



跟我学电脑

王永民

丛书

五笔字型发明人 王永民教授 主编



数据库基础

任宪坤

编著

北京·气象出版社

以平知

PDG

难得一套电脑科普书

前天,纽约时报公布了本周内纽约州畅销书的排行榜。名列榜首的书,是一本理论物理学的科普读物《时间简史》(A Brief History of Time),作者斯蒂芬·霍金(Stephen Hawking),被誉为自爱因斯坦以来当代最伟大的天才理论物理学家。他以残废之身在轮椅上研究著述了 20 多个年头。评论文章称,他的这本书是在世界上引起轰动、在纽约连续 100 个星期销量排名第一的书,发行已超过 100 万册。

我立即到书店花 16 美元买了一本,一口气翻完了 180 页正文。啊!这真是一本我从未见过的令人不忍掩卷的科普书。作者把高深的理论,诸如什么是时间,时间有无头尾,什么是宇宙和黑洞,什么是相对论等等,讲得通俗易懂,趣味盎然!

一本高深理论物理学的科普书居然会如此畅销,的确是发人深省的。

也许,科普书的难点正在于写“深”容易,写“浅”反而难!不是真正精于一门的饱学之士,不是真正了解读者心理的大手笔,便很难写出好的科普书。正所谓“明白不明白的人为什么不明白,才算真明白”。

然而电脑,实在不是一般人容易弄明白的洋机器。

继西方世界全面实现电脑化之后,电脑用于机关,电脑走向民间,在国内已蔚然成风。要让国人明白电脑是怎么一回事,要让普通人学会操作电脑,除了开展正规教育之外,我以为最重要的,恐怕就是编写一套通俗易懂、趣味盎然的自学丛书,满足为数更多的自学者的要求。

事实上,电脑并不高深莫测。不少人对电脑望洋兴叹,常常是因为那些厚厚的叫人眼花缭乱而又枯燥无味的操作手册、用户指南使人望而生畏,不敢问津。

现在,这种情况可望有所改善。我看了中国气象出版社出版的《跟我学电脑》丛书一套 11 册的初稿,觉得这套丛书具有以下特色:

一、**实用性强**。书中介绍的都是最基本的电脑知识,着重于实际应用和操作方法,看了就明白,明白了就能用。

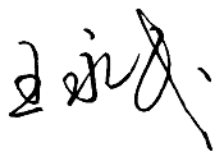
二、**图文并茂**。书中附有大量的电脑屏幕图,以图解文,直观教学,形象生动,另配有许多漫画,可使读者迅速领会,印象深刻。

三、**浅显易懂**。丛书为初学者编写,尽量避免抽象概念,自学者不必死记硬背,只管照章操作,即可熟练掌握,无师自通。

这真是一套难得的电脑科普书。对国内读者来说,可谓是雪中送炭。

而且,这是一套具有《时间简史》一书特色的好书!

我相信,这套丛书也会像《时间简史》在美国受到欢迎一样,在中国乃至国外华人界受到欢迎。特此向中国气象出版社表示祝贺和感谢,是为序。

A handwritten signature in black ink, appearing to be '王永' (Wang Yong), written in a cursive style.

1997 年 5 月 6 日于纽约 Flushing

引 言

近年来,随着微型计算机技术的迅速发展,计算机的应用范围日益扩大,特别是数据库技术的发展,使得计算机在数据管理方面越来越发挥出重要的作用。目前,国内流行的数据库系统大多是关系型数据库系统。

对于非专业用户特别是初学者来讲,一个非常关心的问题就是:如何尽快掌握数据库技术,并利用它进行数据管理工作,开发管理系统软件?这正是编写本书的出发点。

本书特点:

1. 浅显易懂。本书通过介绍 FoxBASE+系统的功能和使用方法,来帮助读者学习数据库技术。FoxBASE+的特点就是简单易学,使用方便,并且与在国内已广为流行的 dBASE III 系统完全兼容,非常适合初学者。

2. 实战性强,由实战例子引出概念和结论,并在实际操练中进一步理解和掌握。

本书结构:

