

一种数据库编程排版方法的研究与实现

白仰东

(陕西省科学技术信息研究所, 陕西 西安 710054)

摘要:采用批处理排版软件对符合结构化数据特性的内容排版时,要求在录入内容的同时输入大量重复的排版命令,效率低下,容易出错。给出了一种数据库编程排版方法,通过一个实例说明了这种方法的实现过程。该方法以 FoxPro 数据库管理系统与方正书版排版系统相结合为例,先将排版内容录入到数据库中,然后通过编程将数据库文件直接转换为方正书版小样文件,排版命令均由程序自动生成,不仅提高了排版效率,而且排出的版面风格统一,版式规范,格式准确。

关键词:FoxPro; 方正书版; 排版; BD 排版语言; 批处理

中图分类号:TP317.2

文献标识码:A

文章编号:1673-629X(2007)07-0226-04

Research and Implementation on a Method of Database Programming Typesetting

BAI Yang-dong

(Institute of Scientific and Technical Information of Shannxi Province, Xi'an 710054, China)

Abstract: When apply the batch processing typesetting system to typeset the content that correspond to the characteristics of the structured data, it is necessary to input a lot of typesetting commands while inputting the content, so the efficiency is low and it is easy to make a mistake. In this paper a method of the database programming typesetting is proposed. By means of a real example, illustrate the implementation procedure of the method. Take FoxPro DBMS and Founder typesetting systems as an example, firstly input the content of typesetting to the database, and then transform the database file by programmed directly into the source file of the Founder typesetting system. All typesetting commands are automatically generated by the program. Not only is the typesetting efficiency improved but also the style is unified. The typesetting style is standard and the format is accurate.

Key words: FoxPro; Founder typesetting system; typesetting; BD typesetting language; batch processing

0 引言

数据库系统把大量的数据按照一定的结构存储起来,在数据库管理系统的集中管理下,实现数据共享。由于数据库具有数据结构化、数据独立性高、数据共享和易于扩充等特点,因此被广泛地应用于各种管理系统中。VisualFoxPro6.0 是美国微软公司推出的关系数据库管理系统,具备友好的界面、强大的功能、便捷的工具、简单的编程机制等特点。VisualFoxPro6.0 及其中文版是可运行于 Windows95 及其以上操作平台的 32 位数据库开发系统,它不仅可以简化数据库管理,而且能使应用程序的开发流程更为合理。VisualFoxPro6.0 提供了一个集成化的系统开发环境,它不

仅支持过程式编程技术,也支持面向对象可视化编程技术^[1],是目前微机上最流行的新一代小型数据库管理系统,深受许多小型数据库应用系统开发者的喜爱^[2]。

方正书版(BD 排版语言)是北大方正电子有限公司自主开发的专业排版软件,采用批处理的排版方式,具有强大的文字处理功能,排版速度快捷,排版效率高,版式多样,适用于期刊、辞书、典籍、科技类和文艺类等书刊的编辑和排版,广泛运用于出版社、杂志社、排版中心、政府办公等领域,并形成批处理排版方式的行业标准^[3]。

排版软件一般可分为两大类:一类是“批处理排版软件”,另一类是“交互式排版软件”。因为这两类软件的使用方法完全不同,各有优缺点,适用范围也各不相同,所以到目前为止这两类排版软件能够长期共存,各有发展。

批处理排版软件也叫“注解式”排版软件。使用批

收稿日期:2006-11-23

基金项目:陕西省二〇〇五年重大科技创新项目计划(2005ZKC(二)10-03)

作者简介:白仰东(1951-),男,高级工程师,主要从事情报技术工作。

处理软件排版是在输入文稿中加入专用的命令(注解)用以说明版面的排法和要求,如文字字体、字号的大小、文章标题的位置及指定一本书的版心尺寸等,这种中间插有排版命令注解的文本文件也叫小样文件。通过运行批处理排版软件,对小样文件先做语法检查,而后进行编译处理,生成一个页面文件专门用于输出,这个页面文件也叫结果文件或大样文件,它就是排版者要得到的版面,操作者可以在计算机屏幕上看到模拟的排版结果。批处理排版软件的优点是高效精确,缺点是不易学习和使用,适合编排内容较长,版式比较简单的图书和刊物。

交互式排版软件可以直观地在计算机屏幕上显示排版版面,操作者通过操纵键盘和鼠标器,可以在屏幕上直接输入、修改版面文字、图片等内容,又可以直观地对版面格式编辑调整,每进行一项操作,都可以在屏幕上看到结果,以“所见即所得”(what are you see is what you get)的方式实现排版。交互式排版软件的优点是直观方便,但在版面编排中需要较多的人工调整,效率较低,适合编排比较复杂的版面^[4]。

