

## Modelagem bayesiana da aprendizagem de estudantes em um AVA

Ademar Crotti Junior<sup>1</sup>, Beatriz Wilges<sup>1</sup>, Silvia Modesto Nassar<sup>1</sup>

<sup>1</sup> Programa de Pós-Graduação em Ciência da Computação – Universidade Federal de Santa Catarina (UFSC)  
Caixa Postal 88.040-900 – Florianópolis – SC – Brasil

{ademar.c@posgrad.ufsc.br},{beaw,silvia@inf.ufsc.br}

**Abstract.** *This paper presents a novel method for learning the structure of Bayesian networks (BN), in the context of a VLE (Virtual Learning Environment). The proposed method applies a fuzzy system to combine different scores, which are used to evaluate BNs. For evaluation, this method was applied to a data set of general exams and also to the exam questions of a VLE. The proposed method modeled the RBs through the datasets. Thus, the extracted knowledge was used to identify possible gaps in content or even in the elaboration of the exam questions of the VLE. This proposal uses a set of common variables between VLE, so it is possible to apply the presented method in any VLE. Furthermore, analysis of the BNs generated by the method were evaluated by the instructors of the VLE.*

**Resumo.** *Este artigo apresenta um novo método para aprendizagem de estrutura de redes bayesianas (RB), inserido no contexto de um AVA. O método proposto aplica um sistema fuzzy para combinar diferentes métricas de pontuação, que são utilizadas para avaliar as RBs. Para verificação, este método foi aplicado em um conjunto de dados de avaliações gerais e, também, em um conjunto de avaliações de questões de um AVA. O método proposto modelou as RBs por meio dos conjuntos de dados. Assim, o conhecimento extraído foi utilizado para identificar possíveis falhas no conteúdo ou, até mesmo, na formulação das questões das avaliações do AVA. Esta proposta utiliza um conjunto de variáveis que são comuns entre os AVA, sendo possível aplicar o método apresentado em qualquer AVA. Além disso, a análise das RBs geradas pelo método foram avaliadas por docentes do ambiente.*

### 1. Introdução

Redes Bayesianas (RB) são modelos gráficos que representam conhecimento e raciocinam por meio de probabilidades. As RBs têm sido utilizadas em diversos tipos de aplicações e, diferentes áreas, como por exemplo: saúde (diagnóstico médico e localização de genes), computação (resolução de problemas e agentes inteligentes) e educação.

Na área da educação, as RBs são aplicadas na modelagem do conhecimento de estudantes, no desenvolvimento de jogos, na identificação de perfis de estudantes em Ambientes Virtuais de Aprendizagem (AVAs). A identificação do perfil do estudante é normalmente tratada pelo mapeamento das informações desse estudante em um Modelo de Aluno (MA). Segundo Seffrin, Rubi e Jaques (2013), uma das formas de mapear o conhecimento é através das RB, por elas serem capazes de tratar a incerteza dos dados.

O processo de construção de uma rede pode ser realizado pelo modelo clássico, que envolve um especialista humano, ou de forma automatizada. O modelo clássico envolve várias etapas onde a aquisição do conhecimento é a principal. Na aquisição do conhecimento o especialista deve expressar e associar valores de probabilidades condicionais para a parte quantitativa da rede, assim como estabelecer as relações que compõem a parte qualitativa.

Em muitos casos, o modelo clássico pode ser demorado, custoso, além de complexo. Por exemplo, a grande quantidade de variáveis e os dados incompletos dificultam a manutenção da estrutura da rede, podendo inviabilizar o processo. Assim, métodos automatizados podem ser potencialmente mais interessantes, porque diminuem o tempo necessário para construção da rede, além de prover uma visualização adequada das informações estruturadas na RB.

Desta forma, este trabalho pretende construir uma rede bayesiana para análise do desempenho da aprendizagem de estudantes considerando recursos pedagógicos do AVA. A estrutura da rede proposta é construída de forma automatizada e permite uma análise que contempla dois enfoques: avaliação geral do desempenho do estudante e outro, específico, no desempenho de cada questão considerada no AVA.

