

一、SQL 语言实现数据操作

现有一个名称为 Amusement 的数据库，其包含如下 3 张表，分别是 Actors、Movies、StarIn，分别表示演员表、电影表和演员参演电影表。一个演员能参演多部电影，一部电影由多位演员完成。数据表具体定义如下：

--演员表：

```
CREATE TABLE Actors(  
    Aid CHAR(10) PRIMARY KEY,    --演员ID  
    Aname CHAR(30) NOT NULL,      --演员姓名  
    Acompany CHAR(8)              --演员所属公司  
)
```

--电影表：

```
CREATE TABLE Movies(  
    Mno CHAR(10) PRIMARY KEY,    --电影编号  
    Mtitle CHAR(8) NOT NULL,     --电影名称  
    Mdate DATETIME              --电影开拍时间  
)
```

--演员参演电影表：

```
CREATE TABLE StarIn(  
    Aid CHAR(10),    --演员ID  
    Mno CHAR(10),    --电影编号  
    Rewards INT      --参演酬金  
)
```

请根据上述信息完成以下题目。

1. 查询 2019 年开始拍摄且电影名称中含有“复仇者联盟”字样的所有电影的名称

```
SELECT Mtitle FROM Movies WHERE Mtitle likes “%复仇者联盟%” AND Mdate = 2019
```

2. 查询 ID 为“A201901”的演员参演的所有电影的名称

```
SELECT Mtitle FROM Movies WHERE Mno IN(  
SELECT Mno FROM StarIn WHERE ID='A201901'  
)
```

3. 查询至少参演了 10 部电影的演员的姓名

```
SELECT Aname FROM Actors WHERE Aid IN (  
SELECT Aid FROM StarIn  
GROUP BY Aid  
HAVING COUNT(*)>=10  
)
```

4. 请说明下面 SELECT 语句完成的功能

```
SELECT Mtitle  
FROM Movies  
WHERE NOT EXISTS(  
    SELECT * FROM StarIn
```

```
WHERE Movies.Mno=StarIn.Mno AND StarIn.Aid IN
    (SELECT Aid FROM Actors WHERE Acompany='新影业公司')
)
```

查询新影业公司的演员没有参演过的电影名称

二、关系代数实现数据库操作分析

下面两张表分别描述了专业和学生的信息，其中“专业编号”是专业表的主码，“学号”是学生表的主码，“专业编号”是学生表的外码，其参照了专业表的主码。

表 1: 专业表

专业编号	专业名称
C01	计算机科学与技术
C02	网络工程
C03	物联网
C04	信息管理与信息系统

表 2: 学生表

学号	姓名	性别	专业编号
14001	张思思	女	C03
14002	李云刚	男	NULL
14003	王琼	男	C01
14004	刘莹	女	C02
14005	杨启明	男	C01

请分析下面给出的关系代数表达式完成的功能是否有区别，并说明其区别。（假设学生表中不存在姓名相同的学生）

- 1) $\prod_{\text{姓名}} (\sigma_{\text{专业名称} = \text{'计算机科学与技术'}} (\text{学生} \bowtie \text{专业})) \setminus$
- 2) $\prod_{\text{姓名}} (\sigma_{\text{专业名称} = \text{'计算机科学与技术'}} (\text{学生} \times \text{专业}))$
- 3) $\prod_{\text{姓名}} (\text{学生} \bowtie \sigma_{\text{专业名称} = \text{'计算机科学与技术'}} (\text{专业}))$
- 4) $\prod_{\text{姓名}} (\text{学生}) - \prod_{\text{姓名}} (\text{学生} \bowtie \sigma_{\text{专业名称} \neq \text{'计算机科学与技术'}} (\text{专业}))$

(1)

姓名
王琼
杨启明

功能：查询专业名称为‘计算机科学与技术’的学生姓名。

(2)

姓名
张思思
李云刚
王琼

刘莹
杨启明

功能：查询学生表中的所有学生姓名。

(3)

姓名
王琼
杨启明

功能：查询专业名称为‘计算机科学与技术’的学生姓名。

(4)

姓名
李云刚
王琼
杨启明

功能：查询专业名称不确定与专业名称为‘计算机科学与技术’的学生姓名。

区别：

这四条关系代数表达式都是对学生姓名进行查询，但限定的条件不一样，导致部分功能不同。虽然关系代数表达式 1 和 3 都是同样的功能，但是 1 是先连接再选择，2 是先选择后连接，明显，后者的效率更高。

