

1、在教学管理数据库中，有如下三个关系表

学生信息表：S(S#, SNAME, AGE, SEX)

课程表：C(C#, CNAME, TEACHER)

选课表：SC(S#, C#, GRADE)

其中 S#、C#为 S、C 表的主码，(S#, C#)是 SC 表的主码，也分别是参照 S、C 表的外码。

请使用关系代数表达式回答下列问题：

- 1) 查询同时学习过课程号为 001 和 002 课程的学生的学号和成绩
- 2) 查询学习过课程号为 001 的学生学号、姓名、及这门课的成绩
- 3) 查询没有学习过课程号为 002 的学生姓名和年龄
- 4) 查询老师 gao 所教过的学生中成绩为 90 分以上(包括 90 分)的学生姓名
- 5) 查询选修了全部课程的学生姓名

2、在工程管理数据库中，包括 S，P，J，SPJ 四个关系模式：

S(S#, SNAME, SCITY)

P(P#, PNAME, COLOR)

J(J#, JNAME, JCITY)

SPJ(S#, P#, J#, QTY)

供应商表 S 由供应商代码 (S#)、供应商姓名 (SNAME)、供应商所在城市 (SCITY) 组成；

零件表 P 由零件代码 (P#)、零件名 (PNAME)、颜色 (COLOR) 组成；

工程表 J 由工程代码 (J#)、工程名 (JNAME)、工程所在城市 (JCITY) 组成；

供应情况表由 SPJ 由供应商代码 (S#)、零件代码 (P#)、工程代码 (J#)、零件供应数量 (QTY) 组成。

其中 S#、P#、J#分别是 S、P、J 表的主码，(S#, P#, J#) 是 SPJ 的主码，也分别是参照 S、P、J 表的外码。

请使用关系代数表达式回答下列问题：

- 1) 检索使用了在北京且供应商名字为 “S1” 的供应商生产的蓝色零件的工程代码
- 2) 检索供应商与工程所在城市相同的工程代码和工程名
- 3) 检索长春的任何工程都不使用的零件代码
- 4) 检索使用了零件号= “P2” 的零件的工程代码及工程名
- 5) 检索为工程代码= “J5” 的工程供应绿色零件的供应商代码和供应商姓名

1、判断下列命题是否成立。若不成立，请给出反例

- (a) $\sigma_{\theta_1}(\sigma_{\theta_2}(R)) = \sigma_{\theta_2}(\sigma_{\theta_1}(R)) = \sigma_{\theta_1 \wedge \theta_2}(R)$
- (b) $\Pi_{L_1}(\Pi_{L_2}(R)) = \Pi_{L_2}(\Pi_{L_1}(R))$
- (c) $\Pi_L(\sigma_\theta(R)) = \sigma_\theta(\Pi_L(R))$
- (d) $\Pi_L(R \cup S) = \Pi_L(R) \cup \Pi_L(S)$ // 假定 R 和 S 满足求并的条件
- (e) $\Pi_L(R \cap S) = \Pi_L(R) \cap \Pi_L(S)$ // 假定 R 和 S 满足求交的条件
- (f) $\sigma_\theta(R \cap S) = \sigma_\theta(R) \cap S = R \cap \sigma_\theta(S)$
- (g) $\sigma_\theta(R - S) = \sigma_\theta(R) - S = R - \sigma_\theta(S)$
- (h) $(R \bowtie_{\theta_1} S) \bowtie_{\theta_2} T = R \bowtie_{\theta_1} (S \bowtie_{\theta_2} T)$
- (i) $(R \bowtie S) \bowtie T = R \bowtie (S \bowtie T)$
- (j) $R \bowtie R = R \cap R$

2、在教学管理数据库中，有如下四个关系表

学生信息表：S(S#, SNAME, AGE, SEX)

课程表：C(C#, CNAME, T#)

选课表：SC(S#, C#, GRADE)

教师表：T(T#, TNAME, ADDR)

请使用SQL语句完成如下的功能操作：

- 1) 查询未讲授“数据库系统”课程的教师号和教室名。
- 2) 查询既讲授了“C1”号又讲授了“C4”号课程的教师姓名。
- 3) 查询至少讲授两门课程的教师号、教师姓名和地址。
- 4) 查询年龄大于20岁男、女同学各有多少人。
- 5) 查询年龄大于所有男同学年龄的女同学的学号、姓名和年龄。
- 6) 查询只选修了两门课程的男同学的学号和姓名。
- 7) 查询“liu”同学不学的课程的课程号。
- 8) 查询每个同学的总分，要求显示学号、分数，并按分数降序排列，分数相同的按学号升序排列。
- 9) 将每一门课的成绩90-100分之间的同学的学号、姓名和性别，插入到另一已存在的基本表SS(SS#, SNAME, SEX)中。
- 10) 将选修了“gao”老师所担任的“数据库系统”课程的所有不及格同学的成绩修改为0分。
- 11) 将“liu”同学的信息从基本表S和SC中删除。

