

复习主要内容

题型：选择题 40 个共 40 分，判断题 10 个共 10 分，关系代数和 SQL 设计 30 分，简答题 2 个共 8 分，数据库规范化 12 分

章节主要复习内容：（不仅仅是掌握概念，还要能灵活应用）

一、概述

- 1) 理解数据、数据库、数据库管理系统、数据库系统
- 2) 实体之间的联系形式主要有哪几种？
1 对 1 1 对多 多对多
- 3) 理解概念模型、逻辑模型、物理模型
- 4) 什么叫数据与程序的物理独立性？什么叫数据与程序的逻辑独立性？为什么数据库系统具有数据与程序的独立性？
- 5) 数据库管理系统的主要管理功能？

数据独立性高

- 数据独立性是指数据的组织和存储方法与应用程序互不依赖、彼此独立的特性。
- 物理独立性
 - 指用户的应用程序与存储在磁盘上的数据库中数据是相互独立的。当数据的物理存储改变了，应用程序不用改变。
- 逻辑独立性
 - 指用户的应用程序与数据库的逻辑结构是相互独立的。数据的逻辑结构改变了，用户程序也可以不变。
- 数据独立性是由DBMS的二级映像功能来保证的

数据由DBMS统一管理和控制

- DBMS提供的数据库控制功能
 - > (1)数据的安全性 (Security) 保护
 - 保护数据, 以防止不合法的使用造成的数据的泄密和破坏。
 - > (2)数据库恢复 (Recovery)
 - 将数据库从错误状态恢复到某一已知的正确状态
 - 比如: 转账->断电
 - > (3)数据的完整性 (Integrity) 检查
 - 将数据控制在有效的范围内, 或保证数据之间满足一定的关系。
 - 比如: 成绩不能是负数
 - > (4)并发 (Concurrency) 控制
 - 对多用户的并发操作加以控制和协调, 防止相互干扰而得到错误的结果。
 - 比如: 多个售票点同时售票

36

二、 关系代数

- 6) 熟练掌握关系代数基本运算
- 7) 关系表有多个候选键, 但只能选定其中一个做为主键
- 8) 投影和选择运算要重点掌握, 投影后, 元组数量可能变化也可能不发生变化, 选择后, 属性数量不发生变化
- 9) 理解自然连接、外连接 (左、右、全外连接)
- 10) 理解主键、外键的概念及关系

三、 SQL

重点掌握

- 11) 要掌握单表查询、多表连接查询、嵌套查询、分组运算、排序、选择、投影等操作 1

三、外连接

- 外连接
 - > 将没有满足连接条件的元组也列出来, 不存在的列值填以NULL值。
- 左外连接
 - > 如果只把左边关系R中要舍弃的元组保留就叫做左外连接(LEFT OUTER JOIN或LEFT JOIN)

```
SELECT      table.column, table.column
FROM  table1 LEFT OUTER JOIN table2
ON      table1.column = table2.column
```
- 右外连接
 - > 如果只把右边关系S中要舍弃的元组保留就叫做右外连接(RIGHT OUTER JOIN或RIGHT JOIN)。

```
SELECT      table.column, table.column
FROM  table1 RIGHT OUTER JOIN table2
ON      table1.column = table2.column
```

23

- 12) 视图的作用? 如何定义视图?

查询选修课程最多的学生学号

```
create view sno_num(sno,count)
as
select sno,count(*) from sc group by sno
```

3.6.4 视图的作用

1. 视图能够简化用户的操作
 2. 视图使用户能以多种角度看待同一数据
 3. 视图对重构数据库提供了一定程度的逻辑独立性
 4. 视图能够对机密数据提供安全保护
 5. 适当的利用视图可以更清晰的表达查询
- 13) 如何去掉查询结果的重复元组（记录） 1
 - 14) HAVING 和 where 语句的用途，HAVING 作用于分组，对分组进行限制输出，Where 作用于整个表，用来选择满足条件的行 1
 - 15) ALTER TABLE 用来修改表的结构（数据库模式）
 - 16) 判断空值用“IS NULL” 1
 - 17) 单列排序、多列排序 1
 - 18) 给定几个关系表，能用关系代数、SQL 实现不同的查询 1
 - 19) Like 查询中“%”与“-”的应用 1
 - 20) 创建索引的语句 1

1.建立索引

- 语句格式
CREATE [UNIQUE] [CLUSTER] INDEX <索引名>
ON <表名>(<列名>[<次序>][,<列名>[<次序>]]...);
 - UNIQUE：每一个索引值只对应唯一的数据记录，不允许表中不同的行在索引列上取相同值。若已有相同值存在，则系统给出相关信息，不建此索引。系统并拒绝违背唯一性的插入、更新。
 - CLUSTER：键值相同的记录在物理上集中存放。一个基本表上只能建一个聚簇索引。适用范围：
 - 很少对基表进行增删操作
 - 很少对其中的变长列进行修改操作
 - ASC/DESC：升 / 降序
- 例如：
CREATE UNIQUE INDEX SC_INDEX ON SC(SNO DESC,CNO ASC);

