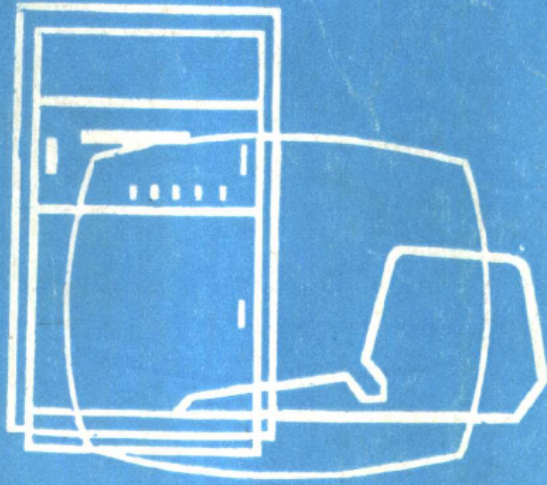


数据库基础及 dBASE III

编者 牛允鹏 宿 华 朱立功



中国计算机函授学院

数据库基础及 dBASE-III

牛允鹏 宿 华 朱立功 编

中国计算机函授学院

1988·合肥

前 言

数据库技术是一门理论性和实践性都很强的课程，针对这种特点，本书首先说明什么是数据库后立即转入dBASE—Ⅲ的学习，深入地讨论了如何在微型计算机环境下使用dBASE—Ⅲ进行应用课题的开发。之后，又比较系统地讨论了数据库的基本概念，详细地阐述了数据库的实现技术，在此基础上介绍了两个有代表性的数据库管理系统IMS和DBTG。

编者力图使本书具有实用性和理论性两方面特点，适宜于管理类专业或函授教学用书。书中附有大量实例和适当练习，便于读者自学。

在编写过程中始终得到了钱洲胜、张宁两同志的大力支持和协作，罗国军同志调试了全部程序，在此深表谢意。

编者水平有限，书中错误之处请批评指正。

编者

1988. 3

目 录

第一章 数据库的初步认识

§ 1—1 数据库系统研究的对象	(1)
§ 1—2 什么是数据库	(3)
§ 1—3 构成数据库系统的三要素	(9)
§ 1—4 数据库种类	(10)
§ 1—5 数据库方式的优缺点	(11)
习题	(12)

第二章 关系式数据库dBASE—Ⅱ概述

§ 2—1 dBASE—Ⅱ的特点和性能	(13)
§ 2—2 dBASE—Ⅱ的常数、变量、表达式、函数、文件	(17)
§ 2—3 dBASE—Ⅱ的命令结构和分类	(31)
§ 2—4 dBASE—Ⅱ的的使用	(33)
§ 2—5 一个数据库操作的简单例子	(43)
习题	(50)

第三章 程序控制语言

§ 3—1 命令文件的建立和执行	(51)
§ 3—2 内存变量的赋值、显示、清除	(52)
§ 3—3 数据输出	(56)
§ 3—4 交互式的数据输入	(59)
§ 3—5 格式输入输出指令	(67)
§ 3—6 停止及注释命令	(71)
§ 3—7 框图介绍和简单程序设计	(71)
§ 3—8 分支	(74)
§ 3—9 循环	(82)
§ 3—10 过程及其调用	(94)
§ 3—11 宏代换	(111)
§ 3—12 打印报表	(115)
§ 3—13 程序(命令文件)设计要点	(127)
习题	(141)

第四章 文件的建立

§ 4—1	数据库及备注文件的建立	(143)
§ 4—2	索引文件的建立及打开	(149)
§ 4—3	命令文件的建立、修改和运行	(154)
§ 4—4	内存文件的建立和调用	(155)
§ 4—5	文本输出文件的建立	(178)
§ 4—6	格式文件的建立及打开	(160)
§ 4—7	报表格式文件的建立与打印	(162)
§ 4—8	标签文件的建立与输出	(170)
	习题	(177)

第五章 数据库操作、维护、组织

§ 5—1	记录指针定位	(178)
§ 5—2	查询命令	(184)
§ 5—3	信息显示	(192)
§ 5—4	数据库的修改	(200)
§ 5—5	记录的添加	(216)
§ 5—6	记录的删除	(220)
§ 5—7	数据库文件的复制	(229)
§ 5—8	从其它文件添加数据	(236)
§ 5—9	数据库组织	(241)
§ 5—10	统计汇总	(247)
§ 5—11	多工作区操作	(252)
	习题	(273)

第六章 函数及系统状态参数设置

§ 6—1	函数	(275)
§ 6—2	dBASE—Ⅱ系统状态设置	(286)