Este artigo está organizado da seguinte forma: a segunda seção apresenta os trabalhos relacionados. A terceira seção expõe a fundamentação teórica. Na quarta seção é apresentada a aprendizagem de RB. Na quinta seção é apresentado os procedimentos metodológicos, a sexta seção relata os resultados obtidos; e por fim, são apresentadas as considerações deste trabalho.

## **2. Trabalhos relacionados**

Muitas pesquisas procuram tratar a incerteza das informações do perfil de estudantes por meio de RB. Assim, diversos trabalhos pesquisam sobre a implementação de um componente denominado Modelo do Aluno, em um AVA, com a utilização de RBs. Além disso, também são desenvolvidas propostas que tratam sobre jogos e analisam os dados gerados por meio da RB para prever tendências de comportamento.

Na pesquisa de Seffrin, Rubi e Jaques (2013) foi implementada uma RB para o mapeamento do conhecimento necessário para resolução de equações do 1º grau. Este mapa foi utilizado em uma modelagem de rede bayesiana para identificar o conhecimento do aluno no domínio dos conceitos necessários para resolver equações de álgebra elementar.

Na pesquisa de Millan et al. (2013) já havia sido apresentado um modelo de RB para o diagnóstico do conhecimento do aluno em equações do 1º grau. Segundo os autores, as unidades de conhecimento estabelecem relações entre si, indicando precedência de conhecimento. As probabilidades dos nodos indicam o quanto o aluno conhece o conteúdo, que pode ser influenciado por conhecimentos prévios.

No modelo proposto por Seffrin, Rubi e Jaques (2013), diferentemente da implementação proposta por Millan et al. (2013), é considerado cada passo no processo de resolução das equações como uma evidência do conhecimento do aluno, ao invés de considerar apenas se o aluno acertou a resposta final. Assim, os autores conseguem uma identificação mais precisa das unidades de conhecimento do estudante.

Na pesquisa de Basto et al. (2012) foi desenvolvido um jogo sobre a logística de execução de compras em um supermercado. A ideia desse trabalho foi avaliar o conjunto de resultados obtidos pelos usuários a partir do jogo em uma rede bayesiana do tipo Naïve Bayes para identificação de crianças e adolescentes com Transtorno de Déficit de Atenção/Hiperatividade (TDAH).

Dias et al. (2008) analisaram os dados de um ambiente de ensino-aprendizagem a partir da aplicação de técnicas de Mineração de Dados para descobrir informações relevantes sobre o perfil do aluno. O trabalho de Dias et al (2008) utilizou árvores de decisão para identificação de padrões referentes ao processo de aprendizagem. Esses padrões foram relacionado com o comportamento do aluno, e a aplicação das redes bayesianas foi utilizada para contabilizar o processo de aprendizagem e o desempenho desses estudantes.

No trabalho de Longhi, Behar e Bercht (2010) foi desenvolvida uma máquina de inferência para identificação dos estados de ânimo dos estudantes em AVAs, por meio de redes bayesianas. A rede bayesiana foi modelada por um especialista, com base em um modelo de teste de personalidade. Assim, a máquina de inferência proposta pelos autores considerou traços de personalidade, fatores motivacionais obtidos através dos padrões de comportamento e a subjetividade afetiva identificada em textos disponibilizados nas funcionalidades de comunicação dos AVAs.

Diversas pesquisas, como Mateus et al. (2011), evidenciam o esforço de encontrar um módulo adaptativo para acompanhar o estudante em um AVA em função da dinâmica de desempenho do mesmo ao longo do curso. Em pesquisas de Wilges et. al (2012) foi estruturada uma modelagem difusa para o AVA em função da mineração dos resultados do desempenho de uma turma.

As propostas de pesquisas relacionadas com este trabalho utilizam diferentes técnicas para identificar o perfil de aprendizagem dos estudantes nos AVA. Esta proposta se destaca porque é capaz de extrair o perfil de aprendizagem dos estudantes não somente na avaliação do desempenho final, mas também no desempenho dos estudantes em cada uma das questões que compõem a avaliação final e que consideram, neste AVA, diferentes níveis de aprendizagem.

