Modelagem do Conhecimento Algébrico dos Estudantes com Redes Bayesianas Dinâmicas

Title: Modelling Student's Algebraic Knowledge with Dynamic Bayesian Networks

Henrique M. Seffrin

Universidade do Vale do Rio dos Sinos
(UNISINOS)
hseffrin@outlook.com

Patrícia A. Jaques

Universidade do Vale do Rio dos Sinos
(UNISINOS)
pjaques@unisinos.br

Resumo

Para que os Sistemas Tutores Inteligentes possam fornecer assistência e instrução personalizada ao aprendiz, eles precisam inferir o seu conhecimento. Redes Bayesianas são comumente empregadas para modelar a probabilidade do aluno saber um determinado tópico (unidade) de conhecimento, uma vez que as mesmas são capazes de lidar com incertezas. Este trabalho apresenta uma modelagem de Rede Bayesiana Dinâmica para a avaliação do conhecimento algébrico do estudante em Sistemas Tutores Inteligentes. O principal diferencial do trabalho proposto é a sua capacidade de avaliar tanto o conhecimento conceitual (declarativo), quanto procedural do estudante, utilizando um modelo de inferência que é independente de problemas (equações). Nesse artigo, são descritas as etapas de levantamento de informações para obtenção das probabilidades desta rede, bem como as etapas das avaliações conduzidas. Os resultados da avaliação mostraram similaridades estatisticamente significativas entre a inferência da rede e os percentuais de acerto por parte dos estudantes, evidenciando que o modelo proposto infere corretamente o conhecimento do estudante.

Palavras-Chave: *Redes Bayesianas Dinâmicas, Sistemas Tutores Inteligentes, Modelo de Aluno, Álgebra*

Abstract

Students' Knowledge Inference is an important component in the construction of Intelligent Tutoring Systems; it gets the basics of each learner knowledge, which allows the tutor to adapt the pedagogical instruction for each student. In the literature, it is common the use of Bayesian Networks to perform this kind of inference, because they are able to deal with uncertainties. This paper presents a Dynamic Bayesian Network modeling for the inference of student's algebraic knowledge. Differently from related works, it evaluates both student's procedural and declarative knowledge, besides being independent of problem. This paper also describes the steps we followed to get the information about the network probabilities, as well the evaluation conducted with the network. The evaluation results showed statistically significant similarities between the network inference and students' performance. This result evidences that the proposed work infers correctly student's knowledge.

Keywords: *Dynamic Bayesian Networks, Intelligent Tutoring System, Student Model, Algebra*

1 Introdução

Uma grande parcela dos estudantes apresentam dificuldades em matemática, em especial na área da Álgebra [1], [2]. Tal fato também é observado na realidade brasileira, conforme apontam os Parâmetros Curriculares Nacionais [3]. Uma possível explicação para isto está no fato da Álgebra introduzir conceitos que destoam do conteúdo aritmético trabalhado até então. Enquanto a aritmética busca a descoberta de um resultado para uma expressão numérica, por exemplo $3+5*4=$, a álgebra busca a descoberta de um valor desconhecido em uma equação, como $2*(x+5)=18$. É introduzida a noção de incógnita: uma letra que representa um determinado valor desconhecido que deve ser descoberto.

Este cenário de baixo rendimento escolar tem motivado pesquisadores a buscar soluções, inclusive tecnológicas, tais como Sistemas Tutores Inteligentes (STIs). STIs são *softwares* voltados para o ensino de conteúdos, de áreas específicas do conhecimento. Diferenciam-se dos demais *softwares* educacionais pela sua capacidade de ensinar e prover assistência de forma personalizada. Através da interação com os alunos, o sistema os avalia, identificando seus conhecimentos e estados afetivos, e adapta a instrução, os exercícios e a assistência de acordo. Os STIs fornecem ensino individualizado (*one-to-one tutoring*), no qual há um professor particular para cada estudante. O *one-to-one tutoring* é uma forma mais efetiva de ensino do que a aula tradicional em sala de aula com um professor para vários alunos [4].

A capacidade adaptativa dos STIs é possível graças ao componente Modelo de Aluno, que é responsável por inferir o conhecimento dos estudantes [5]. Além disso, este componente também armazena esses dados, juntamente com outras informações relativas a cada estudante, como o seu perfil, seus estados afetivos, etc. Estas informações são utilizadas pelo componente Tutor que realizará a adaptação de conteúdo.

No que se refere ao processo de avaliação do conhecimento, também chamado de processo de diagnóstico, ele ocorre durante a resolução dos exercícios, uma vez que as respostas submetidas pelos alunos são as únicas evidências de conhecimento que o sistema dispõe [6]. Desta forma, quanto mais detalhadas elas forem, melhor será a inferência.

Na literatura, é muito comum o uso de Redes Bayesianas (RB) e Redes Bayesianas Dinâmicas (RBD) para a inferência do conhecimento do aprendiz. As RBs são estruturas próprias para o uso com dados que possuem incertezas, característica do conhecimento humano. Através dessas estruturas, o conhecimento é modelado em uma relação de precedência: que conceitos o aluno deve conhecer primeiramente para que possa prosseguir no

estudo. Desta forma, é possível inferir que conteúdos o estudante já está apto para começar a estudar.

Este artigo descreve o desenvolvimento de uma RBD para avaliar o conhecimento algébrico do estudante em Sistemas Tutores Inteligentes, que busca inferir a probabilidade do estudante conhecer os conceitos e as operações relacionadas à resolução de equações de 1º grau, assim como as principais falsas concepções do aprendiz. Este trabalho contou com a colaboração de professores de matemática do ensino fundamental, médio e superior que, através de entrevistas e questionários, contribuíram para a elaboração e evolução da estrutura da RBD proposta.

Este trabalho diferencia-se dos demais pesquisados, por utilizar as etapas da tarefa (ou passos) como evidência para o componente, ao invés da tarefa completa [9], [10]. Esta característica, combinada ao fato de utilizar uma RBD, torna este modelo desenvolvido independente de equação. Deste modo, a rede utilizada não necessita ser reestruturada caso uma nova equação seja inserida no sistema.

Um outro diferencial do trabalho é a sua capacidade de avaliar tanto o conhecimento declarativo (conceitual), quanto procedural do estudante (operações) e as relações entre esses conceitos e operações. Como na álgebra há esta relação de dependência entre as operações e conceitos, essa característica é desejável, pois isso permite a inferência do conhecimento do estudante em habilidades e conceitos não diretamente trabalhados nos problemas.

