



高等院校城市规划与建筑系列教材

城市规划数据库技术

张 军 周玉红 \ 主编



全国优秀出版社
武汉大学出版社





高等院校城市规划与建筑系列教材

城市规划数据库技术

张军 周玉红 \ 主编

图书馆



全国高等院校出版社
武汉大学出版社

图书在版编目(CIP)数据

城市规划数据库技术/张军,周玉红主编. —武汉: 武汉大学出版社,
2004. 9

高等院校城市规划与建筑系列教材

ISBN 7-307-04291-6

I. 城… I. ①张… ②周… III. 数据库系统—应用—城市规划 IV.
TU984

中国版本图书馆 CIP 数据核字(2004)第 056777 号

责任编辑: 解云琳

责任校对: 王 建

版式设计: 支 笛

出版发行: 武汉大学出版社 (430072 武昌 珞珈山)

(电子邮件: wdp4@whu.edu.cn 网址: www.wdp.whu.edu.cn)

印刷: 武汉理工大印刷厂

开本: 787×1092 1/16 印张: 19.125 字数: 458 千字

版次: 2004 年 9 月第 1 版 2004 年 9 月第 1 次印刷

ISBN 7-307-04291-6/TU·47 定价: 29.00 元

版权所有, 不得翻印; 凡购我社的图书, 如有缺页、倒页、脱页等质量问题, 请与当地图书销售门联系调换。

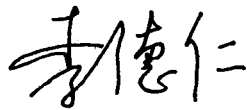
序

随着地理信息系统(GIS)的产生和发展以及在各行各业的不断应用,GIS的理论和技術得到了空前的发展。城市规划作为GIS的重要应用领域之一,为人们普遍关注。自20世纪80年代GIS传入中国后,很多地区都开展了城市信息系统相关项目的建设,一些高等院校也设立了相关的专业方向,因而也吸引了大量的计算机信息业人才从事GIS在城市规划中的应用研究。GIS在这个领域应用的一个最重要的基础是数据库技术。我们究竟应该关心哪些数据以及如何处理这些数据并不是所有的从业人员都明白和掌握了。多数人,包括高等院校的学生都对此缺乏了解,而目前也没有专门的将计算机科学中的数据库技术与城市规划相结合的书籍供参考和使用,《城市规划数据库技术》这本书正是填补了这一空白。

本书的两位编者长期从事城市规划和城市信息系统的研究和实践工作,积累了大量城市规划数据库开发的经验,所提出的一些观点和阐述的内容避免了以往很多城市规划数据库相关书籍注重理论讲解而忽视数据库实践和应用的不足,具有较强的针对性和实用性。本书可以帮助各类城市信息系统的开发人员少走弯路,必将在一定程度上促进城市信息系统开发过程的标准化。

该书不但是编者多年实践的积累,同时也收录了近年来一些相关的理论和技術,如城市规划数据库的内容、空间数据的组织方法、空间数据索引和UML等,这都说明本书的编者既关心城市规划行业的应用问题,也重视数据库技术发展的动态,具有一定的前瞻性,这对于扩充读者的视野是非常有意义的。

随着中国政府对于建设“信息化社会”目标的制定和“数字城市”的深入发展,城市信息系统的发展进程是非常值得关注的,它在一定程度上关系着“信息化”社会目标的最终实现,而作为城市规划信息化的基础——数据库技术也必须引起关注。希望广大从事这项工作的管理人员和技术人员加倍努力,为中国社会的信息化建设做出贡献!



二〇〇四年一月十二日

前 言

伴随着地理信息系统在城市规划领域应用的不断深入,为了满足城市规划与设计专业和相关专业的学科发展及对计算机应用的要求,根据这些学科自身的特点和对计算机应用数据库的要求编写了本教材。

全书共分七章,其主要内容包括:数据库的基础知识、关系数据库设计方法、规划的空间数据的组织方法、城市信息系统的开发设计、城市规划数据库技术的应用及开发、标准化查询语言 SQL 以及城市规划数据库的发展趋势。附录中还以洛阳 UPMIS 中的数据库系统为例说明了 UPMIS 数据库设计的部分内容。

本书着重讲解了规划数据库的相关理论,并配以大量实例进行详细解说,便于课堂教学和自学。本教材可作为城市规划专业本科生及研究生教材,同时也可以供建筑学专业、设计专业和资源环境相关专业教学参考。

