# 第一章

数据库系统概述

数据库的基本概念：DB、DBMS、DBS、DBA

数据管理的发展:人工管理、文件系统和数据库系统

数据库管理系统功能 数据库定义功能；数据组织、存储和管理；数据操纵功能。

据库事务和运行管理；数据库的建立和维护功能。

数据库系统的结构 数据库系统三级模式结构：模式、内模式和外模式

数据库系统的三级模式结构

模式（逻辑模式）

数据库中全体数据的逻辑结构和特征的描述；所有用户的公共数据视图，综合了所有用户的需求；

一个数据库只有一个模式

内模式（存储模式）：是数据物理结构和存储方式的描述；是数据在数据库内部的表示方式

一个数据库只有一个内模式

外模式（子模式或用户模式）：数据库用户使用的局部数据的逻辑结构和特征的描述

数据库用户的数据视图，是与某一应用有关的数据的逻辑表示

一个数据库可以有多个外模式。

数据库系统的二级映象

三级模式是对数据的三个抽象级别，二级映象在DBMS内部实现这三个抽象层次的联系和转换

外模式／模式映象

1. 定义外模式与模式之间的对应关系 2. 保证数据的逻辑独立性

模式／内模式映象

1. 定义了数据全局逻辑结构与存储结构之间的对应关系。 2. 保证数据的物理独立性

数据库系统的特点 数据结构化 数据的共享性高，冗余度低，易扩充 数据独立性高

数据由 DBMS统一管理和控制

数据模型的分两类： 概念模型、逻辑模型和物理模型

数据模型的三要素： 数据结构、数据操作、数据的完整性约束

三种主要数据模型：关系模型、层次模型、网状模型

第二章

关系模型 由关系数据结构、关系操作和关系完整性约束三部分组成。

关系数据结构

关系 二维表，属性是列 ，元组是行

关系模式 对关系的描述 R(U,F)

关系数据库 关系的集合

关系的码

候选码（CK） 关系中能唯一标识一个元组的属性组，称为该关系的候选码

简单情况: 候选码只包含一个属性。

极端情况: 关系的所有属性是关系模式的候选码，称为全码（All-key）

主码（Pk） 若一个关系有多个候选码，则选定其中一个为主码

候选码的诸属性称为主属性。

不包含在任何侯选码中的属性称为非主属性。

外码（FK） 设F是关系R的一个或一组属性，但不是关系R的码。如果F与关系S的主码Ks相对应，则称F是关系R的外码

关系R称为参照关系 关系S称为被参照关系

选修关系的“学号” 与学生关系的主码“学号”相对应

选修关系的“课程号”与课程关系的主码“课程号”相对应

“学号”和“课程号”是选修关系的外码

学生关系和课程关系均为被参照关系

选修关系为参照关系

学生（学号，姓名，性别，年龄） 课程(课程号，课程名，学分)

选修（学号，课程号，成绩）

关系的数据操作

关系数据操作语言

关系代数

传统的集合运算 并、差、交、笛卡尔积

∪ － ∩ ×

专门的关系运算 选择、投影、连接、除

σ (选择) 水平方向选取子集

π (投影）垂直方向选取子集

(连接) 从笛卡尔积中选取子集

÷ （除法）同时从水平、垂直方向选取子集

关系演算 元组演算，域演算

关系数据操作 查询 数据更新 插入、删除、修改

关系操作的特点：一次一集合方式

(操作的对象和结果都是集合)

关系的三类完整性规则 关系的完整性规则是对关系的约束条件

实体完整性 关系的主码不能取空值。

参照完整性 外码必须是另一个表中主码的有效值，或者是空值。

用户定义的完整性

是针对某一具体数据库的约束条件。它反映某一具体应用所涉及的数据必须满足的语义要求。

实体完整性和参照完整性是关系模型必须满足的完整性约束条件，被称作是关系的两个不变性。

第三章

1.定义表

CREATE TABLE <表名> （

<列名1> <类型>[ <列级约束> ,]

<列名2> <类型>[ <列级约束> ,] …………… ，

[<表级完整性约束条件> ] ）；

例 建立 “学生选课表” SC，它由学号Sno、课程号Cno，成绩Grade组成，其中(Sno, Cno)为主码。

CREATE TABLE SC(

Sno CHAR(9) ,

Cno CHAR(3) ,

Grade int CHECK (Grade>=0 and Grade<=100),

Primary key (Sno, Cno),

);

2. 查询语句

SELECT [<DISTINCT>] <目标列表 >

FROM <表 名 >

WHERE <条件>；

[GROUP BY <列名> ...]