Para a avaliação do modelo proposto, foram utilizados dados de interação dos alunos com o STI PAT2Math. O PAT2Math (*Personal Affective Tutor to Math*) é voltado para o ensino de álgebra elementar e trabalha com operações de 1º e 2º grau com uma incógnita [7]. Por ser um STI do tipo *step-based*, é capaz de oferecer assistência individualizada para cada etapa fornecida pelo estudante durante a resolução passo-a-passo das equações. Possui, também, um módulo especialista capaz de identificar as operações aplicadas pelos estudantes e os principais erros cometidos por eles [8].

Os dados do PAT2Math utilizados para avaliação do trabalho proposto foram coletados durante os meses de outubro e novembro de 2014, durante os quais estudantes de uma escola utilizaram o sistema durante 50 minutos semanais. Estes dados foram armazenados, organizados e, posteriormente, inseridos na rede, cujo processo de execução ocorreu da mesma forma que seria se a rede estivesse integrada ao PAT2Math. Ao final desta execução, a rede emitiu um parecer sobre o conhecimento dos aprendizes para cada operação de 1º grau. Estes valores foram comparados com o resultado do pós-teste, realizado pelos mesmos alunos ao final do período de utilização do STI em sala de aula. Os resultados obtidos não foram satisfatórios em uma primeira execução, o que exigiu

ajustes nas políticas de controle da rede. Estabeleceu-se, então, um valor de limiar que determina que, a partir daquele ponto, o aluno possuía o conhecimento. Com essa nova política, foi possível obter uma maior similaridade, estatisticamente significativa, entre a inferência da rede e os resultados do pós-teste.

Este trabalho é uma versão estendida do artigo submetido ao Concurso de Teses e Dissertações do Congresso Brasileiro de Informática na Educação 2015, tendo sido apreciado com o terceiro lugar na categoria mestrado. Esse artigo estende o artigo [17] descrevendo em mais detalhes o processo de elaboração do modelo de inferência, as suas diversas versões, incluindo as entrevistas realizadas e os questionários aplicados em cada uma delas. Foram também adicionadas informações mais detalhadas referentes às etapas do processo de avaliação e validação dos dados, gerados por este modelo.

Este trabalho está organizado como segue. A seção 2 apresenta conceitos sobre RB e RBD, que é a estrutura de dados adotada neste trabalho para a inferência do conhecimento do aluno. Na seção 3 são listados os trabalhos relacionados, que envolvem avaliação de conhecimento usando RBs. O trabalho desenvolvido é descrito na seção 4, que consta a elaboração da rede e as etapas realizadas para o levantamento das probabilidades da mesma. A seção 5 relata as etapas de avaliação desta rede, e, finalmente, na seção 6, as considerações finais e trabalhos futuros são apresentados.

2 Redes Bayesianas

Redes Bayesianas (RB) são modelos probabilísticos utilizados para descrever a relação entre variáveis aleatórias. Representam, de forma compacta, a distribuição conjunta de probabilidade de um conjunto de variáveis [11]. Elas são amplamente utilizadas na modelagem de componentes de STIs, como os Modelos de Aluno, uma vez que estas estruturas são capazes de lidar com incerteza, característica de dados relacionados ao conhecimento humano. Outra vantagem das RBs é a sua capacidade de representação da relação entre as variáveis.

Estruturalmente, as RBs são representadas através de um Grafo Dirigido Acíclico, cujos vértices representam as **variáveis** e as arestas as **relações** de dependência entre as mesmas. Cada variável pode assumir um conjunto finito de valores, denominados **estados**, que possuem determinadas probabilidades de ocorrência. Estas probabilidades estão definidas nas **tabelas de probabilidade**, presentes em cada vértice (ou nodo), e definem a relação entre as variáveis. Por exemplo, se uma variável Y estende uma aresta para X , então as probabilidades de X estão condicionadas a Y , ou seja, sua probabilidade é definida por $P(X|Y)$ [12]. Desse modo, é dito que X é dependente

de Y . Variáveis podem possuir diversos estados, no entanto, é comum encontrar RBs com dois estados.

O processo de inferência em RBs ocorre por meio da apresentação de evidências e da observação da alteração das probabilidades da rede. **Evidências** são certezas que são observadas no ambiente modelado, como: “o estudante acertou um passo na equação, ele aplicou a operação de divisão”. Quando uma evidência é apresentada à rede, o nodo correspondente àquela observação tem o seu estado correspondente alterado para 100% de probabilidade de ocorrência. Deste modo, a respectiva variável é dita **instanciada** [12]. Para o exemplo anterior, o nodo que representa a aplicação da operação de Divisão teria o seu estado “Correto” alterado para 100%.

2.1 Redes Bayesianas Dinâmicas

RBs são empregadas na realização de inferências momentâneas, como verificar o comportamento de uma determinada variável dada uma situação específica. Por exemplo: “Dado que um estudante resolveu corretamente 80% dos exercícios de soma de frações, o quanto, neste momento, é possível afirmar que ele conhece deste assunto?”.

Porém, se é desejada uma observação a longo prazo, ou seja, acompanhar a evolução de algum processo, é necessário usar outro tipo de RB. Por exemplo, “Anteontem um estudante resolveu corretamente 80% das operações de soma de frações, ontem foram 90%, hoje foram 92%, quais as perspectivas para amanhã?”. Para este tipo de inferência ao longo do tempo, nas quais os estados das variáveis se alteram e as ações anteriores influenciam nas ações futuras, há um tipo específico de RB que se mostra mais efetiva: as Redes Bayesianas Dinâmicas (RBDs) ou Redes Bayesianas Temporais.

Elas introduzem o conceito de *time-slice*, ou instante de tempo, que representa a ocorrência de um evento em um momento no tempo. No exemplo acima, um *time-slice* poderia representar cada dia (ou cada passo) em que o estudante resolveu corretamente as operações de soma de frações. De fato, uma RBD pode ser representada como uma sequência de RBs unidas, nas quais cada uma representa um *time-slice*. Este processo é denominado *unrolling*, ou **desenrolar** [11].

Uma característica importante das RBDs é que elas introduzem três novas probabilidades que devem ser definidas: (i) a distribuição de probabilidade do estado inicial, $P(X_0)$; (ii) as probabilidades de transição entre os *slices* $P(X_{t+1}/X_t)$; e (iii) a probabilidade dos **nodossensores** $P(E_t/X_t)$, os nodos que receberão as evidências em cada *slice*.

3 Trabalhos Relacionados

Os Sistemas Tutores Inteligentes diferenciam-se dos demais *softwares* educacionais por serem capazes de ensinar e prover assistência durante a resolução de problemas aos aprendizes de forma personalizada. Para que isto seja possível, eles são capazes de avaliar o conhecimento de cada aluno.