### **3. Redes Bayesianas**

As Redes Bayesianas (RB) são representações gráficas de distribuição de probabilidades. Elas são muito utilizadas como modelos de incerteza e são compostas por grafos acíclicos direcionados, mostrando as relações de causalidade entre os atributos. Nestes grafos, as elipses representam os atributos e as ligações representam os relacionamentos de influência entre os atributos. Através de cálculos estatísticos, cada atributo tem uma tabela de valores de probabilidades. Assim, na utilização de uma ferramenta de análise de RB é possível definir hipóteses sobre um determinado atributo, tendo respostas sobre as influências dele de acordo com as ligações existentes entre outros atributos.

A estrutura de uma RB apresenta as dependências condicionais entre suas variáveis. Os parâmetros apresentam as medidas de probabilidade condicionais, ou seja, quantificam o efeito que um conjunto de variáveis tem sobre suas variáveis relacionadas.

O conhecimento de uma rede bayesiana é extraído da probabilidade condicional de um conjunto de variáveis de consulta, de acordo com as variáveis de evidência. Esta probabilidade é calculada utilizando o Teorema de *Bayes* [Neapolitan, 2004]:

$$P(E|F) = \frac{P(F|E)P(E)}{P(F)}$$

Onde  $E$  e  $F$  são dois eventos, tal que  $P(E|F)$  apresenta a probabilidade a ser descoberta do evento  $E$  dado o evento  $F$ ,  $P(F|E)$  é a probabilidade condicional do evento  $F$  dado  $E$ , e  $P(E)$  e  $P(F)$  são as probabilidades conhecidas dos eventos  $E$  e  $F$ .

#### 4. Aprendizagem de Redes Bayesianas a partir dos dados

Redes bayesianas podem ser modeladas por um especialista ou aprendidas por meio dos um conjunto de dados que são tratados por algoritmos específicos.

O processo de aprendizagem pode ser dividido em duas etapas, a aprendizagem de estruturas e a aprendizagem de parâmetros. A aprendizagem de estruturas identifica as dependências entre as variáveis e a direção da causalidade. A aprendizagem de parâmetros, ou das probabilidades, indica a intensidade dos relacionamentos de cada estrutura. Esta é considerada uma tarefa simples, quando a estrutura da rede é conhecida [Cheng; Greiner, 2001].

Existem duas abordagens principais na aprendizagem de estruturas de redes bayesianas. O método de independência condicional, que faz testes de independência condicional nos dados e busca uma estrutura consistente com as independências observadas. Um problema desta abordagem é o número exponencial de testes de dependência. E também a segunda abordagem, o método de busca heurística, que define uma função que avalia quão bem as dependências de uma estrutura representam os dados, localizando a estrutura mais simples que aumenta o valor desta função. O principal problema desta abordagem é o espaço de busca, que considera todas as possíveis estruturas [Yan; Cercone, 2010]. Essa busca pode ser feita utilizando diversas técnicas, como, por exemplo, algoritmos genéticos, gulosos, ou métodos específicos como Monte Carlo e Cadeias de *Markov*.

##### 4.1. Algoritmo EM-MCMC

O algoritmo EM-MCMC foi proposto por Guo e Li em 2009, e foi baseado no algoritmo MCMC *Metropolis-Hastings*. A aprendizagem dos parâmetros utiliza o algoritmo *Expectation-Maximization* (EM). Assim, o algoritmo EM-MCMC efetua uma busca no espaço de estruturas, criando, removendo e invertendo arestas. O método modifica aleatoriamente a estrutura da rede bayesiana, a cada alteração a estrutura é reavaliada. Ao modificar a estrutura é gerado um número aleatório entre 0 e 1. A estrutura é aceita se o número gerado for menor que a diferença entre os valores encontrados na avaliação da estrutura atual e a próxima estrutura possível, caso contrário, gera-se outro número aleatório entre 0 e 1, e a estrutura é aceita se o número gerado for menor que o fator MCMC dado por Guo e Li (2009):