三、数据库安全分析

图 1 展示了一个权限传播图，描述 DBA、USER1、USER2、USER3 四个用户之间在 Student 表上的操作权限传播情况，编号①-④对应的操作权限授予命令及其执行用户展示在表 3 中。

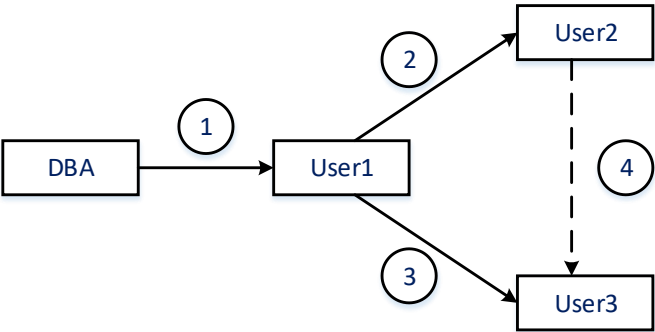


图 1 操作权限传播示意图

表 3 授权命令语句列表

序号	执行者	授权命令语句
①	DBA	GRANT INSERT, DELETE, UPDATE ON Student TO User1 WITH GRANT OPTION
②	USER1	GRANT INSERT, DELETE ON Student TO User2 WITH GRANT OPTION
③	USER1	GRANT INSERT, UPDATE ON Student TO User3 WITH GRANT OPTION
④	USER2	GRANT INSERT ON Student TO User3

若 User1 执行如下语句，

REVOKE INSERT, DELETE ON Student FROM User2 CASCADE //级联回收

请回答 User2、User3 所拥有对 Student 表的操作权限情况，请在表 4 中对应的位置画√和×。

表 4 操作权限结果表

用户	INSERT	DELETE	UPDATE
User2	×	×	×
User3	√	×	√

四、数据库故障恢复分析

- 图 2 给出了多个事务执行的状态图，在 t_1 时刻建立了检查点，并在 t_2 时刻发生了系统故障，请分析 1-6 号事务在进行数据库恢复时应分别采取什么恢复策略，请在表 5 中对应的位置画√？

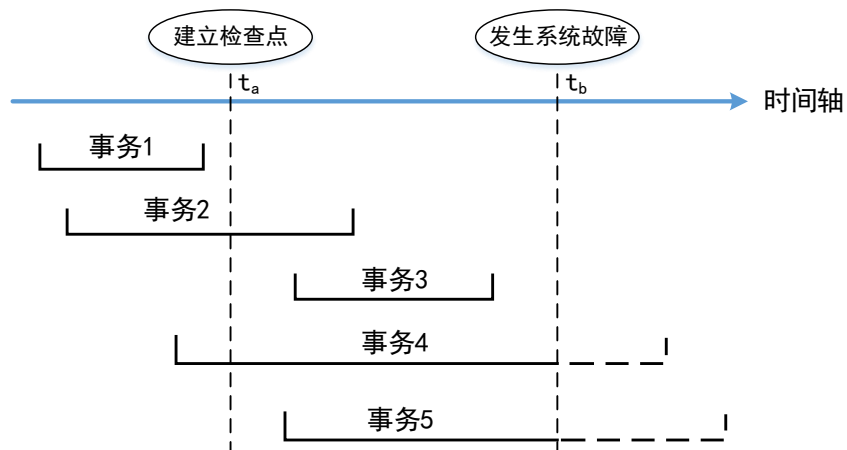


图 2 事务执行状态图

表 5 事务恢复策略信息表

	不处理	UNDO	REDO
事务 1	✓		
事务 2			✓
事务 3			✓
事务 4		✓	
事务 5		✓	

五、并发调度分析

设有两个事务 T_1 和 T_2 ，其操作如下：

T_1 : $X = X + 10$; T_2 : $X = X * 5$;

若 X 的初始值为 2，请说明下面调度是否是正确的调度，若不正确，请指出调度不正确的原因以及使用何种机制能保证调度的正确性。

T_1	T_2
读 $X=2$;	
	读 $X=2$;
$X = X + 10 = 12$	
	$X = X * 5 = 10$;
写 $X=12$;	
	写 $X=10$;

CPU 时间片

运行结果 X=10 是不正确的，这是由于在并发执行过程中出现了丢失修改的错误，T1 写回的数据丢失了，应采用封锁机制实现并发控制，具体修改如下

T1	T2
XLOCK(X)	
R(X)=2	
	XLOCK(X)
X=X+10=12	等待
W(X)=12	等待
COMMIT	等待
UNLOCK(X)	
	获得 XLOCK(X)
	R(X)=12
	X=X*5=60
	W(X)=60
	COMMIT
	UNLOCK(X)

