

UNIVERSIDADE FEEVALE

FERNANDO ALEX HELWANGER

UM EDITOR DE REDES BAYESIANAS COM FOCO EM
USABILIDADE
(Título Provisório)

Anteprojeto de Trabalho de Conclusão

Novo Hamburgo
2016

FERNANDO ALEX HELWANGER

UM EDITOR DE REDES BAYESIANAS COM FOCO EM
USABILIDADE

(Título Provisório)

Anteprojeto de Trabalho de Conclusão de
Curso, apresentado como requisito parcial
à obtenção do grau de Bacharel em
Ciência da Computação pela
Universidade Feevale

Orientador: Paulo Ricardo Muniz Barros

Coorientador: Alessandro Peixoto de Lima

Novo Hamburgo
2016

RESUMO

Existem problemas que por natureza possuem resultados incertos, nos quais as variáveis envolvidas definem a probabilidade de determinado evento acontecer. Um exemplo é o processo de diagnóstico médico, no qual os sintomas do paciente definem a probabilidade da existência de determinada enfermidade. Para este tipo de problema, redes bayesianas podem ser empregadas como uma maneira simples de representação. Com o auxílio desta rede é possível simular diferentes resultados alterando as variáveis de entrada, o que se mostra bastante útil para simuladores no ensino. Porém, os softwares para edição de redes bayesianas existentes mostram-se com uma usabilidade aquém do necessário para boa parte de seus usuários. Como redes bayesianas podem ser empregadas em diversos domínios, os usuários dos softwares para edição nem sempre possuem uma grande experiência com computação e acabam encontrando dificuldades na criação destas redes. O objetivo deste trabalho é fazer uma avaliação da usabilidade dos softwares para edição de redes bayesianas existentes e, a partir dos resultados dessa avaliação, desenvolver um novo software que supra a necessidade de seus usuários sem que eles sejam afetados por uma baixa usabilidade.

Palavras-chave: Redes bayesianas. Usabilidade. Editor.

SUMÁRIO

MOTIVAÇÃO	5
OBJETIVOS	8
METODOLOGIA	9
CRONOGRAMA	10
BIBLIOGRAFIA	11

MOTIVAÇÃO

Existem diversos problemas que devem considerar um nível de incerteza em sua resolução, por possuírem amostras imprecisas ou teorias incompletas. Para este tipo de problema, podem ser usados cálculos probabilísticos, que trazem resultados com um determinado grau de confiança (RUSSEL; NORVIG, 1995).

Para modelar este tipo de problema, é necessário conhecer quais variáveis impactam o domínio observado. O trabalho de Russel e Norvig (1995) mostra um exemplo no qual um dentista recebe um paciente que pode possuir dor de dente e alguma cavidade. Tanto a variável “dor de dente” quanto a variável “cavidade” possuem um domínio booleano. Este problema pode ser modelado criando-se uma tabela que considere todas as combinações de estados possíveis, atribuindo uma probabilidade para que cada uma ocorra.

Dada esta tabela, chamada de *full joint distribution*, é possível responder diversas questões: Qual a probabilidade do paciente possuir uma cavidade caso possua dor de dente? Qual a probabilidade do paciente possuir cavidade ou possuir dor de dente? Porém, problemas reais costumam ter um número grande de variáveis, o que torna esta abordagem bastante trabalhosa e muitas vezes inviável. Considerando apenas variáveis booleanas, o número de estados a serem considerados é 2^n , sendo n o número de variáveis.

Em modelos complexos nem todas as variáveis têm impacto direto nas outras. Algumas variáveis podem ser completamente independentes ou condicionalmente independentes de outras (RUSSEL; NORVIG, 1995). Estes modelos podem ser representados por redes bayesianas.

Redes bayesianas são grafos acíclicos dirigidos (em inglês, *directed acyclic graph*) nos quais cada nodo representa uma variável e as arestas podem ser interpretadas como relações de causa e efeito. Este raciocínio de causa e efeito assemelha-se a maneira que humanos especialistas em alguma área raciocinam (NILSSON, 1998). Por exemplo, “cavidade” é a causa do efeito “dor de dente”.

Por aproveitar a propriedade de independência condicional entre as variáveis, a representação de redes bayesianas é mais simples e menor do que a *full joint distribution*. Essa redução de tamanho pode tornar problemas que seriam inviáveis realizáveis (NILSSON, 1998).

