贝叶斯网及其概率推理在智能教学中的应用

冀俊忠，刘椿年，江 川，杨文盛
（北京工业大学 计算机学院，北京 100022）

摘 要：对贝叶斯网及其概率推理进行了简述，提出并实现了一种以贝叶斯网为学生模型的智能教学系统，并重点介绍了学生模型的结构、功能和概率推理算法．系统运行的实例数据真实地反映了学生模型当前的解题情况．

关键词：贝叶斯网；概率推理；智能教学系统；学生模型
中图分类号：TP 39　　**文献标识码**：A　　**文章编号**：0254-0037(2002)02-0353-05

智能教学系统(ITS)[1]是传统的计算机辅助教学、人工智能和认知心理学的交叉领域；在国际上已经形成较大的研究规模，取得了令人信服的实用成果．尤其是近10多年来，国外ITS发展迅速，并从研发阶段向商品化阶段转化．ITS能够建立并维护学生模型．借助于学生模型，一方面，系统模拟教师的教学决策，对学生进行实时的个性化的辅导；另一方面，系统对一组学生的学生模型的历史数据进行自动分析处理，能帮助教师客观地了解自己的教学对象的整体情况，便于教师适时调整自己的教学计划和教学内容．学生模型能够描述学生的知识水平、学习能力和认知特点；本质上是基于规则的计算机程序，以学生做题的方式来求解实际问题．由于在问题的求解过程中判定学生对规则的掌握存在着很大的不确定性，所以本文的学生模型使用贝叶斯网[2]．近年来，贝叶斯网已成为图形化表示一组变量关系和处理专家系统中不确定性的强大工具，不仅在ITS中最关键的部件(学生评价、个性化的在线帮助、自我解释等)中有建模的成功事例，而且在医疗诊断、文本分类、重要系统的属性预测、软件工程中的因果建模等中也有成功的应用．本文着重介绍贝叶斯网在ITS中建立学生模型方面的应用，下面先对贝叶斯网作一简要介绍．

1 贝叶斯网及其概率推理

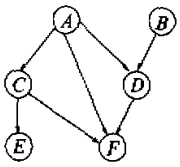
1.1 贝叶斯网概述

贝叶斯网BBN(Bayesian belief net)，是根据随机变量间的概率关系建立的图论模型[3]，它是用有向的无循环图来表示的(见图1)．联合空间中的每一变量(如A、B、C、D、E、F)在BBN中是用节点来表示的，节点的值可以是两值或多值．对于每一节点有两种信息需要指定：首先是网络中的有向弧这样一种变量之间的依赖关系(见图1(a))：如果$X \to Y$，则节点X是节点Y的父母，Y是X的后代，"→"表示了X、Y之间的因果关系；任一变量在给定它的父母节点值时条件独立于它的非后代节点集．其次是每一变量都有条件概率表(见图1(b))，它描述了这个变量在给定它的父母节点值(小写字母表示该变量取真，¬表示变量的非)时的概率分布．在条件独立的前提下，利用条件概率表就能求出BBN的联合概率．

可以说，BBN中有向弧和数字化的概率分布描述了在节点间的概率关系．更精确地说，这种图形结构指定了一组条件独立的声明和用于刻画概率依赖强度的条件概率的数值．BBN用此结构紧凑地表示了网中各变量的联合概率分布．设网上的变量为$Y_1, \cdots, Y_n$，它们分别取值$y_1, \cdots, y_n$时的联合概率为

$$P(y_1, \cdots, y_n) = \prod_{i=1}^{n} P(y_i \mid \text{parents}(Y_i)) \tag{1}$$

收稿日期：2000-10-16．
基金项目：国家自然科学基金资助项目(69883001)．
作者简介：冀俊忠(1969-)，男，工程师，在职博士生；刘椿年(1944-)，男，教授，博士生导师．



(a)各节点的依赖关系

	a,b	$a,\neg b$	$\neg a,b$	$\neg a,\neg b$
d	0.4	0.1	0.8	0.2
$\neg d$	0.6	0.9	0.2	0.8

(b)节点D的条件概率表

图1 一个典型的贝叶斯网

其中：parents(Y_i)是节点 Y_i 的父母节点组，也是网上变量，它们的取值已知，故 $P(y_i|\ \text{parents}(Y_i))$ 可在节点 Y_i 的条件概率表中查到．如果 parents(Y_i)为空集，则称节点 Y_i 是根节点或边界节点．

1.2 概率推理算法

贝叶斯推理[3]提供了一种以概率分布为基础的推理方法．它是基于如下的假设：人们感兴趣的变量值受概率分布的控制，结合观察数据，对这些概率进行推算便可做出最优的决策．贝叶斯学习中的概率推理在机器学习中占有相当重要的地位．