第七章 dBASE—Ⅱ与高级语言间的通讯

§ 7—1	dBASE Ⅱ与高级语言实现数据共享的原则	(295)
§ 7—2	dBASE Ⅱ程序与高级语言程序的连接	(298)
§ 7—3	dBASE Ⅱ与BASIC通讯	(299)
§ 7—4	dBASE ⅡFORTRAN通讯	(304)

第八章 编程技巧

§ 8—1	菜单程序	(308)
§ 8—2	应用系统的保密	(311)

§ 8—3	制作硬盘备份	(313)
§ 8—4	打印区位码表	(316)
§ 8—5	加快程序执行速度	(317)
§ 8—6	避免数据丢失	(319)
§ 8—7	输入代码输出汉字方法	(321)
§ 8—8	库结构操作	(322)
§ 8—9	定义和显示功能键	(323)
§ 8—10	不打印数值 0	(324)
§ 8—11	怎样设计通用数据处理系统	(325)
§ 8—12	验证数据库文件是否存在	(331)
§ 8—13	分析数据库结构	(332)
§ 8—14	对数据库字段名进行编号	(333)
§ 8—15	纵向方式向数据库中输入数据	(336)
§ 8—16	再谈通用制表程序	(338)

第九章 程序实例

§ 9—1	程序设计的步骤和方法	(341)
§ 9—2	图书出借业务管理系统	(343)

第十章 数据库系统的组成

§ 10—1	对数据库的不同观点	(356)
§ 10—2	数据描述语言 DDL	(357)
§ 10—3	数据操纵语言	(358)
§ 10—4	DBMS的组成和作用	(359)
§ 10—5	数据库保护	(361)
§ 10—6	DBMS工作过程	(364)
§ 10—7	数据库管理员 DBA	(365)
	习题	(366)

第十一章 数据模型

§ 11—1	实体及属性	(367)
§ 11—2	信息模型及其图示	(370)
§ 11—3	数据模型	(374)
§ 11—4	层次式数据库模型	(378)
§ 11—5	网状模型	(380)
§ 11—6	关系式数据模型	(382)
§ 11—7	层次网状和关系模型之间的互相转换	(385)
	习题	(385)

第十二章 数据库的物理组织与技术

§ 12—1 引言	(387)
§ 12—2 外存	(388)
§ 12—3 文件组织方法	(393)
§ 12—4 数据库的物理组织	(418)
习题	(423)

第十三章 关系数据库基本理论

§ 13—1 关系	(424)
§ 13—2 关系模型	(426)
§ 13—3 关系模型的物理组织	(428)
§ 13—4 关系数据库语言概述	(428)
§ 13—5 关系代数	(429)
§ 13—6 关系演算	(435)
§ 13—7 关系代数表达式转换为dBASE III 程序	(439)
§ 13—8 关系模型规范化	(448)
习题	(456)

第十四章 IMS和DBTG

§ 14—1 IMS	(459)
§ 14—2 DBTG	(475)

第十五章 数据库设计

§ 15—1 数据库设计过程	(489)
§ 15—2 概念模型设计	(490)
§ 15—3 数据库逻辑设计	(493)
§ 15—4 数据库物理设计	(494)
§ 15—5 数据库的运行和维护	(495)
§ 15—6 数据库设计是一项软件工程	(498)

附录

dBASE III 命令表	(500)
---------------------	---------

第一章 数据库的初步认识

在数据处理领域中，数据库技术变得越来越重要，已经成为当代计算机系统设计及计算机应用中最为关心的课题之一。本章作为绪论，为读者建立关于数据库的初步认识。

§ 1—1、数据库系统研究的对象

数据库技术是一门在计算机环境下研究如何高效地进行数据处理的综合性软件技术。数据是数据处理业务中的基本元素，它蕴含着人对现实世界事物的理解和抽象，是信息的一种量化表示。

一、信息

信息是一个基本词汇，所谓“基本词汇”，意思是说不大容易找到它的同义词，这就是为什么人们频繁地使用它却又难以准确地定义的缘故。

通常认为信息是关于某一事件的情报。例如：

“上海开往北京的22次快车晚点1小时”；

“2月15日，北京下了一场大雪”

但实际上信息含义更广，可以把它认为是一种固有的知识，这种知识可通过人的感官直接或间接地获得，例如：

“水是由氢和氧化合而成”。

可见信息是人对客观事物产生的印象并经过推理升华为知识，而知识的反作用将会能动的改造世界，所以信息是资源。

信息是客观存在的，大千世界信息是无穷的，我们不可能也没有必要都去认识它们，而只需要关心那些我们感兴趣的用得到的信息。银行家关心金融方面的信息，旅客关心车次航班信息；数据库设计关心需求方面的信息。