$$\frac{P(G_i|D)}{P(G_j|D)} = \frac{P(D|G_i)P(G_i)}{P(D|G_j)P(G_j)}$$

A função de pontuação utilizada é chamada *Minimum Description Length* (MDL) [Guo; Li, 2009] e é especificada da seguinte forma:

$$MDL(B, D) = H(B, D) + \frac{K}{2} \log N$$

Sendo:

$$H(B, D) = -N \sum_{i=1}^n \sum_{j=1}^{q_i} \sum_{k=1}^{r_i} \frac{N_{ijk}}{N} \log \frac{N_{ijk}}{N_{ij}}$$

$$K = \sum_{i=1}^n (r_i - 1) \cdot q_i$$

Onde  $N$  é o número de instâncias nos dados  $D$ ; o número de estados de uma variável  $X_i$  é  $r_i$ ; o número de possíveis configurações de pais  $P_{ag}(X_i)$  de  $X_i$  é  $q_i$ ;  $N_{ijk}$  é o número de instâncias nos dados  $D$  onde a variável  $X_i$  possui valor  $x_{ik}$  e o conjunto de variáveis  $P_{ag}(X_i)$  possui valor  $w_{ij}$ ;  $N_{ij}$  é o número de instâncias nos dados onde  $P_{ag}(X_i)$  possui valor  $w_{ij}$ ;  $K$  define a complexidade da rede, sendo usado como um termo de penalização [Guo; Li, 2009].

## 5. Procedimentos metodológicos

Esta seção apresenta um algoritmo para aprendizagem de estrutura de redes bayesianas utilizando o método Monte Carlo e Cadeias de *Markov*. Dessa forma, esta pesquisa aplica um método de busca heurística porque considera importante localizar uma estrutura de RB simples, mas que represente as dependências dessa estrutura de uma forma completa e significativa.

### 5.1. Método proposto

O método proposto nesta pesquisa baseia-se no algoritmo EM-MCMC apresentado na seção anterior. Assim, a partir de um conjunto de dados, a primeira etapa do algoritmo gera uma estrutura aleatória para a rede bayesiana, sendo este o estado inicial da cadeia de Markov.

Na sequência duas variáveis são amostradas aleatoriamente a partir do conjunto de variáveis. A modificação na estrutura é feita gerando um número aleatório por meio de uma distribuição uniforme. As modificações possíveis são adicionar, remover ou inverter um arco entre as variáveis amostradas. Como uma rede bayesiana não pode conter ciclos, antes de avaliar a estrutura, o algoritmo realiza esta verificação.

Geralmente a estrutura é avaliada utilizando uma única métrica, como no algoritmo EM-MCMC. Algumas métricas favorecem estruturas mais complexas, outras levam em consideração a quantidade de parâmetros. Esta proposta combina métricas conhecidas em uma modelagem *fuzzy* para avaliar as estruturas no método MCMC.

A motivação para utilizar a lógica *fuzzy* está na possibilidade de avaliar simultaneamente a estrutura com diferentes métricas. Assim considera-se a imprecisão de seus valores no sistema *fuzzy*. A modelagem *fuzzy* permite combinar, de forma flexível, estas métricas em seu processo de agregação. Neste trabalho, a modelagem *fuzzy* combina três métricas de pontuação, MDL, descrita anteriormente, AIC e BDe.

