

# 数据库技术

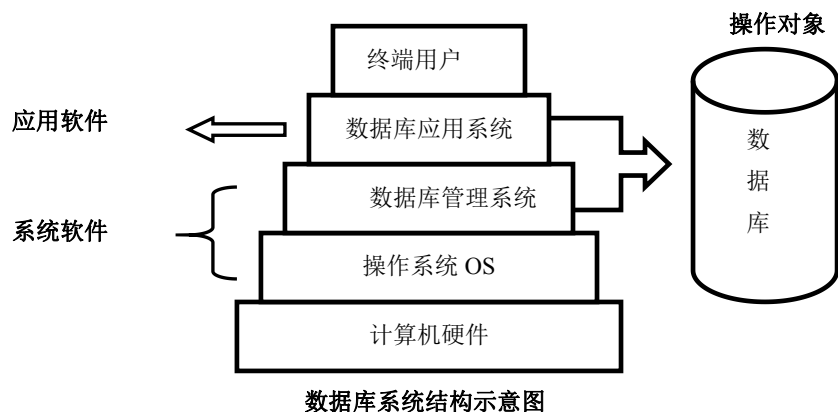
## 第一节 数据库技术的认识

## 一、数据库的基本概念

### （一）基本术语

概念	定义
数据（DATA）	是人类社会的一种重要信息资源，是对现实世界中客观事物的符号。
数据库（DB）	是指按照某种模型组织起来的，可以被用户或应用程序共享的数据的集合。
数据库系统（DBS）	即是采用了数据库技术的计算机系统，是实现有组织地、动态存储大量关联数据，方便多用户访问的计算机软件、硬件和数据资源而组成的系统。一般由数据库、数据库管理系统、应用系统、数据管理员和用户构成。从广义上讲是由数据库、硬件、软件和相关人员组成的。
数据库管理系统（DBMS）	是能够建立数据库、维护数据库及管理数据库的一个开发平台。
数据库应用系统（DBAS）	是应用了数据库的信息系统。
数据库管理员（DBA）	是从事管理和维护数据库管理系统（DBMS）的相关工作人员的统称，主要负责业务数据库从设计、测试到部署交付的全生命周期管理。
数据字典（DD）	是指对数据的数据项、数据结构、数据流、数据存储、等进行定义和描述，其目的是对数据流程图中的各个元素做出详细的说明。

## （二）数据库应用系统、数据库管理系统、数据库三者的区别和联系



说明：数据库系统的核心为数据库管理系统，数据库管理系统的核心为数据库（或数据），数据库技术的根本目的是要解决数据共享问题。

数据库系统是指在计算机系统系统中引入数据库后的系统，一般由数据库、数据库管理系统（及开发工具）、应用系统、数据库管理员构成。

## 二、数据库的发展阶段

### （一）数据管理技术的发展

数据管理是利用计算机硬件和软件技术对数据进行有效的收集、存储、处理和应用的过程。其目的在于充分有效地发挥数据的作用。实现数据有效管理的关键是数据组织。利用数据库进行存储管理大量数据具有能实现高效率的检索等方面的优势。随着计算机技术的发展，数据管理经历了人工管理、文件系统、数据库系统三个发展阶段。

发展阶段	应用范围	软硬件背景	特点
人工管理阶段	科学计算	硬件：无直接存取存储设备，只有磁带、卡片和纸带 软件：没有操作系统	计算机系统不提供对用户数据的管理功能、数据不能共享、不单独保存数据
文件系统阶段	科学计算 数据管理	硬件：出现了磁盘、磁鼓等直接存取存储设备 软件：出现了文件系统	数据可以长期保存、有文件系统管理数据、共享性差、数据独立性差
数据库系统阶段	大规模 数据管理	硬件：出现了大容量磁盘、磁盘阵列 软件：出现了数据库管理系统 (DBMS)	采用复杂的结构化的数据模型、较高的数据独立性、最低的冗余度、数据控制功能