在 BBN 中，某些变量的值经常能被观察到．本文将这些变量称为证据变量，与其相对应的节点称为证据节点．证据节点的取值由于已经确定，故它们取观察值的概率就为 1．概率推理的基本任务就是在给定某些证据变量值的前提下，计算一组查询变量的后验概率分布，即 P(查询变量|证据变量的取值)；正如前面所述 BBN 的网络结构体现了网中各个节点间的依赖关系，因此这些证据势必会通过 BBN 结构传递出去，直接或间接地影响网中的其他节点．证据的传播在 BBN 中一般是用概率推理算法实现的，而且已成为 BBN 中的一个关键问题．

概率推理算法就是借助于概率理论利用 BBN 来推导出目标变量值的方法，当然，如果处理的是随机变量，推导一般不会准确地把目标变量指定为单个确定的值，实际推出的是目标变量的概率分布，即在给定某些变量值的情况下，依一定的推理方法获得目标变量取各种值的概率．通常在给定某些变量的准确值后，推导步骤很直观．一般地，BBN 可推导出网上一组变量的概率分布，即把网上变量的任一子集作为目标变量集；在给定其余变量的任一子集的值或概率分布后，能够推导出目标变量集的概率分布．一般来说，对于任一 BBN 的准确的概率推导是一个 NP-hard 问题[3]．尽管在理论上近似方法也可能是 NP-hard 的，但近似方法已在许多应用中显示了很好的实用性，Pearls 直接采样法就是其中最有效的算法之一．

2 BPU-ITS-Math 系统的学生模型及概率推理算法

2.1 系统概况

本系统是中学数学的智能教学系统——BPU-ITS-Math[4]．系统模拟教师对学生进行个别辅导的经验、方法，为学生提供一个灵活的、个性化的、交互式的、智能化的练习和作业完成环境．

BPU-ITS-Math 的系统功能结构如图 2 所示．学科知识库代表学科规则集；对于题库中的每一问题，推理机(BPU-CLP)利用学科知识库能够自动产生问题的解图．解图和初始模型(学生的作题历史信息)相结合形成动态学生解题模型．行动解释器完成节点信息形式的转换(如表达式的中缀到前缀的转换等)，便于在解图中查询相应节点，并给出输入正确与否的及时反馈．在学生提出帮助时，系统参照学生模型的当前状态，由教学决策模块作决策，给学生提供适时、恰当的帮助．图 2 中虚框部分为本文阐述的重点．

2.2 学生模型的功能

当学生在用户界面上选择题目后，初始学生模型和解图相互匹配结合在一起形成 BBN，即新的学生解题模型；然后系统接受学生的行动，利用概率推理算法计算 BBN 节点的概率，从而更新学生模型的状态．这一部分是整个 ITS 的核心．它记录了学生对学科规则掌握状态的概率估计，学生已经做了哪些题

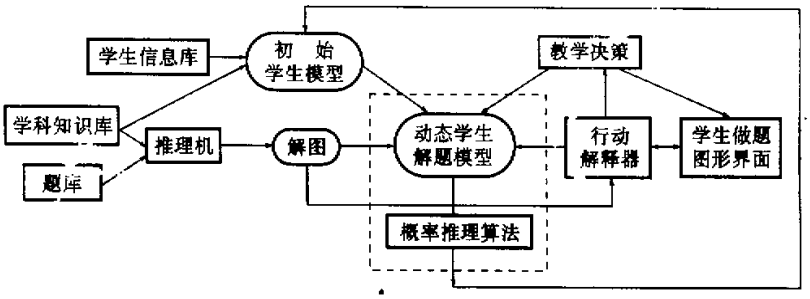


图2　BPU-ITS-Math的系统功能结构图

目，学生在过去接受过哪些帮助等，因此成为个性化的在线帮助、学生评价、教师决策的基础．

2.3　学生模型的贝叶斯网结构

学生模型由静态、动态两部分组成．静态部分是指存放在所有学生模型数据库中的部分；这部分的边际概率随着问题的求解而不断更新，并最终成为对学生了解掌握本学科知识做长期评价的依据．一般来说，初始的学生模型都是静态模型的一个子集．动态部分在学生选择问题后自动产生，当问题解决后学生选择另一道题或退出系统本次动态部分被丢弃．动态部分是由自动推理模块生成的解图转换来的，它包含规则、事实、目标和规则应用4种节点．因为解图结构已经对问题求解的因果结构进行了编码，所以在BBN中仍被维持．贝叶斯网上节点的概率表示学生对节点信息的掌握程度，节点的概率值集合是学生模型当前状态的体现．