A métrica AIC (*Akaike Information Criterion*) possui dois termos, um controlando entropia, baseada em entropia condicional, e outro controlando a complexidade. A entropia é um valor não negativo que mede a incerteza e tende a zero quando o conhecimento é alto. A métrica AIC é dada por [Akaike, 1974]:

$$AIC = H(B, D) + K$$

A métrica BDe (*Bayesian Dirichlet equivalence*) maximiza a probabilidade da estrutura da rede de acordo com os dados. Assim a probabilidade é calculada para diversas estruturas, sendo escolhida a estrutura com maior probabilidade. A métrica BDe é dada por [Bouckaert, 1994]:

$$BDe(B, D) = P(B) \prod_{i=0}^n \prod_{j=1}^{q_i} \frac{\Gamma(r_i)}{\Gamma(r_i + N_{ij})} \prod_{k=1}^{r_i} \frac{\Gamma(r_i + N_{ijk})}{\Gamma(r_i)}$$

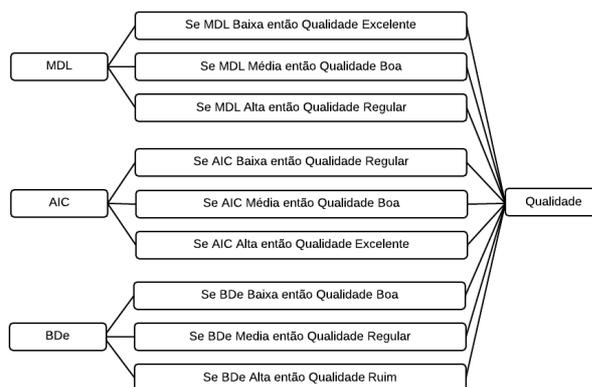
A partir dessas métricas são definidas as variáveis linguísticas para o sistema *fuzzy*. Este sistema é composto por quatro variáveis: três de entrada e uma de saída.

**Tabela 1. Variáveis linguísticas das métricas**

Métrica	Variáveis linguísticas
AIC	Baixa, Média e Alta
MDL	Baixa, Média e Alta
BDe	Baixa, Média e Alta
Qualidade	Ruim, Regular, Boa e Excelente

Cada variável linguística é definida por uma função de pertinência. A definição das funções de pertinência é feita utilizando o maior valor entre as métricas, considerando uma estrutura de rede totalmente conectada e a mesma estrutura totalmente desconectada.

A próxima etapa é a especificação da base de regras do sistema *fuzzy*. As regras foram definidas analisando as propriedades das métricas utilizadas. As métricas MDL e BDe devem ser minimizadas, a AIC maximizada. A Figura 1 apresenta o conjunto de regras.



**Figura 1. Regras para a métrica fuzzy**

As métricas são agregadas com a variável qualidade por truncamento. A parte final é a desfuzzyficação dessa variável em valor escalar. O método selecionado foi o centróide. O algoritmo continua até que a quantidade de iterações seja atingida ou mudanças na estrutura não alterem a variável qualidade.

## 5.2. Origem dos dados

O Ambiente Virtual de Aprendizagem (AVA), de onde os dados desta pesquisa foram extraídos, é utilizado no ensino semipresencial na disciplina de Estatística pelos

estudantes de cursos de graduação em Engenharia da Universidade Federal de Santa Catarina.

Neste AVA tanto a lista de questões, quanto as avaliações consideram diferentes níveis de aprendizagem: básico em um contexto empírico, o intermediário em um contexto pseudo-empírico e avançado em uma abordagem reflexionante.

A base de dados foi pré-processada para aplicação do algoritmo de aprendizagem de estrutura de redes bayesianas. A base de avaliações possui 399 registros. As avaliações são divididas por assunto (QLQL, QTQT e QLQT), tipo (prática e teórica) e nível (básico, intermediário e avançado). As notas e o tempo foram divididos em grupos ordinais baixo, médio e alto. As notas foram definidas por uma variável denominada desempenho. E a variável tempo representa a duração da execução da atividade desenvolvida pelo estudante.

A base de dados das questões possui 4127 registros. No AVA, cada questão pertence a um assunto relacionado à estatística (QLQL, QTQT e QLQT) e possui um tipo que é uma característica da questão: prática ou teórica. As questões são classificadas na base do AVA em níveis: básico, intermediário e avançado. As notas dos alunos nas questões foram discretizadas em cinco categorias (excelente, bom, regular, ruim e fraco) na variável desempenho. A Figura 2 apresenta o esboço dos dados tratados pela RB, sobre a perspectiva de um algoritmo que implementa árvores de decisão (AD).

