

Agentes Pedagógicos Bayesianos

Abstract. *The production rules constitute the major computational model used for the development of the module of taking of decision in Intelligent Tutorial Systems. This article presents an alternative mechanism of inference: a bayesian pedagogical agent to select the learning tactics. The implemented bayesian net in the module of taking of decision of the pedagogical agent uses as evidences, characteristics extracted from the interaction of the student with the STI. In order to show and to validate the inference mechanism considered, a STI was implemented.*

Resumo. *As regras de produção constituem o modelo computacional mais comumente empregado no desenvolvimento do módulo de tomada de decisão em Sistemas Tutores Inteligentes. Este artigo apresenta um mecanismo de inferência alternativo ao propor um agente pedagógico bayesiano para selecionar as táticas de ensino. A rede bayesiana implementada no módulo de tomada de decisão do agente pedagógico utiliza como evidências, características extraídas a partir da interação do aluno com o STI. A fim de exemplificar e validar o mecanismo de inferência proposto, um STI foi implementado.*

1. Introdução

O objetivo fundamental dos STI's é proporcionar uma instrução adaptada ao aluno, tanto no conteúdo como na forma [Viccari 1996]. Segundo Santos [Santos 2005], os Tutores Inteligentes podem adaptar seu conteúdo e apresentação baseado em sua interação com o aluno. Isto é feito com base nas respostas e conhecimentos dos alunos. O tutor altera o material, algumas vezes simplificando-o outras estendendo-o, de forma que possa adaptar-se aos alunos individualmente, facilitando o aprendizado do mesmo.

As possibilidades pedagógicas e didáticas de produtos de *software* com as características dos STI são extensas [Corredor 1993], pois confirmam a importância das características individuais de cada aluno, reconhecendo que a instrução deve ser individualizada, de forma a facilitar ao aluno a criação de estruturas conceituais e metodológicas adequadas a sua capacidade e interesse.

Na busca de inspiração na área da inteligência artificial para aprimorar a atuação desses sistemas, os Agentes Inteligentes têm sido aplicados aos STIs. Os agentes são denominados pedagógicos quando estão inseridos em um ambiente que compõe um sistema de ensino-aprendizagem. Este ambiente pode ser composto por uma sociedade de agentes. O termo agente pedagógico [Giraffa 1999] surgiu devido a muitos sistemas desenvolvidos para fins educacionais adotarem o paradigma de agentes.

Os agentes pedagógicos possuem algumas propriedades fundamentais de agentes inteligentes, como autonomia, capacidade social (interações e comunicação), e adaptabilidade ao ambiente.

A aplicação de agentes pedagógicos tem recebido atenção especial em ambientes educacionais que utilizam o modelo de aluno, definido por suas características psico-pedagógicas, como forma de aperfeiçoar o processo de ensino-aprendizagem. São responsáveis por acompanhar a interação do aluno com o sistema educacional e, desta forma, ser capaz de guiar as ações do aluno para que este tenha um aprendizado eficiente.

A inspiração para a pesquisa apresentada neste artigo surgiu a partir de dois trabalhos anteriormente desenvolvidos. Os trabalhos de [5,6] utilizaram agentes pedagógicos para auxiliarem no processo de ensino-aprendizagem. Nesses trabalhos, dois agentes pedagógicos foram desenvolvidos, a Dóris e o Dimi.

No primeiro trabalho [Silva 2002], o agente Dóris extrai informações a partir da interação do aluno com o sistema. Já o segundo trabalho [Santos 2002] acrescenta o agente Dimi que passa a trabalhar em conjunto com o agente Dóris. O agente Dimi é responsável por utilizar as informações fornecidas pelo agente Dóris para obter características cognitivas do aluno. Originalmente, o agente Dimi possui o módulo de tomada de decisão baseado em regras de produção. Esse módulo era responsável pela seleção de táticas de ensino a serem utilizadas na interação com o aluno.

