



MESTRADO PROFISSIONAL EM COMPUTAÇÃO APLICADA – MPCOMP

Universidade Estadual do Ceará - UECE
Centro de Ciências Tecnológicas - CCT

Instituto Federal de Educação, Ciência e Tecnologia do
Ceará - IFCE
Pró-Reitoria de Pós-Graduação – ProPG

GERMANNO GURGEL DO AMARAL TELES

UM MECANISMO DE APOIO À TOMADA DE DECISÃO
EM AGRAVO DE DENGUE BASEADO EM DADOS PROBABILÍSTICOS

Fortaleza
2013

GERMANNO GURGEL DO AMARAL TELES

***UM MECANISMO DE APOIO À TOMADA DE DECISÃO
EM AGRAVO DE DENGUE BASEADO EM DADOS PROBABILÍSTICOS***

Dissertação apresentada ao Programa de Mestrado Profissional em Computação Aplicada da Universidade Estadual do Ceará e do Instituto Federal de Educação, Ciência e Tecnologia do Ceará, como requisito parcial para obtenção do Grau de Mestre em Computação Aplicada.

Área de concentração: *Redes de Computadores*

Orientador: Profº Dr. Mauro Oliveira

Co - Orientador: Profº Dr. Ronaldo Ramos

Fortaleza

2013

Dados Internacionais de Catalogação na Publicação
Universidade Estadual do Ceará
Biblioteca Central Prof. Antônio Martins Filho
Bibliotecário Responsável – Francisco Welton Silva Rios – CRB-3/919

| | |
|-------|--|
| T269u | <p>Teles, Germanno Gurgel do Amaral</p> <p>Um mecanismo de apoio à decisão em agravo de dengue baseados em dados probabilísticos / Germanno Gurgel do Amaral Teles . – 2013. CD-ROM. 101 f. : il. (algumas color.) ; 4 ½ cm.</p> <p>“CD-ROM contendo o arquivo no formato PDF do trabalho acadêmico, acondicionado em caixa de DVD Slim (19 x 14 cm x 7 mm)”.</p> <p>Dissertação (mestrado) – Universidade Estadual do Ceará, Centro de Ciências e Tecnologia, Instituto Federal de Educação, Ciência e Tecnologia do Ceará, Curso de Mestrado Profissional em Computação Aplicada, Fortaleza, 2013.</p> <p>Área de Concentração: Computação Aplicada.</p> <p>Orientação: Prof. Dr. Antônio Mauro Barbosa de Oliveira.</p> <p>1. Dengue. 2. Tomada à decisão. 3. Dados probabilísticos. I. Título.</p> <p style="text-align: right;">CDD: 004.0151</p> |
|-------|--|

Germanno Gurgel do Amaral Teles


**"UM MECANISMO DE APOIO À TOMADA DE DECISÃO EM AGRAVO DE DENGUE
BASEADO EM DADOS PROBABILÍSTICOS "**


*Dissertação apresentada ao Curso de Mestrado
Profissional em Computação Aplicada da
Universidade Estadual do Ceará, como requisito
parcial para a obtenção do grau de Mestrado
em Computação.*


Defesa em: 31/10/2013

BANCA EXAMINADORA


Marcos José Negreiros Gomes, DSc. (UECE)
Presidente


Antonio Mauro Barbosa de Oliveira, DSc. (IFCE)
Orientador


Luiz Odorico Monteiro de Andrade, DSc.(UFC)
Membro Externo


Antonio de Barros Serra, DSc. (IFCE)
Membro Externo

Dedico este trabalho a todos que contribuíram para a realização do mesmo, à minha esposa, família, amigos e orientador, pelo apoio, força, incentivo, companheirismo e amizade. Sem eles, nada disso seria possível.

AGRADECIMENTOS

À **minha esposa Whana Fechine**, por ter sido meu refúgio e fortaleza nos momentos mais difíceis e em todos os outros.

Aos meus orientadores, Profº Dr. Mauro Oliveira e Profº Dr. Ronaldo Ramos, pelo apoio, paciência, compreensão e companheirismo.

À minha irmã Mirella que passou horas me ajudando com as traduções.

Ao amigo, Gustavo Bastos, pelas horas intermináveis de conversas sobre sistemas especialistas.

À diretoria do CALOUROS 06, que acreditou no trabalho o qual estava envolvido.

Aos meus companheiros de MPCOMP, Michelle, Fred, e Rita e todos os outros, pela força e torcida.

Ao meu amigo Cleisson Santos e Fernando Siqueira, pelas palavras de apoio e incentivo.

À Profª Drª Verônica Pimentel e todo o corpo docente do MPCOMP.

À equipe do Aracati Digital, pela parceria e disponibilidade de ouvir minhas apresentações.

À equipe de apoio administrativo do MPCOMP, Tom, pela presteza e torcida com que nos atende.

Ao Banco do Nordeste pelo seu trabalho de incentivo na realização do mestrado.

A todos os Médicos que me receberam com presteza e me ajudaram na conclusão deste trabalho.

RESUMO

Incertezas envolvidas em processos de tomadas de decisão exigem, cada vez mais, o uso de mecanismos de inferências baseadas em dados probabilísticos no desenvolvimento de sistemas inteligentes. Modelos de grafos têm sido usados com formalismo unificado para muitos modelos probabilísticos.

Redes Bayesianas se adéquam à construção de sistemas que dependem do conhecimento probabilístico. A área da saúde é um exemplo da utilização de redes probabilísticas na modelagem de incertezas em diagnósticos, tais como: Alzheimer, doenças cardíacas, entre outras.

Este trabalho propõe o uso de redes Bayesianas no apoio à tomada de decisão em agravos de dengue para o projeto LARIISA, uma plataforma orientada a contexto (*context-aware concept*) destinada à governança de sistemas de saúde. O trabalho especifica e implementa o LARIISA_Bay, um componente baseado em redes Bayesianas agregado à plataforma LARIISA. Este componente auxilia uma equipe de especialistas para melhor diagnosticar casos de dengue a partir de dados coletados dos usuários do LARIISA, classificando-os em: emergencial (dengue hemorrágica), grave (existência de dengue) ou normal (não existência da enfermidade). A partir desta classificação, a equipe de especialistas poderá tomar a decisão mais conveniente como, por exemplo, a providência imediata de uma ambulância ou a visita ao usuário de uma agente de saúde ou, simplesmente, uma notificação de orientação ao usuário. Um protótipo do LARIISA_Bay é descrito com interfaces para os três principais atores do sistema: paciente, agente de saúde e especialista. Os dados coletados por estas interfaces alimentam a rede Bayesiana ao sistema. Para tanto, o metadado utilizado entre os atores e o sistema contém o SUS_ID (identificação do usuário no SUS), além de outras informações necessárias ao processo de inferência da rede (geolocalização, biométricas, etc).

A rede Bayesiana do LARIISA_Bay leva em consideração uma tabela de probabilidades construída com o auxílio de profissionais da área. Além de permitir o apoio à decisão sobre a gravidade do caso de dengue, filtrando casos prováveis de dengue para a classificação acima citada, o LARIISA_Bay utiliza mecanismos de georreferenciamento que podem ser usados como entrada em sistemas de detecção e monitoramento de epidemias em diversas regiões.

Palavras-Chave: Redes Bayesianas, modelos gráficos, incertezas, dengue.

ABSTRACT

Uncertainties involved in the decision-making processes have increasingly required the use of mechanisms of probabilistic inferences based on data in the development of intelligent systems. Graph models have emerged in a unified formalism for many probabilistic models.

Bayesian Networks are powerful and appropriate methodologies for the construction of systems that rely on probabilistic knowledge. The health system sector is an example for the use of probabilistic networks in modeling of uncertainties in diagnoses such as Alzheimer's, heart disease, among others.

This paper proposes the use of Bayesian Networks to support the decision-making on diseases of *dengue* fever for the Laboratory of Intelligent and Integrated Networks of Health System (LARIISA, in Portuguese) project, a platform for health systems governance. The paper also specifies and implements the LARIISA_Bay, a component based on Bayesian Networks added to LARIISA the platform. This component assists a team of specialists to better diagnose *dengue* cases as collects data from system users, classifying them as: emergency (dengue hemorrhagic fever), severe (presence of dengue fever) or normal (absence of the disease). From this classification, the team of experts can make the decision more accurately, for example, the immediate action of sending an ambulance or to ask a health care agent to visit the patient or, simply, giving instructions on how to deal with the case, accordingly to the classification.

The Bayesian Network of LARIISA_Bay takes into account a probability table built with the help of professionals. It also helps in the decision-making process of the gravity of *dengue* fever cases, clarifying probable cases of *dengue* fever to the classification mentioned above, the LARIISA_Bay uses georeferencing mechanisms that can be used as input to detection systems and monitoring of epidemics in several regions

A prototype described in paper presents interfaces for the three main actors in the system: user, agent and health expert. The collected data feed the Bayesian Network of LARIISA_Bay. Therefore, the metadata used between the people that use it and the system contains, besides what we call a User Identification Number in the national registry Unique Health System (SUS_ID in Portuguese) other information necessary to process inference network (geographic location, biometrics, etc.)

Key words: Bayesian Network, graphical models, uncertainties, dengue fever

LISTA DE FIGURAS

| | |
|--|----|
| Figura 1- Exemplo de Rede Bayesiana aplicada em diagnóstico médico..... | 17 |
| Figura 2 - Conexão Serial: X e Y estão d-separados se V recebeu uma evidência | 19 |
| Figura 3 - Conexão Divergente: X e Y estão d-separados se V recebeu uma evidência..... | 19 |
| Figura 4 - Conexão Convergente: X e Y estão d-separados se nem V nem seus descendentes recebeu uma evidência | 19 |
| Figura 5 - Outro exemplo de Rede Bayesiana(Markov Blanket) | 20 |
| Figura 6 - Modelo de uma Naive Bayes..... | 22 |
| Figura 7 - Modelo Global de Saúde do LARIISA..... | 28 |
| Figura 8 - Modelo Local de Saúde do LARIISA | 29 |
| Figura 9 - Arquitetura do CLARIISA | 30 |
| Figura 10 - Processamento dos Dados | 32 |
| Figura 11 - Cenário do CLARIISA..... | 33 |
| Figura 12 – Modelo preliminar do LARIISA-Bay | 48 |
| Figura 13 – Modelo de informação do LARIISA_Bay..... | 50 |
| Figura 14 - Primeiro modelo da rede criada a partir do Nética | 53 |
| Figura 15 - Segundo modelo da rede criada a partir do Nética | 54 |
| Figura 16 - Modelo final da rede criada a partir do Nética | 54 |
| Figura 17- Arquitetura Funcional do LARIISA | 57 |
| Figura 18: Arquitetura Funcional do LARIISA-BAY..... | 59 |
| Figura 19 – Alto Risco | 64 |
| Figura 20 – Baixo Risco..... | 65 |
| Figura 21 - Arquitetura do sistema proposto..... | 69 |
| Figura 22 – Tela <i>Login</i> | 70 |
| Figura 23 – Perfil Usuário | 71 |
| Figura 24 – Perfil Agente de Saúde | 71 |
| Figura 25 – Perfil Especialista | 71 |
| Figura 26 – Sala de Situação..... | 73 |
| Figura 27 - Decisão do Especialista | 73 |
| Figura 28 – Validação do Especialista..... | 74 |
| Figura 29 – Validação do Especialista recusando o resultado sugerido..... | 74 |
| Figura 30 - <i>Pass Throug</i> | 75 |

LISTA DE TABELAS

| | |
|--|----|
| Tabela 1 – Exemplo Teste/Doença | 7 |
| Tabela 2 - Distribuição de "Idade" | 18 |
| Tabela 3 - Probabilidades Condicionais de "Doença" dado "Sintomas" | 18 |
| Tabela 4 - Tabela de distribuição de probabilidades | 26 |
| Tabela 5 - Exemplos de Sistemas Especialistas Probabilísticos..... | 39 |
| Tabela 6 - Probabilidade <i>a priori</i> | 60 |
| Tabela 7 - Probabilidade Condicional 1..... | 61 |
| Tabela 8 - Probabilidade Condicional 2..... | 61 |
| Tabela 9 - Probabilidade Condicional 3..... | 61 |
| Tabela 10 - Probabilidade Condicional 4..... | 62 |
| Tabela 12 - Casos de uso para rede simplificada | 63 |
| Tabela 13- Casos de uso para a rede completa..... | 64 |

SUMÁRIO

| | |
|---|-----------|
| 1. INTRODUÇÃO | 1 |
| 1.1. Contextualização | 1 |
| 1.2. Motivação do Trabalho | 2 |
| 1.3. Objetivos..... | 3 |
| 1.3.1. Geral..... | 3 |
| 1.3.2. Específico..... | 3 |
| 1.4. Metodologia de Desenvolvimento do Trabalho..... | 3 |
| 1.5. Organização do Trabalho..... | 5 |
| 2. FUNDAMENTAÇÃO TEÓRICA | 6 |
| 2.1. Sistemas Especialistas | 6 |
| 2.2. Sistemas especialistas aplicados à medicina..... | 6 |
| 2.3. Raciocínio probabilístico em medicina | 7 |
| 2.4. Redes Bayesianas | 8 |
| 2.5. Visões sobre os Números de Probabilidade | 10 |
| 2.6. Probabilidade <i>a Priori</i> | 11 |
| 2.7. Probabilidade Condicional..... | 12 |
| 2.8. Regra do Produto..... | 12 |
| 2.9. Axiomas da Probabilidade | 12 |
| 2.10. A Regra de Bayes | 13 |
| 2.10.1. Normalização da Regra de Bayes | 15 |
| 2.11. Redes Bayesianas..... | 16 |
| 2.11.1. Inferência em Redes Bayesianas..... | 20 |
| 2.12. Classificação Bayesianas..... | 22 |
| 2.12.1. Classificador Naive Bayes | 22 |
| 2.12.2. Classificador Noisy-or | 23 |
| 3. PROJETO LARIISA | 27 |
| 3.1. Modelos de Ontologia local e global | 27 |
| 3.2. Cenário de utilização com metadado e geolocalização | 30 |
| 4. TRABALHOS RELACIONADOS | 34 |
| 4.1. Construção de uma Rede Bayesiana Aplicada ao Diagnóstico de Doenças Cardíacas... | 35 |
| 4.2. Um estudo comparativo entre Naive Bayes e Noisy-or..... | 35 |
| 4.3. Utilizando ontologia na inferência para apoio ao diagnóstico | 36 |

| | |
|--|-----------|
| 4.4. Sistemas Especialistas Probabilísticos | 37 |
| 4.5. Exemplos de Sistemas Desenvolvidos | 38 |
| 5. CRIAÇÃO DE REDES BAYESIANAS NO PROJETO LARIISA..... | 47 |
| 5.1. Tratando a incerteza de dados no LARIISA..... | 47 |
| 5.1.1. Incertezas, Ontologia e Contexto..... | 47 |
| 5.1.2. LARIISA_Bay, um componente baseado em redes Bayesianas..... | 48 |
| 5.1.3. Interfaces do LARIISA_Bay..... | 49 |
| 5.2. Criação do modelo da rede Bayesiana do LARIISA_Bay | 51 |
| 5.2.1. Sintomas patológicos observados | 51 |
| 5.2.2. Fatores de risco relacionados à doença..... | 52 |
| 5.2.3. Ferramentas de diagnóstico..... | 53 |
| 5.3. Evolução do modelos construídos | 53 |
| 6. MODELAGEM DO LARIISA E LARIISA_Bay | 56 |
| 6.1. Arquitetura Funcional do LARIISA_Bay | 58 |
| 6.2. Obtenção de probabilidades | 59 |
| 6.2.1. Probabilidades marginais | 60 |
| 6.2.2. Probabilidades Condicionais..... | 60 |
| 6.2.3. Forma utilizadas para a avaliação da rede | 62 |
| 7. AMBIENTE DE TESTE E EXPERIMENTOS | 67 |
| 7.1. Ferramenta de implementação da rede Bayesiana..... | 67 |
| 7.2. Arquitetura do protótipo..... | 68 |
| 7.3. Interfaces do protótipo | 69 |
| 7.4. Simulações com o LARIISA_Bay | 72 |
| 7.5. Interface de Governança | 72 |
| 8. CONCLUSÃO | 76 |
| 8.1. Trabalhos Futuros..... | 77 |
| 9. REFERÊNCIAS BIBLIOGRÁFICAS..... | 79 |
| APÊNDICE A | 83 |
| APÊNDICE B..... | 87 |

1. INTRODUÇÃO

1.1. Contextualização

Com o advento do Programa Saúde da Família (PSF) do Sistema Único de Saúde (SUS), os serviços oferecidos à população na área da saúde sofreram severa descentralização. O atendimento, que antes era oferecido apenas em hospitais e postos de saúde, passa a ser oferecido também na residência do paciente.

Essa mudança aumentou a complexidade da gestão da informação, pois os atendimentos agora estão distribuídos em áreas muito maiores que as dos hospitais e envolvem um conjunto maior de informação (ALCÂNTARA, 2012).

O LARIISA (OLIVEIRA, HAIRON, *et al.*, 2010) é um projeto que usa mecanismos inteligentes para tomada de decisão em governança de sistemas públicos de saúde. Ele obtém dados de usuários do sistema público de saúde em suas residências e os envia para a sua base de dados contribuindo para a construção de sistemas inteligentes capazes de apoiar a tomada de decisão em suas diversas instâncias. Portanto, o LARIISA se apresenta como uma plataforma de apoio à tomada de decisão no âmbito do SUS.

A plataforma LARIISA prevê a representação de informação contextual para ser utilizada por aplicações sensíveis ao contexto (*context-aware concept*) (DEY, SALBER e ABOWD, 2001) e representação de conhecimento mediante o uso de ontologias (VIANA, 2007). O desenvolvimento de aplicações no projeto LARIISA tem certa complexidade por envolver a representação do conhecimento (informações capturadas dos diversos atores do sistema, domínio representado, diversas bases de dados relacionadas, entre outros) e a captura de informações de contexto, via mecanismos de *hardware*, que alimentam os mecanismos de inferência existentes.

Esta complexidade aumenta à medida que a avaliação em sistemas de saúde é vulnerável a incertezas¹, seja na captura de informações subjetivas sobre o usuário, seja no próprio diagnóstico que afeta o consequente processo de tomada de decisão.

Incerteza envolvidas em processos de tomadas de decisão têm exigido, cada vez mais, o uso de mecanismos de inferências baseadas em dados probabilísticos no desenvolvimento de sistemas inteligentes para a área da saúde. Redes Bayesianas se

¹ **Incerteza** é a falta de conhecimento “a priori” do resultado de uma ação ou do efeito de uma condição

apresentam como ferramentas poderosas e adequadas em metodologias para a construção de sistemas que dependem do conhecimento probabilístico.

1.2. Motivação do Trabalho

A dengue representa uma das grandes preocupações do Ministério da Saúde, devido à quantidade de casos notificados todos os anos (MINISTÉRIO DA SAÚDE, 2012). O gerenciamento destes casos está diretamente relacionado tanto com as ações e esforços realizados em postos de saúde, visando diminuir sua incidência, como em toda a cadeia envolvida no processo clínico-epidemiológico e administrativo da dengue.

A disponibilidade de uma ferramenta para gerenciamento e tomada de decisão em casos de dengue, baseada em conhecimentos extraídos por especialistas, poderá facilitar o trabalho de dezenas de profissionais de saúde. Esta ferramenta se faz necessária, quando não indispensável, se considerarmos o atual quadro da saúde pública brasileira em que a demanda por prevenção, diagnóstico e acompanhamento de casos de dengue ultrapassa a real capacidade de atendimento pelo sistema público de saúde.

Junte-se ao problema acima (o desequilíbrio “Demanda X Capacidade de atendimento” do governo) a inexistência do uso de mecanismos computacionais eficientes, para agilizar os processos transacionais complexos no combate a dengue. Se considerarmos, neste contexto, que a atuação de profissionais de saúde é, naturalmente, permeada por certo grau de incerteza (avaliações, diagnósticos, etc.) percebe-se que mecanismos inteligentes capazes de tratar este universo probabilístico são recomendáveis.

Por esta razão, redes Bayesianas, vêm sendo cada vez mais utilizada na construção de sistemas na área da saúde. São exemplos da utilização destas redes probabilísticas, diagnósticos de doença como Alzheimer, doenças cardíacas, entre outras.

A principal motivação deste trabalho é disponibilizar um mecanismo para auxiliar especialistas no diagnóstico em situações de agravo de dengue, em especial, pessoas que apresentam sintomas e podem ser orientadas por profissionais da área de saúde em sua própria residência. Dessa forma, pode ser feita a devida seleção entre os pacientes que devem ser tratados imediatamente (dengue hemorrágica), ou encaminhados a hospitais para maiores análises e tratamento, ou que podem ser tratados em suas próprias residências por apresentarem problemas de menor gravidade.

1.3. Objetivos

1.3.1. Geral

Este trabalho objetiva a definição, especificação e a implementação de um mecanismo inteligente para tratar incertezas nas informações de contexto e na representação do conhecimento no projeto LARIISA. O resultado é o LARIISA_Bay, um componente baseado em redes Bayesianas agregado à plataforma LARIISA. Este componente permite a captura de informações de contexto e produz inferências que auxiliam especialistas a melhor diagnosticar casos de dengue a partir de dados coletados (pacientes, agentes de saúde e especialistas) no LARIISA e de uma tabela de probabilidades de especialistas.

1.3.2. Específico

São objetivos específicos deste trabalho:

- Criação de um modelo baseado em redes Bayesianas que atenda ao objetivo geral;
- Definição dos parâmetros e de uma tabela de probabilidades que ajudam a diagnosticar casos de dengue;
- Criação de uma rede Bayesiana baseada no modelo e parâmetros acima definidos;
- Especificação do metadado e das interfaces com os atores do sistema: paciente, agente de saúde e especialista;
- Implementação do protótipo contendo as três interfaces dos atores envolvidos e a rede Bayesiana.

1.4. Metodologia de Desenvolvimento do Trabalho

Este trabalho dedica-se à criação de um sistema de auxílio à tomada de decisão, utilizando redes Bayesianas destinado à área médica, (PEARL, 1988) (SPIEGELHALTER, 1987) (WETTER, 2002). Ele mostra uma aplicação prática de técnicas de inteligência artificial em uma situação cotidiana de diagnóstico de dengue. Este tem sido um grande desafio aos especialistas, principalmente devido às dificuldades de traduzir o conhecimento médico para uma forma computacionalmente tratável.

Para se atingir este objetivo, foram utilizados os seguintes procedimentos metodológicos:

- Estabeleceu-se, claramente, o objetivo do trabalho: “criar um modelo e uma ferramenta de apoio a especialistas na tomada de decisão em caso de agravo de dengue”. Para tanto, considerou-se o estado da arte do LARIISA que, até então, não tinha um mecanismo definido para tratar incertezas.
- Procurou-se na literatura as principais contribuições acadêmicas quanto ao uso de redes Bayesianas na área de saúde, em especial as contribuições relacionadas com casos de agravos de dengue;
- Especificou-se o modelo LARIISA_Bay, levando-se em consideração o “*approach*” tecnológico e funcional do LARIISA e os objetivos a serem atingidos: um modelo para tratamento de dados probabilísticos;
- Foram entrevistados vários profissionais da área de saúde com o objetivo de se definir os principais parâmetros capazes de auxiliar no diagnóstico de dengue, tanto aqueles a serem fornecidos pelo paciente, quanto os no âmbito do agente de saúde e, principalmente, os do especialista que pode ter acesso ao prontuário eletrônico do usuário a partir do SUS_IS (identificador do usuário) contido no metadado;
- Estruturou-se um metadado a partir dos parâmetros acima citados. Este metadado serve de comunicação entre usuário e sistema, via dispositivos móveis ou TV digital interativa;
- Criou-se uma primeira versão de uma rede Bayesiana à luz dos parâmetros identificados junto a especialistas, contemplando o modelo definido inicialmente, bem como as idiossincrasias do LARIISA;
- Refinou-se a rede Bayesiana criando-se outras versões até a validação da mesma com a ajuda dos especialistas;
- Definiu-se as interfaces dos três principais atores envolvidos no cenário de execução do LARIISA_Bay: interface usuário, interface agente de saúde, interface especialista.
- Definiu-se quais parâmetros são usados pela rede Bayesiana (versão validada por especialistas) criada seriam pertinentes às interfaces dos atores envolvidos;
- Implementou-se o protótipo como prova de conceito do objetivo proposto, integrando-o ao contexto do projeto LARIISA;
- Analisou-se pela possibilidade do LARIISA_Bay também servir de entrada em sistemas de detecção e monitoramento de epidemias em diversas regiões.

1.5. Organização do Trabalho

Esta dissertação está organizada em sete capítulos: O **Capítulo 2** apresenta uma **fundamentação teórica** dos principais conceitos utilizados no trabalho. No **Capítulo 3** é apresentado o **projeto LARIISA**, a plataforma sensível ao contexto e baseada em ontologias que serve de substrato à contribuição deste trabalho. No **Capítulo 4** é abordada uma composição de trabalhos relevantes à solução proposta e que servem de referencial para uma análise crítica ao mesmo. No **Capítulo 5** são descritas as fases de criação e evolução da rede Bayesiana utilizada no modelo. No **Capítulo 6** são descritas as contribuições deste trabalho, com destaque às suas características funcionais, a definição do metadado e o processo de criação do modelo proposto. No **Capítulo 7** é apresentado o protótipo LARIISA_Bay que implementa a comunicação entre o metadado e a rede Bayesiana criada. Finalmente, no **Capítulo 8**, são apresentadas as considerações finais e trabalhos futuros.

2. FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta as principais definições e conceitos relacionados a sistemas especialistas, redes Bayesianas e suas aplicações. O texto tem por objetivo mostrar a fundamentação teórica deste trabalho, mediante a explanação dos conceitos clássicos da disciplina na qual o presente trabalho se insere.

2.1. Sistemas Especialistas

Um especialista é uma pessoa que detém conhecimentos e habilidades específicas capaz de resolver problemas em sua área de conhecimento. Sistemas especialistas são ferramentas computacionais cujo objetivo é simular as decisões que seriam tomadas por especialistas (COWELL ET AL, 1999).

Um sistema especialista, em geral, utiliza uma base de conhecimento específico à área considerada e codificada de alguma forma no sistema. Além disso, o sistema deve possuir um mecanismo de dedução, o qual gera resultados utilizando-se da base de conhecimento e de parâmetros inseridos pelo usuário do sistema. Normalmente, os dois componentes - a base de conhecimento e o mecanismo de inferência – podem ser isolados, de forma que se possa desenvolver cada uma das partes isoladamente.

2.2. Sistemas especialistas aplicados à medicina

A construção de sistemas especialistas na área médica apresenta diversos desafios. Entre os quais se pode destacar a apresentação aos médicos dos métodos utilizados, de forma que estes possam entender o funcionamento dos sistemas construídos e, por conseguinte, julgar com mais lucidez o desempenho dos sistemas. E, em um estágio posterior à construção, uma das maiores dificuldades é promover o uso dos sistemas durante a prática cotidiana.

A utilização dos sistemas deve ser a mais simples possível, de modo que o atendimento médico não sofra uma interrupção inconveniente devido ao uso do sistema especialista. O entendimento e a superação destes desafios são alguns dos principais indicadores de sucesso dos sistemas especialistas.

Uma aplicação que utiliza técnicas de inteligência artificial são os desfibriladores elétricos. Estes aparelhos contêm rotinas inteligentes, de modo que podem ser utilizados por pessoas sem treinamento em paramedicina no atendimento emergencial a pessoas com paradas cardíacas súbitas.

A Rede Bayesiana, como uma técnica para sistemas especialistas, foi criada a partir do Teorema de Bayes, utilizado para cálculos probabilísticos causais. A rede é composta por várias variáveis de entrada e uma variável de saída. (NÓBREGA, 2004).

2.3. Raciocínio probabilístico em medicina

O raciocínio (diagnóstico) médico é sabidamente mais baseado em probabilidades do que em certezas. Sendo o raciocínio probabilístico tão importante no processo de diagnóstico, na definição de regras de predição clínica e mesmo na conduta em casos complexos, é espantoso que quase nada sobre o mesmo, seja ensinado no currículo médico (WASSON, SOX e NEFF, 1985).