全书共分为五个部分。

第一部分——概念篇(第一章到第二章),第一章介绍了数据库系统的基本概念,第二章介绍了 FoxBASE+的基本概念。

第二部分——操作篇(第三章到第六章),介绍了 FoxBASE+的各种操作,并以一个具体的数据库为线索,列举了大量的应用实例。

第三部分——设计篇(第七章到第十章),第七章到第九章介绍了 FoxBASE+程序的设计方法,包括输入输出设计和菜单设计,第十章介绍了两个实用程序。

第四部分——应用篇(第十一章),介绍了应用软件的开发过程,并通过一个人事管理系统软件的研制综合了使用 FoxBASE+开发应用系统的技术。

第五部分——系统篇(第十二章),主要介绍了目前较为流行的几种关系型数据库系统(如 dBASE III、FoxBASE+、FoxPro、INFORMIX、INGRES、ORACLE 和 Sybase 等)的组成、特点和功能等,这将有利于选择合适的数据库管理系统去开发应用系统。

本书阅读方法:

本书内容具有向前封闭性,前面的章节独立于后面的章节,后面的章节依赖于前面的章节。因而本书不适宜跳跃性阅读,请您循序渐进,由简到繁。

目 录

第一部分 概念篇

第1章 数据库系统的基本概念	(3)
1.1 数据库技术的发展	(4)
1.2 信息、数据和数据处理	(6)
1.3 数据库系统	(7)
1.4 数据描述与数据库分类	(9)
第2章 FoxBASE+基本概念	(13)
2.1 FoxBASE+数据类型	(14)
2.2 FoxBASE+的常量和变量	(14)
2.3 运算符和表达式	(15)
2.4 函数与命令	(18)
2.5 赋值命令与显示命令	(20)
2.6 文件	(22)
2.7 FoxBASE+的运行环境	(24)

第二部分 操作篇

第3章 FoxBASE+的基本操作	(27)
3.1 开始建立一个数据库	(28)
3.2 数据库文件的打开与关闭	(35)
3.3 记录的显示与定位	(36)
3.4 记录的插入与删除	(39)
3.5 记录的修改	(43)
3.6 数据库结构的显示与修改	(46)
3.7 数据检索与组织	(47)
3.8 数据统计	(55)
第4章 FoxBASE+函数	(59)
4.1 字符串操作函数	(60)
4.2 数学运算符号	(65)
4.3 日期和时间函数	(68)
4.4 转换函数	(70)
4.5 测试函数	(73)
4.6 标识函数	(80)
4.7 环境函数	(82)
4.8 输入函数	(86)
第5章 多重数据库文件操作	(89)

5.1	选择工作区	(90)
5.2	数据库文件的关联	(91)
5.3	数据库文件的连接	(94)
5.4	数据库文件的更新	(95)
5.5	数据库文件的复制	(96)
第6章	系统其它操作和命令	(99)
6.1	系统环境设置命令	(100)
6.2	磁盘文件操作命令	(107)
6.3	HELP 命令	(108)
6.4	屏幕画面保存与恢复	(110)
6.5	清除和关闭命令	(111)

第三部分 设计篇

第7章	应用程序设计	(115)
7.1	应用程序的建立、修改和执行	(116)
7.2	程序交互性命令	(117)
7.3	程序结构设计	(119)
7.4	自定义函数	(134)
7.5	内存变量	(136)
7.6	数组	(142)
7.7	程序注释	(149)
第8章	输入输出设计	(151)
8.1	报表输出	(152)
8.2	标签输出	(157)
8.3	格式设计命令@	(160)
8.4	利用@命令设计屏幕格式	(167)
第9章	菜单设计	(171)
9.1	POP-UP(上弹式)菜单	(172)
9.2	PULL-DOWN(下拉式)菜单	(173)
9.3	亮带式菜单	(176)
第10章	实用程序	(179)
10.1	编译器 FoxPCOMP	(180)
10.2	过程文件的组合	(181)

第四部分 应用篇

第11章	应用实例	(187)
11.1	应用软件的开发过程	(188)
11.2	人事管理系统	(194)

第五部分 系统篇

第12章	目前常见的关系型数据库管理系统	(211)
-------------	------------------------------	--------------

12.1	dBASE III 数据库系统	(212)
12.2	FoxBASE+ 数据库系统	(215)
12.3	FoxPro 数据库系统	(217)
12.4	INFORMIX 数据库系统	(219)
12.5	INGRES 数据库系统	(222)
12.6	ORACLE 数据库系统	(225)
12.7	Sybase 数据库系统	(229)
附录 A	FoxBASE+ 全屏幕编辑控制键	(233)
附录 B	FoxBASE+ 命令一览	(235)
附录 C	FoxBASE+ 函数一览	(246)
附录 D	FoxBASE+ 错误信息	(252)
附录 E	ASC II 码字符集	(265)

