentation of Dementia in Adults With Down's Syndrome**. The British Journal of Psychiatry 172: 493-498 (1998). The Royal College of Psychiatrists, Department of Psychiatry, University of Cambridge.

[CAOVILLA 2002] Caovilla, V. P.; Canineu, P.R. **Você não está sozinho**. SP: Abrax, 2002.

[CHARCHAT 2001] Charchat, Helenice; Nitrini, Ricardo; Caramelli, Paulo; Sameshima, Koichi. **Investigação de marcadores Clínicos dos Estágios Iniciais da Doença de Alzheimer com testes neuro-psicológicos computadorizados**. Psicologia. Reflexão e Crítica, Porto Alegre, RS, Brasil, V. 14, n. 2, p. 305-316, 2001.

[CHARNIAK 1991] E. Charniak K., **Bayesian networks without tears**. AI magazine, pp.50-63, 1991.

[COWELL 1999] COWELL, R.G. **Probabilistic Networks and Expert Systems**. New York: Springer, p.5-23, 1999.

[COZMAN 2001] Cozman, F. G. The JavaBayes System. The Isba Bulletin, Estados Unidos, v. 7, n. 4, p. 16-21, 2001.

[DAMASCENO 1999] Damasceno, B.P. **Envelhecimento Cerebral: o problema dos limites entre o normal e o patológico**. Arquivos de Neuro-Psiquiatria, v. 57, p. 78-83, 1999

[DAVATZIKOS 2006] Davatzikos, Christos; Fan, Yong; Wu, Xiao Ying; Shen, Dinggang; Resnick, Susan M. **Detection of prodromal Alzheimer's disease via pattern classification of Magnetic Resonance Imaging**. Epub 2006, Dec 14. Section of Biomedical Image Analysis, Department of Radiology, University of Pennsylvania, Philadelphia, PA 19104, USA.

[FLORES 2002] Flores, Cecília Dias; Perotto Filippo; Vicari, Rosa Maria. **Sistemas Baseados em Conhecimento para a Área da Saúde**. In: I Simpósio de Informática da Região Centro do Rio Grande do Sul – SIRC/RS 2002, 2002, Santa Maria. Simpósio de Informática da Região Centro do RS, 2002.

[FRISTON 1995] FRISTON, K.; HOLMES, A. P.; WORSLEY, K.; POLINE, J. B.; FRITH, C. D.; FRACKOWIAK, R. S. J. **Statistical Parametric Maps in Functional Imaging: A General Linear Approach**. Hum Brain Mapp. V.2, p.189 – 210, 1995.

[GOLDMAN 1992] Goldman–Rakic, P.S (1992). **Working Memory and the Mind**. Scientific American 3, 110 – 117. 1992.

[GREENFIELD 2000] Greenfield, S.A. **Cérebro Humano**. 1ª edição: Rocco, 2000.

[IBGE 2000] **Dia do Idoso**. Disponível em: http://www.ibge.gov.br/7a12/voce_sabia/datas/data.php?id_data=47. Acessado em: 13/10/2008.

[IBGE 2004] **Síntese de Indicadores Sociais**. Disponível em: http://www.ibge.gov.br/home/presidencia/noticias/noticia_impressao.php?id_noticia=13. Acessado em: 13/10/2008.

[IDE e COZMAN 2001] IDE, Jaime Shinsuke; COZMAN, F. G. **Probabilist reasoning in Embarked Systems**. Encontro Nacional de Inteligência Artificial – ENIA. Fortaleza, 2001. Annals of XXI the Congress of the Brazilian Society of Computation, 2001. v. 3. : HSCC 2006, LNCS 3927, pp. 79–92, 2006.

[IDE 2001] Ide, Jaime Shinsuke; Cozman, Fabio Gagliardi. **Raciocínio Probabilístico em Sistemas Embarcados**. In: **Terceiro Encontro Nacional de Inteligência Artificial**. 2001, Fortaleza.

[KANDEL 2000] Kandel, Eric R.; Schwartz, James H.; Jessell, Thomas M. **Principles of neural science**. Fourth Edition. McGraw-Hill. 2000.

[KLARMANN 1996] Klarmann, Patricia Cáceres. **Planejamento bayesiano de ensaios clínicos sequenciais**. Dissertação de Mestrado. Universidade de São Paulo, USP, Brasil (1996).

