

Considerações sobre a utilização de Incertezas Bayesianas e Difusas no Apoio ao Diagnóstico de Morbidades Urológicas

Rosangela Westphal da Silva^{1,2}, Priscyla Waleska Targino de Azevedo Simões^{1,2,3},
Merisandra Côrtes de Mattos^{1,2}, Rozenir Ramos^{2,3}, Cristian Cechinel⁴,
Ana Cláudia Garcia Barbosa^{1,2}

¹Curso de Ciência da Computação – Unidade Acadêmica de Ciências, Engenharias e Tecnologias - Universidade do Extremo Sul Catarinense (UNESC) – Criciúma – SC – Brasil - C.P. 3167 - CEP: 88806-000

²Grupo de Pesquisa em Inteligência Computacional Aplicada– Unidade Acadêmica de Ciências, Engenharias e Tecnologias – Universidade do Extremo Sul Catarinense (UNESC) - Criciúma – SC – Brasil C.P. 3167 - CEP: 88806-000

³Curso de Medicina –Unidade Acadêmica de Ciências da Saúde - Universidade do Extremo Sul Catarinense (UNESC) – Criciúma – SC – Brasil - C.P. 3167 - CEP: 88806-000

⁴Curso de Engenharia da Computação – Universidade Federal de Pelotas (UFPel) / Universidade Federal do Pampa (UNIPAMPA / Bagé) Bagé – RS – Brasil - CEP: 96412-420

rosawestphal@hotmail.com, {pri, mem}@unesc.net,
ramos@engeplus.com.br, contato@cristiancechinel.pro.br, agb@unesc.net

Abstract. *This paper describes the preliminary results of a research in progress that presents a hybrid approach used as an alternative to model imprecise and aleatory uncertainties in urology medical area, from fuzzy and bayesian logic.*

Resumo. *O presente artigo descreve os resultados preliminares de uma pesquisa em andamento que apresenta uma abordagem híbrida utilizada como alternativa para modelar incertezas imprecisas e aleatórias na área médica de urologia, a partir das lógicas fuzzy e bayesiana.*

1 Introdução

O conceito de conjuntos difusos foi introduzido por Lotfi A. Zadeh na Universidade de Berkeley, o qual observou que as tecnologias disponíveis na época não possibilitavam a automação de atividades relacionadas a problemas que compreendem situações ambíguas e imprecisas, esses problemas de natureza industrial, biológica ou química.

Segundo Cox (2005), os conhecimentos dos quais se dispõem em um sistema costumam ser incompletos e com diversas fontes de incerteza. Em contrapartida a comunidade de Inteligência Artificial (IA) reconheceu a necessidade de raciocinar probabilisticamente, pois os primeiros sistemas computacionais para suporte à decisão que foram construídos a partir dos anos 60 eram voltados para o problema da diagnose, que tratavam a incerteza aplicando um formalismo matemático restritivo e abrangente denominado Teorema de Bayes. Essa teoria mostra uma maneira de calcular a probabilidade de um evento em particular, dado algum conjunto de observações que se tenha feito [RUSSELL 2004].

As Redes Bayesianas são grafos acíclicos direcionados onde os nós representam as variáveis com medida de incerteza associada, e os arcos representam a existência causal direta entre os nós conectados (parte qualitativa): a força dessa influência é quantificada por probabilidades condicionais [JENSEN 2001].

Conforme Rezende (2003) descreve, o conceito de sistema híbrido (SH) inteligente ou de método híbrido de aquisição de conhecimentos é bastante amplo e pode englobar diferentes tipos de abordagens. De uma maneira mais geral, pode-se descrever que todo o sistema que integre dois ou mais métodos diferentes da IA para a solução de um problema é um SH. Neste sentido, destacam-se várias aplicações desenvolvidas pelos modelos neuro-fuzzy, algoritmos genéticos-fuzzy, árvores de decisão para indexar a base de casos em um sistemas baseado em casos, redes neurais artificiais com o raciocínio simbólico, entre outras.