Na literatura encontram-se diversos trabalhos referentes a esta área de avaliação de conhecimento do estudante. Um dos principais trabalhos é o *Knowledge Tracing*, apresentado inicialmente em [13]. Este trabalho avalia o conhecimento procedural do estudante, ou seja, ele calcula a probabilidade do aluno aplicar corretamente uma regra de conhecimento. Este trabalho utilizou, inicialmente, cálculos de probabilidades e, posteriormente, foi portado para uma estrutura de RBD, conforme mostrado em [14]. Em ambas as versões, ele exige a definição de quatro parâmetros: (i) o conhecimento inicial do estudante ($P(L_0)$), (ii) a probabilidade dele transitar do estado “Não Sabe” para o estado “Sabe”, a taxa de aprendizado ($P(T)$), (iii) a probabilidade de um erro por falta de atenção (*slip*) ($P(S)$), e (iv) a probabilidade de um acerto acidental (*guess*) ($P(G)$). Estudos mais recentes [15] buscam tornar este modelo mais adaptável, através de ajustes do parâmetro $P(T)$, tornando-o mais individualizado para cada estudante.

Para o domínio de Álgebra, encontram-se os trabalhos de Millan [9] e Duijnhoven [10], que utilizam abordagens diferentes para realizar a avaliação do conhecimento dos alunos. O primeiro utiliza uma RB para avaliar o conhecimento declarativo (conceitos algébricos), atribuindo um nodo para indicar o percentual de conhecimento médio do aluno em equações de 1º grau. Este trabalho utiliza como evidência cada exercício resolvido pelos alunos. Deste modo, a rede possui um nodo representando cada atividade desenvolvida, cujos estados indicam se a mesma está “Correta” ou “Incorreta”. Ela é modelada em uma estrutura hierárquica, na qual a base desta hierarquia são as atividades e os conceitos mais específicos. Estes conhecimentos mais básicos, por sua vez, vão influenciar nodos mais genéricos até o topo da hierarquia, “Equações de 1º grau”. Este último mostrará o percentual de quanto o aluno conhece sobre equações de 1º grau.

Por outro lado, o segundo trabalho [10] avalia, também através de uma RB, o conhecimento procedural unido às principais falsas concepções do estudante, classificando-o em um dos três grupos de nível de conhecimento: *High*, *Medium* ou *Low*. Da mesma forma que o primeiro trabalho [9], as evidências desta rede também são as atividades resolvidas pelos estudantes, com o diferencial de também considerar as principais falsas concepções dos aprendizes. Não utiliza uma estrutura hierárquica para estruturar os nodos; ao invés disso, utiliza uma

estrutura onde nodos de conteúdos mais genéricos influenciarão nodos mais específicos, ao contrário do que é realizado no primeiro trabalho.

Comparando-se os três trabalhos, observa-se que estes dois últimos empregam itens de avaliação, ou seja, atividades e exercícios para realizar esta avaliação. Já o *Knowledge Tracing* se foca diretamente na aplicação da regra de conhecimento. Deste modo, a estrutura do mesmo não necessita ser alterada a cada nova tarefa inserida no sistema, o que é uma vantagem sobre os demais. É observado que em todos os trabalhos é considerada a probabilidade do estudante “adivinhar” a resposta e a probabilidade do mesmo errar por falta de atenção. Geralmente, estas probabilidades possuem um teto de 30% [16].

Conforme apresentado anteriormente, o trabalho proposto apresenta uma proposta de RBD para avaliar o conhecimento algébrico do estudante. Esta apresenta como diferencial a sua capacidade de avaliar tanto o conhecimento conceitual (declarativo), quanto procedural do estudante, utilizando um modelo de inferência que é independente de problemas (equações). Conforme apontado, esta independência de problemas também é observada no *Knowledge Tracing*, porém o mesmo não modela a relação entre as unidades de conhecimento; ele as analisa de forma individualizada. Deste modo, o conhecimento de um conceito/operação x_1 não influencia na probabilidade do estudante saber outro conceito/operação x_2 . Como na álgebra há esta relação de dependência entre as unidades de conhecimento, essa característica é desejável, pois isso pode permitir a inferência do conhecimento do estudante em habilidades e conceitos não diretamente trabalhados nos problemas.

4 Metodologia

O presente trabalho consiste em um componente de avaliação do conhecimento algébrico do estudante. Como diferencial, o trabalho infere dois tipos de conhecimento do estudante: o conhecimento declarativo (os conceitos algébricos, tal como entender o significado do sinal de igualdade em uma equação) e o conhecimento procedural (o emprego das operações algébricas, por exemplo, saber somar incógnita) [17]. Esses tipos de conhecimentos estão estruturados em uma RBD e estabelecem relações entre si.

São usadas como evidências as operações algébricas, que são aplicadas pelos estudantes ao resolver uma etapa (passo) de uma equação, e as principais falsas concepções identificadas. Através deste tipo de evidência, é possível obter um componente de inferência que seja independente de equações, ou seja, este modelo pode ser aplicado para avaliar o conhecimento algébrico dos estudantes sem a necessidade de alterá-lo para cada tipo de equação que for resolvida.

Esta capacidade é possível, pois o PAT2Math, o STI ao qual este modelo será integrado, é um *step-based tutor*, ou seja, um tutor que auxilia o estudante a resolver uma tarefa passo-a-passo, oferecendo auxílio em cada uma das etapas. O mesmo possui um sistema especialista que detecta as operações aplicadas pelo estudante ao resolver equações de 1º grau e identifica as suas principais falsas concepções. A operação aplicada pelo estudante é inferida por um sistema especialista que utiliza os dados do passo anterior da equação e do passo atual informado pelo aluno na resolução da equação. Sendo assim, este sistema especialista será o responsável por encaminhar à rede todas as informações necessárias.

