



测绘新技术小丛书



赵 华

测绘数据库

测绘出版社

测绘新技术小丛书

测 绘 数 据 库

赵 华

测 绘 出 版 社

32337

目前,作为计算机科学重要分支之一的数据库技术发展迅速,该技术在测绘领域的应用为测绘资料的管理与保存开辟了一条新的途径。本书除概述数据库的一般知识外,还介绍了测绘数据库的内容及其建库方法,并用实例阐述了在小型机上建立区域性测绘数据库的可行性及建库的具体过程,提出了在高档微型机上建库设想。

本书作为测绘数据库的入门读物,内容通俗,阐述清晰,可供广大测绘专业人员及地质、地理和计算机等专业的有关人员参考。

测绘新技术小丛书

测绘数据库

赵 华

*

测绘出版社出版

张家湾印刷厂印刷

新华书店北京发行所发行·各地新华书店经售

*

开本 787×1092 1/32·印张 4.5·字数 96 千字

1987年3月第一版·1987年3月第一次印刷

印数 0,001—3,500册·定价 1.00 元

统一书号: 15039·新562

目 录

一、数据库的基础知识.....	(1)
(一) 数据库的概念.....	(1)
1. 数据库简介.....	(1)
2. 数据库的组成.....	(2)
3. 数据库系统的发展过程.....	(3)
4. 两种不同类型的数据库.....	(7)
(二) 数据的逻辑组织.....	(8)
1. 数据.....	(8)
2. 文件、记录和数据项.....	(9)
3. 关键字.....	(10)
(三) 数据库的逻辑结构.....	(11)
(四) 数据模型.....	(13)
1. 数据联系的方式.....	(13)
2. 常用的几种数据模型.....	(15)
(五) 数据的物理组织.....	(21)
1. 计算机存贮信息的方法.....	(21)
2. 物理级的文件与记录.....	(26)
3. 顺序文件、索引文件和倒排文件.....	(27)
(六) 数据语言.....	(33)
1. 数据描述语言 DDL.....	(33)
2. 数据操作语言 DML.....	(36)
(七) 数据库管理系统 (DBMS)	(38)
1. DBMS的功能.....	(38)
2. DBMS的组成.....	(39)

3. 数据库管理员 DBA.....	(39)
4. 数据库管理系统与操作系统的关系.....	(40)
5. 数据库的访问过程.....	(42)
(八) 汉字信息处理.....	(43)
1. 汉字输入.....	(44)
2. 汉字库.....	(46)
3. 汉字输出.....	(46)
(九) 数据库设计的一般原则.....	(47)
1. 数据的独立性.....	(47)
2. 数据的冗余和共享.....	(47)
3. 使用方便程度.....	(48)
4. 可靠性、完整性和安全性.....	(48)
5. 数据的重新组织.....	(49)
6. 数据库系统的可扩展性.....	(49)
二、测绘数据库概述.....	(50)
(一) 地理信息数据库的特点.....	(50)
1. 地理信息.....	(50)
2. 地理信息数据库的数据结构.....	(51)
(二) 测绘数据库的内容.....	(55)
(三) 大地测量数据库.....	(56)
1. 大地测量数据库的用途.....	(56)
2. 大地测量数据库的数据结构.....	(57)
(四) 高程数据库.....	(61)
1. 高程数据库的用途.....	(61)
2. 高程数据库的数据结构.....	(62)
(五) 重力数据库.....	(65)
1. 重力数据库的用途.....	(65)

