

4、 //考虑关系 $r_1(A, B, C)$ ， $r_2(C, D, E)$ 和 $r_3(E, F)$ ， A, C, E 分别是其主码，假定 r_1 有 1000 个元组， r_2 有 1500 个元组， r_3 有 750 个元组，估计 $r_1 \bowtie r_2 \bowtie r_3$ 的大小，并给出一个计算连接的高效策略。

一、简答题（24 分）

1、 简述文件系统和数据库系统之间的区别与联系。

5、 证明为什么两阶段封锁协议能够保证事务集合的可串行化。

2、 //简述何时需要建立聚簇索引，以及它的优缺点。

6、 对于下列调度，判断其是否为冲突可串行化？如果是冲突可串行化的，给出与之等价的一个串行调度序列；如果不是，请说明理由。

3、 设关系模式 $R(A, B, C)$ 上有多值依赖 $A \twoheadrightarrow B$ 成立。如果已知 R 的当前关系中存在着三个元组 $(a,b1,c1)$ 、 $(a,b2,c2)$ 、 $(a,b3,c3)$ ，那么这个关系中至少还应该存在哪些元组？

T_1	T_2	T_3	T_4	T_5
	Read(A)			
Read(Q)				
				Read(C)
Write(Q)				
			Read(A)	
	Write(Q)			
Write(B)				
		Read(Q)		
		Read(M)		
			Read(Q)	
			Read(M)	
				Read(M)
				Write(M)

二、综合题（21 分）。

（一）数据库设计（9 分）

某企业的人力资源管理需求如下：

企业有员工，员工需要关注的信息包括：员工号、员工姓名、性别、所在部门、出生日期、年龄、民族、亲属信息；企业有若干部门，部门需要关注的信息包括：部门编号、部门名称、部门地址；企业有若干岗位，岗位需要关注的信息包括：岗位名称、岗位职责；每个员工都隶属于一个部门，一个部门有多个员工，一个部门有一个负责人；一个员工可以承担多个岗位的工作职责，一个岗位可以有多个员工承担；一个部门可以拥有多个岗位，一个岗位也可以在多个部门中存在。员工的亲属信息关注身份证号码、与员工的关系和姓名，员工有多位亲属。

要求：

- 1、用 E-R 图表示上述内容。
- 2、将 E-R 图转换成相应的关系模式。

（二）规范化（9 分）

给定关系模式 $R(P, W, X, Y, Z)$ ，其上的函数依赖集 $F=\{ Z\rightarrow X, X\rightarrow P, XY\rightarrow WP, XYP\rightarrow ZW \}$ ，

要求：1、计算 $(Z)^+$ 。 2、求 F 的正则覆盖 F_c 。

3、求 R 的一个候选码。

4、将 R 分解，使其满足 BCNF 且分解具有无损连接性。

5、将 R 分解，使其满足 3NF 且分解具有无损连接性与保持依赖性。

(三) 证明题（3 分）

给定关系模式 R ，其上的函数依赖集 F ，及其分解 $\{R_1, R_2\}$ ，证明如果 $R_1 \cap R_2 \rightarrow R_1$ 属于 F^+ ，则 R_1 和 R_2 是 R 的无损分解。

三、运算题（共 45 分）

有关系模式 S(sno,sname,dno,sex, dorm, monitor), C(cno,ctitle,credit), SC(sno,cno,sore), D(dno,dname)。
关系 S、C、SC 和 D 分别表示学生信息、课程信息、学生选课情况和院系信息。其属性分别表示如下：sno—学生编号，sname—学生姓名，dno—院系编号，sex—性别，dorm—宿舍，monitor—班长，cno—课程编号，ctitle—课程名称，credit—课程学分，score—成绩，dname—院系名称。

关系模式的实例如下：

S

sno	sname	dno	sex	dorm	monitor
s1	赵晨	d1	M	101	s1
s2	许亮	d2	M	101	
s3	肖莹	d1	F	201	s1
s4	陈静	d3	F	201	
s5	王勤勤	d1	M	101	s1

C

cno	ctitle	credit
c1	计算机组成原理	4
c2	离散数学	3
c3	数据库系统	4
c4	计算机网络	4

SC

sno	cno	score
s1	c1	80
s1	c2	75
s1	c3	45
s2	c2	70
s3	c1	90
s3	c2	60

D

dno	dname
d1	计算机学院
d2	物理学院
d3	化学学院
d4	生命科学学院

（一）针对上述实例，给出下列表达式的结果（只写出结果即可，无需计算过程，每小题 3 分，共 24 分）

1、 $\prod_{\text{sname}} (\sigma_{\text{sex}='M'}(S))$

2、 $\text{snoGcount(cno)}(SC)$

3、 $\prod_{\text{dname}} (D) - \prod_{\text{dname}} (S \bowtie D)$

4、 $\prod_{\text{ctitle}} ((\sigma_{\text{sno}='s1' \wedge \text{score}>60}(SC)) \bowtie (C))$

5、 $\prod_{\text{a.sno}} (\sigma_{\text{a.score}<\text{b.score} \wedge \text{b.sno}='s2' \wedge \text{a.cno}='c2' \wedge \text{b.cno}='c2'} (\rho_{\text{a}}(SC) \times \rho_{\text{b}}(SC)))$