二、数据

电子计算机只能存放数据，为了使计算机能进行数据处理，就得把信息转换为数据。一旦实现这种转换，那么信息处理便等价于数据处理，因此人们通常并不严格区分它们。

透过现象看本质：

数据是表达信息的一种符号。数据反映信息，信息依靠数据表达。

当然，表达信息的符号应该遵从一定的规约，根据约定，现实中的一切信息：文字、声音、图象…等等都可以表示成数据并被计算机加以存贮，这意味着数据不仅指那些具有量值大小的纯数值，更为一般的是能对现实中的一切信息的特征进行某种描述的数值。

信息的多样性决定了数据的多样化，此称为数据类型的不同。在数据库里，数据类型是个重要概念，进行数据库的设计，必须定义数据类型，不同类型数据表达不同信息，不可混淆。

三、数据处理

计算机所以能进行数据处理，是基于如下认识：

电子计算机是一个信息处理机，它接受各种不同形式的输入，经过软件或程序的加工，转换为不同形式的输出。

这说明使用计算机的基本问题是如何组织数据，如何存取及加工利用。不可输入后原封不动输出，若是这样就把计算机当成打印机用了，要充分利用信息资源，也就是必须有一个加工利用过程，这种加工处理归结为两类问题：

数值计算问题

非数值计算问题

前者，指科学计算或工程设计中方程求解，这类问题的特点是计算量大且算法复杂，但是数据类型和数据结构简单，因此这时矛盾的焦点在于计算方法的研究；后者，数据量大，数据类型复杂且数据之间存在着一定的逻辑关系，但计算工作量相对较少，因此这时矛盾的焦点在于合理的组织存贮大量的数据，有效的调度管理这些数据，在于数据管理软件的研究。

所谓数据处理，就是属于这样一种非数值计算问题。这类问题量大面广，例如一批实验数据的分析整理，银行业务，交通调度，工矿企业的生产管理，机关公文报表，物资统计，情报检索乃至图象处理，文字翻译，人工智能等等都是数据处理的典型实例。

下面绘出数据处理的定义：

数据处理泛指数据的收集、存贮、加工、和传播等一系列操作，目的是从大量的杂乱无章的数据中，进行分析推导，提取对人们有价值的成份作为决策的依据。

在数据处理业务中，如何对数据进行有效地管理是个中心问题，通常包括：

- ①数据的组织和存贮
- ②数据的查询检索
- ③数据的维护（指数据更新、增删）
- ④数据的分类、统计汇总
- ⑤制表或其它形式的输出
- ⑥其它

作为例子，让我们考察一个汽车运输企业的经营活动。汽运企业的机构体制，大体可用图1—1所示的网络表示各职能部门之间的业务联系（除人事工会等部门）

图中，部门以方框表示，部门之间以线相连形成一张网状结构图，今后我们会谈到如何以“图”反映信息及其联系，现在暂且留作后话。

由图可见，就整个企业而言，问题相当复杂，因为信息互相关连；货运、客运、机务、财务、车队、仓库等诸方面信息互相牵扯，应该整体范围系统地考虑数据的组织和存贮，而这正是本课程的目的。若就某个局部环节，如单就货运管理而忽略其它业务，

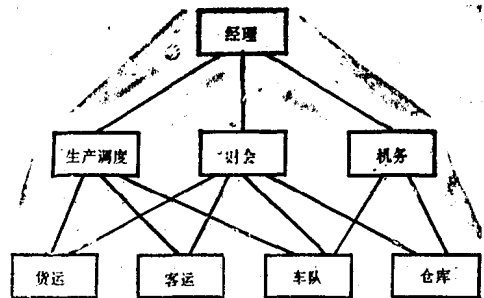


图1—1 汽运企业各职能部门间的信息联系

则问题比较简单。比如，货运部门某年某月承接客户委托，托运货物自甲地发往乙地，一段时间后，便会形成表1—1所示的货运登记表。

表1—1 货运登记表

日期	单位	品名	数差	发地	目的地
1986.3.20	木材加工厂	木材	70	黄山	合肥
1986.3.20	机床厂	钢管	40.5	马鞍山	合肥
1986.3.21	百货大楼	洗衣机	350	合肥	南京
⋮	⋮	⋮	⋮	⋮	⋮

表中的日期、单位、品名、数量、发地、目的地等都是数据，但类型不同，有的是数值型，如数量，有的是字符型，如单位、品名、发地、目的地，有的是日历型，如日期等。这些数据需存贮、有待加工处理：何种货物由何种车型运输，经由何种路线不倒流以便运费最省等等…。以往凭调度员经验办事，但是若信息量很大并且希望科学管理，就得借助于计算机手段进行优化计算供决策依据，其经济效益绝非单纯节省几个人工的问题。