Para entendermos melhor os aspectos quantitativos da lógica diagnóstica, ou mais especificamente o raciocínio probabilístico e o Teorema de Bayes, é necessária uma breve introdução aos conceitos de sensibilidade e especificidade a partir da definição de probabilidade condicional. Dentro deste conceito para dois eventos quaisquer A e B, sendo $P(A) > 0$, definimos a probabilidade condicional de B dado A, que será visto na seção 2.7, como sendo:

$$P(B | A) = \frac{P(A \cap B)}{P(A)} \quad (1)$$

Consideremos agora a situação abaixo, sintetizada na tabela 1, onde n_1 pessoas sabidamente portadoras de uma doença e n_2 pessoas sabidamente livres da doença foram submetidas a um determinado teste clínico:

| <i>Teste/Doença</i> | <i>Positivo</i> | <i>Negativo</i> | <i>Total</i> |
|---------------------|-----------------|-----------------|--------------|
| (D) Presente | a | b | n_1 |
| (d) Ausente | c | d | n_2 |
| Total | $a+c$ | $b+d$ | N |

Tabela 1 – Exemplo Teste/Doença

Temos nessa situação os seguintes resultados possíveis: teste positivo e doença presente: VP (verdadeiro positivo) = a casos; teste negativo e doença ausente: VN (verdadeiro negativo) = d casos; teste negativo e doença presente: FN (falso negativo) = b casos; teste positivo e doença ausente: FP (falso positivo) = c casos.

Ao perguntarmos: se a doença está presente, qual a probabilidade do teste ser positivo, ou melhor, $P(T/D)$? A resposta é dada da seguinte maneira:

$$P(T|D) = \frac{P(T \cap d)}{P(d)} = \frac{a/N}{n_1/N} = \frac{a}{n_1} = VP/(VP + FN) \quad (2)$$

A Eq.(2) é a chamada, a sensibilidade estimada do teste

Ao perguntarmos: Se a doença está ausente, qual a probabilidade do teste ser negativo, ou melhor, $P(T/d)$? A resposta no caso é dada da seguinte maneira:

$$P(T|d) = \frac{P(T \cap d)}{P(d)} = \frac{d/N}{n_2/N} = \frac{d}{n_2} = \frac{VN}{VN+FP} \quad (3)$$

A Eq.(3) é a chamada especificidade estimada do teste.

Em resumo, testes altamente sensíveis são úteis para afastar suspeita de doenças quando negativos e testes altamente específicos são úteis para confirmar doenças quando positivos. Sempre que alterarmos o critério de positividade de um teste no sentido de torná-lo mais específico, estaremos, por outro lado, diminuindo a sua sensibilidade. O oposto é igualmente verdadeiro.

O conceito de prevalência refere-se à frequência da doença na população estudada. Já o conceito de valor preditivo é mais complexo, referindo-se à probabilidade que um indivíduo tem de ter a doença após o resultado alterado de um teste. Só poderemos saber o valor preditivo de um teste se considerarmos conjuntamente informações sobre prevalência, sensibilidade e especificidade desse teste, expressas matematicamente através da fórmula de Bayes.

2.4. Redes Bayesianas

O termo “Bayesiano” é derivado do nome “*Bayes*”, em homenagem a Thomas Bayes, que foi um reverendo presbiteriano que viveu no início do século XVIII (1702 – 1761) na Inglaterra, e suas contribuições culturais e intelectuais podem ser encontradas, hoje, em milhares de artigos científicos (PENA, 2006).

Na estatística, generalizando, existem duas grandes vertentes. Uma delas é a frequentista e a outra é a bayesiana. Essa última aborda aspectos de probabilidade como grau de credibilidade. Pode-se dizer que o grande primórdio da abordagem bayesiana a problemas de inferência partiu da publicação por Richard Price, em 1763, da obra póstuma de Thomas Bayes intitulada “*An essay towards solving a problem in the doctrine of chances*” (Ensaio buscando resolver um problema na doutrina das probabilidades) onde estava presente a demonstração do famoso Teorema de Bayes (PAULINO, TURKMAN e MURTEIRA, 2003).

A ideia de se trabalhar com probabilidades como grau de credibilidade, é de suma importância para compreender a filosofia bayesiana, e remota. Acredita-se que Bernoulli, em sua obra “Ars Conjectandi” (em 1713), foi um dos primeiros autores a definir probabilidade como grau de confiança numa dada proposição, que ainda não se conhece a respeito de sua veracidade (PAULINO, TURKMAN e MURTEIRA, 2003).

O “bayesianismo” tem dois grandes alicerces epistemológicos. O primeiro é ver o universo com base em graus de credibilidade, ao invés da ideia booleana de “tudo-ou-nada”. O outro é uma abordagem matemática que quantifica a mudança das crenças quando de novos dados empíricos.

A modelagem matemática do Teorema de Bayes propõe cálculos de probabilidades, com base em conhecimentos “a priori” e condicionalidades probabilísticas, ou seja, ela liga a inferência racional (probabilidade “a posteriori”) a subjetividade (probabilidade “a priori”) e a experiência empírica (probabilidades condicionais). E como escrito de forma poética por Yudkowsky (YUDKOWSKY, 2009): “O Teorema de Bayes liga a razão humana ao universo físico”.

O conceito de redes bayesianas vem do desenvolvimento “em cadeia” do cálculo de probabilidades de Bayes. Redes bayesianas são grafos acíclicos dirigidos, ou seja, representações de relações causais entre variáveis aleatórias interdependentes em modelos probabilísticos, (LAURIA e DUCHESSI, 2007). Segundo (PENA, 2006) as RBs são diagramas que analisam problemas reais através de um mapeamento probabilístico das relações de causa e efeito entre variáveis.

As aplicações dessa vertente de Inteligência Artificial (IA) em outras áreas como a medicina (PESSETE e VIEIRA, 2002), controle ambiental (ARAUJO e NEHME, 2001), e, de um modo geral, quando se trabalha com incertezas é bastante promissor (MCFARIANE, EVANS e ACKERMAN, 2001). Entretanto, a grande dificuldade de utilização de RBs, está na estimativa dos modelos probabilísticos “a priori” e de modelar quantitativamente o conhecimento empírico em um ambiente não frequencista (ARAUJO e NEHME, 2001).

A Rede Bayesiana (RB) é uma forma de representação de um sistema baseado em conhecimento através de estruturas gráficas e é formalizada pela teoria de probabilidade, também conhecida como Redes Causais ou Redes Probabilísticas. (RUSSELL e NORVIG, 2004)

O raciocínio humano não está baseado somente em sistemas dicotômicos ou na constatação ou não de um fato, em expressões numéricas ou estatísticas, mas também no

quanto ele acredita na ocorrência de um evento, ou na influência que um acontecimento tem sobre outro.

Probabilidade Bayesiana é o método que possibilita representar numericamente o grau de certeza de um determinado evento ocorrer, diante de um cenário de incertezas. A teoria Bayesiana está fundamentada na teoria da probabilidade, porém difere da teoria de Bayes por não adotar o enfoque frequencista (HECKERMAN e M. WELLMAN, 1995).

2.5. Visões sobre os Números de Probabilidade

O estudo sobre a probabilidade nos leva necessariamente a três visões básicas: frequencista, objetivista e subjetivista (RUSSELL e NORVIG, 2004).

A visão frequencista afirma que os números de probabilidade da ocorrência de um dado evento, provêm de experimentos. Isto significa que se testarmos 100 pessoas e descobrirmos que 10 delas têm cárie a probabilidade de ocorrer uma cárie em uma pessoa é aproximadamente de 0,1. A leitura que se faz dessa assertiva é que essa fração seria observada no limite de infinitas amostras.

A visão objetivista afirma que os números de probabilidade ocorrem como aspectos reais do universo. Que seriam as tendências dos objetos a se comportarem de determinadas maneiras e não apenas como uma descrição do grau de crença de um observador. Isso significa que o fato de se lançar uma moeda e ocorrer cara com probabilidade de 0,5 é uma tendência natural da própria moeda. Ela afirma ainda, que as medições frequencistas são meras tentativas de observar essas tendências. Essa visão é compartilhada por muitos físicos quando observam os fenômenos quânticos.

A visão subjetivista descreve os números de probabilidades como uma maneira de caracterizar as crenças de um agente excluindo qualquer significado de natureza física externa. Isso permite que o observador forme esses números quando afirma: “em minha opinião devemos esperar que a probabilidade de cárie seja 0,1”.

No final, até mesmo uma posição frequencista rígida envolve análise subjetiva, e assim as diferenças provavelmente têm pouco ou nenhuma importância prática. Para melhor entendimento, imagine-se que um médico frequencista queira conhecer a chance de um certo paciente ter uma doença específica. O médico busca incluir outros pacientes que tenham características importantes semelhantes, tais como: idade, sexo, hábitos alimentares, sintomas e verifica que proporção deles teve a doença. Entretanto se o médico descesse aos mínimos detalhes sobre as características do paciente, considerando o tamanho e cor do cabelo, altura,

considerando os milímetros e peso, considerando os gramas, com certeza não encontraria outros pacientes exatamente iguais. Daí se infere que não há como criar uma classe de referência para se coletar dados experimentais, o que é um problema grande para a filosofia da ciência.

Entretanto, em 1816, o francês Pierre Simon, Marquês de Laplace, declara que as proposições sintaticamente “simétricas” com relação à evidência devem ser consideradas proposições de igual probabilidade,

A partir desse princípio, muitas lógicas indutivas foram desenvolvidas e todas elas se baseiam em uma distribuição de probabilidades “a priori”, cujos efeitos são reduzidos à medida que são conhecidos os resultados de outras observações.

É um raciocínio do tipo bayesiano que repousa sobre um “*princípio da indiferença*” segundo o qual a atitude “a priori” mais razoável nesta situação é de não dar mais importância a um caso do que a outro.

2.6. Probabilidade *a Priori*

Dada uma proposição **A**, entende-se por probabilidade incondicional ou probabilidade “a priori” associada à proposição **A**, ao grau de crença adotado para a proposição na ausência de quaisquer outras informações. A probabilidade “a priori” é representada por $P(A)$. Assim temos: Se Gravidez denota a proposição de uma determinada mulher estar grávida, então:

$$p(\textit{Gravidez}) = 0.1 \quad (4)$$

significa que, “na ausência de qualquer outra informação“, será atribuída a probabilidade de 0.1 (chance de 10%) ao evento da paciente estar grávida. É importante lembrar que $p(A)$ pode ser usada apenas quando não existe outra informação, na medida em que novas informações são conhecidas sobre o domínio, não pode mais ser usada, posto que essas novas informações impõem um raciocínio condicional da proposição **A**. (RUSSELL e NORVIG, 2004)

2.7. Probabilidade Condicional

Foi visto que para a proposição **A**, na ausência de outras informações, utiliza-se a probabilidade “a priori” como o grau de crença de um observador sobre a referida proposição. Ocorre que no domínio de um problema podem existir outras variáveis que são informações que condicionam a proposição **A**. Neste caso, o raciocínio probabilístico deve levar em consideração essa condicionalidade. Assim, supondo que uma variável **B** surgiu como nova informação no domínio, a notação utilizada será $P(A|B)$, que significa: a probabilidade de **A**, dado a evidência **B**, ou em outras palavras, a probabilidade de **A** dado que tudo o que sabemos é **B**. Neste caso, **A** é a hipótese enquanto **B** é a evidência. A esse raciocínio probabilístico é dado à denominação de Probabilidade Condicional. Assim, fazendo $A = \text{Gravidez}$ e $B = \text{Enjoo}$, e $P(A | B)$ igual a 0,8, tem-se:

$$p(A|B) = 0,8 \text{ ou } p(\text{Gravidez} | \text{Enjoo}) = 0.8 \quad (5)$$

Neste caso, indica que se foi observado que a paciente sente enjoos e ainda não existe disponível nenhuma outra informação, então a probabilidade de que a paciente esteja grávida é 0.8.

2.8. Regra do Produto

As probabilidades condicionais podem ser definidas a partir das probabilidades incondicionais. A Eq.(6) define essa relação:

$$P(A | B) = \frac{P(A \cap B)}{P(B)} \quad (6)$$

Essa equação é válida sempre que $P(B) > 0$

2.9. Axiomas da Probabilidade

Foram vistos, até aqui, os conceitos de probabilidade “a priori” e probabilidade condicional. Serão apresentadas agora as semânticas para as declarações de probabilidade.

Inicia-se pelos axiomas básicos da probabilidade. Eles são descritos conforme as equações 7, 8 e 9.

$$0 \leq P(A) \leq 1 \quad (7)$$

Significa que, para toda proposição **A** todas as probabilidades estão entre 0 e 1.

$$P(Verdadeira) = 1 \text{ e } P(Falso) = 0 \quad (8)$$

Significa que, as proposições verdadeiras, ou seja, válidas, tem probabilidade 1, enquanto as falsas tem probabilidade zero, quando se tratar de eventos dicotômicos.

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad (9)$$

A equação (9) conecta as probabilidades de proposições logicamente inter-relacionadas para dois eventos quaisquer, não exigindo que **A** e **B** sejam mutuamente exclusivos.

2.10. A Regra de Bayes

A partir da regra do produto, Eq.(6), podemos reescrevê-la de duas formas, conforme equações 10 e 11:

$$P(A \cap B) = P(A | B)P(B) \quad (10)$$

$$P(A \cap B) = P(B | A)P(A) \quad (11)$$

Se os dois membros da direita das Eq. (9) e (10) forem igualados e divididos, será obtida a equação (12):

$$P(B | A) = \frac{P(A | B)P(B)}{P(A)} \quad (12)$$

A equação (12) é conhecida como Regra de Bayes, ou Teorema de Bayes ou Lei de Bayes, onde:

$P(B | A)$ = É a probabilidade “a posteriori”, ou a probabilidade condicional de **B** dado **A**.

$P(A | B)$ = É a verossimilhança, ou a probabilidade condicional de **A** dado **B**.

$P(B)$ = É a probabilidade independente, ou a probabilidade “a priori”.

$P(A)$ = É a probabilidade independente de A.

À primeira vista, a regra de Bayes não parece ser de grande utilidade, uma vez que para aplicá-la no cálculo de uma probabilidade condicional $P(A | B)$ são necessárias uma probabilidade condicional $P(B | A)$ e duas probabilidades incondicionais $P(A)$ e $P(B)$. Entretanto, tal fórmula é útil na prática porque, em muitas situações, existem boas estimativas para $P(B | A)$, $P(A)$, $P(B)$ e existe a necessidade de se obter $P(A | B)$.

Em diagnósticos médicos, por exemplo, frequentemente as probabilidades condicionais de relações causais são conhecidas, e se deseja derivar um diagnóstico. Um médico sabe, por exemplo, que meningite causa rigidez no pescoço em 50% dos casos. Ele sabe, também, alguns fatos incondicionais: a probabilidade “a priori” de um paciente ter meningite é $1/50000$ e a probabilidade “a priori” de qualquer paciente sofrer de rigidez no pescoço é $1/20$.

Notando por:

M : o paciente tem meningite

R : o paciente está com o pescoço rígido

tem-se :

$$p(R|M) = \frac{1}{2} \quad p(M) = \frac{1}{50000} \quad p(R) = \frac{1}{20}$$

$$p(M|R) = \frac{p(R|M) \cdot p(M)}{p(R)} = \frac{\frac{1}{2} \cdot \frac{1}{50000}}{\frac{1}{20}} = 0.0002 = \frac{1}{5000} \quad (13)$$

ou seja, a expectativa é de que apenas 1 em 5000 pacientes com rigidez no pescoço tenha meningite.

É importante notar que muito embora a presença de meningite seja um forte indicativo para a rigidez do pescoço $p(R|M) = \frac{1}{2}$, a probabilidade de um paciente com rigidez no pescoço ter meningite permanece baixa. Isso se deve ao fato da probabilidade “a priori” de rigidez no pescoço $p(R) = \frac{1}{20}$ ser muito maior do que a probabilidade “a priori” de meningite $p(M) = \frac{1}{50000}$.

Uma pergunta óbvia que se faz é por que não se fornece diretamente a probabilidade condicional $p(M|R)$, já que se fornece a probabilidade $p(R|M)$ para ser usada em seu cálculo? No caso específico da meningite, o médico talvez já saiba que 1 entre 5000

pacientes com rigidez no pescoço tem meningite, o que torna desnecessário o uso da regra de Bayes.

Entretanto, se acontecer uma súbita epidemia de meningite, por exemplo, a probabilidade “a priori” de meningite – $p(M)$ – irá aumentar. O médico que derivou $p(M|R)$ a partir de observações estatísticas de pacientes, antes da epidemia, não terá a menor ideia de como atualizar este valor. Por outro lado, o médico que calculou $p(M|R)$ a partir das 3 outras probabilidades, verá que $p(M|R)$ aumentará proporcionalmente ao aumento de $p(M)$. Desta forma, a informação $p(R|M)$ permanece inalterada com a epidemia, uma vez que apenas descreve a forma como a meningite atua, que independe da meningite acontecer isoladamente ou em caráter epidêmico.

2.10.1. Normalização da Regra de Bayes

Notando por:

S: o paciente tem sarampo

A: o paciente tem alergia

V: o paciente tem manchas vermelhas pelo corpo

A equação (12) pode ser utilizada para calcular a probabilidade de o paciente ter sarampo, dado que foram observadas manchas vermelhas pelo corpo :

$$p(S|V) = \frac{p(V|S) * p(S)}{p(V)} \quad (14)$$

Entretanto, existe a possibilidade, por exemplo, de que o paciente esteja com alergia, dado que apresenta manchas vermelhas. A probabilidade de A dado V é então:

$$p(A|V) = \frac{p(V|A) * p(A)}{p(V)}$$

$$p(S|V) = 0.9$$

$$p(S) = \frac{1}{9000}$$

$$p(V|A) = 0.6$$

$$p(A) = \frac{1}{1000}$$

Pode-se calcular a Razão Relativa entre Sarampo e Alergia, observadas as manchas vermelhas:

$$\frac{p(S|V)}{p(A|V)} = \frac{p(V|S)*p(S)}{p(V|A)*p(A)} = \frac{0.9*\frac{1}{9000}}{0.6*\frac{1}{1000}} = \frac{1}{6} \quad (15)$$

O que indica que a possibilidade de que o paciente esteja com alergia é 6 vezes maior do que a possibilidade de que ele esteja com sarampo, dado que manchas vermelhas são observadas.

2.11. Redes Bayesianas

Por definição, uma Rede Bayesiana (RB) é composta dos seguintes elementos (JENSEN, 2001)

- I. Um conjunto de variáveis e um conjunto de arestas direcionadas entre as variáveis;
- II. Cada variável tem estados finitos e mutuamente exclusivos;
- III. As variáveis e as arestas direcionadas representam um grafo acíclico direcionado;
- IV. Cada variável A , com pais B_1, B_2, \dots, B_n , possui uma tabela de probabilidades condicionais, $P(A | B_1, B_2, \dots, B_n)$, associada.

Em outras palavras, RB são grafos acíclicos e direcionados que permitem a representação da distribuição conjunta de probabilidades de um conjunto de variáveis aleatórias. Cada vértice do grafo representa uma variável aleatória e as arestas representam dependências entre variáveis. Em uma RB, se há uma aresta direcionada entre os nós A e B , então dizemos que A é pai de B e B é filho de A . Neste trabalho serão tratadas somente as RB com variáveis discretas.

RB com variáveis discretas satisfazem a condição de Markov (NEAPOLITAN, 2004), que é dada por: cada variável da RB é condicionalmente independente do conjunto de todos os seus não-descendentes dado o conjunto de todos os seus pais. Em uma RB, a distribuição conjunta de probabilidades de um conjunto de variáveis discretas $\{X_1, X_2, \dots, X_n\}$ é dada pela Regra da Cadeia,

$$P(X_1, X_2 \dots X_n) = \prod_{i=1}^n P(X_i | Pa_i) \quad (16)$$

Os parâmetros de uma RB são definidos como:

$$\theta_i = P(X_i | Pa_i) \quad i = 1, \dots, n \quad (17)$$

em que, θ_i é uma tabela de probabilidade condicionais de X_i dado Pa_i , onde Pa_i representa os pais de X_i .

Com isso, o conjunto de parâmetros de uma RB é dado por $\Theta_s = \{\theta_1, \theta_2, \dots, \theta_n\}$ e todas são tabelas de probabilidade condicional da RB com variáveis discretas $\{X_1, X_2, \dots, X_n\}$.

Um importante aspecto de uma RB é a sua estrutura (topologia do grafo), que permite a representação de relações complexas entre variáveis de forma gráfica e intuitiva. A estrutura gráfica de uma RB facilita o entendimento das relações entre variáveis do seu domínio, além de permitir o uso combinado de informações obtidas do conhecimento de especialistas com dados históricos para obter a distribuição conjunta de probabilidades da rede.

A estrutura de uma RB pode ser determinada manualmente, com apoio de especialistas, ou pode ser aprendida a partir de bases de dados utilizando algoritmos de aprendizado de estrutura. Os parâmetros de uma RB podem ser obtidos através da elucidação² de probabilidades por especialistas, através do aprendizado a partir de bases de dados ou através da combinação de ambas as abordagens.

Na Figura 1 há um exemplo de RB, que utiliza variáveis discretas, aplicada no diagnóstico de doenças. As variáveis desta RB são {Idade (I), Profissão (P), Clima (C), Doença (D), Sintomas (S)}.

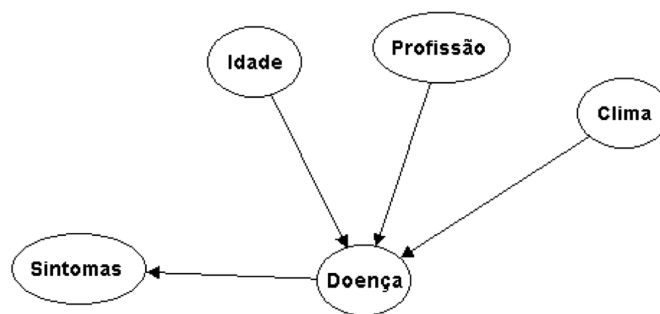


Figura 1- Exemplo de Rede Bayesiana aplicada em diagnóstico médico

Na RB da Figura 1, o nó “Sintomas” tem o nó “Doença” como pai e três ancestrais “Idade”, “Profissão” e “Clima”. Através das premissas de independência condicional, podemos dizer que “Sintomas” é dependente de “Idade”, “Profissão” e “Clima” indiretamente através de sua influência sobre “Doença”. Tomando o nó “Clima”, que não possui pai, então podemos dizer que “Clima” é independente de “Profissão” e “Idade”. Aplicando a Regra da Cadeia, equação (16), a distribuição conjunta de probabilidades desta RB é dada por,

² Elucidação é o procedimento de obtenção de distribuições a partir do conhecimento de especialistas.

$$P(\text{Idade}, \text{Profissão}, \text{Clima}, \text{Doença}, \text{Sintomas}) = P(\text{Idade}).P(\text{Profissão}).P(\text{Clima}). P(\text{Doença} | \text{Idade}, \text{Profissão}, \text{Clima}). P(\text{Sintomas} | \text{Doença}) \quad (18)$$

Utilizando somente as letras iniciais de cada variável a equação (18) pode ser reescrita como,

$$P(I, P, C, D, S) = P(I).P(P).P(C).P(D|I, P, C).P(S|D) \quad (19)$$

Pela Regra da Cadeia, as tabelas de probabilidades condicionais de cada variável da RB precisam ser especificadas para que se obtenha a distribuição de probabilidades conjunta da RB. Com isso, na RB da Figura 1, é necessário especificar as tabelas $P(\text{Idade})$, $P(\text{Profissão})$, $P(\text{Clima})$, $P(\text{Doença}|\text{Idade}, \text{Profissão}, \text{Clima})$ e $P(\text{Sintomas}|\text{Doença})$ para determinarmos a distribuição conjunta de probabilidade $P(\text{Idade}, \text{Profissão}, \text{Clima}, \text{Doença}, \text{Sintomas})$.

Na Tabela 2, temos as tabelas $P(\text{Idade})$ e $P(\text{Sintomas}|\text{Doença})$. As probabilidades apresentadas nestas tabelas também podem ser chamadas de parâmetros, como definimos previamente em (7). Note que a variável “Idade”, que é uma variável contínua, foi discretizada (ou categorizada) para criar uma variável discreta binária.

| | |
|------------|------|
| Idade < 45 | 0.46 |
| Idade ≥ 45 | 0.54 |

Tabela 2 - Distribuição de "Idade"

| Sintomas | Doença | | |
|-----------------|--------------------|---------|---------|
| | Úlcera no Estômago | Infarto | Nenhuma |
| Dor de Estômago | 0.8 | 0.05 | 0.05 |
| Dor no Peito | 0.15 | 0.90 | 0.10 |
| Nenhuma | 0.05 | 0.05 | 0.85 |

Tabela 3 - Probabilidades Condicionais de "Doença" dado "Sintomas"

Se no exemplo das Tabelas 1 e 2, se não fossem utilizadas as suposições de RB e a Regra da Cadeia, ao invés de 5 tabelas seria necessário definir uma grande tabela de probabilidades para obtermos a distribuição conjunta das 5 variáveis. Com isso, a RB fornece uma maneira de simplificar a representação de uma distribuição conjunta de probabilidades.

RB também podem ser utilizadas para calcular novas probabilidades (“a posteriori”) a partir de informações (evidências) sobre uma ou mais variáveis da rede. Em uma RB, um subconjunto de variáveis E com valores conhecidos, $E=e$, em uma dada situação, é conhecido como conjunto de evidência, ou simplesmente evidência, por exemplo, $E=\{X_2=x_2, X_6=x_6\}$. Também podemos dizer que uma variável está instanciada (evidência forte) se conhecemos o estado desta variável.

Um conceito importante em RB é o de d-separação. Segundo (JENSEN, 2001), dizemos que dois vértices distintos X e Y estão d-separados em uma RB se, para todos os caminhos entre X e Y existe um vértice intermediário V (distinto de X e Y) tal que a conexão entre X e Y através de V :

- é serial ou divergente e V recebeu uma evidência ou;
- é convergente e nem V nem algum de seus descendentes receberam uma evidência.

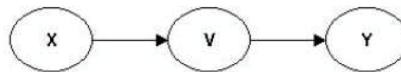


Figura 2 - Conexão Serial: X e Y estão d-separados se V recebeu uma evidência

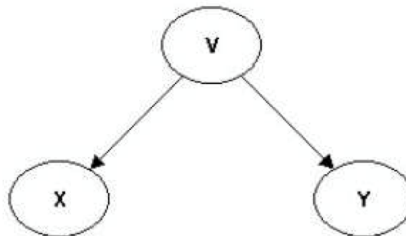


Figura 3 - Conexão Divergente: X e Y estão d-separados se V recebeu uma evidência.

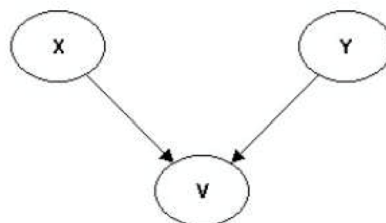


Figura 4 - Conexão Convergente: X e Y estão d-separados se nem V nem seus descendentes recebeu uma evidência

Em RB se dois vértices quaisquer estão d-separados então eles são condicionalmente independentes. Através do conceito de d-separação é possível identificar a condição de independência condicional entre variáveis em uma RB.

Outro conceito importante em RB é o de Markov Blanket. O Markov Blanket de uma variável X é o conjunto das variáveis que são pais de X, filhos de X e as variáveis que compartilham um filho com X (JENSEN, 2001). Com isso, se todas as variáveis do Markov Blanket de X possuem evidências, então X está d-separado de todas as outras variáveis da RB e, conseqüentemente, X é condicionalmente independente de todas as outras variáveis da rede, dado seu Markov Blanket. No exemplo da Figura 5, o Markov Blanket de I é {C, E, K, L, H}.

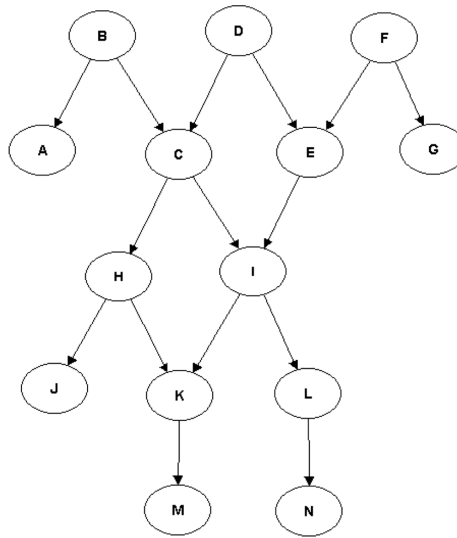


Figura 5 - Outro exemplo de Rede Bayesiana (Markov Blanket)

2.11.1. Inferência em Redes Bayesianas

A Inferência em RB é o processo de atualização das probabilidades “a posteriori” de variáveis dado uma evidência fornecida. Em RB, a evidência pode ser definida para qualquer subconjunto de nós e a probabilidade “a posteriori” pode ser calculada para qualquer outro subconjunto de nós.

