



／
／
／
期
末
不
挂
科

考
前
必
看

数 据 库

期末不挂科 考前必看

2.5h速成课 90+高分秘诀

期末加油站出品



扫码关注公众号

观看更多精彩视频

版权声明

关于合肥科思通途教育科技有限公司作品版权声明：

合肥科思通途教育科技有限公司是“期末加油站系列课程”的合法创作者和所有者，包括但不限于视频内容、音频、图像、讲义、标识、版权标记和其他相关材料。

未经我方明确书面许可，禁止任何个人、机构或实体复制、分发、展示、公开演示、制作衍生作品、搬运或以其他方式未经授权使用“期末加油站系列课程”视频。我方保留对任何侵犯其版权的行为采取法律行动的权利，凡侵犯课程等知识产权的个人、机构，我方将依据国家法律法规，严肃追责，绝不姑息（包括但不限于索赔、起诉和/或追究其法律责任）。

感谢各位小伙伴们对期末加油站的信任与支持，期待在未来的学习旅程中，与大家共同进步。

特此声明！

合肥科思通途教育科技有限公司

2023年11月29日



《数据库系统》讲义

第一章 数据库系统概述

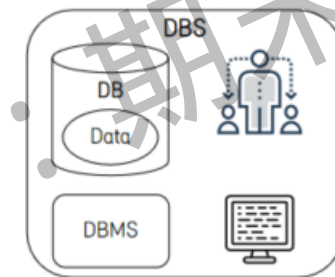
1.1 基本概念

1.1.1 知识点

基本概念：

- 数据（Data）：描述事物的符号记录
- 数据库（DataBase，DB）：长期储存在计算机内、有组织的、可共享的大量数据的集合
- 数据库管理系统（DataBase Management System，DBMS）：位于用户与操作系统之间的一层数据管理软件（计算机基础软件）
- 数据库系统（DataBase System，DBS）：由数据库、数据库管理系统（及其应用开发工具）、应用程序和数据库管理员组成的存储、管理、处理和维持数据的系统

上述几个数据概念之间的关系如下：



数据库的基本特征包括永久存储、有组织、可共享、冗余度小、易扩展、数据独立。

数据库的功能包括数据定义语言（DDL）、数据操纵语言（DML）、数据组织存储和管理、数据库的建立和维护、数据库的事务管理和运行管理

数据库管理技术的三个发展阶段：

1.人工管理阶段：在计算机出现之前，人们运用常规的手段从事记录、存储和对数据加工，也就是利用纸张来记录和利用计算工具（算盘、计算尺）来进行计算，并主要使用人的大脑来管理和利用这些数据。

特点:(1)计算机系统不提供对用户数据的管理功能；(2)数据不能共享；(3)不保存数据。

2.文件系统阶段：

（20 世纪 50 年代后期至 60 年代中期）外部存储器已有磁盘、磁鼓等直接存取的存



储设备；操作系统中的文件系统是专门管理外存的数据管理软件。

特点:数据以“文件”形式可长期保存在外部存储器的磁盘上。由于计算机的应用转向信息管理，因此对文件要进行大量的查询、修改和插入等操作，这些操作由文件系统提供。

缺点:数据冗余;不一致性;数据独立性差。

3.数据库系统阶段:

20 世纪 60 年代后期以来，出现了数据库这样的数据管理技术，满足和解决实际应用中多个用户、多个 应用程序共享数据的要求

特点:(1) 数据结构化。(2) 数据共享性高、冗余少且易扩充。(3) 数据独立性高。(4) 数据由 DBMS 统一管理和控制。

1.1.2 例题

1、DB、DBMS 和 DBS 三者之间的关系是_____。

A、DBS 与 DB 和 DBMS 无关

B、DBMS 包括 DBS 和 DB

C、DB 包括 DBMS 和 DBS

D、DBS 包括 DB 和 DBMS

解: D 数据库系统 (DataBase System) 是由数据库、数据库管理系统 (及其应用开发工具)、应用程序和数据库管理员组成的存储、管理、处理和维护数据的系统

2、数据库系统的基本特征是_____。

A、数据的统一控制

B、数据共享性和统一控制

C、数据共享性、独立性和冗余度小

D、数据共享性和数据独立性

解: C 数据库的基本特征包括永久存储、有组织、可共享、冗余度小、易扩展、数据独立。

1.1.3 课后习题

4、数据库管理系统能实现对数据库中数据的查询、插入、修改和删除等操作，这种功能称为_____。

A、数据操纵功能 B、数据管理功能

C、数据定义功能 D、数据控制功能

解: A

5、数据库系统的数据独立性是指_____。

A、不会因为数据的变化而影响应用程序



B、不会因为系统数据存储结构与数据逻辑结构的变化而影响应用程序

C、不会因为存储策略的变化而影响存储结构

D、不会因为某些存储结构的变化而影响其他的存储结构

解：B

6、数据库管理系统的主要功能？

解：数据库管理系统（DBMS）是位于操作系统与用户之间的一个数据管理软件，它主要功能包括以下几个方面：

数据定义功能 DBMS 提供数据描述语言（DDL），用户可通过它来定义数据。

数据操纵功能 DBMS 还提供数据操纵语言（DML），实现对数据库的基本操作：查询、插入、删除和修改。

数据库的运行管理 这是 DBMS 运行时的核心部分，它包括开发控制，安全性检查，完整性约束条件的检查和执行，数据库的内容维护等。

数据库的建立和维护功能 它包括数据库初始数据的输入及转换，数据库的转储与恢复，数据库的重组功能和性能的监视与分析功能等。

1.2 数据模型

1.2.1 知识点

数据模型的三个组成元素：

- 1、**数据结构**描述数据库的组成对象以及对象之间的联系
- 2、**数据操作**是指对数据库中各种对象的实例允许执行的操作的集合，包括操作及其有关的规则，主要分为查询和更新（插入、删除、修改）两大类操作。
- 3、**数据完整性约束条件**是给定的数据模型中数据及其联系所具有的制约和依存规则，与语义相关。例如，某大学的数据库中规定学生成绩如果有 6 门以上不及格则不能授予学士学位等。

数据模型的分类：

- 1、**概念模型**：按用户的观点来对数据和信息建模，主要用于数据库设计。概念模型的表示方法很多，最常用的为实体-联系方法(Entity-Relationship approach)，该方法用 E-R 图来描述概念模型。
- 2、**逻辑模型**：主要用于数据库管理系统的实现，包括层次模型、网状模型、关系模型、面向对象模型和对象关系数据模型等。

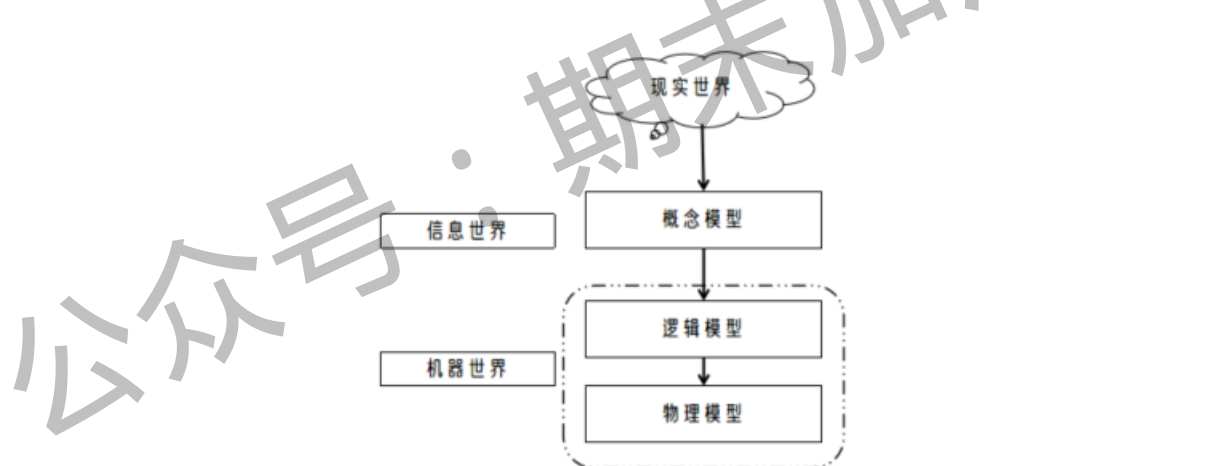


(1) 非关系模型/格式化模型：层次模型(Hierarchical Model)、网状模型(Network Model)。满足下面两个条件的基本层次联系的集合为层次模型（就是树结构）：有且只有一个结点没有双亲结点，这个结点称为根结点；根以外的其它结点有且只有一个双亲结点。满足下面两个条件的基本层次联系的集合（一对多关系，结构上像有向图）：允许一个以上的结点无双亲；一个结点可以有多个的双亲。

(2) 关系模型(Relational Model)：规范化的（二维表），包括关系（一个关系对应通常说的一张表）、属性（表中的一列即为一个属性）、域（属性的取值范围）、元组（表中的一行即为一个元组）、码（表中的某个属性组，它可以唯一确定一个元组）、分量（元组中的一个属性值）、关系模式（对关系的描述，一般表示为 关系名（属性 1，属性 2，.....，属性 n））。

3、物理模型：对数据最底层的抽象，它描述数据在系统内部的表示方法和存取方法，在磁盘或磁带上的存储方式和存取方法，是面向计算机系统的。

三个模型之间的关系：



其中，现实世界到概念模型的映射由数据库设计人员完成，概念模型到逻辑模型的映射由数据库设计人员采用数据库设计工具协助完成，逻辑模型到物理模型的映射由DBMS 完成。