[HAVING <条件表达式>]

[ORDER BY <列名> [ASC|DESC] … ]；

常用查询：

(1) 单表查询 (2) 连接查询 (3) 嵌套查询 (4) 集合查询

(1)单表查询一般形式

SELECT <目标列表 >

FROM <表 名 >

WHERE <条件>；

例1：查询所有年龄在20岁以下的学生姓名及其年龄。

SELECT Sname，Sage

FROM Student

WHERE Sage < 20；

例2：查询选修了课程的学生学号。

SELECT DISTINCT Sno

FROM SC;

例3：查询年龄在20岁以下的姓王的学生姓名及其年龄。

SELECT Sname，Sage

FROM Student

WHERE Sage<20 AND

Sname LIKE ‘王%’;

(2) 连接查询一般形式

SELECT <目标列表 >

FROM <表 名1 > , <表 名2 > ……

WHERE <连接条件>

AND <查询条件>；

外链接查询一般形式

SELECT <目标列表 >

FROM <表 名1 > JOIN <表 名2 > ON <连接条件>

WHERE <查询条件>；

例1：查询选修2号课程且成绩在90分以上的所有学生的学号、姓名

SELECT Student.Sno, student.Sname

FROM Student, SC

WHERE Student.Sno=SC.Sno AND /\* 连接条件\*/

SC.Cno= ‘2’ AND /\* 其他限定条件 \*/

SC.Grade > 90； /\* 其他限定条件 \*/

(3) 嵌套查询一般形式

SELECT－ FROM－ WHERE <比较符/ 谓词>

SELECT－ FROM－ WHERE <比较符/ 谓词>

SELECT－ FROM－ WHERE <比较符/ 谓词>

嵌套查询分类

不相关子查询 子查询的查询条件不依赖于父查询

相关子查询 子查询的查询条件依赖于父查询

例1：不相关子查询 常用IN嵌套

查询所有选修了1号课程的学生姓名。

SELECT Sname /\*外层查询/父查询\*/

FROM Student

WHERE Sno IN

（SELECT Sno /\* 内层查询/子查询\*/

FROM SC

WHERE Cno= ' 1 '）；

例2： 相关子查询,常用EXISTS嵌套

查询所有选修了1号课程的学生姓名。

SELECT Sname

FROM Student

WHERE EXISTS

(SELECT \*

FROM SC /\*相关子查询\*/

WHERE SC.Sno=Student.Sno AND Cno= ‘1’)；

3 .聚集函数

COUNT( ) SUM( ) AVG( ) MAX( ) MIN( )

※ 聚集函数经常在select目标列中或在Group子句的Having条件中使用；

※ 聚集函数不能作为Where子句中的条件表达式

例：查询每个系学生的平均成绩。

SELECT Sdept，AVG(Grade)

FROM Student，SC

WHERE student.Sno=SC.Sno

GROUP BY S.Sdept;

4. 数 据 更 新 （插入数据）

INSERT

INTO <表名> [(<列1> [，<列2 > …)]

VALUES (<常量1> [,<常量2>] … );

例： INSERT

INTO SC (Sno，Cno，Grade)

VALUES (‘ 20095020 ’，‘ 1 ’ ,90)；

注：sql99还支持插入子查询

5. 建立视图

CREATE VIEW <视图名> [(<列名> ，…)]

AS <子查询> ；

例：建立考试成绩为优的学生的成绩单视图，包括学生的学号、课程名和成绩。

CREATE VIEW 成绩单

AS

SELECT SC.Sno，Course.name，SC.Grade

FROM SC , Course

WHERE SC.Cno=Course.Cno AND

SC.Grade>=90；

注：组成视图的属性列名或者全部省略或者全部指定，没第三种选择。

第五章

数据库的安全性是指保护数据库，防止用户非法使用、破坏数据库

数据库的完整性是数据的正确性和相容性。

完整性约束条件作用的对象可以是列、元组、关系

关系模型的完整性在CREATE TABLE语句中定义

例：定义“学生选课表” SC表中 的完整性。

CREATE TABLE SC(

Sno CHAR(9) ，

Cno CHAR(4) ，

Grade int CHECK (Grade>=0 and Grade<=100),

Primary key (Sno, Cno)，

Foreign KEY（Sno） REFERENCES Student (Sno),

Foreign KEY（Cno） REFERENCES Course (Cno));