72

建立索引（续）

[例13] CREATE CLUSTER INDEX Stusname

ON Student(Sname);

- 在Student表的Sname（姓名）列上建立一个聚簇索引

四、 数据库安全性

- 21) 掌握几类数据库安全性控制的常用方法

数据库安全性控制概述（续）

■ 数据库安全性控制的常用方法

- 用户标识和鉴别
- 存取控制
- 视图
- 审计
- 加密存储

22) 如何用 sql 语句定义授权和收回授权

第1种授权（对象权限）

授权:

```
GRANT <特权> ON <表名> TO <受权者>[, <受权者>]  
[WITH GRANT OPTION]
```

特权: **ALL PRIVILEGES** |操作

操作: **SELECT** |**INSERT**|**DELETE**|**UPDATE** [属性表]

受权者: **PUBLIC** |用户标识符

收回特权:

```
REVOKE <特权> ON <表名> FROM <受权者>[, <受权者>]
```

38

五、 数据库完整性

23) 掌握几类数据库完整性

24) 如何用 sql 语句定义数据库完整性

完整性约束命名子句（续）

[例5.11]建立教师表TEACHER，要求每个教师的应发工资不低于3000元。

应发工资是工资列Sal与扣除项Deduct之和。

```
CREATE TABLE TEACHER  
( Eno NUMERIC(4) PRIMARY KEY /*在列级定义主码*/  
  Ename CHAR(10),  
  Job CHAR(8),  
  Sal NUMERIC(7,2),  
  Deduct NUMERIC(7,2),  
  Deptno NUMERIC(2),  
  CONSTRAINT TEACHERFKKey FOREIGN KEY (Deptno)  
    REFERENCES DEPT(Deptno),  
  CONSTRAINT C1 CHECK (Sal + Deduct >= 3000)  
);
```

```
ALTER TABLE Student  
DROP CONSTRAINT C1;  
ALTER TABLE Student  
ADD CONSTRAINT C1 CHECK (Sno BETWEEN 900000 AND 999999),  
ALTER TABLE Student  
DROP CONSTRAINT C3;  
ALTER TABLE Student  
ADD CONSTRAINT C3 CHECK(Sage < 40);
```

六、 关系数据理论

- 25) 掌握第一范式、第二范式和第三范式、BC 范式定义及判断
- 26) 给定一个不规范的应用，画出 E-R 图，能标注表的函数依赖关系，能判断属于第几范式，如果不规范能将表规范到第三范式。 ? ? ? ? ?

七、 数据库设计

- 27) 数据库的设计过程有哪些？每个阶段主要任务是什么？

7.1.3 数据库设计的基本步骤

- 数据库设计分6个阶段
 - 需求分析
 - 概念结构设计
 - 逻辑结构设计
 - 物理结构设计
 - 数据库实施
 - 数据库运行和维护
- 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

- 28) 理解 E-R 图图示的表示方法
- 29) 能画出简单应用的 E-R 图
- 30) 局部 E-R 图合并成全局 E-R 图时可能出现的合并冲突有哪几种。