在实际工作中有时需要对大量排版样式相同的内容进行排版印刷,比如图书目录、成果汇编、产品目录等,这些内容都有一个共同的特点,即每条信息都具有相同的排版样式,只是数据内容不同,也就是说这些数据具有明显结构化特性。由于每条数据的排版样式相同,这就要求每个排版人员的排版效果要达到一致,如果采用批处理排版方式,在排版录入时不但要录入文字内容,同时还要录入相同的命令(注解),排版工作非常繁琐,人为出错的几率也很高。事实上要做到排版效果完全的一致是有困难的,只能是尽量的一致,这样就会使最终的排出来的效果不能够完全保持一致。如果采用交互式排版方式,要人工逐条定义字体、字号等,工作效率不高,排版效果更难保持完全一致。

对于结构化数据,使用关系型数据库管理系统(RDBMS)进行管理是目前最为流行的一种方式,一般排版软件没有排序、分类、查重等功能,不能对录入的内容进行有效管理。数据库编程排版是将编排的内容录入到数据库中,通过程序将其转换为符合方正书版要求的小样文件,这种方法将数据库系统和方正书版系统相结合,充分发挥两个系统的特点,不仅可以有效地组织数据,大大减低排版的工作量,而且排出的版面风格统一,版式规范,格式准确^[5]。

1 实现的方法

首先按照内容要求用 FoxPro 建立数据库,定义表结构。表结构根据所要编排的内容定义,对于长度较

短的内容如标题、作者等可选择字符型字段,对文字内容超过 254 个字符的内容,如文摘、项目简介等应选择 memo 备注型字段,因为在数据库文件中可以定义的字符型的字段最多可以存放 254 个字符,超过这个限制,字符型字段变量就无法实现,而 memo 备注型字段的基本容量是 512 个字节,其实际长度根据存放的信息量以 512 个字节的倍数是可变的,没有限制,完全可以满足处理大段文字的要求。建立数据库,定义表结构过程和一般建立数据库的过程是相同的,不同的是在表结构定义好以后要在各字段间插入注解字段,注解字段应选用字符型。

数据库结构定义好以后就可以将数据录入到数据库中。为了录入方便,可以编写一个简单的编排内容录入程序,录入程序不包括注解字段,这样可以减少录入者对数据库物理结构的依赖,使录入者将注意力集中到编排内容上。录入程序还应该具有检索功能,以便在以后的修改中能够迅速找到要修改的记录,检索功能在 FoxPro 中是很容易实现的。录入完成后,可以对数据库进行排序、查重、校对等整理工作。

最后编写转换程序。先用方正书版系统根据要求排出一段示范小样文件,根据示范小样文件中排版注解的特征编写程序中的插入排版注解部分。可以看出,对于不同的排版要求,程序有所不同^[6],但实现插入排版注解的原理是相同的。转换程序生成的小样文件要经过方正书版软件的验证,这个过程有时可能要反复几次,直到编排的结果满意为止。

下面通过一个编排文献资料目录的例子来说明程序的工作原理。文献资料目录的内容字段包括【正题名】、【作者】、【单位】、【分类号】、【关键词】和【文摘】,要求【正题名】左对齐单独排一行,字体为 5 号书宋,【作者】和【单位】排一行,【分类号】和【关键词】排一行,字体为 5 号楷体,【文摘】另起一行顺延排版,字体为 5 号书宋。其表结构如下所示,ZTM 为【正题名】、ZZH 为【作者】、DW 为【单位】、FLH 为【分类号】、GJC 为【关键词】、WZH 为【文摘】,其中 ZJ1 到 ZJ6 为方正书版的注解字段,可以看出,方正书版的注解字段分布在内容字段之间,这样做的目的是为了满足不同 BD 排版语言的要求在内容中插入方正书版注解命令。

表结构: E:\FOXPRO\PB1.DBF

数据记录数: 1000

最近更新的时间: 10/18/06

备注文件块大小: 64

代码页: 936

字段	字段名	类型	宽度
1	ZJ1	字符型	30

2	ZTM	字符型	50
3	ZJ2	字符型	40
4	ZZH	字符型	30
5	ZJ3	字符型	30
6	DW	字符型	30
7	ZJ4	字符型	30
8	FLH	字符型	10
9	ZJ5	字符型	30
10	GJC	字符型	20
11	ZJ6	字符型	30
12	WZH	备注型	4
* * 总计 * *			335