上述思想是把问题局限于某个局部范围分别对待，但是企业是个整体，各部门间有着千丝万缕的联系。一个数据多个部门都要用到，需要互相传递，这个问题在信息量剧增的现代，变得更加突出。比如经过货运处理后要下达调度方案，调某驾驶员承接某项货物运输，这就需要了解有关驾驶员情况，要了解驾驶员级别、开的是什么车型、车况如何以及他完成当年定额情况等，如表1—2所示。

表1—2 驾驶员登记表

姓名	年龄	级别	车型	车况	出车标志	完成客额
张三	45	2	解放	良好	途中	47%
李四	27	4	丰田	大修	/	35%
王二	25	4	黄河	良好	待命	32%
⋮	⋮	⋮	⋮	⋮	⋮	⋮

这需要事先将有关驾驶员信息存入计算机，存在一个查询检索问题。进而月终结算支付驾驶员工资及超额奖，又得查驾驶员工资表，对照他完成定额情况支付奖金等等。显然一个数据要在各部门互相传递利用，如果不统一地系统考虑问题，将势必造成数据冗余，即一个数据重复存贮，这是很不科学的。数据库技术有效地解决了这个问题，所以说数据库技术是迄今为止计算机用于数据处理领域的最新技术成果。

§ 1—2、什么是数据库

数据库这个词不知是谁首先提出来的，它形象但不准确，但是仍然被人们所接受并广泛地流传开来。数据库其实是个复杂的系统，难以用简练的语言准确地概括其全部特征，而且从不同角度人们对数据库有着不同的理解，所以至今对于什么是数据库仍无一个统一的公认的定义。因此我们打算从数据管理的演变及进展谈起，从对比中突出数据库的特点，对数据库给出一个轮廓性的描述。

一、数据管理技术的进展

数据管理指的是数据的组织、存贮、检索、维护等，是数据处理的核心问题。数据管理随计算机软件和硬件的发展，经历了如下三个阶段：

自由管理方式

文件系统方式

数据库系统方式

1、自由管理方式

早期的计算机，除了硬件之外，没有必要的软件支持，摆在用户面前的是一台裸机。在裸机上进行数据处理，用户必须自行考虑数据的物理存贮位置，用户程序不仅要设计数据处理手段，还得记住数据在内外存的地址，对数据的存取是按物理地址进行的。图1-2表明自由管理方式程序高度依赖于数据。

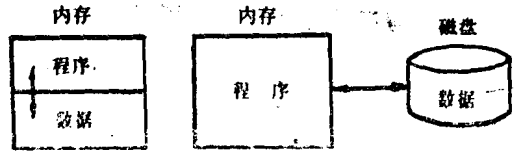


图1-2 程序与数据紧密相关

这种方式，用户负担极重，要直接与数据的物理地址打交道，数据稍有变动，便会导致程序全部作废，乃至会破坏其它用户数据，极不安全可靠。

2、文件管理方式

文件管理方式是有关数据独立出来组织成文件形式，在操作系统支持下容许对文件命名并通过文件名存取文件中的数据，而数据的实际“物理”存贮用户不必介入，由操作系统负责其间的映射。这就使得程序与数据相互间有一定的独立性，应用程序可以集中精力于算法而不必过多地考虑物理细节。后文要用到“逻辑”与“物理”这二个概念，在此先作一说明：

“逻辑数据”指的是用户看到的和理解的并通过数据名引用的数据形式；“物理数据”指的是实际存放在存贮介质上的数据，这二者之间可能差别极大，这二个词今后将一再出现。

文件管理是操作系统的一个功能模块，对于非计算机专业读者，可能比较陌生，有必要对操作系统略加介绍。

操作系统 (Operating System)

OS是计算机系统中最主要的直接与硬件接触的核心软件，它充当用户与裸机之间的界面，使得计算机呈现在用户面前的是一台比裸机功能更强、使用更加方便的虚拟机，如图1-3所示。

在计算机系统中，OS占据举足轻重地位，计算机愈复杂，其重要性愈突出。由图可见，OS是核心软件，外层软件依靠内核支撑。OS如同一个总管家，管理计算机中的所有软硬件设备资源，包括：