A construção de redes bayesianas envolve duas partes: uma qualitativa e outra quantitativa. A parte qualitativa é a estrutura do grafo que representa a rede. A parte quantitativa

pode ser representada pelas tabelas de probabilidade condicional associadas a cada variável. Essa especificação pode ser inferida de dados, construída manualmente por um especialista ou uma mistura das duas alternativas (WIEGERINCK; KAPPEN; BURGERS, 2010).

O processo de diagnóstico médico é um problema que envolve raciocínio probabilístico, no qual o paciente possui uma série de sinais e sintomas que indicam uma probabilidade de ele possuir uma certa enfermidade (MARQUES; DUTRA, 2002). O uso de software que interprete estes sintomas e calcule as probabilidades pode ser utilizado para auxiliar o diagnóstico, tanto na prática quanto no ensino (SEIXAS et al., 2002).

No ensino da medicina, o aluno precisa praticar a construção de modelos hipotéticos que relacionem doenças com suas causas e sintomas. É preciso que ele possa avaliar a solução e tomar ações no decorrer deste processo. Adotando uma abordagem construtivista de aprendizagem, redes bayesianas podem ser aplicadas. Na teoria construtivista, o conhecimento deve ser construído por cada indivíduo, e não transmitido de uma pessoa para outra (SEIXAS et al., 2002). Neste contexto, o uso de um software no qual o aluno possa simular situações reais e testar suas ações em tempo real mostra-se bastante útil.

Atualmente, existem projetos que utilizam redes bayesianas com este intuito. Alguns dos projetos estudados para este trabalho foram o AMPLIA (SEIXAS, 2005), o SimDeCS (BARROS et al., 2012) e o Health Simulator (LIMA et al., 2015). A seguir eles serão brevemente apresentados, bem como a tecnologia utilizada para empregar o uso de redes bayesianas.

O AMPLIA é um projeto no qual o aluno representa seu raciocínio de diagnóstico por meio de uma rede bayesiana e este raciocínio é avaliado por agentes inteligentes (SEIXAS, 2005). O editor de redes bayesianas do AMPLIA é baseado no editor gráfico para redes bayesianas SEAMED, aplicação desktop desenvolvida em Delphi 6.0 (FLORES, 2005).

O SimDeCS é um simulador no qual um especialista constrói uma rede bayesiana baseada em uma diretriz clínica. Após isso, o professor monta um caso clínico que utiliza o conhecimento modelado nesta rede bayesiana. Posteriormente, na forma de um jogo sério, o aluno de medicina realiza um atendimento a um paciente virtual conforme descrito neste caso clínico. Este processo é avaliado por agentes inteligentes na forma de uma negociação

pedagógica (FONSECA, 2013). Neste projeto, a inferência de redes bayesianas é feita pelo UnBBayes¹ e a edição delas é feita pelo software Hugin².

O Health Simulator é um software do tipo paciente virtual, com o intuito de auxiliar no ensino na medicina. A sua estruturação quanto a modelagem do problema é semelhante ao SimDeCS, no qual um especialista cria uma rede bayesiana que é utilizada pelo professor na montagem de um caso clínico para a simulação. Porém, segundo Lima et al. (2015, p. 286),

a experiência adquirida no desenvolvimento do SIACC e SimDeCS mostrou que era necessário desenvolver um simulador mais realístico, que forneça mais liberdade ao aluno para testar suas hipóteses diagnósticas e experimentar na prática o que este encontraria na área da saúde.

Conforme relatado por Maroni (2013), o fato do aluno usuário do AMPLIA precisar utilizar um editor de redes bayesianas para modelar suas hipóteses muitas vezes atrapalha o objetivo da aprendizagem, pois ele encontra dificuldades de utilização. Isto é um indicativo de que o editor utilizado no AMPLIA possui problemas de usabilidade.

Para Reiss (2015, p. xvii, tradução nossa), “Usabilidade lida com a habilidade de um indivíduo realizar uma tarefa específica ou meta mais ampla enquanto “usa” o que você estiver investigando, melhorando ou desenvolvendo”. Caso este indivíduo não consiga atingir seu objetivo, ou tiver dificuldades durante o processo, caracteriza-se um problema de usabilidade.

Apesar de o SimDeCS ter removido do aluno a tarefa de edição de redes, o especialista na área ainda precisa criá-las e editá-las. Como a montagem do caso clínico e a simulação do atendimento médico dependem dessa edição, é essencial que ela esteja correta.

Como o Health Simulator mantém o conceito de modelagem do conhecimento do especialista utilizando redes bayesianas, surge a necessidade de um editor com foco em usabilidade, para que possa ser utilizado por especialistas que não possuam necessariamente um conhecimento aprofundado em computação.