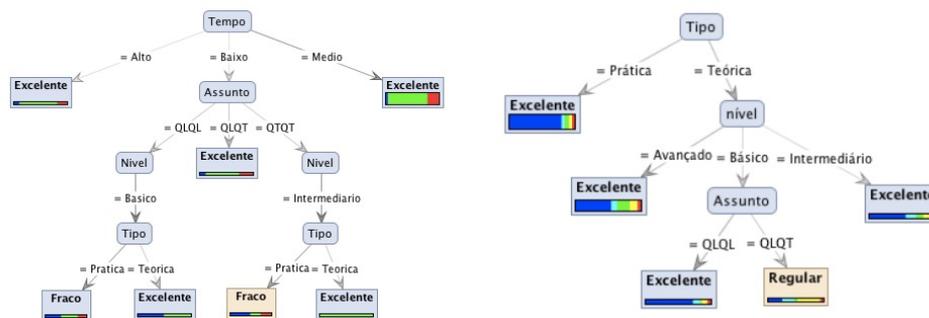


Figura 2. Conjunto de dados das avaliações e questões.

As estruturas apresentadas na Figura 2 permitem observar o conjunto de dados processado pelo algoritmo que gerou as RBs propostas nesta pesquisa. Nestas ADs o *label* utilizado é o desempenho. No caso da avaliação geral são considerados os conjuntos excelente, regular e fraco. No caso das questões são considerados os grupos excelente, bom, regular, ruim e fraco. Desta forma, é possível verificar o comportamento dos dados nas diferentes situações em que o estudante é avaliado: avaliação geral (AD da esquerda) e avaliação nas questões (AD da direita). Assim, por meio das ADs é possível perceber a divisão do conjunto de dados considerando o tipo, o assunto, o tempo e o nível.

### 5.3. Avaliação

As redes bayesianas descobertas pelo método proposto foram avaliadas por um grupo de professores. Esta verificação permitiu uma análise concentrada especificamente no desempenho dos estudantes e também uma percepção do nível de elaboração das atividades propostas tanto nas questões quanto nas avaliações.

Assim, foi possível que os docentes concentrassem esforços na reelaboração das atividades presentes no AVA a fim de elevar o desempenho dos estudantes. Com isso, os professores foram capazes de potencializar as atividades propostas e melhorar a concepção dos conteúdos apresentados no ambiente.

## 6. Resultados

Foram considerados dois conjuntos de avaliações de desempenho de estudantes no AVA: avaliação geral e avaliação específica nas questões. A Figura 3 apresenta a RB gerada a partir dos dados das avaliações dos estudantes.

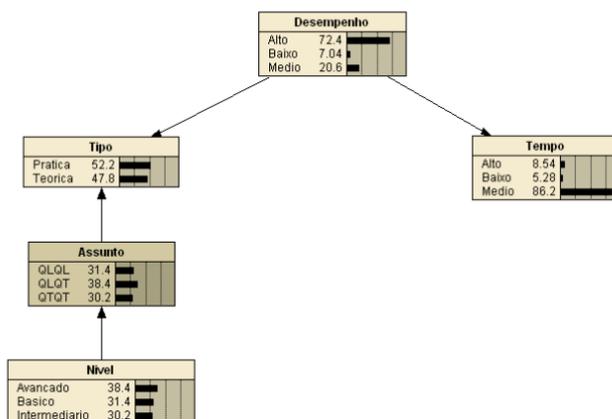


Figura 3. Estrutura da RB para o desempenho nas avaliações

Na Figura 3 observa-se que, na definição das probabilidades do desempenho do aluno, as variáveis principais são o tempo e o tipo da avaliação. Conforme observado, as variáveis nível e assunto alteram as probabilidades da variável tipo. A RB apresentada identificou dois enfoques: no docente para as variáveis tipo, assunto e nível, e nos estudantes para as variáveis tempo e desempenho. Isso porque, o conjunto de questões e avaliações são elaborados por meio das definições do corpo pedagógico e docente. E o enfoque no estudante está no desenvolvimento das atividades que foram concebidas pelos professores.