[LEÃO 1990] leão, B. Rocha, A.F, 1990. **Proposed Methodology for Knowledge Acquisistion: a study on congenital heart disease diagnosis**. Methods Inf Med. 1990; 29(1): 30-40.

[MARQUES 2000] Marques, Roberto Ligeiro; Dutra, Inês. **Redes Bayesianas: o que são, para que servem, algoritmos e exemplos de aplicações**. Notas de aula, UFRJ, 2000.

[MAURES 1997] Maures, Konrad; Volk, Stephan; Gerbaldo, Hector. **August D and Alzheimer's disease**. Neuropsychologia. V. 38, l. 2, Feb. 2000, P. 213-224.

[MCCRONE 2002] Mccrone, John. **Como o cérebro funciona**. 1ª edição: Publifolha, 2002.

[MILHO 2000] Milho, Isabel; Fred, Ana; Albano, Jorge; Baptista, Nuno; Sena, Paulo. **An Auxiliary System for Medical Diagnosis Base Don Bayesian Belief Networks**. In: Proceedings of the 11th Portuguese Conference on Pattern Recognition – RECPAD 2000, Porto, Portugal, May, 2000.

[NITRINI 2005] Nitrini, Ricardo; Caramelli, Paulo; Bottino, Cássio Machado de Campos; Damasceno, Benito Pereira; Brucki, Sonia Maria Dozzi; Anghinah, Renato. **Diagnóstico de doença de Alzheimer no Brasil – Critérios diagnósticos e exames complementares**. Recomendações do Depto. Científico de Neurologia Cognitiva e do Envelhecimento da Academia Brasileira de Neurologia. Arquivos de Neuro-Psiquiatria. São Paulo, v. 63, 1. 3-A, P. 713-719, 2005.

[OLIVEIRA 2005] Oliveira, M. F.; Ribeiro, M.; Borges, R.; Luginger, S. **Doença de Alzheimer: Perfil neuropsicológico e tratamento 2005**. Universidade Lusíada do Porto, Porto, Abril de 2005. Psicologia.com.pt - O Portal dos Psicólogos. Disponível em: http://www.psicologia.com.pt/artigos/ver_artigo.php?codigo=A0121&area=d7&subarea=. Acesso em 11/08/2008.

[PASCALE 2002] Pascale, Maria Aparecida. **Ergonomia e Alzheimer: a contribuição dos fatores ambientais como recurso terapêutico nos cuidados de idosos portadores da demência do tipo Alzheimer**. Tese de Doutorado, UFSC, 2002.

[PEARL 1988] Pearl, Judea; **Probabilistic Reasoning in Intelligent Systems**. Morgan Kaufmann; San Francisco, Califórnia, revised second printing, 1988 [144 – 150; 184 – 189].

[PEREIRA 2001] Pereira, Basílio de Bragança. **Estatística em psiquiatria**. Revista Brasileira de Psiquiatria, São Paulo, v. 23, n. 3, p. 168-170, 2001.

[PEROTTO 2001] Perotto, Filipo Studzinski. **Modelagem do conhecimento, sistemas especialistas e o projeto SEAMED**. Revista Eletrônica de Iniciação Científica, v. I, 2001.

[PESSETE 2002] Pessete, Renato S.; Vieira, Kleber Magno Maciel. **Redes Bayesianas no Diagnóstico Médico**. (UFSC), Santa Catarina, 2002.

[PSYCHIATRIC 1987] American Psychiatric Association. **Diagnostic and statistical manual of mental disorders**. 3rd E. Washington, D.C.: American Psychiatric Association, 1987.

[PSYCHIATRIC 1994] American Psychiatric Association. **Diagnostic and statistical manual of mental disorders**. 4rd E. Washington, D.C.: American Psychiatric Association, 1994.

[RAMOS 2001] Ramos, Luiz Roberto; Macedo, Maria Betriz. **Como diagnosticar e tratar distúrbios da memória e demência**. Revista Ciber Saúde. Julho, 2001.

[RAZZOUK 1995] Razzouk, Denise. **Construção de uma base de conhecimento de um sistema de apoio à decisão no diagnóstico dos transtornos do espectro da esquizofrenia**. Tese de doutorado, Universidade Federal de São Paulo, UNIFESP, Brasil, 2001.