2. 重力数据库的数据结构·····	(66)
(六) 美国 NGS 的大地、高程、重力数据库···	(67)
1. 大地测量数据管理系统 GDMS 的概况·····	(67)
2. NGS 的地形高程数据库·····	(68)
3. NGS 的重力数据库·····	(69)
(七) 地名数据库·····	(74)
1. 地名普查工作·····	(74)
2. 地名数据库的建库方法·····	(75)
(八) 地图数据库·····	(77)
1. 地形图数据库·····	(77)
2. 专题地图数据库·····	(78)
三、小型数据库的建库方法·····	(92)
(一) 区域性测绘数据库的作用·····	(92)
(二) 小型计算机的磁盘功能·····	(94)
1. 磁盘机概况·····	(94)
2. 磁道、段地址和控制信息·····	(96)
3. 信息存贮方式·····	(98)
4. 程序文件的读/写 (DC 型)·····	(99)
5. 数据文件的读/写 (DC 型)·····	(101)
6. 数据文件内逻辑记录的跳越·····	(102)
7. 绝对寻址法数据的读/写 (DA 型)·····	(105)
8. 数据压缩技术·····	(107)
(三) 小型测绘数据库的应用实例·····	(109)
1. 区域性高程数据库 EDB 系统·····	(109)
2. 区域性重力数据库 GDB 系统·····	(115)
3. 区域性大地测量数据库 GDL 系统·····	(118)
4. 机助编图与信息管理系统 AUXPLOT···	(121)

（四）微型机上建立测绘数据库的前景·····	（ 127 ）
1．通用数据库 dBASE II 的概况·····	（ 128 ）
2．dBASE II 的建库过程·····	（ 129 ）
3．dBASE II 的运算功能·····	（ 134 ）
4．dBASE II 的检索功能·····	（ 135 ）

一、数据库的基础知识

(一) 数据库的概念

1、数据库简介

数据库 (*Data Base*) 是计算机技术的一个重要分支,是六十年代初兴起的一门新学科。目前,数据库技术虽然在工程实践中已经相当成熟,但它的概念和方法仍在不断发展。另外,数据库技术的实用性很强,不同的应用对象,其内容和结构差别很大。因此,很难用简练的语言来概括其全部特征,至今还没有一个完整、确切、通用并得到公认的定义。在许多有关数据库的文献中,人们都是根据数据库的现状和对它的认识给出各自的定义和解释的。

通俗地说,数据库技术就是把各种有用的信息按照一定的数据模型组织起来,并以一定的物理布局存入计算机的外存介质(如磁盘、磁带等),通过软件进行控制管理,用户可以随时迅速而准确地检索到所需要的信息,并按照不同的使用目的,通过应用程序对它们进行各种加工处理,最后选择一定的方式(数据、表格或图件)输出结果。

一种比较概括的定义则是:数据库是以一定的组织形式存贮在一起的互相有关联的数据集合。值得注意的是,数据库是大型集中的数据集合,但并不是简单的相加。它是通过重新组织信息,去掉其重复多余部分,增加数据间关系的描述,使其能以多种组合方式为多个用户所共享。数据库管理也不是简单地存取信息,更重要的是加工信息,这不仅能扩展

人们的记忆,而且能帮助人们去控制与存贮信息相关的事物。

对于不熟悉数据库的人来说,把数据库比作图书馆更易于理解。图书馆并不仅仅是一个存放图书的书库,而且还是一个管理图书的机构。它有一套完整的书卡作为藏书的模型(相当于逻辑结构),又有一种书卡与书架的对应关系使书库里的书按照一定的顺序规则存放(相当于物理布局)。此外,图书馆还制订了严密的借阅制度,规定了借阅权限、借还手续等。借书时读者出示借书证,图书馆管理员验明身份和借阅权限(相当于用户使用数据库时要在终端上键入身份密码进行“注册”);读者填写借书单(相当于“访问请求”);管理员根据书卡与书架的对应关系表查找数据(相当于“检索”);找出书籍交付给读者(相当于“响应”);最后填写借阅登记(相当于“日志”)。而图书管理员的工作规程则相当于数据库软件。

2、数据库的组成

无论什么数据库都包括两个最基本的组成部分:一是素材库,即数据库本身,它存放全部原始信息,是数据库系统加工的对象;二是程序库,即数据库软件,它存放管理和控制信息的各种程序,是数据库系统加工的手段。