Com o propósito de entender o aprendizado do aluno no AVA com mais detalhes, o desempenho dos mesmos estudantes nas diferentes questões e situações de aprendizagens específicas foi analisado. Assim, o método proposto foi aplicado neste conjunto de dados e gerou a RB apresentada na Figura 4.

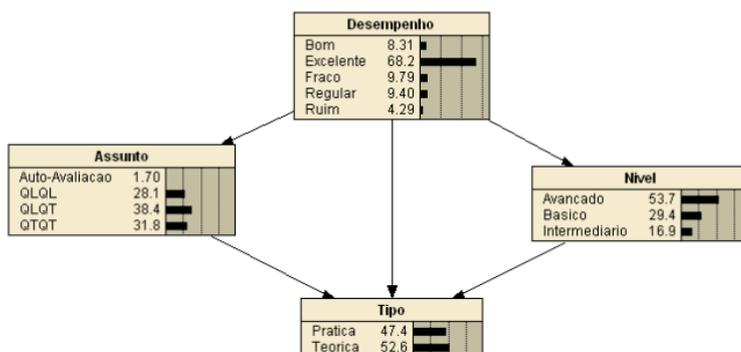


Figura 4. Estrutura da RB para o desempenho nas questões

Os resultados da Figura 4 expressam o comportamento do desempenho considerando as questões e não os resultados finais das avaliações, conforme apresentado na Figura 3. A análise desta rede mostra que as variáveis tipo, nível e assunto possuem influência sobre o desempenho nas questões, que pode ser observado na variável desempenho da Figura 4.

As RBs também podem ser utilizadas para inferência. A Figura 5 mostra a RB quando as variáveis de entrada, assunto, tipo e nível, estão observadas. Assim, percebe-se que as questões de assunto QLQL, de nível intermediário, do tipo teórica foram as que apresentaram menor probabilidade de aprovação. Segundo os professores que avaliaram o modelo, isso pode indicar uma falha na elaboração destas questões, ou nos conteúdos destas avaliações.

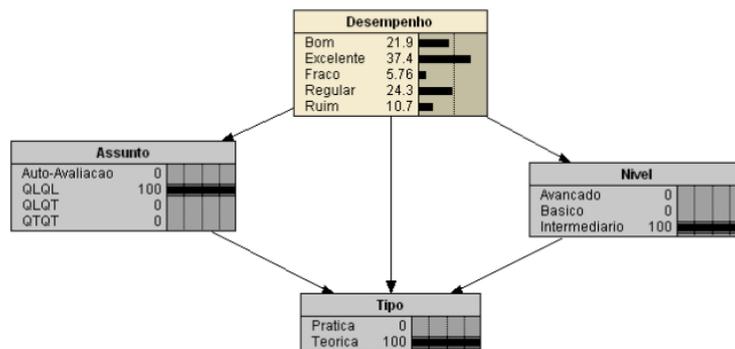


Figura 5. Inferência da RB das questões

Ao instanciar a variável tipo como prática na RB da Figura 5, foi possível verificar que para questões práticas, de qualquer assunto ou nível, a probabilidade dos alunos terem desempenho excelente é alta.

## 7. Considerações

Este trabalho apresenta um diferencial quanto à aplicação da RB em um AVA, porque considera os aspectos que envolvem a natureza das avaliações realizadas no AVA: teórica e prática. Trata-se de uma pesquisa que busca a identificação do desempenho do estudante quando ele está inserido em um contexto de avaliação onde as avaliações e questões possuem diferentes níveis de aprendizagem.

O método proposto nesta pesquisa estrutura o conhecimento do processo de aprendizagem do estudante sobre dois enfoques: o desempenho final na avaliação e o desempenho nas diferentes questões que compõe a avaliação. Assim, é possível uma visualização que contempla diferentes granularidades na rede bayesiana proposta. Um com enfoque de alta granularidade, ou seja, com menos detalhes observados. Neste caso, considerando o comportamento da RB ao receber os dados da avaliação geral dos estudantes. E, também, um enfoque com uma baixa granularidade, onde mais detalhes são observados no desempenho do estudante. Neste caso, considerando cada questão respondida pelos estudantes nos diferentes níveis adotados pelo AVA.