Desta forma, uma vez integrado ao PAT2Math, o processo de avaliação de conhecimento ocorrerá durante a resolução das tarefas, conforme o exemplo a seguir:

1. O modelo, isto é a RBD, é inicializado em t_0 ;
2. O sistema apresenta uma equação para o aluno resolver: por exemplo, $2x+3=0$;
3. O aluno resolve uma etapa (passo), $2x=-3$, e submete ao tutor para correção;
4. O sistema especialista, do STI, identifica a aplicação da operação algébrica de princípio aditivo por parte do estudante;
5. A RBD recebe do sistema especialista a informação da correta aplicação do princípio aditivo;
6. Na rede, em t_1 , esta evidência é apresentada, ou seja, o nodo responsável pelo princípio aditivo tem o seu estado “Correct” alterado para 100% de probabilidade de ocorrência;
7. O estudante recebe um *feedback* positivo;
8. O estudante submete a próxima etapa: $x=2/-3$;
9. O sistema especialista identifica uma falsa concepção no princípio multiplicativo;
10. Na rede, em t_2 , são apresentadas as evidências de ocorrência da falsa concepção no princípio multiplicativo, e a evidência de que o mesmo foi aplicado incorretamente (estados “True”, para a falsa concepção, e “Incorrect”, para o princípio multiplicativo, em 100%);
11. O estudante recebe um *error feedback*, indicando um erro no princípio multiplicativo;
12. O estudante corrige a etapa: $x=-3/2$;
13. O sistema especialista identifica a correta aplicação do princípio multiplicativo;

14. Na rede, em t_3 , esta evidência é apresentada, com o respectivo nodo com o estado “Correct” em 100%;

15. O estudante recebe um *feedback* positivo;

16. A resolução da equação é encerrada.

Após as n etapas até a solução, o *time-slice* t_n representará a inferência atual da rede quanto ao conhecimento do aluno. Cada *time-slice* da RBD representa o conhecimento do aluno em cada momento da resolução, conforme apontado pelas evidências até aquele momento.

Caso o estudante omitisse algum passo ou inserisse apenas a resposta final do exercício, as informações quanto aos passos intermediários seriam inferidas pelo sistema especialista. Como há equações com mais de um caminho de solução, em uma situação destas, o sistema especialista optaria por um caminho de resolução que pode, ou não, ser o que o aluno utilizou. Deste modo, quanto mais detalhada for a resolução, quanto mais passos forem demonstrados pelo estudante, melhor e mais precisa será a inferência do modelo.

4.1 Estrutura da Rede Bayesiana

Esta rede foi elaborada segundo a estrutura causal das RBs. Isto significa que a relação entre os nodos define que o conhecimento de um conteúdo/operação implica na possibilidade do aluno compreender um conteúdo seguinte.

Inicialmente, esta rede foi definida como um mapa conceitual, que foi elaborado com a colaboração de uma professora de matemática. Este mapa descrevia a relação entre as operações de 1º grau; nesta etapa do desenvolvimento apenas as operações foram consideradas. Este mapa, por sua vez, deu origem a primeira versão da rede, a Rede I. As avaliações conduzidas com esta rede mostraram resultados interessantes, uma vez que o resultado da inferência foi similar aos pareceres da professora [18]. Nesta versão, a rede era dependente de exercício e foi modelada como uma rede Bayesiana.

Figura 1: Relação Temporal entre os nodos

Tendo como objetivo torná-la independente de exercício, a estrutura da Rede I foi portada para uma estrutura de rede Bayesiana Dinâmica, tendo como evidência a aplicação (in)correta de uma operação para resolver um passo de uma equação de 1º grau. Por se tratar de uma RBD, algumas probabilidades extras foram necessárias; adotou-se a estrutura temporal apresentada por Reye [14] para o *knowledge tracing*. Esta estrutura é apresentada na Figura 1 e descreve a relação entre os nodos do tempo t_{n-1} e t_n . Esta relação descreve que a probabilidade do estudante compreender uma nova operação é dada por $P(O_n/O_{n-1}, P_{n-1})$, ou seja, a probabilidade do estudante compreender a operação O, dado que ele já possuía o conhecimento, ou não, e que o passo P foi aplicado (in)corretamente. A seção seguinte abordará em mais detalhes esta questão das probabilidades.

Uma vez convertida em uma RBD, a Rede II foi submetida à mesma avaliação realizada com a Rede I, porém a Rede II não apresentou resultados relevantes [19]. Deste modo, foram conduzidas entrevistas com professores de matemática do ensino superior com mais de cinco anos de experiência, a fim de revisar a estrutura da Rede II.

Destas entrevistas constatou-se que: (i) somente as relações entre as operações de 1º grau não são suficientes para indicar se um aluno conhece ou não uma operação, e (ii) foi apontada a dificuldade dos estudantes em compreender os conceitos básicos da álgebra, como o conceito de equação, igualdade e o princípio por trás da operação inversa.

Figura 2: Estrutura da Rede Bayesiana

Com base nas entrevistas realizadas e trabalhos consultados, a Rede II foi reestruturada para comportar tanto as operações, quanto os conceitos algébricos e as falsas concepções dos estudantes. Esta nova versão, a Rede III, passou por avaliações de professores de matemática até assumir a forma apresentada na Figura 2, conforme descrito em [20].

Nesta modelagem, por exemplo, se um aluno conhece a regra do Mínimo Múltiplo Comum (MMC), então ele

possui um certo grau de probabilidade de compreender a regra da Soma de Frações (AD/SB). Da mesma forma, se ele conhece o conceito “Igualdade dos Membros”, ele pode ser capaz de conhecer os princípios aditivo (PAd) e multiplicativo (PMult). O oposto também é verdade; se o aluno conhece a Divisão de Frações (DV $\rightarrow$ Frações), ele também conhece a Multiplicação de Frações (MT $\rightarrow$ DV), uma vez que esta segunda é requisito para a primeira.

4.2 Probabilidades Condicionais

Conforme apresentado na subseção anterior, a rede é estruturada de modo a definir a precedência entre as operações. Deste modo, as probabilidades condicionais definem as chances que um aluno tem de compreender determinado conteúdo, dados os pré-requisitos deste conteúdo. Por exemplo, as probabilidades dos princípios aditivo (PAd) e multiplicativo (PMult) estão diretamente relacionadas à probabilidade do estudante compreender o conceito “Igualdade dos Membros”.

Para obter estes dados, foram realizadas entrevistas com dois professores de matemática que atuam no ensino fundamental e médio e com mais de cinco anos de experiência no meio. Para facilitar a obtenção destes dados, utilizou-se um modelo de questionário apresentado em [21]. Este modelo apresenta uma escala numerada com as probabilidades e termos que expressam os valores das probabilidades a serem obtidas, conforme é mostrado na Figura 3. Acompanhando esta escala, é apresentada uma descrição textual da probabilidade a ser obtida com o questionário; por exemplo: “Considerando que um estudante consiga realizar a multiplicação de frações, quais as chances dele resolver uma divisão de frações?”

Figura 3: Escala utilizada no questionário, baseada em [21]

Conforme [21], esta escala auxilia o entrevistado a “melhor visualizar a posição do intervalo que expressa a probabilidade que ele tem em mente”.

