

基层报表统计处理系统

朱卫东 (北方交通大学计算中心)

一、概述

“七五”国家科技攻关项目《基层报表统计处理系统》是受国家统计局委托而开发的。该系统的任务是将“七五”国家科技攻关项目基层年报表录入到计算机中,进行各种统计分析和评估处理,最后将结果进行计算机排版印刷成册。

“七五”攻关项目合同数共有 4664 个,每年合同执行单位都要填报一份基层报表。五年共有 2 万余份报表。每份报表包括九种类型的记录,120 多种数据项,而且,由于各种原因每年的基层报表格式都不一样,总的数据量大约两年多万字节。要求对这些报表进行录入、校对、统计、评估和处理后生成 200 余种报表,最后生成的报表排版印刷。因此“七五”国家科技攻关项目基层报表的处理不但数据处理量十分巨大,而且非常复杂。传统的报表处理软件只能处理单一记录结构或固定长的二维表结构,而且所生成的报表不能直接用于电子排版印刷。为了满足要求,我们在 IBM PC 机上用 C 语言开发了这套软件系统,该系统能完成原始报表的录入、检错、统计分析和评估处理,并将最后生成可直接在激光照排机上制版的文件。在实际应用中,取得了令人满意的效果。

二、系统结构

本系统由数据录入统计处理、报表排版输出三部分组成,系统的处理流程如图 1 所示:

1. 数据录入部分

数据录入部分完成“七五”基层报表的录入、复录、改错以及数据录入操作的技术统计,该部分由下述几部分组成:

(1) 定义数据字典。数据字典包括了对“七五”基层报表的结构和内容的描述。这些描述有报表中数据项的名字、数据在记录中的位置、长度和值域。在报表的录入及后期处理中都要借助这个字典读取所需的数据。

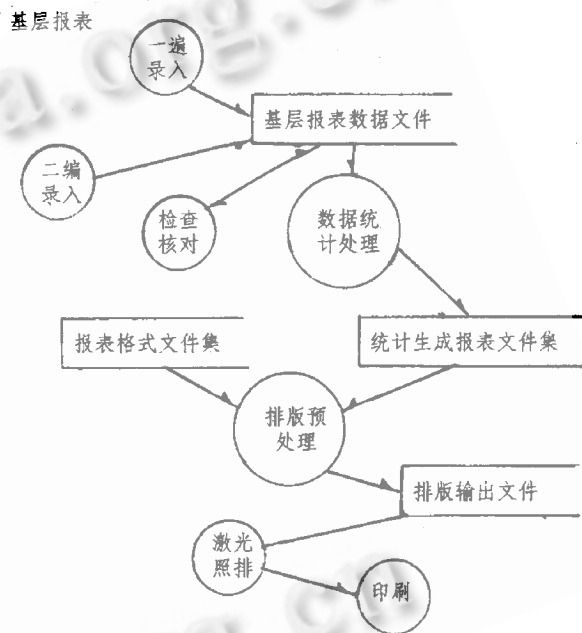


图 1

(2) 定义报表数据文件格式。报表数据文件是基层报表录入后的磁盘文件,由于“七五”基层报表格式复杂,数据种类多样,如何组织记录格式对磁盘空间的利用和其后的处理是非常重要的。本系统中采用不定长记录存放报表,所有数据项都用字符方式存放。每份基层报表在文件中占若干个记录,记录的基本格式如下:

公共区	记录类型	记录区
-----	------	-----

- ① 公共区 每份基层报表中所共同的数据描述,公共区包括四种记录项:专题合同编号,主持部门,单位类型,地区代码。
- ② 记录类型 本系统中,将每份基层分成九种记录类型,用数字 1~9 来表示这些记录类型。

③记录区 每一种记录类型所包含的数据项。
九种记录格式举例如下:

公共区

1	2	3	4	5	7	8	9	11	12	13	14	15
NN1	NN2	NN3	RECTYPE	ZCBM	DWLX	QDM						
项目号	课题数	专题合同数	记录类型	主管部门	单位类型	地区代码						

记录类型1格式(专题基本情况)

16	20	26	30	34	38	42	51	57
JA1	JA2	JA3	JA4	JA5	JA6	JA7	JA8	JA9
合同签订日期	计划完成日期	负责人年龄文化程度职称	拨款类型	偿还拨款金额	已还金额	专题活动类型	专题任务形式	合同签订形式