A pesquisa relatada neste artigo, originou-se a partir das deficiências encontradas na seleção de táticas de ensino pelo agente Dimi, constatou-se que o módulo de tomada de decisão original apresentava deficiências inatas às regras de produção. Sendo assim, o módulo de tomada de decisão do agente Dimi foi totalmente refeito, utilizamos uma rede bayesiana que recebe como evidências, características extraídas pelo agente Dóris a partir da interação do aluno com o STI. O módulo de tomada de decisão, composto pelo mecanismo de inferência bayesiano possibilitou uma seleção mais adequada das táticas de ensino. A descrição do processo de re-implementação do mecanismo de inferência é descrito ao longo desse artigo.

Esse artigo está estruturado em cinco seções: a primeira, apresenta a introdução, a segunda apresenta os problemas relacionados à utilização das Regras de Produção em mecanismos de tomada de decisão. A terceira seção descreve brevemente as Redes Bayesianas e fornece indicativos de como uma RB pode ser utilizada como mecanismo de inferência. A quarta seção é dividida em três sub-seções e descreve os aspectos centrais da pesquisa: como pode ser feito o mapeamento de um Rede Bayesiana (RB) a partir de um conjunto de Regras de Produção (RP); qual foi a RB gerada e finalmente como ocorre a inferência no módulo de tomada de decisão proposto. A quinta e última seção apresenta as conclusões obtidas.

2. Deficiências encontradas na utilização de Regras de Produção

Os sistemas baseados em Regras de Produção podem falhar ao lidar com incertezas. No trabalho [Ladeira 2004], o autor justifica as deficiências das regras de produção com o seguinte argumento: “Para tratar a incerteza, os sistemas de regras de produção associam fatores de certeza às regras e definem um mecanismo para combiná-los. Formalmente, a incerteza é considerada como um valor verdade generalizado associado às fórmulas (derivadas da lógica clássica), sendo que a incerteza de qualquer fórmula pode ser calculada como uma função da certeza das suas sub-fórmulas. Esse enfoque é eficiente em computador, mas sem semântica clara e pode produzir resultados

inesperados e não intuitivos, porque a dependência entre eventos muda com o conhecimento de outros eventos. Portanto sistemas que são baseados em regras de produção para representar seu conhecimento, como os Sistemas Especialistas, dificilmente conseguem atribuir fatores de incerteza as suas explicações.”

O raciocínio coerente sob incerteza é sensível ao contexto diferentemente das regras de produção que são livres de contexto. Considere a seguinte regra de produção:

Se A então B.

De acordo com esta regra pode-se concluir B quando A ocorre na base de conhecimento, independentemente de outras informações na base de conhecimento (princípio da localidade) e independentemente de como A foi derivado (princípio da separabilidade). Princípios como “localidade” e “separabilidade” costumam falhar ao se lidar com incerteza em um sistema baseado em regras. Estas falhas podem ser demonstradas no exemplo a seguir.

Considere a seguinte regra (R1):

“se a grama está molhada de manhã, então assuma que choveu a noite (com fator de certeza FC1)”

E considere o fato (F1):

“O esguicho (de jardim) estava ligado ontem a noite”

Analisando esta regra, considerando o princípio da localidade, constata-se que não é possível confiar em R1, pois ela desconsidera a hipótese da grama estar molhada devido ao fato (F1), do esguicho estar ligado na noite passada. Dessa forma, observa-se que mesmo uma regra de produção tendo um fator de certeza absoluta, pode ter uma conclusão errada. Além disso, F1 por si só não suporta ou refuta a possibilidade de ter chovido!

Pelo princípio da separabilidade, observando-se A, pode-se concluir C, aplicando-se em cadeia as regras de produção se A então B e se B então C. Uma proposição derivada (B) pode disparar uma regra (se B então C) com o mesmo vigor do que uma proposição observada. Esse fato, em geral, não se mantém ao se lidar com incerteza. Seja R2: “se o esguicho estava ligado, então assuma que a grama está molhada de manhã (com fator de certeza FC2)” e R1: “se a grama está molhada de manhã, então assuma que choveu a noite (com fator de certeza FC1)”. A observação do esguicho ligado não permite concluir que “choveu a noite”.