Deste modo, a probabilidade de um nodo de conhecimento K é igual ao somatório dos pesos dos nodos que o influenciam, conforme mostra a Equação 1.

$$P(K | K_1, \dots, K_n) = \sum_{i \in I} w_i, \quad I = \{i \in \{1, 2, \dots, n\} | K_i = \text{Knows}\}$$

Equação 1: Fórmula que define as probabilidades entre os conceitos e operações

Logo, a probabilidade condicional do nodo de Adição (AD) é igual a soma dos pesos das propriedades Comuta-

tiva e Associativa; pesos estes obtidos através dos questionários.

No entanto, por tratar-se de uma RBD, também são necessárias as probabilidades relacionadas à parte temporal, bem como as probabilidades relacionadas às evidências. Para isto, foram definidos parâmetros de configuração, que foram baseados no *Knowledge Tracing*:

- o conhecimento inicial - $P(L_0)$
- a taxa de *guess* - $P(G)$
- a taxa de *slip* - $P(S)$
- a taxa de transição do estado “Não Aprendeu” para o estado “Aprendeu”, a Taxa de Aprendizado [14] - $P(T)$
- a Taxa de Lembrança, que representa as chances do aluno reter o conhecimento na transição de um *time-slice* para o outro

Com base nestes parâmetros, foram calculadas as seguintes probabilidades para configurar a parte temporal da rede:

- $P(O_n/L_{n-1})$ = A chance de o aluno acertar um passo, uma vez que ele conheça o assunto.
- $P(O_n/\neg L_{n-1})$ = A chance de o aluno acertar um passo, uma vez que ele não conheça o assunto.
- $P(L_n/L_{n-1}, O_n)$ = A chance de o aluno conhecer determinado conteúdo, dado que ele já conhecia este conteúdo e respondeu corretamente o passo, ou seja, a sua taxa de lembrança - $\rho(O_n)$.
- $P(L_n/\neg L_{n-1}, O_n)$ = A chance de o aluno conhecer determinado conteúdo, dado que ele não conhecia este conteúdo antes e ele respondeu corretamente o passo, ou seja, a sua taxa de aprendizado - $\lambda(O_n)$ ou $P(T)$.

Também foram definidos parâmetros para determinar a força da influência dos requisitos quando estes são conceitos e operações. Isso permite configurar na rede quais os conceitos (o conhecimento declarativo) ou as operações (o conhecimento procedural) que exercem maior influência no atual conhecimento do aluno.

No desenvolvimento desta rede, estes níveis de influência foram ajustados para 60% (conceitos) e 30% (operações). Quanto aos demais parâmetros, estes foram ajustados para: $P(L_0)=0,5$; $\lambda(O_n)=0,2$; $P(G)=0,3$; $P(S)=0,3$ e $\rho(O_n)=0,99$. Este último foi mantido próximo a 1, de modo a permitir uma propagação de conhecimento entre os *slices* sem haver perda considerável de conhecimento. $P(G)$ e $P(S)$ foram definidos com valores pesquisados na literatura [16]. Já, $\lambda(O_n)$ ou $P(T)$, foi definido

com um valor inferior ao utilizado na literatura [14], uma vez que com o valor referenciado (0,4) a rede apresentou um conhecimento com uma taxa elevada de crescimento. Como a rede desta referência não possui relação entre outros nodos de conhecimento, optou-se por utilizar a metade deste valor, que apresentou resultados aceitáveis, com um crescimento de conhecimento mais moderado.

5 Avaliação da Rede

Estando modelada a rede, a etapa seguinte foi verificar a sua capacidade de inferência. Como não havia uma integração entre a rede e o STI PAT2Math até o momento em que este foi aplicado em sala de aula, os dados foram coletados nesta escola pelo STI, assim como armazenados em seu banco de dados para posterior emprego na rede. Este processo foi similar ao que ocorreria se houvesse uma integração entre a RBD e o STI. A rede, então, emitiu um parecer referente ao percentual de conhecimento de cada aluno em cada conceito e operação de 1º grau. Este diagnóstico da rede foi comparado com um pós-teste, realizado pelos mesmos alunos, de modo a verificar a similaridade entre os dados.

5.1 Obtenção dos Dados

No segundo semestre de 2014, o projeto PAT2Math fechou uma parceria com uma escola de ensino fundamental e médio da cidade de São Leopoldo, no Rio Grande do Sul, na qual o STI seria utilizado na disciplina de matemática das turmas de 7º ano.

Ficou definido que o STI seria utilizado pelas duas turmas de 7º ano daquela escola, totalizando 59 alunos. O uso do STI, efetivamente, ocorreu nos meses de Outubro e Novembro de 2014. Deste modo, uma vez por semana, os alunos se dirigiam ao laboratório de informática da escola para resolver equações de 1º grau no PAT2Math, durante um período de aula (50 minutos). As 148 equações inseridas no sistema foram obtidas com a própria professora, ou seja, eram equações que são utilizadas no dia a dia, em sala de aula. Estas equações foram agrupadas em 9 planos de aula, de acordo com o seu nível de dificuldade.

Todas as respostas submetidas pelos estudantes foram registradas no banco de dados do sistema. Posteriormente, estes dados foram analisados e organizados em arquivos *Comma-separated values* (CSV), dois por aluno (operações e falsas concepções), totalizando 118 arquivos. Para cada submissão do aluno, foram coletados os seguintes dados:

- A equação inicial: aquela que é apresentada ao aluno no início do exercício;

- O passo (a resposta) do aluno: a equação que o aluno supõe que seja o próximo passo de resolução da equação;
- A operação aplicada pelo estudante no passo. Pela resposta do aluno, o sistema especialista do PAT2Math infere qual operação ele aplicou.
- Um valor booleano indicando se o passo do estudante está correto ou não;
- Uma mensagem do sistema;
- O *error feedback*, em caso de erro.