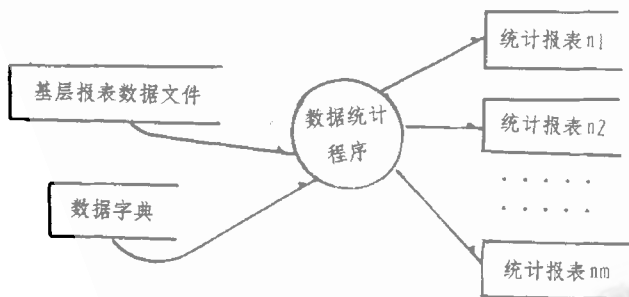


图 2

(3)数据录入模块。数据录入模块的主要功能是:根据数据字典将要输入的基层报表格式显示在屏幕上,接受数据录入人员的输入。该模块提供下述功能:

“修改”,提供对已录入报表数据的查找和修改功能;

“复录”,为了保证录入质量,对已录入的报表进行重复录入,边录入边和上次录入的数据进行比较,若不一致,则需从新核实才能录入,这样可找出大部分错误。

“统计”,提供录入操作工作量的统计信息,便于考核录入人员的工作情况。

(4)数据的检查及校正模块。该模块的功能是对已

录入的报表进行平衡关系检查,并将不平衡的报表按格式打印或显示,供有关人员找出错误原因。另外,由于一些特殊的原因需要对已录入部分进行人工核实和修改时,完成数据的提取。

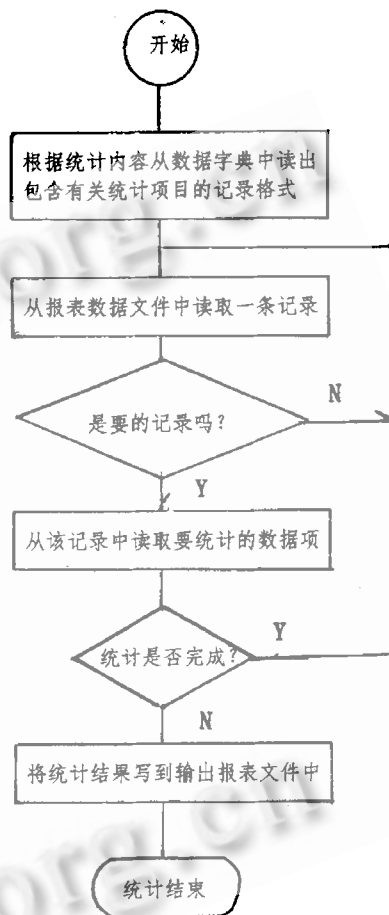


图 3

2.数据统计处理部分

该部分的主要功能是将已录入的基层报表数据文件按一定要求进行分门别类的统计。“七五”报表要求统计的项目近 300 种,统计工作量是很大的。“七五”统计要求按四种统计类型统计,它们分别是:按 36 个主持部门,76 个分类项目,6 种专题活动类型,4 种单位类型。统计项目可分五大部分。

(1)综合报表

(2)专题基本情况

(3)专题投入情况

(4)专题成果

(5)专题成果应用情况

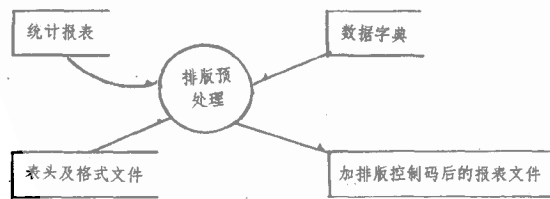


图 4

由于统计报表的数量很大,在实际处理中,将要统计的报表进行分类,将相近的放在一组,每组用 C 语言编写一个数据统计程序进行处理,处理过程如图 2。

这样运行一次统计程序 N 就可生成 M 个报表。数据统计处理流程如图 3。

3. 报表排版输出

根据统计局的要求,所有统计处理后生成报表都要排印成书。也就是说最后要进行一次排版。通常的计算机排版系统需要用文字编辑系统对要排版输出的文件逐行加排版控制码,然后再经排版产生排版输出文件,最后

进行激光照排制版。由于“七五”统计所产生的结果都是各种表格形式,表格处理排版是最复杂最费时的。要对“七五”统计输出的几百种报表用常规的方法进行排版至少要用几个月的时间,而且易出错,不能按期完成任务。为了解决这一难题,我们专门开发了一套报表排版预处理软件。该软件能取代人工排版操作,将由统计程序生成的各种报表,按要求的输出格式加上相应的制表控制码,所生成的文件可直接用于排版。该软件的处理过程如图 4。

最后的排版过程如图 5。

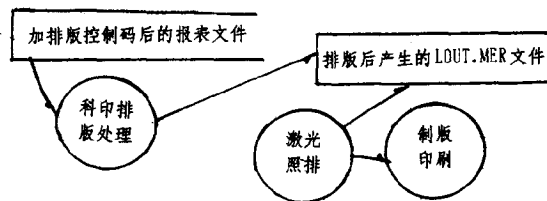


图 5