A partir dos conceitos apresentados, busca-se com essa pesquisa oferecer um sistema híbrido fuzzy-bayesiano, esperando-se dessa forma, contribuir para essa área de estudo de forma a explorar algumas possibilidades de combinação da teoria bayesiana com a lógica fuzzy para melhor representar um domínio de conhecimento da área da saúde, a qual geralmente é pautada pelo processo de tomada de decisão considerando-se a presença de diferentes tipos de incertezas.

Assim, os próximos itens apresentam a metodologia do desenvolvimento dessa pesquisa, além das conclusões preliminares.

2 Metodologia

A Metodologia dessa pesquisa é baseada nas seguintes etapas: levantamento bibliográfico; levantamento dos trabalhos correlatos; pesquisa sobre o raciocínio bayesiano e lógica fuzzy; definição da ferramenta para o desenvolvimento da inferência bayesiana; e, definição de uma problemática da área da saúde que apresente raciocínio sobre incertezas híbridas. Finalizadas essas etapas busca-se demonstrar a modelagem e o raciocínio matemático do sistema híbrido, finalizando-se com a validação do sistema.

Neste contexto, as etapas já concluídas referem-se ao levantamento bibliográfico, incluindo-se os trabalhos correlatos, a partir dos quais verificou-se a existência de diversas pesquisas que utilizam variadas formas de hibridismo [THÉ 2001] [BRIGNOLI 2001].

Em seguida, definiu-se como ferramenta para o desenvolvimento da inferência bayesiana a shell Netica, por ser bastante utilizada no ensino de redes bayesianas nas universidades brasileiras, por possuir API para integração com diferentes ambientes de programação, e pelo fato da instituição onde essa pesquisa está sendo desenvolvida possuir licença de utilização.

Com relação à definição da problemática foi escolhida a área de Urologia, em específico o diagnóstico de câncer de próstata, por ser uma das morbidades mais prevalentes na população do sexo masculino, e por apresentar a possibilidade da modelagem híbrida do conhecimento. Assim, a modelagem do conhecimento bayesiano e difuso convergiu para o diagnóstico de câncer de próstata (CP) e hiperplasia prostática benigna (HPB).

2.1 Aquisição e Representação do Conhecimento

Nesta etapa foram definidas as hipóteses diagnósticas e as evidências necessárias para chegar ao diagnóstico. Considerando que as hipóteses diagnósticas estabelecidas foram HPB e CP, os sinais e sintomas definidos foram: disúria, febre, obstrução com sensação de não esvaziamento vesical, urinar com intervalo menor que duas horas, jato urinário fraco, acordar várias vezes à noite para urinar, e para auxiliar, alguns exames que são usados para o diagnóstico como PSA total, densidade do PSA, toque retal, e Ultra-som.

Com relação ao hibridismo definiu-se que o raciocínio aleatório seria utilizado em todas as evidências com exceção do PSA total e densidade do PSA que apresentou o raciocínio por imprecisão.

2.1.1 Aquisição e Representação do Conhecimento Bayesiano

A definição das probabilidades condicionais da RB se deu por meio entrevistas com o especialista. Essa técnica foi escolhida em virtude do especialista não contar com uma base de dados estruturada que poderia ser utilizada para aprendizagem bayesiana.

Definidas as probabilidades das hipóteses diagnósticas e evidências, foi realizada a representação do conhecimento por meio de uma rede bayesiana construída na *shell* Netica e ilustrada na Figura 1. Essa rede é composta de um nó para representar as hipóteses diagnósticas (CP, HPB), e outros nove nós (disúria, febre, obstrução, esforço para urinar, jato urinário fraco, intervalo para urinar menos de 2 horas, toque retal, cultura da urina, ultra-som) para representar os sinais e sintomas a serem considerados no diagnóstico.