素材库中存放的全部数据除了要遵循一定的组织方式(逻辑组织和物理组织)以外,在入库前还要经过一系列的预处理。例如,列为数据项的数据应当有独立的意义,凡是通过计算可以取得结果的,一般不入库体;对原始信息要经过规范化,各种代码、量级都要有规范性的定义和描述;对于连续性的变量要进行数字化;对同一数据的命名要避免出现二义性等。

程序库存放的各种程序，可以用指令、汇编语言、数据库专用语言或其他通用的程序设计语言来编写。数据库软件的内容十分丰富，其中最基本的是数据库管理系统、操作系统、各种语言处理和应用程序。

除了上述两个基本组成部分以外，数据库系统当然需要相应的硬设备。一般的要求是：

(1) 大容量的内存：用以存放操作系统、数据库管理系统的例行程序、应用程序、系统缓冲区和用户工作区等；

(2) 具有直接存取功能的大容量外存，目前最常用的外存介质是磁盘；

(3) 中央处理机要有较强的数据处理能力；

(4) 有较高的通道能力，能带多个终端供多个用户共享；

(5) 能够根据不同的用途配置相应的外部设备。如绘图仪、制表机和汉字处理机等。

3、数据库系统的发展过程

数据库技术出现以来经历了三个时期，其发展过程可描述为：

文件管理系统 → 数据管理系统 → 数据库管理系统。

(1) 文件管理系统FHS(*File Handling System*)

出现于六十年代初，一般以顺序式的组织方式建立在磁带上。记载信息的文件在外存介质上的物理结构与用户观念中的逻辑结构完全一致，每个数据文件由相应的一个程序进行控制。图1-1是一个工厂的文件管理系统的部分结

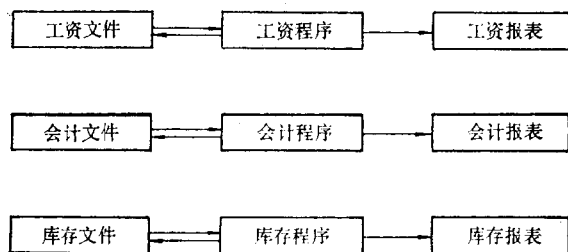


图 1 - 1 文件管理系统结构示意图

构。

如图所示，工资程序处理工资文件而产生工资报表；会计程序处理会计文件而产生会计报表；库存程序处理库存文件而产生库存报表。显然，如果再需要另一类文件的数据报表，就要再编新的程序。在这种只有简单的对应关系的结构中，一个文件只适应于一种应用。因此，同一个数据可能出现在不同的文件里，数据被大量重复存贮，即产生不必要的冗余。另外，各个文件彼此独立，没有反映出数据间的关系，一个文件中的数据不能为多个用户、多种应用服务，即没有实现数据共享。

(2) 数据管理系统 DMS (Data Management System)

六十年代中期出现了直接存取存贮设备 DASD (Direct Access Storage Device) 之后，人们力求通过对数据的集中存贮和控制实现共享。于是就进行了大型集中管理数据文件系统的研制，这项工作六十年代后期取得

成功，这就是数据管理系统，即初期的数据库系统。在这种系统中，文件的物理结构与逻辑结构之间已有所区别，对文件中的数据作了进一步的细部构造，如在文件的物理结构中引入了链接和索引等形式，因而对文件的记录可顺序地和随机地访问，并提供了对文件进行检索、分类、排序及处理用户查询命令等功能。在这种系统里，库内的全部数据由DMS统一管理，而不是由个别程序去处理，任何一个应用系统都能调用所有的数据。仍以前面的工厂管理系统为例，其结构如图1-2所示。