表结构定义好以后,就可以按照 FoxPro 的要求录入数据,也可以将其它数据库的数据转入到 FoxPro 中。数据录入时仅对内容字段录入,ZJ1 到 ZJ6 注解字段中的内容由程序自动生成。数据录入工作完成后,可以采用 sort on 命令对记录进行排序整理,比如在本列中可以按分类号排序,FoxPro 的排序功能是很强的,可以指定一个和多个关键字,这样在录入数据时可以不考虑录入顺序,录入完成可对数据库进行排序处理。

下面给出的是上述例子的转换程序,该程序已通过 Microsoft visual foxpro v6.0 中文版和北大方正书版 9.1 版验证,运行程序就可以直接将 pb1 表中的数据转换成符合上述要求的方正书版小样文件 abc.fbd,转化过程完全自动执行,无须人工干预。

```

set default to e: \ foxpro
if file("abc.fbd")
    delete file abc.fbd
endif
copy file begin.txt to abc.fbd

use pb1
repl all zj1 with "⌘⌘【HT5H】正题名:【HT5SS】"
repl zj1 with "⌘【HT5H】正题名:【HT5SS】" for
recno() = 1
repl all zj2 with "⌘【HT5H】作 = 者:【HT5K】"
repl all zj3 with "【DW】【HT5H】单 = 位:【HT5K】"
repl all zj4 with "⌘【HT5H】分类号:【HT5K】"
repl all zj5 with "【DW】【HT5H】关键词:【HT5K】"
repl all zj6 with "⌘【HT5H】文 = 摘:⌘【HT5K】"

renn = recc()
ren = 1
do while ren <= renn
    if file("temp.txt")
        delete file temp.txt
    
```

```

endif
copy to temp.txt for recno() = ren sdf
go ren
copy mem wzh to temp.txt addi
run copy abc.fbd + temp.txt abc.fbd
ren = ren + 1
enddo
run copy abc.fbd + end.txt abc.fbd
return
    
```

程序的开始是初始化部分,begin.txt 是一个方正书版的起始小样文件,这个文件可以根据要求存放排版小样文件起始注解。开始先将 begin.txt 拷贝到 abc.fbd,生成一个方正书版的起始小样文件,在本列中这个文件仅存放位标【WB】注解,以便确定后边的对位【DW】注解的位置。end.txt 文件为方正书版的结束小样文件,这个文件一般用来存放排版要求的结束注解,在本列中仅有小样文件结束符。引入起始小样文件和结束小样文件,使得最后生成的小样文件无须任何人工修改,就可以直接通过方正书版程序编译,形成满足显示和输出的大样文件。由于在其它系统下无法输入 BD 排版语言的特殊字符,起始小样文件和结束小样文件应该在方正书版环境下建立。

程序的第二部分是对每个记录生成排版注解,采用 replace all 命令将所有记录的注解字段的内容替换成所需的方正书版排版注解。

程序的第三部分是文件转换,可以看出,转换程序的作用实际上是按字段顺序和记录顺序将整个数据库文件转换成文本文件。首先采用 copy to 命令将当前记录中除 memo 备注型字段的内容拷贝到临时文件 temp.txt 中,然后采用

copy memo wzh to temp.txt addi 命令将该记录中的 memo 字段的内容追加拷贝到临时文件 temp.txt 后,最后调用 DOS 命令

run copy abc.fbd + temp.txt abc.fbd 将 temp.txt 文件追加到 abc.fbd 后,转换的过程是逐个记录循环拷贝,直到所有记录拷贝完成,最后将结束小样文件 end.txt 追加到 abc.fbd 后,这样就得到了完整的方正书版小样文件 abc.fbd。直接用方正书版程序打开 abc.fbd,即可直接显示和输出的大样文件。

需要指出的是 copy memo 命令是 FoxPro 的一条命令,这是 FoxPro 的前身 FoxBASE 和 dBASE III 所没有的,它的作用是将当前记录的备注字段的内容复制到一个文件中,但这条命令只能将当前记录的备注字段内容复制出来,而 copy to 命令执行后记录指针总是指向最后一个记录,所以循环采用记录号进行控制^[7]。

方正书版注解中有一些特殊符号,如换行符、换段符等在 FoxPro 环境下无法输入,解决的办法是在方正书版环境下输入这些符号,然后用复制粘贴的方法拷贝到 FoxPro 环境中,尽管这些符号在 FoxPro 环境中无法正常显示,但不会影响程序的功能。