Durante o período em que o PAT2Math foi utilizado na escola, foi observado o comportamento dos estudantes ao interagir com o sistema, e, desta forma, as seguintes dificuldades, em álgebra, foram observadas:

- **Somas e Subtrações de incógnitas com constantes:** O estudante realiza uma operação de soma/subtração entre incógnitas e constantes, operações estas que são válidas apenas na multiplicação e na divisão. Exemplo: $3x+4=0 \rightarrow 7x=0$.
- **Sinal negativo:** O aluno opera de forma incorreta os termos da operação que possuem sinal negativo ou simplesmente os ignora. Exemplo: $4x-3x=0 \rightarrow 7x=0$.
- **Operação inversa é aplicada incorretamente:** Ao aplicar a operação inversa para as operações soma ou subtração, “movendo um termo para o outro lado da igualdade da equação”, a operação que acompanha o termo é mantida. Exemplos: $-4x-3=0 \rightarrow -4x=-3$, e $-4x=-3 \rightarrow x=-3*-4$.
- **Operação inversa aplicada corretamente, porém o sinal é trocado:** Ao aplicar a operação inversa para as operações de multiplicação e divisão, “movendo um termo para o outro lado da igualdade da equação”, o aluno inverte incorretamente o sinal do termo que foi movido. Exemplo: $-4x=-3 \rightarrow x=-3/4$.
- **Desconhecimento da operação inversa dos termos:** Ao aplicar a operação inversa, “movendo um termo para o outro lado da igualdade da equação”, é aplicada uma operação incorreta, ou seja, um termo que está multiplicando/dividindo a incógnita, passa a somar/subtrair no outro lado da igualdade. Exemplo: $-4x=-3 \rightarrow x=-3-4$.

- **Operação inversa correta, porém posição dos termos incorreta:** Ao aplicar a operação inversa da multiplicação, “movendo um termo para o outro lado da igualdade da equação”, o termo passa a ser o numerador da fração, em vez de ser o denominador. Exemplo: $-4x = -3 \rightarrow x = -4/-3$.
- **Erros diversos:** Erros, ou deslizos, ao realizar as operações básicas, somas, subtrações, multiplicações e divisões incorretas. Exemplo: $2x = 26 + 30 \rightarrow 2x = 66$.

Tais dificuldades foram, também, inferidas pelo sistema especialista do PAT2Math e, desta forma, foram considerados pela rede.

No final do período de utilização do sistema, na primeira semana de dezembro, os alunos resolveram um pós-teste como parte da avaliação do STI. Este visou verificar se o PAT2Math contribuiu positivamente para o aprendizado dos estudantes. Este teste também foi considerado para validar a inferência da rede. Para isto, o teste foi analisado, com o objetivo de obter o percentual de acerto de cada estudante, em cada operação de 1º grau. Nesta análise, constatou-se a aplicação das operações de Adição, Subtração, Multiplicação, Propriedade Distributiva, Simplificação de frações, Princípio Aditivo, Princípio Multiplicativo, MMC e, muito raramente, a Soma de Frações. Obteve-se um total de 49 pós-testes, dos 59 alunos que utilizaram o PAT2Math.

Figura 4: Histograma de Execuções da Rede

5.2 Resultados

Calculados os percentuais de acerto de cada aluno, nos pós-testes, os seus respectivos arquivos CSV de registro foram selecionados e submetidos à análise da RBD. O processo buscou simular a forma como a rede funcionaria caso tivesse sido integrada ao PAT2Math. Deste modo, cada linha do arquivo representou uma submissão de resposta do estudante. Estas “submissões” foram repassadas à rede da seguinte forma:

- Se foi um **acerto**, então o nodo de evidência da respectiva operação (nome iniciado por “Passo_”) é instanciado com o estado “Correct”, e os nodos das falsas concepções correspondentes (nome iniciado por “Misc_”) são instanciados com o estado *False*, indicando a não ocorrência de uma falsa concepção.

- Se for um **erro**, então o nodo da Falsa Conceção identificadora é instanciado com o estado “True”; o nodo de evidência da operação correspondente a esta Falsa Conceção é instanciado com o estado “Incorrect”. As demais Falsas Conceções, relacionadas a esta operação, também são instanciadas com o estado “False”.

Após a execução da rede, os percentuais obtidos nos nodos representando o conhecimento do estudante para cada operação algébrica foram comparados com os percentuais dos pós-testes. Uma vez que o pós-teste reflete o atual conhecimento do aluno, a rede deveria chegar a uma conclusão similar, analisando os resultados das interações com o STI. Deste modo, elaborou-se um histograma com as principais operações observadas nos pós-testes, como mostra a Figura 4(a). Neste histograma foi calculada a diferença absoluta entre os resultados da rede e o percentual de acerto dos estudantes. Quanto menor o valor, menor a diferença entre a rede e o atual conhecimento do aluno, refletido pelo pós-teste.

Observa-se neste histograma (Figura 4(a)) que não há uma grande presença de valores no intervalo de 5%, que seria o desejável. Isto ocorre devido às limitações da rede, que necessitam que a mesma seja reiniciada, ou seja, voltar a t_0 após 20 *time-slices*, por questões de desempenho. Quanto maior o número de *time-slices*, mais tempo a rede leva para calcular as probabilidades. Vinte foi o número escolhido para o número máximo de *time-slices*, pois foi observado que, ao resolver uma equação, o sistema especialista leva menos de 20 passos para resolver uma equação mais complexa. Considerando que ele resolve uma equação de forma bem detalhada, resolvendo uma operação por vez, este número foi considerado aceitável. Foi verificado que, caso a rede não fosse reiniciada, o número de *time-slices* necessários para a rede comportar todos os passos submetidos por um estudante seria, em média, 390 *time-slices*. Esse elevado número faria com que a rede demorasse alguns minutos para realizar a inferência, o que não é aceitável em um sistema web online, visto que esta inferência ocorrerá no momento em que o aluno submeter um passo da equação para correção. Logo, não deve demorar mais do que alguns poucos segundos.

No processo de reinício da rede, somente a informação do último *time-slice* é passada adiante. Visando reduzir esta perda de informações, adotou-se uma política de limiares: se uma porcentagem em algum nodo, com exceção dos nodos de evidência e falsa concepção, ultrapassar este valor de limiar, então a variável é instanciada com o estado “Knows” no *time-slice* correspondente.

Definida esta nova política, mais duas execuções da rede foram realizadas, cada uma com um valor diferente

de limiar. Gerou-se os histogramas correspondentes, da mesma forma que a primeira execução. A Figura 4(b) apresenta o histograma com um limiar de 96% e a Figura 4(c) o histograma de limiar 98%. Tais valores foram utilizados por serem valores próximos a 100%. Desta forma, a rede não considera que o estudante conhece determinado conceito/operação muito rapidamente.