1、设有如下实体:

图书: 书号、书名、出版日期、出版社

读者: 借书证号、姓名、年龄、性别、家庭住址

书架: 书架号、房间号

上述实体中存在如下联系:

- a) 一本书只能放在一个书架上, 一个书架可以放多本书;
- b) 一位读者可以借多本书, 一本书只能被一位读者借阅

试完成如下工作:

(1) 设计该图书管理系统的 E-R 图; (7 分)

(2) 将该 E-R 图转换为等价的关系模式表示的数据库逻辑结构。

2、设有如下实体:

学生: 学号、单位名称、姓名、性别、年龄、选修课名

课程: 编号、课程名、开课单位、任课教师号

教师: 教师号、姓名、性别、职称、讲授课程编号

单位: 单位名称、电话、教师号、教师姓名

上述实体存在如下联系:

- (1) 一个学生可以选多门课程, 一门课程可被多名学生选修
- (2) 一个教师可讲授多门课程, 一门课程可由多个教师讲授
- (3) 一个单位可有多个教师或学生, 一个教师或学生只属于一个单位
- (4) 一个单位可开设多门课程, 一门课程只能属于一个单位

完成如下工作:

设计该系统的 ER 图, 并写出对应的关系模式, 标明主码

3、工厂需建立一个管理数据库存储以下信息:

1) 工厂: 厂名、厂长姓名;

2) 车间: 车间号、车间主任姓名、地址、电话;

3) 仓库: 仓库号、仓库主任姓名、电话;

4) 零件: 零件号、重量、价格;

5) 产品: 产品号、价格;

上述实体存在如下联系:

- a) 一个工厂内有多个车间和多个仓库, 一个车间或一个仓库都只能属于一个工厂;
- b) 一个车间生产多种产品, 每种产品只能产自一个车间;
- c) 一个车间生产多种零件, 一种零件也可能为多个车间所制造;
- d) 一个产品由多种零件组成, 一种零件也可装配出多种产品;
- e) 产品和零件均存入仓库。

根据上述要求, 完成如下工作:

画出该系统的 E-R 图, 并写出对应的关系模式, 标明主码。

4、举例说明 3 个实体型之间的三元联系型与这 3 个实体型之间的 3 个二元联系型表示的意义通常是不同的。在什么情况下 3 个实体型之间的三元联系型与这 3 个实体型之间的 3 个二元联系型表示的意义会是相同的?

1. 已知学生关系模式:

$S(Sno, Sname, SD, Sdname, Course, Grade)$

其中: Sno 学号、Sname 姓名、SD 系名、Sdname 系主任名、Course 课程、Grade 成绩。

(1) 写出关系模式 S 的基本函数依赖和主码。

(2) 原关系模式 S 为几范式? 为什么? 分解成高一级范式, 并说明为什么?

(3) 将关系模式分解成 3NF, 并说明为什么?

2. 设关系模式 $R(ABCD)$, F 是 R 上成立的 FD 集, $F=\{AB \rightarrow CD, A \rightarrow D\}$ 。

(1) 试说明 R 不是 2NF 模式的理由;

(2) 试把 R 分解成 2NF 模式集;

3. 设关系模式 $R(ABC)$, F 是 R 上成立的 FD 集, $F=\{C \rightarrow B, B \rightarrow A\}$ 。

(1) 试说明 R 不是 3NF 模式的理由;

(2) 试把 R 分解成 3NF 模式集;

4. 设关系模式 $R(A, B, C, D)$, 其函数依赖集为:

$F=\{A \rightarrow C, C \rightarrow A, B \rightarrow AC, D \rightarrow AC\}$

(1) 求 R 的极小函数依赖集 (最小覆盖)

(2) 求 R 的候选码

5. 设有关系模式 $R\{A, B, C, D, E\}$, 其上有函数依赖集:

$F=\{A \rightarrow C, C \rightarrow D, B \rightarrow C, DE \rightarrow C, CE \rightarrow A\}$

(1) 求所有候选键。

(2) 判断 $\rho=\{AD, AB, BC, CDE, AE\}$ 是否是无损连接分解?