第一部分

概念篇

第1章

数据库系统的基本概念



- 数据库技术的发展
- 信息、数据和数据处理
- 数据库系统
- 数据描述与数据库分类

本章主要介绍数据库的一些基本概念,如数据、数据描述、数据管理和数据库系统等。熟悉这些概念,将有助于掌握数据库技术,充分发挥所运用的数据库管理系统(如 FoxBASE+)功能。

1.1

数据库技术的发展

自 1946 年第一台电子计算机问世以来,在不到 50 年的时间内,计算机系统得到飞速发展,同时为数据处理进入全自动化的电子数据处理提供了有利条件。最初计算机主要用于解决科学或工程计算问题。50 年代计算机在数据处理技术上的突破,使计算机的应用范围从单纯的数值计算扩展到解决数据处理等非数值计算问题,如辅助教学、人事管理、库存管理、工资管理等。

随着科学技术的不断发展,计算机愈来愈显示出它在信息处理方面的优越性。它处理数据速度快,能为生产过程提供实时控制,为各种生产管理部门处理大量即时信息;存储容量大,能长期保存大量有价值信息,这是以往任何工具很难做到的;输入输出灵活,而且将人工干预操作减少到最低程度。计算机在国民经济各个领域的广泛应用,已经取得了巨大成就,充分显示了它的广阔的发展和应用前景。

数据管理技术的发展大体上经历了三个阶段:

1. 人工管理阶段

50 年代中期以前,数据管理处于人工管理阶段。在这一阶段,计算机主要应用于科学计算。计算机硬件功能弱,输入输出设备简单,主要是纸带、卡片和磁带。软件方面没有操作系统及进行数据管理的系统软件。数据管理基本上是手工进行,不需要长期存储数据。

这一阶段数据管理的基本特点是数据不独立,它由应用程序管理。用户在进行数据处理时,除了考虑应用程序、数据的逻辑定义和组织外,还要物理结构。应用程序和数据相互结合成一个整体,互相依赖,如图 1-1。

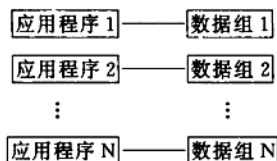


图 1-1 人工管理阶段的数据与程序

数据面向特定应用,一组数据对应于一个程序。数据之间、程序之间相互独立,因此数据不能共享。

在这种管理方式下,程序设计员在应用程序中包含数据结构(包括存储结构,存取方法和输入输出格式)。如果数据结构和外存设备发生变化,应用程序就必须随之被修改。

2. 文件系统阶段

50 年代后期到 60 年代中期,数据管理处于文件系统阶段。这一阶段,计算机不仅用于科

学计算,而且大量用于数据管理。

硬件方面,先是使用磁带作为输入输出装置或主存储装置的辅助存储装置;然后又出现了磁盘、磁鼓等直接存取的存储设备,存储能力大大增强。

软件方面,出现了操作系统,而且有了专门管理数据的软件——文件系统,具有了文件管理和一定的数据管理功能。

这一阶段数据管理的基本特点是数据脱离应用程序,不再由其管理,而改由专门的软件进行管理。数据是有组织、有结构地构成文件形式,由文件系统完成数据的逻辑结构与物理结构的映射,将文件自动存放在磁带、磁盘上。文件管理系统是数据文件和应用程序之间的接口,应用程序必须通过文件管理系统才能建立和存储文件以及检索数据文件中的数据,如图 1-2。

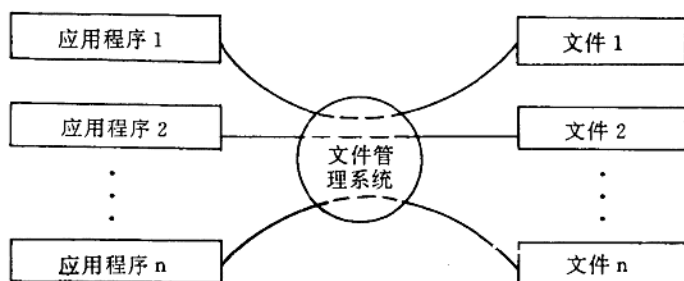


图 1-2 文件系统阶段的数据与程序

数据依然是面向特定应用的。在数据管理上,虽然需要通过一个文件管理系统对数据文件进行统一管理,但文件基本上还是对应于一个或几个特定的应用程序。

数据共享性差,冗余大。数据只能在文件级共享。当不同的应用程序所需数据存在相同之处时,仍要建立各自的数据文件,而无法实现数据共享。

3. 数据库系统阶段

从 60 年代后期开始,数据管理进入数据库系统阶段。在这一阶段,随着社会和生产的高速发展,数据处理的规模越来越大,需要处理的数据量急剧增长,数据共享的要求也日益迫切,而文件系统已远远不能满足这些要求。另一方面,计算机在软、硬件方面也有了很大发展。硬件方面,出现大容量、高速度硬盘,可靠性增强,而且硬件价格大幅度下降。软件方面,设计出了面向数据管理的数据库管理系统。