A menos que se tomem cuidados especiais durante a construção da base de regras e se limite o tipo de inferência (causal ou diagnóstico), regras de produção apresentam as seguintes deficiências:

- a) Tratamento impróprio de fontes de evidências correlacionadas;
- b) Manipulação inadequada de inferências bidirecionais (ocorrência de A e de B, onde A é evidência para B e vice versa);

- c) Dificuldades em explicar as conclusões obtidas.

3. As Redes Bayesianas

Uma rede Bayesiana captura relações (que podem envolver incertezas, funções estatísticas ou mesmo serem imprecisas) entre um grupo de variáveis que são relevantes para um problema. Essas relações podem ser importantes porque elas são observáveis ou porque o seu valor é necessário para a tomada de uma decisão ou para justificar um resultado ou mesmo porque ajudam a expressar relacionamentos entre outras variáveis.

Quando uma rede Bayesiana é construída, um nodo é usado para cada variável. Os nodos são então conectados através de enlaces direcionados. Se existe um link (enlace) que inicia a partir do nodo A e termina no nodo B, então o nodo A é chamado de “pai” do nodo B, e este é denominado “filho” de A. Este conceito é recursivo e aplica-se aos demais nodos da rede. Um link partindo de A para B indica que A causa B, que A parcialmente causa ou predispõe B, que A e B são funcionalmente relacionados ou que A e B são estatisticamente correlacionados. O estágio final de uma rede Bayesiana inclui relações probabilísticas que são fornecidas para cada nodo, as quais expressam as probabilidades de cada estado do nodo assumir cada um dos valores, condicionado aos valores do nodo “pai”.

Após a sua construção, uma rede Bayesiana pode ser aplicada a um caso particular funcionando como um mecanismo inferencial. Para cada variável que saibamos o valor, o valor representa uma evidência que produz uma inferência probabilística para encontrar crenças para todas as outras variáveis. Dependendo da estrutura da rede e de quais nodos recebem evidências ou mostram crenças, é possível calcular valores para diagnóstico, previsão, classificação, cálculos lógicos ou aritméticos ou uma combinação de todos para atingir inferências probabilísticas. As crenças calculadas são chamadas de probabilidades a posteriori (uma vez que as probabilidades a priori são aquelas que existiam antes de qualquer evidência ser apresentada a rede). Uma inferência probabilística que ocorra usando uma rede Bayesiana é chamada de atualização de crenças. Na prática, uma inferência probabilística resulta em um conjunto de crenças para os valores de cada nodo. A estrutura da rede Bayesiana não é alterada. Se as evidências que foram fornecidas são exemplos verdadeiros, elas podem dar alguma indicação de casos que poderiam ser vistos no futuro. Além disso, essas evidências mudam a base de conhecimento assim na próxima vez que a rede for usada; as probabilidades condicionais existentes entre os nodos da rede refletirão o mundo real de forma mais acurada.

Vantagens de uma rede Bayesiana:

- a) permite a representação e manipulação de incertezas com base em princípios matemáticos;
- b) modelam o “expertise” de um especialista de um modo intuitivo;
- c) é um das técnicas que pode efetuar qualquer tipo de inferência probabilística, isto é, causalidade, diagnóstico ou uma combinação desses.

3.1 Comparação entre Redes Bayesianas e Regras de Produção

De acordo com [Silvestre 2004], na comparação entre Sistemas baseados em Regras de Produção (RP) e Redes Bayesianas (RB), percebe-se que a última apresenta as seguintes vantagens sobre a primeira:

- a) *Apresenta um formalismo matemático:* Este formalismo é baseado na teoria clássica da probabilidade e consegue resolver o problema de evidências locais influenciando outras evidências não previstas pelo especialista criador do domínio (o que pode ocorrer com alguns sistemas baseados em regras).
- b) *Estrutura o conhecimento de forma acíclica:* Esta estrutura evita a confusão que pode ser gerada por uma regra de produção quando filhos influenciam pais, que por sua vez realimentam artificialmente as crenças sobre seus filhos.
- c) *Raciocínio causal é mais fácil de estimar do que o raciocínio Diagnóstico:* As probabilidades costumam ser mais fáceis de estimar utilizando o raciocínio causal ao invés do raciocínio diagnóstico. Por exemplo, médicos são muito bons em criar estimativas para “se um paciente tem câncer no pulmão, quais são as chances de um raio X ser anormal?”, melhor do que se utilizar o raciocínio Diagnóstico, “se o raio X do paciente é anormal, quais são as chances de um câncer no pulmão ser a causa?”
- d) *Resultados sempre serão obtidos:* Uma RB apresenta resultados mesmo com informações incompletas, pois quando não há evidências para a rede, as conclusões serão as probabilidades a priori dos nodos. Nas regras de produção quando não há sentenças criadas para todas as combinações possíveis dos valores das variáveis, pode-se não encontrar conclusões.

4. O Agente Pedagógico Bayesiano

As deficiências ou vulnerabilidades encontradas nas regras de produção originaram a idealização de um mecanismo de inferência alternativo para o agente Dimi. A pesquisa efetuada não teve como objetivo questionar a base de conhecimento implícita nas regras de produção do módulo de tomada de decisão do agente Dimi, mas sim criar um mecanismo que utilizasse outro modo de inferência que não possuísse os problemas destacados na seção 2. A experiência de um dos autores em outro projeto [Schreiber 2003] influenciou na opção por redes bayesianas.

Na próxima sub-seção são apresentadas as etapas na migração das regras de produção para as redes bayesianas.

4.1 Mapeamento: Regras de Produção para Rede Bayesiana

A definição de uma RB a partir de um conjunto de regras de produção implicou na substituição das regras de produção presentes no módulo de tomada de decisão do agente Dimi por uma RB. A figura 1 demonstra esse processo genericamente.

Uma metodologia de mapeamento RP para RB composta por três etapas foi criada com a finalidade de possibilitar a substituição das RPs por uma Rede Bayesiana:

- a) *Escolher as proposições e conclusões relevantes de uma RP e transformá-las em nodos*: esta é a principal etapa e demandou um grande esforço da equipe do projeto. A determinação da relevância de uma regra existente implicou em refazer os passos e compreender as decisões tomadas pelo engenheiro de conhecimento do projeto original;

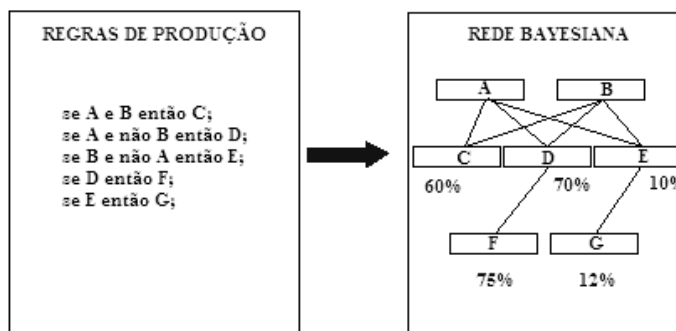


Figura 1 – Mapeamento genérico de Regras de Produção por uma Rede Bayesiana.

- b) *Estimar as probabilidades a priori dos nodos*: nesta etapa foi efetuada a análise dos arquivos de log existentes a fim de verificar a frequência com que cada regra de produção era disparada no sistema anterior;
- c) *Estabelecer relações de probabilidade condicional entre os nodos*: após a definição dos nodos utilizou-se o *expertise* da equipe do projeto para determinar os valores da tabela de distribuição conjunta de probabilidade.