合并分E-R图，生成初步E-R图（续）

- 冲突的种类
 - 属性冲突
 - 命名冲突
 - 结构冲突

1. 属性冲突

■ 两类属性冲突

➤ 属性域冲突

- 属性值的类型
- 取值范围
- 取值集合不同

例1，由于学号是数字，因此某些部门（即局部应用）将学号定义为整数形式，而由于学号不用参与运算，因此另一些部门（即局部应用）将学号定义为字符型形式。

例2，某些部门（即局部应用）以出生日期形式表示学生的年龄，而另一些部门（即局部应用）用整数形式表示学生的年龄。

➤ 属性取值单位冲突

例：学生的身高，有的以米为单位，有的以厘米为单位，有的以尺为单位。

88

属性冲突（续）

■ 属性冲突的解决方法

➤ 通常用讨论、协商等行政手段加以解决。

➤ 满足整体需求基础上的方便原则

3. 结构冲突

■ 三类结构冲突

➤ 同一对象在不同应用中具有不同的抽象

例，“课程”在某一局部应用中被当作实体

在另一局部应用中则被当作属性

- 解决方法：通常是把属性变换为实体或把实体变换为属性，使同一对象具有相同的抽象。

92

3. 结构冲突

■ 三类结构冲突

➤ 同一实体在不同分E-R图所包含的属性个数和属性排列次序不完全相同

- 产生原因：不同的局部应用关心的是该实体的不同侧面。

- 解决方法：使该实体的属性取各分E-R图中属性的并集，再适当设计属性的次序。

实体和联系转化为表的方式

八、 不考

九、 查询优化

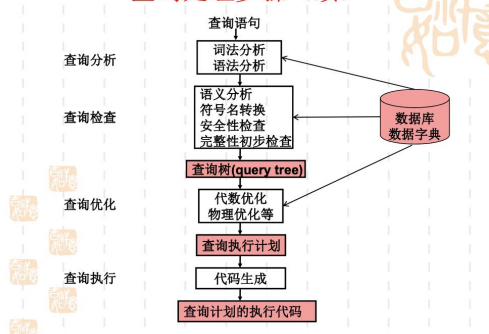
31) 数据库查询处理的主要步骤及内容。

9.1.1 查询处理步骤

■ 关系数据库管理系统查询处理阶段：

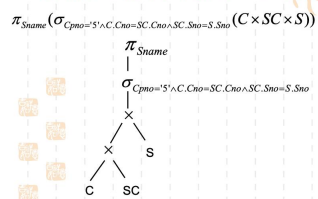
1. 查询分析
2. 查询检查
3. 查询优化
4. 查询执行

查询处理步骤（续）

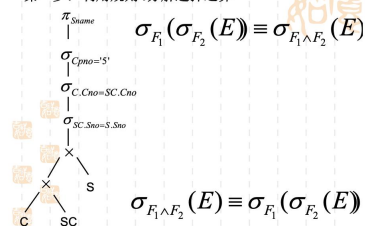


32) 代数优化的一般原则

1. 先分开

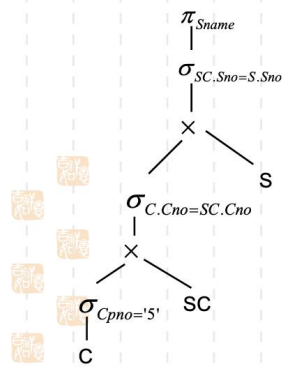


第一步：利用规则4分解选择运算

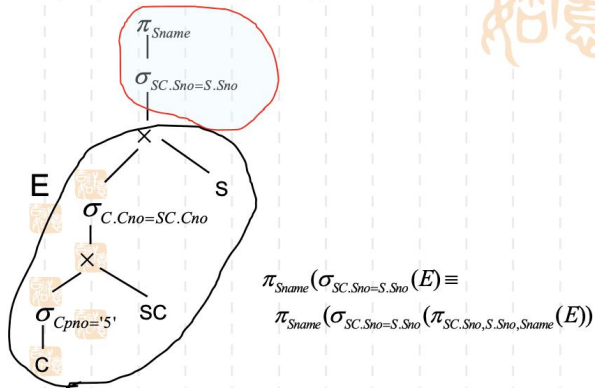


2. 先做投影运算

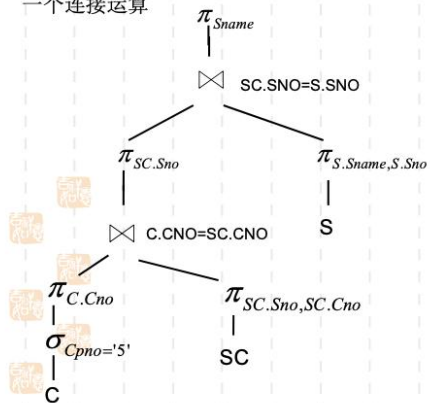
第二步(2)：下放完成后：



第三步：尽量下放投影运算



第五步：把选择同在它前面要执行的笛卡尔积结合起来成为一个连接运算



用投影属性消除无用的属性

代数优化的若干方法（总结）

■ 一般规则

小关系先做join

选择条件向树叶方向下压

连接和选择同时做

加投影

减小中间表大小

先过滤部分记录，减小中间表

减少产生一个额外的中间表

过滤不起作用的字段

- 可能会减少代价，但不总是正确，有异常的情况，因为代数优化没有考虑到实际的数据分布情况。

33) 什么是物理优化？

- 实践证明：合理选择存取路径，往往能够收到显著的优化效果，应成为优化的重点。
- 物理优化就是要选择高效合理的操作算法或存取路径，求得优化的查询计划

34) 哪些情况对属性建立索引比较合适？哪些情况不适合建立索引？

查询结果 $\ll 1$ 或者小于百分之十的时候

35) 笛卡尔运算切记不要滥用，有可能使查询很长时间

十、 数据库恢复

36) 掌握事务的概念及事务的四个特性

37) 什么是事务故障、系统故障和介质故障？产生原因，如何进行故障恢复？

■ 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

38) 重点掌握具有检查点的恢复技术及恢复步骤

39) 常见的数据库保护措施有安全性保护，完整性保护，并发控制及其故障恢复

十一、 数据库并发控制

- 40) 数据库并发操作可能带来的数据不一致的几种形式
- 41) 事务的定义、提交和回滚操作
- 42) 封锁机制是数据库并发控制的主要方式
- 43) 锁的主要类型及相容矩阵
- 44) 重点掌握三级封锁协议

4. 封锁协议小结

- 三级协议的主要区别
 - 什么操作需要申请封锁以及何时释放锁（即持锁时间）
- 不同的封锁协议使事务达到的一致性级别不同

	X锁		S锁		一致性保证		
	操作结束释放	事务结束释放	操作结束释放	事务结束释放	不丢失修改	不读“脏”数据	可重复读
一级封锁协议		√			√		
二级封锁协议		√	√		√	√	
三级封锁协议		√		√	√	√	√

表11.1 不同级别的封锁协议和一致性保证

- 45) 什么是两段锁协议 2PL
- 46) 满足两段锁协议是数据库并发调度可串行化的充分条件。