Como uma RB contém a distribuição de probabilidade conjunta de todas as variáveis do seu domínio, então é possível obter a distribuição de probabilidades de qualquer variável do seu domínio a partir da Regra da Probabilidade Total. No exemplo de RB da Figura 1, a distribuição da variável “Doença” (D), pela Regra da Probabilidade Total, é dada (inferida) por,

$$P(D = d) = \sum_i \sum_p \sum_c \sum_s P(I = i, P = p, C = c, D = d, S = s) \quad \forall d \quad (20)$$

Pela Regra da Cadeia, a equação (19) pode ser escrita como,

$$P(D = d) = \sum_i \sum_p \sum_c \sum_s P(I = i).P(P = p).P(C = c).P(D = d|I = i, P = p, C = c).P(S = s|D = d) \quad \forall d \quad (21)$$

em que, i , p , c , d e s representam, respectivamente, cada estado das variáveis “Idade”, “Profissão”, “Clima”, “Doença”, “Sintoma”.

Segundo Zhang e Poole (1996) Inferência em uma RB se refere ao processo de cálculo da probabilidade “a posteriori” $P(X|Y = Y_o)$ de um conjunto de variáveis X depois de obter algumas observações (evidência) $Y=Y_o$. Aqui Y é uma lista de variáveis observadas e Y_o é a lista de valores observados. Pelo Teorema de Bayes, $P(X|Y = Y_o)$ é dada por,

$$P(X|Y = Y_o) = \frac{P(X, Y=Y_o)}{P(Y=Y_o)} \quad (22)$$

Com isso, $P(X|Y = Y_o)$ é obtida a partir da distribuição marginal $P(X, Y)$, que por sua vez é calculada a partir da distribuição conjunta $P(X_1, X_2, \dots, X_n)$ pela soma de probabilidades de todas as variáveis fora do domínio de $X \cup Y$ uma por uma. No entanto, isso não é viável, pois esta soma fora do domínio de $X \cup Y$ requer um número exponencial de adições (ZHANG e POOLE, 1996).

Para exemplificar o procedimento de inferência utilizaremos o exemplo de RB da Figura 1. Dada uma evidência $E=e=\{Idade='<45'; Sintoma='Dor de estômago'\}$, queremos obter distribuição “a posteriori” da variável “Doença” (D) que é,

$$P(D = d|E = e) = \frac{P(D=d, E=e)}{P(E=e)} = \frac{P(D=d, I='<45', S='Dor de estômago')}{P(I='<45', S='Dor de estômago')} \quad \forall d \quad (23)$$

que é igual a,

$$P(D = d|E = e) = \frac{\sum_p \sum_c P(I='<45', P=p, C=c, D=d, S='Dor de estômago')}{\sum_p \sum_c \sum_d P(I='<45', P=p, C=c, D=d, S='Dor de estômago')} \quad \forall d \quad (24)$$

Aplicando a Regra da Cadeia o numerador da equação (24) pode ser escrito como, $\sum_p \sum_c P(I='<45').P(P = p).P(C = c).P(D = d|I='<45', P = p, C = c).P(S = 'Dor de estômago|D = d')$ (25)

O denominador da equação (24) pode ser escrito como,

$$\sum_p \sum_c \sum_d P(I='<45').P(P = p).P(C = c).P(D = d|I='<45', P = p, C = c).P(S = 'Dor de estômago|D = d') \quad (26)$$

em que p , c , d representam, respectivamente, cada estado das variáveis “Idade”, “Profissão”, “Clima” e “Doença”.

Para diminuir o tempo necessário para o cálculo das probabilidades “a posteriori”, diversos algoritmos de Inferência têm sido propostos. Existem dois tipos de algoritmos de Inferência em RB: os exatos e aproximados. Entre os algoritmos exatos destaca-se o algoritmo Junction Tree proposto por Jensen (2001), que é baseado em Teoria dos Grafos. Os algoritmos de Inferência exatos podem não ser eficientes em redes com um grande número de nós e arestas, pois o problema de Inferência é NP-hard (ZHANG e POOLE, 1996). Para estes

casos, podem ser utilizados algoritmos aproximados, baseados em simulação estocástica, tais como: *Forward Sampling*, *Likelihood Weighting*, *Gibbs Sampling*, *Metropolis-Hasting* (NEAPOLITAN, 2004).

Apesar da complexidade inerente do procedimento de inferência, sistemas que possuem RB como base de conhecimento têm se mostrado muito eficientes e têm sido muito difundidos ganhando importância inclusive em áreas comerciais. Os sistemas *Hugin*, *Netica* e *JavaBayes* são exemplos destes sistemas.

2.12. Classificação Bayesiana

Redes Bayesianas (RB) podem ser utilizadas em problemas de classificação de uma maneira clara e direta e as RB utilizadas em problemas de classificação de dados são chamadas de Classificadores Bayesianos. Nos Classificadores bayesianos com variáveis discretas $\{A_1, A_2, \dots, A_n, C\}$, uma delas, C , é a variável classe (variável resposta) e as demais, $\{A_1, A_2, \dots, A_n\}$, são os atributos (variáveis preditoras).

2.12.1. Classificador Naive Bayes

Os Classificadores Bayesianos mais simples conhecidos são os chamados Naive Bayes. Os Classificadores Naive Bayes partem da hipótese que todos os atributos são independentes, dada a variável classe, e sua representação gráfica é dada na Figura 6.

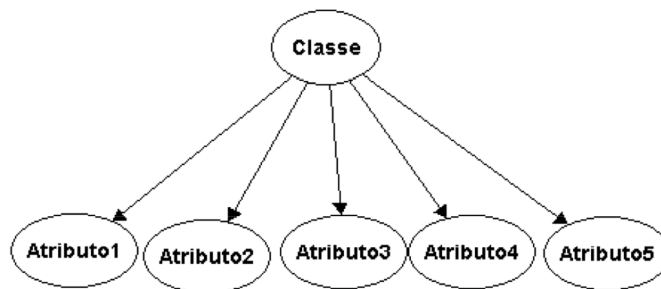


Figura 6 - Modelo de uma Naive Bayes

Sob a hipótese de independência condicional entre atributos dada a classe, aplicando a Regra da Cadeia é obtida a distribuição conjunta de probabilidades do Classificador Naive Bayes dada por,

$$P(A_1, \dots, A_n, C) = P(C) \prod_{k=1}^n P(A_k | C) \quad (27)$$

Em um Classificador bayesiano, com atributos discretos e classe C , assumindo valores $\{0,1\}$, a probabilidade de classificarmos um novo caso, $\{A_1 = a_1, \dots, A_n = a_n\}$, em $C=1$ é:

$$P(C = 1|A_1 = a_1 \dots, A_n = a_n) = \frac{P(C=1).P(A_1=a_1 \dots, A_n=a_n|C=1)}{P(A_1=a_1 \dots, A_n=a_n)} \quad (28)$$

E a probabilidade de classificarmos um novo caso em $C=0$ é,

$$P(C = 0|A_1 = a_1 \dots, A_n = a_n) = \frac{P(C=0).P(A_1=a_1 \dots, A_n=a_n|C=0)}{P(A_1=a_1 \dots, A_n=a_n)} \quad (29)$$

Com isso, uma nova observação (caso), $\{A_1 = a_1, \dots, A_n = a_n\}$, é classificada na classe $C=1$ segundo o seguinte critério:

$$\frac{P(C=1|A_1=a_1, \dots, A_n=a_n)}{P(C=0|A_1=a_1, \dots, A_n=a_n)} \geq 1 \quad (30)$$

O critério descrito em (29) que pode ser escrito como,

$$\frac{P(C=1)}{P(C=0)} \cdot \frac{P(A_1=a_1, \dots, A_n=a_n|C=1)}{P(A_1=a_1, \dots, A_n=a_n|C=0)} \geq 1 \quad (31)$$

No caso do Classificador bayesiano *Naive Bayes*, um novo caso $\{A_1 = a_1, \dots, A_n = a_n\}$ é classificado em $C=1$ segundo o seguinte critério:

$$\frac{P(C=1)}{P(C=0)} \cdot \prod_{i=1}^n \frac{P(A_i=a_i|C=1)}{P(A_i=a_i|C=0)} \geq 1 \quad (32)$$

O Classificador *Naive Bayes* é conhecido por sua simplicidade e eficiência, pois apresentam estrutura fixa e parâmetros ajustáveis. Embora sua suposição de independência seja problemática, pois esta hipótese raramente se verifica no mundo real, os Classificadores *Naive Bayes* têm apresentado um bom desempenho em um grande número de aplicações, especialmente naquelas onde as variáveis preditoras não são fortemente correlacionadas (CHENG e GREINER, 2001).

2.12.2. Classificador Noisy-or

O excessivo número de probabilidades necessárias para a quantificação de uma rede é uma das maiores dificuldades na aplicação prática de redes Bayesianas. Em um nó X com k categorias o número de probabilidades a serem especificadas é:

$$N^o \text{ de elementos da tabela}(X) = P(x_i) \prod_{pa(X)} n_{categoria \ pa(X)} \quad (33)$$

No entanto, para determinados tipos de nós, estas probabilidades podem ser calculadas a partir de outras em vez de serem especificadas diretamente. O modelo *noisy-or* (HECKERMAN e BREESE, 1994) permite tal cálculo, com a restrição que os pais do nó possam contribuir independentemente para a probabilidade do nó em que é aplicado o *noisy-or*, e que o efeito combinado de diversos pais contribua cumulativamente na probabilidade do nó. Neste trabalho, também é necessário que o nó em questão seja binário, isto é, possua somente duas categorias, com uma categoria para representar verdadeiro e outra para falso.

Utilizando nomes adequados para nós em redes modelando problemas médicos, tem-se um nó D com categorias d e \bar{d} representando uma doença, cujas causas são R_1, R_2, \dots, R_n , as probabilidades para D são dadas pela tabela conjunta de probabilidades condicionais $P(D/R_1, R_2, \dots, R_n)$.

O modelo *noisy-or* permite o cálculo da tabela conjunta de probabilidades condicionais a partir das probabilidades $P(D/R_i)$ para cada nó pai R_i , desde que seja respeitada a restrição em relação aos pais posta anteriormente.

No caso de ter-se somente nós binários R_i , com estados r_i (verdadeiro) e \bar{r}_i (falso), estas probabilidades condicionais são conhecidas como sensibilidade $P(d/r_i)$ e especificidade $P(\bar{d}|\bar{r}_i)$ que, muitas vezes, estão disponíveis em estudos populacionais. Já as probabilidades conjuntas $P(D/R_1, R_2, \dots, R_n)$ são de mais difícil obtenção a partir de especialistas ou de informações bibliográficas por envolverem um alto número de combinações de condições.

A expressão que representa a probabilidade de cada fator causar a doença independentemente é calculada como:

$$p_i = P(d|r_i \text{ apenas}) = P(d|\bar{r}_1, \bar{r}_2, \dots, r_i, \dots, \bar{r}_n) \quad (34)$$

As probabilidades p_i são conhecidas como probabilidades de ligação e cada valor da probabilidade conjunta é calculado como:

$$P(d|H) = 1 - (\prod_{R_i \in H^+} [1 - p_i]) \quad (35)$$

onde H é uma configuração de (R_1, R_2, \dots, R_n) e H^+ é o subconjunto de nós definidos como verdadeiros.

A maneira proposta por (NIKOVSKI, 2000) para se calcular as probabilidades de ligação parte das sensibilidades e especificidades para cada nó pai. A probabilidade de ligação p_i é estimada considerando-se que D pode ser causada por um fator R_i ou pela atuação combinada de todos os outros fatores que levam a D , isto é, os outros R . Esta outra causa pode ser chamada de R_{tudo} , com uma probabilidade de ligação p_{tudo} . Assim,

$$\begin{aligned} P(d|r_i) &= p_i + p_{tudo} - p_i * p_{tudo} \\ P(d|\bar{r}_i) &= p_{tudo} \end{aligned}$$

Simplificando,

$$p_i = \frac{P(d|r_i) - P(d|\bar{r}_i)}{1 - P(d|\bar{r}_i)} \quad (36)$$

Após calcular todos os p_i , é possível obter os valores para todas as hipóteses H .

A aplicação do *noisy-or* como apresentado acima desconsidera fatores não modelados. Em qualquer representação do mundo real, existirão alguns fatores que não foram modelados. A falta desses fatores pode ser intencional, para evitar um modelo

demasiadamente complexo, ou accidental, por mero desconhecimento de quais seriam estes fatores.

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, distribuímos os valores para as probabilidades de não apresentar febre dadas as evidências de apresentar ou não, determinadas características (onde encontramos \neg lê-se “não apresenta”) este exemplo está descrito em (RUSSELL e NORVIG, 2004)

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

A tabela 4 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): multiplica-se a $P(\neg \text{febre} \mid \neg \text{resfriado}, \neg \text{gripe}, \text{malária}) = 0,1$, $P(\neg \text{febre} \mid \neg \text{resfriado}, \text{gripe}, \neg \text{malária}) = 0,2$ e $P(\neg \text{febre} \mid \text{resfriado}, \neg \text{gripe}, \neg \text{malária}) = 0,6$, logo chega-se no resultado que a $P(\neg \text{febre} \mid \text{resfriado}, \text{gripe}, \text{malária})$ é de 0,012), portanto a de apresentar febre ($P(\text{febre} \mid \text{resfriado}, \text{gripe}, \text{malária})$) é de 0,988 e assim é feito sucessivamente.

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

Tabela 4 - Tabela de distribuição de probabilidades

Geralmente esses tipos de distribuições são usados 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, 2004).

3. PROJETO LARIISA

“O LARIISA é um *framework* para tomada de decisão em governança para sistemas públicos de saúde” (OLIVEIRA, HAIRON, *et al.*, 2010). Ele integra dados sobre residências de famílias em um novo sistema inteligente de informação para cuidados de saúde.

Com base em cinco domínios de governança (conhecimento, normativo, clínico-epidemiológico, administrativo e gerenciamento compartilhado), o *framework* contará com uma infraestrutura de comunicação óptica e sem fio (WiMAX), o Cinturão Digital do Ceará, que segundo a Etice³ (2011), oferecerá banda larga de alta velocidade cobrindo 82% da população urbana do Estado.

São mais de 2.500Km de fibra ótica de alta velocidade (1 a 2Gbps) conectando escolas, hospitais, postos de saúde, delegacias e demais órgãos públicos. Servirá para telemedicina, educação a distância, TV digital, videoconferência e outros serviços para o desenvolvimento do Estado e a qualidade de vida da população.

Caracterizada pela informação em tempo real e sistemas de inferência baseada em um modelo de ontologias, a plataforma LARIISA está centrada no conceito de informação de contexto de entidades em um sistema de saúde. Uma entidade é, por exemplo, um membro da família, um agente de saúde, gestor da saúde, entre outros, que são considerados relevantes para as interações entre um usuário e um sistema de saúde, a fim de fornecer sistemas inteligentes.

3.1. Modelos de Ontologia local e global

O LARIISA define um contexto formal em saúde, a fim de facilitar a representação do contexto, compartilhamento e interoperabilidade semântica no sistema de governança da saúde. Para tanto, o LARIISA define duas ontologias OWL-DL (OLIVEIRA, HAIRON, *et al.*, 2010): modelagens de informações de contexto de saúde local e global.

Contexto de saúde local descreve a situação de qualquer entidade interagindo com o sistema de governança, tais como usuários finais (pacientes), gestores de saúde, agentes de saúde, etc. Essas informações são utilizadas para a definição de regras de decisão locais em saúde e para construir o contexto de saúde global.

³ A Empresa de Tecnologia da Informação do Estado do Ceará (ETICE) é uma empresa pública responsável pelos serviços de TICs no Governo do Estado do Ceará, entre eles o Cinturão Digital.

O contexto de saúde global descreve informações de alto nível (figura 7), derivado do contexto de saúde local, e é utilizado para tomada de decisão em governança de saúde. Por exemplo, ele descreve o número de casos de dengue confirmados em uma região (ex: bairro, cidade, comunidade), durante um determinado período de tempo (ex: um dia, uma semana). Portanto, essas informações podem ser vistas como indicadores globais utilizados para melhorar as decisões de governança.

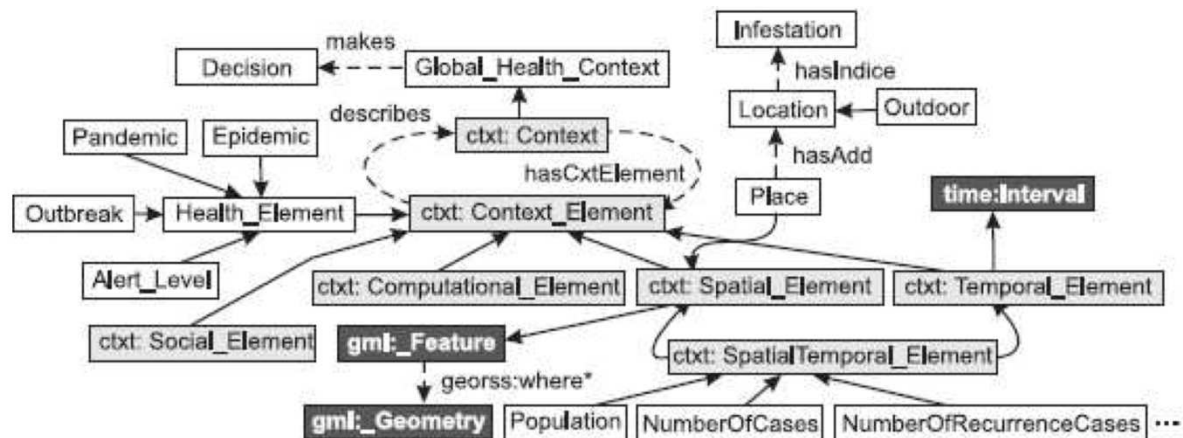


Figura 7 - Modelo Global de Saúde do LARIISA

Os contextos de saúde local e global são classificados em seis dimensões:

- Espacial – quaisquer informações que caracterizem a situação da dimensão espacial (ex: localização, local, coordenadas GPS).
- Temporal – quaisquer informações que caracterizem a situação da dimensão do tempo (ex: instante, intervalo, período do dia, período do mês, período do ano, estação).
- Espaço-Temporal – quaisquer informações que caracterizam a situação que é dependente tanto da dimensão espacial quanto da dimensão temporal (ex: condições climáticas, temperatura, ruído, luminosidade).
- Social – quaisquer informações que caracterizem a situação dos relacionamentos sociais.
- Computacional – quaisquer informações que descrevem a situação das características computacionais (ex: configuração de dispositivos do usuário).
- Elemento de saúde – classifica o contexto da informação a partir do ponto de vista da saúde (ex: batimento cardíaco, pulso, pressão sanguínea).

O LARIISA reutiliza conceitos do *Geographically Encoded Objects for Really Simple Syndication feeds* (GeoRSS), uma marcação com informação de localização para descrição de coordenadas e relações geo-espaciais, assim como o OWLTime, utilizado para representar conteúdo temporal (OLIVEIRA, HAIRON, *et al.*, 2010).

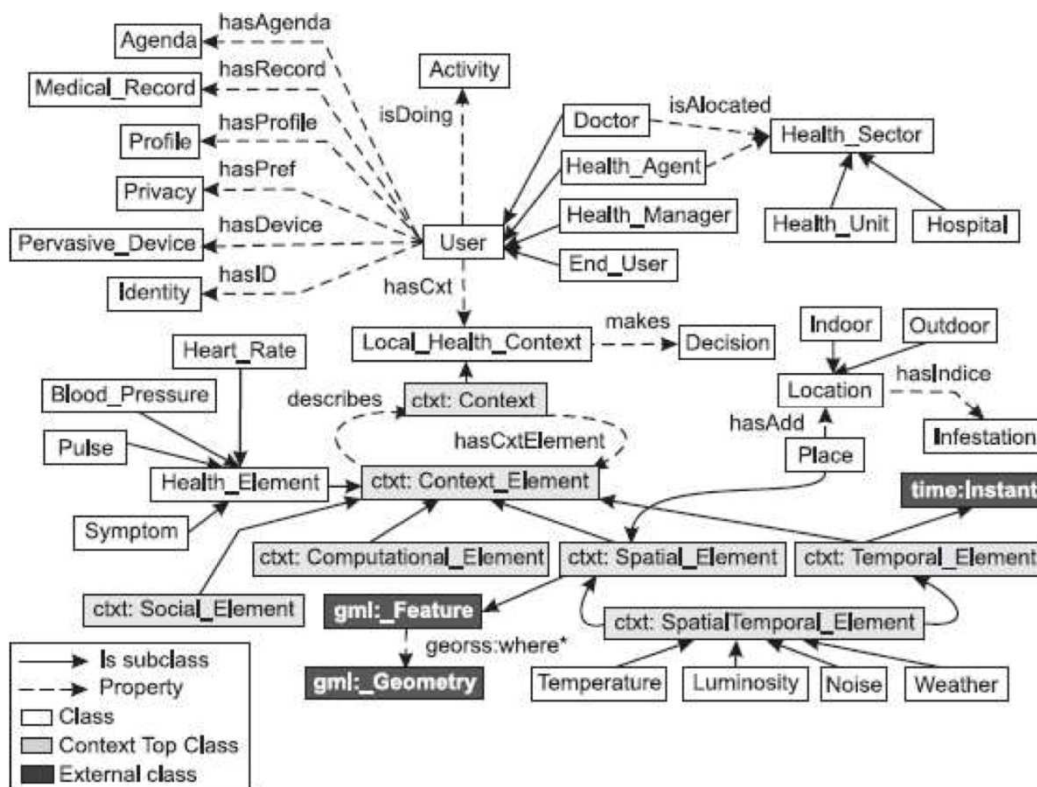


Figura 8 - Modelo Local de Saúde do LARIISA

O LARIISA define as classes *Global_Health_Context* (Figura 7) e *Local_Health_Context* (Figura 8) e. Elas capturam do contexto quaisquer informações para caracterizar uma situação que é relevante para contribuir com decisões em governança de saúde, isto é, que podem ser utilizadas para definir regras de decisão local e global. O *framework* utiliza a base do modelo *Event-Condition-Action* (ECA) para descrever regras de decisão local e global que são traduzidas dentro de regras utilizando a *Semantic Web Rule Language* (SWRL), “uma linguagem com sintaxe de abstração de alto nível para regras da OWL” (HORROCKS, 2002). Um evento representa a identificação de mudanças no contexto. Uma condição descreve um conjunto válido de restrições de contexto, e uma ação descreve uma decisão.

3.2. Cenário de utilização com metadado e geolocalização

A Figura 9 apresenta uma visão geral do sistema proposto no cenário de geolocalização, dividido em três partes principais: Aquisição de dados, Processamento de Dados e Publicação.

Aquisição de dados refere-se à aplicação do sensor, a informação (por exemplo, identificação de usuário e sintomas) adicionada pelo usuário, e os dados adquiridos.

Depois disso, todos os dados adquiridos serão enviados para a etapa de Processamento de Dados. Nesta parte, o sistema utiliza os dados brutos, a fim de capturar os dados de diagnóstico necessários (GARDINI, BRAGA, *et al.*, 2013).

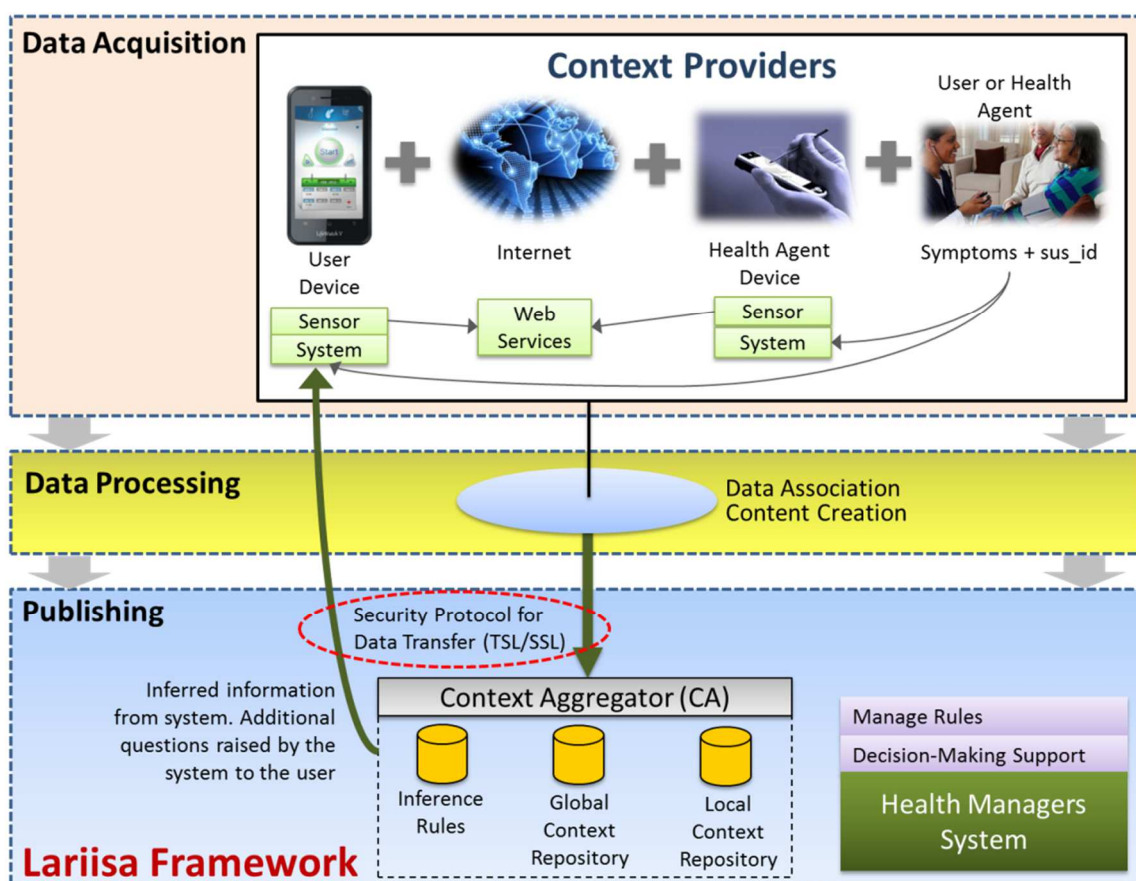


Figura 9 - Arquitetura do CLARIISA

Como pode ser observado na figura 9, as informações de saúde para alimentar o LARIISA podem ser originadas de dispositivos móveis de pacientes ou de agentes de saúde. Uma camada de segurança para transmissão dos dados pela *Internet* é implementada no cenário proposto. O protocolo TLS (*Transport Layer Security*) foi utilizado para prover integridade e confidencialidade das informações, protegendo os dados de saúde do paciente contra acessos ou modificações não autorizadas.

Identificar qual usuário está enviando dados de saúde enriquecidos para o banco de dados do LARIISA é crucial para o *framework* proposto. Sem esta identificação, não é possível determinar quem está enviando dados vitais de saúde para o sistema. Um identificador único precisa ser informado pelo usuário no momento que ele/ela inicia um novo diagnóstico de saúde do seu dispositivo móvel. O identificador único escolhido foi o Número do SUS, um número de identificação designado para todo cidadão brasileiro como parte de um registro nacional de usuários – para consolidação do Sistema Único de Saúde do Brasil (SUS). Neste contexto, o Número SUS pode ser utilizado para registrar uma nova coleta de dados de saúde e também para pesquisar informações de outros bancos de dados tal como a base de dados do SUS, base de dados de hospitais, etc.

Enquanto um usuário está registrando seu diagnóstico remoto, sensores coletam informações de contexto, tal como posição geográfica, direção, data, e clima. As informações coletadas são manipuladas pela segunda parte da arquitetura proposta (Processamento dos Dados) para então serem tratadas e enviadas para a base de dados do LARIISA.

Uma das principais características de sistemas sensíveis a contexto é a localização. O sistema proposto também utiliza esta característica. O sensor GPS coleta a localização geográfica de um diagnóstico realizado, e então adiciona esta informação no metadado, como pode ser observado na figura 10. Esta ação é importante para ajudar a geração de conteúdo bem como, para adquirir novas informações (ex: nome da localidade, clima, etc) de um diagnóstico que está sendo realizado.

A arquitetura observada na figura 9 considera a utilização de ambos dispositivos móveis: dispositivos de usuários e dispositivos móveis de agentes de saúde.

A etapa de Processamento de Dados é responsável por incrementar a funcionalidade do sistema. Esta etapa do sistema associa e organiza as informações de forma a fornecer uma estrutura compreensível para ser publicada no banco de dados do LARIISA.