本书是在武汉大学城市规划专业和原武汉测绘科技大学城市规划与设计专业及城镇建设专业本科生和研究生多年教学中使用的自编教材的基础上扩充、改造而成的,并吸收了作者多年来的教学实践经验和科研成果,同时,借鉴了荷兰国际航空航天测量与地学学院相关课程和教材的内容。

参加本书编写的人员有张军(第一章、第二章、第三章、第五章、附录二),周玉红(第四章、第六章、第七章、附录一)。书中插图由杨满伦、郭汝、王芳、张悦毅、陈庆利、周蕊、汪华绘制。由张军负责全书的统稿、定稿工作。

本书在编写过程中得到了中国科学院院士、中国工程院院士李德仁教授,武汉大学徐肇忠教授、蓝运超教授的大力支持和指导,他们都对本书提出了宝贵意见,在此一并致以谢忱!

同时,感谢武汉大学教务部及武汉大学城市建设学院给予该书出版经费上的资助以及城市建设学院的领导和老师在教材编写过程中提出的建议和给予的帮助,感谢武汉大学出版社解云琳编辑为该书的出版付出的艰辛劳动。

由于受编写时间和作者水平的限制,书中难免存在缺点和错误,敬请读者批评指正。

编者

2004 年 9 月

目 录

第一章 概论	1
第一节 基本概念	1
第二节 数据库的发展历史	4
第三节 数据库技术在城市规划中的应用	15
第二章 关系数据库基本原理	19
第一节 关系表	19
第二节 完全规范化表	31
第三节 实体关系模型	33
第四节 数据库设计	56
第五节 UML 及其在数据库设计中的应用	61
第三章 关系数据库标准查询语言	68
第一节 数据定义	69
第二节 查询	71
第三节 数据更新	81
第四节 视图	83
第四章 城市规划空间数据的组织	88
第一节 城市规划空间数据概述	88
第二节 空间数据的组织方法	97
第三节 空间数据索引	132
第五章 城市规划管理信息系统	137
第一节 信息系统	137
第二节 信息系统开发方法	140
第三节 结构化系统开发生命周期法	145
第四节 快速原型法	148
第五节 城市规划管理信息系统设计	151
第六章 城市规划数据库的设计	168

第一节	城市规划数据库设计概述	168
第二节	概念化设计、逻辑设计与实施规划	184
第三节	城市规划数据库的内容与设计示例	218
第七章	城市规划数据库的发展趋势	244
第一节	数据库技术的发展	244
第二节	面向城市规划领域的数据库新技术	255
附录一	数据库系统设计	268
附录二	数据库设计范例	288
参考文献		296

第一章 概 论

数据库技术产生于 20 世纪 60 年代中期，是数据管理的新技术。作为计算机科学的重要分支，随着计算机应用在各行各业的深入发展，数据库技术也得到了长足的发展。当然，它也应用于城市规划领域，并形成工程数据库的一部分。本章介绍数据库的有关概念以及为什么要发展数据库技术。

第一节 基本概念

一、数据

人类所处的世界是物质世界，同时也是信息世界。信息伴随着物质的存在而存在。信息是物质的反映，而表示信息最准确的是数据。因此，数据（Date）是物质世界中信息的表示形式。

说起数据，人们首先想到的是数字。数字表示信息，有时是自然的，有时是人为的。例如：一个城市的人口为 50 万，一块城市用地面积为 5 公顷，一栋楼房有 20 层，这些都是自然数字表示的数据。又如：楼房的外墙面是红色或绿色，其中的红色或绿色无法直接用数字表示，可定义 1 表示红色，2 表示绿色，这里的数字 1 或 2 则是人为数字表示的数据。另外，除了以上提到的数字，日常生活中的文字、图形、图像、声音等都是数据。

为了认识物质世界并交流信息，人们需要描述客观事物。数据实际上是描述事物的符号记录。在计算机中，为了存储和处理这些事物，就必须将这些事物令人感兴趣的或必要的特征组织成一组记录来描述。例如：管理人事档案时，人们可能感兴趣的是人的姓名、出生年月、性别、籍贯、工作时间、基本工资、职称等特征，那么可以采用如下的方式进行描述：