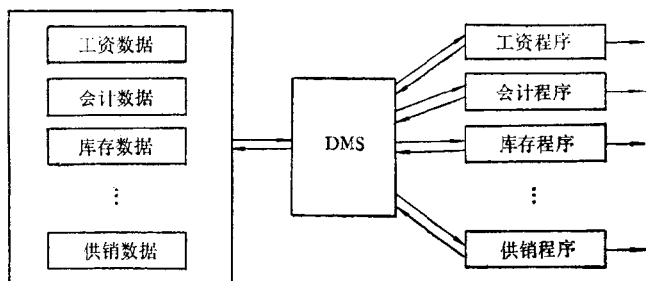


图 1-2 数据管理系统结构示意图

在这种系统中，组成文件的最小单位是记录，但记录之间仅存在顺序关系而无其他联系。另外，在数据管理系统的管理软件中还增加了安全、保密检查措施。

（3）数据库管理系统 DBMS (*Data Base Management System*)

数据库管理系统大约在六十年代末形成，它具有以下几个特点：

①对用户方面的数据进行严格细致的描述，使得文件→记录→数据项等数据单位之间的联系清晰，结构简单；

②允许用户以记录或数据项作单位进行访问，允许以多个约束条件检索和进行文件之间的交叉访问；

③数据的应用独立于数据的存贮。即改动数据的物理位置和存贮结构时不必修改或重写应用程序。

关于这些特点以及它们涉及的一些新的概念将在后面的章节中进一步讨论。

数据库技术仍在不断地发展，当今的数据库往往不是孤立存在的，而是更庞大的信息控制系统的一个组成部分。广义地说，数据库的功能是控制这样一个循环：即从观察客观事物出发，经过对信息的采集、存贮、加工，找出其内在的变化规律，再通过数学模型的分析，得出预测的趋势，从而指导修订对策，再返回来控制客观事物的发展。其过程如图 1-3 所示。

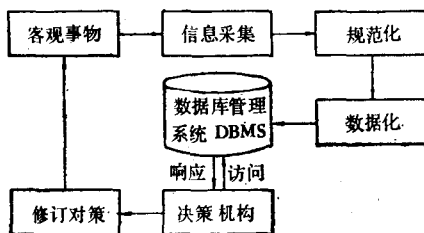


图 1-3 数据库系统的功能

随着数据库技术的发展和渐趋完善，数据库的应用愈来愈广，规模也愈来愈大。它已经从一个厂矿、一个企业的经营扩大到一个地区、一个国家的资源管理；从中、小型计算机的单一系统发展到由大型、巨型机和远程终端组成的信息网络；从一般事务处理发展到空间定位的地理信息管理；从

静态取样发展到动态监控；从一般统计分析发展到多因素的综合评价和决策。近几年来，数据库与数据通讯技术相结合，产生了分布式数据库系统；数据库技术与人工智能研究相结合，产生了知识库、智能数据库等。另外，随着微型计算机技术的飞速发展，有一种明显的趋势：数据库也要向微型化发展。

4、两种不同类型的数据库

从数据库技术在地学领域中的应用这个角度来看，数据库按存贮信息的特征（定性或定位）可以分为两大类，一类是事务管理数据库，另一类是地理信息数据库。

（1）事务管理数据库

事务管理数据库即商用数据库。最初用于商业，后来逐步扩大到交通、邮电、图书、情报、教育、医疗、物资、人事以及企业和国家行政机构的各种事务管理。这类数据库的管理对象大多是人、事、物一类的社会信息，它所包含的各种数据之间的联系是其属性间的抽象的逻辑关系，它们的记载方式主要是文件和数表。因为在内容和形式上有较强的通用性，所以在许多数据库文献书籍中把它视作通用数据库，如果不加专门说明，指的就是这类数据库。

（2）地理信息数据库

地理信息就是与空间位置相关的信息（大多是自然信息），地理信息数据库就是管理地理信息的数据库，有的文献上也把它叫做定位数据库。地理信息数据库的数据之间除了抽象的逻辑关系外，还建立了严谨的空间几何关系（实体的位置、形状和范围等），这是它与事务管理数据库的根本差别，由于这种差别，从数据结构到管理方法，都有其自身

的理论和模式。当然这两种数据库也有许多共同的基础，而且有一部分地理信息可以直接采用事务管理数据库来管理。

目前，国外地理信息数据库已广泛应用于测绘资料的自动管理、专题地图的机助编绘、地质资料的整理分析、资源与环境信息的监控预报和国土经济的综合研究等许多方面。近几年来，国内各有关部门也十分重视这项新技术的发展，正处在从考察、研究试验向工程实践过渡的阶段。