1.2.2 例题

3、关系模型和网状模型的数据结构分别是_____和_____。

解：二维表，有向图

4、以下是信息世界的模型，且实际上是现实世界到机器世界的一个中间层次的是（ ）。



A、数据模型 B、概念模型 C、关系模型 D、E-R 图

解：B

1.2.3 课后习题

1、数据库常用的数据模型有_____、层次模型、_____。

解：网状模型、关系模型

2、数据模型是由_____、_____、和_____三部分组成的。

解：数据结构、数据操作、数据完整性约束条件

3、构成 E—R 模型的三个基本要素是_____。

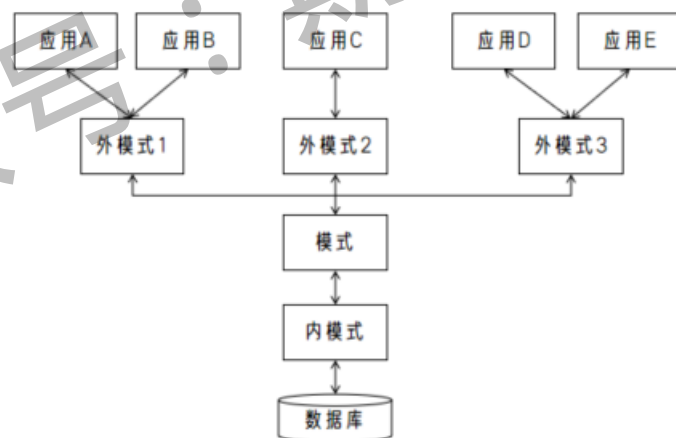
A、实体、属性、联系； B、实体、属性值、关系；
C、实体、实体集、联系； D、实体、实体集、属性

解：A

1.3 数据库系统结构

1.3.1 知识点

数据库三级模式结构：模式（Schema）、外模式（External Schema）、内模式（Internal Schema）



模式（也称逻辑模式）：①数据库中全体数据的逻辑结构和特征的描述；②所有用户的公共数据视图，综合了所有用户的需求。模式是数据库系统模式结构的中间层，与数据的物理存储细节和硬件环境无关，与具体的应用程序、开发工具及高级程序设计语言无关。

外模式（也称子模式或用户模式）：①数据库用户（包括应用程序员和最终用户）使用的局部数据的逻辑结构和特征的描述；②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。外模式的用途：①保证数据库安全性的一个有力措施 ②每个用户只能看见和访问所对应的外模式中的数据

内模式（也称存储模式）：①是数据物理结构和存储方式的描述②是数据在数据库内部的表示方式，包括记录的存储方式（顺序存储，按照 B 树结构存储，按 hash 方法存储）、索引的组织方式、数据是否压缩存储、数据是否加密和数据存储记录结构的规定。

内模式和模式之间的映射关系称为“内模式/外模式映像”，一一对应关系。外模式和模式之间的映射关系称为“外模式/模式映像”，多对一关系，任意一个外模式变化后，模式需要对应更新。

1.3.2 例题

5、在数据库的三级模式结构中，用来描述数据库中全体数据的全局逻辑结构和特征的是_____。

解：模式/逻辑模式

1.3.3 课后习题

7、数据库系统中有哪三层模式结构？采用三层模式结构有什么好处？

解：数据库系统三级模式结构是指数据库系统是由外模式、模式和内模式三级构成。

数据库系统的三级模式是对数据的三个抽象级别，它把数据的具体组织留给 DBMS 管理，使用户能逻辑地抽象地处理数据，而不必关心数据在计算机中的具体表示方式与存储方式。为了能够在内部实现这三个抽象层次的联系和转换，数据库管理系统在这三级模式之间提供了两层映像：外模式 / 模式映像，模式 / 内模式映像，正是这两层映像保证了数据库系统中的数据能够具有较高的逻辑独立性和物理独立性。



第二章 关系数据库

2.1 数据结构及形式化定义

2.1.1 知识点

关系数据库的数据结构是关系，一张扁平的二维表

关系数据库的相关形式化定义：

1、域（D）：域是一组具有相同数据类型的值的集合

2、笛卡尔积：笛卡尔积是域上面的一种集合运算，笛卡尔积中每个元素称为一个 n 元组或简称**元组**，笛卡尔积可表示为一个二维表，表中每行对应一个元组，元组中的每一个值称为**分量**，表中每一列的值来自一个域。

3、关系（Relation）： $D_1 \times D_2 \times \dots \times D_n$ 的子集叫作在域 D_1, D_2, \dots, D_n 上的关系，表示为 $R(D_1, D_2, \dots, D_n)$ R ：关系名； n ：关系的目或度（Degree）。

- 关系也是一个二维表，每行对应一个元组，每列对应一个域，每列称为**属性**
- 若关系中的某一属性组的值能唯一地标识一个元组，则称该属性组为**候选码**（Candidate key）。
- 若一个关系有多个候选码，则选定其中一个或多个为**主码**（Primary key）。
- 候选码的多个属性称为主属性（Prime attribute），不包含在任何候选码中的属性称为非主属性
- 最简单的情况下，候选码只包含一个属性，如果关系模式的所有属性都是这个关系的候选码，则称为**全码**（All-key）

2.1.2 例题

无

2.1.3 课后习题

无

2.2 关系操作

2.2.1 知识点

关系操作中包含：

1、查询操作：选择（Select）、投影（Project）、连接（Join）、除（Divide）、并（Union）、差（Except）、交（Intersection）、笛卡尔积。选择、投影、并、差、笛卡尔积是 5 种基本操作。



2、插入、删除、修改操作

关系数据语言包括三种：

- 1、关系代数语言
- 2、关系演算语言：元组关系、域关系
- 3、结构化查询语言（SQL）

关系模型的完整性规则是对关系的某种约束条件，关系模型中有三类完整性约束：

- 实体完整性（Entity Integrity）：关系模型以主码作为唯一性标识，主码中的属性即主属性不能取空值
- 参照完整性（Referential Integrity）：外码，反映和其他表之间的关联关系
- 用户自定义完整性（User-defined Integrity）：反映某一具体应用所涉及的数据必须满足的语义要求

2.2.2 例题

1、关系数据语言可以分为三类：_____、_____和关系演算语言。

解：关系代数语言、SQL

2、关系数据库中，实现表与表之间的联系是通过_____。

- A、实体完整性规则 B、值域
C、用户自定义完整性 D、参照完整性规则

解：D

2.2.3 课后习题

2、对关系的完整性约束通常包括_____三种。

- A、实体完整性、属性完整性、关系完整性；
B、实体完整性、参照完整性、用户定义完整；
C、实体完整性、属性完整性、用户定义完整；
D、实体完整性、属性完整性、参照完整性；

解：B

2.3 关系代数

2.3.1 知识点

关系代数的运算对象：关系

关系代数的运算结构：关系



关系代数的运算分类：

- 传统的集合运算：并 $R \cup S$ 、差 $R - S$ 、交 $R \cap S$ 、笛卡尔积 $R \times S$
- 专门的关系运算
 - 选择： $\sigma_F(R)$ 从关系 R 中选取使逻辑表达式 F 为真的元组，是从行的角度进行的运算
 - 投影： $\Pi_A R$ 从列的角度进行运算，但投影成功之后不仅取消了原关系中的某些列，而且还可能取消某些元组（避免重复行）
 - 连接： $R \bowtie_{A\theta B} S$ 从两个关系的笛卡尔积中选取属性间满足一定条件的元组
 - ◆ 自然连接：同名属性值相等；去除重复（悬浮元组、左外连接、右外连接、外连接）
 - ◆ 等值连接：属性值相等
 - 除： $R \div S$ 同时从行和列的角度进行运算的

两个关系 R 和 S 在做自然连接时，关系 R 中某些元组有可能在 S 中不存在公共属性上值相等的元组，从而造成 R 中这些元组在操作时被舍弃了，这些被舍弃的元组称为悬浮元组。

如果把悬浮元组也保存在结果关系中，而在其他属性上填充值(Null)，就叫做外连接；左外连接(LEFT OUTER JOIN 或 LEFT JOIN)：只保留左边关系 R 中的悬浮元组；右外连接(RIGHT OUTER JOIN 或 RIGHT JOIN)：只保留右边关系 S 中的悬浮元组

2.3.2 例题

3、在关系代数运算中，五种基本运算为_____。

- A、并、差、选择、投影、连接
- B、并、交、选择、投影、笛卡尔积
- C、并、差、选择、投影、笛卡尔积
- D、并、除、投影、笛卡尔积、选择

解：C 选择、投影、并、差、笛卡尔积是 5 种基本操作。

4、设关系 R 有 K_1 个元组，关系 S 有 K_2 个元组，则关系 R 和 S 进行笛卡尔积操作后的结果关系中的元组数目是_____。

- A、 K_1+K_2 B、 $\geq K_1+K_2$



C、 $K1 \times K2$ D、 $\leq K1 \times K2$

解：C

5、根据下面示例的关系 R 和 S，分别计算 $\sigma_{C>10}(R)$ ， $\Pi_B(R)$ ， $R \bowtie_{R.B=S.B} S$ ， $R \bowtie S$