（张泰，1965/5，男，山东青岛，1988/9，465，工程师）

又如：规划管理中的红线，其必要的特征有红线编号、所属单位、地块面积、审批日期、经办人等，那么可以描述如下：

（2001001，柳州市自来水公司，800，2001/8/6，张华）

二、数据库

收集大量数据后，应将其保存起来以供进一步加工处理和提取有用的信息。数据保存的方法有人工保存、存放在文件里、存放在数据库里等多种形式。其中，数据库（Database，DB）是存放数据的最佳场所。

所谓数据库，是指长期存储在计算机中的有组织、可共享的数据集合。数据库中的数

据按一定的数据模型组织、描述和存储，具有较小的冗余、较高的数据独立性和易扩展性，并可为各种用户共享。

三、数据库管理系统

数据库管理系统 (Database Management System, DBMS) 是位于用户与操作系统之间的一层数据管理软件，同时它也是数据库系统的核心软件，是数据处理技术各种先进思想的集中，是一种综合的、通用的大型系统软件。

数据库在建立、运行和维护时由 DBMS 统一管理、统一控制。DBMS 使用户能方便地定义数据和操作数据，并能保证数据的安全性、完整性、多用户对数据的并发使用及发生故障后的系统恢复。

数据库技术自出现开始，DBMS 从初级到高级，其功能越来越强大，技术与理论也越来越完善。根本的发展是数据独立性越来越强，用户接口趋向于简单、方便、直观，用户操作非过程化、智能化。尤其是从层次模型、网状模型发展到关系模型以后，DBMS 的目标更加明确。总之，DBMS 包括的最基本部分有：定义数据库结构的语言，即数据定义语言；操纵数据库语言，即数据操纵语言；数据库系统的各种应用程序，包括装配程序、控制程序、维护程序、故障恢复程序和其他实用程序。

四、数据库系统

数据库系统一般由数据库、数据库管理系统及其开发工具、应用系统、数据库管理员和用户构成。当然，从其运行环境来看，还应包括硬件资源。应当指出，数据库的建立、使用和维护等工作只靠一个 DBMS 是不够的，还需要专门的人员来完成，这些人员就是数据库管理员 (Database Administrator, DBA)。

数据库系统可以用图 1-1 来表示。

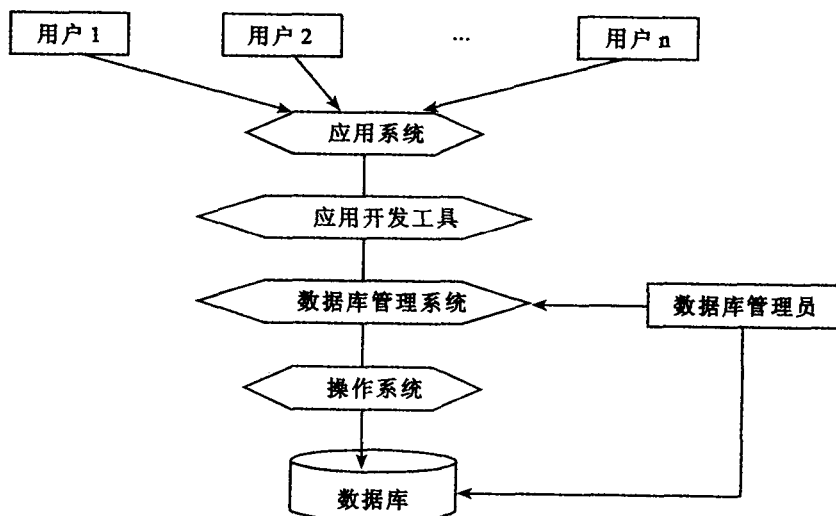


图 1-1 数据库系统

下面对数据库系统组成部分的硬件、软件及数据库管理员分别加以介绍。

1. 硬件

这里的硬件指数据库的硬件资源，它包括 CPU、内存、磁盘、终端显示器、键盘、打印机等外围设备。

(1) CPU

CPU 决定计算机的运算速度。数据库处理中数据检索的速度与计算机本身的运算密切相关，因而，CPU 决定了数据库处理的速度。同时，数据库的检索速度也与 I/O 操作所占用的时间有关。