CPU管理

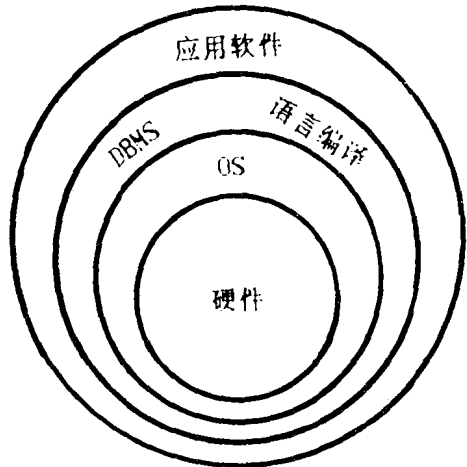


图1-3 硬件、软件层次

内存管理
设备管理
作业管理
文件管理

以作业管理为例，它使多用户多道作业能同时在一台计算机中运行而无需人工干预，在用户看来似乎他独占资源；

内存管理解决内存空间分配，避免程序干涉，及时回收无用空间，使系统更加安全可靠，避免用户直接介入内存分配造成误操作；

设备管理解决输入/输出设备驱动，省去用户直接管理外设的沉重负担；

下面较详细地说一下操作系统对文件的管理。

OS的文件管理

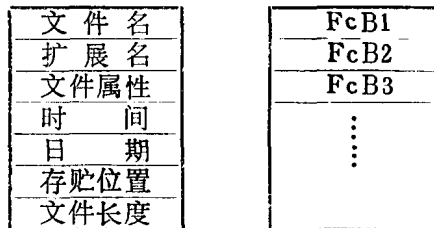
文件管理是OS对信息管理的一种手段。所谓文件就是把数据（或命令）按一定的格式组织起来构成的数据集合（或命令集合），文件是文件管理的基本单位。那么，OS是怎样对文件进行管理的呢？

①对每个文件赋予一个文件名，文件名便标识了该文件；

②每个文件名对应一个文件控制块FCB（File Control Block），在FCB中登记有关文件的参数，见图1—4（a）；

④由各个FCB构成文件目录存放在磁盘目录区，见图1—4（b）；

⑤对文件中数据单元的存取，通过文件名实现



a 文件控制块FcB b 文件目录

图1—4 文件结构

鉴于文件的重要性，下面给出文件及其有关的某些术语的确切定义：

字段（Field）

字段是表征现实信息的最小数据单位，在数据文件中是一个不可再分割的基本数据项，在COBOL语言中称为初等项。例如为了描述一个人的基本情况，通过姓名、年龄、性别、职务、文化程度，工资级别等来刻画，这些都是不可再分割的最小数据单位。

记录（Record）

字段的有序集合称为记录。

文件（File）

记录的集合称为文件。

在数据的文件管理方式中又可分为独立文件方式和共享文件方式。

独立方式，各个用户建立各自的数据文件，数据的逻辑定义、组织方式和存取方法

仍需程序决定，仅不过由OS的文件管理系统充当接口，通过它实现具体控制，如图1—5所示。

在上一节我们所举的汽运企业计算机管理的例子，就是属于这一种。这种管理方式，数据文件仍与程序紧密相关，一旦文件离开了使用它的应用程序，便失去了存在的全部意义。一个程序建立的文件，别的程序根本无法享用。尤其重要的是数据的组织仍是松散的结构，不能反映信息之间的客观联系，而这恰恰是数据应具有至关重要的性质。

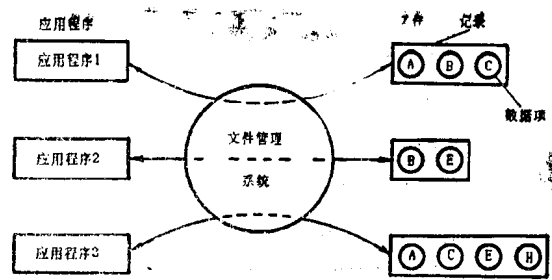


图1—5 独立文件方式示意图

共享文件方式是在独立文件方式基础上发展起来的，这种方式由一个专用程序负责不同应用程序共享不同文件，也可实现不同文件中必要的数据库联系，即通过专用程序能将不同文件中的数据组织起来形成一个整体结构信息，如图1—6所示。所有这些特性，标志着共享文件方式已经有了数据库的某些特点，但它只是数据库的雏形，还不是真正的数据库。

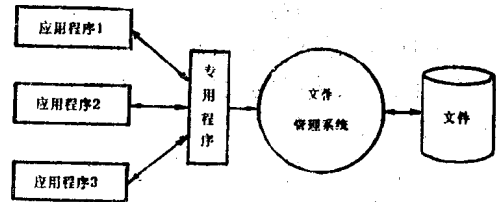


图1—6 共享文件方式示意图

文件方式藉助于OS的支持，允许通过文件名访问数据文件，提供了从逻辑到物理的映射，不再需要考虑数据的物理地址，从而使数据处理技术跨越了一大步。但是文件方式仍存在很大缺点，主要表现在数据面向具体应用，不是以整体观点设计数据模型，而仍然是孤立地考虑问题，因此必然存在以下弊端：