A etapa de Processamento de Dados é responsável por incrementar a funcionalidade do sistema. Esta etapa do sistema associa e organiza as informações de forma a fornecer uma estrutura compreensível para ser publicada no banco de dados do LARIISA. Fazendo uso dos dados adquiridos e organizados por “*tags*”, a etapa de Processamento de Dados é iniciada no sistema. A principal ideia é usar informação de contexto de diagnósticos remotos para suportar a tomada de decisão, considerando ambos os cenários, Global e Local – como comentado no início deste capítulo. O sistema fornece diagnóstico de saúde com dados enriquecidos baseado nas informações adquiridas pela aplicação instalada no dispositivo móvel. Por exemplo, se um usuário inicia um diagnóstico de saúde remoto, o sistema gerará

um novo diagnóstico com as coordenadas S 3° 45' 48.6429", W 38° 36' 28.7434" às 17h na data de 03/02/2013. Além disso, a aplicação do dispositivo móvel captura dados dos sensores médicos (temperatura corporal, pressão sanguínea, batimentos cardíacos, etc.), e o usuário é requisitado a adicionar seus sintomas e também seu Número do SUS. De acordo com a figura 10, todos os dados são agrupados e posteriormente transferidos para a base de dados do LARIISA.

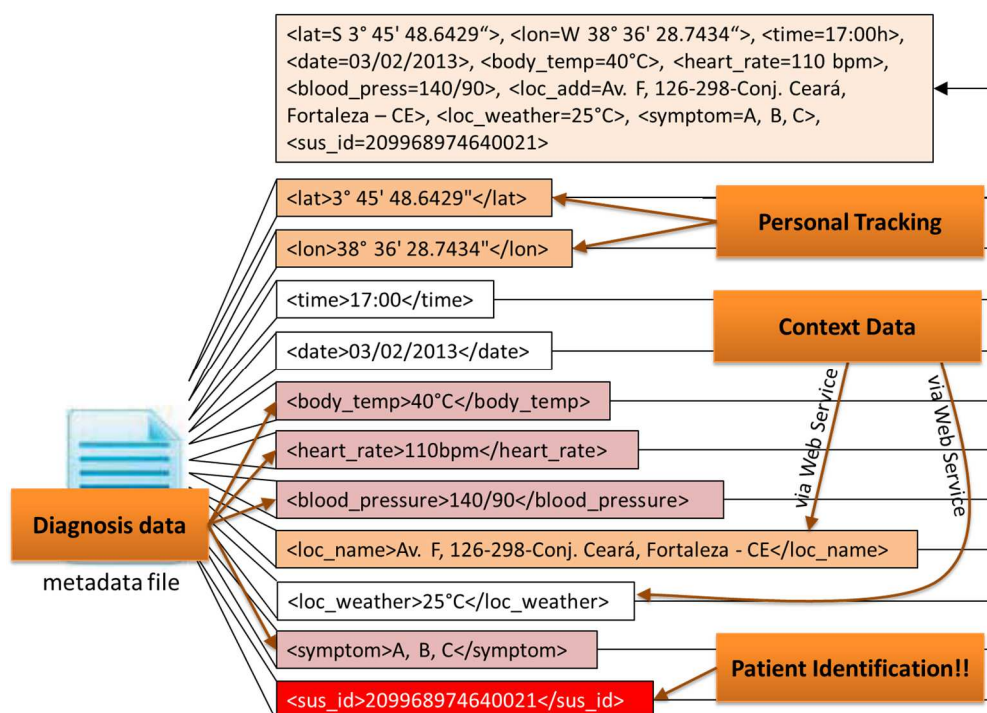


Figura 10 - Processamento dos Dados

Após a publicação de dados sobre o estado de saúde do paciente no banco de dados do LARIISA, este sistema é capaz de levantar perguntas adicionais para o usuário. Estas perguntas adicionais são levantadas pelo sistema com base nas informações inferidas que foram previamente analisadas pelo sistema utilizando conceitos de ontologia. Depois que um paciente responde às questões levantadas pelo LARIISA, conforme mostrado na figura 11, uma decisão é tomada para atendimento ou acionamento de alguma ação por parte do paciente, como por exemplo, solicitando que o mesmo se desloque para o posto de saúde mais próximo.

Em cenários epidemiológicos, o sistema faz inferências com base em diagnósticos similares vindos de localizações geográficas próximas umas das outras. Por exemplo, se existem muitos casos de dengue nos últimos dias e se mais de 90% dos casos vêm da mesma região geográfica (via GPS), então tem-se um cenário favorável de epidemia de dengue. Neste caso o sistema é capaz de direcionar um número razoável de agentes de saúde ou médicos

para aquela determinada região, ou também acionar a secretaria de saúde para compra de medicamentos que atendam aquela região específica, não causando assim indisponibilidade de remédios necessários para tratamento de uma epidemia de dengue.

A figura 11 mostra o cenário de aplicação do LARIISA onde podemos identificar:

- Lado esquerdo, na parte superior, temos a captura de dados do paciente;
- Parte superior no meio, temos a representação do envio dos dados com o uso do metadado;
- Lado direito, temos a regras de inferência do LARIISA;
- Parte inferior no meio, podemos visualizar as regras;
- Lado esquerdo, na parte inferior, vemos os cenários, envio de ambulância, envio de um agente de saúde.

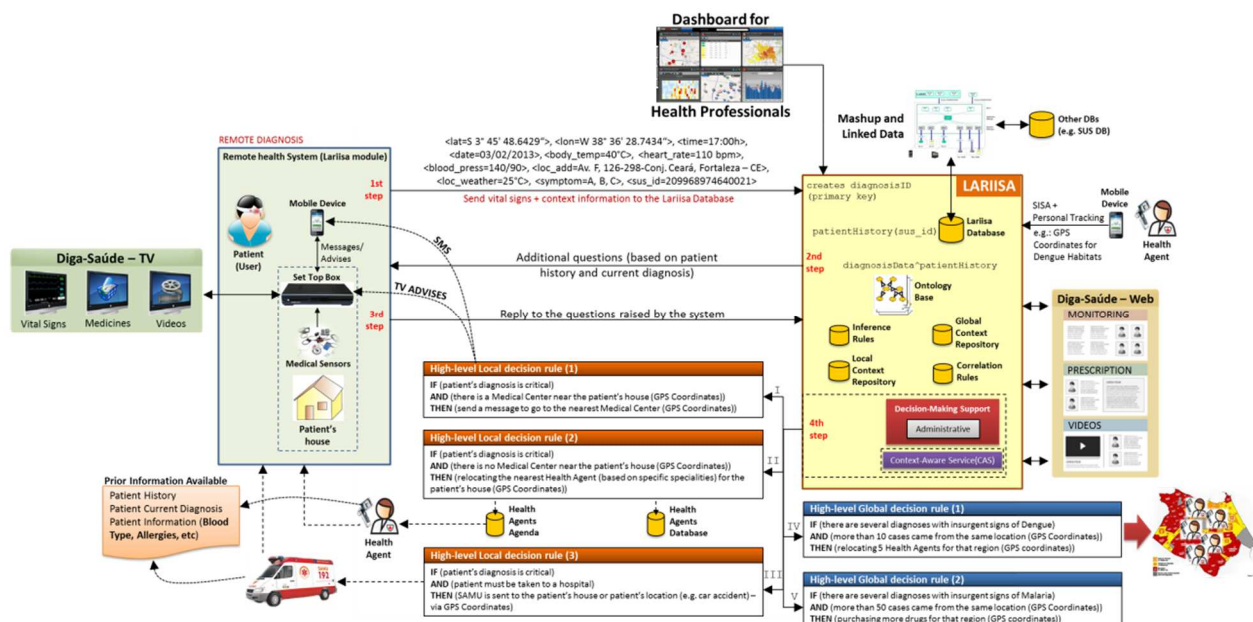


Figura 11 - Cenário do CLARIISA

O trabalho apresentado nesta dissertação tem o objetivo de tratar, dentro do contexto da figura 11, o mecanismo de inteligência do CLARIISA. Esta parte tem papel fundamental para o restante do sistema, uma vez que a falta de precisão desta parte comprometerá a funcionalidade de todas as outras partes do sistema. No capítulo 5 será apresentado o mecanismo de inteligência que auxilia as inferências realizadas pelo CLARIISA - aumentando sua precisão e eficiência na tomada de decisões por parte dos gestores de saúde.

4. TRABALHOS RELACIONADOS

Este capítulo apresenta um levantamento da literatura evidenciando trabalhos relacionados com Redes Bayesianas.

O uso de Redes Bayesianas para gerar sistemas da informação visando à estimativa de probabilidade tem sido cada vez mais utilizado. Bastos (2010) propõe um sistema de diagnóstico de perdas comerciais e técnicas nas distribuidoras de energia elétrica, baseado em Redes Bayesianas.

(FLORES e HÖHER, 2000), apresentam um sistema de apoio a decisão médica baseado na aplicação de redes probabilísticas (redes Bayesianas). A tecnologia utilizada pelos autores se mostra ideal para o tratamento de incerteza, muito comum na área médica e, além disso, modela o conhecimento do especialista do domínio de uma forma intuitiva. O sistema proposto destina-se ao diagnóstico médico, não prevê a utilização para redução dos índices de sinistralidade.

(MATOS, 2003), propõe o desenvolvimento de um sistema de apoio à decisão (SAD) para planejamento em saúde a partir dos indicadores de saúde padronizados no documento Indicadores e Dados Básicos para a Saúde (IDB, 2000), padronizadas pelo Ministério da Saúde. Trata-se do protótipo de um Sistema de Apoio à Decisão utilizando rede Bayesiana (RB) que tem como base o raciocínio probabilístico.

(KOEHLER, 2002), utiliza a abordagem bayesiana para tratamento da incerteza em sistemas inteligentes, uma alternativa para resolver o problema da incerteza em sistemas de apoio à decisão na área da Saúde. Modela um diagnóstico de Desnutrição Infantil baseado na idade, dados antropométricos e sinais clínicos.

(FLORES, PEROTTO e VICARI, 2002), apresentam um sistema que promove a melhoria na qualidade do diagnóstico médico, a redução do tempo necessário para realização do diagnóstico e a redução dos custos em exames sofisticados. Os autores apresentam as técnicas de representação do conhecimento incerto bem como a descrição de ferramentas computacionais desenvolvidas no projeto SEAMED/UFRGS.

(NUNES, 2013), esclarece uma abordagem bayesiana para epidemiologia com o uso de mapas de índices ou taxas de incidências o qual permite a análise da dispersão e dos padrões espaciais do risco de ocorrência de um determinado evento. Um grande problema associado ao uso de taxas brutas de ocorrência é a alta instabilidade que elas possuem para expressar o risco de eventos raros em regiões de população pequena. Alternativamente, são apresentadas as Taxas Bayesianas Empíricas, que utilizam informações de toda a região ou da vizinhança para estimar o risco de ocorrência do evento em cada área.

Observando o trabalho realizado pelo SISA, acrônimo de Sistema de Saúde adaptado ao contexto de Gestão de Saúde, traz o uso de dispositivos simples, como *smartphones*, entre o agente de endemia e o LARIISA.

O SISA trata de um projeto piloto de implementação de componentes, de Gestão do Conhecimento e de Clínica-Epidemiológica que estão presentes na arquitetura do projeto LARIISA. Este projeto piloto está sendo concebido com o propósito de agilizar o combate à dengue, bem como otimizar as medidas preventivas que possam contribuir com a diminuição de casos da doença.

4.1. Construção de uma Rede Bayesiana Aplicada ao Diagnóstico de Doenças Cardíacas

Foi visto no trabalho de Saheki (SAHEKI , 2005) a construção de um sistema especialista aplicado ao diagnóstico de doenças cardíacas, usando como ferramenta computacional redes Bayesianas. O trabalho envolveu a interação entre diferentes áreas do conhecimento, engenharia e medicina, com maior foco na metodologia da construção de sistemas especialistas. São apresentados os processos de definição do problema, modelagens qualitativa e quantitativa, e avaliação. Neste trabalho, os processos de modelagens e avaliação foram realizados com o auxílio de um especialista médico e de dados bibliográficos. São apresentados como resultados a rede Bayesiana construída e um *software* para manipulação de redes Bayesianas denominado iBNetz.

No trabalho de Saheki (SAHEKI , 2005) revela que desenvolvimento da estrutura de uma rede Bayesiana, com o auxílio de especialistas médicos, pode ser feito independentemente das probabilidades da rede, até que se chegue a uma rede cujos nós e arcos reflitam um consenso entre as pessoas envolvidas no trabalho.

4.2. Um estudo comparativo entre Naive Bayes e Noisy-or

A criação de um classificador é uma tarefa central para muitos problemas de aprendizagem de máquina. No artigo de David (CHEN, 2000) foram descritos dois classificadores de rede Bayesiana, os modelos Naive Bayes e as Noisy-or. Eles são de interesse particular por causa das suas estruturas simples. Onde foi comparado em duas dimensões: o poder expressivo e capacidade de aprender.

(CHEN, 2000) constata que Naive Bayes e Noisy-or têm expressividades equivalentes. Descrevendo em derivações matemáticas como transformar um modelo no outro modelo.

Esses modelos diferem em sua capacidade de aprender. O autor descreve um experimento confirmando a “intuição” de que Naive Bayes executa melhor do que o modelo Noisy-or, quando os dados se encaixam no conceito de independência. No entanto, ainda não tem um conjunto claro de critérios para determinar exatamente em que condições cada modelo se destaca.

Um estudo mais aprofundado dos pontos fortes e fracos de cada modelo onde se deve fornecer uma visão mais profunda sobre como melhorá-los. Uma proposta seria a de combinar o Naive Bayes e Noisy-or, de modo que a rede retratasse mais de perto a relação real entre todos os atributos.

4.3. Utilizando ontologia na inferência para apoio ao diagnóstico

A análise da problemática envolvendo o processo de diagnóstico, observam-se oportunidades de aprimorar os modelos correlatos, recentemente desenvolvidos, que procuram auxiliar esta tarefa intensiva em conhecimento. O artigo apresentado por (LOPES, GONÇALVES e TODESCO, 2011) apresentam trabalhos analisados e fundamentados em ontologias, mostrando a eficiência na manipulação das incertezas envolvidas no processo de diagnóstico aplicando técnicas apropriadas (lógica *fuzzy*, redes neurais, raciocínio baseado em casos) e verificou-se a deficiência na geração de novos conhecimentos, a fim de atualizar a base de conhecimento.

O problema enfrentado foi como trabalhar de forma adequada com as informações e dados obtidos pelos diagnósticos realizados, de forma que novos conhecimentos pudessem ser gerados e contribuíssem ao aprimoramento do diagnóstico.

O Cálculo Probabilístico, no trato às incertezas da base de conhecimento, atua de forma complementar ao processo de diagnóstico, fornecendo diagnósticos probabilísticos bem como, destaca informações relevantes para a geração de possíveis novos conhecimentos. O diagnóstico, na forma probabilística, parece ser a maneira mais adequada de apresentar resultados em domínios onde a subjetividade e incertezas fazem parte do processo.

Assim, tem-se um modelo que emprega metodologia, métodos e técnicas disponíveis e possibilita, através da integração entre eles, construir um Sistema de Conhecimento que minimiza os riscos de implantação e, com sua utilização, fornece diagnósticos probabilísticos bem como, propõe possíveis novos conhecimentos para apoio ao diagnóstico (LOPES, GONÇALVES e TODESCO, 2011).

4.4. Sistemas Especialistas Probabilísticos

Os sistemas especialistas são desenvolvidos para resolver uma grande variedade de problemas, sendo que o conhecimento necessário para solucionar estes problemas, geralmente, não é precisamente definido. Podemos constatar que muitos problemas na vida real são repletos de incerteza. Especialistas humanos nestes domínios são capazes de formar julgamentos e tomar decisões baseados em informação incerta, incompleta e, algumas vezes, até mesmo contraditória.

Para que um sistema especialista seja confiável, o mesmo deve lidar com este tipo de informação com a mesma facilidade que o especialista humano. O desenvolvimento de formalismos para tratar incerteza e, os algoritmos, para manipular informação com incerteza tornaram-se um tópico maior de pesquisa em IA chamado Raciocínio com Incerteza.

A partir da metade da década de 80, a pesquisa sobre raciocínio probabilístico em sistemas especialistas resultou na introdução de Redes de Crença Bayesianas. Estas redes têm sua fundamentação na teoria da probabilidade e são caracterizadas por um poderoso formalismo que representa o conhecimento no domínio e as incertezas por aleatoriedade associadas a este domínio. Mais especificamente, o formalismo proporciona uma representação concisa de uma distribuição conjunta de probabilidades considerando um grupo de variáveis estocásticas.

Associados a este formalismo estão os algoritmos para calcular eficientemente as probabilidades associadas a eventos no domínio de aplicação e para processar as evidências (*input*) relevantes da base de conhecimentos; estes algoritmos constituem as partes básicas para o raciocínio com o conhecimento assim representado. Desde sua introdução, a estrutura de redes Bayesianas vem rapidamente ganhando popularidade e agora está começando a mostrar o seu valor em domínios complexos. Aplicações práticas estão sendo desenvolvidas, por exemplo, para diagnóstico e prognóstico médico e para recuperação probabilística de informação (*data mining*).

Como a Teoria da Probabilidade é uma teoria matematicamente bem fundamentada, possuindo uma longa tradição de pesquisa e experiência, não é surpreendente que esta teoria conquiste lugar proeminente na pesquisa de raciocínio com incerteza em sistemas especialistas. Porém, aplicar esta teoria no contexto de um sistema especialista não é tão fácil quanto possa parecer à primeira vista. A aplicação direta dos conceitos básicos da teoria da probabilidade leva a problemas de complexidade computacional, isto é, a representação explícita de uma distribuição conjunta de probabilidade requer espaço exponencialmente proporcional ao número de variáveis consideradas na rede Bayesiana.

4.5. Exemplos de Sistemas Desenvolvidos

Com o objetivo de mostrar algumas aplicações desenvolvidas utilizando a abordagem bayesiana a seguir são comentados alguns exemplos apresentados na tabela 5 (PENZ , 2011).

- **Uma Abordagem Híbrida Fuzzy-Bayesiana para Modelagem de Incertezas.**

Nesta pesquisa investigou-se a modelagem conjunta de dois tipos de incertezas: a aleatória e a difusa em sistemas especialistas. Dentre as diferentes abordagens encontram-se as Redes Bayesianas que são utilizadas para tratar a incerteza por aleatoriedade (probabilidade) e os Sistemas Especialistas Fuzzy os quais tratam a incerteza por imprecisão (possibilidade). No entanto, em alguns domínios do mundo real, coexistem esses dois tipos de incerteza: a imprecisão e a aleatoriedade. Assim, propõe-se uma abordagem híbrida Fuzzy-Bayesiana, para modelar incertezas aleatórias e imprecisas em sistemas especialistas, baseando-se na união da Teoria dos Conjuntos Fuzzy e da Teoria da Probabilidade. O hibridismo Fuzzy-Bayes permite trabalhar com a incerteza por imprecisão nas entradas de uma rede Bayesiana, formada por nodos dicotômicos ou não dicotômicos. As variáveis com características de imprecisão são “fuzzificadas”, e propagadas utilizando o Teorema de Bayes modificado, gerando uma redistribuição do vetor de probabilidades da saída da rede híbrida. Esta redistribuição ocorre devido ao reconhecimento da incerteza presente nos dados de entrada, encontra-se fundamentada matematicamente e foi avaliada considerando o problema de distúrbio nutricional (TIBIRIÇÁ, 2005).

| Nome | Local | Ano | Autor |
|--|--|------|--------------------------|
| Procedimentos para Prover Confiabilidade ao Uso de Inteligência Artificial em Ensaio de Desempenho de Compressores Herméticos de Refrigeração | Depto de Engenharia Mecânica da UFSC | 2011 | Cesar Alberto Penz |
| Abordagem Bayesiana para Redes sem Fio Cognitivas | Depto de Informática e de Estatística da UFSC | 2010 | Rafael Mendes |
| Uma Abordagem Híbrida Fuzzy-Bayesiana para Modelagem de Incertezas | Depto de Informática e de Estatística da UFSC | 2005 | Carlos Augusto Tibirica |
| ABS-Bayes: Sistema Especialista Probabilístico para Apoio ao Diagnóstico de Absenteísmo | Depto de Informática e de Estatística da UFSC | 2003 | Nelso Antônio Losi |
| Modelagem Multiagente Escalonável para Integração de Sistemas Especialistas - CLIDENP | Depto de Informática e de Estatística da UFSC | 2003 | Michel Barnad Geszychter |
| Agente Bayesiano de Apoio à Vigilância de Infecção Hospitalar - SAVIH | Depto de Informática e de Estatística da UFSC | 2002 | Marcos Venicius |
| Sistema especialista probabilístico para apoio ao diagnóstico de potencial econômico – SEPE | Depto de Informática e de Estatística da UFSC | 2002 | Hamilton Pasini |
| Sistema de Avaliação da Evasão Escolar: E-Bayes | Depto de Informática e de Estatística da UFSC | 2002 | Eugênio Rovaris Neto |
| Sistema Especialista de Apoio à Nutrição Enteral Pediátrica - SANEP | Depto de Informática e de Estatística da UFSC | 2001 | Lisiane de Albuquerque |
| Sistema de Avaliação do Crescimento Infantil - SACI | Depto de Informática e de Estatística da UFSC | 2001 | Priscyla W. T.A. Simões |
| Modelo Híbrido Difuso Probabilístico | Depto de Informática e de Estatística da UFSC | 2001 | Juliano Brignolli |
| Sistema Especialista Probabilístico: base de conhecimento dinâmica-SEDIN | Depto de Informática e de Estatística da UFSC | 2000 | Carlos Stein |
| Baseline Bayesiana para Gerência de Redes de Computadores | Depto de Informática e de Estatística da UFSC | 2000 | Cleverson Veronez |
| Sistema de Apoio ao Diagnóstico Diferencial de Cefaléia | Depto de Informática da PUC, Curitiba, Paraná | 2000 | Renata V. A. Toledo |
| SISPAN | Depto de Informática e de Estatística da UFSC | 1998 | Cristiane Koehler |
| Lógica Autônoma de Controle - Lockheed Martin | <i>Lockheed Martin Marine Systems and Artificial Intelligence Center - Sunnyvale, Califórnia</i> | 1998 | ---- |
| Microsoft Pregnancy and Child Care | <i>Microsoft e Knowledge Industries (KI)</i> | 1998 | ---- |
| DIAVAL | Depto de Informática e Automática - U.N.E.D. - Madrid - Espanha | 1994 | Fernandez J. Díez |
| General Electric Aircraft Engines Services | <i>Knowledge Industries (KI) , ENTER Software, Califórnia</i> | 1989 | ---- |
| DXplain | Laboratório de Ciência da Computação (Hospital Geral de Massachusetts) | 1987 | ---- |

Tabela 5 - Exemplos de Sistemas Especialistas Probabilísticos

- **Procedimentos para Prover Confiabilidade ao Uso de Inteligência Artificial em Ensaios de Desempenho de Compressores Herméticos de Refrigeração.**

Nesta tese de doutorado foi investigado o uso de técnicas de IA para auxiliar a previsão de resultados de ensaios em compressores contribuindo para a diminuição do tempo de avaliação. Foram utilizadas Redes Neurais e Redes Fuzzy-Bayesianas. O uso de diferentes técnicas de IA provê confiabilidade aos resultados, consequentemente levando à maior aceitação pelas comunidades técnico-científica e industrial (PENZ, 2011).

- **Abordagem Bayesiana para Redes sem Fio Cognitivas.**

Nesta pesquisa foi investigado o uso de Redes Bayesianas e Teoria da Decisão para a autoconfiguração e reconfiguração dos equipamentos de rede sem fio. Foram propostas duas Redes Bayesianas, a primeira dedicada ao sensoriamento do espectro e a outra dedicada à caracterização e decisão no espectro com o objetivo de, respectivamente, aumentar a acurácia na detecção de usuários primários e gerenciar a conectividade da rede entre usuários primários, secundários, proprietários dos equipamentos e secundários não proprietários. Desta forma, também busca fornecer bases tecnológicas para o compartilhamento altruísta da conectividade à *Internet*, visando sua aplicação com fins sociais (MENDES, 2010).

- **Uma Abordagem Híbrida Fuzzy-Bayesiana para Modelagem de Incertezas**

Esta pesquisa aprofundou a investigação da Lógica Fuzzy quando aplicada a nodos de uma Rede Bayesiana que apresentam transição gradual entre suas classes (TIBIRIÇÁ, 2005).

- **ABS-Bayes: Sistema Especialista Probabilístico para Apoio ao Diagnóstico de Absenteísmo.**

Este sistema busca oferecer apoio a área de recursos humanos de uma empresa detectando os principais problemas de absenteísmo por doenças e suas principais características. Gera dinamicamente uma base de conhecimentos por estar ligado ao banco de dados de exames periódicos dos funcionários e ao banco de dados de atestados médicos apresentados por funcionários que faltam o trabalho. O usuário pode configurar as hipóteses diagnósticas ou deixar que o próprio sistema se encarregue de configurar com as doenças mais prevalentes na empresa (LOSI, 2003).

- **Modelagem Multiagente escalonável para Integração de Sistemas Especialistas.**

Foi desenvolvido um protótipo de um sistema Multiagente – CLIDENP – Clínica Virtual para Diagnóstico e Educação em Nutrição Pediátrica, que integrou os sistemas

especialistas probabilísticos que apoiam o diagnóstico de desnutrição infantil: SISPAN, SACI e SANEP. O sistema CLIDENP foi desenvolvido numa estrutura cliente-servidor. A integração desses sistemas na forma de um sistema Multiagente teve como objetivo principal a manutenção da escalabilidade, assim podendo ampliar globalmente a base de conhecimentos pela inserção de novos agentes especialistas (GESZYCHTER, 2003).

- **Sistema especialista probabilístico para apoio ao diagnóstico de potencial econômico – SEPE**

Este sistema foi desenvolvido para apoiar o planejamento mercadológico a partir do conhecimento do potencial econômico dos municípios de Santa Catarina. Na sua elaboração trabalha-se com dados incertos sobre o mercado, ou seja, tomam-se decisões sem ter as informações necessárias. Para que o diagnóstico seja realizado de uma forma mais eficiente, é necessário que os técnicos visualizem os dados relacionados ao potencial de uma forma gráfica e que sejam investigadas relações entre variáveis endógenas e exógenas aos Correios, por meio de evidências observadas. O sistema SEPE explora o raciocínio probabilístico em sistemas especialistas, com uma base de conhecimentos atualizável. Ainda, emprega técnicas de mineração de dados no processo de aquisição e atualização de conhecimento de uma base de dados. Foi desenvolvido para a plataforma *Windows*, no ambiente *Visual Basic 5*, utilizando as DLL's da shell Nética (PASINI, 2002).

- **Agente Bayesiano de Apoio à Vigilância de Infecção Hospitalar – SAVIH**

Este sistema utiliza uma rede bayesiana para apoiar o gestor de um hospital a avaliar o risco de um paciente à infecção hospitalar. Mostra o cenário de infecção nas diversas unidades de internação permitindo fazer a previsão de risco a partir das características epidemiológicas e de doença do paciente. Recupera casos armazenados que forem similares ao paciente ingressante na unidade. Foi desenvolvido utilizando a shell Nética (VENICIUS, 2002).

- **Sistema de Avaliação da Evasão escolar: E-Bayes.**

Este sistema utiliza uma rede bayesiana para apoiar o gestor de uma universidade a avaliar a evasão escolar de um curso universitário. Mostra o cenário da evasão discente permitindo fazer a previsão da permanência de um aluno no curso a partir de suas características sócio-demográficas. Busca na base de dados quais os alunos do curso que são similares às características do aluno em questão. Explora a aplicação de

sistemas especialistas em sistemas de informação. Foi desenvolvido utilizando a shell Nética (ROVARIS , 2002).

- **Sistema Especialista de Apoio à Nutrição Enteral Pediátrica – SANEP**

Este sistema identifica a dieta adequada a ser administrada a uma criança desnutrida e calcula toda a fórmula nutricional a partir dos diferentes tipos de leites disponíveis na farmácia de um serviço de atendimento pediátrico. Por questões de segurança, mantém numa base de dados todas as informações de entrada e de saída de um caso consultado (ALBUQUERQUE, 2001).