(2) 内存

在数据库系统下，数据库管理系统、操作系统、应用程序、系统缓冲区、目标模式和子模式等多个程序都要常驻内存，因而要求具有比较大的内存空间。

(3) 外存

外存是存放数据库的仓库，它的大小至少应该容纳下数据库中的数据，否则，数据库系统将无法运行。外存分为两级，即联机 and 脱机。联机查询的数据必须存放在磁盘上；而用于维护、恢复的副本、日志等则一般脱机存放在磁带上。因此，两种外存通常都需要配备，而且容量要足够大。

2. 软件

指数据库系统的软件资源，包括操作系统、主语言（即高级语言）、DBMS、应用软件。

(1) 操作系统

操作系统（OS）是 DBMS 的主要支持软件，DBMS 对数据库的存取操作最终是由 OS 完成的，因此 OS 直接影响数据库的检索速度和数据库系统的功能。

(2) 主语言

DBMS 只提供最基本的数据操作命令，而其他一些功能语句，如计算、判断、条件等语句则由支持 DBMS 的 OS 支持下的高级语言提供。用户在编写应用程序时，可以通过调用 DBMS 的数据操作命令查找数据，进入 DBMS 的功能程序，完成后再返回。高级语言与 DBMS 的接口方式通常有两种：应用程序通过普通的过程调用来启动 DBMS 的命令执行程序；数据操作语言的命令作为主语言的扩展，用户将操作命令嵌入到主语言中去。

(3) 应用软件

应用软件不属于 DBMS。DBMS 只能向用户提供最基本的数据存取功能，不能提供满足用户特定要求的检索和特殊查询功能；如果需要，就必须使用高级语言和数据操作语言编写程序完成。另外，对于用户的具体业务管理程序，如规划管理部门的红线管理程序、银行的出纳和记账程序等，也需要通过高级语言和数据操作语言编写程序完成。以上所提到的程序就是应用程序（软件）。

随着数据库技术的发展，应用程序的发展出现了两种趋势：一是用 DBMS 和高级语言将大量重复的工作编制成通用程序，组成应用程序软件包，如报表生成程序、统计计算程序等；二是 DBMS 向着非过程化发展，即将一部分共性应用软件纳入自己的范畴，如求平均值、求和、求方差等模块。

3. 数据库管理员

数据库系统中除了集中管理数据的系统软件外，还有专门负责整个系统的建立、维

护、协调工作的人员，这些人被称为数据库管理员（DBA）。DBA 不仅要熟悉系统软件，还要熟悉相关部门的业务工作。其任务是：

- (1) 进行需求调查，决定数据库的信息内容；
- (2) 进行数据库的逻辑设计、物理设计，描述数据库的结构，即形成模式、子模式和物理模式；
- (3) 与用户取得联系，随时调整和修订系统；
- (4) 定义用户的使用权限；
- (5) 负责维护、恢复工作。

第二节 数据库的发展历史

一、数据库技术的发展

数据库技术是应数据管理任务的需要而产生的。数据管理是指对数据进行分类、组织、编码、存储、检索和维护，它是数据处理的中心问题。随着计算机硬件和软件的发展，数据管理经历了人工管理、文件系统和数据库系统三个阶段。

1. 人工管理阶段

20 世纪 50 年代中期以前，计算机主要用于科学计算。当时的硬件状况是：外存只有纸带、卡片、磁带，而没有直接存取的存储设备；软件方面没有操作系统，没有管理数据的软件；数据处理方式是批处理。

人工数据处理具有如下特点：

- (1) 数据不保存。
- (2) 数据需要由应用程序自身管理。应用程序中不仅要规定数据的逻辑结构，而且要设计相应的物理结构，包括存储结构、存取方法、输入方式等。
- (3) 数据不共享。数据都是面向应用的，一组数据只能对应一个程序。当一些相同的数据要被几个应用程序使用时，必须各自重新定义。
- (4) 数据不具有独立性。数据的逻辑结构或物理结构发生变化时，必须对相应的应用程序做出修改。

人工管理阶段应用程序与数据之间的对应关系可用图 1-2 表示。

2. 文件系统阶段

20 世纪 50 年代后期，随着计算机的发展，计算机不仅用于科学计算，而且大量用于管理。这时硬件有了磁盘、磁鼓等直接存取设备；软件方面，操作系统中已经有了专门的数据管理软件，一般称为文件系统；处理方式上不仅有了文件批处理，而且能够联机实时处理。