A	B	C
a1	b1	5
a1	b2	6
a2	b3	8
a2	b4	12

B	E
b1	3
b2	7
b3	10
b3	2
b5	2

解：

(1) 是选择操作，选择 R 关系中 $C>10$ 的元组，可以得到结果：

A	B	C
a2	b4	12

(2) 是投影操作，投影 R 中 B 属性，可以得到结果：

B
b1
b2
b3
b4

(3) 是等值连接操作，不需要去除重复值，且需要保留两个关系中的属性列，可以得到结果：

A	R.B	C	S.B	E
a1	b1	5	b1	3
a1	b2	6	b2	7
a2	b3	8	b3	10
a2	b3	8	b3	2

(4) 是自然连接操作，需要去除重复值，仅保留一行，可以得到结果：

A	B	C	E
a1	b1	5	3



a1	b2	6	7
a2	b3	8	10
a2	b3	8	2

6、有学生-课程关系数据库，关系表描述如下，用关系代数表达式表示下列问题：

(1) 检索选修了教师王敏所授课程的学生姓名和学号

(2) 检索选修了课程号是 C3 或者 C5 的学生学号和成绩。

列名	说明	数据类型	约束
Sno	学号	字符(8)	主键
Sname	姓名	字符(12)	非空，唯一

Student

列名	说明	数据类型	约束
Cno	课程号	字符(8)	主键
Teacher	姓名	字符(12)	非空，唯一

Course

列名	说明	数据类型	约束
Sno	学号	字符(8)	主键，参照 Student 表
Cno	课程号	字符(12)	主键，参照 Course 表
Grade	成绩	短整型	0-100

SC

解：(1) $\Pi_{Sno, Sname}(Student \bowtie (\sigma_{Teacher='王敏'}(SC \bowtie Course)))$

(2) $\Pi_{Sno, Grade}(\sigma_{Cno='C3'}(SC)) \cup \Pi_{Sno, Grade}(\sigma_{Cno='C5'}(SC))$

2.3.3 课后习题

1、有学生-课程关系数据库，关系表描述如下，用关系代数表达式表示下列问题：

(1) 检索选修了教师李华所授课程的学生姓名和学号

(2) 检索选修了课程号是 C5 的学生学号、年龄和院系。

列名	说明	数据类型	约束
Sno	学号	字符(8)	主键
Sname	姓名	字符(12)	非空，唯一
Sage	年龄	整型	15-36
Sdept	院系	字符(20)	

Student

列名	说明	数据类型	约束
----	----	------	----



Cno	课程号	字符(8)	主键
Teacher	姓名	字符(12)	非空，唯一
Cname	课程名	字符(12)	非空

Course

列名	说明	数据类型	约束
Sno	学号	字符(8)	主键，参照 Student 表
Cno	课程号	字符(12)	主键，参照 Course 表
Grade	成绩	短整型	0-100

SC

解：(1) $\Pi_{Sno, Sname}(Student \bowtie (SC \bowtie (\sigma_{Teacher='李华'}(Course))))$

(2) $\Pi_{Sno, Sage, Sdept}(Student \bowtie (\sigma_{Cno='C5'}(SC)))$

3、自然连接是构成新关系的有效方法。一般情况下，当对关系 R 和 S 使用自然连接时，要求 R 和 S 含有一个或多个共有的_____。

A、行 B、属性 C、记录 D、元组

解：B



第三章 SQL 语言

3.1 基本概念

3.1.1 知识点

SQL：结构化查询语言 Structured Query Language；标准语言；关系数据库语言

功能：数据查询、数据操纵、数据定义和数据控制

特点：综合统一；高度非过程化；面向集合的操作方式；语言简洁、易学易用；一种语法结构多种使用方式

SQL 中的几种数据：

- (1) 基本表：本身独立存在的表
- (2) 视图：虚表，从一个或几个基本表（或视图）导出的表，数据库只存放视图的定义而不存放视图对应的数据，这些数据仍存放在原来的基本表中
- (3) 索引：关系数据库的内部实现技术，属于内模式的范畴。建立索引是加快查询速度的有效手段，用户（数据库管理员或建表者）可以在基本表上建立一个或多个索引，以提供多种存取路径，系统在存取数据时会自动选择合适的索引作为存取路径
 - 唯一索引：UNIQUE 关键字
 - 非唯一索引：不使用 UNIQUE 关键字
 - 聚簇索引：指索引项的顺序与表中记录的物理顺序一致的索引组织，一个表只能建立一个聚簇索引，一般默认是主键

SQL 对关系数据库的支持：

- (1) 一个关系对应一个基本表
- (2) 一个或多个基本表对应一个存储文件
- (3) 一个表可以带若干索引，索引在存储文件中
- (4) 外模式：若干视图和部分基本表
- (5) 模式：若干基本表
- (6) 内模式：若干存储文件

视图的作用

①简化用户的操作：用户所做的只是对一个虚表的简单查询，而这个虚表是怎样得来的，用户无须了解。



②使用户能以多种角度来看待同一数据：当许多不同种类的用户共享同一个数据库时，这种灵活性是非常有必要的。

③对重构数据库提供了一定的逻辑独立性：即使重构数据库也不一定需要修改应用程序。

④能够对机密数据提供安全保护：只允许用户查询提供给他视图而不是直接查询表，可以隐藏表中的机密数据

⑤适当的利用视图可以更清晰地表达查询语句：利用视图来表达查询语句

3.1.2 例题

1、SQL 的视图是从_____中导出的。

- A、基本表 B、视图
C、数据库 D、基本表或视图

解：D

2、SQL 语言中的“视图(VIEW)”对应于数据库系统三级模式结构中的_____。

- A、内模式 B、外模式
C、模式 D、都不对应

解：B

3.1.3 课后习题

3.2 数据定义

3.2.1 知识点

1、模式的创建和删除语法

操作方式	前置条件：权限；定义内容：命名空间
创建	create schema <模式名> authorization <用户名> [<表定义子句> <视图定义子句> <授权定义子句>]
删除	drop schema <模式名> <cascade restrict>

2、基本表的创建、修改和删除

操 作 方式	语 法	备 注



创建	<pre>create table <模式名.表名> (<列名><数据类型>[列级完整性约束条件],, <列名><数据类型>[列级完整性约束条件], [表级完整性约束条件])</pre>	数据类型： bit int decimal float varchar char date datetime bigint 约束类型： Check, Unique, Primary key, not null, Foreign key
修改	<pre>alter table <表名> [add [column] <新列名><数据类型>[完整性约束]] [add <表级完整性约束>] [drop [column] <列名> [cascade restrict]] [drop constraint <完整性约束名> [restrict cascade]] [alter column <列名><数据类型>]</pre>	1、增加新列； 2、增加表级完整性约束； 3、删除某一列； 4、删除完整性约束； 5、修改某一列。
删除	<pre>drop table <表名>[restrict cascade]</pre>	

3、索引的创建、修改和删除

操作方式	语法	备注
创建	<pre>create [unique] [cluster] index <索引名> on <表名>(<列名>[<次序>],.....)</pre>	次序：asc desc
修改	<pre>alter index <旧索引名> rename to <新索引名></pre>	
删除	<pre>drop index <索引名></pre>	

索引的几种类别：

- 顺序文件上的索引：按指定属性值升序或降序存储的关系
- B+树索引：索引属性组织为 B+树形式；动态平衡
- 散列索引：若干个桶，根据散列函数值映射到桶中；查找速度快
- 位图索引：位向量记录索引属性中可能出现的值

4、视图的创建和删除

操作方式	语法
创建	<pre>CREATE VIEW <视图名> [(<列名> [, <列名>]...)] AS <子查询> [WITH CHECK OPTION]</pre>
删除	<pre>drop view <视图名> [cascade]</pre>



3.2.2 例题

3、写出下列情况的 SQL 语言

- (1) 为用户 ZHANG 创建一个模式 TEST
- (2) 删除模式 ZHANG 以及该模式内的所有数据库对象

解：(1) `CREATE SCHEMA TEST AUTHORIZATION ZHANG;`
(2) `DROP SCHEMA ZHANG CASCADE;`

4、写出下列情况的 SQL 语言

- (1) 建立一个学生表 Student
- (2) 向 Student 表中增加“入学时间”列，数据类型为日期型
- (3) 将年龄的数据类型由字符型改为整数
- (4) 增加课程名为唯一值的约束条件
- (5) 删除 Student 表

解：(1) `create table Student(Sno char(9) primary key, Sname char(20) unique, Ssex char(2), Sage char(2), Sdept char(20));`
(2) `alter table Student add Sentrance date;`
(3) `alter table Student alter column Sage int;`
(4) `alter table Student add unique(Cname);`
(5) `drop table Student cascade;`

5、写出下列情况的 SQL 语言

- (1) 为学生表按照学号升序建立唯一索引
- (2) 将学生表的 SCno 索引名修改为 SCSno
- (3) 删除 Student 表的 Stusname 索引

解：(1) `create unique index SCno on Student(Sno asc);`
(2) `alter index SCno rename to SCSno;`
(3) `drop index Stusname;`

