

基于关系数据库的智能教学系统知识库设计

Design Knowledge Base in Intelligent Tutoring System Based on Relation Database

罗梅 王万森 (首都师范大学信息工程学院 100037)

摘要:本文通过对知识库和关系数据库的比较,说明了基于关系数据库设计知识库的可行性,并以智能教学系统中知识库的构建过程为例,介绍了知识表示方法,具体给出了基于关系数据库建立知识库的方法,并简要说明了知识库管理和知识获取方法。

关键词:智能教学系统 知识库 关系数据库 知识表示

1 引言

智能教学系统(Intelligent Tutoring System, ITS)是涉及人工智能、计算机科学、认知科学、教育学、心理学和行为科学的综合性课题,其研究目的是由计算机系统承担起人类教育的主要责任,即赋予计算机系统以智能,由计算机系统在一定程度上代替人类教师实现最佳教学。研究 ITS 的意义在于减轻人类教师的教学工作量、提高教学质量、增进人类对其自身认知过程的了解和带动相关学科的发展。这就要求计算机系统具备领域知识、教学知识、了解学生的能力以及跟学生交互的能力。

传统上,智能教学系统可按功能划分成四大模块(如图 1):学生模型(学生的知识水平和学习能力)、教师模型(教学策略)、知识库(领域知识和经验知识描述)、人机接口(在智能教学系统和教师及学生之间建立友好的通讯交流方式)^[1]。

其中,知识库作为 ITS 的重要组成部分,主要解决教什么的问题,表示教学领域的相关知识,以及教学专家能够基于这些知识解决有关问题的问题求解知识。知识库是智能教学系统质量是否优越的关键所在,即知识库中知识的质量和数量决定着智能教学系统的质量水平。为了便于知识存储和管理,本文将关系数据库技术引入智能教学系统研究领域。

2 知识库与关系数据库的比较

知识库中不仅包含显式表达的事实知识,而且包

含显式表达的包括常识、经验等在内的领域知识和启发式知识等。关系数据库采用二维表表示一些基本的事实数据,具有处理海量数据的能力,但无表达和处理规则等知识的能力,如在某确定的规则条件下,不易推导出新的数据项,不支持专门的推理机制和带控制的搜索,其推理能力较差。

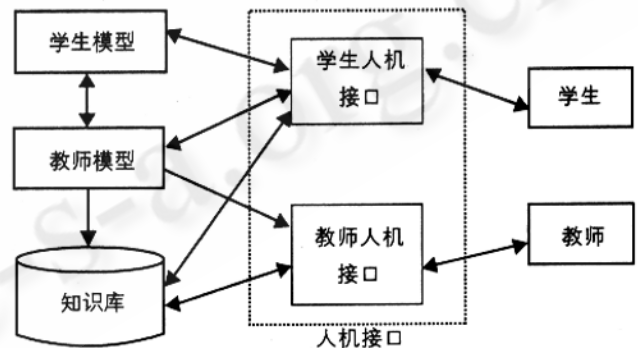


图 1 ITS 系统结构

知识库的基本内容可描述为 $KB = F + R$, 其中 KB 代表知识库; F 代表事实集, 相当于数据库中实体及其属性所对应的数据集合; R 代表规则集, 其中包含领域知识或启发式知识。数据库中的规则通常是隐含于数据结构或完整性约束中, 而知识库中的规则主要通过某种知识表示方式, 与事实一样显式地表达出来。相对于数据库, 知识库的知识较少改动, 而且知识库应具有拒绝插入不一致的知识的能力。将数据库技术与知识库技术结合起来, 构造大规模的知识库系统, 可存储与管理海量条规则与事实, 而且从知识库中存取知识

如同从数据库中存取数据一样迅速方便。由此,数据库中所包含的明显表达的关系约束条件,可看作某种低层次的知识;知识库中所包含的有明确定义的事实和通过规则表达的关系也是数据库的内容。数据库和知识库内容的比较如表 1^[2]。

表 1 数据库和知识库内容比较

数据库	知识库
存储大量数据,少量完整性条件	存储大量事实、规则及大量完整性约束
结构简单、数据类型少	结构复杂、数据类型多
规则性	不规则性
一般人员可更新	知识工程师更新
即时查询	频繁的规则存取
模式修改少	经常修改模式
精确、频繁修改	不精确、不经常修改
表示事实	表示事实和规则
管理事实	管理事实、完整性约束、规则
推理能力差	推理能力强
不支持专门的推理规则	支持专门的推理规则
不支持受控制的搜索	支持受控制的搜索
历史数据、静态数据	动态数据,既有历史又有现在

通过以上比较,可以发现为满足智能教学系统领域的应用要求,必须寻找知识库与关系数据库结合的方法,把事实数据与抽象知识结合起来,把数据处理能力和基于知识的处理能力结合起来,形成一个广义的知识库系统。知识库系统应提供对数据和知识有效一致的管理。

