

知识题库信息处理系统的设计与实现

王伯健 (解放军 57319 部队)

摘要: 本文介绍了知识题库信息处理系统的数据结构、功能划分、工作原理、操作流程及所采用的技术措施。

一、系统概述

知识题库包括了各种题类(如填空、单项选择、多项选择、问答、实验等)和各个范围(语文、物理、化学、英语等)的考题及相应答案。知识题库信息处理系统由 Foxpro2.5.C、汇编三种语言联合开发设计,适应性很强,处理速度快,可运行于与 IBMPC 相兼容、CPU 为 386 档、1 兆内存以上的微机。其主要功能:根据用户的出题意向及出题原则,随机生成考卷及相应答案考卷;能对题库中考题进行调阅、查询、打印;能对题库中考题进行增、删、改,并能自动实现考题指针转换,变系统对不定长文本格式文件的管理为对考题指针索引库的管理。满足了数据库管理系统对管理对象的数据格式要求。

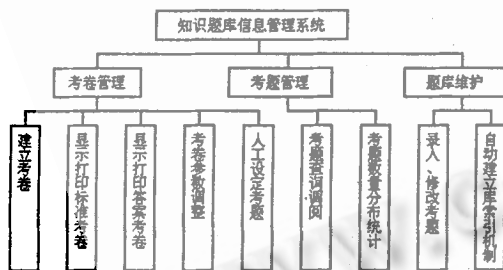


图 1

系统共分为两级目录,三个管理级功能模块,考卷管理、考题管理、题库维护,九个功能级子模块。

1.建立考卷。通过良好的交互式人机界面环境,系统接受用户出题意向,并将其转换成系统能够识别的考卷生成条件。

2.显示打印标准考卷。用户通过窗口选择某一考卷,系统根据该考卷的生成条件,随机生成标准考卷,并将其显示在屏幕上,需要时,可打印之。

3.显示打印答案考卷。在生成标准考卷的基础上,生成与之相对应的答案考卷,并显示在屏幕上,需要时,可打印之。

4.考卷参数调整。用户在生成考卷后,对考卷中题目的分布结构有不满意时(如填空题太多、选择题太少等),可通过此模块对上次用户输入的出题意向进行修改,系统根据新意向生成新考卷,循环往复,直至用户满意为止。

5.人工设定考题。对系统自动生成考卷中的个别考题有不满意时,或是用户刻意要在考卷中出某题时,可通过该模块在题库中选定若干个考题,设定在考卷中。

6.考题查询调阅打印。对题库中的考题可按题类、范围等条件进行查询调阅,并输出打印,以供参考。

7.考题数量分布统计。可按题类、范围等条件查询考题的分布及数量情况。

8.录入、删除、修改考题。该模块可利用一些通用的文本编辑器,对题库进行增、删、改,以实现用户对题库不断的维护。

9.自动建立题库索引。以全透明方式自动建立指针转换机制。

二、工作原理

指针转换机制是整个系统的运行基础。其工作原理如图 2 所示。用户通过一般文本编辑器,按人们习惯的文书格式建立,修改原题库,考题指针转换机制作用于原知识题库,生成考题指针序列索引库。

在建立了考题指针序列索引库的基础上,考卷自动生成系统依据用户出题意向,按系统所设定的规则和方式,自动生成标准考卷和答案考卷。其工作原理如图 3 所示:首先,出题意向转换器接受用户输入的出题意向

(如:化学类,填空题,10道...)并将其转换成查询条件因子的数据格式参数,该数据参数格式能满足查询子程序的入口参数格式要求;查询子程序、随机数产生机制以及出题规则算法同时作用于考题指针序列索引库,其输出为考卷中考题指针序列列表;考题指针的逆转换机制结合该指针序列列表作用于知识题库,将指针序列转换成考题和答案系列,并生成符合打印格式要求的标准文本文件,将考卷和答案显示在屏幕上,同时,可直接打印输出。用户通过调整考题子模块对考卷中考题指针序列列表进行修改。通过出题意向子模块可对某考卷的出题意向参数进行修改,系统再次生成新的标准考卷和答案,循环往复,直至用户满意为止。

