

# 智能教学中的知识模型与教学评价的研究

## Research on Knowledge Model and Knowledge Diagnosis in ITS

夏丽华 (上海杉达学院计算机系 上海 201209)

齐景嘉 (哈尔滨金融高等专科学校计算机系 哈尔滨 150030)

**摘要:**目前,基于 Web 的智能教学系统是计算机应用于教育领域的研究方向之一。本文从知识点及其关系出发构建一个知识网络模型,并基于该模型介绍如何进行领域知识表示及考核试题设置,然后设计了试题难度系数的动态更新方法,还提出一种简捷的知识诊断方法,准确地将学生学习缺陷定位到具体的知识点,得出学习路径方面的建议,从而增强了教学的针对性。

**关键词:**知识网络 智能教学 难度系数 知识诊断

### 1 概述

教学评价是教学中至关重要的一个环节,它所起的作用可以概括为五个方面:反馈调节功能、诊断指导功能、强化激励功能、教学提高功能、目标导向功能。但是目前几乎所有的网络智能教学系统的教学评价部分都只是通过考试成绩本身正确评估学生对领域知识的掌握程度,缺乏对学生薄弱环节的诊断指导,不能对学生的测试结果信息深入挖掘,提出改进学习的方向和优化教学的建议<sup>[1]</sup>。针对这一现状,本文从知识点及其关系出发构建一个知识网络模型,研究怎样科学地分析学生的考试结果,准确发现学生自身的薄弱环节并指出改进的学习方向,从而帮助教师提高教学效率,有针对性地实施因材施教,同时能对试题库的关键属性进行动态更新。

### 2 知识网络

一门学科的整体知识体系是相互关联前后融合的,为了反映整个知识体系的全貌,教学内容编排应该按顺序分层次设置。常见的方法是将所有课时的内容分为章(或单元等)、节、知识点三个基本层次,章内包含节,节内包含知识点。整个教学内容的文件组织形式采用树型分级目录,按照逻辑顺序将各章节统一排列,各章内部按节进行逻辑排序,节内内容以知识点为中心按照知识展开顺序编制。

知识点是教学活动过程中传递教学信息的基本单元,它包括理论、原理、概念、定义、范例和结论等。知识点可分为复合知识点和元知识点。复合知识点是一组相关知识点的集合,组成复合知识点的知识点可以是教学元知识点也可以是教学复合知识点,元知识点是在结构上不可再分的知识点。根据课程特点,对课程知识点进行分解和分类,将复合知识点细分为几个较小的知识点,每个小知识点还可以划分为更小的知识点,直至在结构上为不可再分的元知识点。知识点必须由教学专家根据学科的特点和实际教学过程的特点,来进行系统、科学的划分,划分的基本原则是保证知识内容的局部完整性及全局统一性。

为了实现教学评价,还需要赋予每个知识点相关的属性,实现教学诊断及试题库属性的更新。通常知识点的属性项包括知识点的编号、种类、难易重要程度、要求掌握程度、知识点通过阈值等。其中知识点种类分为记忆型知识、理解型知识、陈述型知识和过程型知识;难度及重要程度分为 4 级,分别为容易题、一般题、较难题及难题;要求掌握程度分为记忆、理解、应用等;知识点通过阈值是指学生对知识点掌握程度的最低标准,只有当学生掌握程度超过该阈值时,才能离开该知识点的学习。

通过对知识点的分析,可知知识点之间的关系主要包括父子关系和依赖关系。如图 1 所示,知识点  $K_1$ 、 $K_{21}$  为复合知识点,知识点  $K_{22}$ 、 $K_{23}$ 、 $K_{31}$ 、 $K_{32}$  为不可再分

\* 基金项目:上海市优秀青年教师科研专项基金目(050703)

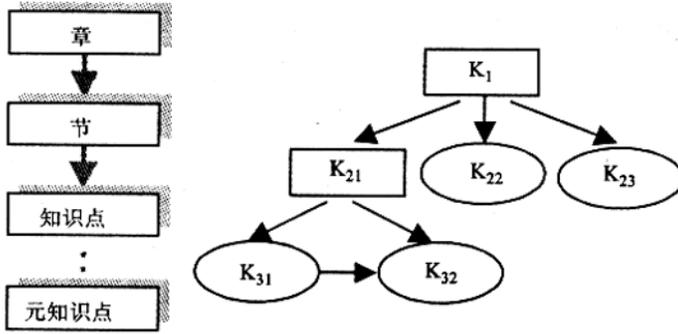


图 1 知识网络图

的元知识点。根据课程教学内容，将复合知识点细分为子知识点，则复合知识点与子知识点的关系即为父子关系，如复合知识点  $K_1$  按层次关系划分为子知识点  $K_{21}$ 、 $K_{22}$  和  $K_{23}$ 。在学习某一知识点之前必须先学习相关的另一知识点，这两者之间的关系为依赖关系，如知识点  $K_{31}$  是知识点  $K_{32}$  的前序知识，知识点  $K_{32}$  是知识点  $K_{31}$  的后序知识。图 1 按父子关系搭建成一棵知识树，按依赖关系组成了知识网络，从而可以建立起教学评价系统的领域知识库<sup>[2,3]</sup>。