文件系统管理数据具有以下特点：

- (1) 数据可以长期保存。由于计算机大量用于数据处理，数据需要长期保存在外存上，以便反复进行查询、修改、插入、删除等操作。
- (2) 有专门的文件系统管理数据，程序与数据之间由软件提供的存取方法进行转换，使得应用程序与数据间有一定的独立性。数据存储上的改变不一定反映在应用程序上。

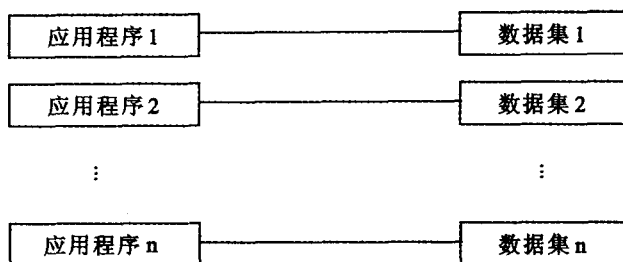


图 1-2 人工管理阶段应用程序与数据之间的关系

(3) 数据独立性低。文件系统中的文件是为某一特定应用服务的，数据逻辑结构对该应用是最优的，但是要增加数据将非常困难，系统不易扩张。同时，数据逻辑结构的改变也会引起应用程序的修改。虽然数据与程序之间有了独立性，但独立性非常低。

(4) 数据共享性差。在文件系统中，一个文件基本上对应一个应用程序。当不同应用程序具有部分相同的数据时，必须建立各自的文件，数据冗余大，存储空间浪费大。

文件系统阶段应用程序与数据之间的关系可用图 1-3 表示。

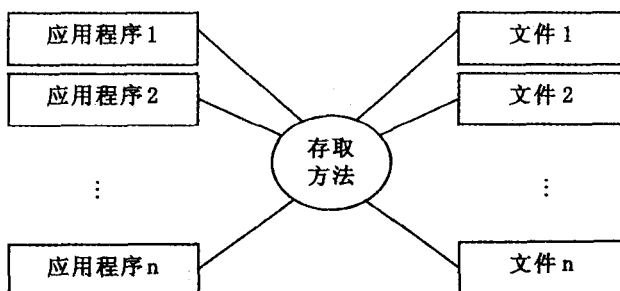


图 1-3 文件系统阶段应用程序与数据之间的关系

3. 数据库系统阶段

20 世纪 60 年代后，计算机用于管理的应用越来越广泛，同时，多种应用、多语言互相覆盖地共享数据集合的要求越来越强烈。这时，硬件已有大容量磁盘；编制和维护系统软件及应用程序所需的成本相对增加；在处理方式上，联机实时处理要求越来越多，并开始提出和考虑分布式处理。为满足多用户、多应用共享数据的需求，使数据为尽可能多的用户服务，出现了数据库技术及统一管理数据的专门软件系统——DBMS。

数据库系统管理数据具有如下特点：

(1) 数据结构化

在文件系统中，相互独立的文件记录本身是有结构的。最常见的一种文件形式，同时也是最简单的文件形式是等长同格式的记录集合。例如：建筑工程申请表中的每条记录可采用如图 1-4 所示的记录格式。

项目编号	建筑名称	建筑结构	高度	层数	底层面积	建筑面积	单位造价	合计金额
------	------	------	----	----	------	------	------	------

图 1-4 建筑工程申请记录

对于任何建筑工程，以上 9 项中除了建筑名称外，都是必需的而且基本上是等长的，建筑名称则随着信息量的变化而变化较大。如果采用等长的记录形式来存储这些数据，每条建筑项目记录的长度必须等于信息量最多的记录的长度，因而造成大量的存储空间被浪费。所以，最好使用变长记录或主记录与详细记录相结合的形式建立文件，即将建筑名称作为详细记录，其他项作为主记录，每条记录具有如图 1-5 所示的记录格式。

项目编号	建筑结构	高度	层数	底层面积	建筑面积	单位造价	合计金额
------	------	----	----	------	------	------	------

↓

项目编号	建筑名称	备注
------	------	----

图 1-5 记录格式

通过这样的方式可以节省许多存储空间，文件的灵活性也有相对提高。但是这样建立的文件仍然有局限性，因为其灵活性仅限于一个应用而言。对于一个规划管理部门而言，将涉及许多应用。在数据库系统中，不仅要考虑某个应用的数据结构，还要考虑整个组织的数据结构。例如：对于一个规划管理信息系统，还要考虑规划审批管理、用地审批管理、竣工验收管理等。具体规划管理系统的设计将在以后章节中详述。