4.2 O Módulo de Tomada de Decisão do Agente Pedagógico

Esta seção apresenta a rede bayesiana implementada no módulo de tomada de decisão do agente bayesiano. A rede bayesiana foi desenvolvida com a ferramenta *Shell Nética* [Norsys 2004]. Esta ferramenta permite criar nodos e estabelecer relações de causalidade entre esses nodos de uma forma intuitiva. Depois de desenvolvida a RB no *Shell Nética*, a ferramenta permite salvar a rede em arquivo. O *Nética* disponibiliza uma API para a linguagem *Java* e através dessa API é possível carregar a RB e realizar inferências. Via de regra, os nodos utilizados na RB são as variáveis (proposições e conclusões) encontradas nas regras de produção originais do agente Dimi. A quantidade de estados que constituirão um nodo é determinado pela quantidade de valores que a variável da RP pode assumir. Quando uma variável de uma RP pode assumir somente dois valores (binária), cria-se um nodo com apenas dois estados, por exemplo, verdadeiro e falso. Caso contrário, o nodo terá o mesmo número de estados que os valores possíveis que a variável na RP pode assumir.

A base de regras originalmente existente no módulo de tomada de decisão do agente Dimi deu origem a rede bayesiana mostrada na figura 2. Neste artigo a RB é didaticamente dividida em três partes com a finalidade de facilitar a sua compreensão.

A primeira parte corresponde ao perfil do aluno. Essa parte da rede é responsável por inferir os possíveis perfis do aluno. Como pode ser visto na figura 2, os nodos “Percepções”, “Ações Mobilizadas por” e “Compreensão de Leitura” estão conectados

com o nodo “Perfil”. Isso indica que o nodo “Perfil” possui três relacionamentos de causalidade, ou seja, as probabilidades dos quatro estados do nodo “Perfil” podem ser alteradas no momento em que se tem evidências para os outros nodos. A rede foi projetada dessa forma para que o perfil do aluno fosse determinado de acordo com suas características. Essas características dizem respeito à percepção, mobilização e compreensão de leitura do aluno. O perfil do aluno irá indicar quais elementos poderão ser apresentados a ele nas páginas das aulas.

A segunda parte da rede determina se as táticas de ensino atuais devem ser trocadas ou não. Nesta parte da rede bayesiana o nodo “Trocar Táticas” possui dois relacionamentos, indicando que depende condicionalmente dos nodos “Motivo Pouco ou Muito Tempo” e “Gostou dos Elementos”. Portanto, os motivos do aluno permanecer pouco ou muito tempo na página ou o fato dele estar gostando ou não dos elementos que estão sendo expostos na página, irão indicar se as táticas de ensino devem ou não ser trocadas.