这一阶段数据管理的基本特点是数据由数据库管理系统 DBMS(Data Base Management System)管理,具有完整的数据结构,数据之间存在联系,避免了数据重复。同时,数据完全从应用程序中独立出来,如图 1-3。

数据面向系统或数据本身,而不再面向特定的一个或几个应用,因而共享性好。另外,由于计算机的共享性是并发的,即许多用户同时使用数据库。为了有效地实现数据共享,系统还提供了对数据的多种维护能力,如安全性,完整性和并发控制等。

(1) 安全性

安全性是指保护数据不被破坏,不被无关人员接触或窃取。一般采用规定用户对数据的操作权限,用密码编码数据等方法进行保护。

(2) 完整性

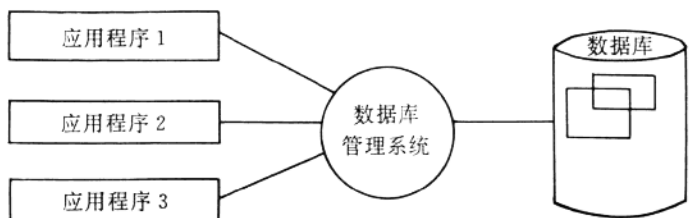


图 1-3 数据库系统阶段的数据与程序

完整性是指数据的正确性、一致性,保证数据不出现与实际不符的情况。一般采用数据检验来保证数据的完整性。检验即对数据库中数据的值做某种限制性的检查,监视数据是否在合理范围内。

(3) 并发控制

并发控制是指避免当多个用户同时修改、存取数据时,由于互相干扰而产生数据不一致的现象。一般采用加锁和发通知的办法来实现控制。

总的来说,数据库是存储在一起的相互关联的综合性的数据集合,这些数据没有不必要的冗余,可供各种用户共享;数据独立性和结构化很强;对数据库进行插入、修改或检索都能通过 DBMS 按一种公用的和可控制的方法进行。同时 DBMS 可保证数据的安全性、完整性和并发控制。这些特点为用户正确、便利地处理数据提供了基础。

1.2

信息、数据和数据处理

在日常生活、工作中,人们会碰到大量的、各种各样的数据。数据被视为人类社会极为重要的资源。如股市行情数据、企业管理数据、商店销售情况数据等等。那么,什么是数据呢?一般来说,任何一种用来反映客观事物的性质、属性以及相互关系的字符、文字、数字和图形都可以称为数据。这些数据之所以重要,是因为它们表示出信息的内容,是对事物进行的描述。换言之,数据是用以载荷信息的字母、数字和符号,是表达和传递信息的工具。例如“某一工程的总投资为 3000 万人民币”,这“3000 万”是一个数据,同时它又表示这一工程项目的总投资额这一信息。可见,信息是可以用数据来表示的,二者不可分离又有一定区别。简单地说,信息是在人们头脑中对现实世界事物的抽象反映,它依赖于数据而存在,是数据加工的结果。

自从原始人类的结绳记数开始,就揭开了人类社会的数据处理的序幕。所谓数据处理,是指对各种数据进行收集、加工、存储和传播等活动的总和。

数据处理的主要目的是把数据转换成便于观察、分析及进一步处理的形式;把数据加工成对正确决策有用的数据;借助计算机将有价值数据存储起来,以便人们充分利用这些宝贵的资源。

数据处理的过程一般分为:数据的收集、转换、组织、输入、处理、输出、存储和反馈等几个阶段。

1.3

数据库系统

一、数据库系统的构成

数据库系统是为了更好地组织和存取大量数据而将数据库引入计算机后形成的管理系统。它由数据库、数据库管理系统、计算机系统和有关人员组成。引入数据库之后的计算机系统软、硬件层次如图 1-4。