- **Sistema de Avaliação de Crescimento Infantil – SACI.**

Esta pesquisa volta-se ao desenvolvimento de um sistema especialista probabilístico para o diagnóstico de alguns distúrbios associados à falência do crescimento em crianças com até 2 anos. O sistema possui uma rede bayesiana composta de um nodo para representar as hipóteses diagnósticas (Deficiência de Ferro, Má Absorção e Etiologia a Esclarecer), e outros três nodos para representar as evidências necessárias para chegar às hipóteses diagnósticas (Diarreia, Taxa de Ferritina e D-Xilosemia). O sistema ainda oferece o acompanhamento do crescimento da criança graficamente, pois para que o diagnóstico seja realizado de uma forma mais eficiente, é necessário que os médicos visualizem os dados relacionados à evolução do crescimento de uma forma gráfica e que sejam investigadas suspeitas de perda de peso e altura da criança, por meio de evidências observadas (SIMÕES, 2001).

- **Modelo Híbrido Difuso Probabilístico**

Nesta pesquisa foi explorada a modelagem de duas formas de incerteza: por aleatoriedade e por imprecisão, utilizando-se a teoria da probabilidade e a teoria dos conjuntos difusos (fuzzy sets). Foi desenvolvida uma abordagem para redes bayesianas considerando apenas variáveis dicotômicas. Aplicou-se esta abordagem ao problema de diagnosticar danos à saúde considerando a presença de antibióticos no leite, primeiramente foi implementada uma rede bayesiana e, após, foi realizado um ajuste nas probabilidades do diagnóstico utilizando graus de pertinência relativos à presença de antibióticos (BRIGNOLLI, 2001).

- **Sistema Especialista Probabilístico: base de conhecimento dinâmica.**

Nesta pesquisa foi tomada a base de conhecimentos do sistema SISPAN (KÖEHLER, 1998) e foi desenvolvido um sistema que permite que um usuário possa: incluir/excluir nodos e alterar valores de probabilidade da rede bayesiana. Dessa

forma, pode ser avaliada a acurácia de um sistema especialista probabilístico e pode ser utilizado como um ambiente de ensino por simulação (STEIN, 2000).

- **Baseline Bayesiana para Gerência de Redes de Computadores.**

Este sistema é um módulo inteligente que utiliza informações de uma MIB (*management information base*) para avaliar dinamicamente o tráfego em certo trecho de uma rede de computadores. Oferece como saída às probabilidades das hipóteses, diagnósticas do tráfego na rede estar: normal, leve, moderado e intenso a partir dos dados iniciais de hora, dia fornecidos pelo administrador e as informações buscadas na MIB. Foi desenvolvido utilizando a *shell* Nética (VERONEZ, 2000).

- **Sistema de Apoio ao Diagnóstico Diferencial de Cefaleia**

É um sistema especialista médico que oferece apoio a médicos generalistas, plantonistas ou residentes no diagnóstico diferencial de cefaleias (dores de cabeça). Este sistema utiliza o enfoque de rede bayesiana. A base de conhecimentos foi construída considerando os Critérios de Classificação da Sociedade Internacional de Cefaleias (IHS), levando em conta os sinais e sintomas dos pacientes e nos valores estimados de probabilidades fornecidos pelos especialistas que participaram do projeto. Foi desenvolvido utilizando a *shell* Nética. A avaliação deste sistema foi realizada comparando as respostas fornecidas pelos especialistas e as respostas fornecidas pelo sistema, a partir de um conjunto de fichas médicas de pacientes com cefaleia selecionados aleatoriamente pelos especialistas do projeto. Os resultados experimentais indicam que o sistema foi capaz de fornecer os mesmos diagnósticos que os especialistas do projeto em 95% dos casos. Por outro lado, os mesmos casos clínicos foram avaliados por médicos generalistas e verificou-se que estes obtiveram um percentual de acerto de 53%. Portanto, o sistema desenvolvido apresenta um desempenho muito bom ao realizar diagnóstico diferencial de cefaleias (TOLEDO, 2000).

- **SISPAN - Sistema Especialista Probabilístico de Apoio a Avaliação Nutricional.**

Este sistema realiza a avaliação nutricional de crianças com até dois anos de idade. Considera como *inputs* os dados relativos a: peso, altura, idade, cabelo, pele, unha, hipotrofia e edema. Classifica o estado nutricional da criança em: sem desnutrição ou com desnutrição, podendo ser leve, moderada ou severa quanto à intensidade, e marasmo, kwashikor ou marasmo-kwashikor quanto ao tipo. Foi desenvolvido utilizando a *shell* Nética, com o módulo de interface com o usuário desenvolvido em C++ (KÖEHLER, 1998).

- **DIAVAL.**

É um sistema especialista destinado a ajudar médicos no diagnóstico de doenças cardíacas a partir, principalmente, de dados ecocardiográficos. Este *software* faz a introdução dos dados, apresentação dos resultados do diagnóstico e explicações sobre o raciocínio, destacando as características que diferenciam o DIAVAL de outros sistemas especialistas e de outras redes bayesianas. Este *software* é o resultado da tese de doutorado do Prof. Fernandez J. Díez e tem 2 objetivos principais: primeiro, contribuir para o estudo de redes bayesianas e, segundo, aplicar redes bayesianas na construção de um sistema especialista para diagnóstico de ecocardiografia, chamado DIAVAL. Este sistema consiste em calcular a probabilidade “a posteriori”, selecionar a mais provável e a doença mais relevante, gerando um relatório impresso. Outras informações podem ser obtidas no endereço: <http://www.dia.uned.es/~fjdiez/papers/diaval.html>

- **Dxplain.**

É um sistema de apoio a decisão que utiliza um conjunto de dados clínicos (sinais, sintomas, dados laboratoriais) para produzir uma lista de diagnósticos associados às manifestações clínicas. Este *software* possui uma grande base de dados sobre 5000 manifestações clínicas associadas com mais de 2000 diferentes doenças. O Laboratório de Ciência da Computação (do Hospital Geral de Massachusetts) vem desenvolvendo DXplain há dez anos; o sistema tem sido utilizado por milhares de médicos e estudantes de medicina. O banco de dados e o sistema estão sendo continuamente aperfeiçoados e atualizados. DXplain vem sendo utilizado em vários hospitais e escolas de medicina para educação clínica e como um auxílio educacional na solução de problemas clínicos. DXplain tem as características de livro-texto eletrônico e um manual de referência médica. Como um livro-texto médico, DXplain pode oferecer uma descrição de mais de 2000 doenças diferentes, com ênfase nos sinais e sintomas que ocorrem em cada doença, a etiologia, a patologia e o diagnóstico. DXplain também oferece mais de 10 referências para cada doença, dando ênfase para revisões clínicas onde estas estão disponíveis. Outras informações podem ser obtidas no endereço: <http://www.medlib.iupui.edu/dxplain.html>.

- **General Electric Aircraft Engines Services.**

O diagnóstico de falhas em turbinas de aviões em grandes empresas aéreas implica em grandes custos, pois os dados retirados da análise possuem muito ruído e não podem atuar para um defeito específico. As solicitações mecânicas para efetuar os testes e

inspeções, assim como a excessiva verificação do motor podem duplicar os custos de manutenção. Para um 757 típico, este custo pode chegar a 500.000 dólares. Contratadas pela General Electric, a *Knowledge Industries* (KI) e a *ENTER Software*, de Menlo Park, CA, uma empresa líder em análise de desempenho em tempo real de turbinas a gás para geradoras de energia elétrica, desenvolveram modelos de redes bayesianas que guiam um mecânico de motores através de uma sequência otimizada de testes e passos de inspeção. O sistema de recomendação de testes utiliza os resultados de um modelo de desempenho termodinâmico do motor da aeronave, que processa os dados vindos de sensores montados sobre o motor no estande de teste. Mais informações podem ser conseguidas na home-page da KI, cujo endereço é: <http://www.kic.com>.

- **Lógica Autônoma de Controle**

A *Lockheed Martin Marine Systems*, em *Sunnyvale*, Califórnia, em conjunto com o Artificial Intelligence Center em Palo Alto, vem desenvolvendo um sistema de controle para o UUV (*Unmanned Underwater Vehicle*). Este veículo submarino poderá permanecer semanas, submerso, sem necessidade de intervenção humana. O objetivo da pesquisa é desenvolver um *software* que seja capaz de controlar eventos planejados e eventos imprevistos, de modo a diminuir o risco de perda do veículo e maximizar a probabilidade de sucesso da missão. O sistema possui maior capacidade de manusear a grande gama de anomalias e alarmes falsos que ocorrem em missões de longa duração, tornando-o bastante superior aos sistemas atuais. Quando completamente desenvolvido, este sistema permitirá ao UUV monitorar o progresso de sua missão, analisar o estado de seu equipamento, detectar e analisar eventos que causam impacto nos objetivos da missão, tomar decisões e efetuar ações para compensar os eventos, além de modificar o plano da missão quando o atual estiver completamente fora de alcance. Mais informações podem ser obtidas no site da Hugin: www.hugin.dk.

- **Microsoft Pregnancy and Child Care.**

A *Microsoft* e a *Knowledge Industries* (KI) uniram-se para criar um serviço *on-line* de informação sobre saúde, com capacidade de multimídia, que emprega raciocínio bayesiano. O produto é chamado de *Microsoft Pregnancy and Child Care*, e está disponível no endereço da Microsoft Network. Redes bayesianas são o componente central da opção *Find by Symptom* (Busca por Sintoma). Sintomas que ocorrem frequentemente são conectados a módulos especialistas que guiam os pais da criança através de uma série de questões. O sistema repetidamente determina qual a melhor

pergunta a se fazer aos pais, mostrando as apresentações em multimídia que correspondem às exatas condições de saúde da criança. Os pais podem escolher entre artigos ilustrados para obter informação detalhada sobre os sintomas e a doença da criança. Os artigos também contêm *links* para outras fontes de informação relacionadas, incluindo outros *Web sites*. O Microsoft *Pregnancy and Child Care* foi criado pela Divisão de Consumo da Microsoft. A KI desenvolveu e testou as bases de conhecimentos em redes bayesianas, trabalhando com os especialistas Dr. Richard Behrman, Professor de Pediatria na *Stanford University* e na *University of California*, San Francisco, e Dr. David Kerns, Chefe de Pediatria no *Santa Clara Valley Medical Center*. Para maiores informações, consulte www.kic.com

5. CRIAÇÃO DE REDES BAYESIANAS NO PROJETO LARIISA

5.1. Tratando a incerteza de dados no LARIISA

5.1.1. Incertezas, Ontologia e Contexto

A ideia central do projeto LARIISA consiste em inferir sobre os dados coletados do paciente e os existentes no seu prontuário, identificado pelo SUS_ID, a partir do conhecimento de especialistas da área de saúde representados no sistema.

Estas inferências podem apoiar a tomada de decisão de gestores (governador, prefeitos, diretor de hospital, médicos, agentes de saúde, etc.) em cinco domínios da saúde: Gestão de Conhecimento, Normativo Sistêmico, Clínico/Epidemiológico, Administrativo e Gerenciamento Compartilhado, (ANDRADE, 2009).

O LARIISA faz uso de ontologias (VIANA, 2007) na representação do conhecimento e de informações de contexto (*context-aware concept*) (DEY, SALBER e ABOWD, 2001). Como mostrado no capítulo três, as figuras 7 e 8 apresentam as ontologias *Global_Health_Context* e *Local_Health_Context*. A agregação de informações de georreferenciamento à *template* do metadado (figura 10), além de outros dados do usuário (SUS_ID, biométricos, etc.) enviado ao “core” do LARIISA teve como resultado uma versão mais bem elaborada do projeto, denominada CLARIISA (figura 11) (GARDINI, BRAGA, *et al.*, 2013).

Neste contexto evolutivo do projeto LARIISA, este trabalho preocupa-se com o tratamento da incerteza da informação. Portanto, este trabalho se relaciona tanto com a representação da informação sensível ao contexto (dados coletados) quanto ao conhecimento de especialistas, representação no LARIISA por ontologia. Portanto, incerteza na aquisição de dados e/ou na representação do conhecimento ainda não tinha sido tratada pelo LARIISA.

Este trabalho descreve o LARIISA_Bay, um componente baseado em redes bayesianas que permite a captura de informações de contexto e produz inferências que auxiliam especialistas a melhor diagnosticar casos de dengue a partir de dados coletados (usuários, agentes de saúde e especialistas) no sistema e de uma tabela de probabilidades sobre a ocorrência de sintomas, construída por profissionais da área. A figura 12 mostra o modelo preliminar do LARIISA_Bay.

5.1.2.LARIISA_Bay, um componente baseado em redes Bayesianas

Conforme pode ser verificado na figura 12, o componente CLARIISA_Bay é constituído das seguintes funcionalidades:

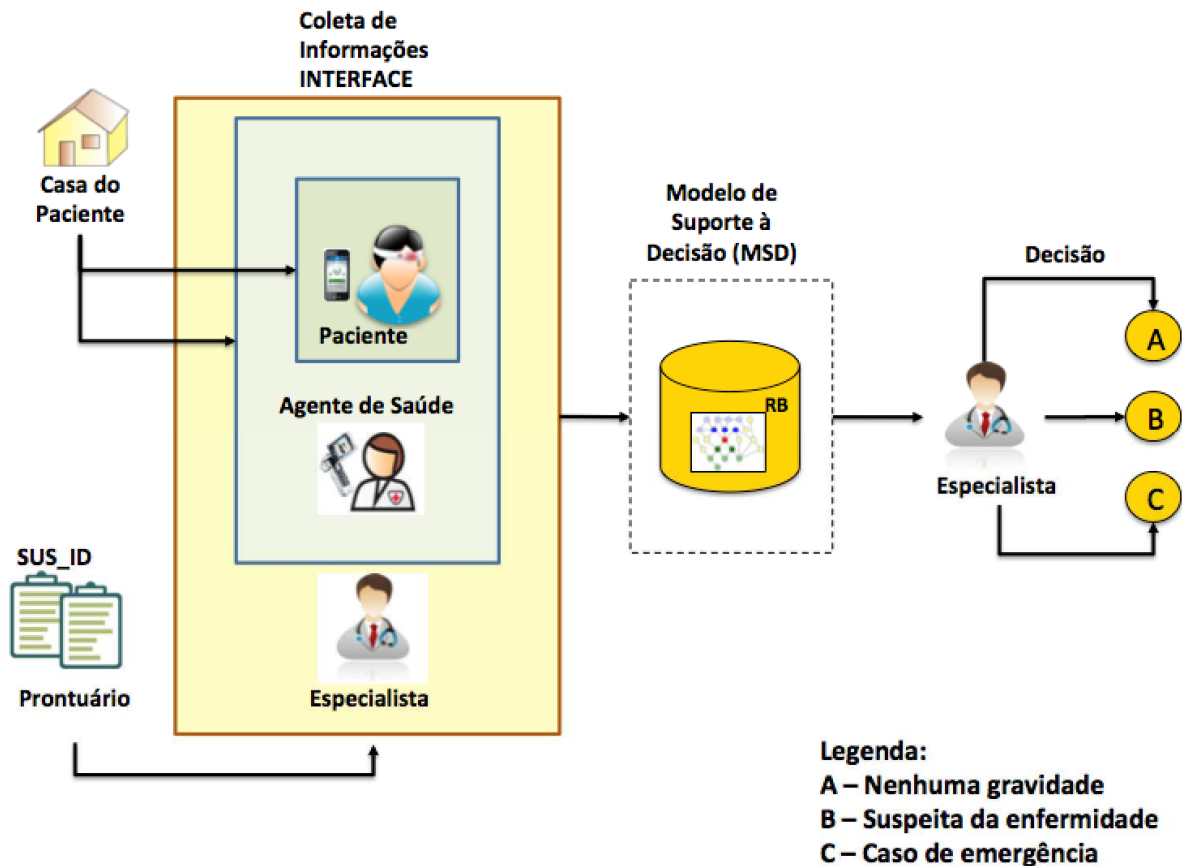


Figura 12 – Modelo preliminar do LARIISA-Bay

- *Interfaces* com três tomadores de decisão do CLARIISA: paciente, agente de saúde e especialista. Percebe-se que a *Interface* Especialista contém a *interface* Agente de Saúde, que contém a *interface* usuário;
- O Modelo de Suporte à Decisão (MSD) apoio na tomada de decisão de uma equipe médica. O MSD baseado em redes bayesianas tem duas finalidades:
 - Apoiar o diagnóstico da equipe médica, filtrando casos prováveis de dengue em três níveis de classificação:
 - Paciente normal (nenhuma gravidade)
 - Paciente com suspeita de dengue
 - Paciente com suspeita de dengue hemorrágica

- Apoiar o diagnóstico de epidemias em determinadas regiões de um município
- Uma equipe de saúde com especialistas capazes de melhor avaliar as seleções feitas pelo MSD. Compete a esta equipe o diagnóstico que resultará em um dos três procedimentos a seguir:
 - (A) Enviar mensagem de orientação ao usuário
 - (B) Enviar um agente de saúde prioritariamente à casa do usuário
 - (C) Enviar uma equipe de urgência à casa do usuário

5.1.3. Interfaces do LARIISA_Bay

Como já observado, as *interfaces* do LARIISA_Bay tem complexidade crescente, sendo a *interface* do paciente um subconjunto da *interface* do agente de saúde. Esta última, por sua vez, é um subconjunto da *interface* do médico:

- A *interface* PACIENTE contém os seguintes campos:
 - Identificação única do usuário (SUS_ID), fornecida manualmente pelo usuário;
 - Texto informado manualmente pelo usuário sobre sintomas (140 caracteres)
 - Dados biométricos complementares, também informados manualmente pelo usuário (idade, peso, altura, etc);
 - Dados coletados via sensores (temperatura, batimento cardíaco, pressão sanguínea, etc.)
 - Dados obtidos via *webservice* (latitude, longitude, data, hora, endereço aproximado de localização);
- A *interface* AGENTE DE SAÚDE contém os seguintes campos, além dos acima citados:
 - **SIM/NÃO** - Apresenta algum tipo de sangramento, vias respiratórias e/ou sistema digestivo?
 - **SIM/NÃO** - Está com dificuldade de se manter alerta apresentando sonolência e/ou irritabilidade?
 - **SIM/NÃO** - Reside em local sem saneamento básico?
 - **SIM/NÃO** - Localidade possui coleta de lixo regular?

- A *interface* ESPECIALISTA contém os seguintes campos, além dos acima citados
 - **SIM/NÃO** - Paciente apresenta queda abrupta de plaquetas
 - **SIM/NÃO** - O paciente apresenta um aumento do tamanho do fígado (Hepatomegalia dolorosa)
 - **SIM/NÃO** - O paciente apresenta dificuldade de manter a temperatura corpórea
 - **SIM/NÃO** - O exame do hemograma encontra-se alterado?
 - **SIM/NÃO** - O paciente não consegue ficar em ortostase?
 - **SIM/NÃO** - Existe algum desequilíbrio proporcional de água e células sanguíneas (hematócrito)?
 - **SIM/NÃO** - Apresenta diminuição da diurese?
 - **SIM/NÃO** - O exame de sorologia encontra-se alterado?

Uma vez que as informações do usuário são coletadas pelas *interfaces* USUÁRIO e AGENTE DE SAÚDE, elas são disponibilizadas ao ESPECIALISTA, o qual pode acessar o prontuário eletrônico do usuário a partir do SUS_ID, contido no metadado enviado pelo usuário ao sistema. Todas estas informações coletadas (figura 11) servem de entrada ao MSD, o bloco de apoio à decisão baseado em redes bayesianas desenvolvido neste trabalho.

A partir das decisões obtidas, estas serviriam de entrada para detecção e monitoramento de epidemias em diversas regiões, como mostrado na figura 13. Esta figura mostra o modelo de informação do LARISSA_Bay

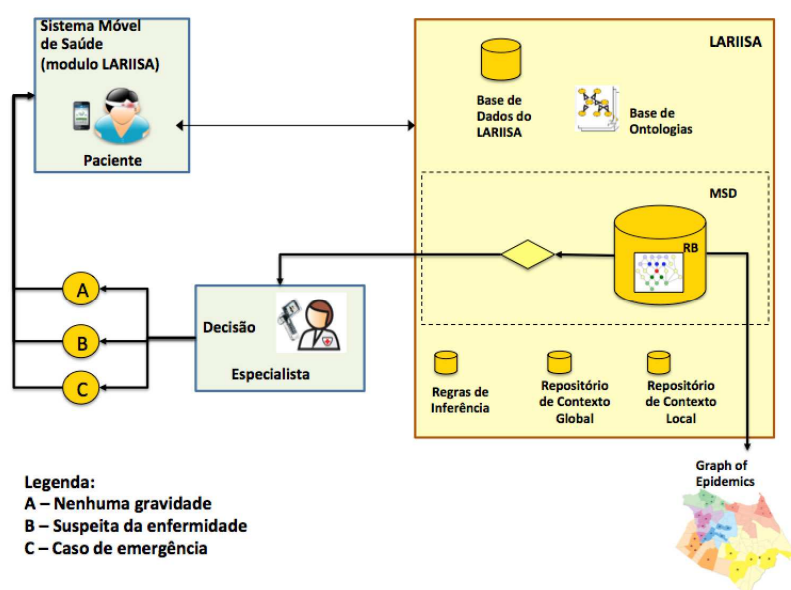


Figura 13 – Modelo de informação do LARISSA_Bay

5.2. Criação do modelo da rede Bayesiana do LARIISA_Bay

Informações relativas ao saneamento básico e as patologias em casos de dengue foram essenciais na construção da rede bayesiana utilizada neste trabalho, destinado ao auxílio à tomada de decisão no projeto LARIISA. Dessa forma, juntamente com a obtenção de dados bibliográficos, foram realizadas reuniões com a participação de um grupo de médicos especialistas, que auxiliaram na construção da estrutura da rede e sua posterior quantificação. Foram discutidos aspectos do domínio médico, com o intuito de estabelecer um vocabulário comum entre todos os envolvidos.

A obtenção do modelo estrutural da rede, usando-se o conhecimento de médicos e fontes bibliográficas, buscou identificar quais informações relativas ao problema médico poderiam ser representadas como variáveis na rede, assim como as relações causais entre estas variáveis. Para a realização desta tarefa foram obtidas informações médicas específicas ao problema que, posteriormente, foram colocadas na versão inicial da estrutura da rede. O refinamento iterativo da estrutura inicial, que envolveu a inserção e remoção de nós e arcos, entre outras modificações, levou à estrutura final para quantificação.

5.2.1. Sintomas patológicos observados

Usando-se o conhecimento de médicos e fontes bibliográficas, considera-se caso suspeito de dengue todo paciente que apresente doença febril aguda, com duração máxima de sete dias, acompanhada de pelo menos dois dos sinais ou sintomas como cefaleia⁴, dor retro-orbitária, mialgia⁵, artralgia⁶, prostração⁷ ou exantema⁸, associados ou não à presença de sangramentos ou hemorragias, com história epidemiológica positiva, tendo estado nos últimos 15 dias em área com transmissão de dengue ou que tenha a presença do *Aedes aegypti* (MINISTÉRIO DA SAÚDE, 2012).

Outros sintomas patológicos observados são:

- Dor abdominal intensa e contínua

O paciente não suporta palpação abdominal superficial, podendo impedir os

⁴ Termo médico para dor de cabeça

⁵ Termo utilizado para caracterizar dores musculares em qualquer parte do corpo

⁶ Dor em uma ou mais articulações

⁷ Estado de abatimento extremo, físico e psíquico, que se traduz por imobilidade total e ausência de reações às solicitações exteriores

⁸ É uma erupção geralmente avermelhada que aparece na pele devido à dilatação dos vasos sanguíneos ou inflamação

movimento de sentar e andar

- Hipotensão postural e/ou lipotímia;
O paciente não consegue ficar em ortostase⁹.
- Vômitos persistentes;
O paciente não tolera a ingestão de nenhum tipo de substância inclusive água
- Hepatomegalia dolorosa;
O paciente apresenta um aumento do tamanho do fígado
- Sangramento de mucosa
O paciente apresenta sangramento em mucosa respiratórias e/ou sistema digestivo
- Sonolência e/ou irritabilidade;
O paciente com dificuldade de se manter alerta e apresenta agitação psicomotora
- Diminuição da diurese;
Devido à desidratação provocada pelos vômitos e/ou pela hipotensão arterial
- Hipotermia;
Dificuldade de manter a temperatura corpórea.
- Aumento repentino do hematócrito;
O desequilíbrio proporcional de água e células sanguíneas
- Queda abrupta de plaquetas;
Diminuição do número de plaquetas.
- Desconforto respiratório
O paciente apresenta dificuldade de respirar, podendo ser decorrente dos sangramentos ou edema pulmonar.

5.2.2.Fatores de risco relacionados à doença

A ocorrência de dengue está diretamente ligada a alguns fatores de riscos onde os principais são:

- locais sem saneamento adequado
- coleta de lixo precária

⁹ Ato de ficar com o pé no chão e pernas estendidas

5.2.3. Ferramentas de diagnóstico

O diagnóstico preciso de dengue é realizado por meio de sorologia. No entanto, estes exames são realizados em laboratórios e hospitais. Por esse motivo, o diagnóstico rápido é importante a fim de que seja feito o encaminhamento do paciente para tratamento. Os principais exames são:

- Sorologia;
- Hemograma.

5.3. Evolução dos modelos construídos

A construção da rede foi realizada em duas etapas de uma forma iterativa. Com o resultado obtivemos dois modelos, descritos abaixo.

O primeiro modelo construído é mostrado na Figura 14 este modelo contém parte apenas as patologias diretas sobre dengue.

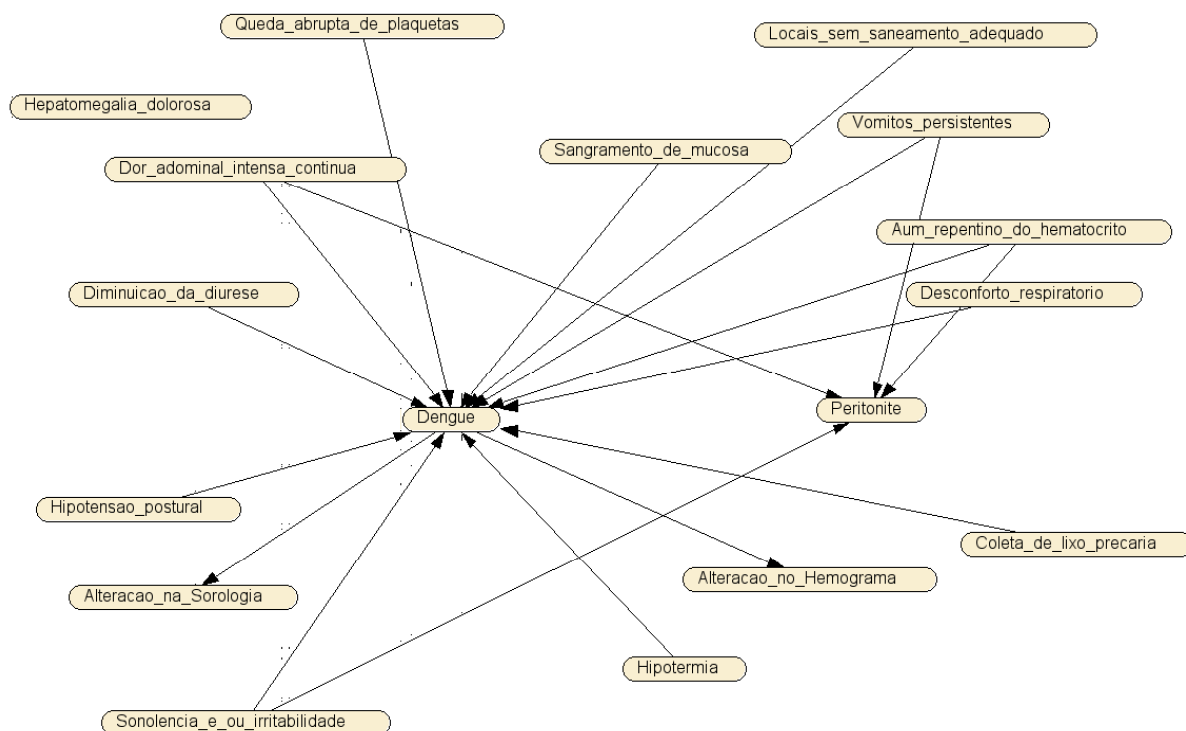


Figura 14 - Primeiro modelo da rede criada a partir do Nética

O principal objetivo deste modelo foi pré selecionar quais nós da rede e mostrar que tais nós podem indicar outro sintoma. Foi acompanhado o modelo e decidiu-se retirar o nó Peritonite da rede, pois não fornecia nenhum valor na descoberta do diagnóstico.

A partir desse estudo trabalhou-se para aprimorar a rede buscando identificar e classificar os pacientes com sintoma de dengue comum ou dengue hemorrágica, o que pode ser visualizado na figura 15.