[RAZZOUK 2003] Razzouk, Denise; Mari, Jair de Jesus; Shirakawa, Itiro; Wainer, Jacques; Sigulen, Daniel. **Knowledge acquisition in schizophrenia: clinical reasoning patterns among three experts**. Schizophr Res. 2003; 63(3): 295-6.

[RAZZOUK 2006] Razzouk, Denise; Mari, Jair de Jesus; Shirakawa, Itiro; Wainer, Jacques; Sigulen, Daniel. **Decision Support System for the Diagnosis of Schizophrenia Disorders: a helpful tool to clinical making-decisions in the differential diagnostic of psychoses**. Brazilian Journal of Medical and Biological Research, São Paulo, v. 39, p. 119-128, 2006.

[RAZZOUK 2007] Razzouk, Denise; Mari, Jair de Jesus; Shirakawa, Itiro; Wainer, Jacques; Sigulen, Daniel. **How Experts Recognize Schizophrenia: the role of the disorganization symptom**. Revista Brasileira de Psiquiatria (São Paulo), São Paulo, v. 28, n. 1, p. 5-9, 2006.

[RAZZOUK 2000] Razzouk, Denise; Shirakawa, Itiro; Mari, Jair de Jesus. **Sistema inteligentes no diagnóstico da esquizofrenia. Revista Brasileira Psiquiatria.** Revista Brasileira de Psiquiatria, 2000; 22(Suppl): 35-7.

[RUSSELL E NORVIG 1995] RUSSELL, S. J.; NORVIG, P. **Artificial Intelligence: a Modern Approach.** Upper Saddle River, Prentice Hall, New Jersey, 1995.

[SAHEKI 2003] Saheki, André Hideaki; Sharovsky, Rodolfo; Cozman Fabio Gagliardi; Coupé, Veerle M. H. **Construção de uma RB aplicada ao diagnóstico de doenças cardíacas.** In: Encontro Nacional de Inteligência Artificial (ENIA), 2003.

[SAMPAIO 2007] Sampaio, Rudini Menezes; VALENTIM, F.L. **Inference Algorithms for Systems of Medical Diagnosis Aid based.** INFOCOMP (UFLA), v. 7, p. 1-8, 2008.

[TAILARACH 1988] J. Talairach and P. Tournoux, "Co-planar Stereotaxic Atlas of the Human Brain: 3-Dimensional Proportional System - an Approach to Cerebral Imaging", Thieme Medical Publishers, New York, 1988

[THOMAZ 2004] THOMAZ, C E. **Maximum entropy covariance estimate for statistical pattern recognition.** 2004. 152 f. Tese de Doutorado – *Department of Computing Imperial College London, University of London, London, 2004.*

[TIBIRIÇÁ 2006] Tibiricá, Carlos Augusto Gonçalves ; NASSAR, S. M. ; PIRES, Maria Marlene de Souza . **Sistema Bayesiano de Avaliação Nutricional via Web.** In: II Congresso Sul Catarinense de Computação, 2006, Criciúma. II Congresso Sul Catarinense de Computação, 2006.

[VENTURA 2007] Ventura, M. M.; Bottino, C. M. C. **Estudo de confiabilidade da versão em português de uma entrevista estruturada para o diagnóstico de demência.** Ver. Ass. Med. Brasileira, 2001; 47: 110-116.

[WHO 1993] World Health Organization (WHO). **The ICD-10 classification of mental and behavioral disorders. Diagnostic criteria for research.** Geneva: World Health Organization, 1993.

[ZHANG 1996] Zhang, Nevin Lianwen; Poole, David; **Exploiting Causal Independence in Bayesian Network Inference.** Journal of Artificial Intelligence Research 5; december,1996.