A descoberta automática das estruturas de RB pelo método proposto permite modelar a aprendizagem do aluno, além de explicitar os desafios pedagógicos concebidos pelo docente. O modelo da rede pode ser utilizado para observar a dependência entre as variáveis, as probabilidades dos eventos acontecerem e também como um modelo de consulta. A partir do conhecimento extraído e estruturado em uma

rede *bayesiana*, o AVA pode ser modificado, observando conteúdos que apresentam deficiência no aprendizagem do aluno.

## Referências

- Akaike, H. A New Look at the Statistical Model Identification. IEEE Transactions on Automatic Control. 1974, 716-723.
- Bastos, Angela Paula Zão; Santos, Fábio; Andrade, Leila; Mattos, Paulo. Utilização de um Jogo Sério e Naïve Bayes para Auxiliar na Avaliação Cognitiva do Transtorno de Déficit de Atenção/Hiperatividade. Anais do Simpósio Brasileiro de Informática na Educação. Vol. 23, n. 1. 2012.
- Bouckaert, Remco R. Probabilistic Network Construction Using The Minimum Description Length Principle. Technical Report, Utrecht University, 1994.
- Cheng, Jie; Greiner, Russel. Learning Bayesian Belief Network Classifiers: Algorithms and System. Alberta: Proceedings of the 14th 10Biennial Conference of the Canadian Society on Computational Studies of Intelligence: Advances in Artificial Intelligence, 2001.
- Dias, Maxwel Macedo; Filho, Luiz Alberto da Silva; Lino, Adriano Del Pino; Favero, Eloi Luiz; Ramos, Edson Marcos Leal Soares. Aplicação de Técnicas de Mineração de Dados no Processo de Aprendizagem na Educação a Distância. Anais do Simpósio Brasileiro de Informática na Educação. Vol. 1, n. 1, p. 105-114, 2008.
- Guo, P. and Li, N. (2009). An em-mcmc algorithm for bayesian structure learning. In Computer Science and Information Technology, 2009. ICCSIT 2009. 2nd IEEE International Conference on, pages 158–162. IEEE.
- Longhi, Magalí T; Behar, Patricia A; Bercht, Magda. Máquina de inferência baseada na teoria bayesiana para identificar os estados de ânimo do aluno em um ambiente virtual de aprendizagem. RENOUE, vol. 8, n.3, 2010.
- Mateus, Gustavo P.; Wilges, Beatriz; Nassar, S. M.; Wronski, V.; Bastos, Rogério. Um módulo adaptativo para AVA baseado no desempenho do estudante. In: Simpósio Brasileiro de Informática na Educação (SBIE), 2011, Aracaju-SE. Anais do XXII SBIE - XVII WIE, 2011.
- Millan, E.; Descalço, L.; Castillo, G.; Oliveira, P.; Diogo, S. Using Bayesian networks to improve knowledge assessment. Computers & Education, 60 (1): 436–447. 2013.
- Neapolitan, R. E. (2004). Learning bayesian networks. Pearson Prentice Hall Upper Saddle River.
- Seffrin, Henrique ; Rubi, Geiseane; Jaques, Patricia. Uma Rede Bayesiana aplicada a Modelagem do Conhecimento Algébrico do Aprendiz. Anais do Simpósio Brasileiro de Informática na Educação. Vol. 24, n. 1, 2013.
- Wilges, Beatriz ; Mateus, Gustavo P.; Nassar, S. M.; Bastos, Rogério. Fuzzy Modeling Built Through a Data Mining Process. Revista IEEE América Latina, v. 10, p. 1622-1626, 2012.
- Yan, Lisa Jing; Cercone, Nick. Bayesian network modeling for evolutionary genetic structures. Computers & Mathematics with Applications 59.8: 2541-2551, 2010.