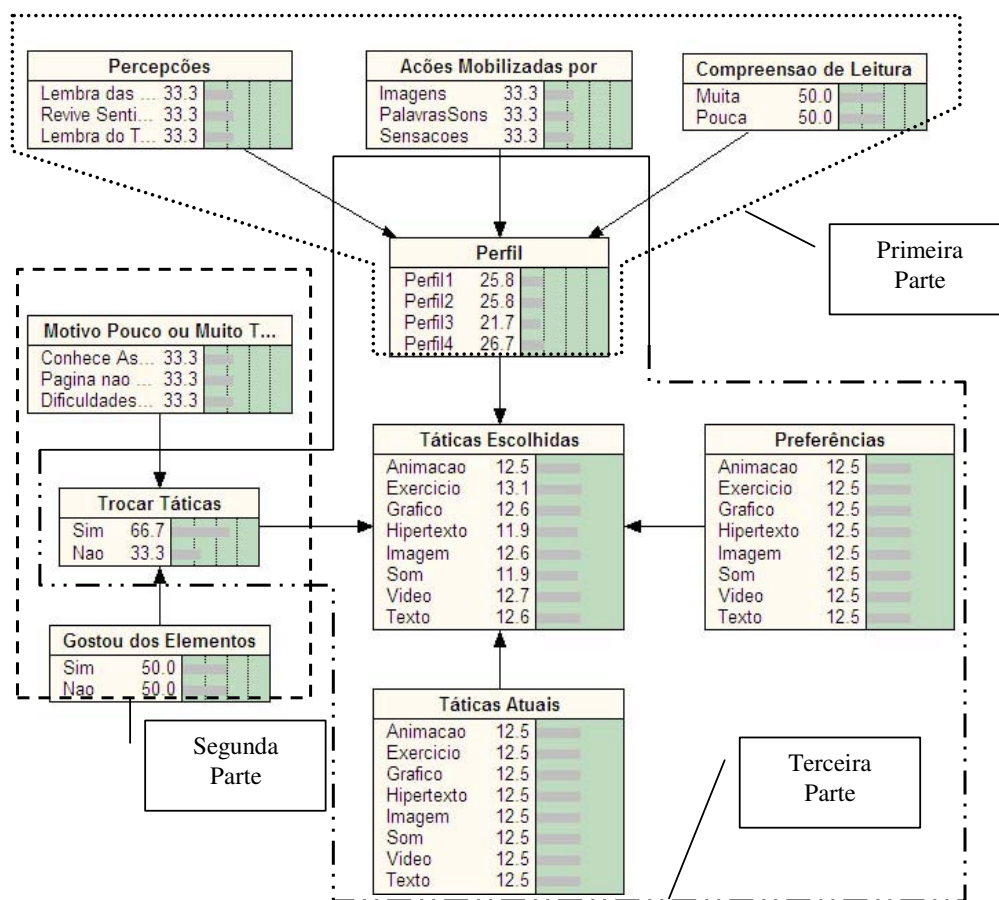


Figura 2 – Rede Bayesiana existente no módulo de tomada de decisão

Na terceira parte da RB, os estados do nodo “Táticas Escolhidas” representam as táticas de ensino que podem ser escolhidas para o aluno. Esse nodo possui quatro relacionamentos de causalidade, ou seja, depende condicionalmente dos nodos “Perfil”, “Trocar Táticas”, “Preferências” e “Táticas Atuais”. Dessa forma, o perfil do aluno, suas

preferências quanto aos elementos expositores, suas táticas de ensino atuais e a necessidade de trocar ou não essas táticas, irão determinar as novas táticas de ensino para o aluno.

4.3 Inferência do Agente Bayesiano Dimi

Durante a interação do aluno com o STI, a agente Dóris interage com o aluno questionando-o sobre diversos aspectos. A inferência ocorre ao se transferir as respostas fornecidas pelo aluno ao agente Dóris para o agente Dimi. A RB implementada no módulo de tomada de decisão do agente Dimi interpreta as respostas fornecidas à Doris como evidências e a partir dessas calcula a probabilidade a posteriori dos diversos estados da rede.

Por exemplo, na primeira vez que o aluno utiliza o ambiente do STI, é feito um questionário para determinar suas características. As respostas desse questionário servirão de evidências para os nodos “Percepções”, “Ações Mobilizadas por” e “Compreensão de Leitura” (primeira parte da RB). As evidências fornecidas para esses nodos determinarão as probabilidades para os prováveis perfis do aluno.

As evidências para a segunda parte da rede bayesiana são fornecidas no momento em que o aluno responde a Dóris se está gostando ou não dos elementos utilizados na página. A figura 3 exibe a Dóris questionando o aluno se ele gostou dos elementos utilizados na página.

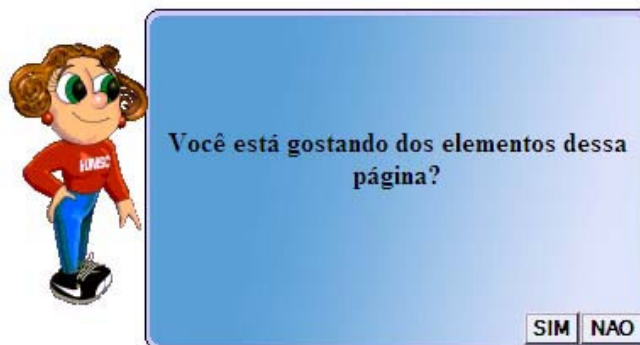


Figura 3 – Dóris Questionando se o Aluno Gostou dos Elementos da Página.

Em poucas palavras: o agente Dóris intervém durante a interação do aluno com o sistema, resultando em evidências que são fornecidas aos demais nodos da RB possibilitando o agente pedagógico inferir continuamente. Todas essas evidências coletadas pela Dóris e utilizadas pelo Dimi na RB determinarão as novas táticas de ensino a serem empregadas ao aluno.