### （二）文件系统阶段与数据库系统阶段的比较

阶段	独立性	冗余度	共享性	安全性与完整性	数据结构化程度
文件系统阶段	差	高	差	差	低
数据库系统阶段	高	低	高	高	高

说明：文件系统阶段与数据库系统阶段根本区别在数据的结构化程度高低；数据库技术的应用领域，主要应用于数据密集型领域。

## 三、数据库系统的分类

数据库系统一般简称数据库，分为关系型数据库（RDBM）和非关系型数据库(NoSQL)。

### （一）常见的关系数据库管理系统

#### 1.关系数据库分类

类别	对应数据库
大中型数据库	Oracle（Oracle 公司）、DB2（蓝色巨人 IBM）、SQL Server（微软公司） Sybase（Sybase 公司，已被 SAP 收购）
小型数据库	Foxpro（原为 Fox Software 公司的产品，后被微软收购）、Access（微软） MySQL（开发者为瑞典 MySQL AB 公司，最后被 Oracle 收购）
嵌入式数据库	SQLite（开源）、Firebird（开源） Berkeley DB（开源，Sleepycat Software，2006 年 2 月，Oracle 收购了 Sleepycat Software。）

	SQLCE (SQL Server 的 CE 版本)
--	----------------------------

## 2.主要关系数据库的比较

数据库	开放性	伸展性, 并行性	安全性	性能	客户端支持 及应用模式	操作简便	使用风险
SQL Server	只运行在 Windows	有限	无安全证书	不佳	C/S 结构	简单	有风险
DB2	跨平台	很好	ISO 标准认证	较高	多层结构	简单	风险小
Oracle		良好		最高		较复杂	无风险

## 3.关系数据库的选择

- (1) 一般小型的 C/S 程序多使用 Access 数据库;
- (2) 大型的 C/S 和 B/S 程序多使用 SQL Server、Oracle、DB2 数据库;
- (3) 小型的网站多采用开源的 MySQL 数据库 (Apache+MySQL+PHP)。

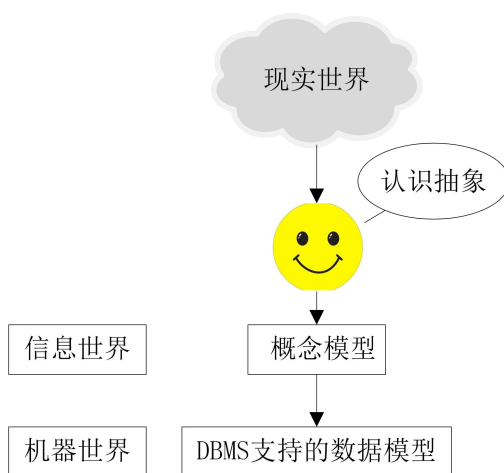
## 第二节 数据模型

### 一、数据模型

#### （一）基本概念

数据模型应满足三方面要求：一是能比较真实的模拟现实世界；二是容易为人所理解；三是便于在计算机上实现。

为了满足这三个方面的要求，人们将现实世界中的具体事物抽象为信息世界的某一种信息结构，这种信息结构并不依赖于任务的计算机系统，它不是某一个 DBMS 支持的数据模型，而是概念模型；然后再把概念模型转换为计算机上的某一 DBMS 所支持的数据模型。其过程如下：



#### （二）术语对照

现实世界	信息世界	机器世界
对象	实体	记录
特征	属性	字段
（对象）总体	实体集	二维表
关键特征	键	关键字
事物及其联系	概念模型	数据模型

### 二、概念模型

这是一种独立于任何计算机系统的模型，完全不涉及信息在计算机系统中的应用，用于建立信息世界的数据库模型，是现实世界的第一层抽象，是用户和数据库设计人员进行交流的工具，其中最著名的模型是“实体联系模型”（E-R 模型）。

E-R 模型直接从现实世界中抽取出实体类型及实体间联系图（E-R 图）表示数据模型。一般遇到实际问题时，总是先设计一个 E-R 模型，然后再把 E-R 模型转换成与 DBMS 关联的数据模型。