**8 APÊNDICE 1 – PROGRAMA JAVABAYES DA REDE
BAYESIANA PARA DOENÇA DE ALZHEIMER**

Este código fonte é onde são lançados todos os valores de probabilidade tanto a priori como condicional e é parte integrante do programa JavaBayes [JAVABAYES 2007].

```
// Bayesian network
network "InternalNetwork" { //14 variables and 14 probability distributions
}
variable "Alzheimer" { //2 values
  type discrete[2] { "Nao_apresenta" "Apresenta" };
  property "position = (390, 44)" ;
}
variable "Historico" { //2 values
  type discrete[2] { "Nao_apresenta" "Apresenta" };
  property "position = (131, 180)" ;
}
variable "Sintomas" { //2 values
  type discrete[2] { "Nao_apresenta" "Apresenta" };
  property "position = (390, 176)" ;
}
variable "Neuroimagem" { //2 values
  type discrete[2] { "Nao_apresenta" "Apresenta" };
  property "position = (704, 177)" ;
}
variable "Hereditariedade" { //2 values
  type discrete[2] { "Nao_apresenta" "Apresenta" };
  property "position = (28, 317)" ;
}
variable "Idade" { //2 values
  type discrete[2] { "Media" "Avançada" };
  property "position = (152, 317)" ;
}
variable "Escolaridade" { //2 values
  type discrete[2] { "Alta" "Baixa" };
  property "position = (201, 317)" ;
}
variable "Afasia" { //2 values
  type discrete[2] { "Nao_apresenta" "Apresenta" };
  property "position = (299, 317)" ;
}
variable "Perda_de_memoria" { //2 values
  type discrete[2] { "Nao_apresenta" "Apresenta" };
```

```

    property "position = (389, 317)" ;
}
variable "Praxia" { //2 values
    type discrete[2] { "Nao_apresenta" "Apresenta" };
    property "position = (535, 317)" ;
}
variable "Cingulo_Posterior" { //2 values
    type discrete[2] { "Nao_apresenta" "Apresenta" };
    property "position = (577, 360)" ;
}
variable "Hipocampo_Giroparahipocampal" { //2 values
    type discrete[2] { "true" "false" };
    property "position = (662, 307)" ;
}
variable "Ventriculo" { //2 values
    type discrete[2] { "Nao_apresenta" "Apresenta" };
    property "position = (781, 352)" ;
}
variable "Cortex_parietal" { //2 values
    type discrete[2] { "Nao_apresenta" "Apresenta" };
    property "position = (869, 354)" ;
}
}
probability ( "Alzheimer" "Historico" "Sintomas" "Neuroimagem" ) { //4
variable(s) and 16 values
    table
        1 0.2 0.2 0.04 0.2 0.04 0.04 0.008 0 0.8 0.8 0.96 0.8 0.96 0.96
        0.992;
}
probability ( "Historico" "Hereditariedade" "Idade" "Escolaridade" ) { //4
variable(s) and 16 values
    table
        1 0.6 0.7 0.42 0.7 0.42 0.49 0.294 0 0.4 0.3 0.58 0.3 0.58 0.51
        0.706;
}
probability ( "Sintomas" "Afasia" "Perda_de_memoria" "Praxia" ) { //4 variable(s)
and 16 values
    table
        1 0.5 0.4 0.2 0.5 0.25 0.2 0.1 0 0.5 0.6 0.8 0.5 0.75 0.8 0.9;
}
probability ( "Neuroimagem" "Cingulo_Posterior"
"Hipocampo_Giroparahipocampal" "Ventriculo" "Cortex_parietal" ) { //5
variable(s) and 32 values
    table
        1 0.3 0.4 0.12 0.5 0.15 0.2 0.06 0.6 0.18 0.24 0.072 0.3 0.09

```

```

0.12 0.036 0 0.7 0.6 0.88 0.5 0.85 0.8 0.94 0.4 0.82 0.76 0.928
0.7 0.91 0.88 0.964;
}
probability ( "Hereditariedade" ) { //1 variable(s) and 2 values
    table
        0.1;          0.9 // p(Nao_apresenta | evidence )
                        // p(Apresenta | evidence );
}
probability ( "Idade" ) { //1 variable(s) and 2 values
    table
        0.909;       0.091 // p(Media | evidence )
                        // p(Avançada | evidence );
}
probability ( "Escolaridade" ) { //1 variable(s) and 2 values
    table
        0.914;       0.086 // p(Alta | evidence )
                        // p(Baixa | evidence );
}
probability ( "Afasia" ) { //1 variable(s) and 2 values
    table
        0.1;          0.9 // p(Nao_apresenta | evidence )
                        // p(Apresenta | evidence );
}
probability ( "Perda_de_memoria" ) { //1 variable(s) and 2 values
    table
        0.1;          0.9 // p(Nao_apresenta | evidence )
                        // p(Apresenta | evidence );
}
probability ( "Praxia" ) { //1 variable(s) and 2 values
    table
        0.1;          0.9 // p(Nao_apresenta | evidence )
                        // p(Apresenta | evidence );
}
probability ( "Cingulo_Posterior" ) { //1 variable(s) and 2 values
    table
        0.1;          0.9 // p(Nao_apresenta | evidence )
                        // p(Apresenta | evidence );
}
probability ( "Hipocampo_Giroparahipocampal" ) { //1 variable(s) and 2 values
    table
        0.1;          0.9 // p(true | evidence )
                        // p(false | evidence );
}