图 1-4 引入数据库后计算机系统的软硬层次

(1)数据库

以一定的组织方式存储在一起的相关数据的集合,供多用户共享。

(2)数据库管理系统 DBMS

DBMS 是联系用户与数据库的纽带,它管理和维护数据库中的数据,维持数据库系统的正常活动。使用了 DBMS,用户则可以不必注重于数据的逻辑和物理细节,而只需注重通过数据所表达的信息内容。

(3)计算机系统

计算机系统包括用来管理数据库的计算机硬件和软件。硬件指的是 CPU、存储设备和输入输出设备;软件指的是操作系统、数据库管理系统 DBMS、高级语言如 COBAL、FORTRAN、C 及其编译系统、应用开发程序。

(4)有关人员

有关人员指的是管理、开发和使用数据库的人员,包括:

数据库管理员 DBA

DBA 是系统工作人员,负责对整个数据库的管理工作,保证数据库系统的正常运行。

应用程序员

他们是专业用户,负责应用程序的编写和维护工作。

终端用户

这类用户多是非计算机专业人员。他们通过联机终端设备,使用由基本命令组成的询问语

言对数据库进行检索,删除和插入等操作。

二、数据库系统的结构

数据库系统有一个严谨的体系结构,它分为三级:外模式、概念模式和内模式,同时也是定义了数据库的三个层次,反映了看待数据库的三个不同角度。如图 1-5。

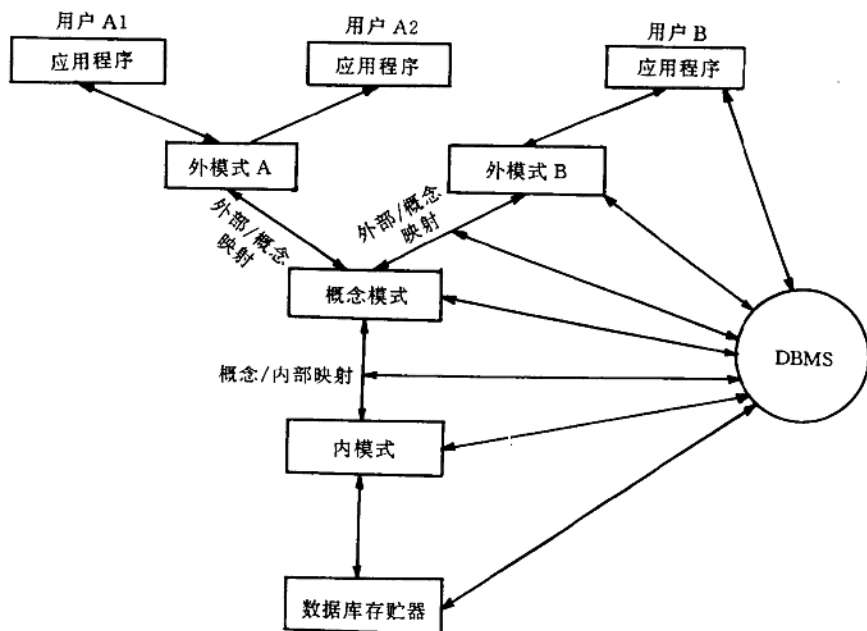


图 1-5 数据库系统的结构

(1) 外模式

外模式是数据库与用户的接口,是应用程序可见到的数据描述。

(2) 概念模式

概念模式是数据库的整个逻辑描述,它说明一个数据库所采用的数据模型。

(3) 内模式

内模式描述了数据的物理结构、数据在存储器上的实际安排与存取方式。

数据库的三级结构是靠两个映象来联结的。在外模式与概念模式之间和概念模式与内模式之间各有一个映象,实现双方的对应和转换。数据库系统的三级模式和两级映象为实现数据共享和增强数据的独立性提供了可能。

一个数据库系统可以有多个外模式,但只能有一个概念模式和一个内模式。

1.4

数据描述与数据库分类

一、现实世界的描述

在信息系统中,信息从客观事物出发,流经数据库,通过控制决策机构(由人或自动控制系统组成),最后回到客观事物。信息的这一循环经历了三个不同领域:现实世界、信息世界和数据世界(计算机世界)。

1. 现实世界

现实世界是存在于人们头脑之外的客观世界,事物及其相互联系就处在这个世界之中。一个客观存在并且可以识别的事物称为个体。个体可以是某一具体事物,如一名职工,一本书等,也可以是某一个抽象事物,如一次会议,一次订货等。