6、当对视图进行 UPDATE、INSERT 和 DELETE 操作时，为了保证被操作的行满足视图定义中子查询语句的谓词条件，应在视图定义语句中使用可选择项

解：with check option



3.2.3 课后习题

4、下列 SQL 语句中，修改表结构的是_____。

A、ALTER B、CREATE C、UPDATE D、INSERT

解：A

3.3 数据操作

3.3.1 知识点

1、查询操作：

查询语句的语法：根据 WHERE 子句的条件表达式，从 FROM 子句指定的基本表或视图中找出满足条件的元组，再按 SELECT 子句中的目标表达式，选出元组中的属性值形成结果表。如果有 GROUP BY 子句，则将结果按<列名 1>的值进行分组，该属性列值相等的元组称为一个组。如果 GROUP BY 子句带 HAVING 短语，则只有满足指定条件的组才予以输出。如果有 ORDER BY 子句，则结果表还要按<列名 2>的值的升序或降序排列

SELECT [ALL | DISTINCT] <目标列表达式>[别名][, <目标列表达式>] ...
FROM <表名或视图名>[别名]
[, <表名或视图名> [别名]]
(<select 语句> [AS] <别名>)] ...
[WHERE <条件表达式>]
[GROUP BY <列名 1> [HAVING <条件表达式>]]
[ORDER BY <列名 2> [ASC|DESC]];

查询条件：

查询条件	内容
比较	=, >, <, >=, <=, !=, <>, !>, !<, not + 运算符
范围	between ..and.., not between ..and..
集合	in, not in
字符匹配	like ‘<匹配串>’, not like ‘<匹配串>’ => 完整字符串；通配符（% 可以代表一个或多个字符；_ 只能代表一个字符）
空值判断	is null, is not null
逻辑运算	and, or, not

GROUPBY 子句将查询结果按某一列或多列的值分组，值相等的为一组；分组的目的是为了细化聚集函数的作用对象，如果未对查询结果分组，聚集函数将作用于整个查询结果分组后聚集函数将作用于每一个组，即每一个组都有一个函数值。聚合函数的类



别：

聚合函数	功能
COUNT (*)	统计元组个数
COUNT ([DISTINCT ALL] <列名>)	统计一列中值的个数
SUM ([DISTINCT ALL] <列名>)	计算一列值的总和（此列必须为数值型）
AVG ([DISTINCT ALL] <列名>)	计算一列值的平均值（此列必须为数值型）
MAX ([DISTINCT ALL] <列名>)	求一列中的最大值
MIN ([DISTINCT ALL] <列名>)	求一列中的最小值

若一个查询同时涉及两个以上的表，则称之为连接查询，连接查询的分类：

查询类型	特点
等值与非等值连接查询 (自然连接)	连接查询的 WHERE 子句中用来连接两个表的条件称为链接条件或连接谓词，其一般格式如下： [<表名 1>]<列名 1> <比较运算符> [<表名 2>]<列名 2> 运算符为 = 时称为等值连接，否则称为非等值连接 自然连接（在等值连接中把目标列中重复的属性列去掉）
自身连接	表与自己进行连接；需要为表取别名
外连接	LEFT OUT JOIN; RIGHT OUT JOIN
复合条件连接	INNER JOIN（内连接,或等值连接）：获取两个表中字段匹配关系的记录。 LEFT JOIN（左连接）：获取左表所有记录，即使右表没有对应匹配的记录。 RIGHT JOIN（右连接）：与 LEFT JOIN 相反，用于获取右表所有记录，即使左表没有对应匹配的记录。

where 语句的查询范围是由查询语句构造的查询语句称为嵌套查询，在子查询中，程序先运行在嵌套在最内层的语句，再运行外层。因此在写子查询语句时，可以先测试下内层的子查询语句是否输出了想要的内容，再一层层往外测试，增加子查询正确率。否则多层的嵌套使语句可读性很低。嵌套查询中子查询的 SELECT 语句中不能使用 ORDER BY 子句，ORDER BY 子句只能对最终查询结果排序。子查询的查询条件依赖于父查询，则称这类子查询为相关子查询，否则称不相关子查询。具体包括以下四类：

- 带有 IN 谓词的子查询：WHERE XXX IN (SELECT XXXXX)
- 带有比较运算符的子查询：WHERE XXX > (SELECT XXXXX)
- 带有 ANY(SOME)或 ALL 谓词的子查询：子查询如果返回多值，就需要使用 ANY 或 ALL 谓词，其语义为：>ANY 表示大于子查询结果中的某个值；>ALL 表示大于子查询结果中的所有值
- 带有 EXISTS 谓词的子查询：带有 EXISTS 谓词的子查询不返回任何数据，只产



生逻辑值” true” 或” false”

集合查询：多个 select 语句的结果进行集合查询；集合操作包括并操作 UNION、交操作 INTERSECT 和差操作 EXCEPT；参加集合操作的各查询结果的列数必须相同，对应项的数据类型也必须相同

2、插入操作：

插入：

元组：INSERT INTO <表名> [(<属性列 1>[, <属性列 2>...]) VALUES (<常量 1>[, <常量 2>] ...)

子查询结果：INSERT INTO <表名> [(<属性列 1>[, <属性列 2>...]) VALUES (子查询)

3、修改操作：

UPDATE <表名> SET <列名>=<表达式>[, <列名>=<表达式>]... [WHERE <条件>];

4、删除操作：删除数据，表定义依然存在

DELETE FROM <表名> [WHERE <条件>];

3.3.2 例题

7、有学生-课程关系数据库，各关系表描述如下，使用 SQL 语句实现下列问题

(1) 创建 Student 表

(2) 检索姓"李"且倒数第 2 个字为"华"或"涛"字的学生的姓名和学号

(3) 向 Student 表中插入元组 “S07060，王梅”，并写出该记录的详细信息。

(4) 删除选课成绩在 45 分以下的学生记录。

(5) 把“网络工程”课程成绩提高 8%。

(6) 求出女同学的每一年龄组（超过 4 人）有多少人？要求查询结果按人数升序排列，人数相同的按年龄降序排列

列名	说明	数据类型	约束
Sno	学号	字符(8)	主键
Sname	姓名	字符(12)	非空，唯一
Ssex	性别	字符(2)	男/女，默认女
Sage	年龄	短整型	16-40，默认 21
Sdept	院系	字符(20)	默认电子系

Student

Cno	Cname	Teacher
-----	-------	---------



C1	操作系统	李明
C2	大学物理	赵华
C3	电子商务	周亮

Course

Sno	Cno	Grade
S07001	C1	76
S07002	C2	82
S07003	C3	65

SC

解：（1）创建表

create table Student(Sno char(8) primary key,

Sname char(12) not null unique,

Ssex char(2) check (Ssex in ('男','女')) default '女',

Sage smallint check (Sage >=16 and Sage <=40) default 21,

Sdept char(20) default '电子系');

（2）select Sno,Sname from Student where Sname like '李%[华涛]_';

（3）insert into Student(Sno,Sname) values('S07060','王梅');

该记录的详细信息是（S07060,王梅,女,21,电子系）

（4）DELETE FROM Student WHERE SNO IN (SELECT SNO FROM SC WHERE GRADE<45);

（5）UPDATE SC SET Grade=Grade*1.08 WHERE Cno IN (SELECT Cno FROM Course WHERE Cname='网络工程');

（6）SELECT Sage, COUNT(Sno) FROM Student WHERE Ssex='女' GROUP BY Sage HAVING COUNT(Sno)>4 ORDER BY COUNT(Sno) ASC, Sage DESC;

8、用如下的 SQL 语句创建一个 Teacher 表

CREATE TABLE Teacher (TNO Char(6) NOT NULL,

NAME Char(8) NOT NULL,

SEX Char(2), SAL INT,

SDETP char(12) check (SDEPT IN('IS' , ' MA' , ' CS')))

可以插入到 Teacher 表中的元组是_____。

A、（ 'T0203' , '刘刚' , NULL, NULL, 'IS' ）



B、(NULL, '王芳', '女', '1200', 'CS')

C、('T0111', NULL, '男', 23, 'MA')

D、('T0101', '李华', 男, 2100, 'EN')

解：A

3.3.3 课后习题

1、SQL 语言中，条件年龄 BETWEEN 18 AND 30 表示年龄在 18 至 30 之间，且_____。

A、包括 30 岁但不包括 18 岁 B、包括 18 岁和 30 岁

C、包括 18 岁但不包括 30 岁 D、不包括 18 岁和 30 岁

解：B

2、允许取空值但不允许出现重复值的约束是_____。

A、NULL B、PRIMARY KEY

C、UNIQUE D、FOREIGN KEY

解：C

3、删除基本表用_____语句，删除基本表的元组用_____语句

解：drop, delete

5、有学生-课程关系数据库，各关系表描述如下，使用 SQL 语句实现下列问题

(1) 写出创建 Student 表的 SQL 语句

(2) 检索课程名以"DB_"开头，且倒数第 3 个字符为 k 的课程的具体情况：

列名	说明	数据类型	约束
Sno	学号	字符(8)	主键
Sname	姓名	字符(12)	非空，唯一
Ssex	性别	字符(2)	男/女，默认女
Sage	年龄	短整型	15-36，默认 22
Sdept	院系	字符(20)	默认计算机系

Student

列名	说明	数据类型	约束
Cno	课程号	字符(8)	主键
Cname	课程名	字符(8)	非空
Teacher	教师名	字符(10)	唯一

Course



列名	说明	数据类型	约束
Cno	课程号	字符(8)	主键，参照 Student 表，级联更新和删除
Sno	学号	字符(8)	主键，参照 Student 表，级联更新，拒绝删除
Grade	成绩	短整型	0-100