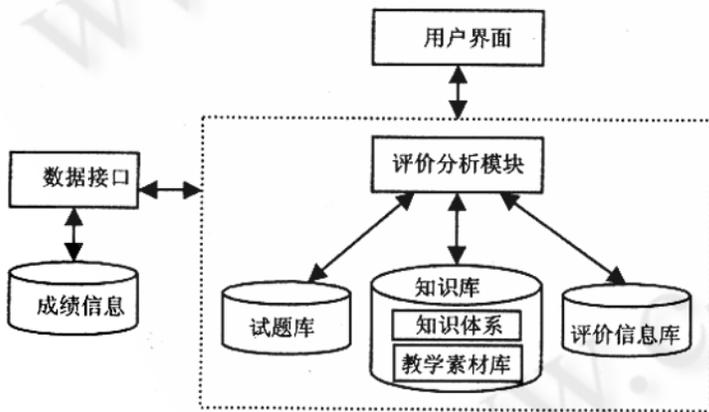


图 2 评价系统体系结构图

### 3 基于 web 的教学评价系统的总体设计目标

基于 web 的教学评价系统的总体结构，如图 2 所示，数据接口模块负责从成绩信息库中导入数据；评价分析模块结合评价信息库和知识体系库对数据进行分析，然后将得到的数值结果图形化处理生成评价报告，并针对学生薄弱环节从教学素材库中提取出需要继续学习的内容材料；用户界面模块负责显示评价分析的

结果及教学学习的改进方向。该教学评价系统主要解决以下两个目标：

(1) 标准试题库关键属性的动态更新<sup>[4]</sup>。试题库试题的难度系数是由有经验的教育专家指定的，一经指定就不再变化。由于靠经验指定必然存在偏差，即使是最有经验的教师，由于各种主观和客观原因，他对试题的难度系数估计也难十分准确，同时考试群体可接受情况也在不断的变化，因此随着系统的运行，结合学生测试成绩的统计结果，需要对试题的难度系数进行动态更新，使之更加科学化。

(2) 学习障碍的诊断。通常一个试题考察的内容是多个知识点的综合要求，如果测试者在做这个试题后，教学评价系统仅仅是发现对试题的解答有误，从而认为测试者在试题考察的所有知识上都存在缺陷是不合适的。目前的教学系统的评价机制，都是建立在这种不够细致和深入的数据上，进行统计分析和数据挖掘，产生的评价结果与测试者的真实情况很可能存在偏差<sup>[5]</sup>。本文在利用知识网络表示知识和设置试题的基础上，通过分析学生在试题的表现和试题与知识点的关联关系，可以获知学生在知识点上达到教学目标的范围和程度，对学生的薄弱环节精确定位，从而实现学生个性化的学习要求。

## 4 教学评价系统的设计与实现

### 4.1 标准试题库难度系数的动态更新

一般而言，基于题库的 web 教学评价系统中，试题所具有的基本属性项有：试题编号、试题类型、考察知识点、试题难度系数、区分度、题干、答案、试题解析答题时间、建议分数、使用总次数、上次使用时间、出题人、出题日期等。其中，最重要的属性是试题难度系数，它是衡量题目难易水平的数量指标，通常以题目的答错比率来表示。

#### 4.1.1 理论依据—贝努利大数定律

设  $n_A$  是  $n$  重贝努利试验中事件  $A$  出现的次数， $P$  是事件  $A$  在每次试验中发生的概率，则对任意正数  $\epsilon > 0$ ，有

$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{n_A}{n} - P \right| < \epsilon \right\} = 1 \quad (1)$$

因为  $n_A$  是  $n$  重贝努利试验中事件  $A$  出现的次数，所以  $n_A/n$  表示事件  $A$  发生的频率。贝努利大数定律

表明事件 A 发生的频率  $n_A/n$  依概率收敛于事件 A 的概率 P。也就是说,当 n 很大时,事件发生的频率与概率有较大偏差的可能性很小,因此可以用事件发生的频率来代替事件的概率。