六、完整性约束分析

有如下建表语句，请结合表 6（学生表），指出表中存在的违背完整性约束的情况。

```
CREATE TABLE Student
```

```
( Sno CHAR(6),
```

```
 Sname CHAR(20) NOT NULL,
```

```
 Sage INT
```

```
 CONSTRAINT C1 CHECK (Sage < 30),
```

```
 Ssex CHAR(2)
```

```
 CONSTRAINT C2 CHECK (Ssex IN ('男', '女')),
```

```
 CONSTRAINT C3 PRIMARY KEY(Sno)
```

```
);
```

表 6 学生表

序 号	Sno	Sname	Sage	Ssex
1	90001	王刚	35	男
2	90002	周敏	20	女

3	90002	张宁	18	女
4	90004	赵全	25	Male

1.2 号与 3 号的 Sno 相同，违背了实体完整性约束

2.1 号 Sage>30,违背了自定义的年龄必须小于 30 的完整性约束

3.4 号 Ssex 没有在男、女之间取值，违背了用户自定义的完整性约束

七、求解最小函数依赖集

设有关系模式 $R(U, F)$ ，其中 $U = \{X, Y, Z, M\}$ ， $F = \{XY \rightarrow Z, Z \rightarrow X, Y \rightarrow XZ, M \rightarrow X, M \rightarrow Y\}$ ，请完成如下操作：

(1) 求属性 M 关于函数依赖集 F 的属性闭包 $(M)_F^+$ 。(5 分)

(2) 请给出 F 的最小函数依赖集。(5 分)

(1) $XYZM$

(2) $F_m = \{Y \rightarrow Z, Z \rightarrow X, M \rightarrow Y\}$

八、关系模式评价与分解

现有如下关系模式，用于记录学生参加竞赛信息：

参加竞赛（竞赛编号，竞赛名称，竞赛组织者，竞赛开始日期，学号，学生姓名，获奖等级）

若有以下规定，

(1) 所有竞赛的编号是唯一的；

(2) 一名学生可以在不同日期参加不同的竞赛，但在同一日期只能参加一个竞赛；一个竞赛可有多名学生参加；

(3) 每名学生参加一次竞赛，会获得一个不同等级的奖项。

请完成如下问题：

1) 请写出该关系模式的码。(5 分)

2) 该关系模式满足第几范式？请说明理由。(5 分)

3) 请给出关系模式的一个有效分解，使得分解后的关系模式均满足第三范式？(5 分)

(1) 竞赛编号，学号

(2) 满足第一范式；存在非主属性（竞赛名称）对码（竞赛编号，学号）的部分函数依赖

(3)