¹ O UnBBayes é um framework para montagem e inferência em redes bayesianas desenvolvido pela Universidade de Brasília (SANTOS, 2009).

² O Hugin é uma ferramenta de propósito geral para modelos probabilísticos gráficos (MADSEN, 2003).

OBJETIVOS

Objetivo geral

Criar um editor de redes bayesianas com foco em usabilidade que possa ser empregado em diferentes aplicações de diferentes domínios.

Objetivos específicos

- Realizar pesquisa bibliográfica sobre redes bayesianas;
- Realizar pesquisa bibliográfica sobre conceitos de usabilidade;
- Realizar pesquisa sobre editores de redes bayesianas existentes;
- Identificar problemas de usabilidade em editores de redes bayesianas existentes;
- Especificar um novo editor que atenda aos problemas identificados;
- Desenvolver um novo editor conforme a especificação;
- Validar parcialmente os conceitos aplicados no novo editor.

METODOLOGIA

Quanto à natureza, este trabalho caracteriza-se como pesquisa aplicada, pois tem como objetivo gerar conhecimento para resolver um problema específico envolvendo verdades e interesses locais (PRODANOV; FREITAS, 2013).

Quanto ao objetivo, fica caracterizado como pesquisa exploratória (PRODANOV; FREITAS, 2013), pois serão levantados dados para apontar quais melhorias podem ser feitas nas soluções existentes e, a partir desses dados, será desenvolvida uma alternativa. Isso caracteriza que a pesquisa está em fase preliminar e, a partir de estudos, será feita a definição do que será preciso desenvolver para atender aos requisitos necessários.

Os procedimentos técnicos utilizados serão pesquisa bibliográfica e pesquisa experimental. A pesquisa bibliográfica tem como fonte de conhecimento materiais publicados, como livros, revistas, publicações em periódicos e artigos científicos, monografias, dissertações, teses, entre outros. A pesquisa experimental é o tipo de pesquisa na qual as variáveis que podem influenciar o objeto de estudo são manipuladas e os resultados são observados (PRODANOV; FREITAS, 2013).

O trabalho terá início através de pesquisas bibliográficas a respeito de redes bayesianas e usabilidade. A partir do conhecimento obtido nessas pesquisas, serão levantadas melhorias que podem ser aplicadas nas soluções existentes de editores de redes bayesianas. Dada essa definição, um novo editor será desenvolvido e seus resultados serão avaliados.

A validação será realizada por meio de testes e pesquisa com usuários dos editores de redes bayesianas. Estes usuários irão utilizar tanto as ferramentas existentes estudadas quanto a nova que será desenvolvida. Durante o uso, eles serão observados presencialmente. Além disso, eles responderão um questionário no qual serão informados resultados pertinentes à usabilidade de cada editor. Com base nesses dados, será feita uma avaliação utilizando uma abordagem qualitativa.

CRONOGRAMA

Trabalho de Conclusão I

Etapa	Meses			
	Mar	Abr	Mai	Jun
Escrita do anteprojeto	■	■		
Pesquisa bibliográfica sobre redes bayesianas	■	■		
Pesquisa bibliográfica sobre usabilidade	■	■		
Pesquisa sobre trabalhos correlatos		■	■	
Levantamento de problemas nos editores existentes			■	■
Escrita do TC1		■	■	■

Trabalho de Conclusão II

Etapa	Meses			
	Ago	Set	Out	Nov
Especificação do novo editor	■	■		
Desenvolvimento do novo editor		■	■	
Validação parcial do novo editor			■	■
Escrita do TC2		■	■	■