综上所述，事务管理数据库是通用性较强的数据库，对于地学工作者来说，可以把它看作数据库技术入门的基础；地理信息数据库是具有定位特点的数据库，是我们研究的主要对象。为了便于叙述，我们在这本小册子的第一部分以事务管理数据库为主体概略地介绍数据库的基本理论；在第二、三部分中以测绘应用为主体着重讨论地理信息数据库的建库方法。

(二) 数据的逻辑组织

1、数据(Data)

数据是对外界客观事物的信息的一种描述方式，是不带任何理论和观点的原始素材。数据有两种基本类型，即数字型和文字型。此外，还可以有由两者混合组成的字符串。例如，在一个三角点的记录中，“坐标”、“高程”等是数字型数据；“点名”、“觐标类型”等是文字型数据（可以是英文、中文或其他文字）；“所在图幅的国际分幅编号”则是数字与文字混合组成的字符串。

所有数据都包含下列三个内容：

(1) 主题：即研究对象，是一个客观存在的事物，数

数据库术语中称之为实体 (Entity) ;

(2) 特征: 即对实体的描述, 数据库术语中称之为属性 (Attribute)。显然, 一个实体可有多个属性;

(3) 值 (Occurrence): 即每个属性的具体内容或数量。

在上面的例子中, 一个三角点为一个实体, 点名、高程、觐标类型及所在图幅编号等都是三角点这个实体的多个属性, 而“红石山”、“1569. 3”、“寻常标”及“K-48-103-A”等就是这些属性的值。

2、文件、记录和数据项 (File, Record and Item)

在数据库术语中, 把描述实体的数据称为记录, 而把描述属性的数据称为数据项。由于一个实体包含若干个属性, 因此, 一个记录由若干个数据项组成。在前面的例子中, 对一个三角点的描述就是一个记录, 其中“点名”、“点号”、“坐标”、“高程”等都是该记录所包含的数据项。

数据项还可细分为初等项和组合项。初等项是可命名数据的最小单位, 如“点号”、“点名”等; 组合项是由若干初等项或其他组合项组成的数据单位, 如“大地纬度”中可包括“度、分、秒”等。

记录和数据项都有型与值之分。型是框架, 值是内容。例如“年龄”的型可以这样描述: “名称为年龄, 数据特征是不大于三位十进制的正整数”, 而它们的值为“1岁, 2岁……”。值一般应当有定义范围, 称为值域。例如年龄的值域为 [1, 200], 超过200岁, 就不符合定义了, 在操作中可以看作是一种出错信息。

文件是记录的集合。一个文件可以包括若干个记录，一个记录可以包括若干个数据项，三者的基本关系如图1-4所示。

	数据项 1	数据项 2		数据项 n	
文 件	红石山	1569.3		K-48-103-A	记录 1
	李 村	1831.7		K-49-092-B	记录 2
	⋮	⋮		⋮	⋮
	白马岭	1918.2		K-50-037-C	记录 m

图 1-4 文件的逻辑结构

数据项、记录、文件都是数据的重要单位，统称为逻辑数据。当把它们贮存到计算机的存贮介质上时，则称为物理数据。

3、关键字 (Key)

在记录中，一般都把第一个数据项设计为实体的编号。例如，学生的学号、职工的工作证号、三角点的点号等。用户提取数据时并不限于以编号提取。有些数据项的值也可用来确定所指的记录，但不是所有的项，例如，指出某个学生的学号值“84001”，就确定了所指的记录是“84001，王平，21岁，男，江苏……”。但若给出的是另一个数据项如性别值“男”，则无法确定所指的记录是哪一个。因此，在设计数据库时，通常把其值能标识记录的一个或多个数据项定义为该文件的关键字。其中，又把用于组织文件的，能够唯一标识记录的一个关键字定义为主关键字，而把其余的关键