#### （一）E-R 简介

##### 1. E-R 模型三种图素



用矩形框表示实体

用菱形框表示实体之间的联系

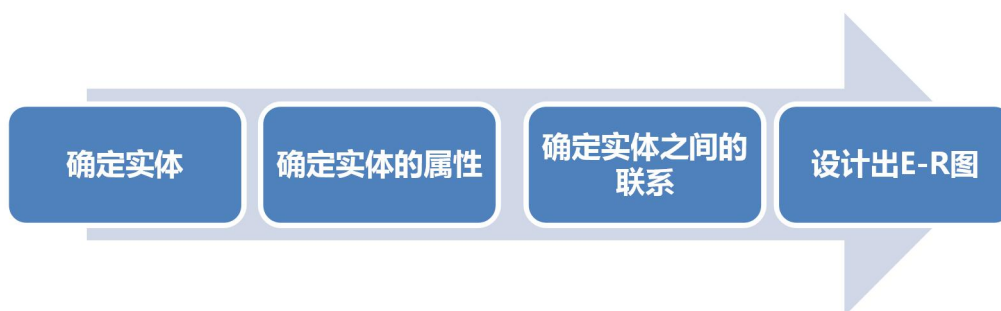
用椭圆形框表示属性

## 2.两个实体型之间的联系

有如下 3 种类型：

类型	概念
一对一联系	实体集 A 中的一个实体至多与实体集 B 中的一个实体相对应，反之亦然，则称实体集 A 与实体集 B 为一对一的联系。
一对多联系	实体集 A 中的一个实体与实体集 B 中的多个实体相对应，反之，实体集 B 中的一个实体至多与实体集 A 中的一个实体相对应。
多对多联系	实体集 A 中的一个实体与实体集 B 中的多个实体相对应，反之，实体集 B 中的一个实体与实体集 A 中的多个实体相对应。

## 3.E-R 模型作图



E-R 模型	将 E-R 模型转换成关系数据模型的步骤
	第一步：实体→二维表 模型中有两个实体，“职工”和“产品”所以转成 2 张表： 职工表（ <u>职工号</u> ，姓名，年龄） 产品表（ <u>产品号</u> ，产品名，价格）
	第二步：实体间联系→二维表 负责（职工号，产品号）
	第三步：合并 合并方案 1：“负责”与“职工”合并： 职工（ <u>职工号</u> ，姓名，年龄，产品号） 产品（ <u>产品号</u> ，产品名，价格） 合并方案 2：“负责”与“产品”合并： 职工（ <u>职工号</u> ，姓名，年龄）

	产品（ <u>产品号</u> ，产品名，价格，职工号）
	第一步：实体→二维表 学生（ <u>学号</u> ，姓名，班级） 课程（ <u>课程号</u> ，授课老师，课程名）
	第二步：实体间联系→二维表 选课（ <u>学号</u> ， <u>课程号</u> ，成绩）
	第三步：合并 “选课”和“学生”、“课程”间是多对多关系，不能合并

## （二）最常用的数据模型

数据模型	说明	结构示意图
层次模型	用树型结构来表示各类实体以及实体间的联系，在层次模型中结点表示实体集，连线表示相连两个实体间的联系，这种联系只能是一对多的。	
网状模型	像一张网，网上结点之间是平等的，结点之间是多对多的关系。	
关系模型	用满足一定条件的二维表来表示实体及实体之间的联系。	

数据模型	说明	结构示意图
面向对象模型	是数据库技术与面向对象程序设计方法结合的新一代数据模型，它能适应更复杂数据的管理。	

### 三、关系模型基础

关系数据库是应用关系数据模型来建立和处理数据库中的数据。这其中主要涉及几个重要的概念。

#### （一）关系模型的概念

概念	定义
关系	二维表在关系数据库中称作关系
字段	二维表中的列称为字段
记录	二维表中行称为记录，可以用关键字来标识（从第二行开始）
值	值是二维表中的一个具体数据项，是数据库中最小的数据单位
关键字	关键字也称作键，是指能唯一标识一条记录的字段
域	域是属性所有可能取值的集合
关系模式的结构	关系名（属性名 1，属性名 2，...，属性名 n）

书 号	书 名	作 者	价 格	出版日期	是否借出
20061001	计算机图形学	朱铨	¥10.00	2007-3-22	<input type="checkbox"/>
20061002	操作系统	朱铨	¥7.00	2007-3-23	<input type="checkbox"/>
20061003	ASP 网络编程	朱铨	¥14.00	2007-3-24	<input type="checkbox"/>
20061004	现代数学方法选讲	谢季坚	¥10.00	2007-3-25	<input type="checkbox"/>

关键字（主键）  
 第一行为：字段名  
 第二行开始每一行为一条记录  
 每一列为一个字段

#### （二）关系模型的键码

键码	定义
候选码（Candidate Key）	在关系中能唯一标识元组的属性或属性集
主码（primary key）	如果一个关系中有多个候选键，可以从中选择一个作为查询、插入或删除元组的操作变量，被选用的候选键称为主码，或称主键、主关系键、关系键、关键字
主属性（Prime Attribute）与非码属性（Non-Prime Attribute）	主属性：包含在主码中的各属性称为主属性；非码属性：不包含在任何候选码中的属性称为非码属性
外码（Foreign key）	若模式 R 中某属性集是其他模式的候选键，那么该属性集对模