①数据冗余

数据冗余是孤立考虑问题的必然结果，同一数据重复存贮。例如驾驶员姓名，在工资管理程序中，人事档案程序中，货运调度程序中都要用到，由于孤立地考虑问题，不得不各自建立文件，因而造成驾驶员这个数据出现在不同文件中，使空间利用率大大降低。

②可能引起数据不相容

这是数据冗余的必然结果。一个数据存贮在不同文件中极易引起数据不相容，特别当数据更新时最容易发生矛盾，造成同一数据具有不同数值。

③程序不灵活

文件方式仍然是数据面向应用，即程序是针对文件编制的，一旦写成便难以修改或扩充，而建立新的文件，又将导致程序无效。

④不能反映现实世界各种事物之间客观存在的联系

3、数据库方式

数据库方式从整体观点组织数据，数据是结构化的，面向整个系统，因而能满足所有用户的不同需求。应用程序不再与具体的物理文件相对应，而是取自整体数据的某个子集作为逻辑文件与应用程序相对应，通过一个软件建立逻辑文件与物理文件之间的映射关系。这样，文件仅仅是逻辑性的出现在程序中，如图1—7所示。

有了这种初步认识，下面给出数据库的定义，已如前述，对于什么是数据库，从不同的角度有着不同定义。

二、数据库定义

Martin James 的定义：

数据库是一组相关数据的集合，这些数据避免了有害的数据冗余，能为不同用户所共享。数据的存贮独立于程序，对数据的插入、更新、检索等操作均能以一种通用的控制方法进行。

C、J、Date从三个方面对数据库进行了描述：

- ① 存贮在磁盘或磁鼓介质上的数据库本身；
- ② 批处理程序或远程终端用户以常规方式（包括查询、插入、修改、删除）使用数据库中的数据；

③ 数据库中的数据是集成的，包括了许多不同用户的数据，允许不同用户使用的数据相互重叠，同一片数据为多用户所共享。

图1—8是Date对数据库描述的形象表示。

· 其它的定义：

数据库是集中统一地保存一个单位的所有有用信息的数据系统，能反映信息及其联系并能提供一切必须的存取路径存取其中任一数据满足不同用户需求。

不论哪一种定义，无不直接或间接的作了如下论断：

① 数据库中的数据含有许多不同用户所需要的成份，而且是统一地组织起来，因此这个数据系统是集成化、结构化的。

② 因此，所有用户有关的数据都能从数据库中导出，为多用户所共享。

③ 不同用户对数据库有不同的理解，他所看到的只是与他的程序有关的那部分数据，有他自己的数据视图。好有一比：如同“瞎子摸象”，一个人摸到象腿，说它是一棵树；另一人摸针象尾，觉得它象一根绳子；摸到了象耳，觉得它是把扇子。

④ 数据之间的联系，是数据库最重要的特点，是数据库方式与文件方式根本区别所在。现实世界各种事物之间的内在联系，在数据库中表现为数据之间的联系。因此，数据库的数据结构比较复杂，表现在描述数据时不仅要描述它本身，还要描述数据之间的联系。

⑤ 在数据库中的数据联系是通过存取路径实现的。

为了加深对数据库的理解，让我们考察一个机床厂的产品生产模型，从中体会数据库的含义。

设某机床厂生产若干种机床，所需另件除了本厂制造外，多数由外协单位供应，一部机床由许多另部件装配起来的，是由不同部门不同车间工人生产出来的，等等…。以上描述用图表示最直观，图1—9就是所述的各种数据以及数据之间存在的自然联系。

