

管理信息系统中数据库设计

国家教委高等工程教育培训中心 徐炳亭

管理信息系统的设计从基本上说包括两方面,即数据库设计和程序模块设计,人们比较注重于程序模块编程技巧的提高和交流,这无疑是很重要的,但是也不可忽视数据库是程序设计的基础,数据库设计不但关系到程序设计的繁简,而且也关系到整个管理信息系统性能和功能的扩展。因此本文试图从数据库设计方面谈些粗浅的体会。

1.数据库应体现数据冗余度小、数据共享的特点

数据冗余就是数据信息有重复,如果想取得某种信息可以从数据库不同字段提取或推算出来就可以说该数据库所存数据有冗余,如工资管理库中已有应发工资和应扣款字段,再设实发工资字段;商品物资库有数量和单价字段,再加金额字段;如此等等,这些都可以说是数据信息的冗余。从程序设计角度来看这样作似乎可以简化程序设计,但却给数据库维护造成隐患。有的单位或部门进行不同的工作,单独建立仅与该工作有关的数据库,特别是再设立一些与原有数据库内容重复的字段,作为管理系统来说也是不可取的。专家们对关系数据模型定义了四级范式,一般认为只有四型范式关系才是性能最好的关系模型,这主要就是指数据冗余度应该最小。

2.数据库字段中存记的信息应尽量丰富、便于维护

有的职工档案库设诸如年龄、工龄等字段,这样每年必然有一次数据更新的麻烦,若设出生日期、参加工作日字段,其年龄、工龄可以通过函数 YEARO、DATEO 算出来,数据维护量会大大减少;“人员档案库设”是否华侨港澳台”(L型)字段,不如设为 C 型字段,明确填写华侨或港胞、澳胞、台胞;如此等等,这种设置字段的方法显然其信息要丰富的多,数据的利用率会更大,管理信息也易于扩展功能。

3.数据库的数据应尽量规范、统一、标准化

人员档案库、商品物资库等一般设编号字段作为关键字建立索引文件,这样就可很迅速、方便地检索其指定

编号记录的内容。编号不但区分开了同名的不同记录对象,而且统一了同一记录对象的不同写法,如“李明”等在计算机中就是不同的名字,只能根据编号来判别是否是同一个人。编号最好有一定的意义的规律,数字位数不要过长,便于逻辑记忆和使用。一些常用、公用的数据,如省市、学校、民族、专业、课程等名称应进行标准化,建立标准字典库,用编号与其名称对应起来(有的已有相应的国家或部委或行业编码标准),当需要这些数据时,如籍贯、家庭住址、毕业学校等字段,就可只存记其编号代码,不但方便查询检索,而且会大大节约数据输入的时间、数据存储所占的空间。由于 DBASEIII 可以同时打开多个数据库,又有库逻辑连接命令(SET RELAT O),这些编号代码还原为汉字信息也是不难的。

4.管理信息系统应围绕主数据库建立多种辅助数据库

DBASEIII 同一数据库中各记录的字段个数相同,且在同一字段上留的宽度也相等。但现实世界中事物的个体及其属性是复杂、多样的,描述某事物的属性有时需要的字段数量不一样多,同一字段内所存储的内容多少也不相同,若按最多字段、最大宽度设计数据库必然造成存贮空间的浪费,还会影响程序运行速度。一般数据库以 20~40 个字段为宜,这就需要建立与主数据库相对应的关联数据库,存放一些与主数据库数据紧密相连的数据,主数据库与关联数据库的记录既可以多对一的通过索引文件连接起来(此时关联数据库多为标准字典库),也可以一对一的逻辑连接,还可以一对多的索引连接。比如:学生档案库和学生成绩库可以采用下面两种关联方式:

对数据库进行统计汇总供查询显示或打印报表,对这些经过统计汇总产生的数据应该建立统计汇总数据库。当多次查询这些统计资料时,就不必再花时间重新进行统计,可以重复调用这些数据资料。只有当主数据

库内容有了变化,再及时进行新的统计,更新统计汇总数据库中的内容。

记录一对多的关联

学生档案库记录:----->学生成绩库(一)记录(按学号索引)

学号	姓名	性别	...	学号	课名	成绩
89001	王小明	男	...	89001	数学	90
89002	李红霞	女	...	89001	英语	65
			记录一对一的关联↓	89001	政治	74