(3) 把 R 分解为 BCNF, 并具有无损连接性。

1、利用可扩展 hash 方法对以下记录进行 hash 存储：

3, 5, 7, 12, 16

设 hash 函数 $h(x) = x \bmod 8$ ，其中散列函数 $h(k)$ 是一个 b (足够大) 位二进制序列，序列的前 d 位用作索引，来区分每个元素属于哪个桶。

现要求每个桶至多包含 2 个元素，以上元素按从左往右的顺序依次添加。开始时只使用序列的前 1 位作索引 (即 $d=1$)，当桶满时进行分裂， d 相应增大。请画出添加完以上所有元素后，最终的索引结构。

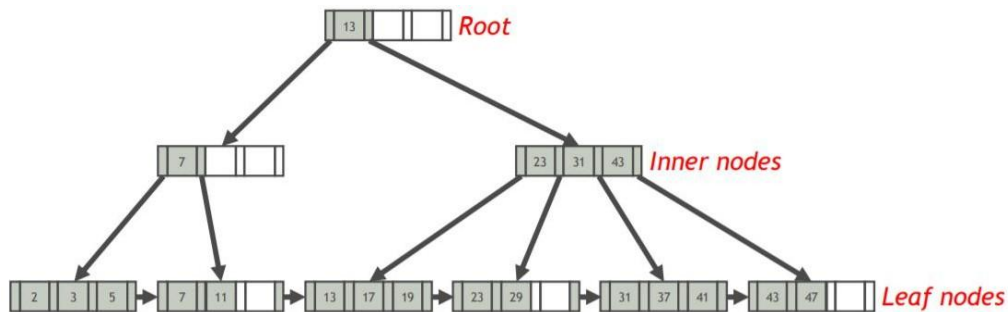
2、利用可扩展 hash 方法对以下记录进行 hash 存储：

3, 16, 20, 21, 30, 44, 47, 61

设 hash 函数 $h(x) = x \bmod 16$ ，其中散列函数 $h(k)$ 是一个 b (足够大) 位二进制序列，序列的前 d 位用作索引，来区分每个元素属于哪个桶。

现要求每个桶至多包含 2 个元素，以上元素按从左往右的顺序依次添加。开始时只使用序列的前 1 位作索引 (即 $d=1$)，当桶满时进行分裂， d 相应增大。请画出添加完以上所有元素后，最终的索引结构。

3、已知有如下 b+树：



回答下列问题：

- (1) 插入键值为 35 的索引项(index entry)后，该 B+树变成什么样？请绘制出来；
- (2) 删除键值为 13 的索引项(index entry)后，该 B+树变成什么样？请绘制出来；

4. 设教学管理数据库有如下 3 个关系模式：

$S(S\#, SNAME, AGE, SEX)$

$C(C\#, CNAME, TEACHER)$

$SC(S\#, C\#, GRADE)$

其中 S 为学生信息表、 SC 为选课表、 C 为课程信息表； $S\#$ 、 $C\#$ 分别为 S 、 C 表的主码， $(S\#, C\#)$ 是 SC 表的主码，也分别是参照 S 、 C 表的外码

用户有一查询语句：

Select SNAME

From S, SC, C

Where $SC.S\#=S.S\#$ and $SC.C\#=C.C\#$ and CNAME= “数据库”

检索选学 “数据库” 课程的学生的姓名。

- (1) 写出以上 SQL 语句所对应的关系代数表达式。

- (2) 画出上述关系代数表达式所对应的查询计划树。使用启发式查询优化算法，对以上查询计划树进行优化，并画出优化后的查询计划树。
- (3) 设 SC 表有 10000 条元组，C 表有 50 条元组，S 表中有 1000 条元组，SC 中满足选修数据库课程的元组数为 150，计算优化前与优化后的查询计划中每一步所产生的中间结果大小

5. 已知关系 $R(w,x), S(x,y), T(y,z)$ 的块数分别为 5000, 10000, 10000。我们准备执行关系代数查询 $(R \bowtie S) \bowtie T$ 。假设缓冲池中有 $M = 101$ 个页可用， R, S, T 上均无索引且未按连接属性排序。请回答下列问题。

- (a) 使用什么算法执行 $R \bowtie S$ 最适合? 说明理由。
- (b) 使用(a)中选择的算法执行 $R \bowtie S$ 的 I/O 代价是多少?
- (c) 如果 $R \bowtie S$ 的结果不超过 49 块，那么在使用(a)中选择的算法执行 $R \bowtie S$ 时， $R \bowtie S$ 的结果是否需要物化(materialize)到文件中? 说明理由。
- (d) 如果 $R \bowtie S$ 的结果不超过 49 块，那么使用什么算法将 $R \bowtie S$ 的结果与 U 进行自然连接最合适? 说明理由。
- (e) 使用(d)中选择的算法计算连接结果的 I/O 代价是多少?
- (f) 如果 $R \bowtie S$ 的结果大于 49 块，那么使用什么算法将 $R \bowtie S$ 的结果与 U 进行自然连接最合适? 说明理由。
- (g) 使用(f)中选择的算法计算连接结果的 I/O 代价是多少?