Intervalo de Equivalência					
	0,05	0,1	0,15	0,2	0,3
AD	p=0,999	p=0,908	p=0,246	p=0,004	-
SB	p=0,203	p=0,003	-	-	-
MT	p=0,531	p=0,121	p=0,009	-	-
DM	p=0,165	p=0,015	-	-	-
SP	p=0,25	p=0,028	-	-	-
PA	p=0,086	p=0,002	-	-	-
PM	p=0,108	p=0,001	-	-	-
MM	p=0,196	p=0,021	-	-	-
AF	p=1	p=1	p=1	p=1	p=1

Tabela 1: Resultados do Teste de Equivalência para a versão sem limiar

Intervalo de Equivalência					
	0,05	0,1	0,15	0,2	0,3
AD	p=0,192	p=0,002	-	-	-
SB	p=0,586	p=0,044	-	-	-
MT	p=0,402	p=0,045	-	-	-
DM	p=0,798	p=0,29	p=0,03	-	-
SP	p=0,888	p=0,346	p=0,024	-	-
PA	p=0,725	p=0,163	p=0,006	-	-
PM	p=0,77	p=0,129	p=0,002	-	-
MM	p=0,272	p=0,039	-	-	-
AF	p=1	p=1	p=1	p=1	p=1

Tabela 2: Resultados do Teste de Equivalência para a versão com 96% de limiar

Intervalo de Equivalência					
	0,05	0,1	0,15	0,2	0,3
AD	p=0,999	p=0,888	p=0,215	p=0,003	-
SB	p=0,083	p=0,001	-	-	-
MT	p=0,334	p=0,4	-	-	-
DM	p=0,502	p=0,096	p=0,005	-	-
SP	p=0,803	p=0,24	p=0,013	-	-
PA	p=0,337	p=0,024	-	-	-
PM	p=0,545	p=0,038	-	-	-
MM	p=0,191	p=0,022	-	-	-
AF	p=1	p=1	p=1	p=1	p=1

Tabela 3: Resultados do Teste de Equivalência para a versão com 98% de limiar

Observa-se nestes histogramas que o limiar de 96% produziu um resultado mais desejável, ou seja, maior presença de valores no intervalo de 5%. Ressalta-se que a Soma de Frações (AF) foi desconsiderada neste histograma uma vez que a mesma não ocorreu de forma frequente nos pós-testes.

Por fim, foi necessário verificar se os resultados da rede e dos pós-testes são equivalentes estatisticamente. Aplicou-se, então, o *two one-sided test* (TOST) [22], para o qual define-se um intervalo entre δ_1 e δ_2 e aplica-se dois testes de hipótese. Para que os dados sejam equivalentes, as hipóteses nulas $H_{0a}(\mu \leq \delta_1)$ e $H_{0b}(\mu \geq \delta_2)$ devem ser rejeitadas.

O símbolo μ representa a diferença entre as amostras da rede e do pós-teste. Define-se, também, a hipótese alternativa $\delta_1 \leq \mu \leq \delta_2$. Logo, se as hipóteses nulas forem rejeitadas, μ estará dentro do limite δ_1 e δ_2 e os dados serão considerados equivalentes. Para que as hipóteses nulas sejam rejeitadas, é necessário que os valores P_1 e P_2 sejam menores que α , que, neste trabalho foi de 0,05, portanto um intervalo de confiança de 95%. A Tabela 1 mostra os resultados do teste de equivalência para a execução da rede sem a política de limiares, já as Tabelas 2 e 3 mostram os resultados do teste para o limiar de 96% e 98%. Da mesma forma que o histograma, não foi possível mostrar a equivalência para a operação de soma de frações por não haver dados suficientes no pós-teste. Observa-se que neste teste a execução com o limiar de 96% também obteve o melhor resultado pois, com exceção da operação de Soma de Frações, a sua equivalência foi comprovada com intervalos menores. A equivalência foi comprovada com um δ máximo de 0,15, o que não ocor-

reu nas outras execuções, nas quais ambas obtiveram um δ máximo de 0,2.

Verificada a equivalência entre os dados da rede e os do pós-teste, o próximo passo consistiu em confirmar a melhora na capacidade de inferência desta rede. Para isso, comparou-se os resultados obtidos com esta rede (Rede III) com os resultados obtidos com a sua versão anterior (Rede II). Foi aplicado um teste *t* não pareado sobre esses resultados, cuja hipótese nula (H_0) afirma que a Rede III tem uma capacidade de inferência inferior à Rede II. Em contrapartida, a hipótese alternativa (H_1) define que a capacidade de inferência da Rede III é superior.

Com um intervalo de confiança de 95% ($\alpha=0,05$), obteve-se $p=0,0028$. Sendo este menor que α , H_0 foi rejeitada e H_1 pode ser assumida. Portanto, foi possível afirmar que há indícios de uma melhora significativa na capacidade de inferência por parte da Rede III em comparação à Rede II.

6 Conclusão

Ser capaz de identificar o conhecimento de cada aluno permite aos STIs adequarem a sua instrução aos mais diversos tipos de estudantes, direcionando o ensino de acordo com as características, habilidades e dificuldades de cada aluno. Tal tarefa é desempenhada pelo componente Modelo de Aluno que, além de realizar esta inferência, também é responsável por armazenar diversos tipos de informações referente ao aluno, tais como perfil, estados afetivos, entre outros. Estas informações permitem ao STI prover uma instrução mais individualizada.

Este trabalho propôs um modelo de inferência do conhecimento algébrico do estudante através de Redes Bayesianas Dinâmicas. Atualmente, diversos trabalhos utilizam as Redes Bayesianas para realizar este tipo de inferência devido a sua capacidade de lidar com incertezas. O grande desafio deste trabalho foi modelar o conhecimento algébrico, de forma que a rede seja capaz de inferir tanto os conceitos algébricos, quanto as operações e as principais falsas concepções dos estudantes.

A independência de exercício também é uma característica importante do trabalho. Nos trabalhos pesquisados, os problemas eram utilizadas como evidências, exigindo, portanto, que a cada novo exercício inserido no sistema, novos nodos também fossem integrados à rede. Esses nodos de evidência representavam a (in)correta resolução de cada exercício. Ao adotar a estrutura dinâmica e as operações como evidências, cada *time-slice* representa a (in)correta aplicação de uma operação de 1º grau. Como toda equação de 1º grau é resolvida por um subconjunto destas operações, não há informação nova para ser inserida na rede.

De modo a obter um modelo mais condizente com a realidade, foram realizadas diversas entrevistas com professores do ensino fundamental, médio e superior. Tais entrevistas auxiliaram tanto na estruturação da rede quanto na definição das probabilidades entre os nodos de conhecimento.

Outra questão que foi considerada no desenvolvimento da rede foi o desempenho da mesma, uma vez que, ao aumentar, consideravelmente, o número de *time-slices* da rede, o seu desempenho diminui drasticamente. Desta forma, outras estratégias foram empregadas para contornar esta queda de desempenho, como o limite máximo de *time-slices* e a política de limiares. O desempenho é um aspecto importante em STIs que devem realizar a inferência em tempo real a fim de personalizar a assistência ao estudante. Deste modo este processo deve ocorrer em um tempo aceitável, o que não pode durar mais de alguns segundos.