SC

- (1) 写出创建 Student 表的 SQL 语句
- (2) 检索课程名以"DB_"开头，且倒数第 3 个字符为 k 的课程的具体情况；
- (3) 向学生选修课程表中插入元组“学生 S10 选修课程号 C4”。
- (4) 删除选课成绩在 55 分以下的学生记录；
- (5) 把“高等数学”课程成绩提高 8%；
- (6) 求出女同学的每一年龄组（超过 3 人）有多少人？要求查询结果按人数降序排列，人数相同的按年龄升序排列；
- (7) 定义视图 S_age，其中包括学生学号以及每个学生选修课程的门数(要求成绩非空)和平均成绩；
- (8) 在 Course 表中增加一个类型为 char(10)的职称(pro)列；

解：（1）创建表

```
create table Student(Sno char(8) primary key,
Sname char(12) not null unique,
Ssex char(2) check (Ssex in ('男','女')) default '女',
Sage smallint check (Sage >=15 and Sage <=36) default 22,
Sdept char(20) default '计算机系');
```

- (2) select * from Course where Cname like 'DB_%k_';
- (3) INSERT INTO SC (Sno, Cno) VALUES('S10','C4');
- (4) DELETE FROM Student WHERE SNO IN (SELECT SNO FROM SC WHERE GRADE<55); 或者：

```
DELETE FROM Student WHERE Student.SNO = SC.SNO AND SC.GRADE<55 ;
```

- (5) UPDATE SC SET Grade=Grade*1.08 WHERE Cno IN (SELECT Cno FROM Course WHERE Cname='高等数学');

```
(6) SELECT Sage, COUNT(Sno) FROM Student WHERE Ssex='女' GROUP BY Sage
HAVING COUNT(Sno)>3 ORDER BY COUNT(Sno) DESC, Sage ASC;
```



```
( 7 ) CREATE VIEW S_age (Sno, Cno_num, Avg_grade) AS SELECT  
Sno,COUNT(Cno),AVG(grade) FROM SC WHERE grade IS NOT NULL GROUP BY Sno;
```

```
(8) alter table Course add pro char(10);
```

3.4 查询处理和优化

3.4.1 知识点

查询处理的步骤：

- 查询分析：判断是否符合 SQL 语法规则。
- 查询检查：根据数据字典对合法的查询语句进行语义检查
- 查询优化：从多种可供选择的执行策略中选择一个高效执行的查询处理策略
- 查询执行：代码生成器生成执行这个查询计划的代码

查询优化：用户不必考虑如何最好地表达查询以获得较好的效率，系统可以比用户程序的“优化”做得更好

- 代数优化：

关系代数表达式等价交换规则

查询树的启发式优化

- 物理优化：

基于启发式规则的存取路径选择优化

基于代价估算的优化

3.4.2 例题

无

3.4.3 课后习题

无



第四章 关系数据库理论

4.1 基本概念

4.1.1 知识点

数据依赖：一个关系内部属性与属性之间的一种约束关系（完整性约束的表现形式）；数据内在的性质；语义的体现。有多种类型的数据依赖，其中最重要的是函数依赖（Functional Dependency, FD）和多值依赖（Multivalued Dependency, MVD）

函数依赖 FD：自变量 x 确定后，相应的函数值 y 也随之确定；来源于现实世界

- 非平凡函数依赖： $X \rightarrow Y$, 但 $Y \not\subseteq X$
- 平凡函数依赖： $X \rightarrow Y$, 但 $Y \subseteq X$ （自己与自己的关系）
- 完全函数依赖： Y 完全依赖 X ， X 的任何一个真子集都不能推导 Y
- 部分函数依赖： Y 不完全依赖与 X

4.1.2 例题

1、在下列关于函数依赖的叙述中，不正确的是_____。

- A、若 $A \rightarrow B, A \rightarrow C$, 则 $A \rightarrow BC$
- B、若 $A \rightarrow C$, 则 $AB \rightarrow C$
- C、若 $A \rightarrow B, DB \rightarrow C$, 则 $DA \rightarrow C$
- D、若 $AB \rightarrow C$, 则 $A \rightarrow C, B \rightarrow C$

解：D

4.1.3 课后习题

4.2 范式

4.2.1 知识点

规范化理论：规范化理论是用来改造关系模式，通过分解关系模式来消除其中不合适的数据依赖，以解决数据冗余、插入异常、更新异常、删除异常这些问题。

规范化：一个低一级范式的关系模式通过模式分解可以转换为若干个高一级范式的关系模式的集合，这个过程称为规范化

$$1NF \supset 2NF \supset 3NF \supset 4NF \supset 5NF$$

1、第一范式：

如果一关系模式 $r(R)$ 的每个属性对应的域值都是不可分的(即原子的)，则称 $r(R)$ 属于第一范式



第一范式的目标是：将基本数据划分成称为实体集或表的逻辑单元，当设计好每个实体后，需要为其指定主码。

第一范式是对关系模式的最起码的要求,不满足第一范式的数据库模式不能称为关系数据库

2、第二范式：

若关系模式 $R \in 1NF$ ，并且每一个非主属性都完全函数依赖于任何一个候选码，则 $R \in 2NF$

第二范式的目标：将只部分依赖于候选码（即依赖于候选码的部分属性）的非主属性移到其他表中。

2NF 范式虽然消除了由于非主属性对候选码的部分依赖所引起的冗余及各种异常，但并没有排除传递依赖。

3、第三范式：

若 $R \in 3NF$ ，则每一个非主属性既不部分依赖于码也不传递依赖于码

第三范式的目标：去掉表中不直接依赖于候选码的非主属性

4、BC 范式

BCNF 是修正的第三范式，有时也称为扩充的第三范式

一个满足 BCNF 的关系模式有：

所有非主属性都完全函数依赖于每个候选码

所有的主属性都完全函数依赖于每个不包含它的候选码

没有任何属性完全函数依赖于非码的任何一组属性

BCNF 范式排除了：

任何属性(包括主属性和非主属性)对候选码的部分依赖和传递依赖；

主属性之间的传递依赖。

关系模式规范化的基本步骤：





4.2.2 例题

2、关于 BC 范式下列说法正确的是_____。

- A、如果 $R \in 3NF$ ，则 R 一定是 BCNF
- B、若 $R \in 3NF$ ，且不存在主属性对非码的函数依赖，则其是 BCNF
- C、如果 $R \in BCNF$ ，则 $R \in 3NF$
- D、以上说法都不对

解：B

3、下面关于关系数据库范式的描述中，错误的是_____。

- A) 关系满足的范式越低，则更新操作的代价就越高
- B) 关系满足的范式越高，则更新操作的代价就越低
- C) 关系满足的范式越高，则查询操作的代价就越高
- D) 关系满足的范式越低，则查询操作的代价就越高

解：D

4、假设某商业集团数据库有一关系模式 R 如下：（10 分）

R （商店编号，商品编号，数量，部门编号，负责人）现规定：

- (1)每个商店的每种商品只在一个部门销售。
- (2)每个商店的每个部门只有一个负责人。
- (3)每个商店的每种商品只有一个库存数量。

回答下列问题：

- ①写出 R 中的存在的基本函数依赖（3 分）
- ②找出关系模式 R 的候选键（3 分）



③关系模式 R 最高能不能达到第 3 范式，若不能如何分解？（4 分）

解：（1）（商店编号，商品编号）->部门编号

（商店编号，部门编号）->号负责人

（商店编号，商品编号）->数量

（2）关系模式 R 的候选键为：（商店编号，商品编号）

（3）非主属性为：部门编号，负责人，数量。

它们对候选键都是完全函数依赖关系，所以 R 属于 2NF。

但是：

（商店编号，商品编号）→（部门编号，商店编号）

（商店编号，部门编号）→（负责人）

所以非主属性负责人对候选键传递依赖，那么 R 不满足第三范式，因此 R 最高满足第二范式。

4.2.3 课后习题

1、如果一个关系模式 R 是 1NF，并且关系的每个决定因子都是候选码，那么 R 至少应该是_____范式

解：BC 范式

2、在关系数据库规范化理论的研究中，在函数依赖的范畴内，_____范式达到了最高的规范化程度。

解：BC 范式

3、现有如下关系模式：借阅（图书编号，书名，作者名，出版社，读者编号，读者姓名，借阅日期，归还日期）。其中规定：图书编号唯一；读者编号唯一；读者在归还某一本书后还可以其他时间再次借阅。回答下列问题

（1）写出该关系模式中存在的函数依赖

（2）求出该关系模式的候选键

（3）该关系模式最高满足第几范式？为什么？

解：（1）图书编号→(书名，作者名，出版社)

读者编号→(读者姓名，借阅日期，归还日期)