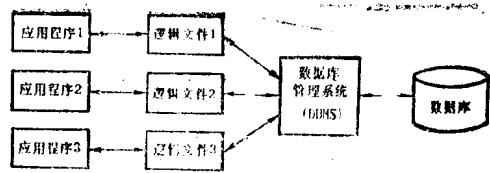


图1—7 数据库方式示意图

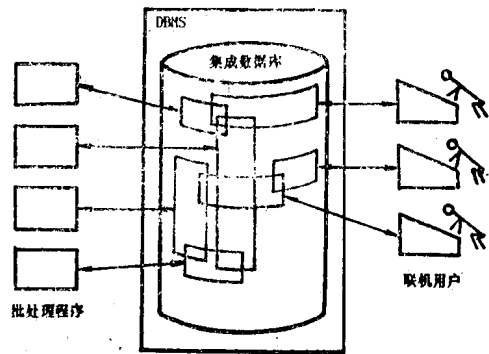


图1—8 数据库系统示意图

在数据库中就是要把各个数据以及它们之间的联系组织起来，形成结构化的数据系统，

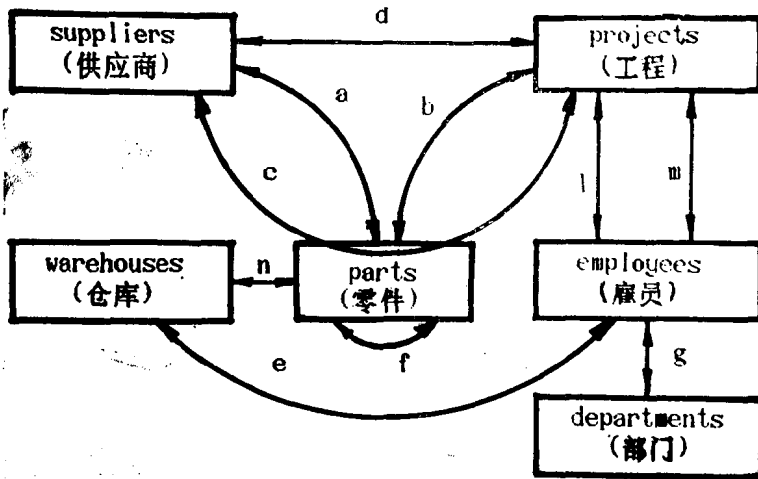


图 1-9 数据联系图

否则无法反映客观现实，所述模型将不复存在，可见联系是多么重要。

图中，客观事物视作实体，分别用方框表示，实体之间的联系用双向有向线段表示。

对此图说明如下：

在实体联系中有多种不同情况，这是根据问题的性质和自然情况决定的。具体地说：有二个实体之间的联系（如a和b）；也有三个实体之间联系（如c）；还可能存在二个实体间二种以上的联系（如l和m）；以及一个实体类型内部数据间的联系（如f）等等…。

a将Supp和Part两个实体联系起来，一个箭头方向表示某供应商供应哪些另件，反向表示某另件由不同供应商供给；

b将Proj和Part两个实体联系起来，表示某工程用到哪些另件，反向表示一种另件被哪些工程所使用；

c将Supp、Part和Proj三个实体联系起来，表示某供应商供应若干另件给哪些工程，反之表示每个工程项目使用的每种另件可由不同的供应商供给；

上述几种情况，显然c比a和b表示更多的信息，因为仅从a和b，导不出c。证明如下：

假定从a的联系中得知 S_2 供应 P_4 ；从b的联系中得知 P_4 用于 J_3 ，这时你不能断言“ S_2 向 J_3 供应 P_4 ”，因为供应 P_4 者不仅是 S_2 ，还可能有 S_y ，而 P_4 不仅仅用于 J_3 ，还可能用于 J_x ，显然 S_2 不一定是 S_y ， J_3 不一定是 J_x 。所以只能从三个实体联系中，才能表示出“某某供应商向某某工程供应某某另件”。

l和m是Proj和emp二实体间的二种不同的联系，l表示某工程由哪些工人施工；m表示某工程的指挥者是谁，它们表示的信息是不相同的。

f表示Part内部的数据间联系，因为在Part中有的是另件，有的是部件，而部件是由另件装配起来的。

对于上述的例子，在数据库中要把各个实体的有关信息以及各个实体之间存在的自然联系，统一地组织起来，所以我们说数据库是一个集成化的数据系统。对数据的存取路径问题，沿着哪一条路径，通过哪些联系才能找到我们所需要的数据，是个重要问题，有关这方面的知识后继章节介绍。

§ 1—3、构成数据库系统的三要素

数据库是个复杂的系统，并非单指数据库本身，而是指计算机引进数据库技术后的整个系统，它由三部分构成：数据、硬件和软件。

1、数据

这里的数据指的是数据本身和数据之间的联系，联系也是数据。数据是数据库工作的对象，前而已经相当详细地讨论过了。

2、硬件

硬件资源包括CPU、内存、磁盘及其它外部设备。在数据处理中，影响处理速度的决定因素是 I/O时间而非CPU速度。认识这一点很重要，它说明了在数据处理业务中，必须很好地研究数据的组织和存贮而且主要是研究数据如何在外存的组织和存贮，否则将导致 I/O次数的增加，会严重影响效率。此外，数据库系统要求较大的内存空间，因为数据库运行，除操作系统外，还有数据库管理系统、应用程序以及缓冲区均占据内存。例如dbase-III要求内存不少于256k，汉字dbase-III不少于512k。