	式 R 而言就是外码或外部关键字
--	------------------

## 四、关系模型

关系模型是目前最流行的一种数据模型，它是用二维表格结构表示实体集，关键码表示实体间的联系。

关系模型有 3 个部分构成：

数据结构：关系模型采用的数据结构是关系。

关系操作：关系模型提供一组完备的关系运算，以支持对数据库的各种操作。关系运算的理论是关系代数和关系演算。

关系的完整性：在关系模型中，数据的约束条件通过三类完整性约束条件来描述。即：

（1）实体完整性

要求关系中的元组的主键值不能是空值。

（2）参照完整性

要求在关系中不允许引用不存在的实体。

（3）用户定义的完整性

这是针对某一具体数据的约束条件，由应用环境决定，例如属性的值限制。



## 第三节 关系数据库

### 一、基本概述

关系数据库是应用关系数据模型来建立和处理数据库中的数据。这其中主要涉及几个重要的概念。

#### （一）关系模型

概念	定义
关系	二维表在关系数据库中称作关系
字段	二维表中的列称为字段
记录	二维表中行称为记录，可以用关键字来标识（从第二行开始）
值	值是二维表中的一个具体数据项，是数据库中最小的数据单位
关键字	关键字也称作键，是指能唯一标识一条记录的字段
域	域是属性所有可能取值的集合
关系模式的结构	关系名（属性名 1，属性名 2，...，属性名 n）

书号	书名	作者	价格	出版日期	是否借出
20061001	计算机图形学	朱铨	¥10.00	2007-3-22	<input type="checkbox"/>
20061002	操作系统	朱铨	¥7.00	2007-3-23	<input type="checkbox"/>
20061003	ASP 网络编程	朱铨	¥14.00	2007-3-24	<input type="checkbox"/>
20061004	现代数学方法选讲	谢季坚	¥10.00	2007-3-25	<input type="checkbox"/>

#### （二）关系模型的键码

键码	定义
候选码（Candidate Key）	在关系中能唯一标识元组的属性或属性集
主码（primary key）	如果一个关系中有多个候选键，可以从中选择一个作为查询、插入或删除元组的操作变量，被选用的候选键称为主码，或称主键、主关系键、关系键、关键字
主属性（Prime Attribute）与非码属性（Non-Prime Attribute）	主属性：包含在主码中的各属性称为主属性；非码属性：不包含在任何候选码中的属性称为非码属性
外码（Foreign key）	若模式 R 中某属性集是其他模式的候选键，那么该属性集对模式 R 而言就是外码或外部关键字

### 二、关系模型

关系模型是目前最流行的一种数据模型，它是用二维表格结构表示实体集，键码表示实体间的联系。

关系模型由 3 个部分构成：

数据结构：关系模型采用的数据结构是关系。

关系操作：关系模型提供一组完备的关系运算，以支持对数据库的各种操作。关系运算的理论是关系代数和关系演算。

关系的完整性：在关系模型中，数据的约束条件通过三类完整性约束条件来描述。即：

### （一）实体完整性

要求关系中的元组的主键值不能是空值。

### （二）参照完整性

要求在关系中不允许引用不存在的实体。

### （三）用户定义的完整性

这是针对某一具体数据的约束条件，由应用环境决定，例如属性的值。

## 三、关系代数

关系查询语言根据其理论基础的不同分成两大类：

（一）关系代数语言：查询操作是以集合操作为基础的运算。

（二）关系演算语言：查询操作是以谓词演算为基础的运算。

其中，关系代数是集合代数为基础发展起来的，它是以关系为运算对象的一组高级运算的集合。

关系代数的运算可分为两类：

1.基本运算操作：并、差、笛卡尔积、投影和选择。

2.组合运算操作：交、连接、自然连接和除。

以下对几种常用的关系运算作一个简单的介绍。

由于传统的集合运算，只是从行的角度进行，而要灵活地实现关系数据库多样的查询操作，必须引入专门的关系运算。专门的关系运算：不仅涉及行运算，也涉及列运算，这种运算是为数据库的应用而引进的特殊运算。包括选择、投影、连接和除法等运算。