（2）关系模式 R 的候选键为：（图书编号，读者编号）

（3）非主属性为：书名，作者名，出版社，部分函数依赖主码图书编号



非主属性为：读者姓名，借阅日期，归还日期，部分函数依赖主码读者编号
它们对候选键都是完全函数依赖关系，所以 R 属于 2NF。

公众号：期末加油站



第五章 数据库安全性和完整性

5.1 安全性理念

5.1.1 知识点

安全性的定义：数据库的安全性是指保护数据库以防止不合法使用所造成的数据泄
漏、更改或破坏

不安全因素：

非授权用户对数据库的恶意存取和破坏

数据库重要或敏感数据泄漏

安全环境的脆弱性

安全性保护手段（从外而内）：用户身份鉴别-》多层存取控制-》审计-》视图机制
-》数据加密

5.1.2 例题

1、数据库的安全性保护措施有_____。

解：用户身份鉴别、多层存取控制、审计机制、视图机制、数据加密技术

5.1.3 课后习题

5.2 安全性保护措施

5.2.1 知识点

用户身份鉴别措施：DMS 的最外层安全保护措施，防止不可信用户使用系统

用户标识=用户名+用户标识号（UID，唯一）

用户身份鉴别方法：

- 静态口令鉴别
- 动态口令鉴别
- 生物特征鉴别
- 智能卡鉴别

存取控制措施：确保授权给有资格的用户访问数据库的权限

存取控制机制的组成：

1、定义用户权限，并将用户权限登记到数据字典中，形成安全/授权规则：

用户对某一数据对象的操作权力称为权限（管理问题）；

2、合法权限检查：根据安全规则进行合法权限检查



存取措施机制的分类：

1、自主存取控制（Discretionary Access Control, DAC）：用户对于不同的数据库对象有不同的存取权限，不同的用户对同一对象也有不同的权限，而且用户还可以将其拥有的存取权限转授给其他用户，非常灵活。

主要通过 SQL 的 GRANT 语句（赋予）和 REVOKE（收回）语句来实现

GRANT <权限>[,<权限>]...ON <对象类型> <对象名>[,<对象类型> <对象名>]... TO <用户>[,<用户>]...[WITH GRANT OPTION];

REVOKE <权限>[,<权限>]...ON <对象类型> <对象名>[,<对象类型> <对象名>]... FROM <用户>[,<用户>]...[CASCADE|RESTRICT];

2、强制存取控制（Mandatory Access Control, MAC）：每一个数据库对象被标以一定的密级，每一个用户也被授予某一个级别的许可证。对于任意一个对象，只有具有合法许可证的用户才可以存取。较为严格。

主体是系统中的活动实体，客体是系统中的被动实体，受主体操纵

敏感度标记（Label）：对于主体（许可证级别）和客体（密级），TS（绝密）>=S（机密）>=C（可信）>=P（公开）

- 仅当主体的许可证级别大于或等于客体的密级时，该主体才能读取相应的客体
- 仅当主体的许可证级别小于或等于客体的密级时，该主体才能写相应的客体

视图机制：为不同的用户定义不同的视图，把数据对象限制在一定的范围内。

审计：把用户对数据库的所有操作自动记录下来放入审计日志中。

数据加密：根据一定的算法将原始数据变换为不可直接识别的格式。

5.2.2 例题

2、SQL 的 GRANT 和 REVOKE 语句主要用来维护数据库的_____。

- A、安全性 B、完整性
C、可靠性 D、一致性

解：A GRANT 和 REVOKE 语句属于自主存取控制措施，安全性保护手段

3、对下列两个关系模式：学生（学号，姓名，年龄，姓名，住址，班级号）；班级（班级号，班级名，班主任，班长），使用 GRANT 语句完成下列授权：

- （1）授权用户 U1 对两个表的所有权限，并可给其他用户授权
- （2）授权用户 U2 对学生表具有查看权限，对家庭住址具有更新权限



(3) 将对班级表查看的权限授予所有用户

解：(1) grant all privileges on table Student,Classes to U1 with grant option;

(2) grant select,update(address) on table Student to U2;

(3) grant select on table Classes to public;

5.2.3 课后习题

1、什么是数据库的安全性？DBMS 有哪些安全性措施？

解：数据库的安全性是保护数据库以防止不合法使用所造成的数据泄露、更改或破坏。

DBMS 主要通过以下方法来保护数据库的安全的：通过定义用户标识对用户身份进行鉴定，只允许合法的用户才能进入系统；通过定义用户的访问权限，对于合法的用户只能访问它有权访问的数据；通过定义视图，屏蔽掉用户不能访问的数据；通过对数据加密以及通过审计的功能等方法来保护数据库中的安全性的。

5.3 完整性要求

5.3.1 知识点

定义：数据库的完整性是指数据的正确性和相容性

- 数据的正确性：数据符合显示世界语义、反映方式实际状况
- 数据的相容性：数据库同一对象在不同关系表中的数据是符合逻辑的

目标：防止数据库中存在不符合语义的数据，即防止数据库中存在不正确的数据。

5.3.2 例题

4、数据库的_____是指数据的正确性和相容性。

A、恢复 B、安全性 C、并发控制 D、完整性

解：D

5.3.3 课后习题

2、数据库的完整性是指数据的_____和_____。

(1)正确性 (2)合法性 (3)不被非法存取 (4)相容性 (5)不被恶意破坏

A、(1)和(3) B、(2)和(5)

C、(2)和(4) D、(1)和(4)

解：D



5.4 完整性约束手段

5.4.1 知识点

1、实体完整性：关系模型的实体完整性在 CREATE TABLE 中用 PRIMARY KEY 定义。一种是定义为列级约束条件，另一种是定义为表级约束条件。多个属性构成的码只能通过表级约束实现

检查和违约处理：

- (1) 检查主码值是否唯一，如果不唯一则拒绝插入或修改
- (2) 检查主码的各个属性是否为空，只要有一个为空就拒绝插入或修改

2、参照完整性：关系模型的参照完整性在 CREATE TABLE 中用 FOREIGN KEY 短语定义哪些列为外码，用 REFERENCES 短语指明这些外码参照哪些表的主码。

检查和违约处理：对被参照表和参照表进行增删改操作时都会进行检查

- (1) 拒绝(NO ACTION)执行
- (2) 级联(CASCADE)操作
- (3) 设置为空值

3、用户自定义完整性：针对某一具体应用的数据必须满足的语义要求

属性上的约束条件：NULL/NOT NULL、DEFAULT、UNIQUE、CHECK

- (1) 列值非空 (NOT NULL)
- (2) 列值唯一 (UNIQUE)

元组上的约束条件：使用 CHECK 短语定义元组上的约束条件

违约处理：拒绝执行

5.4.2 例题

5、将 Student 表中 Sno 属性对应为码；将 SC 表中的 Sno、Cno 属性组定义为码

解：create table Student(Sno char(12) primary key,...);或者

create table Student(Sno char(12) not null,...,primary key(Sno));

create table SC(Sno char(12) not null,Cno char(12) not null,...,primary key(Sno,Cno));

6、关系数据库中，实现表与表之间的联系是通过_____。

- A、实体完整性规则 B、值域
- C、用户自定义完整性 D、参照完整性规则

解：D 外键



5.4.3 课后习题

无

公众号：期末加油站



第六章 数据库恢复和并发控制

6.1 事务基本概念

6.1.1 知识点

事务：用户定义的一个数据库操作序列

特点 ACID：

- 原子性 Atomicity: 事务是数据库的逻辑工作单位,事务中包括的操作要么都做,要么都不做
- 一致性 Consistency: 事务执行的结果必须是使数据库从一个一致性状态变到另一个一致性状态。事务执行过程中出现故障则称这时的数据库处于不一致性状态。
- 隔离性 Isolation: 一个事务的执行不能被其他事务干扰,并发执行的各个事务之间不能互相干扰
- 持续性 Durability: 一个事务一旦提交,它对数据库中数据的改变就应该是永久性的。

事务的 ACID 特性可能遭到破坏的因素有：

- (1) 多个事务并行运行时,不同事务的操作交叉执行
- (2) 事务在运行过程中被强制停止

6.1.2 例题

1、一个事务执行过程中,其正在访问的数据被其他事务所修改,导致处理结果不正确,这是由于违背了事务的何种特性而引起的

A、隔离性 B、一致性 C、原子性 D、持久性

解：A

6.1.3 课后习题

1、事务的特性有：_____、原子性、_____和持久性。

解：隔离性、一致性

2、“一个事务中的诸操作要么全做,要么都不做”,这是事务的什么特性_____。

A、原子性 B、一致性

C、隔离性 D、持久性

解：A



3、什么是数据库中的事务？它有哪些特性？这些特性的含义是什么

解：事务是用户定义的一个数据库操作序列，这些操作要么全做要么全不做，是一个不可分割的工作单位。例如在关系数据库中，一个事务可以是一条 SQL 语句、一组 SQL 语句。

事务的特性：原子性（Atomicity）、一致性（Consistency）、隔离性（Isolation）、持续性（Durability），简称 ACID

1.原子性：事务是数据库的逻辑工作单位，事务中包括的操作要么都做，要么都不做

2.一致性：事务执行的结果必须是使数据库从一个一致性状态变到另一个一致性状态。事务执行过程中出现故障则称这时的数据库处于不一致性状态。

3.隔离性：一个事务的执行不能被其他事务干扰，并发执行的各个事务之间不能互相干扰

4.持续性（永久性）：一个事务一旦提交，它对数据库中数据的改变就应该是永久性的。

6.2 数据库恢复技术

6.2.1 知识点

数据库面临的故障种类：

1.事务内部的故障：

2.系统故障：如操作系统故障，CPU 故障，系统断电

3.介质故障：如磁盘损坏、磁头碰撞、瞬时强磁场干扰等

4.计算机病毒

数据库恢复技术有以下几种：

1、数据转储：转储即 DBA 定期地将整个数据库复制到磁带或另一个磁盘上保存起来的过程。这些备用的数据称为后被副本。

转储又分为静态转储和动态转储：静态转储必须等待正在运行的用户事务结束才能进行；动态转储是指转储期间允许对数据库进行存取或修改，即转储和用户事务可以并发执行。

也可分为海量转储和增量转储：海量转储即每次转储全部数据库，增量转储即每次只转储上一次转储后更新的数据。



2、登记日志文件 **Logging**：日志文件是用来记录事务对数据库的更新操作的文件。不同数据库系统采用的日志文件格式并不完全一样，主要有两种格式：以记录为单位的日志文件和以数据块为单位的日志文件