3 基于关系数据库技术的智能教学系统知识库设计

在设计智能教学系统知识库之前,必须先将该系统中的需要存储和管理的知识以一定的方法表示出来。下面介绍知识表示概念及智能教学系统中的知识表示方法。

3.1 知识表示

3.1.1 知识表示的概念

对于知识表示,可以从其所起的作用这个角度来定义。知识表示有如下五个作用^[3]:

(1) 知识表示是一个推理代理。

(2) 知识表示是一系列的本体任务。

(3) 知识表示是智能推理的子理论。

(4) 知识表示是一种高效计算方法。

(5) 知识表示是一种人类表达方法。

3.1.2 三种知识表示观^[4]

认识论表示观:表示是对自然世界的描述,表示自身不显示任何智能行为,其唯一的作用就是携带知识,表示研究与“启发式”研究无关。

本体论表示观:表示是对自然世界的一种近似,它规定了看待自然世界的方式,即一个约定的集合。表示只是描述了在这个世界中,观察者当前所关心的那部分,其它部分则被忽略。

知识工程表示观:表示是对自然世界描述的计算机模型,它应该满足计算机这一实体的具体限制,因此,表示可以理解为一类数据结构及在其上的一组操作。

知识工程表示观在两个重要方面区别于前两种表示观:一是知识工程表示观一般将表示理解为一类数据结构及在其上的操作;二是对知识的内容更强调与领域相关的、那些只适合于这个领域的、来自领域专家经验的知识。这两个特点说明这类表示观更强调其工程实现性,而不太关心对其行为的科学解释。

本文介绍的智能教学系统中的知识表示方法即是基于知识工程表示观的。

3.2 智能教学系统中的知识分类

按照知识的性质,可分为两类:^{[5][6]}

(1) 领域知识(事实性知识):指系统所要教授的课程内容及所附习题等。领域知识模型应当能够反映出学习的内容和学习的过程。

(2) 经验知识(启发性知识):是教师在教学过程中积累的教学经验即教学策略,这类知识是基于教师对课程结构、教学理念、方法的认识而形成的,因此,不可避免地带有教师自身知识的痕迹。只有以高水平的教师(领域专家)为前提才能构成有效的智能教学系统。

3.3 智能教学系统知识库设计

由于智能教学系统中的知识可分为两类,所以在设计知识库时采用不同的表示方式,最终都设计成关系数据库模型。

3.3.1 领域知识库设计

对于领域知识,采用超文本表示方法,根据教学要

求把课程内容分解成小单元,抽象成知识点,每个知识点通过相互链接构成一个网状结构。因此经分析可知领域知识库中应包含知识点属性和知识点链接关系。知识点属性是知识表示的基础,知识点链接是进行推理的基础^{[5][7]}。

首先定义知识点间的三种关系:

- 先行关系:① → ② 要学习知识点②必须先学习知识点①
- 示例关系:① → ② 知识点②是对知识点①的示例
- 顺序关系:① → ② 学习了知识点①之后可以学习知识点②

下面给出领域知识库中通用的所需的各表的设计,在设计实现某门课程的教学系统时可根据需要稍做修改:

① 教学内容表 content 定义如下:

字段名	类型	长度	意义
知识点编号	字符		具有唯一性,用于检索和查找 (长度根据实际情况确定)
知识点类型	字符	1	记忆型,问题求解型,认知策略型, 用数字 1~3 表示
知识点内容	字符		根据实际情况确定
难度	字符	1	易、一般、难、很难 4 级,用数字 1~4 表示
重要程度	字符	1	选学、一般、重要 3 级,用数字 1~3 表示
认知类型	字符	1	记忆、理解、应用和综合等 4 级, 用数字 1~4 表示
相关例题	字符		表示对此知识点提供的例题的编号, 0 表示无例题
相关习题	字符		表示对此知识点提供的习题的编号, 0 表示无习题

② 习题表 exercise 定义如表 2。

③ 示例表 example 定义同表 exercise

④ 知识点关系表 relation 定义如下:

字段名	类型	长度	意义
知识点 1 编号	字符		同 content
知识点 2 编号	字符		同 content
关系	字符	1	先行、示例、顺序三种 关系,用数字 1~3 表示

3.3.2 经验知识库(教学策略库)设计

教学策略库用来存放所有形成教学策略的知识,

形成教学策略的知识是用产生式规则表示的,其形式为:规则号:IF <条件> THEN <结论>。共有二类规则:评价规则和学习方法指导规则分别存放到教学策略库中的评价表和学习方法指导表中^[8]。

表 2

字段名	类型	长度	意义
习题编号	字符		具有唯一性,用于检索和查找 (长度根据实际情况确定)
习题内容	字符		描述一个题目
参考答案	布尔	1	True 表示有参考答案, False 表示无参考答案
知识点编号	字符		此题涉及的知识点的编号
解题标识	布尔	1	True 表示解题过程字段有内容, False 表示无
解题过程	字符		描述题目的解题过程
难度	字符	1	易、一般、难、很难 4 级, 用数字 1~4 表示
认知类型	字符	1	记忆、理解、应用和综合等 4 级, 用数字 1~4 表示
题型	字符	1	填空、单选、多选、判断、解答四种, 用数字 1~4 表示