### （一）投影

该操作是对关系进行垂直分割，消去某些列，并重新安排列的顺序，再删去重复元组（列的运算）。

例如：

<u>A</u>	<u>B</u>	<u>C</u>		<u>A</u>	<u>B</u>
b	2	z		b	2
b	8	y		b	8
a	5	x		a	5
关系R					

### （二）选择

这个操作是根据某些条件对关系作水平分割，即选择符合条件的元组。条件可用命题公式  $F$  表示， $F$  中的运算对象是常数（用引号括起来）或元组分量（属性名或列的序号）。运算符有算术比较运算符（ $\leq$ ,  $<$ ,  $\geq$ ,  $>$ ,  $=$ ,  $\neq$ ）和逻辑运算符（ $\wedge$ ,  $\vee$ ,  $\neg$ ）（行的运算）。例如：

<u>A</u>	<u>B</u>	<u>C</u>		<u>A</u>	<u>B</u>	<u>C</u>
b	2	z				
b	8	y				
a	5	x		b	2	z
关系R				$\sigma_{A=b \wedge B=2}(R)$		

### （三）自然连接

自然连接要求两个关系中进行比较的分量必须有相同的属性组，并且在结果中把重复的属性除去。

例如：

<u>A</u>	<u>B</u>	<u>C</u>		<u>B</u>	<u>C</u>	<u>D</u>	<u>E</u>		<u>A</u>	<u>B</u>	<u>C</u>	<u>D</u>	<u>E</u>
a1	b1	c2		b1	c2	d1	e1						
a2	b2	c1		b3	c1	d2	e2						
a3	b3	c3		b3	c3	d3	e3		a1	b1	c2	d1	e1
a4	b3	c5		b2	c2	d4	e4		a3	b3	c3	d3	e3
a5	b4	c1		b3	c1	d5	e5						
关系R				关系S					自然连接				

## 第四节 SQL 语言

### 一、SQL 基本概述

结构化查询语言 SQL(Structured Query Language)是关系数据库的标准语言。

SQL 是与 DBMS 进行通信的一种工具，它与 DBMS 的其他组件组合在一起，使用户可以方便地进行数据库的管理和数据的操作。

#### （一）SQL 语言的功能

##### 1.数据定义

SQL 能让用户自己定义所存储数据的结构，以及数据各项之间的关系。

##### 2.数据更新

SQL 提供了对数据的删除、修改和添加的更新操作。

##### 3.数据查询

SQL 提供了根据用户需要查询数据的功能。

##### 4.数据安全

SQL 提供访问、添加数据等操作的权限控制。

##### 5.数据完整性

SQL 可以定义约束规则，定义的规则存储在数据库内部。

##### 6.数据库结构的修改

SQL 允许用户通过数据库或应用程序修改数据库的结构。

#### （二）SQL 语言的特点

SQL 之所以能够使众多用户或者企业所接受，其根本原因就是它功能的强大和优点的突出。

##### 1.一体化

SQL 集 DDL、DML、DCL（数据控制语言）等功能于一体，语言风格统一，为数据库应用系统开发提供了良好的环境。

##### 2.面向集合的操作方式

SQL 采用集合操作方式，其查找结果和插入、删除、更新操作的对象都是元组集。

##### 3.高度非过程化

SQL 是一种描述型语言，用户只需提出“做什么”，而不必指明“怎么做”

##### 4.一套语法、二种使用方式

SQL 既是自含式语言，又是嵌入式语言。

##### 5.语言简明、易学易用

SQL 功能极强，但由于设计巧妙，语言十分简洁，完成数据定义、数据操纵、数据控制的核心功能只用了九个动词。

### 二、SQL 语句

功能	语句
----	----

创建数据库	CREATE DATABASE database-name
删除数据库	drop database dbname
备份数据库	BACKUP DATABASE pubs TO testBack
创建新表	create table tabname(col1 type1 [not null] [primary key],col2 type2 [not null],..)
根据旧表创建新表	create table tab_new like tab_old
删除表	drop table tabname
添加主键	Alter table tabname add primary key(col)
创建索引	create [unique] index idxname on tabname(col...)
创建视图	create view viewname as select statement
选择	select * from table1 where 范围
删除	delete from table1 where 范围
更新	update table1 set field1=value1 where 范围
查找	select * from table1 where field1 like '%value1%'
排序	select * from table1 order by field1,field2 [desc]
总数	select count * as totalcount from table1
求和	select sum(field1) as sumvalue from table1
平均	select avg(field1) as avgvalue from table1
最大	select max(field1) as maxvalue from table1
最小	select min(field1) as minvalue from table1