DBMS的完整性控制机制具有3方面功能

定义功能。提供定义完整性约束条件的机制

检查功能。提供完整性检查的方法，检查用户发出的操作请求是否违背了完整性约束条件。

违约处理。用户的操作请求违背完整性约束条件时DBMS应采取一定的动作来保证数据的完整性。

第六章 关系数据理论

1. 规范化的基本思想

关系数据库的规范化理论是数据库逻辑设计的工具。

一个关系只要其分量都是不可分的数据项，它就是规范化的关系，但这只是最基本的规范化。

采用“一事一地”的模式设计原则，让一个关系描述一个概念、一个实体或者实体间的一种联系。 若多于一个概念就把它“分离”出去。

2. 数据依赖

一个关系内部属性与属性之间的约束关系 是现实世界属性间相互联系的抽象

是数据内在的性质 是语义的体现

3. 数据依赖的类型

函数依赖（ 简记为 FD） 多值依赖（ 简记为 MVD） 其他依赖

4.数据依赖

设R(U)是一个属性集U上的关系模式，X和Y是U的子集。

若对于R(U)的任意一个可能的关系r，r中不可能存在两个元组在X上的属性值相等， 而在Y上的属性值不等， 则称 “X函数确定Y” 记作X→Y。

设K为关系模式R<U,F>中的属性或属性组合。

若K ｆ U，则K称为R的一个侯选码(CK)。

5.关系模式规范化过程

1NF（关系中每一属性是不可再分的数据项）

↓ 消除非主属性对码的部分函数依赖

消除决定属性 2NF（关系中每一非主属性都完全函数依赖码）

集非码的非平 ↓ 消除非主属性对码的传递函数依赖

凡函数依赖 3NF（关系中每一非主属性都不传递依赖码）

↓ 消除主属性对码的部分和传递函数依赖

BCNF(关系中每一主属性都不部分、传递依赖码)

↓ 消除非平凡且非函数依赖的多值依赖

4NF （关系中不存在非平凡、非函数的多值依赖）

↓ 消除不为候选码的连接依赖

5NF

例：

1. 在关系模式R（A，B，C，D）中，存在函数依赖关系{A→B，A→C，A→D，（B，C）→A}，则候选码是A ，（B，C） ，关系模式R 属于2NF 。

2.关系模式R中的属性全是主属性，则R的最高范式是（D）。

A.1NF B.2NF C.3NF D.BCNF

3. 设有关系模式R（S，D，M），F＝{S→D，S→M}，则关系模式R的规范化程度最高达到（ D ）。

A.1NF B.2NF C.3NF D.BCNF

第七章

1. 数据库的设计过程

需求分析（现实世界的需求） 概念结构设计（信息世界模型）

逻辑结构设计（机器世界模型） 物理设计 数据库实施 数据库运行维护

需求分析和概念设计独立于任何数据库管理系统，逻辑设计和物理设计与选用的DBMS密切相关。设计一个完善的数据库应用系统往往还会有许多反复。

2. 数据库各级模式的形成

需求分析阶段 综合分析用户的应用需求。

概念设计阶段 形成独立于机器特点、独立于DBMS产品的概念模式(E-R图)。

逻辑设计阶段 将E-R图转换成具体的数据库产品支持的数据模型如关系模型，形成数据库逻辑模式。然后根据用户处理的要求，再建立必要的视图（VIEW），形成数据库的外模式 。