- 对以记录为单位的日志文件，日志文件中需要登记的内容包括：各个事务的开始(BEGIN TRANSACTION)标记、各个事务的结束(COMMIT 或 ROLLBACK)标记、各个事务的所有更新操作。以上每一条内容记为一个日志记录(log record) 每个日志记录的内容主要包括：
 - 事务标识（标明是哪个事务）
 - 操作的类型（插入、删除或修改）
 - 操作对象（记录内部标识）
 - 更新前数据的旧值（对插入操作而言，此项为空值）
 - 更新后数据的新值（对删除操作而言，此项为空值）
- 对以数据块为单位的日志文件，日志记录的内容包括事务标识和被更新的数据块。由于将更新前的整个块和更新后的整个块都放入日志文件中，操作的类型和操作对象等信息就不必放入日志记录中了。

日志文件的作用：

日志文件用于事务故障恢复和系统故障恢复，并协助后备副本进行介质故障恢复。具体作用如下：

- 事务故障恢复和系统故障恢复必须用日志文件。
- 在动态转储方式中必须建立日志文件，备份副本和日志文件结合起来才能有效地恢复数据库。
- 在静态转储方式中，也可以建立日志文件。

为保证数据库是可恢复的，登记日志文件时必须遵循两条规则：

- (1) 登记的次序严格按并发事务执行的时间次序
- (2) 必须先写日志文件，后写数据库

数据库的恢复策略包括以下几种：

- **REDO**：重做，正向扫描日志文件，对每个 REDO 事务重新执行日志文件登记的操作



- **UDNO**: 撤销，反向扫描日志文件，对每个 **UNDO** 事务的更新操作执行逆操作
- **COMMIT**: 提交，将事务中所有对数据库的更新写回到磁盘上的物理数据库中，事务正常结束
- **ROLLBACK**: 回滚，事务运行的过程中发生了某种故障，事务不能继续执行，系统将事务中对数据库的所有已完成操作全部撤销，回滚到事务开始时的状态
事务故障的恢复：

- (1) 反向扫描日志文件（即从最后向前扫描日志文件），查找该事务的更新操作
- (2) 对该事务的更新操作执行逆操作。（来得及或者未来得及写入数据库都没关系）

- (3) 继续反向扫描日志文件，查找该事务的其他更新操作，并做同样处理
- (4) 如此继续，直到读到该事务的开始标记

系统故障的恢复：

(1) 正向扫描日志文件，找出在故障发生前已经提交的事务（这些事务既有 **BEGIN TRANSACTION** 记录，也有 **COMMIT** 记录），将其事务标记记入 **REDO** 队列；同时找出故障发生时尚未完成的事务（这些事务只有 **BEGIN TRANSACTION** 记录，无相应的 **COMMIT** 记录），将其事务标记记入 **UNDO** 队列

- (2) 对撤销队列中的各个事务执行 **UNDO** 操作
- (3) 对重做队列中的各个事务执行 **REDO** 操作

介质故障的恢复：

- (1) 装入最新的数据库后备副本，使数据库恢复到最近一次转储时的一致性状态
- (2) 装入相应的日志文件副本（转储结束时刻的日志文件副本），重做已完成的事务，即扫描日志文件找出需要重做和撤销的事务

6.2.2 例题

2、数据库系统中可能发生各种各样的故障，引起故障的原因大致可以分为几类，分别为：_____、计算机病毒、_____和介质故障。

解：事务内部故障、系统故障

3、写一个修改到数据库中，与写一个表示这个修改的记录到日志文件中的操作，应该（ ）。

- (A) 前者先做 (B) 由程序员安排 (C) 后者先做 (D) 由系统决定哪一个先做



解：C

4、采用数据库镜像技术，主要是为了有效解决_____的问题。

- A、复制故障 B、系统故障
C、事务故障 D、介质故障

解：D

6.2.3 课后习题

6.3 并发控制理念

6.3.1 知识点

事务是并发控制的基本单位，事务读数据 x 一般记为 $R(x)$ ，写数据 x 一般记为 $W(x)$

并发控制的目标：保证事务的隔离性和一致性；使一个事务的执行不受其他事务的干扰

并发操作带来的不一致性：

- 丢失修改：两个事务读入同一数据并修改，其中一个事务的修改会丢失
- 不可重复读：事务 T_1 读取数据后， T_2 执行更新操作，使 T_1 无法再现前一次读取结果
- 读“脏”数据：“脏”数据指事务 T_1 修改某一数据，并将其写回磁盘，事务 T_2 读取同一数据后， T_1 由于某种原因被撤销，则 T_2 读取到的数据就为“脏”数据，即不正确的数据。

主要技术：封锁、时间戳、乐观控制法、多版本并发控制 MVCC

6.3.2 例题

5、设有两个事务 T_1 、 T_2 ，其并发操作如图 1 所示，下面描述正确的是_____。

- A、该操作不存在问题 B、该操作不能重复读
C、该操作丢失修改 D、该操作读“脏”数据

时刻	事务 T_1	事务 T_2	数据库 A 的值
T0	read (A)		A=10
T1	$A=A*A$		
T2	write (A)		A=100
T3		read (A)	A=100
T4	rollback		



解：D 因为出现了回滚操作

6.3.3 课后习题

6.4 锁

6.4.1 知识点

封锁即事务 T 在对某个数据对象例如表、记录等操作之前，先向系统发出请求，对其加锁。

基本的封锁类型：

- 排它锁（写锁，X 锁）：保证了其他事务在 T 释放 A 上的锁前不能再读取和修改 A
- 共享锁（读锁，S 锁）：保证了其他事务可以读 A，但在 T 释放 A 上的 S 锁之前不能对 A 做任何修改

封锁可能会出现两种异常状态：

- 活锁：一个事务可能永远等待，可以采用先来先服务的策略解决。
- 死锁：两个事务互相申请对方锁住的资源导致两个事务永远不能结束。
 - 预防策略：一次性封锁法和顺序封锁法
 - ◆ 一次性封锁要求每个事务必须一次将所有要使用的数据全部加锁。缺点：势必扩大封锁的范围，从而降低了系统的并发度；
 - ◆ 顺序封锁法：预先对数据对象规定一个封锁顺序，所有事务都按这个顺序实行封锁。缺点：难以事先确定封锁顺序
 - 解除策略：超时法和等待图法

- ◆ 超时法：如果一个事务的等待时间超过了规定的时限，就认为发生了死锁；缺点：可能误判死锁，事务因为其他原因使等待时间超过时限；若时限设置得太长，死锁发生后不能及时发现
- ◆ 等待图法：并发控制子系统周期性地生成事务等待图，并进行检测，如果发现图中存在回路，则发生了死锁。如果检测到死锁，一般采用的方法是选择一个处理死锁代价较小的事务，将其撤销，释放此事务持有的所有的锁，之后对撤销的事务所执行的数据修改操作必须加以恢复。

6.4.2 例题

6、关于死锁，下列说法正确的是（ ）



- (A) 死锁是操作系统中的问题，数据库中不存在
- (B) 数据库中防止死锁的方法是禁止两个用户同时访问数据库
- (C) 当两个用户竞争相同的资源时不会发生死锁
- (D) 只有出现并发操作时，才有可能出现死锁

解：C

6.4.3 课后习题

4、什么是数据库中的死锁？如何解决死锁问题

解：死锁即两个事务互相申请对方锁住的资源导致两个事务永远不能结束

解决死锁有以下方案：

- (1) 一次性封锁法：一次性封锁发要求每个事务必须一次将所有要使用的数据全部加锁。缺点：势必扩大封锁的范围，从而降低了系统的并发度；
- (2) 顺序封锁法：预先对数据对象规定一个封锁顺序，所有事务都按这个顺序实行封锁。缺点：难以事先确定封锁顺序

公众号：期末加油站



第七章 数据库设计

7.1 设计流程

7.1.1 知识点

数据库设计可以分为以下 6 个阶段：

- 需求分析：准确了解和分析用户需求（包括数据和处理）
- 概念结构设计：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
- 逻辑结构设计：将概念结构转换为某个 DBMS 所支持的数据模型，并对其进行优化
- 物理结构设计：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：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，组织数据入库并试运行
- 数据库运行和维护：在数据库投入正式使用后不断地对其进行评价、调整与修改

7.1.2 例题

1、规范化数据库设计分为以下六个设计阶段：需求分析阶段、_____、_____、_____、数据库实施阶段、数据库运行和维护阶段。

解：概念结构设计阶段、逻辑结构设计阶段、物理结构设计阶段

7.1.3 课后习题

7.2 需求分析

7.2.1 知识点

任务：准确了解和分析用户需求（信息要求、处理要求、安全性和完整性要求）

方法：结构化分析方法（SA）；自顶向下、逐层分解

输出：数据字典（DD，数据库中关于数据的描述，元数据）+数据流图（DFD）

- 数据字典：包括数据项、数据结构、数据流、数据存储、处理过程 5 个部分

1.数据项：数据项描述 = {数据项名，数据项含义说明，别名，数据类型，取值范围.....}

2.数据结构：数据结构反映了数据之间的组合关系。数据结构描述 = {数据结构名，



含义说明，组成：{数据项或数据结构}}

3.数据流：数据流是数据结构在系统内传输的路径。数据流描述 = {数据流名，数据流来源，数据流去向，平均流量，高峰期流量，组成：{数据结构}}

4.数据存储：数据存储是数据结构停留或保存的地方。数据存储描述 = {数据存储名，说明，编号，输入的数据流，输出的数据流，组成：{数据结构}，数据量，存取频度，存取方式}