#### 4.1.2 难度系数动态更新算法

由难度系数的定义可知,难度系数正比于失分概率,所以可以通过分析失分概率计算难度系数。设采用答对得分答错失分的阅卷模式,  $P_e$  为该题的失分率,  $N_e$  为累计答错次数,  $N$  为该题被累计选中次数,则  $P_e = N_e/N$ 。随着考试次数的增加,根据贝努利定理,  $P_e$  将依概率 1 收敛于失分概率 P。若一道题有多个知识点,则可将该题的失分率定义为该题所有知识点失分率的平均值。设 m 为知识点数,  $P_{ei}$  为该题第 i 个考核点的失分率,则

$$P_e = \frac{1}{m} \sum_{i=1}^m P_{ei} \quad (2)$$

#### 4.2 学习障碍的诊断

试题属性是进行考核和评估的重要依据,为了实现学习障碍的诊断,要在试题中添加评价用的属性项:备选错误现象集合。错误现象是测试者每次做完一个试题后,可能会出现的错误的特征表现的集合。对于客观试题,由计算机直接进行批改,判断正误并选择错误现象;如选择试题,答错者的选项对应内容,就视为测试者出现的错误现象。如果是主观试题,需要由人工进行批改;对没有得到满分的试题,试题批改人需要在备选错误现象中,结合测试者的解题情况,选择可能错误现象。每一个试题的错误现象集合是可以增加的。

##### 4.2.1 知识诊断

评价信息库是教学评价系统的评价标准和参照。在评价信息库中最重要的是错误原因。错误原因的内容要包含错误原因编号、文字描述、与之相关的知识点集合  $K(k_1, k_2, k_3 \dots k_i)$ 。对于试题中的每一个错误现象,都和评价信息库中一个错误原因对应,多个错误现象可以对应到同一个错误原因上。每一个错误原因,就是一个或若干个可能没有掌握的元知识点。因此通过错误现象和错误原因的对应关系,就可以确定测试者哪些知识点已经达到教学要求,哪些知识点还需要加强训练,这种方法比常规的教学评价系统更准确地说明了测试者的学习缺陷。一门课程的所有错误

原因,是这门课程上所有可能出现的缺陷的汇总,是一个比较庞大的集合。随着教学评价系统的使用,这个集合中的数据将不断地丰富。

教学素材库是指导测试者改进错误、提高能力所需数据的来源,每一个学习素材都和相关知识点建立关联。根据评价分析结果,对于没有达到教学要求的知识点,可以从教学素材库提取出与该知识点相关的继续学习内容,帮助学生纠正错误理解和观念,解决自身的学习缺陷。

教学评价要对每一个学生的测试情况给出一个评价报告,评价报告的内容有 2 个部分。第 1 部分为整个试卷所有试题的批改情况,这部分和通常的教学评价系统的评价结果类似,给出每一个试题的得分情况和整个试卷的得分情况,然后在整体分析的基础上,给出该试卷的优良中差的比例和人数,将得到的数值结果进行图形化处理,使测试者更客观和更具体地了解自己在群体中的位置和水平。第 2 部分是本系统特有部分,给学生罗列出存在的错误现象,确定每个错误现象的错误原因,汇总没有掌握的知识点,在此基础上提出个性化的学习指导意见,并给出错题题目的例题讲解以及需要继续学习的素材。

##### 4.2.2 案例分析

表 1 试题一元知识点关系表

试题编号 \ 元知识点	$K_{22}$	$K_{23}$	$K_{31}$	$K_{32}$
$Q_1$	0	7	3	0
* $Q_2$	3	1	5	6
$Q_3$	5	0	5	0
$Q_4$	0	0	2	8
* $Q_5$	3	1	3	6
$Q_6$	0	0	8	2
* $Q_7$	2	4	0	4
$Q_8$	0	3	6	1
知识点累计 Sum( $k_i$ )	13	16	32	27
错误知识点累计 Error( $k_i$ )	8	6	8	16
错误知识点比率 R( $k_i$ )	62%	38%	25%	59%

由于每道试题考查的侧重点不同,因此每个元知识点与试题相关联的程度也不相同。我们可以用一个

试题与元知识点的相关度(也称知识点在习题中的权重)加以表示。这里,我们以 1-10 的整数值来表示知识点在试题中的权重,取值如果为 0,则表示两者之间没有关系。此值的设置以试题难度及覆盖知识点范围为参考,需要有关教育专家或教师协助设置,如表 1 列出元知识点  $K_{22}$ 、 $K_{23}$ 、 $K_{31}$ 、 $K_{32}$  在 8 个题目中的相关度。表中项目  $\text{Sum}(k_{ij})$  表示知识点  $k_{ij}$  在测试表单中的总权重,即  $k_{ij}$  的所有对应各试题相关度值之和。 $\text{Error}(k_{ij})$  是与知识点  $k_{ij}$  相关的不正确答案的总权重,即学生答错题目所对应知识点  $k_{ij}$  的相关度值之和,因此该项取值要在得到学生提交的答案后计算得到。 $R(k_{ij}) = \text{Error}(k_{ij}) / \text{Sum}(k_{ij})$  表示不正确答案所占比率。