Na avaliação da rede, foram coletados dados de resolução de equações, provenientes da interação de alunos do 7º ano com o STI PAT2Math. O PAT2Math é um STI que fornece assistência individualizada aos estudantes durante a resolução passo-a-passo de equações. Ele provê auxílio em cada passo das equações, o que permite ao sistema obter evidências do conhecimento de cada estudante. Isto permite ao modelo de aluno inferir o conhecimento de cada aluno através do trabalho desenvolvido.

Para a avaliação, dados de 49 alunos foram submetidos à rede. Os resultados da rede para cada aluno foram comparados com os resultados da correção do pós teste, realizado pelos estudantes ao final do período de interação com o STI. Desta comparação extraiu-se a diferença absoluta. Este processo foi repetido três vezes, uma vez sem a política dos limiares e outras duas com os limiares de 96% e 98%.

Os resultados mostraram um melhor desempenho na rede em que foi aplicada um limite menor de limiar, de 96%, cuja equivalência foi mostrada através de testes estatísticos. Optou-se por não selecionar um valor muito baixo de limiar, pois ocorreria do aluno ser considerado conhecedor de determinado conteúdo muito rapidamente. Por isto, foi usado um valor relativamente alto para o limiar, mas que apresentasse resultados interessantes.

Como trabalhos futuros, pretende-se integrar este componente ao STI PAT2Math, de modo que seja possível realizar avaliações em sala de aula. Busca-se igualmente incluir outros conceitos à rede, que foram observados durante as interações dos alunos com o sistema, tais como, o conceito de número fracionário (fração); a operação fração inversa, aplicada na divisão de frações junto à multiplicação de frações; etc. Futuramente, objetiva-se igualmente agregar as operações e conceitos relativos a

equações de 2º grau, dado que na versão atual o domínio limita-se à resolução de equações de 1º grau.

Agradecimentos

O presente trabalho foi realizado com o apoio dos seguintes órgãos de fomento à pesquisa: CNPq e FAPERGS.

Referências

- [1] T. P. Carpenter, H. S. Kepner, M. K. Corbitt, M. M. Lindquist, and R. E. Reys, "Student Performance in Algebra: Results from the National Assessment," *Sch. Sci. Math.*, vol. 82, no. 6, pp. 514–531, 1982.
- [2] N. C. on Excellence in Education, "A nation at risk: The imperative for educational reform," Washington, DC, 1983.
- [3] Brasil, *Parâmetros curriculares nacionais: Matemática*. Rio de Janeiro, RJ: DP & A, 2000.
- [4] B. S. Bloom, "The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring," *Educ. Res.*, vol. 13, no. 6, pp. 4–16, 1984.
- [5] B. P. Woolf, *Building intelligent interactive tutors*, vol. 3. Morgan Kaufmann, 2008.
- [6] K. Vanlehn, "The behavior of tutoring systems," *International Journal of Artificial Intelligence in Education*, vol. 16, no. 3, pp. 227–265, 2006.
- [7] P. A. Jaques, H. Seffrin, G. Rubi, F. de Moraes, C. Guilardi, I. I. Bittencourt, and S. Isotani, "Rule-based expert systems to support step-by-step guidance in algebraic problem solving: The case of the tutor PAT2Math," *Expert Systems with Applications*, vol. 40, no. 14, Apr. 2013.
- [8] H. Seffrin, G. Rubi, and P. Jaques, "O Modelo Cognitivo do Sistema Tutor Inteligente PAT2Math," *Anais do Simpósio Brasileiro de Informática na Educação*, pp. 10–19, 2011.
- [9] E. Millán, L. Descalço, G. Castillo, P. Oliveira, and S. Diogo, "Using Bayesian networks to improve knowledge assessment," *Computers & Education*, vol. 60, no. 1, pp. 436–447, 2013.
- [10] J. Van Duijnhoven, "Knowledge Assessment using Bayesian Networks: A case study in the domain of algebraic expectation," Utrecht, 2003.
- [11] S. J. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach (2nd Edition)*. Prentice Hall, 2002.
- [12] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Francisco, CA: Morgan Kaufmann, 1988.

- [13] A. Corbett and J. R. Anderson, “Knowledge tracing: Modeling the acquisition of procedural knowledge,” *User Model. User-adapt. Interact.*, 1994.
- [14] J. Reye, “Student modelling based on belief networks,” *International Journal of Artificial Intelligence in Education*, vol. 14, pp. 1–33, 2004.
- [15] M. V. Yudelson, K. Koedinger, and G. J. Gordon, “Individualized Bayesian Knowledge Tracing Models,” *Proceedings of International Conference on Artificial Intelligence on Education. 2013*, vol. 7926, pp. 171–180, 2013.
- [16] R. S. J. d. Baker, A. T. Corbett, and V. Aleven, “More Accurate Student Modeling through Contextual Estimation of Slip and Guess Probabilities in Bayesian Knowledge Tracing,” in *Proceedings of International Conference on Intelligent Tutoring Systems*, 2008, vol. 5091, pp. 406–415.
- [17] H. M. Seffrin and Patricia A. Jaques, “Avaliando o conhecimento algébrico dos estudantes através de Redes Bayesianas Dinâmicas,” in *Anais do Simpósio Brasileiro de Informática na Educação*, 2015.
- [18] H. Seffrin, G. Rubi, and P. Jaques, “Uma Rede Bayesiana aplicada à Modelagem do Conhecimento Algébrico do Aprendiz,” in *Anais do Simpósio Brasileiro de Informática na Educação*, 2013.
- [19] H. Seffrin, G. Rubi, and P. Jaques, “A Dynamic Bayesian Network for Inference of Learners Algebraic Knowledge,” in *Proceedings of the 29th Symposium On Applied Computing*, 2014.
- [20] H. Seffrin and P. Jaques, “Modelando o conhecimento algébrico do estudante através de Redes Bayesianas Dinâmicas,” in *Anais do Simpósio Brasileiro de Informática na Educação*, 2014.
- [21] L. C. van der Gaag, S. Renooij, C. L. M. Witteman, B. M. P. Aleman, and B. . Taal, “How to elicit many probabilities,” in *Proceedings of 15th conference on Uncertainty in artificial intelligence, UAI*, 1999, no. 1999–15, pp. 647–654.
- [22] M. Meyners, “Equivalence tests – A review,” *Food Qual. Prefer.*, vol. 26, no. 2, pp. 231–245, 2012.