① 评价规则及评价表定义:

评价规则是对学生学习过程中的表现进行评价的规则。其条件包括学生在记忆、理解、应用和综合等方面的表现。每项在学生在学习过程中都可得到一个分值,系统根据各项所得的分值,根据评价规则进行推理,对学生的学习进行评价。如可设定一条规则为:

IF (70 < 记忆 < 80) and (60 < 理解 < 70) and (50 < 应用 < 60) and (40 < 综合 < 60) THEN (评价结论为:记忆力不错,理解和应用能力一般,分析综合能力较差。)

评价规则表 (assessment) 定义如下:

字段名	类型	长度
规则号	字符	具有唯一性,长度根据规则总数确定
记忆	整型	3
理解	整型	3
应用	整型	3
综合	整型	3
结论	字符	根据实际情况确定

② 学习方法指导规则及学习方法指导表定义:

学习方法指导规则是以学生学习过程中各知识点的成绩作为依据,先通过一个对各分值加不同权的算法($S = \sum_{i=1}^4 N_i P_i$,其中, N_i 是分值, P_i 是该分值的权, $\sum P_i = 1$)将其转换成百分制评价成绩,然后以此分数作为学习方法指导规则的条件,对学生的学习进行指导。学习方法包括:重新学习本知识点;如有先行知识点,则学习先行知识点后再学习本知识点;继续下一知识点的学习。可定义如下规则:

IF (评价成绩 < 50) and (先行知识点) THEN (学习先行知识点)

学习方法指导表(instruction)定义如下:

字段名	类型	长度
规则号	字符	具有唯一性,长度根据规则总数确定
评价成绩	整型	3
先行知识点	布尔	1
结论	字符	根据实际情况确定

3.4 知识库管理

知识库的管理包括知识的获取、查询、修改和维护等功能,这些功能在关系数据库环境中更容易实现。

(1) 知识获取^[9]:其任务是将人类已有的知识从领域专家的大脑中或书本上抽取出来,加以归纳总结,编码成计算机可以识别的形式,为计算机完成领域专家所能的工作提供知识基础。有两类知识获取方法:

- 直接方式:领域专家向智能教学系统提供数据和资料,应用统计归纳、因果推理和人工神经网络等技术从中提取出所需知识。

- 间接方式:目前应用较多也较为成熟的方式。首先,领域专家将自己的知识用语言及书面的形式整理出来;然后,知识工程师在专家的帮助下对所提供的知识进行分析、抽象及简化,编码成计算机能识别的形式,通过知识编辑器之类的工具,将知识输入到智能教学系统的知识库中。

直接方式类似于数据库领域的数据挖掘过程,因此可以在其基础上实现;间接方式在数据库的逻辑设

计中主要表现为如何向知识库中添加记录,很容易在关系数据库中实现。

(2) 知识查询:提供了对规则及规则元素的查询。对规则可按规则号及规则所包含的规则元素进行查询。

(3) 知识修改:知识的修改在一致性检查发现错误时,提供一个修正错误规则的机会。同时,对修改后的知识进行一致性检测。

(4) 知识维护:知识库中的知识可能遭到破坏,使知识库中的知识不一致,为保证知识库的一致性及完整性,要定期对知识进行一致性与完整性维护。

4 结语

随着社会的进步,计算机辅助教育及远程教育越来越受重视,而这也促进了对智能教学系统的研究。可以预计,随着智能教学系统研究的不断深入和发展,对知识库的性能要求会越来越高。而将知识库建立在关系数据库基础之上,充分利用关系数据库系统的强大管理功能和便捷的开发环境即可最大程度地满足这个要求。

参考文献

- 1 莫赞等,智能教学系统的发展与前瞻,计算机工程与应用,2002.6。
- 2 黄万里等,基于数据库的大坝安全监控专家系统知识库设计,大坝与安全,2004.3。
- 3 王珏等,关于知识表示的讨论,计算机学报,1995.3。
- 4 Randall Davis等,What is a Knowledge Representation? AI Magazine, 14(1):17-33, 1993。
- 5 张燕姑等,智能教学系统知识表示的设计与实现,温州师范学院学报,2003.10。
- 6 陈品德等,智能化远程学习支持系统——领域建模方法的分析与比较,中国电化教育,2004.1。
- 7 史悦等,基于知识库的智能教学系统建模,贵州师范大学学报,2003.8。
- 8 韩燮等,基于超媒体智能CAI系统的研究,计算机工程与应用,2002.10。
- 9 吴海桥等,基于关系数据库的知识库的建立,微型电脑应用,2001年第17卷第11期。