表结构中的注解字段可以事先定义好,也可以在录入完成后通过修改数据库结构的方法插入这些字段,注解字段要留有足够的宽度。有时需要对数据库中记录的顺序进行调整,而现有的字段又不能满足排序要求,这时可以增加一个专门用来排序的临时字段,排序完成后将该字段删除,以免影响排版结果。

对文字进行校对工作时,可以用数据库直接打印出校对稿,也可以先运行转换程序把数据库文件转换成小样文件,然后利用方正书版软件编译生成的编排结果打印出校对稿,但对原始数据的修改工作都应该在数据库中进行。

2 结 论

对于信息量大、规律性很强的出版物的排版制作,应该尽量交给计算机来自动完成,这样才能够压缩制作的周期,同时保证排版效果的完全一致,而且人为出错的几率也会大大降低。数据库编程排版方法改变了传统将内容和排版注解录入小样文件的排版方法,而是将内容录入到数据库中,通过程序自动添加排版注解,将数据库文件转换为符合方正书版要求的小样文件,最后通过批处理排版软件编译处理,获得用来输出的排版结果文件。与传统的排版方法相比,这种方法具有以下特点:

1)将大量重复性容易出错的工作交给程序完成,不仅减少了工作量,提高了工作效率,而且排除的版面一致性好。

2)采用关系型数据库管理系统(RDBMS)管理编排数据,可以有效地管理数据,充分利用数据库的功能

对数据进行分类、排序、查重等整理工作,非常适合边收集资料边进行编排的工作。

3)在实现的过程中,不仅完成了编排任务,同时还建成了数据库,这个数据库还可以用作其他用途。

4)将 FoxPro 命令与 DOS 命令相结合的方法,有效地解决了对 memo 备注字段内容的连续转换,满足了对文摘、内容简介等大段文字内容的排版要求,扩大了适用范围。

5)编排的信息虽然统一地存放 to 同一个数据库中,但在编排的制作时可以通过数据库的选择功能对数据进行选择输出,以满足不同的需求。

总之,数据库编程排版方法的实现过程灵活、方便,对系统的环境要求不高,可以在没有网络环境下的单台微机上实现,也可以在网络环境中应用。数据库编程排版方法也可以推广应用到其它批处理或叙述标记系统(descriptive markup system)^[8]的排版软件中,如 LaTeX 排版系统等等。

参考文献:

- [1] 徐容宽. Visual FoxPro 6.0 简明教程[M]. 南京:东南大学出版社,2005.
- [2] 梁锐城. Visual FoxPro 数据库应用教程[M]. 北京:科学出版社,2005.
- [3] 北京大学计算机科学技术研究所,北京大学新技术公司培训部. BD 排版语言[M]. 北京:[出版者不详],1992.
- [4] 杨 勇,周庆才. 电子排版技术:方正飞腾 4.0[M]. 北京:电子工业出版社,2004.
- [5] 易桂生. 计算机文字排版技术[M]. 北京:科学出版社,2003.
- [6] 丁革建. BD 排版语言与数据库输出表格设计[M]. 微机发展,1996,6(6):44-45.
- [7] 仲秋雁,李 弘. FoxPro 3.0 实用程序设计[M]. 大连:大连理工大学出版社,1994.
- [8] 罗振东,葛向阳. 排版软件 LATEX 简明手册[M]. 第2版. 北京:电子工业出版社,2004.

(上接第 225 页)

位人才,为营造多渠道就业环境搭建了一个好的平台。通过该方法,对高校招生、教职工科研方面进行了分析,并取得了许多有价值的信息。

通过实例也可以看到,利用可辨识矩阵和逻辑运算的约简算法,比较繁琐,下一步工作是对该属性约简算法做进一步改进,提高挖掘效率。

参考文献:

- [1] Pawlak Z. Rough Sets—Theoretical Aspects of Reasoning About Data[M]. Dordrecht: Kluwer Academic Publishers, 1991.

- [2] 王国胤. Rough 集理论与知识获取[M]. 西安:西安交通大学出版社,2001.
- [3] Skowron A, Rauszer C. The discernibility matrices and functions in information systems[C]//In: Slowinski R. Intelligent Decision Support - Handbook of Applications and Advances of the Rough Sets Theory. [s.l.]: Kluwer Academic Publishers, 1992.
- [4] 叶东毅,陈昭炯. 一个新的差别矩阵及其求核方法[J]. 电子学报, 2002, 30(7):1086-1088.
- [5] 王国胤,于 洪,杨大春. 基于条件信息熵的决策表约简[J]. 计算机学报, 2002, 25(7):759-766.
- [6] 王国胤. 决策表核属性的计算方法[J]. 计算机学报, 2003, 26(5):611-615.