以上这种数据组织方式为各个管理提供必要的记录，使整个管理的数据结构化。这就要求在描述数据时，不仅要描述数据本身，还要描述数据之间的联系。因而数据库系统实现整体数据的结构化，是数据库的主要特征之一。

在数据库系统中，不仅数据是结构化的，而且存取数据方式也很灵活，可以存取数据库中的某一个数据项、一组数据项、一条记录或一组记录。而在文件系统中，数据的最小存储单位是记录，粒度不能细化到数据项。

(2) 数据的共享性好，冗余度低

数据的共享程度直接影响数据的冗余度。数据的共享程度越高，数据的冗余度就越低；反之，数据的共享程度越低，数据的冗余度就越高。数据库系统从整体角度看待和描述数据，数据不再面向某个应用，而是面向整个系统。图 1-6 所示的例子中，项目基本记录可以被多个应用共享使用。这样既可以大大减少数据冗余，节约存储空间，又能够避免数据之间的不相容性与不一致性。所谓数据的不一致性，是指同一数据不同拷贝的值不同。采用人工管理或文件管理系统时，由于数据被重复存储，当使用和修改不同的数据拷贝时，就容易造成数据的不一致。

(3) 数据库的独立性强

数据库系统提供了两方面的映像功能，从而使数据既具有物理独立性，又具有逻辑独立性。

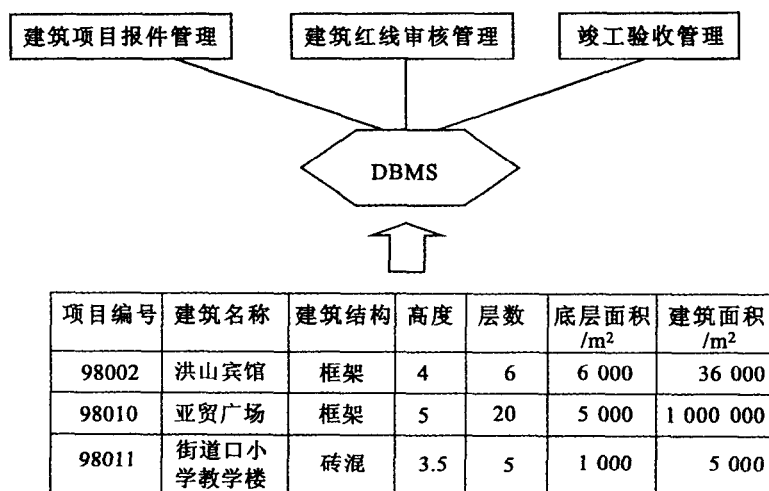


图 1-6

数据库系统的一个映像功能是数据的总体逻辑结构与某类应用所涉及的局部逻辑结构之间的映像，这种映像被称为逻辑变换。逻辑变换保证当数据总体逻辑结构改变时，通过对映像的相应改变可以保证数据的局部逻辑结构不变。由于应用程序是按照数据的局部逻辑结构编写的，因而无须改变应用程序。这就是数据与程序的逻辑独立性，简称为数据的逻辑独立性。

数据库系统的另一个映像功能是数据的存储结构与逻辑结构之间的映像或转换功能，被称为物理变换。物理变换保证当数据存储结构改变时，通过对映像的相应改变可以保证数据的逻辑结构不变，从而应用程序也无须改变。这就是数据与程序的物理独立性，简称为数据的物理独立性。

数据的独立性使得可以把数据的定义和描述从应用程序中分离开。同时，数据的存取由 DBMS 管理，用户不需要考虑存取路径等细节，从而简化了应用程序的编制，有利于应用系统的维护与修改。

(4) 数据由 DBMS 统一管理和控制

数据由 DBMS 统一管理，因此使用数据库时方式很灵活。可以获取整体数据的各种合理子集，用于不同的应用系统。当系统需求或功能发生改变时，只需重新选取不同的数据子集或加上相应的一部分数据，便可以有更多的用途，满足系统的新要求。

除数据管理功能外，DBMS 还必须提供数据控制功能，如数据的安全性（Security）、数据的完整性（Integrity）、并发控制（Concurrency）、数据库恢复（Recovery）。