A separação dos nós que são informados a partir dos usuários, agentes de saúde e dos especialistas também fica clara na figura 15.

Com o andamento das reuniões conseguimos criar o modelo final da rede, figura 16, que foi utilizada como base das inferências utilizadas no LARIISA_Bay.

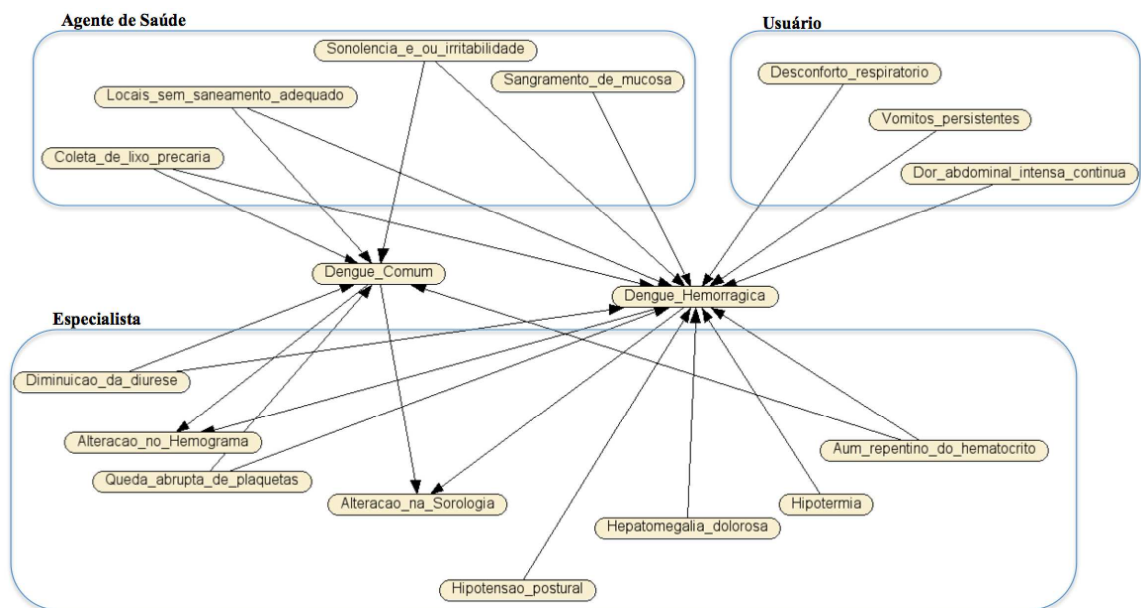


Figura 15 - Segundo modelo da rede criada a partir do Nética

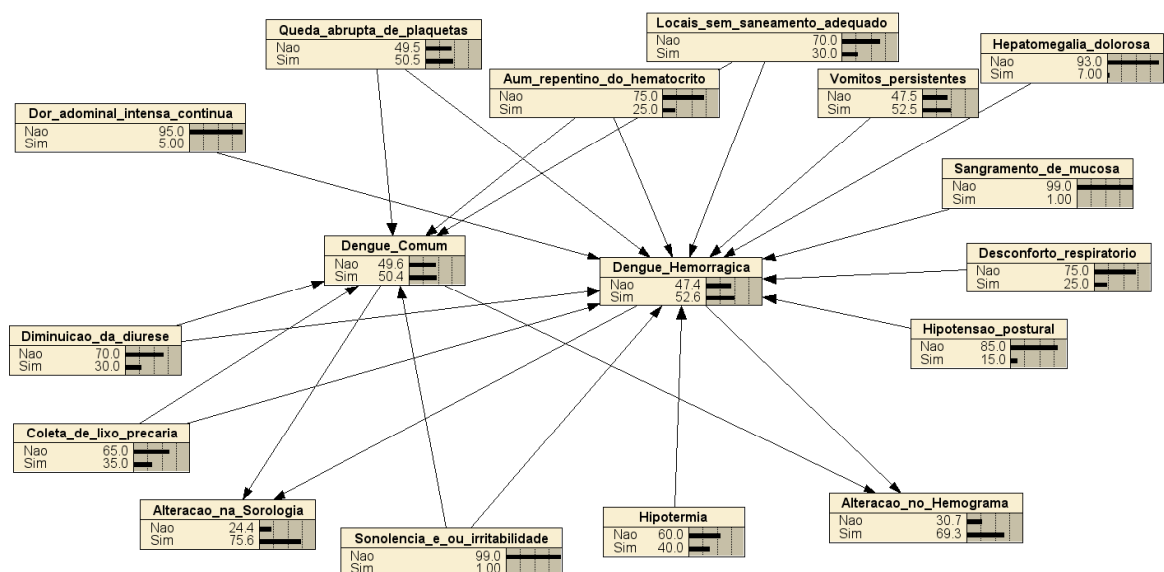


Figura 16 - Modelo final da rede criada a partir do Nética

Observou-se quem nos modelos criados, algumas patologias poderiam indicar a gravidade de dengue. Em relação a categorias, para a todos nós foram definidas apenas duas: *Sim* e *Não*. Tal representação é adequada quando existe uma doença com diversos fatores de causas ou uma patologia. Por exemplo, para o nó *Dengue_Hemorrágica* da rede construída neste trabalho, os sintomas são:

Queda_abrupta_de_plaquetas, Hepatomegalia_dolorosa, Dor_abdominal_intensa_contínua, Diminuição_da_diurese, Hipotensão_postural, Sonolência_e_ou_irritabilidade, Hipotermia, Coleta_de_lixo_precária, Desconforto_respiratório, Aum_repentino_do_hematócrito, Vômitos_persistentes, Locais_sem_saneamento_adequado e Sangramento_de_mucosa. Neste trabalho, o modelo *noisy-or* foi utilizado para os nós *Dengue_Hemorrágica* e *Dengue_Comum* devido ao excessivo número de probabilidades necessárias para a quantificação da rede, descrito na seção 2.12.2.

6. MODELAGEM DO LARIISA E LARIISA_Bay

LARIISA_Bay é um mecanismo de apoio à tomada de decisão baseado em dados probabilísticos para agravo de dengue. Trata-se de um componente agregado à arquitetura do projeto LARIISA que tem como objetivo apoiar especialistas no diagnóstico em agravo de dengue. .

6.1. Nova Arquitetura Funcional do LARIISA

Uma das contribuições deste trabalho foi a criação de uma nova arquitetura para o LARIISA. Como mostrado na figura 17, esta arquitetura é formada pelos seguintes componentes:

- Entrada do Sistema;
 - *Interface* composta pelo paciente, agente de saúde e o especialista;
 - Sensores de sinais vitais do paciente;
 - Outros provedores de contexto;
 - Metadado criado a partir da *interface* e sensores acima;
- LARIISA_Bay;
 - Módulo de Inferência do LARIISA_Bay: contém a rede Bayesiana criada a partir de uma tabela construída com a ajuda de especialistas
 - Sala de Situação: contém a *interface* que alimenta o módulo de decisão. Neste módulo residem os especialistas que recebem o apoio para a tomada de decisão.
- Saída do Sistema;
 - Diálogo com o paciente;
 - Acionamento do Agente de Saúde;
 - Acionamento da Ambulância;
 - Exibição do gráfico de epidemia.

A entrada do sistema (*interface* do paciente, do agente de saúde e do especialista; os sensores e outros provedores de contexto) já servia de insumo para as bases de ontologia e de outros mecanismos de inferência. Agora, esta entrada passa também a alimentar o módulo de inferência do LARIISA_Bay. Este módulo lê as informações do metadado e cria uma tabela de probabilidades que servirá de entrada para a Sala de Situação.

Na Sala de Situação, um responsável pela gestão macro do sistema (Agente de Governança não representado na figura 17) poderá escolher qual dos três cenários disponíveis (Decisão do especialista, Validação do Especialista ou Pass Through) é o mais adequado à realidade de uso do sistema. Por exemplo, um município que não possua especialistas para a análise/validação dos resultados fornecidos pela interface do módulo de decisão poderá escolher a opção 3 (Pass through).

Na figura 17 podem ser observadas duas linhas pontilhadas. Elas representam o resultado da rede Bayesiana alimentando o modelo de ontologia e a própria base de ontologia servindo de entrada à Sala de Situação.

Finalmente, a figura 17 apresenta uma saída que envolve todos os atores relacionados ao LARIISA. Neste módulo são executados os procedimentos resultantes da tomada de decisão na Sala de Situação. Como já descrito, esta tomada de decisão de especialistas (diálogo com o paciente; acionamento do Agente de Saúde e/ou ambulância) é apoiada pelo novo componente LARIISA_Bay, agregado ao sistema. O LARIISA_Bay também facilita a geração de um gráfico epidemiológico, a partir da informação do campo SUS_ID existente na estrutura do metadado.

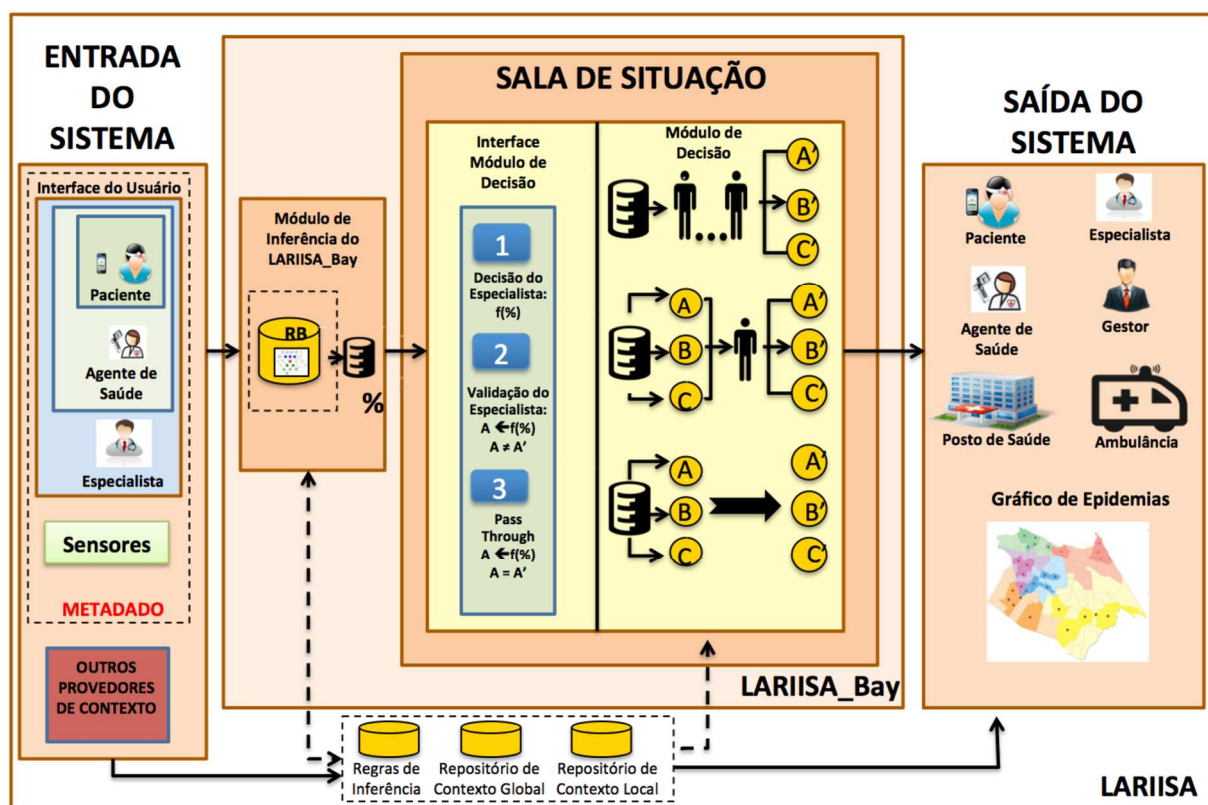


Figura 17- Arquitetura Funcional do LARIISA

6.2. Agente de Governança e o Módulo de Decisão do LARIISA_Bay

A arquitetura funcional do LARIISA_Bay é constituída por 4 módulos (figura 18), extraída da arquitetura do LARIISA (figura 17). Nela merece destaque o Módulo de Decisão, onde estão definidos os cenários de aplicação. Cabe a um Agente de Governança (não representado nas figuras 17 e 18) a seleção de um dos três cenários a seguir:

- Decisão do Especialista: cenário que considera a existência de uma equipe de especialistas capaz de melhor diagnosticar o agravo de dengue a partir da informação recebida da Interface Módulo de Decisão do LARIISA_Bay. A partir desta informação, resultado das funcionalidades do Módulo de Inferência, a equipe tomaria a decisão mais adequada no trato com determinado paciente.
- “Pass Thought”: cenário diametralmente oposto ao anterior no que diz respeito disponibilidade de uma equipe especializada para a tomada de decisão a partir da análise da informação fornecida pela *Interface* Módulo de Decisão. Neste caso, compete ao Agente de Governança avaliar se é melhor delegar a tomada de decisão ao próprio Módulo de Inferência do LARIISA_Bay (Pass Through) ou se é preferível monitorar o sistema enquanto aguarda a disponibilidade de especialistas para a tomada de decisão.
- Validação do Especialista: cenário intermediário entre os dois acima citados. Neste cenário estima-se que uma insuficiência de profissionais especialistas para a tomada de decisão. No entanto, considera-se que as informações fornecidas pela *Interface* Módulo de Decisão possam ser filtradas/validadas ao invés de analisadas como no primeiro cenário.

Percebe-se nesta versão da arquitetura do LARIISA a importância do novo Agente de Governança, o que implica em outras questões que fogem ao contexto técnico. Por exemplo, a questão do “Pass Thought” enfrenta divergências nas visões éticas que costumam aquecer o debate sobre o uso de procedimentos não humanos na tomada de decisão na área da saúde, o que foge do escopo deste trabalho.

É de se destacar que em qualquer um dos cenários acima descritos, o componente LARIISA_Bay atua com um redutor do espaço amostral.

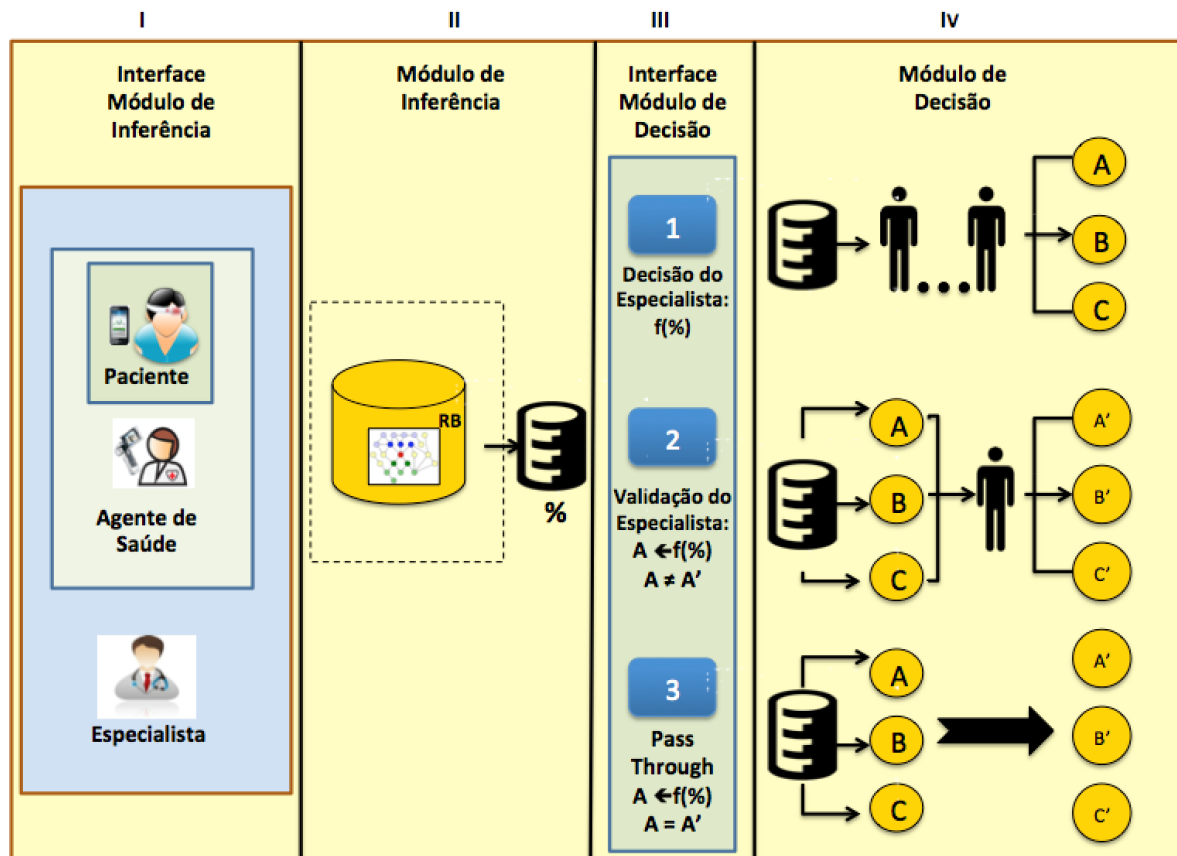


Figura 18: Arquitetura Funcional do LARISSA-BAY

6.3. Obtenção de probabilidades

O aprendizado em redes Bayesianas (HECKERMAN, GEIGER e CHICKERING, 1995) pode ter como objetivo a obtenção das probabilidades da rede, da estrutura da rede, ou ambos. Os objetos utilizados no aprendizado são um algoritmo de aprendizado e uma base de dados que corresponde ao domínio do problema modelado pela rede. No caso de problemas na área de saúde, a base corresponde, na maior parte dos casos, a dados clínicos de casos anteriores.

Há também a possibilidade de combinar-se o conhecimento de especialistas com técnicas de aprendizado. Por exemplo, a estrutura inicial da rede pode ser obtida a partir de especialistas e ser usada como ponto de partida para algoritmos de aprendizado. Ou então, pode-se obter as probabilidades por aprendizado, que seriam posteriormente verificadas por especialistas. De qualquer forma, o objetivo final é obter as probabilidades que melhor correspondam à realidade do problema modelado pela rede Bayesiana.

O principal obstáculo em utilizar o aprendizado (em redes Bayesianas para auxílio à decisão) é a dificuldade na obtenção da base de dados, pois a coleta de dados (que podem

auxiliar a construção de sistemas especialistas) não é uma prática comum no atendimento de na área de saúde.

Para a quantificação da rede Bayesiana apresentada neste trabalho foi utilizada a especificação direta de probabilidades, utilizando tabelas preenchidas pelo médico especialista, a partir de sua experiência profissional e dados bibliográficos. Tal método possibilitou a obtenção das probabilidades de uma forma rápida, seguida de correções nos valores de algumas probabilidades que se mostraram incoerentes durante as avaliações realizadas.

6.3.1. Probabilidades marginais

As probabilidades marginais são as mais simples de serem obtidas e correspondem a nós sem pais. Estas probabilidades correspondem à prevalência da doença na população em questão, isto é, nos usuários que se apresentam em postos de saúde. Graficamente, estas probabilidades marginais são representadas na parte superior da rede apresentada na Figura 16. As tabelas, a seguir, contêm as probabilidades marginais da rede Bayesiana desenvolvida no LARIISA_Bay. A primeira linha de cada tabela representa as categorias do nó, e a segunda a prevalência para cada categoria.

| Nó | SIM | NÃO |
|--------------------------------|-----|-----|
| Queda_abrupta_de_plaquetas | 1% | 99% |
| Hepatomegalia_dolorosa | 7% | 93% |
| Dor_adominal_intensa_contínua | 5% | 95% |
| Diminuição_da_diurese | 30% | 70% |
| Hipotensão_postural | 15% | 85% |
| Sonolência_e_ou_irritabilidade | 1% | 99% |
| Hipotermia | 40% | 60% |
| Coleta_de_lixo_precária | 35% | 65% |
| Desconforto_respiratório | 25% | 75% |
| Aum_repentino_do_hematócrito | 25% | 75% |
| Vômitos_persistentes | 5% | 95% |
| Locais_sem_saneamento_adequado | 30% | 70% |

Tabela 6 - Probabilidade a priori

6.3.2. Probabilidades Condicionais

Apenas três nós aparecem com probabilidade condicional em relação aos

respectivos pais na rede do LARIISA_Bay. Para a especificação destas probabilidades, os nós foram divididos em dois grupos: exames diagnósticos e sintomas patológicos. Os nós do grupo sintomas patológicos representam características da doença que, apesar de difícil medição, são indispensáveis para a modelagem.

a) exames diagnósticos

| Categorias das variáveis condicionantes | <i>Alteracao_no_Hemograma</i> | |
|---|-------------------------------|------------|
| | <i>Sim</i> | <i>Não</i> |
| Dengue = Sim | 87% | 13% |
| Dengue = Não | 62% | 38% |

Tabela 7 - Probabilidade Condicional 1

| Categorias das variáveis condicionantes | <i>Alteracao_na_Sorologia</i> | |
|---|-------------------------------|------------|
| | <i>Sim</i> | <i>Não</i> |
| Dengue = Sim | 99% | 1% |
| Dengue = Não | 9% | 91% |

Tabela 8 - Probabilidade Condicional 2

b) sintomas patológicos

As tabelas as seguir referem-se aos nós *Dengue_Comum* e *Dengue_Hemorrágica*. Para estes dois nós, foi utilizado o modelo *noisy-or*, apresentado na Seção 2.12.2. As tabelas necessárias para o *noisy-or* contêm as probabilidades condicionais do nó em questão em relação a cada um de seus pais, de forma não combinada, isto é, como se o nó tivesse apenas um pai de cada vez.

| Categorias das variáveis condicionantes | <i>Dengue_Comum</i> | |
|---|---------------------|------------|
| | <i>Sim</i> | <i>Não</i> |
| Queda_abrupta_de_plaquetas=Sim | 63% | 37% |
| Queda_abrupta_de_plaquetas=Não | 27% | 73% |
| Diminuicao_da_diurese=Sim | 73% | 27% |
| Diminuicao_da_diurese=Não | 92% | 8% |
| Aum_repentino_do_hematócrito=Sim | 60% | 40% |
| Aum_repentino_do_hematócrito=Não | 65% | 35% |

Tabela 9 - Probabilidade Condicional 3

| Categorias das variáveis condicionantes | <i>Dengue Hemorrágica</i> | |
|---|---------------------------|------------|
| | <i>Sim</i> | <i>Não</i> |
| Queda_abrupta_de_plaquetas=Sim | 98% | 2% |
| Queda_abrupta_de_plaquetas=Não | 2% | 98% |
| Diminuição_da_diurese=Sim | 97% | 3% |
| Diminuição_da_diurese=Não | 3% | 97% |
| Aum_repentino_do_hematócrito=Sim | 97% | 3% |
| Aum_repentino_do_hematócrito=Não | 3% | 97% |
| Queda_abrupta_de_plaquetas=Sim | 99% | 1% |
| Queda_abrupta_de_plaquetas=Não | 2% | 98% |
| Hepatomegalia_dolorosa=Sim | 74% | 26% |
| Hepatomegalia_dolorosa=Não | 15% | 85% |
| Dor_abdominal_intensa_contínua=Sim | 68% | 32% |
| Dor_abdominal_intensa_contínua=Não | 10% | 90% |
| Diminuicao_da_diurese=Sim | 96% | 4% |
| Diminuição_da_diurese=Não | 3% | 97% |
| Hipotensão_postural=Sim | 98% | 2% |
| Hipotensão_postural=Não | 2% | 98% |
| Desconforto_respiratório=Sim | 81% | 19% |
| Desconforto_respiratório=Não | 2% | 98% |

Tabela 10 - Probabilidade Condicional 4

6.3.3. Forma utilizada para a avaliação da rede

O Teorema de Bayes é empregado frequentemente na realização de teste de diagnóstico ou triagens. A triagem é a aplicação de um teste em indivíduos que não apresentam qualquer sintoma clínico para classificá-los com relação às probabilidades de terem em uma doença particular. Os que apresentem resultado positivo são considerados mais prováveis de terem a doença e normalmente são submetidos a procedimentos de diagnósticos adicionais ou a tratamentos. A triagem é frequentemente mais utilizada por profissionais na

área de saúde em situações nas quais a detecção previa de doença contribui para prognósticos mais favoráveis ao indivíduo ou para a população em geral.

Neste trabalho é apresentada uma triagem em relação aos diagnósticos fornecidos por uma comissão de médicos que criticaram as principais formas de se identificar sintomas de dengue. Para a criação da rede das informações coletadas foram consideradas 2 partes: a primeira, onde o próprio paciente poderá informar dados sintomas e a segunda, onde o profissional de saúde pode completar o restante dos dados de uma forma mais técnica.

A primeira parte é para a avaliação da rede utilizando apenas informações fornecidas pelo paciente. Três casos foram analisados representando: baixo, alto e médio risco.

| | <i>1 - Baixo Risco</i> | <i>2 - Alto Rico</i> | <i>3- Médio Risco</i> |
|--------------------------------|------------------------|----------------------|-----------------------|
| REDE | 9.89% | 92.2% | 24.8 |
| MÉDICO | 5% | 86% | 23% |
| CONCORDA? | Parcialmente | Sim | Sim |
| Dor_abdominal_intensa_contínua | Não | Sim | Sim |
| Desconforto_respiratório | Não | Sim | Não |
| Vômitos_persistentes | Não | Sim | Sim |
| Sangramento_de_mucosa | Não | Sim | Sim |
| Sonolência e/ou irritabilidade | Não | Sim | Sim |
| Locais_sem_saneamento_adequado | Sim | Não | Não |
| Coleta_de lixo_precária | Sim | Não | Não |

Tabela 11 - Casos de uso para rede simplificada

A linha *Rede* da tabela acima representa o resultado fornecido pela rede para o nó dengue. A linha *Médico* é a probabilidade dada pelo médico para o mesmo nó, e a linha *concorda?* Representa a opinião do médico em relação ao desempenho da rede.

Já a tabela abaixo, mostra o resultado fornecido pela Rede dos dados fornecido pelo Agente de Saúde e enviado para a rede criada.

| | <i>1 - Baixo Risco</i> | <i>2 - Alto Rico</i> | <i>3- Médio Risco</i> |
|----------------------------|------------------------|----------------------|-----------------------|
| REDE | 1.63% | 99.7% | 11.1% |
| MÉDICO | 5% | 96% | 27% |
| CONCORDA? | Parcialmente | Sim | Sim |
| Queda_abrupta_de_plaquetas | Não | Sim | Não |
| Hepatomegalia_dolorosa | Não | Não | Não |

| | | | |
|------------------------------|-----|-----|-----|
| Hipotermia | Não | Sim | Não |
| Alteração_no_Hemograma | Não | Não | Sim |
| Hipotensao_postural | Não | Sim | Sim |
| Alteração_na_Sorologia | sim | Sim | Não |
| Aum_repentino_do_hematócrito | Não | Sim | Sim |
| Diminuição_da_diurese | Não | Sim | Não |

Tabela 12- Casos de uso para a rede completa

Foi feita uma última avaliação com casos de uso. Nesta avaliação foram utilizados casos mais variados, simulado em ambientes onde estas informações nem sempre estão disponíveis para o agente de saúde.

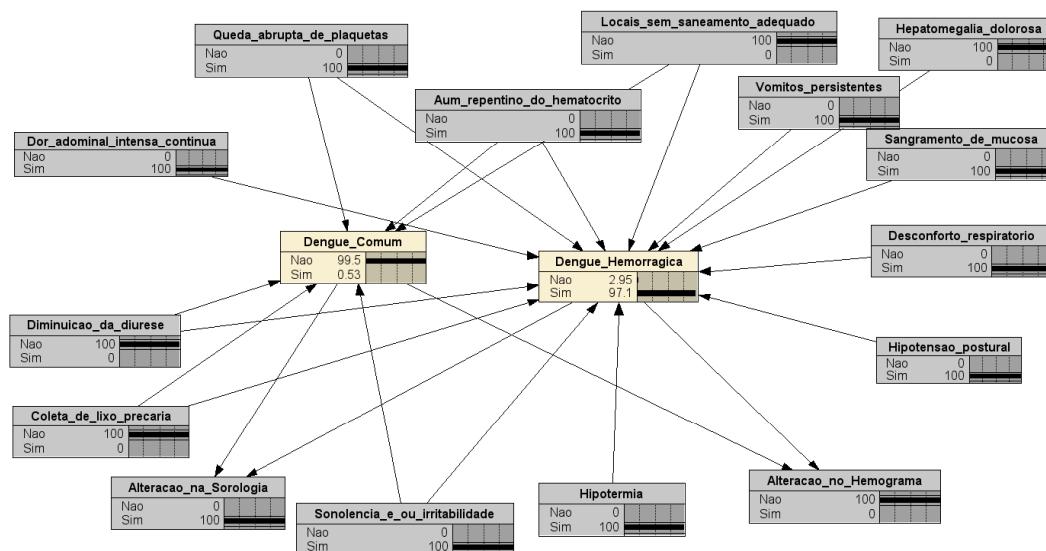


Figura 19 – Alto Risco

A figura 19 apresenta um alto risco de Dengue Hemorrágica com as alterações na Sorologia, sangramento da mucosa, entre outros sintomas apresentados na figura. Observa-se que este seria um caso onde os exames de sangue teriam uma influência significativa.