假设学生在考核中试题  $Q_2$ 、 $Q_5$ 、 $Q_7$  错误,按照常规的评价方式,如果测试者的某个试题错误,就认为测试者在试题考察的所有知识上都存在错误,则可计算出  $R(k_{22}) = \text{Error}(k_{22}) / \text{Sum}(k_{22}) = 8/13 = 62\%$ ;  $R(k_{23}) = \text{Error}(k_{23}) / \text{Sum}(k_{23}) = 6/16 = 38\%$ ;  $R(k_{31}) = \text{Error}(k_{31}) / \text{Sum}(k_{31}) = 8/32 = 25\%$ ;  $R(k_{32}) = \text{Error}(k_{32}) / \text{Sum}(k_{32}) = 16/27 = 59\%$ 。将计算结果按照大小进行排序,评价结果发现测试者有 4 个缺陷,其顺序为  $k_{22}$ 、 $k_{32}$ 、 $k_{23}$ 、 $k_{31}$ 。

表 2 测试结果分析

试题编码	改进前		改进后		
	知识点集合	错误集合	错误现象	错误原因编码	错误集合
$Q_2$	$k_{22} k_{23} k_{31} k_{32}$	$k_{22} k_{23} k_{31} k_{32}$	$k_{23} k_{31} k_{32}$ 概念理解错	$E_1$	$k_{23} k_{31} k_{32}$
$Q_5$	$k_{22} k_{23} k_{31} k_{32}$	$k_{22} k_{23} k_{31} k_{32}$	$k_{23} k_{31} k_{32}$ 综合应用没有掌握	$E_1$	$k_{23} k_{31} k_{32}$
$Q_7$	$k_{22} k_{23} k_{32}$	$k_{22} k_{23} k_{32}$	$k_{22} k_{23} k_{32}$ 综合应用没有掌握	$E_2$	$k_{22} k_{23} k_{32}$

采用本评价系统进行评价,如表 2 所示,发现测试者的错误原因有 2 个( $E_1$ 、 $E_2$ ), $E_1$  与知识点  $k_{22}$  无关,即试题  $Q_2$ 、 $Q_5$  的错误与知识点  $k_{22}$  无关,因此  $R(k_{22}) = \text{Error}(k_{22}) / \text{Sum}(k_{22}) = 2/13 = 15\%$ 。测试者的学习缺陷按照重要程度顺序为  $k_{32}$ 、 $k_{23}$ 、 $k_{31}$ 、 $k_{22}$ 。如果系统规定知识点通过阈值  $\theta$  用于表示每个知识点可接受的错误率,当  $R(k_{ij}) < \theta$ ,说明学生已经学会知识点  $k_{ij}$ ,否则仍需要加强该知识点的学习,这个知识点将被加入到“加强学习的知识点集”中。假定  $\theta$  值为 0.3,则“加强学习的知识点集”为  $\{k_{23}, k_{32}\}$ ,按照“先前序再后序”的顺序建立学习路径,如果存在多条学习路径,将路径上各知识点  $R(k_{ij})$  值由大到小的顺序排序,则选择加强学习的路径为  $k_{32} \rightarrow k_{23}$ 。

组织、领域知识的统一表示,需教学专家协助进行。如果再结合一种有效的学生模型和教学策略,就可以构造一个高效的基于 Web 的智能教学系统的模型,从而真正实现智能化教学。

参考文献

- 1 余胜泉、李世亮、谢晓林,在线测试分析与数据挖掘系统的设计[J],开放教育研究,2005,11(5).
- 2 温常君、刘曙光、杨一平,智能教学中的知识表示与知识诊断研究[J],微电子学与计算机,2005,(09).
- 3 Juntang Zeng, Yiping Yang. Information Retrieving based on Conceptual Network[J], 2003 International Conference on Natural Language Processing and Knowledge Engineering Proceedings. 2003, 380 - 387.
- 4 李一波、张森悦,试题库试题难度系数自适应学习整定[J],计算机工程,2005,31(12):181-1820.
- 5 刘小明、汤士平、樊孝忠,基于资源信息细分的学习评价系统的设计与实现[J],计算机工程与设计,2006,27(1).

5 结束语

本文利用知识网络进行了领域知识表示及考核试题设置,建立了教学材料和领域知识的统一表示及管理模式,并在此基础上设计了对试题库的难度属性进行动态更新的方法,以及提出发现学生学习薄弱环节的诊断方法,准确地将学生学习症结定位到具体的知识点,从而增强了教学的针对性。但有关教学材料的