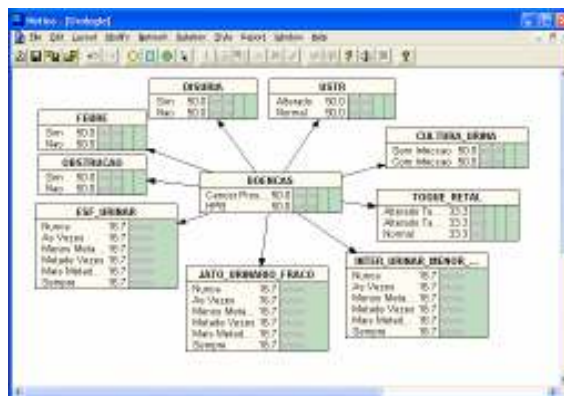


Figura 1. Representação do conhecimento por meio de Rede Bayesiana

2.1.2 Aquisição e Representação do Conhecimento Fuzzy

Seguindo-se com a aquisição do conhecimento, passou-se à definição dos conjuntos difusos necessários para representar o PSA Total e a densidade do PSA, que conforme descrito anteriormente apresentou imprecisão podendo ser alterado ou normal.

Assim, as funções de pertinência L e trapezoidal foram escolhidas para o PSA total por apresentarem um resultado satisfatório para esta pesquisa. O PSA Total apresenta dois conjuntos difusos (normal e alterado), considerando-se que o primeiro representado pela função L que varia entre 0 e 2,5 ng/ml, e o segundo representado pela função trapezoidal que varia entre 2,00 e 9,99 ng/ml.

Para o exemplo da modelagem matemática fuzzy considerando o valor de entrada do PSA Total de 2,4 ng/ml esse valor é fuzzificado pela função de pertinência conforme as equações demonstradas abaixo. Aplicando as fórmulas das funções tem-se:

$$f(\mu(Normal)) = \begin{cases} 1, x \leq a \\ (b-x)/(b-a), a < x < b \\ 0, x \geq b \end{cases} \quad f(\mu(Alterado)) = \begin{cases} 0, x \leq a \\ (x-b)/(a-b), a < x < b \\ 1, b \leq x \leq c \\ (d-x)/(d-c), c < x < d \\ 0, x \geq d \end{cases}$$

$$\mu(Normal) = \frac{2.5 - 2.4}{2.5 - 2.0} = \frac{0.1}{0.5} = 0.20$$

$$\mu(Alterado) = \frac{2.4 - 2.5}{2.0 - 2.5} = \frac{-0.1}{-0.5} = 0.20$$

3 Conclusão

Como essa pesquisa encontra-se em andamento, tem-se como próximas etapas a modelagem matemática do conhecimento híbrido na área de urologia, desenvolvimento do sistema inteligente, e validação do mesmo.

Considerando-se as etapas finalizadas, pode-se concluir que a temática do hibridismo fuzzy-bayesiano é recente no campo de estudo da IA, apresentando variados estudos no Brasil, destacando-se os desenvolvidos na Universidade Federal de Santa Catarina em aplicações voltadas à área da saúde, em específico ao diagnóstico médico.

Referências

- BRIGNOLI, J. T. **Modelo Híbrido Difuso-Probabilístico: uma alternativa para sistemas especialistas**. Dissertação- Pós-Graduação em Ciência da Computação, Universidade Federal de Santa Catarina, Florianópolis, Santa Catarina, 2001.
- COX, E. **Fuzzy modeling and genetic algorithms for data mining and exploration**. California: Morgan Kaufmann, 2005.
- JENSEN, F. V. **Bayesian networks and decision graphs** New York: Springer, 2001
- REZENDE, S. O. **Sistemas inteligentes: fundamentos e aplicações**. São Paulo: Manole, 2003.
- RUSSELL, S. J.; NORVIG, P. **Inteligência artificial. Tradução da segunda edição**. Rio de Janeiro: Elsevier, 2004.
- THÉ, M. A. L. **Raciocínio baseado em casos uma abordagem fuzzy para diagnóstico nutricional**. Tese – Pós graduação em Engenharia de Produção, Universidade Federal de Santa Catarina, Santa Catarina, 2001.