Já na figura 20, apresentada abaixo, mostra um caso de baixo risco, não apresentando nenhum outro sintoma, apenas informações sobre o saneamento básico e a coleta de lixo em sua região.

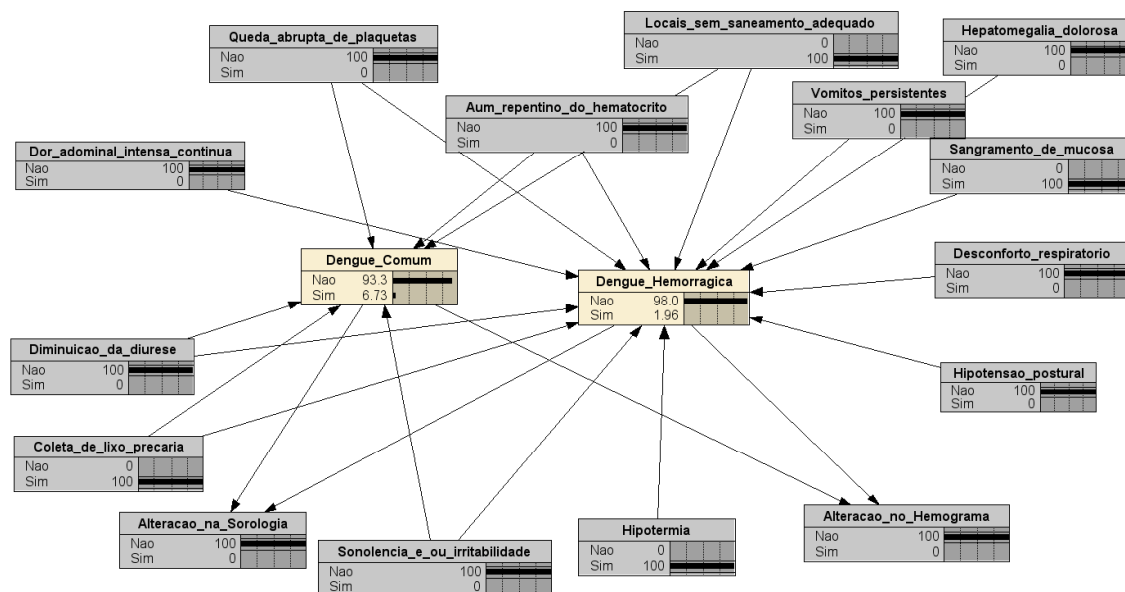


Figura 20 – Baixo Risco

O desempenho da rede obtido na última etapa da avaliação com casos de uso foi considerado adequado pelo especialista, ou seja, as probabilidades fornecidas pela rede correspondem às expectativas do especialista. Tal fato representa principalmente a adequação das probabilidades condicionais especificadas no Capítulo 6 em relação ao funcionamento geral da rede. É possível que outro especialista, que não tenha participado da construção da rede, possa discordar dos resultados.

Não foi notado nenhum prejuízo no desempenho da rede devido à utilização do modelo *noisy-or*, visto que não foram apontadas incoerências nas probabilidades dos nós onde o modelo foi aplicado.

O segundo tipo de avaliação foi feito de forma qualitativa, a fim de obter a opinião de outros especialistas médicos quanto à estrutura da rede. Esta avaliação considerou apenas a estrutura de nós, incluindo suas categorias, e arcos da rede. Os aspectos relativos às probabilidades e desempenho da rede não fizeram parte desta avaliação.

Para a realização desta avaliação, foram utilizados questionários, que foram preenchidos e comentados por um especialista convidado. As questões fornecidas e as respostas obtidas são apresentadas nesta seção.

Questionário

Local de trabalho: Hospital São José de doenças infectocontagiosas

- a) Você concorda com a organização dos arcos da rede? ☐

R. Sim.

- b) Os nós que representam causas diretas ou indiretas nos sintomas de dengue são:
Vômitos_persistentes, Sangramento_de_mucosa, Sonolência e/ou irritabilidade,

Desconforto_respiratório, Hipotensão_postural, Aum_repentino_do_hematócrito, Diminuição_da_diurese, Dor_abdominal_intensa_contínua, Hipotermia. Quais destes nós não deveriam estar presentes e quais outros fatores deveriam ser adicionados? Por quê?

R. Os nós citados são pertinentes e o que se poderia acrescentar seria que pacientes podem evoluir para o choque sem que sejam observados sangramentos ou prova do laço positiva, como também presença de derrames cavitários. Isso porque são sinais de alarme de piora e evolução rápida para o óbito.

- c) A mesma questão para os nós referentes ao fator de risco relacionado à doença: *Locais sem saneamento adequado, Coleta de lixo precária*

R. A presença de casos semelhantes na família, vizinhança, escola ou creches. Existência de deslocamento, nos últimos 15 dias a locais de transmissão de dengue

- d) A mesma questão para os nós referentes a resultados de exames: *Sorologia e Hemograma*

R. Histórico de infecção pregressa por dengue com ou sem confirmação sorológica.

- e) Qual a sua opinião sobre o nível de complexidade da rede?

R. Satisfatório

7. AMBIENTE DE TESTE E EXPERIMENTOS

Este capítulo descreve a implementação do protótipo LARIISA_Bay. Ele apresenta a biblioteca Nética-Java API utilizada na construção das redes Bayesianas, e é mostrada, arquitetura do protótipo, além das três interfaces implementadas e a Sala de Situação.

A API Nética-Java está disponível a partir da versão Windows 95, *Windows* NT 4.0 e Macintosh; tendo uma versão para UNIX, esta API é de propriedade da empresa Norsys *Software* Corp. que disponibiliza uma versão gratuita e uma paga, onde a gratuita com um limite do número de nós.

7.1. Ferramenta de implementação da rede Bayesiana

Os exemplos da rede Bayesiana concebidas no LARIISA_Bay foram gerados utilizando o Nética-Java API, uma biblioteca de classes em Java. Esta biblioteca utiliza redes bayesianas para realizar vários tipos de inferência usando algoritmos modernos e rápidos. Desenvolveu-se, assim, um modelo com a ajuda do API Nética-Java, tornando possível a implementação e a acoplagem deste ao protótipo.

A construção das redes Bayesianas foi realizada por meio de chamadas a procedimentos da API Nética-Java. Uma vez construída a rede, foi armazenada em um arquivo para uso posterior. O apêndice A apresenta o código para a construção de uma rede Bayesiana com a estrutura da rede definida na figura 20. Ao final, toda a estrutura da rede é salva no arquivo “LARIISA.dne”, cuja estrutura é apresentada no apêndice B. Este arquivo é utilizado para todas as atividades relacionadas com a manipulação da rede, como inferência probabilística e atualização das probabilidades condicionais.

Após uma rede Bayesiana ser construída, ela pode ser utilizada para a execução de inferências probabilísticas. Cada valor conhecido de variável é informado como uma evidência e a rede se encarrega de encontrar as crenças para todas as outras variáveis.

Dado um novo caso, que o usuário tem conhecimento limitado, a API Nética-Java encontrará os valores ou probabilidades apropriadas para todas as variáveis desconhecidas. O caso em evidência pode ser convenientemente salvo em um arquivo e depois incluído dentro da rede (ou em uma rede diferente) para incrementar a consulta, ou para trazer uma nova informação sobre o este caso.

Benefícios advindos da utilização do API Nética-Java:

- Gerar uma apresentação gráfica de qualidade, que pode ser incorporada dentro de outros documentos,

- Poder encontrar decisões ótimas para problemas de decisão sequencial,
- Solucionar diagramas de influência,
- Poder aprender relações probabilísticas através de dados,
- Permitir fácil atualização da rede de crença e dos diagramas de influência, incluindo: excluir, colar e duplicar nodos da rede de crença e dos diagramas de influência,
- Manter diagramas complexos ordenados,
- Permitir comentários,
- Permitir desfazer e refazer comando ilimitadamente,
- Permitir a entrada de relações probabilísticas através de equações, com uma extensa biblioteca de funções probabilísticas e funções matemáticas,
- Realizar com facilidade a discretização de variáveis contínuas.

7.2. Arquitetura do protótipo

Geolocalização, outro conceito importante no protótipo, é uma tecnologia baseada em GPS, frequentemente usada em dispositivos móveis. Esta tecnologia identifica com precisão a posição onde se encontra o dispositivo a partir de dados geográficos (latitude & longitude). Essa posição também pode ser feita pela identificação de computadores conectados à *Internet*, a partir do endereço IP.

O protótipo fornece uma interface cujos metadados enviam informações de contexto do paciente: sua geolocalização, o seu número do Cartão Nacional de Saúde do paciente (SUS ID), além dos dados coletados de seus sinais vitais do paciente. Estes metadados alimentam o LARIISA com informações necessárias ao desenvolvimento de inferências que resultem na melhor atuação do Agente de Saúde no combate à dengue.

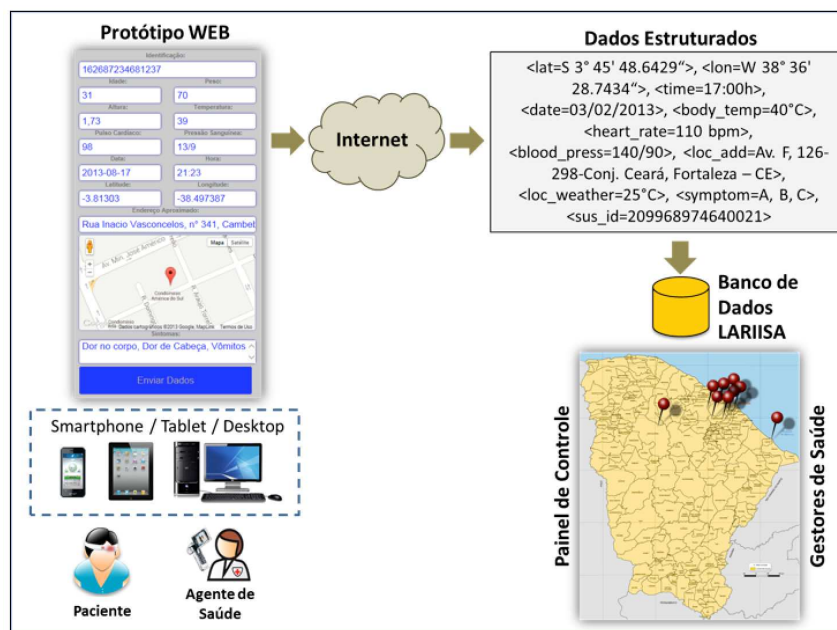


Figura 21 - Arquitetura do sistema proposto

O protótipo do LARIISA_Bay pode ser classificado como um componente de gestão do conhecimento no domínio Clínico-Epidemiológica no contexto do projeto LARIISA.

Na figura 21 pode-se observar uma visão geral de funcionamento do protótipo. Ela mostra a *interface web* com os seus autores, a representação dos dados através do metadado, o banco de dados do LARIISA e a representação do painel de controle gerenciados pelos agentes de governança.

7.3. Interfaces do protótipo

O protótipo do LARIISA_Bay, componente agregado à plataforma LARIISA, pode ser observado em seu funcionamento sob dois enfoques:

- No primeiro, uma *interface* do usuário identifica (figura 22) quem está utilizando o aplicativo: o usuário, o agente de saúde ou o especialista. Assim, de acordo com o perfil selecionado, LARIISA_Bay envia um grupo de perguntas específicas à ser respondido pelo utilizador do sistema.
- No segundo enfoque, o protótipo utiliza as informações recebidas para alimentar as probabilidades à rede Bayesiana no LARIISA_Bay.

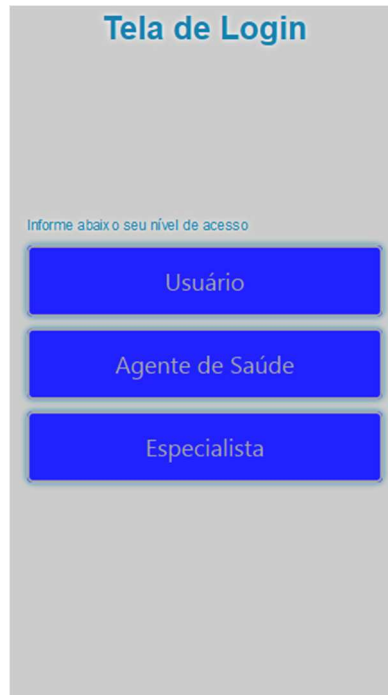


Figura 22 – Tela Login

Uma vez realizado estes dois passos, a rede Bayesiana do LARIISA_Bay pode agora dar sequência a uma série de ações de inferência que resultem na classificação do caso em análise. Uma vez feita a classificação, os resultados estarão disponíveis para a tomada de decisão pela equipe de especialistas, dentro do preconizado no modelo proposto neste trabalho, ilustrado na figura 18 (arquitetura do LARIISA_Bay).

Também mostrado na figura 18, a *interface* do usuário é um subconjunto da *interface* do agente de saúde. Esta última, por sua vez, é um subconjunto da *interface* do especialista. As figura 21, 22 e 23 mostra as telas e os relatórios associados aos perfis do usuário, do agente de saúde e do especialista, respectivamente.

Uma vez respondido e submetido às informações dos relatórios pela *interface* de entrada é gerado um arquivo chamado “relatorio.txt” com as todas as informações armazenadas que serão consumidas pelo componente LARIISA_Bay.

Identificação:

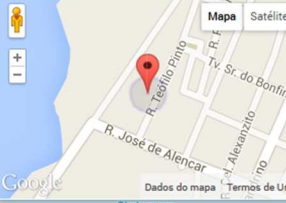
Insira sua identificação do SUS

Idade: Ex: 27 anos Peso: Ex: 70,6 Kg

Altura: Ex: 1,75 m Temperatura: Ex: 37,0 °C

Endereço Aproximado em Raio de 43 Metros:

Rua Teófilo Pinto, 136-378, Aracati

 Mapa Satélite

Sintomas:

Resuma em 140 caracteres seus atuais sintomas

Relatório do Usuário

Mova o cursor para esquerda para Sim, direita para Não e centralize caso não saiba responder.

Possui dor abdominal intensa e continua?

Esta sentido algum desconforto respiratório ou dificuldade de respirar?

A ingestão de algum tipo de substância inclusive água provoca vômitos?

Figura 23 – Perfil Usuário

Identificação:

Insira sua identificação do SUS

Idade: Ex: 27 anos Peso: Ex: 70,6 Kg

Altura: Ex: 1,75 m Temperatura: Ex: 37,0 °C


Pulso Cardíaco: Ex: 80 bpm Pressão Sanguínea: Ex: 12/8

Data: 2013-09-24 Hora: 22:30

Latitude: -4.568539299999999 Longitude: -37.7764713

Endereço Aproximado em Raio de 43 Metros:

Rua Teófilo Pinto, 136-378, Aracati - CE, 62800-000, República

 Mapa Satélite

Sintomas:

Resuma em 140 caracteres seus atuais sintomas

Relatório do Agente de Saúde

Mova o cursor para esquerda para Sim, direita para Não e centralize caso não saiba responder.

Apresenta algum tipo de sangramento, vias respiratórias e/ou sistema digestivo?

Esta com dificuldade de se manter alerta apresentando sonolência e/ou irritabilidade?

Reside em local sem saneamento básico?

Localidade possui coleta de lixo regular?

Figura 24 – Perfil Agente de Saúde

Identificação:

Insira sua identificação do SUS

Idade: Ex: 27 anos Peso: Ex: 70,6 Kg

Altura: Ex: 1,75 m Temperatura: Ex: 37,0 °C


Pulso Cardíaco: Ex: 80 bpm Pressão Sanguínea: Ex: 12/8

Data: 2013-09-24 Hora: 22:30

Latitude: -4.568539299999999 Longitude: -37.7764713

Endereço Aproximado em Raio de 43 Metros:

Rua Teófilo Pinto, 136-378, Aracati - CE, 62800-000, República

 Mapa Satélite

Sintomas:

Resuma em 140 caracteres seus atuais sintomas

Relatório do Especialista

Mova o cursor para esquerda para Sim, direita para Não e centralize caso não saiba responder.

Paciente apresenta queda abrupta de plaquetas?

O paciente apresenta um aumento do tamanho do fígado (Hepatomegalia dolorosa)?

O paciente apresenta dificuldade de manter a temperatura corpórea?

O exame do hemograma encontra-se alterado?

O paciente não consegue ficar em ortostase?

Existe algum desequilíbrio proporcional de água e células sanguíneas (hematócrito)?

Apresenta diminuição da diurese?

O exame de sorologia encontra-se alterado?

Figura 25 – Perfil Especialista

Na implementação das *interfaces* foram usadas as seguintes ferramentas:

- *JavaScript*
- HTML5
- CSS3

7.4. Simulações com o LARIISA_Bay

Como pode ser observado nas figuras 21,22 e 23, as telas dos relatórios a serem enviados ao LARIISA variam de acordo com o perfil envolvido, onde todas essas informações servem de entrada para o processo de inferência a ser realizado pelo mecanismo de suporte à decisão.

As perguntas utilizadas nas interfaces do LARIISA_Bay foram extraídas da experiência técnica de profissionais experientes da área de saúde. Os questionários abordam perguntas desde saneamento básico existente no local onde reside o usuário (*interface* usuário) a exames sanguíneos (*interface* agente de saúde), até avaliações sobre plaquetas (*interface* dos especialistas).

Foram realizadas simulações preliminares que servem como prova de conceito da proposta. Nelas pode-se constatar que a ausência das informações enviadas a base de dados traz, naturalmente, algumas discrepâncias no resultado da rede.

7.5. Interface de Governança

A *interface* do usuário envia um arquivo .txt com as respostas dos relatórios junto com o metadado para o servidor do LARIISA. Como foi descrito no Módulo de Decisão a partir da figura 18 o agente de governança escolhe o cenário apropriado a sua necessidade (figura 26).

A escolha do cenário feita pelo agente dispara uma rotina no LARIISA onde o arquivo relatorio.txt é lido e executado por meio de chamadas a um *WebService* na API Nética-Java (Apêndice A) para assim obter a probabilidade e/ou os relatórios, exibindo ao agente de governança.

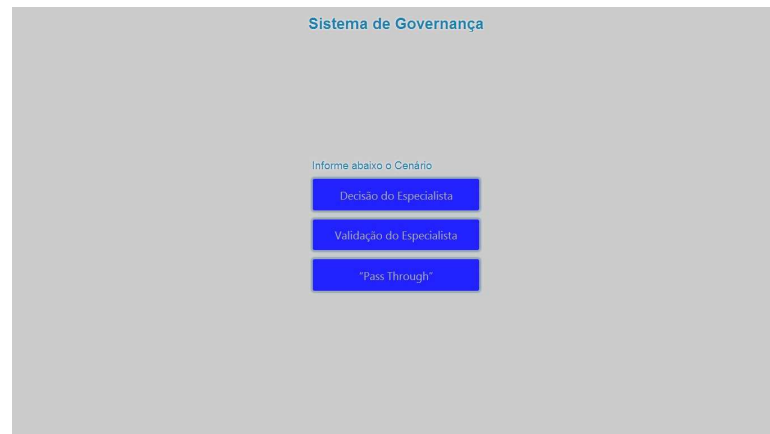


Figura 26 – Sala de Situação

A sala de situação é composta de três telas:

- Decisão do Especialista;
- Validação do Especialista;
- “Pass Through”

A tela de Decisão do especialista visualizada na figura 27 mostra a identificação do usuário onde será feita a busca pelo SUS_ID exibindo as probabilidades, calculadas pela rede Bayesiana, do paciente possuir os sintomas de dengue comum ou Hemorrágica junto com todas as informações do arquivo metadado que contém as informações do paciente e do arquivo TXT que contém as respostas dos relatórios.

Neste cenário o responsável (Médico) pode colocar o seu diagnóstico a partir da avaliação dos dados exibidos e disponíveis no prontuário do paciente, já que possui seu SUS_ID, onde o número elevado de profissionais de saúde em atividade pode dar vazão a esse tipo de cenário.

Identificação:

Probabilidade de Dengue Comum: 4%
Probabilidade de Dengue Hemorrágica: 0.1%

Sugestão do Agente de Governança:

Resultados do Relatório:

- Possui dor abdominal intensa e continua? **Sim**
- Localidade possui coleta de lixo regular? **Não**
- Reside em local sem saneamento básico? **Não Sei**
- O exame do hemograma encontra-se alterado? **Não**
- O paciente não consegue ficar em ortostase? **Sim**
- Apresenta diminuição da diurese? **Não Sei**
- Paciente apresenta queda abrupta de plaquetas? **Sim**
- O exame de sorologia encontra-se alterado? **Sim**
- O paciente apresenta dificuldade de manter a temperatura corpórea? **Não**

Aceitar

Idade: 25 **Peso:** 72
Altura: 1,75 **Temperatura:** 38
Pulso: 80 **Pressão:** 12/8
Data: 17/12/2013 **Hora:** 19:30
Sintomas: Febre, dor no corpo, dor de cabeça, tosse, indisposição
Endereço: Rua Miguel Filismino, 143-247, Aracati - CE, 62800-000, República Federativa do Brasil

Mapa:

Figura 27 - Decisão do Especialista

A tela de Validação do Especialista, figura 28 e 29 possui as mesmas informações já descritas, mas se diferencia pela exibição de uma sugestão(Ex.: Normal) do sistema sobre o estado do paciente dado as informações contidas em seu relatório.

Com isso, fica a critério do especialista aceitar ou recusá-la, como podemos verificar na figura 28. Caso o especialista recuse ele deve informar um novo diagnóstico. Este cenário deve ser utilizado quando o quantitativo de profissionais de saúde em atividade é insuficiente.

Identificação: Insira a identificação do SUS

Probabilidade de Dengue Comum: 4%

Probabilidade de Dengue Hemorrágica: 0.1%

Resultado para avaliação: Normal

Resultados do Relatório:

- Possui dor abdominal intensa e contínua? **Sim**
- Localidade possui coleta de lixo regular? **Não**
- Reside em local sem saneamento básico? **Não Sei**
- O exame do hemograma encontra-se alterado? **Não**
- O paciente não consegue ficar em ortostase? **Sim**
- Apresenta diminuição da diurese? **Não Sei**
- Paciente apresenta queda abrupta de plaquetas? **Sim**
- O exame de sorologia encontra-se alterado? **Sim**
- O paciente apresenta dificuldade de manter a temperatura corpórea? **Não**

Idade: 25 **Peso:** 72

Altura: 1,75 **Temperatura:** 38

Pulso: 80 **Pressão:** 12/8

Data: 17/12/2013 **Hora:** 19:30

Sintomas: Febre, dor no corpo, dor de cabeça, tosse, indisposição

Endereço: Rua Miguel Filismino, 143-247, Aracati - CE, 62800-000, República Federativa do Brasil

Figura 28 – Validação do Especialista

Identificação: Insira a identificação do SUS

Resultado para avaliação: Normal Recusado

Sugestão do Agente de Governança:

Resultados do Relatório:

- Possui dor abdominal intensa e contínua? **Sim**
- Localidade possui coleta de lixo regular? **Não**
- Reside em local sem saneamento básico? **Não Sei**
- O exame do hemograma encontra-se alterado? **Não**
- O paciente não consegue ficar em ortostase? **Sim**
- Apresenta diminuição da diurese? **Não Sei**
- Paciente apresenta queda abrupta de plaquetas? **Sim**
- O exame de sorologia encontra-se alterado? **Sim**
- O paciente apresenta dificuldade de manter a temperatura corpórea? **Não**

Idade: 25 **Peso:** 72

Altura: 1,75 **Temperatura:** 38

Pulso: 80 **Pressão:** 12/8

Data: 17/12/2013 **Hora:** 19:30

Sintomas: Febre, dor no corpo, dor de cabeça, tosse, indisposição

Endereço: Rua Miguel Filismino, 143-247, Aracati - CE, 62800-000, República Federativa do Brasil

Cancelar **Submeter**

Figura 29 – Validação do Especialista recusando o resultado sugerido

A tela *Pass Throug* é o último cenário, figura 30, mostra os dados do paciente e o resultado da avaliação da rede, nesta tela o agente de governança não interfere no diagnóstico, um situação característica onde existe escassez do profissional de saúde.

Identificação:

Idade: 25 **Peso:** 72

Altura: 1,75 **Temperatura:** 38

Pulso: 80 **Pressão:** 12/8

Data: 17/12/2013 **Hora:** 19:30

Sintomas: Febre, dor no corpo, dor de cabeça, tosse, indisposição

Endereço: Rua Miguel Filismino, 143-247, Aracati - CE, 62800-000, República Federativa do Brasil

Resultado para avaliação: Emergência

Mapa: Google Maps showing the location in Aracati, CE.

Figura 30 - Pass Throug

8. CONCLUSÃO

A pesquisa na área de sistemas inteligentes tem como um de seus principais objetivos, senão o principal, produzir resultados que possam ser aplicados em situações práticas. Temos como exemplos, robôs autônomos e sistemas de auxílio à tomada de decisão.

A tomada de decisão em sistemas de saúde torna-se um processo complexo por envolver diversas variáveis permeadas por incertezas. Estas são intrínsecas tanto aos vários autores do sistema (usuário, agente de saúde, médicos, diretores de hospital, secretários, etc.) quanto a fatores marginais a estes atores (sazonalidades epidemiológicas, aspectos econômicos e sociais).

O projeto LARIISA se apresenta como uma plataforma hábil no apoio à tomada de decisão na governança em saúde em seus cinco domínios de atuação (Gestão de Conhecimento, Normativo Sistêmico, Clínico/Epidemiológico, Administrativo e Gerenciamento Compartilhado). Para tanto, o LARIISA faz uso de ontologias na representação do conhecimento e de informações de contexto (*context-aware concept*) das diversas variáveis citadas (atores e fatores marginais).

A contribuição deste trabalho diz respeito ao tratamento da incerteza da informação. Portanto, ele se relaciona com ambos os conceitos-chave do LARIISA: a representação da informação sensível ao contexto e a representação do conhecimento.

O projeto LARIISA não tinha, até então, tratado a questão da incerteza. Este trabalho representa, então, um primeiro esforço nesta direção: a definição, especificação e a implementação de um modelo para tratar incertezas nas informações de contexto e na representação do conhecimento no projeto LARIISA.

A contribuição do trabalho é melhor evidenciada com a introdução de um novo componente no mecanismo de inferência do projeto LARIISA. Trata-se do LARIISA_Bay, um componente baseado em redes bayesianas que permite a captura de informações de contexto e produz inferências que auxiliam especialistas a melhor diagnosticar casos de dengue a partir de dados coletados (usuários, agentes de saúde e especialistas) no sistema e de uma tabela de probabilidades sobre a ocorrência de sintomas, construída por profissionais da área.

Três enfoques merecem registro na concepção e desenvolvimento do LARIISA_Bay, proposto neste trabalho como complemento à plataforma LARIISA.

- Este componente pode auxiliar uma equipe de especialistas a melhor diagnosticar casos de dengue classificando-os em: emergencial (dengue hemorrágica), grave (existência de dengue) ou normal (não existência da

enfermidade). A partir desta classificação, a equipe de especialistas poderá tomar a decisão mais conveniente como, por exemplo, a providência imediata de uma ambulância ou a visita ao usuário de uma agente de saúde ou, simplesmente, uma notificação de orientação ao usuário, respectivamente à classificação.

- A criação da estrutura da rede Bayesiana e a quantificação das probabilidades foram feitas com o auxílio de especialistas. A estrutura da rede foi obtida a partir de dados bibliográficos e reuniões com um especialista médico. As probabilidades foram obtidas através de técnicas de elucidação, questionários e posteriores correções dos valores iniciais.
- A prototipação do modelo proposto envolve desde *interfaces* móveis para a veiculação de metadados capazes de alimentar as tabelas de probabilidades à rede bayesiana construída a partir da consulta á diversos profissionais especialistas em agravos de dengue. As três *interfaces* desenvolvidas (usuário, agente de saúde de especialista) são concernentes a três diferentes metadados que traduzem os diferentes níveis de acesso à informação e capacidade de decisão. Ou seja, embora a tomada de decisão final (diagnóstico da dengue) seja de competência do especialista, o agente de saúde e o próprio usuário são também tomadores de decisão em instâncias mais restritas.