数据库对外存的要求更加严格，必须有大容量的磁盘或磁带设备的支持。外存是一种非易失性的存贮介质，掉电不丢失数据而且容量大，这正是数据库系统所必须的。特别是磁盘设备是一种能直接进行存取的外设，不同于磁带只能顺序存取，它可以随机访问库中的任一数据，这是数据库运行绝对不可缺少的重要性质，可以说它是数据库赖以存在的物质条件。不过磁带存贮数据稳定可靠且成本低，通常用作维护和恢复数据库作为辅助外存用。结论如下：磁盘配以磁带作为数据库的外存是较为理想的方案。

3、软件

软件资源包括：

OS

主语言（COBOL、FORTRAN等）

数据库管理系统DBMS

应用程序等。

其中，DBMS是数据库管理系统的简称，全文是：

Data Base Management System

今后一律用DBMS缩写替代数据库管理系统。

DBMS是数据库系统软件，它对数据的运行进行统一控制管理，允许用户逻辑地、抽象地使用数据而不必涉及这些数据在计算机中是怎样存放的，它负责数据的映射、负责数据处理和维护，它充当用户与数据库的介面。简言之，DBMS是用户使用数据库的工具。图1—10是一个微型计算机数据库系统配置。

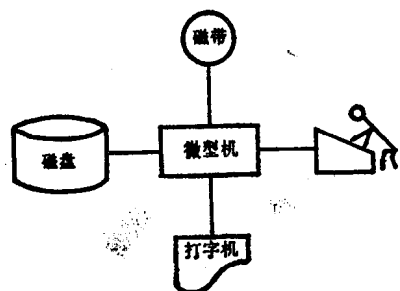


图 1—10 微型计算机数据库系统硬件配置

综上所述，数据库系统包括：

- ①数据：数据库工作对象
- ②硬件：物理支撑环境
- ③软件：主要指DBMS

平时，说到数据库，多半是指DBMS而言，今后我们还要进一步讨论。

§ 1.4、数据库种类

根据数据间联系所采用的方法不同，人们把数据库分成三种类型，当前软件公司出售的商品软件DSMS基本上概括在这三类中：

- 层次式数据库
- 网状数据库
- 关系式数据库

1、层次式数据库

层次式数据库，呈树状结构，其名称由来是因为把数据库中的各个数据组织成一棵倒挂的树形，由根长出若干分支，每一分支又长出若干子树，如图1—11所示。

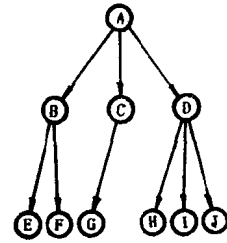


图1—11 层次式数据库呈树形结构

这种结构，数据之间的联系表现为有向线段，在数据库中则藉助于有逻辑关系的数据之间的指针，提供必须的存取路径。

2、网状数据库

网状结构反映数据的错综复杂关系，各数据间不仅表现为一种层次联系，也发生横向联系，如图1—12所示。

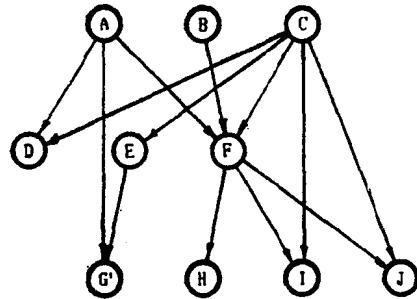


图1—12 网状数据库

比较图1—11和图1—12，不难发现网状结构任一结点上的数据都可与其它任意结点上的数据发生联系，而层次式结构只有相邻两级的结点数据才能相连，这是两者的根本区别。至于联系的手段，大多也是通过数据指针实现，从这一角度讲，层次式与网状式的结构，其实质是一样的。

3、关系数据库

“关系”是个数学名词，通俗地讲它是一张二维表格，用二维表格反映数据及其联系，是关系式数据库本质所在。如图1—13所示的关于另件的有关数据。

零件号	零件名	规格	重量	单价

图1—13 关系式数据库结构

用关系反映‘另件’这个实体很容易理解，图1—13通过各个栏目的设置，便将另件的有关数据及其特征刻划了出来，问题是如何表示数据之间的联系，即两个不同的关系怎样实现联系？

不同于层次、网状数据库，关系数据库是通过在不同的关系中那些含有相同栏目的数据项来实现的。换句话说关系数据库不是人为的设置指针，而是由数据本身自然地建立起联系，通过关系运算就可存取任一数据。