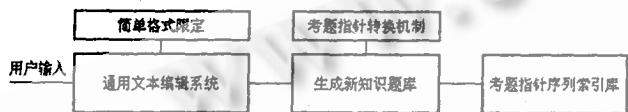


图 2

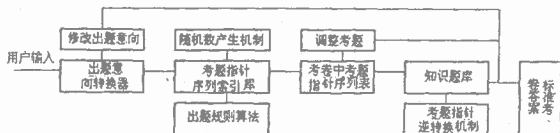


图 3

三、技术难点及采取的技术措施

1.知识题库信息处理系统的处理对象——知识题库,是由一般文本文件编辑器编辑产生的,具有不规则、不定长、未加任何标志等特点,而管理系统所要求的是按记录和字段有规则排列的信息,如何能在原文本格式文件的基本上,不加任何标志,以全透明的方式建立该转换机制是实现考题、答案、考卷管理的难点和关键所在。为了实现对上述文本格式信息的数据库形式管理,我们采用了指针转换技术,即必须在知识题库与管理系统之间,建立一种符合库格式要求转换机制,该机制的入口是不规则知识题库,它的出口是一种能够进行数据库形式管理的库格式文件,管理系统通过它能间接实现对文本格式信息的管理。通过采用特征字符串判别,对各种类别(如:填空、选择、简答、试验等)考题及相应答案进行文件

位移长度的指针定位,为每个考题在转换机制库中建立一个记录,将该考题所对应的考题指针、若干个答案指针(试验考题有若干个答案)、题类标志、题号等信息存于该记录,变系统对文本文件的管理为对标准库文件格式的指针管理。在这个转换过程中,由于文本文件的不规则性,给指针定位造成了很大困难。通过对题库中所有考题进行全面分析,我们得到了如下规律:

(1)考题类别特性:考题类别具有一定的延续性,即若干个不同类别的考题在题库中是连续出现的,因此,只要能抓住决定考题类别的类别标题,既可判别该类别标题之后,下一类别标题之前的所有考题的考题类别。

(2)考题首指针分为两类:①一般类题目首指针具有普遍性,除试验类外的其它类考题具有相同的特征字符串。②试验类具有特殊性,一个试验现象(即考题的题目部分)对应若干个答案,每个答案代表一个考题,这若干个考题拥有相同的题目,因此,试验类的题目不带题号,而以另空 4+“试验现象”开头的字符串。

(3)考题的答案指针分为两类:①一题对应于一个答案。②对应多个答案(如一道试验类考题的答案分为 2 个部分:试验分析部分和试验结论部分),存放在不一定连续的两处。

通过上述分析,我们共总结出 10 个特征字符串。通过采用指针定位技术,确定特征字符串的精确位置。系统首先确定考题类别,依据它,即可转向某个相应特征字符串的判别程序,相继确定考题的类别、首指针、答案指针、副答案指针,并将考题的前 20 个汉字一并填入考题指针序列列表库中该考题对应的记录中。

2.在指针转换机制的特征字符串判断过程中,题库中的特征字符串是由人工输入,易出错。如:“多项选择”错写为“多向选择”等等,给特征字符串定位带来很大困难。为此,我们将模糊技术引用到程序设计中,极大地提高了程序的容错性。

3.为满足系统在生成考卷时出题的随机性和考题在考卷中存在的唯一性,我们编制了题号随机发生器。该发生器应具备两个功能:随机数产生能力;所产生的题号序列不能重。采取措施如下:将满足用户条件的所有考题(如有机化学和无机化学类中所有的填空题)在转换机制库中所对应的记录取出并形成一临时数据库,该库的总题数为随机数产生的范围,每道题的物理记录号为该

题的物理题号。由随机数发生器产生一物理题号。从题中提取该物理题号考题的所有参数并将该题列入考卷中,从临时库中删除该题,保证下次不再产生该题,随机数产生范围相应减 1,继续由随机数发生器产生物理题号,依次循环,直到全部考题随机产生。

四、系统特点

1.系统操作透明性好,从考题指针索引库的建立到用户出题意向的输入、考卷建立、直至生成标准考卷和标准答案,其操作极符合人工出题的操作步骤,且自动化程

度高。

2.系统适应性强,在系统程序设计中,因充分考虑到系统对环境的适应性,故对汉字操作系统无特殊要求,只要基本内存在 400K 以上,即可运行该系统。生成的考卷既能以随机方式产生,也能以人工方式设定,基本满足了各种考试出题需要。

3.题库维护方便,由于系统对题库没有任何标志要求,只有一些简单并符合人工录入习惯的格式要求,且容错性好,故用户可通过任何标准文本编辑器对题库进行编辑、修改。系统能自动对修改好的题库进行索引处理。