物理设计阶段 根据DBMS特点和处理的需要，进行物理存储安排，设计索引，形成数据库内模式。

3.需求分析

（1）任务 充分了解原系统工作概况，明确用户的各种需求， 在此基础上确定新系统的功能。

（2）重点 收集与分析用户在数据管理中的信息要求、处理要求、安全性与完整性要求。

（3）方法 自顶向下的结构化分析方法（简称SA方法）并用数据流图和数据字典描述系统。

（4）数据字典的内容 数据项 数据结构 数据流 数据存储 处理过程

4.概念结构设计（ E-R图设计）

概念结构是对现实世界的一种抽象

从实际的人、物、事和概念中抽取所关心的共同特性，忽略非本质的细节

把这些特性用各种概念精确地加以描述，这些概念组成了E-R模型

概念模型的三种常用数据抽象:分类、聚集、概括

分类 （E-R模型中的实体）

定义某一类概念作为现实世界中一组对象的类型

聚集 （E-R模型中的属性）

定义某一类型的组成成分

概念模型的表示方法 (E-R方法)

用E-R图来描述现实世界的概念模型（也称为E-R模型）

实体 用矩形表示，矩形框内写明实体名。

属性 用椭圆形表示，并用无向边将其与相应的实体连接起来

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，同时在边旁标上联系的类型（1:1、1:n或 m:n）

联系的属性 联系本身也是一种实体型，也可以有属性

概念结构设计的步骤

（1）抽象数据并设计局部视图

标定局部应用中的实体、属性、码、实体间的联系,

用E-R图描述出来

（2）集成局部视图，得到全局概念结构,消去冲突（属性冲突、命名冲突、结构冲突），形成整体E-R图

（3）验证整体概念结构

5.逻辑结构设计的步骤

(1) 将概念结构转化为一般的关系（网状、层次）模型

(2) 将转化来的关系模型（网状、层次）向特定DBMS支持下的数据模型转换

(3) 对数据模型进行优化

6. E-R图向关系模型的转换原则

⒈ 一个实体型转换为一个关系模式。关系的码就是实体型的码。

⒉一个1:1联系可以与任意一端对应的关系模式合并。每个实体的码均是该关系的候选码。

⒊ 一个1:n联系 可以与n端对应的关系模式合并。合并后n端关系的码不变， 1端关系的码是它的外码。

⒋一个m:n联系必须转换为一个独立的关系模式。 该关系的码是两端实体的码的组合。

（1:1联系、 1:n联系也可以转换为一个独立的关系模式。）

7. 数据库的物理设计

数据库在物理设备上的存储结构与存取方法称为数据库的物理结构，它依赖于给定的计算机系统。

为一个给定的逻辑数据模型选取一个最适合应用环境的物理结构的过程，就是数据库的物理设计。

物理设计的步骤

(1) 确定数据库的物理结构.

(2) 对物理结构进行分析评价。（评价的重点是时间和空间效率）

物理设计的内容

（1） 为关系模式选择存取方法(建立存取路径)

（2） 设计关系、索引等数据库文件的物理存储结构

例：某工厂有多个部门，每一个部门有多名职工，且每一名职工只属于一个部门；工厂承揽多项工程，每一项工程有多名职工参加，每一名职工可参加多项工程。

解：

（1）步 抽象数据并设计局部视图

① 部门由 部门号，部门名称，电话…………等数据描述；

② 职工由 职工号，姓名，性别，年龄 ………等数据描述；

③ 工程由 工程号，工程名，经费…………等数据描述；

（2）步 集成局部视图，得到全局概念结构(整体E-R图)

该工厂的E-R模型

（3）步 验证整体概念结构

（4）转换成关系模型：

① 部门（部门号，部门名称，电话）

② 工程（工程号，工程名，部门号，经费）

③ 职工（职工号，姓名，性别，年龄，部门号）

④ 施工（工程号，职工号，工时）

（5） 定义该工厂关系模型中部门关系、工程关系的结构。

解：部门关系的结构：

CREATE TABLE 部门（

部门号 CHAR(10) PRIMARY KEY,

部门名称 CHAR(20) NOT NULL,

电话 CHAR(12))；

工程关系的结构：

CREATE TABLE 工程（

工程号 CHAR(7) PRIMARY KEY,

工程名 CHAR(20) NOT NULL,

部门号 CHAR(10)，

经费 NUMERIC(7,2)，

FOREIGN KEY (部门号) REFERENCES 部门(部门号));