5.处理过程：处理过程的具体处理逻辑一般用判定表或判定树来描述。处理过程描述 = {处理过程名，说明，输入：{数据流}，输出：{数据流}，处理：{简要说明}}概念结构设计

- 数据流程图：数据流图表达了数据和处理的关系。用处理、外部实体、数据流以及数据存储来表示系统需求的图表

7.2.2 例题

2、在数据库设计中数据流图（DFD）和数据字典(DD)主要用来描述结构化方法中的_____阶段的工具。

- A、概念结构设计 B、需求分析
C、可行性分析 D、逻辑结构设计

解：B

7.2.3 课后习题

7.3 概念结构设计

7.3.1 知识点

任务：将需求分析得到的用户需求抽象为信息结构（概念模型）的过程

方法：概念模型 E-R 模型

输出：E-R 图

- 实体型用矩形标识，矩形框内写明实体名
- 属性用椭圆形标识，无向边连接实体和属性
- 联系用菱形标识，菱形框内写明联系名，无向边连接实体，注明类型(1:1;1:n;m:n)

子系统的 E-R 图合并时会产生冲突，包括以下三类冲突：

- 属性冲突：属性域冲突、属性取值单位冲突
- 命名冲突：同名异义、异名同义



- 结构冲突：
 - 同一对象在不同应用中具有不同的抽象
 - 同一实体在不同子系统的 E-R 图中所包含的属性个数和属性排列次序不完全相同
 - 实体间的联系在不同 E-R 图中是不同的类型

7.3.2 例题

3、某高校教务处教务管理中需要如下信息：

教师：教师编号、姓名、性别、职务、职称、电话、E-mail

学生：学号，姓名、性别、年龄、所学专业、籍贯

课程：课程号、课程名、开课学期、所属类别

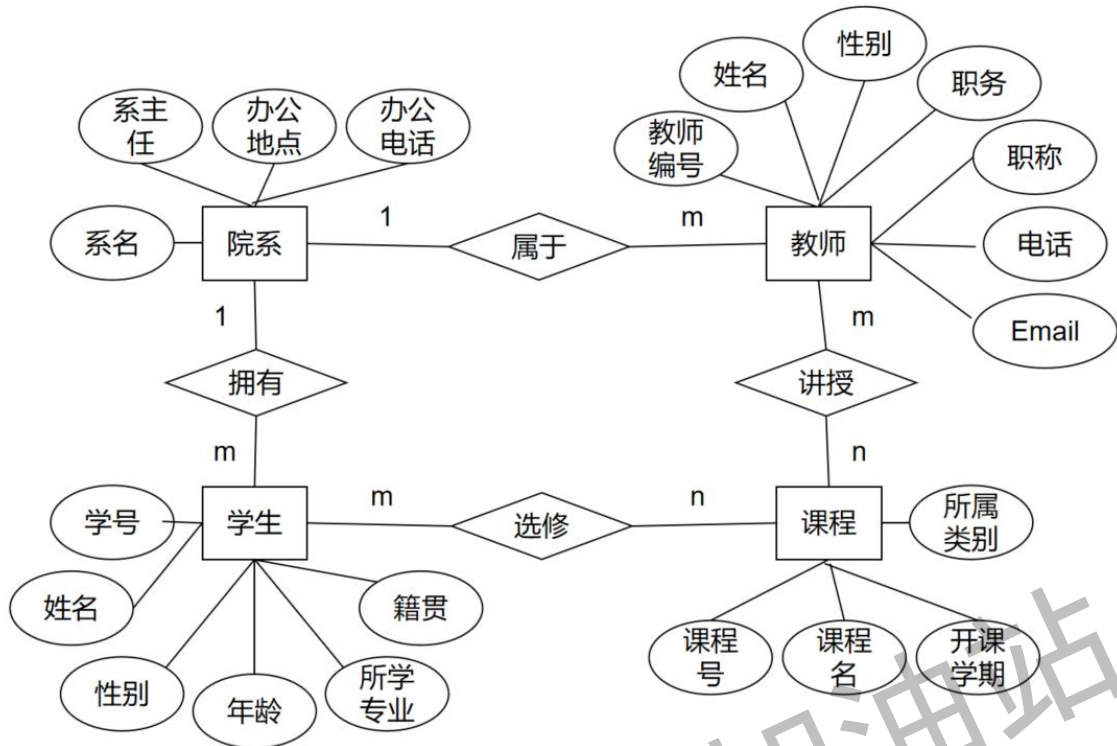
院系：系名、系主任、办公地点、办公电话

其中：①一个系可以有多个学生，一个学生可以选修多门课程，而一门课程可以为多个学生选修；②一个单位(系)可以有多个教师，一个教师只能在一个单位(系)工作，一个教师可以讲授多门课程，一门课程可以为多个教师讲授，完成如下设计：

- (1) 根据院系，教师，课程，学生等实体，设计该教务管理系统局部 E-R 模型
- (2) 将该 E-R 图转换为关系模型结构，并用下划线标明各关系中的主关键字

解：(1) E-R 图如下所示：





(2) 四个实体是四个关系模型

教师（教师编号，姓名，性别，职务，职称，电话，E-mail）

学生（学号，姓名，性别，年龄，所学专业，籍贯）

课程（课程号，课程名，开课学期，所属类别）

院系（系名，系主任，办公地点，办公电话）

四个关系可以转换为四个关系模型，分别是：

拥有（系名，学号）

选修（学号，课程号）

讲授（教师编号，课程号）

属于（系名，教师编号）

5、当将局部 E-R 图集成为全局 E-R 图时，如果同一对象在一个局部 E-R 图中作为实体，而在另一个局部 E-R 图中作为属性，这种现象称为_____。

A、属性冲突 B、命名冲突 C、结构冲突 D、语义冲突

解：C

7.3.3 课后习题

2、数据库设计的概念结构设计阶段，表示概念结构的常用方法和描述工具是_____。



- A、数据流程分析法和数据流程图 B、实体—联系方法和 E—R 图
C、结构分析法和模块结构图 D、层次分析法和层次结构图

解：B E-R 模型

3、在数据库概念结构设计中，各分 E-R 模型之间的冲突主要有：属性冲突、
_____和结构冲突。

解：命名冲突

7.4 逻辑结构设计

7.4.1 知识点

任务：将概念结构转换为某个 DBMS 所支持的数据模型，并对其进行优化

目标：E-R 图向关系数据模型的转换

输出：关系模型

- 一个 1:1 联系可以转换为一个独立的关系模式，也可以与任意一端对应的关系模式合并。
- 一个 1: n 联系可以转换为一个独立的关系模式，也可以与 n 端对应的关系模式合并。
- 一个 m: n 联系转换为一个关系模式。
- 三个或三个以上实体间的一个多元联系可以转换为一个关系模式。
- 具有相同码的关系模式可合并。

7.4.2 例题

4、从 E-R 模型关系向关系模型转换时，一个 M:N 联系转换为关系模式时，该关系模式的码是_____。

- A、M 端实体的码 B、N 端实体的码
C、M 端实体码与 N 端实体码组合 D、重新选取其他属性

解：C

6、在 E-R 模型中，如果有 3 个不同的实体集，3 个多对多的联系，则应转换为
() 个关系模式。

- (A) 3 (B) 6 (C) 9 (D) 4

解：B 三个实体+三个联系



7.4.3 课后习题

1、在数据库设计过程中将 E-R 模型转换为关系数据模型的转换原则？

解：（1）一个 1:1 联系可以转换为一个独立的关系模式，也可以与任意一端对应的关系模式合并。如果转换为一个独立的关系模式，则与该联系相连的各实体的码以及联系本身的属性均转换为关系的属性，每个实体的码均是该关系的码；如果与某一端实体对应的关系模式合并，则需要在该关系模式的属性中加入另一个关系模式的码和联系本身的属性。

（2）一个 1: n 联系可以转换为一个独立的关系模式，也可以与 n 端对应的关系模式合并。如果转换为一个独立的关系模式，则与该联系相连的各实体的码以及联系本身的属性均转换为关系的属性，而关系的码为 n 端实体的码

（3）一个 m: n 联系转换为一个关系模式。与该联系相连的各实体的码以及联系本身的属性均转换为关系的属性，各实体的码组成关系的码或关系码的一部分

4、若两个实体间的关系是 1:m，则进行逻辑设计时实现 1:m 联系的方法是（ ）

- （A）在“m”端实体转换的关系中加入“1”端实体转换的关系的码
- （B）将“m”端实体转换的关系的码加入“1”端实体转换的关系中
- （C）在两个实体转换的关系中，分别加入另外一个关系的码
- （D）将两个实体转换成一个关系

解：A

7.5 实施与维护

7.5.1 知识点

- 数据的载入和应用程序的调试
- 数据库的试运行
- 数据库的运行和维护
 - 1.数据库的转储和恢复
 - 2.数据库的安全性、完整性控制
 - 3.数据库性能的监督、分析和改造
 - 4.数据库的重组织和重构造



7.5.2 例题

7.5.3 课后习题

公众号：期末加油站