BIBLIOGRAFIA

- BARROS, Paulo Ricardo Muniz et al. Um simulador de casos clínicos complexos no processo de aprendizagem em saúde. **RENOTE - Revista Novas Tecnologias na Educação**, Porto Alegre, v. 10, n. 1, p. 234, jun. 2012.
- BONDY, John Adrian; MURTY, Uppaluri Siva Ramachandra. **Graph theory with applications**. London: Macmillan, 1976.
- CHARNIAK, Eugene. Bayesians Networks without Tears. **AI Magazine**, Palo Alto, v. 12, n. 4, p. 50. 1991.
- FAGUNDES, Moser Silva. **Integrating BDI model and Bayesian networks**. 2007. 104 f. Dissertação (Mestrado em Ciência da Computação) – Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, 2007.
- FLORES, Cecília Dias. **Negociação pedagógica aplica a um ambiente multiagente de aprendizagem colaborativa**. 2005. 121 f. Tese (Doutorado em Ciência da Computação) – Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, 2005.
- FLORES, Cecília Dias et al. **Negociação pedagógica no ambiente AMPLIA**. In CONGRESSO DA SOCIEDADE BRASILEIRA DE COMPUTAÇÃO, v. 23, p. 617-626. 2003.
- FONSECA, João Marcelo Lopes. **Descrição de um simulador baseado em redes bayesianas como método de avaliação do aprendizado de diretrizes clínicas em ensino a distância para medicina de família e comunidade**. 2013. 118 f. Dissertação (Mestrado em Ciências da Saúde) – Universidade Federal de Ciências da Saúde de Porto Alegre, Porto Alegre, RS, 2013.
- GARRETT, Jesse James. **The elements of user experience: user-centered design for the web and beyond**. 2. ed. Berkley: New Riders, 2011. 172 p.
- KRUG, Steve. **Não me faça pensar!:** uma abordagem de bom senso à usabilidade na web. 2. ed. Rio de Janeiro: Alta Books, 2006. 201 p.
- LIMA, Alessandro et al. Projeto para desenvolvimento do Simulador Health Simulator. **Anais do Computer on the Beach**, Florianópolis, 279-288. 2015.
- LUGER, George F.; CHAKRABARTI, Chayian. **Knowledge-based probabilistic reasoning from expert systems to graphical models**. Disponível em: <<http://www.cs.unm.edu/~luger/23-Luger-Chakrabarti.pdf>>. Acesso em: 20 abr. 2016.
- MADSEN, Anders L. et al. The Hugin tool for learning Bayesian networks. In: **Symbolic and quantitative approaches to reasoning with uncertainty**. Springer Berlin Heidelberg, 2003. p. 594-605.
- MARQUES, Roberto Ligeiro; DUTRA, Inês. **Redes Bayesianas:** o que são, para que servem, algoritmos e exemplos de aplicações. Coppe Sistemas – Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil, 2002.

MARONI, Vinícius. **Construção de um motor de inferência para análise de desempenho em ambientes virtuais de aprendizagem aplicados ao ensino da medicina de família e comunidade.** 2013. 103 f. Dissertação (Mestrado em Ciências da Saúde) – Universidade Federal de Ciências da Saúde de Porto Alegre, Porto Alegre, RS, 2013.

NILSSON, Nils J. **Artificial intelligence: a new synthesis.** San Francisco: Morgan Kaufmann, 1998. 513 p.

PEARL, Judea. **Probabilistic reasoning in intelligent systems: networks of plausible inference.** San Francisco: Morgan Kaufmann, 1988. 552 p.

PRODANOV, Cleber Cristiano; FREITAS, Ernani Cesar de. **Metodologia do trabalho científico.** 2. ed. Novo Hamburgo: Feevale, 2013. 276 p.

REISS, Eric. **Usable usability: simple steps for making stuff better.** Indiana: John Wiley & Sons, 2012. 230 p.

RODRIGUES, Fabrício Henrique. **Extração de modelos bayesianos a partir de ontologias construídas para diagnóstico médico.** 2012. 256 f. Trabalho de Conclusão de Curso (Monografia) – Curso de Sistemas de Informação, Universidade Feevale, Novo Hamburgo, RS, 2012.

RUSSELL, Stuart; NORVIG, Peter. **Artificial Intelligence: A modern approach.** 3. ed. New Jersey: Prentice-Hall. 1995. 1132 p.

SANTOS, Felipe Scuciatto dos. **Desenvolvimento de módulo de redes bayesianas para o amplia.** 2009. 70 f. Trabalho de Conclusão de Curso (Monografia) – Curso de Ciência da Computação, Universidade Feevale, Novo Hamburgo, RS, 2009.

SEIXAS, Louise Jeanty de et al. **An architecture for an intelligent learning environment with a constructivist approach.** ITS2002. San Sebastian, Spain. 2002.

SEIXAS, Louise Marguerite Jeanty de. **Estratégias pedagógicas para um ambiente multi-agente probabilístico inteligente de aprendizagem – AMPLIA.** 2005. 147 f. Tese (Doutorado em Informática na Educação) – Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, 2005.

WIEGERINCK, Wim; KAPPEN, Bert; BURGERS, Willem. **Bayesian networks for expert systems: Theory and practical applications.** In: Interactive collaborative information systems. Springer Berlin Heidelberg, 2010. p. 547-578.