8.1. Trabalhos Futuros

Com o LARIISA_Bay, o novo componente baseado em redes bayesianas, será possível ao LARIISA apoiar a detecção e monitoramento de epidemias em diversas regiões. Provavelmente, uma tecnologia inteligente de aprendizagem (redes neurais, por exemplo), se adequaria na implementação desta funcionalidade. Embora fora do escopo deste trabalho, esta funcionalidade se apresenta como um trabalho futuro, dada sua relevância para o sistema único de saúde (SUS).

Finalmente, é relevante destacar que a mesma solução proposta neste trabalho poderá ser analisada e avaliada tomando-se como referência um modelo de ontologia de domínio construído para atender aos objetivos estratégicos da área de estudo. A procura de um mecanismo híbrido, utilizando ontologia e redes Bayesianas, torna a predição mais refinada além de possibilitar inferências indiretas de difícil obtenção sem uma modelagem baseada em ontologia. Neste caso, o resultado da rede Bayesiana poderia alimentar o modelo

de ontologia, ou seja, a sua base de dados e, a partir daí, gerar as inferências resultantes na ontologia do LARIISA, utilizando-se estratégias de busca semântica na base de informação modelada com a ontologia. A inferência obtida com as redes Bayesianas não monta o grafo de inferências indiretas, podendo estas inferências resultantes serem usadas tanto no atendimento do usuário (escopo local - Figura 7) como na conclusão se aquele local possui ou não uma epidemia (escopo global Figura 8).

Ainda na perspectiva de trabalhos futuros, propõe-se a realização de testes práticos e análise de sensibilidade. Esta análise determina a influência que as alterações nas probabilidades dos nós têm no comportamento da rede.

Além de ter contribuído conceitualmente na evolução do projeto LARIISA, o LARIIS-BAY deverá ser utilizado sendo realizado pelo DATASUS no município de Tauá, Ceará.

A expectativa final deste trabalho é ser adotado, juntamente com o LARIISA, como uma solução que ajude efetivamente o Ministério da Saúde no combate da dengue no País.

9. REFERÊNCIAS BIBLIOGRÁFICAS

- ALBUQUERQUE, L. **Sistema Especialista de Apoio à Nutrição Enteral Pediátrica – SANEP**. Dissertação de Mestrado, Programa de Pós-Graduação em Ciência da Computação. [S.l.]. 2001.
- ALCÂNTARA, T. P. D. A. **PAOLA: Uma Plataforma Para O Desenvolvimento De Aplicações Baseadas Em Ontologias Para O Projeto Lariisa**. MESTRADO PROFISSIONAL EM COMPUTAÇÃO APLICADA. [S.l.]. 2012.
- ANDRADE, L. O. M. **Sistemas integrados de saúde, redes e linhas de Cuidado: construção de modelagem para estudo Comparado Brasil Canadá**. Projeto Pós-Doutoral na Universidade de Quebec. [S.l.]. 2009.
- ARAUJO, P. H. M.; NEHME, C. C. **Uso de modelos de incerteza no controle ambiental do processo erosivo na bacia do Alto Taquari**. In Anais do III Brazilian Symposium on GeoInformatics. Rio de Janeiro: In Anais do III Brazilian Symposium on GeoInformatics. 2001. p. 112.
- BASTOS, P. R. F. D. M.; FERREIRA, N. R.; SOUZA, B. A. D. **Diagnóstico das perdas não técnicas empregando redes Bayesianas**. Cidel Argentina. [S.l.]: [s.n.]. 2010.
- BRIGNOLLI, J. **Modelo Híbrido Difuso Probabilístico**. Dissertação de Mestrado, Programa de Pós-Graduação em Ciência da Computação. [S.l.]. 2001.
- CHEN, D. **Sisterhood of Classifiers: A Comparative Study of Naive Bayes and Noisy-or Networks**. University of California. [S.l.]. 2000.
- CHENG, J.; GREINER, R. **Learning Bayesian Belief Network Classifiers: Algorithms and System**. Proceedings of 14th Biennial Conference of the Canadian Society for Computational Studies of Intelligence. [S.l.]: [s.n.]. 2001.
- COWELL ET AL, R. G. **Probabilistic Networks and Expert Systems**. New York: [s.n.], 1999. 5-23 p.
- COIERA, E. **Guide to: Medical Informatics, The Internet and Telemedicine**. **Guide to: Medical Informatics, The Internet and Telemedicine**, 2007. Disponível em: <<http://www.coiera.com/>>. Acesso em: 23 nov. 2012.
- DEY, A. K.; SALBER, D.; ABOWD, G. D. **A Conceptual Framework and a Toolkit for Supporting the Rapid Prototyping of Context-Aware Applications**. **Article for Special Issue on Context-Awareness**, 2001. 97–166.
- FLORES, C. D.; HÖHER, C. L. **Uma Experiência do Uso de Redes Probabilísticas no Diagnóstico Médico**. Argentine Symposium on Artificial Intelligence. Porto Alegre: [s.n.]. 2000.
- FLORES, C. D.; PEROTTO, F.; VICARI, R. M. **Sistemas Baseados em Conhecimento para a Área da Saúde**. Porto Alegre: [s.n.]. 2002.

GARDINI, L. M. et al. **Clariisa, a Context-Aware Framework Based on Geolocation for a Health Care Governance System**. 5TH International Conference on e-Health Networking, Application & Services. [S.l.]: [s.n.]. 2013.

GESZYCHTER, M. B. **Modelagem Multiagente escalável para Integração de Sistemas Especialistas**. Dissertação de Mestrado, Programa de Pós-Graduação em Ciência da Computação, Depto. de Informática e de Estatística, UFSC. [S.l.]. 2003.

HUANG, W. R. et al. Schedule of review of waste lubricant recycling system, 2006.

HECKERMAN, D.; BREESE, J. **A new look at causal independence**. roceedings of the 10th Conference on Uncertainty in Artificial Intelligence. [S.l.]: Morgan Kaufmann. 1994. p. 226-286.

HECKERMAN, D.; GEIGER, D.; CHICKERING, D. Learning Bayesian networks: The combination of knowledge and statistical data, 20, 1995. 197-243.

HECKERMAN, D.; M. WELLMAN. **Bayesian Networks**. Comm. ACM. [S.l.]: [s.n.]. 1995. p. 27-30.

JENSEN, F. V. **Bayesian Networks and Decision Graphs**. New York: Springer, 2001. 268 p.

KÖEHLER, C. **Uma Abordagem Bayesiana para Sistemas Especialistas**. Dissertação de Mestrado, Programa de Pós-Graduação em Ciência da Computação. [S.l.]. 1998.

KOEHLER, C. **Desenvolvimento de um Sistema Inteligente para Apoio à Decisão em Saúde**. [S.l.]: [s.n.]. 2002.

LAURIA, E. J. M.; DUCHESSI, P. J. A methodology for developing bayesian networks: an application to information technology (IT) implementation. **European Journal of Operational Research**, NY, v. 179, n. 1, p. 234- 252, maio 2007.

LOPES, L. F.; GONÇALVES, L.; TODESCO, L. Um Modelo De Engenharia Do Conhecimento Baseado Em Ontologia E Cálculo Probabilístico Para Apoio Ao Diagnóstico. **Revista Eletrônica Sistemas & Gestão** , Florianópolis, 2011. 272-293.

LOSI, N. A. **ABS-Bayes: Sistema Especialista Probabilístico para Apoio ao Diagnóstico de Absenteísmo**. Dissertação de Mestrado, Programa de Pós- Graduação em Computação, Depto. de Informática e de Estatística, UFSC. [S.l.]. 2003.

NUNES, G. **Abordagem bayesiana na epidemiologia espacial de notificações de casos de dengue no Estado de Goiás**. Simpósio Brasileiro de Sensoriamento Remoto. Foz do Iguaçu: [s.n.]. 2013.

NEAPOLITAN, R. E. **Learning Bayesian Networks**. New Jersey : [s.n.], 2004. 674 p.

NIKOVSKI, D. **Constructing Bayesian networks for medical diagnosis from incomplete and partially correct statistics**. IEEE Transactions on Knowledge and Data Engineering. [S.l.]: [s.n.]. 2000. p. 509-516.

NÓBREGA, R. D. D. S. Sistema de decisão mista, Agosto 2004. Disponível em: <<http://www.dei.isep.ipp.pt/~paf/proj/Set2004/SistemasDecisaoMista.pdf>>. Acesso em: 22 Agosto 2013.

MATOS, N. P. D. **Sistema de Apoio à Decisão para Planejamento em Saúde**. Paraná: [s.n.]. 2003.

MCFARIANE, S. A.; EVANS, K. F.; ACKERMAN, A. S. **Liquid water cloud retrievals – a bayesian approach**. Eleventh ARM Science Team Meeting. Atlanta: Eleventh ARM Science Team Meeting. 2001. p. 8.

MENDES, R. S. **Abordagem Bayesiana para Redes sem Fio Cognitivas**. Dissertação de Mestrado, Programa de Pós-Graduação em Ciência da Computação. [S.l.]. 2010.

MINISTÉRIO DA SAÚDE. Programa Nacional de Controle da Dengue. **Portal da Saúde**, 2012. Disponível em: <http://portal.saude.gov.br/portal/saude/visualizar_texto.cfm?idtxt=23614>. Acesso em: 22 ago. 2013.

OLIVEIRA, M. et al. **A context-aware framework for health care governance decision-making systems: A model based on the Brazilian Digital TV**. IEEE INTERNATIONAL SYMPOSIUM ON A WORLD OF WIRELESS MOBILE AND MULTIMEDIA NETWORKS. Montreal: Anais. 2010. p. 1-6.

PAULINO, C. D.; TURKMAN, M. A. A.; MURTEIRA, B. **Estatística bayesiana**. 1. ed. Lisboa: Fundacao Caloustre Gulbenkian, v. 1, 2003.

PASINI, H. **Sistema Especialista Probabilístico para Apoio ao Diagnóstico de Potencial Econômico - SEPE**. Dissertação de Mestrado, Programa de Pós- Graduação em Ciência da Computação, Depto. de Informática e de Estatística, UFSC. [S.l.]. 2002.

PEARL, J. **Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference**. San Mateo: Morgan Kaufmann, 1988.

PENZ, C. A. **Procedimentos para Prover Confiabilidade ao Uso de Inteligência Artificial em Ensaios de Desempenho de Compressores Herméticos de Refrigeração**. Tese de Doutorado, Programa de Pós-Graduação em Engenharia Mecânica. [S.l.]. 2011.

PENA, S. D. Bayes: o “cara”! **Revista Ciência Hoje**, v. 38, p. 22-29, 2006.

PESETTE, R. S.; VIEIRA, K. M. M. Redes bayesianas no diagnostico medico. **Redes bayesianas no diagnostico medico**, 2002. Disponível em: <http://www.inf.ufsc.br/~barreto/trabaluno/IA_kleber_Renato_RB.pdf>. Acesso em: 17 set. 2012.

SAHEKI, H. **Construção de uma rede bayesiana aplicada ao diagnóstico de doenças cardíacas**. Escola Politécnica da Universidade de São Paulo. [S.l.]. 2005.

SAMPAIO, L. M. D. **Análise e Classificação das Operadoras de Saúde Suplementar**. UFRJ. [S.l.]. 2008.

SIMÕES, P. W. T. A. **Sistema de Apoio na Avaliação da Falência do Crescimento Infantil – SACI**. Dissertação de Mestrado, Programa de Pós-Graduação em Ciência da Computação. [S.l.]. 2001.

SPIEGELHALTER, D. J. Probabilistic expert systems in medicine. **Statistical Science**, v. 2, p. 3-44, 1987.

STEIN, C. **Sistema Especialista Probabilístico: base de conhecimento dinâmica**. Dissertação de Mestrado, Programa de Pós-Graduação em Ciência da Computação. [S.l.]. 2000.

RUSSELL, S.; NORVIG, P. **Inteligência Artificial**. 2. ed. Rio de Janeiro : Elsevier, 2004. 480, 567, 579, 580 p.

ROVARIS , E. N. **Sistema Especialista de Apoio à Gestão da Evasão Escolar – E- Bayes**. Dissertação de Mestrado, Programa de Pós-Graduação em Ciência da Computação. [S.l.]. 2002.

TIBIRIÇÁ, C. A. **Uma Abordagem Híbrida Fuzzy-Bayesiana para Modelagem de Incertezas**. Dissertação de Mestrado, Programa de Pós-Graduação em Ciência da Computação, Depto. de Informática e de Estatística. [S.l.]. 2005.

TOLEDO, R. V. **Sistema de Apoio ao Diagnóstico Diferencial de Cefaléia**. Dissertação de Mestrado, Programa de Pós-Graduação em Ciência da Computação. [S.l.]. 2000.

VENICIUS, M. **Agente Bayesiano de Apoio à Vigilância de Infecção Hospitalar - SAVIH**. Dissertação de Mestrado, Programa de Pós-Graduação em Ciência da Computação. [S.l.]. 2002.

VERONEZ, C. **Baseline Bayesiana para Gerência de Tráfego em Redes de Computadores**. Dissertação de Mestrado, Programa de Pós-Graduação em Ciência da Computação. [S.l.]. 2000.

VIANA, W. A semantic approach and a web tool for contextual annotation of photos using camera phones. In **WISE**, 2007. 225– 236.

WASSON, J. H.; SOX , H. C.; NEFF , R. K. clinical prediction rules. applications and methodological standards. **N Engl J Med.**, 1985.

WETTER, T. E. Lessons learnt from bringing knowledge- based decision support into routine use. **Artificial Intelligence in Medicine** , v. 24, p. 195-203, 2002.

YUDKOWSKY, E. S. An Intuitive Explanation of Bayes' Theorem. **Eliezer S. Yudkowsky Research Fellow - Singularity Institute for Artificial Intelligence**, 2009. Disponível em: <<http://yudkowsky.net/rational/bayes>>. Acesso em: 02 dez. 2012.

ZHANG, N.; POOLE , D. Exploiting Causal Independence in Bayesian Network Inference. **Journal of Artificial Intelligence Research** , v. 5, p. 301-328 , 1996.

APÊNDICE A

```
import norsys.netica.*;

public class InferenceForDengue_Comum {

    public static void main (String[] args){
        try {
            Environ env = new Environ (null);

            // Read in the net created by the BuildNet.java example program
            Net net = new Net (new Streamer ("Data Files/LARISSA.dne"));

            LeituraDoArquivo leituraDoArquivo = new LeituraDoArquivo();

            Node queda_abrupta_de_plaquetas      = net.getNode("queda_abrupta_de_plaquetas");

            Node diminuicao_da_diurese            = net.getNode( "diminuicao_da_diurese"      );

            Node coleta_de lixo_precaria          = net.getNode( "coleta_de_lixo_precaria"    );

            Node sonolencia_e_ou_irritabilidade   = net.getNode(
"sonolencia_e_ou_irritabilidade" );

            Node aum_repentino_do_hematocrito     = net.getNode( "aum_repentino_do_hematocrito" );

            Node locais_sem_saneamento_adequado  = net.getNode(
"locais_sem_saneamento_adequado" );

            Node dengue_comum                     = net.getNode( "dengue_comum" );

            net.compile();

            double belief = dengue_comum.getBelief ("sim");
            System.out.println ("\n A probabilidade de dengue é " + belief);

            if(leituraDoArquivo.validaItem()){
                queda_abrupta_de_plaquetas.finding().enterState ("sim");
                belief = dengue_comum.getBelief ("sim");
                System.out.println ("\n Dado que possui um queda abrupta de
palquetas,\n"+
                                "A probabilidade de dengue é " + belief);
            }

            if(leituraDoArquivo.validaItem()){
                diminuicao_da_diurese.finding().enterState ("sim");
                belief = dengue_comum.getBelief ("sim");
                System.out.println ("\nDado que possui um queda abrupta de palquetas e
diminuicao_da_diurese,\n" +
                                "A probabilidade de dengue é " + belief + "\n");
            }

            if(leituraDoArquivo.validaItem()){
                coleta_de_lixo_precaria.finding().enterState ("sim");
                belief = dengue_comum.getBelief ("sim");
                System.out.println ("\n Dado que possui um queda abrupta de
palquetas,\n"+
                                "A probabilidade de dengue é " + belief);
            }

            if(leituraDoArquivo.validaItem()){
                sonolencia_e_ou_irritabilidade.finding().enterState ("sim");
                belief = dengue_comum.getBelief ("sim");
                System.out.println ("\nDado que possui uma sonolencia_e_ou_irritabilidade e
diminuicao_da_diurese,\n" +
                                "A probabilidade de dengue é " + belief + "\n");
            }

            if(leituraDoArquivo.validaItem()){
                aum_repentino_do_hematocrito.finding().enterState ("sim");
                belief = dengue_comum.getBelief ("sim");
                System.out.println ("\n Dado que possui um queda abrupta de
```

```

palquetas,\n"+
                                "A probabilidade de dengue é " + belief);
    }

    if(leituraDoArquivo.validaItem()){
        locais_sem_saneamento_adequado.finding().enterState ("sim");
        belief = dengue_comum.getBelief ("sim");
        System.out.println ("\nDado que possui um queda abrupta de palquetas e
diminuicao_da_diurese,\n" +
                                "A probabilidade de dengue é " + belief + "\n");
    }

    net.finalize(); // not strictly necessary, but a good habit
}
catch (Exception e) {
    e.printStackTrace();
}
}
}

import norsys.netica.*;

public class InferenceForDengue_Hemorragica {

    public static void main (String[] args){
        try {
            Environ env = new Environ (null);
            LeituraDoArquivo leituraDoArquivo = new LeituraDoArquivo();
            // Read in the net created by the BuildNet.java example program
            Net net = new Net (new Streamer ("Data Files/LARISSA.dne"));

            boolean checaRelatorio = false;

            Node queda_abrupta_de_plaquetas = net.getNode("queda_abrupta_de_plaquetas");

            Node diminuicao_da_diurese = net.getNode( "diminuicao_da_diurese" );

            Node coleta_de_lixo_precaria = net.getNode( "coleta_de_lixo_precaria" );

            Node sonolencia_e_ou_irritabilidade = net.getNode(
"sonolencia_e_ou_irritabilidade" );

            Node aum_repentino_do_hematocrito = net.getNode( "aum_repentino_do_hematocrito"
);

            Node locais_sem_saneamento_adequado = net.getNode(
"locais_sem_saneamento_adequado" );

            Node dor_abdominal_intensa_continua = net.getNode(
"dor_abdominal_intensa_continua" );
            Node desconforto_respiratorio = net.getNode( "desconforto_respiratorio" );
            Node vomitos_persistentes = net.getNode( "vomitos_persistentes" );
            Node sangramento_de_mucosa = net.getNode( "sangramento_de_mucosa" );

            Node dengue_hemorragica = net.getNode( "dengue_hemorragica" );
            Node alteracao_na_Sorologia = net.getNode( "alteracao_na_Sorologia" );
            Node alteracao_no_Hemograma = net.getNode( "alteracao_no_Hemograma" );
            Node Hepatomegalia_dolorosa = net.getNode( "Hepatomegalia_dolorosa" );
            Node Hipotermia = net.getNode( "Hipotermia" );
            Node Hipotensao_postural = net.getNode( "Hipotensao_postural" );

            net.compile();

            double belief = dengue_hemorragica.getBelief ("sim");
            System.out.println ("\n A probabilidade de dengue é " + belief);

            if(leituraDoArquivo.validaItem()){
                queda_abrupta_de_plaquetas.finding().enterState ("sim");
                belief = dengue_hemorragica.getBelief ("sim");
                System.out.println ("\n Dado que possui um queda abrupta de

```



```

palquetas,\n"+
                                "A probabilidade de dengue é " + belief);
    }
    if(leituraDoArquivo.validaItem()){
        aum_repentino_do_hematocrito.finding().enterState ("sim");
        belief = dengue_hemorragica.getBelief ("sim");
        System.out.println ("\n Dado que possui
aum_repentino_do_hematocrito,\n"+
                                "A probabilidade de dengue é " + belief);
    }
    if(leituraDoArquivo.validaItem()){
        locais_sem_saneamento_adequado.finding().enterState ("sim");
        belief = dengue_hemorragica.getBelief ("sim");
        System.out.println ("\n Dado que possui um
locais_sem_saneamento_adequado,\n"+
                                "A probabilidade de dengue é " + belief);
    }
    if(leituraDoArquivo.validaItem()){
        vomitos_persistentes.finding().enterState ("sim");
        belief = dengue_hemorragica.getBelief ("sim");
        System.out.println ("\n Dado que possui vomitos_persistentes,\n"+
                                "A probabilidade de dengue é " + belief);
    }
    if(leituraDoArquivo.validaItem()){
        Hepatomegalia_dolorosa.finding().enterState ("sim");
        belief = dengue_hemorragica.getBelief ("sim");
        System.out.println ("\n Dado que possui uma
Hepatomegalia_dolorosa,\n"+
                                "A probabilidade de dengue é " + belief);
    }
    if(leituraDoArquivo.validaItem()){
        sangramento_de_mucosa.finding().enterState ("sim");
        belief = dengue_hemorragica.getBelief ("sim");
        System.out.println ("\n Dado que possui um sangramento_de_mucosa,\n"+
                                "A probabilidade de dengue é " + belief);
    }
    if(leituraDoArquivo.validaItem()){
        Hipotensao_postural.finding().enterState ("sim");
        belief = dengue_hemorragica.getBelief ("sim");
        System.out.println ("\n Dado que possui uma Hipotensao_postural,\n"+
                                "A probabilidade de dengue é " + belief);
    }
    if(leituraDoArquivo.validaItem()){
        Hipotermia.finding().enterState ("sim");
        belief = dengue_hemorragica.getBelief ("sim");
        System.out.println ("\n Dado que possui uma Hipotermia,\n"+
                                "A probabilidade de dengue é " + belief);
    }
    if(leituraDoArquivo.validaItem()){
        alteracao_no_Hemograma.finding().enterState ("sim");
        belief = dengue_hemorragica.getBelief ("sim");
        System.out.println ("\n Dado que possui uma
alteracao_no_Hemograma,\n"+
                                "A probabilidade de dengue é " + belief);
    }
    if(leituraDoArquivo.validaItem()){
        alteracao_na_Sorologia.finding().enterState ("sim");
        belief = dengue_hemorragica.getBelief ("sim");
        System.out.println ("\n Dado que possui uma
alteracao_na_Sorologia,\n"+
                                "A probabilidade de dengue é " + belief);
    }
    if(leituraDoArquivo.validaItem()){
        desconforto_respiratorio.finding().enterState ("sim");
        belief = dengue_hemorragica.getBelief ("sim");
        System.out.println ("\n Dado que possui um
desconforto_respiratorio,\n"+
                                "A probabilidade de dengue é " + belief);
    }
    if(leituraDoArquivo.validaItem()){
        dor_abdominal_intensa_continua.finding().enterState ("sim");
        belief = dengue_hemorragica.getBelief ("sim");
        System.out.println ("\n Dado que possui um
dor_abdominal_intensa_continua,\n"+
                                "A probabilidade de dengue é " + belief);
    }
    if(leituraDoArquivo.validaItem()){

```

```

        sonolencia_e_ou_irritabilidade.finding().enterState ("sim");
        belief = dengue_hemorragica.getBelief ("sim");
        System.out.println ("\n Dado que possui uma
sonolencia_e_ou_irritabilidade,\n"+
                                "A probabilidade de dengue é " + belief);
    }
    if(leituraDoArquivo.validaItem()){
        diminuicao_da_diurese.finding().enterState ("sim");
        belief = dengue_hemorragica.getBelief ("sim");
        System.out.println ("\n Dado que possui uma diminuicao_da_diurese,\n"+
                                "A probabilidade de dengue é " + belief);
    }
    if(leituraDoArquivo.validaItem()){
        coleta_de_lixo_precaria.finding().enterState ("sim");
        belief = dengue_hemorragica.getBelief ("sim");
        System.out.println ("\n Dado que possui uma
coleta_de_lixo_precaria,\n"+
                                "A probabilidade de dengue é " + belief);
    }

    net.finalize(); // not strictly necessary, but a good habit
}
catch (Exception e) {
    e.printStackTrace();
}
}
}

```

APÊNDICE B

```
// ~->[DNET-1]->~

// File created by an unlicensed user using Netica 5.04 on 10/02/13 at 10:15:21.

bnet Lariisa_bay {
  autoupdate = FALSE;

  node Queda_abrupta_de_plaquetas {
    kind = NATURE;
    discrete = TRUE;
    chance = CHANCE;
    states = (nao, sim);
    parents = ();
    probs =
      // nao      sim
      (0.99,      0.01);
    title = "Queda abrupta de plaquetas";
  };

  node Dor_abdominal_intensa_continua {
    kind = NATURE;
    discrete = TRUE;
    chance = CHANCE;
    states = (nao, sim);
    parents = ();
    probs =
      // nao      sim
      (0.95,      0.05);
    title = "Dor abdominal intensa continua";
  };

  node Desconforto_respiratorio {
    kind = NATURE;
    discrete = TRUE;
    chance = CHANCE;
    states = (nao, sim);
    parents = ();
    probs =
      // nao      sim
      (0.75,      0.25);
    title = "Desconforto respiratorio";
  };

  node Vomitos_persistentes {
    kind = NATURE;
    discrete = TRUE;
    chance = CHANCE;
    states = (nao, sim);
    parents = ();
    probs =
      // nao      sim
      (0.95,      0.05);
    title = "Vomitos persistentes";
  };

  node Sangramento_de_mucosa {
    kind = NATURE;
    discrete = TRUE;
    chance = CHANCE;
    states = (nao, sim);
    parents = ();
    probs =
      // nao      sim
      (0.99,      0.01);
    title = "Sangramento de mucosa";
  };

  node Sonolencia_e_ou_irritabilidade {
    kind = NATURE;
```

```

discrete = TRUE;
chance = CHANCE;
states = (nao, sim);
parents = ();
probs =
    // nao      sim
    (0.99,      0.01);
title = "Sonolencia e/ou irritabilidade";
};

node Locais_sem_saneamento_adequado {
    kind = NATURE;
    discrete = TRUE;
    chance = CHANCE;
    states = (nao, sim);
    parents = ();
    probs =
        // nao      sim
        (0.7,       0.3);
    title = "Locais sem saneamento adequado";
};

node Coleta_de lixo_precaria {
    kind = NATURE;
    discrete = TRUE;
    chance = CHANCE;
    states = (nao, sim);
    parents = ();
    probs =
        // nao      sim
        (0.65,      0.35);
    title = "Coleta de lixo precaria";
};

node Diminuicao_da_diurese {
    kind = NATURE;
    discrete = TRUE;
    chance = CHANCE;
    states = (nao, sim);
    parents = ();
    probs =
        // nao      sim
        (0.7,       0.3);
    title = "Diminuicao da diurese";
};

node Aum_repentino_do_hematocrito {
    kind = NATURE;
    discrete = TRUE;
    chance = CHANCE;
    states = (nao, sim);
    parents = ();
    probs =
        // nao      sim
        (0.75,      0.25);
    title = "Aumento repentino do hematocrito";
};

node Dengue_hemorragica {
    kind = NATURE;
    discrete = TRUE;
    states = (nao, sim);
    parents = (Dor_abdominal_intensa_continua, Desconforto_respiratorio,
Vomitos_persistentes, Sangramento_de_mucosa, Sonolencia_e_ou_irritabilidade,
Locais_sem_saneamento_adequado, Coleta_de_lixo_precaria, Queda_abrupta_de_plaquetas,
Diminuicao_da_diurese, Aum_repentino_do_hematocrito);
    title = "Dengue Hemorragica";
};

node Dengue_comum {
    kind = NATURE;
    discrete = TRUE;
    states = (nao, sim);
    parents = (Sonolencia_e_ou_irritabilidade, Locais_sem_saneamento_adequado,
Coleta_de_lixo_precaria, Queda_abrupta_de_plaquetas, Diminuicao_da_diurese,
Aum_repentino_do_hematocrito);
    title = "Dengue comum";
};

```

```

    };

    node Alteracao_na_Sorologia {
        kind = NATURE;
        discrete = TRUE;
        states = (nao, sim);
        parents = (Dengue_hemorragica, Dengue_comum);
        title = "Alteracao na Sorologia";
    };

    node Alteracao_no_Hemograma {
        kind = NATURE;
        discrete = TRUE;
        states = (nao, sim);
        parents = (Dengue_hemorragica, Dengue_comum);
        title = "Alteracao no Hemograma";
    };
};

```