学生成绩库(二)记录(按学号索引)←

学号	数学	英语	政治	...	学号	数学	成绩
89001	90	65	74	...	89002	数学	80
89002	80	75	92	...	89002	日语	75
				...	89002	政治	92

DBASEIII 备注型(.DBT)文件适于存放内容较长、且各记录相应内容长度不一的有关数据,如个人简历、内容摘要等。还有一些公用和专用数据,数据类型又不甚一致,不便建库,则采用内存变量文件(.MEM)存记,根据需要随时调入内存使用或更新内容。

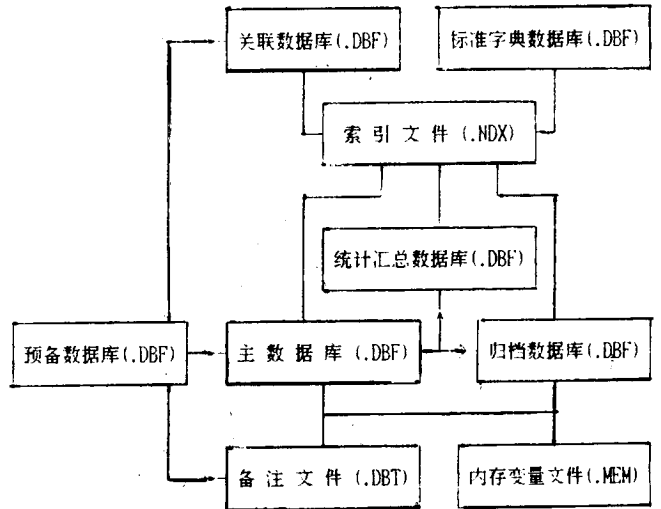
为了保证数据库中数据准确、维护方便,当向数据库输入大量数据时,应先输入到预备数据库中,经过认真地校对、核对无误后,再加入到主数据库或其关联数据库中;同样主数据库应存记常用记录的数据,一些不经常调用或要删除的记录,应存入归档数据库后给予删除,以保证管理信息系统运行的效率。

作为管理信息系统,各种数据库文件关系以及数据的流向可以使用下图表示出来。

5.数据库文件名及字段名应便于使用

文件名及字段名一般采用汉字、拼音字母或其字头、字母加字段顺序号、英文字等方式命名。初学者或编写教材时多用汉字名,如“职工档案、DBF”“代号”、“姓名”等,文件和字段内容意义直观,容易记忆,但有的名字就难于用五个汉字表示清楚,只好略掉一些字,采用难于理解的简称,此外在编写程序时,拼写汉字也不方便。汉语拼音熟练者多用拼音字母或其字头命名,如“ZGDA、DBF”“T”“DH”、“XM”等,通过汉语拼音联想其内容也比较直观,编写程序就较汉字名字方便了许多,但有时会拼出重名字段名,要设法避免。使用英文名字,如“DOSSIER、DBF”和“cobe”、“name”等,并且也常采用英文字头简化,需要懂英文,不懂英文者就颇为费解了,如果编写的管理信息系统仅供国内使用,最好不采用这种命名方法。字母加序号(如 A1、A2、A3 等)命名法仅

看名字难以知道其实际存入的内容,不直观,但易于编写出通用的程序,程序便于移植。目前多用拼音字母命名,而当字段描述的事物属性类型相同时,则采用拼音字母字头加序号命名法。



说明:——表示数据文件间联系; ·→表示数据的流向

数据库文件名中若含有其内部记录的共同属性,还可以省略某些字段,以节约存储空间,如成绩库名为 CJ9220、DBF,其中 92 为入学年度,20 为班级代号,则其库结构中就可省去入学年度、班级字段的设置。统计汇总数据库的字段名最好取自被统计库中的有关信息,比如按工作部门统计职工的籍贯分布情况,统计汇总库的字段名就可以以省市名称命名(设只取前两个汉字),取数值型字段,则统计程序编写起来很简单,其程序示意如下(为了便于阅读,这里使用了汉字名字):

```

USE 籍贯统计
SELE 2
USE 职工档案
INDE ON 工作部门 TO 工作部门
GO TOP
DO WHIL、NOT、EOFO
ELE 1
APPE BLAN
REPL 工作部门 WITH B->工作部门
SELE 2
DO WHIL 工作部门=A->工作部门、AND、NOT、EOFO
    
```

```
JG=SUBS(籍贯,1,4)
SELE 1
REPL § JG 量 WITH § JG+1
SELE 2
SKIP
ENDDO
ENDDO
NSE
```