## 第五节 ACCESS 数据库基础

### 一、ACCESS 基本概述

#### （一）基本概念

ACCESS 是一种小型数据库管理系统，其数据库文件的扩展名为\*.mdb。启动与退出的方法与 Word 的方法是相同的。

ACCESS 数据库是由表、查询、窗体、数据访问页、报表、宏和模块七个对象组成的。

##### 1.表

一个 Access 数据库中 can 存储多张表，表之间还可以建立关系，以便把不同表中的数据联系起来。最多可以同时打开 1024 个表。

##### 2.查询

查询是 Access 数据库中的一种重要的对象，它是一种虚拟表，即不会用来存储数据，而是按照一定的条件或准则从一个或从个数据表中映射出的虚拟视图。查询对象为用户更方便地查看、分析和更改数据库中的数据提供了一种直观的视图。查询最多可从 16 个表中选择数据，字段最多为 255 个。

##### 3.窗体

窗体为用户查看和编辑数据库中的数据提供了一种友好的交互式界面。在 Access 2003 中，用户可以使用各种图形化的工具和向导快速地制作出用来显示和操作数据的窗体。

##### 4.报表

报表为打印输出数据库中的数据或数据的处理结果提供了一种便捷的方式。用户可以将一个或多个表和查询中的数据以一定的格式制作成报表，还可以将数据处理的结果或各种图表插入到报表中。

##### 5.页

页是 Access 2003 中唯一一个独立于 Access 数据库文件之外的对象。用户可以直接打开数据访问页，并通过它对数据库中的数据进行各种操作。

##### 6.宏

宏是一种为实现较复杂的功能而建立的可定制的对象，它实际上是一些列操作的集合，其中每个操作都能实现特定的功能。

##### 7.模块

模块是 Access 数据库中最复杂也是功能最强大的一种对象。在 Access 2003 中，使用其内置的 Visual Basic for Application 来建立和编辑模块对象，一个模块对象一般是一组相关功能的集合。分为标准模块、类模块和窗体模块。

#### （二）数据类型

数据类型	使用对象	大小
文本	文本或文本与数字的组合。Access 只保存输入到字段中的字符，不保存文本字段中未用位置上的空字符。	最长为 255 个字符
备注	保存长度较长的文本及数字。	最长为 64000 个字符
数字	可用来进行自述计算的数字数据，设置“字段大小”属性，定义一个特定的数字类型。	1、2、4 或 8 个字节，与“字段”大小的属性

		定义有关
<b>日期/时间</b>	日期及时间。	8 个字符
<b>货币</b>	货币值。使用货币数据类型可以避免计算时四舍五入引起的计算误差。精确度为小数点左方 15 位数及右方 4 位数。	8 个字符
<b>自动编号</b>	在添加记录时自动插入的唯一顺序（每次递增 1）或随机编号。	4 个字节
<b>是/否</b>	只包含两种值的一种。	1 位
<b>OLE 对象</b>	在其他使用 OLE 协议程序创建的对象。	最大为 1GB
<b>超级链接</b>	保存超级链接的字段。	最长为 64000 个字符
<b>查阅向导</b>	创建字段，该字段将允许使用组合框来选择另一个或一个列表中值。	通常为 4 个字节

## 二、Access 基本操作

### （一）启动与退出

#### 1.启动 Access

- （1）开始菜单->程序->Microsoft Office 找到 Access，单击即可。
- （2）桌面上的 Access 快捷方式，双击打开。
- （3）在“我的电脑”或“资源管理器”窗口中，双击 Access 的可执行文件。

#### 2.退出 Access

- （1）单击 Access 工作区窗口标题栏右端的关闭按钮。
- （2）单击“文件”菜单中的“退出”选项。
- （3）使用 Alt+F4 组合键。
- （4）双击菜单左上角的 Access 图标。