竞赛（竞赛编号，竞赛名称，竞赛组织者，竞赛开始日期）

学生（学号，姓名）

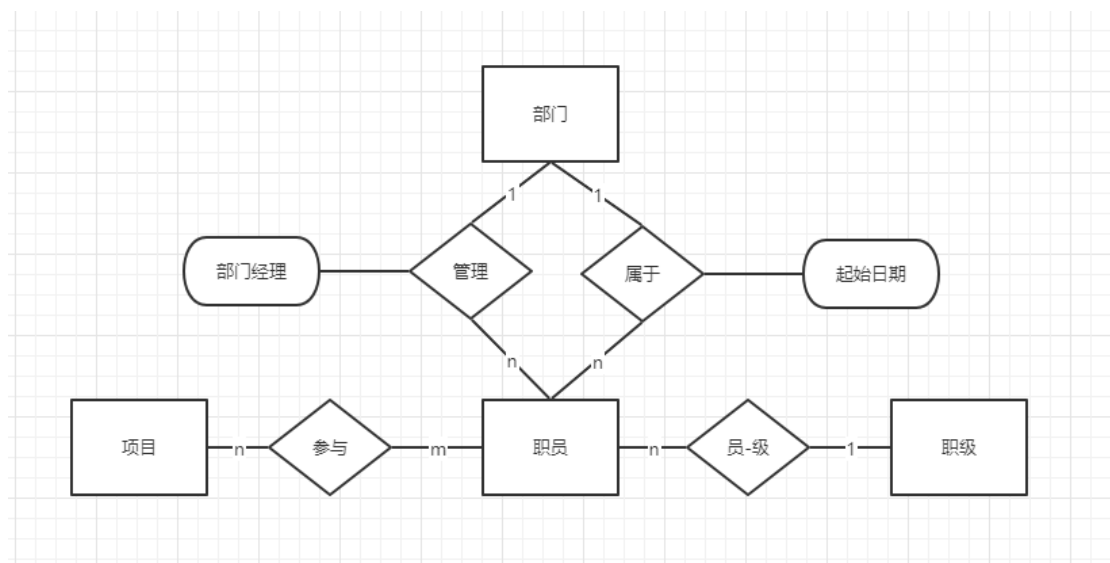
获奖记录（学号，竞赛编号，获奖等级）

九、数据库设计与优化

软件公司下设多个部门，包括技术开发部门、产品维护部门、财务部门等。每个部门有若干职员，显然每位职员只能在一个部门工作，且有一个加入部门工作的起始日期。每位职员会关联一个职级，其记录了职员的级别、月薪等信息。每个部门中会有一名员工被任命为部门经理。公司会承接各类项目，由不同部门的职员合作完成，一名职员可以参与多个项目。

“部门”的属性有部门编号、部门名称和联系电话，“职员”的属性有职员编号、职员姓名；“职级”的属性有职级编号、职级描述、月薪；“项目”的属性有项目编号、项目名称、项目金额。请根据上面的内容完成如下题目：

- (1) 请根据给定的需求，使用 ER 图给出该数据库的概念模型（只需画出联系的属性，不必画出实体的属性）；



- (2) 请将(1)中得到的概念模型转换为关系模式集合，并指出每一个关系模式的主码和外码。

项目（项目编号，项目名称，项目金额）

项目承接记录（项目编号，职员编号） 外码：项目编号，职员编号

部门（部门编号，部门名称，联系电话）

职员（职员编号，部门编号，职级编号，部门经理，起始日期，职员姓名）

外码：部门编号 职级编号

职级（职级编号，职级描述，月薪）

（3）若需求分析存在如下结论，

“统计指定职员参加的各类项目的名称”是频繁执行的查询，且职员参加项目的信息表数据量很大；

请问应在数据库存储结构设计方面如何考虑，实现查询性能优化，请给出初步的设计思路。

使用聚簇功能将每一个职员参加的所有项目集中存放，使得读取每一个物理块可以得到多个满足查询条件的元组，减少访问磁盘的次数。

综合题

假设某商业集团数据库中有一关系模式 R 如下：

R（商店编号，商品编号，数量，部门编号，负责人）

如果规定：

- （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。

试回答下列问题：

- （1）根据上述规定，写出关系模式 R 的基本函数依赖；

（商店编号，商品编号）→部门编号

（商店编号，商品编号）→数量

（商店编号，部门编号）→负责人

- （2）找出关系模式 R 的候选码；

侯选码（商店编号，商品编号）

- （3）试问关系模式 R 最高已经达到第几范式？为什么？

2NF，存在传递函数依赖。

- （4）如果 R 不属于 3NF，请将 R 分解成 3NF 模式集。

R1 (商店编号, 商品编号, 数量, 部门编号); R2 (商店编号, 部门编号, 负责人)

知识点

- 1) 笛卡尔积定义
- 2) 码，主码，候选码，主属性，非主属性定义
- 3) 选择 投影 连接（自然连接，等值连接） 除运算 供应商习题
- 4) Sql 增删改查
- 5) 集合查询 GROUP BY、AVING、IN、 EXISTS 的使用
- 6) 视图操作（建立视图，视图的作用）
- 7) 自主存取控制的授权与收权
- 8) 强制存取控制
- 9) 视图机制
- 10) 实体完整性，参照完整性，用户自定义的完整性（CONSTRAINT CHECK）
- 11) 1 范式，2 范式，3 范式，BC 范式的定义
- 12) 求闭包、求候选码、求最小函数依赖
- 13) 模式分解（无损连接性）（具体问题具体分析）
- 14) E-R 图的设计→找出候选码
- 15) 索引的概念，什么情况下建立索引，聚簇的概念
- 16) 事务的四个特性
- 17) 事务内部故障，系统故障，介质故障的分析，以及如何修复
- 18) 并发操作的问题：丢失修改，不可重复读，读脏数据 的判定
- 19) 如何解决并发操作的问题？封锁机制
- 20) X 锁，S 锁
- 21) 并发调度的可串行性（分析这里的并发调度是否可串行，为什么？）