数据库管理阶段应用程序与数据之间的对应关系可用图 1-7 表示。

综上所述，数据库是存储在计算机中有组织的、大量的、共享的数据的集合。它可供多种用户共享，具有最小冗余度和较高的数据独立性。DBMS 在数据库的建立、运行和维护时对数据库进行统一控制，以保证数据的完整性、安全性，并在多用户同时使用数据库时并发控制，在发生故障后对系统进行恢复。

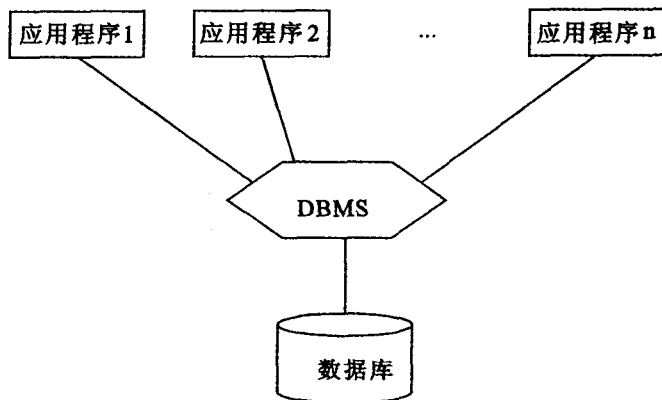


图 1-7 应用程序与数据库的对应关系

二、数据库的理论知识

1. 数据模型及其发展

数据库不仅要反映数据本身的内容，而且要反映数据之间的联系。在数据库中用数据模型来抽象表示和处理现实世界中的数据和信息。

数据模型应满足三方面的要求：比较真实地模拟现实世界；容易被人理解；便于在计算机上实现。数据库系统原则上应很好地满足这三方面的要求，针对不同的使用对象和应用目的，采用不同的数据模型。不同的数据模型具有不同的数据结构形式，其发展经历了层次模型（Hierarchical Model）、网状模型（Network Model）、关系数据模型，正在走向面向对象的数据模型等非传统数据模型。其中，层次模型和网状模型统称为格式化模型（或非关系模型），这种格式化模型的数据库系统基本上已经被关系模型的数据库系统取代，而面向对象的数据模型还不甚成熟，因而这里着重介绍关系模型，对其他两种模型只进行简单说明。

2. 数据模型的要素

数据模型常由数据结构、数据操作和数据的约束条件三个要素组成。

（1）数据结构

数据结构用于描述系统的静态特性，是所研究的对象类型（Object Type）的集合。这些对象是数据库的组成部分，它们包括两类：一是与数据类型、内容、性质有关的对象；二是与数据之间的联系有关的对象。

数据结构是刻画数据模型性质的重要方面，因此在数据库系统中，通常按照其数据结构的类型来命名数据模型。例如：层次结构、网状结构和关系结构的数据模型分别命名为层次模型、网状模型和关系模型。

（2）数据操作

数据操作用于描述系统的动态特性，是指对数据库中各种对象（类型）的实例（值）允许执行的操作的集合，包括操作及其操作规则。数据库主要有检索和更新（包括插入、删除、修改）两大类操作。数据模型必须定义这些操作的确切含义、操作符号、操作规

则以及实现操作的语言。

(3) 数据的约束条件

数据的约束条件是一组完整性规则的集合。完整性规则是给定的数据模型中数据及其联系所具有的制约和存储规则,用以限定符号数据模型的数据库状态以及状态的变化,以保证数据的正确、有效和相容。

数据模型应该反映和规定本数据模型定义的完整性约束条件;同时,数据模型还应该提供定义完整性约束条件的机制,以反映具体应用所涉及的数据必须遵守的、特定的语句约束条件。

3. 概念模型的概念及表示方法

为了把现实世界中具体事物抽象为 DBMS 所支持的数据模型,人们常常先将现实世界抽象为信息世界,然后将信息世界转换为计算机世界。也就是说,首先把现实世界中的客观对象抽象为一种信息结构,这种信息结构不依赖于具体的计算机系统,而是概念模型;然后再将概念模型转换为计算机中 DBMS 支持的数据模型。从这个过程中可以看出:概念模型是现实世界到计算机世界的一个中间层次。