2.4　Pearls直接采样法[5]

BPU-ITS-Math系统对BBN概率的推理采用Pearls直接采样法．假设用如下的符号表示BBN的信息：W为节点集，w为W的实例化；$E \subseteq W$为证据节点集$\{E_1, \cdots, E_n\}$，其观察值向量为$e = \{E_1 = 1/0, \cdots, E_n = 1/0\}$；$X$为任一非证据节点，$x$为$X$的实例化；$W_x$表示除$X$的节点集，$U_x$表示$X$的父母节点集，$u_x$为$U_x$的实例化；$Y_x$为$X$的子女节点集$\{Y_1, \cdots, Y_m\}$，$Y_j$为$X$的第$j$个子女，$y_j$为$Y_j$的实例化；$F_j$为$Y_j$的父母节点集，$f_j(x)$为$F_j$的实例化，求$X$对证据的后验概率分布$P(x|e)$．该算法如下：

1)将节点集W初始实例化为w，证据节点实例化为观察值(证据的概率为1，永远不变)，非证据节点实例化为任意值或以该节点的概率随机采样得到的值．

2)对每一非证据节点X，计算在其他节点给定值的情况下X的后验概率分布：

$$P(x|W_x) = \alpha P(x|U_x)\prod_j P(y_i|f_j(x))$$

该式表明了本节点的概率仅与其父母节点、子女节点及其子女节点的父母节点有关，α为节点X取1或0的加权因子，$P(x|U_x)$、$P(y_i|f_i(x))$为条件概率，这些非根节点的条件概率表可在线求出．

3) 按此概率分布对X随机采样，结果作为X的新实例化，反复进行第2)步，直到近似过程达到稳定状态(设进行了m次后，满足了精度要求)，这时

$$P(x|e) = \frac{1}{m}\sum_{i=1}^{m} P_i(x|W_x)\ .$$

3　应用实例

根据上述的设计思想，实现了BPU-ITS-Math系统，并对函数求值域问题进行了测试．下面举例说明：已知函数$y = 1/\lg(3x-2)$的定义域为$x > 1$，求值域．

在这道例题中涉及知识表达中关于函数求值域问题的几条具体规则，如一次规则、对数规则、反比例规则、复合对数规则、复合反比规则等．这些规则随着学生选题，由系统自动从知识库中调出，与题目的 prolog 描述一起作为 BPU-CLP 推理机的输入，经推理机推理后形成了学生当前的贝叶斯网解题模型，见图 3．在图中的变量都是命题变量，故其值只取 true 或 false；true 表示学生对该命题已掌握，false 表示没掌握．学生在首次使用某些规则时，系统自动给根节点赋概率初值 0.5(若原来已使用过该规则，系统会从学生的静态模型中读取概率值，其值越接近 1，表明对该规则掌握程度越高)．此处 0.5 表示系统对学生是否掌握该规则很模糊，故取其平均值，其他节点的概率值有待于学生输入解题行为后，使用 Pearls 算法求出．

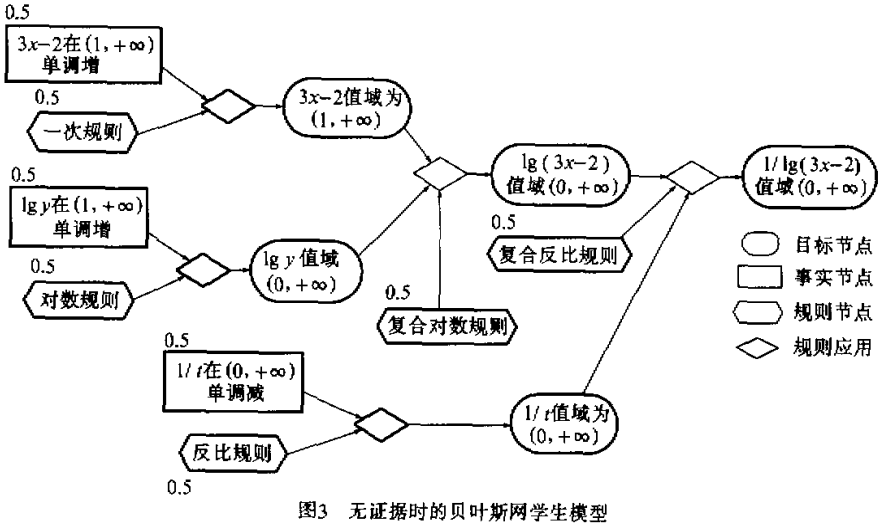


图3 无证据时的贝叶斯网学生模型

如果学生输入解步“$3x-2$ 值域为 $(1,+\infty)$”和“lg y 值域为 $(0,+\infty)$”，这两个目标节点(图 4 中的粗框)被系统匹配上，成为证据节点；其为 true 的概率值为 1，经 Pearls 算法完成概率推理．由于受 BBN 模型因果关系的制约，其他节点的概率也随之更新，即可得到在证据节点已知的情况下这些节点取 true 时的后验概率，由此反映出这些证据节点对学生解题模型的影响．此时学生解题模型的状态如图 4 所示．