### （二）创建空数据库的方法

创建数据库的步骤主要有三步：创建空数据库→建立数据表结构→输入记录数据。

- （1）使用“向导”方式创建数据库。
- （2）进入 ACCESS 系统后，利用“文件”菜单下的“新建”。
- （3）单击数据库工具栏“新建”按钮。

### （三）创建数据表

#### 1.创建数据表的方法

- （1）使用设计器创建数据表。
- （2）使用向导创建表。
- （3）通过输入数据创建表。

#### 2.创建数据表的操作步骤

- （1）设置数据表结构（字段名、属性，主键）。
- （2）输入记录数据。

#### 3.修改表结构

- （1）打开数据库窗口，单击“对象”下的“表”选项卡，在数据库对象列表中选中对象，单击数据库窗口的“设计”按钮。
- （2）打开数据库窗口，单击“对象”下的“表”选项卡，在数据库对象列表中选中对象，并点击

鼠标右键，在弹出的快捷菜单中选择“设计”命令。

#### （四）查询

##### 1.定义

查询是指在数据库中查找满足条件的记录项。

##### 2.查询分类

Access 一共有 5 种查询类型：“选择查询”、“参数查询”、“交叉表查询”、“操作查询”和“SQL 查询”。

（1）选择查询：是最常见的查询类型，它从一个表或多个表中检索数据，并按照用户所需要的排列次序以数据表的方式显示结果。还可以使用“选择查询”来对记录进行分组，并且对记录进行总计、计数、平均值以及其他类型的总和计算。

（2）参数查询：在执行时会显示一个对话框，要求用户输入参数，系统根据所输入的参数找出符合条件的记录。

（3）交叉表查询：交叉表查询显示来源于表中某个字段的汇总值（合计、计算以及平均等），并将它们分组，一组行在数据表的左侧，一组列在数据表的上部。

（4）操作查询：操作查询是在一个记录中更改许多记录的查询，查询后的结果不是动态集合，而是转换后的表。它有 4 种类型：生成表查询、追加查询、更新查询和删除查询。

（5）SQL 查询：是用户使用 SQL 查询语句创建的查询。

##### 3.查询与筛选的区别

（1）查找范围不同：查询能在整个数据库中查找，而筛选仅能在当前数据表中查找。

（2）显示的字段不同：查询能显示指定字段，而筛选则显示当前数据表中的所有字段。

##### 4.注意事项

在多表查询时，要先建立好数据表之间的关系，否则在显示结果时将会出现数据的冗余显示的现象。



#### （五）统计

要对数据库里的数据实现统计功能，可以使用 Access 提供“合计”功能来实现，常见的合计功能有：分组、求和、平均、最小值、最大值、统计记录数、按条件统计。

类型	作用
求和（Sum）	计算字段中值的总和
最小值（Min）	搜索该字段的最小数值
条数（Count）	计算记录条数
平均（Avg）	计算平均值



<b>最大值 (Max)</b>	搜索该字段的最小数值
<b>条件 (Where)</b>	设置计算条件
<b>分组 (Group By)</b>	按某一字段将记录进行分组

#### (六) 数据的导入、链接与导出

数据操作	说明
<b>导入</b>	将外部数据转换并复制到当前数据库中，可被 Access 利用的文件类型有： <b>mdb、dbf、db、xls、txt</b>
<b>链接</b>	引用外部数据到当前数据库中
<b>导出</b>	将数据输出到其他数据库或其他应用程序当中

## 第六节 数据库设计

### 一、数据库设计阶段

按照规范设计的方法，考虑数据库及其应用系统开发全过程，将数据库设计分为以下六个阶段：

设计阶段	主要功能
需求分析阶段	需求收集和分析，结果得到数据字典描述的数据需求(和数据流程图描述的处理需求)。
概念结构设计阶段	通过对用户需求进行综合、归纳与抽象，形成一个独立于具体 DBMS 的概念模型，可以用 E-R 图表示。
逻辑结构设计阶段	将概念结构转换为某个 DBMS 所支持的数据模型(例如关系模型)，并对其进行优化。
数据库物理设计阶段	为逻辑数据模型选取一个最适合应用环境的物理结构(包括存储结构和存取方法)。
数据库实施阶段	运用 DBMS 提供的数据库语言(例如 SQL)及其宿主语言(例如 C)，根据逻辑设计和物理设计的结果建立数据库，编制与调试应用程序，组织数据入库，并进行试运行。
数据库运行和维护阶段	数据库应用系统经过试运行后即可投入正式运行。在数据库系统运行过程中必须不断地对其进行评价、调整与修改。