1、考虑下面的三个事务和它们的一个调度 S（时间从上往下依次增大）。判断 S 是否是冲突可串行化的调度？要求画出优先图并给出判断依据。

T0	T1	T2
r0(A)		
w0(A)		
		r2(A)
		w2(A)
	r1(A)	
r0(B)		
		r2(B)
w0(B)		
		w2(B)
	r1(B)	

2、设 T1、T2、T3 是如下三个事务：

T1: $A := A + 4$

T2: $A := A * 3$

T3: $A := A^2$

初始 $A = 2$

(1) 设三个事务都遵守两段锁协议，按 T2-T3-T1 的顺序执行，请给出一个不产生死锁的可串行化调度，并给出最终 A 的结果

(2) 若这三个事务都遵循两段锁协议，请给出一个产生死锁的调度。

3、考虑两个事务 T1, T2。其中，T1 显示账户 A 与 B 的内容：

T1: Read(B);
Read(A);
Display(A+B).

T2 表示从账户 B 转 50 美元到账户 A, 然后显示两个账户的内容：

T2: Read(B);
 $B := B - 50$;
Write(B);
Read(A);
 $A := A + 50$;
Write(A);
Display(A+B).

请给出一个满足时间戳协议的一个可能的调度。（注：Display(A+B)仅表示显示账户 A 和账户 B 的内容）

4、一个带检查点的日志内容如下，结束处发生了故障，请简述恢复算法的过程，并给出 Undo-List 和 Redo-List，以及数据库系统恢复后的 A,B,C 的值。

Start of the logs
...
<T4 start>
<T4, B, 1000, 1700>
<T5 start>
<checkpoint {T4, T5}>
<T5, C, 300, 100>
<T5 commit>
<T6 start>
<T6, A, 700, 600>
<T4, B, 1000>
<T4 abort>
← System crash, start recovery

5、设一个数据库系统启动后中，执行 4 个事务 T0、T1、T2 和 T3。四个事务的内容如下：

T0: $A := A + 20$ (读入数据库元素 A 的值，加上 20 后，再写回 A 的值)

T1: $B := B - 10$ (读入数据库元素 B 的值，减去 10 后，再写回 B 的值)

T2: $C := C * 2$ (读入数据库元素 C 的值，乘以 2 后，再写回 C 的值)

T3: $D := D + 15$ (读入数据库元素 D 的值，加上 15 后，再写回 D 的值)

除了这四个事务外，系统中无其他事务执行。设四个事务开始前，数据库元素 A、B、C、D 的值分别为 $A = 50$ ， $B = 30$ ， $C = 35$ ， $D = 15$ 。在执行这四个事务的过程中，系统发生了故障。系统重启后，经故障恢复，数据库元素 A、B、C、D 的值被恢复为 $A = 50$ ， $B = 20$ ， $C = 70$ ， $D = 15$ 。故障恢复时，数据库系统日志文件中包含如下 12 条日志记录，这里只给出部分日志记录。已知该数据库管理系统使用基于 undo-redo 日志的故障恢复技术，这段日志中仅有 1 个不停机检查点 (又称模糊检查点)。

1	
2	
3	
4	<start checkpoint (T0, T2)>
5	<end checkpoint>
6	
7	
8	
9	
10	
11	
12	

请根据上述信息，回答下列问题：

a.将日志文件补充完整，直接上面的日志文件中填写。一个事务 T 启动时向日志文件中写入日志记录<T, start>;提交时向日志文件中写入日志记录<T, commit>;

中止时向日志文件中写入日志记录<T, abort>; 对数据库元素 X 进行修改时向日志文件中写入日志记录<T, X,X 的旧值, X 的新值>。

b.在故障恢复过程中，哪些事务需要 redo，哪些事务需要 undo。说明理由。

c.在故障恢复过程中，还会向日志文件添加什么日志记录？说明理由。