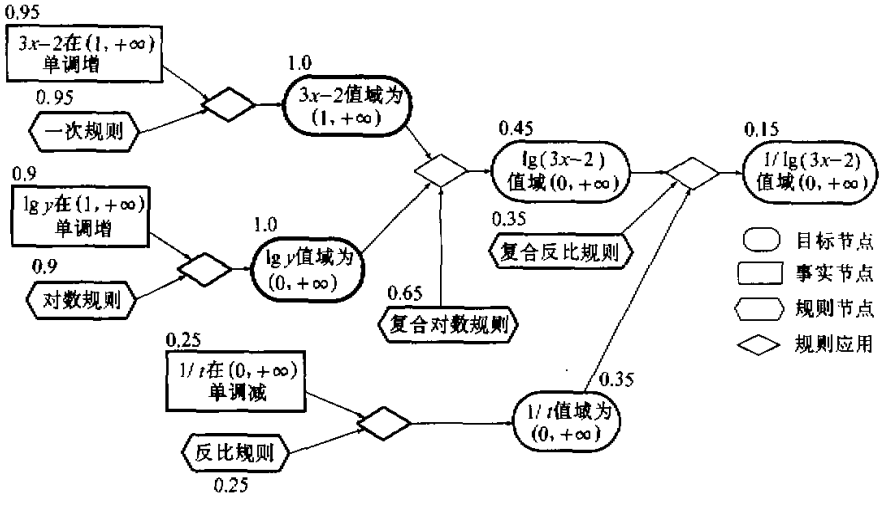


图4 有证据后当前学生模型的新状态

由图 4 中的概率值可见，在学生写完上面两步后，学生接下来最可能写的解步是“$\lg(3x-2)$ 值域 $(0,+\infty)$”(0.45)；此时学生最可能不清楚的规则是“反比规则”(0.25)，这些都符合有经验的老师对学生解题

思路的判断，可见，基于BBN的学生解题模型不仅能记录学生的当前行为，而且能对学生的未来行动作出正确的预测和判断．当学生做完此题或中途退出系统时，该题所涉及的规则节点概率值将回写到学生模型的数据库，用以更新该生的静态模型．

4 结束语

近10年来，随着专家系统和不确定性掀起的人工智能领域中的科研巨浪，BBN已成为一种解决专家系统中不确定性的有效手段．作者在介绍BBN及其概率推理的基础上，对BBN在智能教学中的应用进行了研究，通过系统的运行实例证明了用BBN建立学生解题模型的正确性、有效性和实用性．这种设计思想为其他系统的建模提供了值得借鉴的方法．进一步的工作包括试用其他的概率推理算法[6-8]，并与本文采用的方法进行比较．

参考文献：

[1] MURRAY T. Authoring intelligent tutoring systems: an analysis of the state of the art[J]. Int J Artif Intell Educa, 1999, 10: 98-129.

[2] MARTIN J, VANLEHN K. Student assessment using Bayesian nets[J]. Int J of H-C S, 1995, 42: 575-591.

[3] MITCHELL T. Machine Learning[M]. New York: The McGraw-Hill Companies, 1997.

[4] LIU Chun-nian, ZHENG Lei, JI Jun-zhong, et al. Electronic homework on the WWW[A]. Zhong N. WI 2001[C]. Berlin: Springer-Verlag, 2001. 540-548.

[5] PEARL J. Evidential reasoning using stochastic simulation of causal models[J]. Artif Intell, 1987, 32: 245-257.

[6] HRYCEJ T. Gibbs sampling in Bayesian networks[J]. Artif Intell, 1990, 46: 351-363.

[7] COUSINS S, CHEN W, FRISSE M. A tutorial introduction to stochastic simulation algorithms for belief networks[J]. Artif Intell Med, 1993(5): 315-340.

[8] LAURITZEN S L, SPIEGELHALTER D J. Local computations with probabilities on graphical structures and their application to expert systems[J]. J Royal Stat Soc, 1998, 50: 157-224.

Application of Bayesian Network and its Probability Reasoning to Intelligent Tutoring System

JI Jun-zhong, LIU Chun-nian, JIANG Chuan, YANG Wen-sheng

(College of Computer Science, Beijing Polytechnic University, Beijing 100022, China)

Abstract: The authors firstly outline BBN and probability reasoning, then propose and implement our intelligent tutoring system that uses BBN as student models, and emphasize the structure and functions and probability reasoning algorithm of student models. The data of practical examples in the operation of the system really reflect how the student model solves problems currently.

Key words: Bayesian belief networks(BBN); probability reasoning; intelligent tutoring system; student model