如果数据库只包含成功事务提交的结果，就说数据库处于一致性状态

保证数据一致性是对数据库的最基本的要求。

事务是数据库的逻辑工作单位。

DBMS保证一切事务的原子性、一致性、隔离性和持续性

恢复中最经常使用的技术：数据库转储和登记日志文件

恢复的基本原理：利用存储在后备副本、日志文件和数据库镜像中的冗余数据来重建数据库

常用恢复技术

事务故障的恢复 UNDO

系统故障的恢复 UNDO + REDO

介质故障的恢复 重装备份并恢复到一致性状态 + REDO

提高恢复效率的技术

检查点技术 可以提高系统故障的恢复效率

镜像技术 可以改善介质故障的恢复效率

数据库的并发控制以事务为单位，多事务执行方式有：

(1)串行执行 (2)交叉并发 (3)同时并发

事务的ACID特性

原子性、一致性、 隔离性、持续性

数据库并发控制的主要技术是封锁

最常用的基本封锁是排它锁(X锁)和共享锁(S锁)。

数据库的价值在很大程度上取决于它所能提供的数据共享度

数据共享在很大程度上取决于系统允许对数据并发操作的程度

数据并发程度取决于数据库中的并发控制机制

并发事务操作是否正确的判别准则是可串行性

并发操作的正确性则通常由两段锁协议来保证。

两段锁协议是可串行化调度的充分条件但不是必要条件 事务遵守两段锁协议 事务可串行化的调度

冲突可串行化调度是可串行化调度的充分条件，但不是必要条件。

冲突可串行化调度 可串行化的调度

对数据对象施加封锁，可能引起死锁、活锁问题

避免活锁方法：先来先服务

解决死锁的方法

死锁预防

一次封锁法

顺序封锁法

死锁的诊断与解除

超时法

等待图法

一、单选题

1. SQL语言的一次查询结果是一个( D ) 。

A. 数据项 B.记录 C. 元组 D. 表

2.若要撤销数据库中已经存在的表S，可用（C ）。

A.DELETE TABLE S B.DELETE S

C.DROP TABLE S D.DROP S

3.描述数据库全体数据的全局逻辑结构和特性的是（B ） 。

A.模式 B.内模式 C.外模式 D. 三层模式

4.关系数据库中的投影操作是指从关系中（ B ）。

A.选取特定记录 B.选取特定字段

C.建立相应的影像 D.建立相应的图形

5.下列四项中，不属于数据库系统特点的是（ C ） 。

A.数据共享 B.数据完整性

C.数据冗余度高 D.数据独立性高

二、填空题

1.数据库的完整性是指数据的\_\_\_\_\_\_和 \_\_\_\_\_\_。

2.关系代数运算中，专门的关系运算有\_\_\_\_、\_\_\_\_和连接。

3. 数据模型的三要素是指 ， ， 。

4.实体之间的联系有\_\_\_\_\_\_\_\_、\_\_\_\_\_\_\_\_ 和 三种。

5.人们已经提出了多种类型的数据依赖，其中最重要的是\_\_\_\_\_\_\_和\_\_\_\_\_\_\_\_。

三、求解题

针对教材第3章（P82）学生—课程数据库，完成下列各题：

1.查询年龄在20～23岁之间的学生的姓名、系别和年龄。

2.查询选学了课程号为‘2’的学生姓名和所在系。

3.查询年龄小于20岁的所有学生学号、姓名和年龄，查询结果按年龄升序排列。

4. 删除Student中姓’李’的所有学生记录。

5. 建立考试及格学生的成绩视图(名为‘成绩单’)，包括姓名、课程名和成绩。

四、综合题

1.简述数据库设计的主要过程。

答：1) 需求分析

2) 概念结构设计

3) 逻辑结构设计

4) 数据库物理设计

5) 数据库实施

6) 数据库运行和维护

这是一个完整的实际数据库及其应用系统的设计过程。不仅包括设计数据库本身，还包括数据库的实施、数据库运行和维护。设计一个完善的数据库应用系统往往是上述六个阶段的不断反复

2. 简述数据库系统的主要特点。

一、数据结构化数据库系统实现整体数据的结构化

二、数据的共享性高，冗余度低，易扩充数据库的数据不再面向某个应用而是面向整个系统，因此可以被多个用户、多个应用、用多种不同的语言共享使用。由于数据面向整个系统，是有结构的数据，不仅可以被多个应用共享使用，而且容易增加新的应用，这就使得数据库系统弹性大，易于扩充。