```

```
probability ( "Ventriculo" ) { //1 variable(s) and 2 values
  table
    0.1;          0.9 // p(Nao_apresenta | evidence )
                  // p(Apresenta | evidence );
}
probability ( "Cortex_parietal" ) { //1 variable(s) and 2 values
  table
    0.1;          0.9 // p(Nao_apresenta | evidence )
                  // p(Apresenta | evidence );
}
}
```

**9 APÊNDICE 2 - TRABALHO SOBRE VISÃO
COMPUTACIONAL RELACIONADO À DOENÇA DE
ALZHEIMER REALIZADO DURANTE OS CRÉDITOS
DESTE MESTRADO**

Primeiro experimento utilizando neuroimagem

Para darmos os primeiros passos à análise das imagens e constatação de alterações em algumas regiões citadas anteriormente, realizamos um teste utilizando algumas imagens (figuras A2.1 e A2.2) descritas anteriormente, onde pudemos constatar o alargamento do ventrículo em pessoas com DA. Para isso, foi implementado o algoritmo “Boundary-Fill” que preenche com pixel de cor vermelha (para maior destaque) a área determinada através de um código desenvolvido em Java.

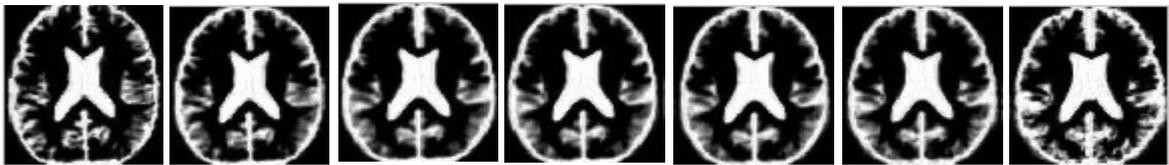


Figura A2.1: Imagens iniciais usadas como base p/ preenchimento e contagem de pixels vermelhos.

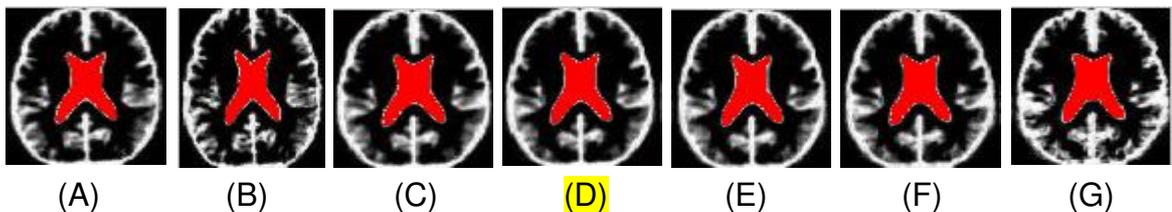


Figura A2.2: Imagens com a área ventricular preenchida com pixels vermelhos.

O algoritmo proposto neste experimento, além de preencher a área com pixel vermelho, também faz a contagem desses pixels para conseguirmos obter informações importantes quanto ao aumento de volume da região estudada.

Considerando a imagem (D) como o limiar de controle, ou seja, como sendo um parâmetro para guiar-nos nos resultados, temos que, as imagens ((A), (B) e (C)) caracterizam ausência de DA e as imagens ((E), (F) e (G)) presença de DA, conforme mostra a tabela A2.1:

Tabela A2.1: Tabela com resultado da figura A2.2.

Imagem	Resultado obtido
A	558
B	578
C	602
D	602
E	614
F	618
G	646

Detectando a diferença na região do ventrículo por meio desse simples algoritmo partimos para as análises mais detalhadas com software especializado em análise univariada descritos a seguir, para certificarmos das diferenças entre

diversas regiões do cérebro, sempre comparando pacientes com controles.