由于概念模型用于信息世界的建模,是现实世界到信息世界的抽象,是用户与数据库设计人员之间进行交流的语言,因此概念模型一方面应该具有较强的语义表达能力,能够方便、直接地表达应用中的各种语义知识,另一方面它还应该简单、清晰、易于理解。

(1) 基本概念

概念模型的概念中包括了现实世界的与信息世界的一些概念,它们存在着一定的对应关系。

现实世界的概念有:

① 实体 (Entity)

客观存在并且相互区别的物体叫做实体。实体既可以是具体的人、事、物,如一个职员、一个学生、一块地、一个街区等,也可以是抽象的概念或关系,如部门的一次订货、居民与街区的关系等。

② 实体集 (Entity Set)

性质相同的同类实体的集合叫做实体集。如全体职员就是一个实体集。

③ 属性 (Attribute)

实体所具有的某一特性叫做属性。一个实体可以由若干属性来刻画。例如:学生的姓名、性别、年龄、籍贯、成绩等,一块地的编号、面积、用地性质、权属单位等,这些都表示了实体的固有特征。

④ 实体标识符 (Entity Mark)

现实世界中的任何两个实体各不相同,也就是说,不可能有两个属性完全相同的实体。那么,将能够区分同类实体中各个实体的属性集称为实体标识符。

⑤ 域 (Domain)

属性的取值范围称为该属性的域。例如:性别的域为(男,女),学号为8位整数,地块面积的域为大于0的实数等。

⑥ 实体型 (Entity)

具有相同属性的实体必然具有共同的特征和性质。用实体名及其属性名集合来抽象和

表示同类实体,称为实体型。例如,学生(姓名,性别,年龄,籍贯,成绩)就是一个实体型。

⑦ 联系 (Relationship)

现实世界中,事务内部以及事务之间存在联系,这些联系在信息世界中反映为实体内部的联系和实体之间的联系。实体内部的联系通常是指组成实体的各属性之间的联系。两个实体型之间的联系将在第二章中进行详细论述。

信息世界的概念则有:

① 字段 (Field)

标记实体属性的符号集叫做数据项或字段。字段与现实世界中的属性相对应,字段的命名通常和属性的名称相同。

② 记录 (Record)

字段的有序集合叫记录。记录对应现实世界中的实体,它是用来描述实体的,因而又可定义为描述一个实体集的所有符号集。例如一个地块的信息、一次借书登记。

③ 文件 (File)

同类记录的集合叫做文件。文件对应现实世界中的实体集,用来描述实体集,因而又可定义为描述一个实体集的所有集合符号集。例如一个学校所有的学生信息表构成学生文件。

④ 关键字 (Key)

能惟一表示一条记录的字段集叫做关键字。关键字对应现实世界中的实体标识符,用它来描述实体标识符,例如学生的学号、地块的编号等。

(2) 表示方法

概念模型的表示方法很多,其中最常见的是实体-联系方法 (Entity-relationship Approach)。该方法用 E-R 图来描述现实世界的概念模型。

E-R 图用来表示实体型、属性和联系的方法如下:

① 实体型:用矩形表示,矩形框内写明实体名。

② 属性:用椭圆形表示,并用无向边将其与实体联系起来。

③ 联系:用菱形表示,菱形框内写明联系名,并用无向边分别与有关实体连接起来,同时在无向边旁标上联系的类型 ($1:1$, $1:n$, $m:n$)。

注意,联系本身也是一种实体型,可以有属性。如果一个联系具有属性,那么这些属性也须用无向边与该联系相连。

图 1-8 用 E-R 图描述了实体之间的三种典型联系。

图 1-8 中的三个 E-R 图用非常简单的方法描述了现实世界的客观实际。如果用这种方法表示了现实世界中某一个具体的管理事务,则就形成了这个事件的概念模型。下面以城市土地权属管理(部分)为例加以说明。

为了清晰起见,图 1-9 中将复杂的管理事件用两张 E-R 图表示,分别为实体及其联系图和实体及其属性图。当然也可以用一张复杂的 E-R 图来表示,即将图 1-9 中的图 (a) 和图 (b) 合并,将实体所有的属性都表示在图 1-9 (a) 中。

实体-联系方法是抽象和描述现实世界的有力工具。用 E-R 图表示的概念模型独立于具体的 DBMS 所支持的数据模型,是各种数据模型的共同基础。因此,数据模型更一般、