6.数据库字段类型选择应灵活实用

对于进行数值计算的数据,如库存物资数量、工资、利息之类的字段应采用 N 型字段;对于确切知道年月日的日期,如出生日期、入党日期等应采用 D 型字段;对于只有两种状态的数据可以采用 L 型字段;字符型数据最好控制在 254 字符以内,这时可以采用 C 型字段;M 型字段实际上也是字符型数据,每个记录最少占用 512 个字节,最多可以占用 4096 字节,但 M 型字段的内容输入及调用均不太方便,因此能够压缩为 254 个字符以下,还是用 C 型字段为好。

除以上通常情况外,有时需要根据应用及编程方便来选定字段类型。有些字段内容虽由数字组成,但不进行数值计算,如编号、电话号码等,采用 C 型字段会带来某些使用上的方便;纯百分记分制的学习成绩采用 N 值型字段计算时有方便之处,但若同时登记五级分制成绩或合格不合格之类的成绩,这时就只能采用 C 型字段;对于仅存年月而不计具体日期的数据,若采用 D 型字段输入数据时,其具体时期不能跳过不写,因此不如采用 C 型字段方便。实际上,由于 DBASE 备有各种转换函数,几乎所有的数据都可以采用字符型数据表示,屏幕显示或打印使用最方便的还是字符型数据。

1 20 24 26 28 32 (字段字节序号)

课程名称	成绩	学分	补考	(成绩字段填写格式)
管理信息系统	90	T	4	(具体记录填写实例)

↑ 登记补考成绩
 成绩通过方式——(提前通过)

有时数据库字段数过多,影响运行速度,可以采用组合字段,即将内容相关、并经常在一起连用的几种数据放

在一个字段中,按规定位置存放。单独使用时,可以使用取子串函数 SUBS()取出。比如学生成绩,一门课程成绩至少需要记载其课程名称、成绩、补考成绩等内容,为了避免数据库的字段过多,就可以采用组合字段,一门课程设一个字段,填写成绩时按格式填写,如上图所示:

7.管理系统本身也要建立数种不同功能的管理数据库

管理信息系统主要以对用户的种数据、信息的管理、处理为目的,但其本身甚至程序模块中也牵涉到许多数据,若存入相应数据库中,使可以使程序与数据相对独立,提高管理信息系统的通用性、可移植性、可扩展性、易维护性、操作方便性、安全保密性等。比如:设置操作登记库,内设口令操作者姓名、操作日期和时间等字段,口令预先登记,各操作者使用自己的口令启动系统,顺便就在操作登记库中进行了登记,变更口令也方便,不必“惊动”程序模块。建立使用说明库,存记操作使用说明,屏幕显示和打印使用说明等模块设计起来就简单多了。各级功能菜单可以存入功能菜单数据库,内设菜单代号、功能、级别、调用模块等字段,只编写一个通用菜单模块即可实现对整个管理信息系统各级功能菜单的管理和显示,扩展系统功能也不必“打扰”原设计的系统模块,只是在通用菜单数据库中按格式要求填加一个记录即可。输入、修改、显示等屏幕格式也可建立相应的屏幕格式库,存记屏幕各行的显示格式和相应的数据库名、字段名及其屏幕显示坐标位置序列等内容,只要编写一个通用屏幕格式显示程序,调用该数据库中的参数形成所需的屏幕显示即可,程序员可以大大减轻程序模块的设计工作量,用户可以根据需要,自行向屏幕格式库增加记录,扩展屏幕格式显示功能。此外还可以建立相应的数据库存储整个管理信息系统的库结构和文件,加必要的汉字说明和注,必将给管理信息系统设计和使用带来更多方便。

DBASEPLUS 及 FoxBASE+相对于 DBASEIII 增加了许多函数,如 DBF()、FIEL()、NDX()、STUF()、IIF()等函数,给数据库设计带来了更多方便,使程序中的语句表达更为简洁。在数据库设计中注意前述的问题,在程序设计中灵活运用各种函数及编程技巧,就一定会开发出实用、通用而又理想的管理信息系统。