四、数据由DBMS统一管理和控制数据库的共享是并发的共享，即多个用户可以同时存取数据库中的数据甚至可以同时存取数据库中同一个数据。为此，DBMS必须提供统一的数据控制功能，包括数据的安全性保护，数据的完整性检查，并发控制和数据库恢复。

三、数据独立性高数据独立性包括数据的物理独立性和数据的逻辑独立性。数据库管理系统的模式结构和二级映象功能保证了数据库中的数据具有很高的物理独立性和逻辑独立性。

3.简述关系模型完整性约束规则。

关系模型的完整性规则是对关系的某种约束条件。关系模型中可以有三类完整性约束：实体完整性、参照完整性和用户定义的完整性。其中实体完整性和参照完整性是关系模型必须满足的完整性约束条件，被称作是关系的两个不变性，应该由关系系统自动支持。

1) 实体完整性规则：若属性A是基本关系R的主属性，则属性A不能取空值。

2) 参照完整性规则：若属性（或属性组）F是基本关系R的外码，它与基本关系S的主码Ks相对应（基本关系R和S不一定是不同的关系），则对于R中每个元组在F上的值必 须为： ? 或者取空值（F的每个属性值均为空值）； ? 或者等于S中某个元组的主码值。

3) 用户定义的完整性是针对某一具体关系数据库的约束条件。它反映某一具体应用所涉及的数据必须满足的语义要求。

4.简述DBMS完整性控制机制的基本功能。

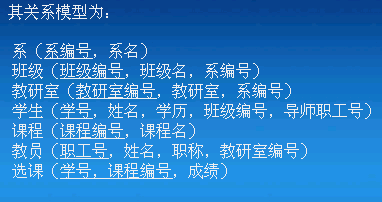
**答：**RDBMS的完整性控制机制应具有三个方面的功能：

（1）定义功能。即提供定义完整性约束条件的机制。

（2）检查功能。即检查用户发出的操作请求是否违背了完整性约束条件。

（3）违约处理。如果发现用户的操作请求使数据违背了完整性约束条件，则采取一定的动作来保证数据的完整性。

5. 针对教材第7章（P234）习题12、18、19。完成概念模型设计再转换成关系模型，并定义表的结构和完整性约束。



**学校**

**属于**

**系**

**组建**

**教研室**

**课程**

**设置**

**教员**

**指导**

**选课**

**组成**

**班级**

**有**

**学生**

n

k

n

s

m

n

m

m

1

1

1

1

1

1



转换成关系模型：

产品（产品号,产品名称）

零件（零件号，零件名，零件量）

产品组成（产品号，零件号，使用零件量）

零件存储(零件号，仓库号，储存量)

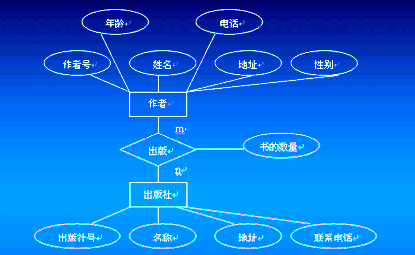
原材料（原材料号，原材料名，类别，仓库号）

FK=仓库号，参照仓库关系

仓库（仓库号，仓库名）

零件制造（零件号，原材料号，使用原材料量）

**第七章习题18**



**转换成关系模型：**

作者（**作者号**,姓名，年龄，性别，电话，地址）

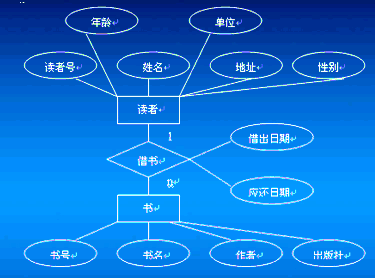
出版社（**出版社号**，名称，地址，联系电话）

出版（**作者号，出版社号**，书的数量）

PK=（**作者号，出版社号**）

FK=作者号，参照作者关系，FK=出版社号，参照出版社关系

**第七章习题19**



**转换成关系模型：**

读者（**读者号**,姓名，年龄，性别，单位，地址）

书（**书号**，书名，作者，出版社）

借书（**书号**，**读者号** ，借出日期，应还日期）

PK=（**书号**，**读者号）**

FK=书号，参照书关系，FK=读者号，参照读者关系