5. Conclusão

Atualmente, a pesquisa sobre tutores inteligentes é realizada com a construção de ambientes que habilitam um aprendizado mais eficiente. Neste contexto, o uso de agentes inteligentes proporciona o desenvolvimento de diferentes instâncias e a integração de várias ações para alcançar o aprendizado efetivo.

A presença de agentes pedagógicos pode prover ambientes que são mais eficientes no que tange o aprendizado. Este artigo apresentou uma completa modelagem de um agente pedagógico bayesiano utilizado em um ambiente de tutoria inteligente destacando a sua estrutura de nodos. Um dos agentes pedagógicos desenvolvidos é responsável por coletar informações relevantes sobre o educando, fornecendo essas informações a rede Bayesiana como uma evidência, e este aspecto configura o desafio central desenvolvido nesta pesquisa.

Um protótipo com o domínio focado em Sistemas Especialistas foi desenvolvido para avaliar o comportamento do agente DIMI. Durante a interação dos alunos com o tutor, a agente DÓRIS interagiu fornecendo evidências ao agente bayesiano DIMI. Os resultados demonstraram a superação das deficiências anteriormente encontradas no agente DIMI, como por exemplo, que decisão tomar diante de uma situação não pré-programada via regras de produção? Uma rede bayesiana possui como característica inerente a tomada de decisão, independente de possuir valor para todas as variáveis envolvidas. O próximo passo previsto para esse projeto é avaliar a qualidade do aprendizado dos estudantes ao utilizar um mecanismo bayesiano como selecionador das táticas de ensino.

Referencias

- CORREDOR, M. V. (1993) “La Inteligencia Artificial y la Education: lo Aprendido y las Futuras acciones.” *Informática Educativa*, Colômbia, v.6, n. 3, p. 235-242.
- GIRAFFA, L.M.M. (1999) “Uma arquitetura de tutor utilizando estados mentais.” Porto Alegre: CPGCC da UFRGS. (Tese de Doutorado).
- NORSYS (2004) “Introduction to Bayes Nets.” Disponível em: http://www.norsys.com/tutorials/netica/secA/tut_A1.htm.
- LADEIRA, Marcelo. Redes (2004) “Bayesianas Multiagentes”. Disponível em: <http://www.sbc.org.br/reic/edicoes/2002e1/tutoriais/RedesBayesianasMultiagentes.pdf>.
- SANTOS (2005) “Tutores Inteligentes”, Nilson M. Disponível em: <http://www.din.uem.br/ia/tutores/tutores.html>.
- SANTOS, Cássia; Frozza, Rejane; Paschoal, Luciano; Dahmer, Alessandra (2002): “Dóris - Pedagogical Agents for Intelligent Tutoring System”. 6th International Conference on Intelligent Tutoring System, Lecture Notes in Computer Science, Biarritz, France, June.
- SCHREIBER, Jacques (2003) ”Análise do tempo de navegação na composição de Um Modelo para Hiperídia Adaptativa”. Tese de Doutorado, CPGEPS, UFSC.
- SILVA, LUÍS F. R. (2002) “Dimi. Um Agente Selecionador de Estratégias de Ensino em Sistemas Tutores Inteligentes.” Trabalho de conclusão Ciência da Computação. Santa Cruz do Sul: Unisc.

SILVESTRE, André (2004) “Raciocínio Probabilístico Aplicado ao Diagnóstico de Insuficiência Cardíaca Congestiva (ICC).” Disponível em: <http://www.inf.ufrgs.br/pos/ppgc/semanacademica/artigos/Art.AndreSilv.pdf>.

VICCARI, R. M.;GIRAFFA,L.M.M (1996). “*Sistemas Tutores Inteligentes: abordagem tradicional x abordagem de agentes*” Brazilian Symposium on Artificial Intelligence,13.,1996.Proceedings... Curitiba: Springer Verlag.