个体都有自己的特征,用以区别其他个体,例如职工有姓名、年龄、性别、职称、工资等不同特征来标识自己。但在我们具体研究某一类个体时,往往只选择其中与研究问题有意义的特征。例如在描述一个人健康情况时,可以选择身高、体重和血压等特征,而在说明订购情况时,则可以选择货名、订购数量和订购单位等特征。我们把具有相同特征要求的个体称为同类个体,所有同类个体的集合称为总体。例如所有的“职工”就是一个总体。

现实世界中的客观事物是信息的源泉,是设计数据库的起点。

2. 信息世界

现实世界的事物反映到人的头脑中来,人的大脑对于这些事物有个认识过程,经过选择、加工,将有意义的对象进行命名、分类后形成信息,即进入信息世界。在信息世界中每一个被认识了个体称为实体,它是信息世界的基本单位。反映实体的特征称为属性。例如一本书就是一个实体,可以通过书号、书名、作者、出版社和单价等属性来表示。

我们把具有相同属性的实体称为同类实体,同类实体的集合为实体集。在信息世界里,主要研究的不是个别实体,而是它们的共性,即同类实体和实体集。

3. 数据世界(计算机世界)

数据世界是信息世界中信息的数字化。信息是用数字、字符、符号、图像或声音表示的。如职工的姓名和年龄,表示路标的符号,红绿灯等。但是,这些信息在计算机世界中只能用二进制表示,即一切信息在进入计算机时必须数字化。在数据世界中,记录是实体的数据表示;数据项是实体属性的数据表示;文件是同类记录的集合。

现实世界中,个体与个体之间总是存在某些联系。这些联系,反映到世界中,就是实体与实体之间的联系,即构成实体模型;反映到信息数据世界中,就是记录与记录之间的联系,即构成数据模型;实体模型的数据化表示,就是数据模型。

现实世界、信息世界和计算机世界是由客观到认识、由认识到使用管理的三个不同层次,后一领域是前一领域的抽象描述。这三个不同层次之间的术语对应关系如图 1-6 所示。

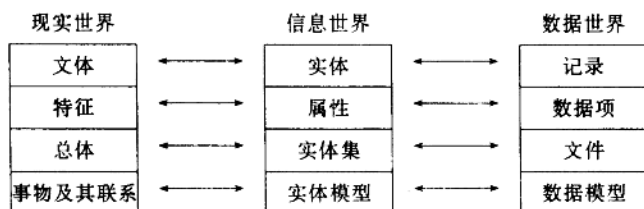


图 1-6 三个世界的术语对应关系

二、实体模型

实体模型表示信息世界中实体及实体间的联系。实体之间的关系可归结为三类：

1—1 关系

如果两个实体集 A 和 B 的每一个实体至多和另一实体集中的一个实体有联系，则称 A 和 B 是 1—1 关系。这是实体间最简单的联系，如学校与校长，工厂和厂长之间的关系。

1—m 关系

现有两个实体集 A 和 B，如果 B 中每个实体和 A 中任意个实体有关，而 A 中每个实体至多和 B 中一个实体有关，则称 A 和 B 是 1—m 关系。1—m 关系是比较普遍的关系，如厂长和职工，班长和同学之间的关系。

m—m 关系

如果两个实体集 A、B 中的每个实体都和另一个实体集中任意个实体有关，则称 A 和 B 是 m—m 关系，即“多对多关系”。m—m 关系是更一般的实体间的关系，例如图书与借书人，师生关系等都是多对多关系。

三、数据模型与数据库分类

数据库系统的核心就是数据模型。数据模型是表示实体与实体之间联系的数据库的数据结构，是数据库中的数据全局逻辑结构的描述。

目前流行的数据模型有层次模型，网状模型和关系模型等。层次模型和网状模型又称为格式化模型。这两种模型使用有向图的概念来构造数据模型。图的结点表示实体集合，方向弧表示实体间的联系，关系模型是以数学理论为基础构造的。它把数据模型看成是关系的集合，不同的数据模型将决定其不同的数据库操作语言的结构。

层次模型

层次模型又称为树状模型，它是以记录类型为结点的有根的定向树或森林。在树中，把没有双亲的记录称为根记录，其它记录称为从属记录。从属记录中，只有父记录而无子女记录的称为叶记录。从根记录开始，一直到最低下一层叶记录为止，所具有的层次称为该数据模型的层次。同一层上的记录之间没有联系。如图 1-7。