设计一个完善的数据库应用系统不可能一蹴而就，它往往是上述六个阶段的不断反复。

### 二、关系数据库范式理论

#### （一）函数依赖

数据依赖是现实世界中属性间联系和约束的抽象，是数据的内在性质。

函数依赖(functional dependency, FD)是一种最重要、最基本的数据依赖。其具体定义如下：

设有关系模式  $R(U)$ ， $X$  和  $Y$  是属性集  $U$  的子集， $FD$  是行为  $X \rightarrow Y$  的一个命题，只要  $r$  是  $R$  的关系，对  $r$  中任意两个元组都有“ $X$  值相等蕴涵  $Y$  值相等”，那么函数依赖  $X \rightarrow Y$  在关系模式  $R(U)$  中成立。

$FD$  与候选键之间的关系：若存在  $X \rightarrow U$ ，并且不存在  $X$  的任意真子集  $X_1$ ，使得  $X_1 \rightarrow U$  成立，那么就称  $X$  为关系的一个候选键。

删除操作异常是指执行删除操作时将不应该删除的数据删除的情形；插入异常是指执行插入操作时应该插入的数据无法插入的情形。

为了使数据库设计的方法走向完备，人们研究了规范化理论，指导我们设计规范的数据库模式。按属性间依赖情况来区分，关系规范化的程度为第一范式、第二范式、第三范式、BCNF 范式等。

#### （二）范式

范式（Normal Form, NF）是衡量关系模式的优劣的标准。范式有很多种，与数据依赖有着直接的联系。

##### 1. 第一范式

第一范式规定表的每个列的值都是不可分的简单数据项。在任何一个数据库中，第一范式都是一个最基本的要求。例如：表 1 不符合第一范式，表 2 符合第一范式。

表 1

学号	姓名	性别	联系电话
----	----	----	------

			宿舍电话	家庭电话
20161108	刘庆	男	80687614	85987328
20161209	崔清波	男	80687615	85984328
20161310	李易	女	80687616	85988345

表 2

学号	姓名	性别	宿舍电话	家庭电话
20161108	刘庆	男	80687614	85987328
20161209	崔清波	男	80687615	85984328
20161310	李易	女	80687616	85988345

## 2.第二范式

第二范式有两项要求：

(1) 所有表必须符合第一范式；

(2) 表中每一个非主键所在的列都必须完全函数依赖于主键（关键字）。

注意：表的不完全函数依赖关系将导致很多问题的发生。例如：

①数据冗余：如果外语系有 700 个学生，就得重复 700 次相同系名。

②更新异常：如果“外语系”升格成了“外语学院”，那么所有原外语系学生的元组都要更新，很难保证修改不出错误。

③插入异常：如果增加了一个新系，但该系还没有学生，那么这个系将不能入数据库。

例如：表 3 不符合第二范式，可以将表 3 拆分成表 4 和表 5。

表 3

学号	姓名	性别	课程号	分数
20161108	刘庆	男	01	90
20161209	崔清波	男	01	85
20161209	崔清波	男	02	88
20161310	李易	女	02	90

表 4

学号	课程号	分数
20161108	01	90
20161209	01	85
20161209	02	88
20161310	02	90

表 5

学号	姓名	性别
20161108	刘庆	男
20161209	崔清波	男
20161310	李易	女

### 3.第三范式

第三范式有两项要求：

（1）所有表必须符合第二范式。

（2）表中每一个非主键所在的列对主键都不存在传递依赖，而应是直接依赖。

传递依赖的存在与违背第二范式一样，也将会导致添加、删除、更新、冗余等问题。

表中出现上述问题的原因是对主键依赖的传递，解决的方法同样是在遵守无损分解的条件下，将表分解成多个表，从而消除传递依赖的情况。例如：表 6 不符合第二范式，可以拆分成表 7 和表 8。

表 6

学号	系名	宿舍楼
20161108	中文	6 号
20161209	计算机	7 号
20161310	外语	8 号
20161311	外语	8 号

表 7

学号	系名
20161108	中文
20161209	计算机
20161310	外语

表 8

学号	宿舍楼
20161108	6 号
20161209	7 号
20161310	8 号

### 4.BC 范式 BCNF

若关系模式 R 属于 1NF，且每个属性都不传递依赖于 R 的候选关键字，则称 R 属于 BC 范式。

由上可知，4 种范式之间的关系：BCNF<3NF<2NF<1NF。