姓名

学号

级

专业

学院

山东大学 2016-2017 学年 1 学期 数据库系统 课程试卷

6、 $\Pi_{\text{sname}}(\Pi_{\text{sno},\text{cno}}(\text{SC}) \div \Pi_{\text{cno}}(\sigma_{\text{sno}=\text{'s3'}}(\text{SC}))) \bowtie \text{S}$

7、 $\{ t \mid \exists r \in C (t[\text{ctitle}] = r[\text{ctitle}] \wedge r[\text{cno}] = \text{'c1'}) \}$

8、 $\{ t \mid \exists r \in D (t[\text{dno}] = r[\text{dno}]) \wedge \neg \exists v \in S (v[\text{dno}] = t[\text{dno}]) \}$

(二) 对于上述关系模式，用 SQL 完成下列操作（只写出 SQL 即可，不需要执行结果，每小题 3 分，共 21 分）

1、查询化学学院学生姓名，按照学生姓名降序排序。

2、查询没有选修“c1”课程的学生学号。

3、查询与自己的班长住在同一个宿舍的学生学号。

4、查询获得“计算机网络”课程最高分的学生学号。

5、查询学生人数最少的院系名称。

6、查询被所有学生都选修的课程名称。

7、如果学生某门课的考试成绩低于该门课的平均成绩，则将该学生该门课的成绩提高 5%。

一、简答题（24 分）

1、简述文件系统和数据库系统之间的区别与联系。

区别：

文件系统	数据库系统
程序和数据有一定的联系	程序和数据分离
用操作系统中的存取方法对数据进行管理	用 DBMS 对数据进行管理
实现以文件为单位的数据共享	实现以记录和字段为单位的数据共享
用文件将数据长期保存在外存上	用数据库统一存储数据

联系：均为数据组织的管理技术，均有数据管理软件管理数据，程序与数据之间用存取方法进行转换，数据库系统是基于文件系统的。

评分细则：本题 4 分，区别和联系各占 2 分，回答不完全者酌情给分

2、简述何时需要建立聚簇索引，以及它的优缺点。

如果经常需要按照某搜索码（在文件中查找记录的属性或属性集）做查找或者排序操作，可在该搜索码上建立聚簇索引，使得数据文件的记录顺序按照搜索码的顺序排序。

优点：利用搜索码上的聚簇索引，可以基于搜索码属性做高效的查找和排序操作；聚簇索引可以采用稀疏方式存储，节省索引的存储成本。

缺点：聚簇索引要求数据文件的记录顺序按照搜索码的顺序排序，因而在一个关系上只能建立一个聚簇索引。如果基于非聚簇索引（辅助索引）的搜索码进行查找操作，成本会相对高一些。

评分细则：本题 4 分，优点和缺点各占 2 分，回答不完全者酌情给分

3、设关系模式 $R(A, B, C)$ 上有多值依赖 $A \twoheadrightarrow B$ 成立。如果已知 R 的当前关系中存在着三个元组 $(a, b1, c1)$ 、 $(a, b2, c2)$ 、 $(a, b3, c3)$ ，那么这个关系中至少还应该存在哪些元组？

$(a, b1, c2)$, $(a, b2, c1)$, $(a, b1, c3)$, $(a, b3, c1)$, $(a, b2, c3)$, $(a, b3, c2)$

评分细则：本题 4 分，回答不完全者酌情给分

4、考虑关系 $r_1(A, B, C)$ ， $r_2(C, D, E)$ 和 $r_3(E, F)$ ， A, C, E 分别是其主码，假定 r_1 有 1000 个元组， r_2 有 1500 个元组， r_3 有 750 个元组，估计 $r_1 \bowtie r_2 \bowtie r_3$ 的大小，并给出一个计算连接的高效策略。

连接 r_1 和 r_2 会产生一个至多 1000 个元组的关系，因为 C 是 r_2 的主码。同样，将该结果与 r_3 相连，会产生一个至多 1000 个元组的关系，因为 E 是 r_3 的主码。这样最终的结果关系会至多 1000 个元组。

计算连接的一个有效的策略是在 C 属性上为关系 r_2 创建索引，在 E 属性上为 r_3 创建索引，这样对于 r_1 中的任一个元组，使用 r_2 的索引来查找匹配 r_1 中的 C 值的至多一个元组；使用 r_3 的索引查找匹配 r_2 中 E 值的至多一个元组。

评分细则：本题 4 分，估计大小和连接策略各占 2 分，回答不完全者酌情给分

5、证明为什么两阶段封锁协议能够保证事务集合的可串行化。

采用反证法，假定两阶段封锁协议不能保证可串行化，则存在一个遵循两阶段封锁协议的事务集合 T_0, T_1, \dots, T_{n-1} 产生了一个非串行的调度，那么在其优先图中会有一个环。不失一般性，假定优先图中的环是 $T_0 T_1 T_2 \dots T_{n-1} T_0$ ，让 a_i 对应事务 T_i 获得最后一个锁的时间（即事务 T_i 的封锁点）。因而对于所有的事务 $T_i < T_j$ ，都有 $a_i < a_j$ ，对于这个环，我们就有 $a_0 a_1 a_2 \dots a_{n-1} a_0$ ，存在矛盾，因此假设不成立。

所以，两阶段封锁协议能够保证事务集合的可串行化。

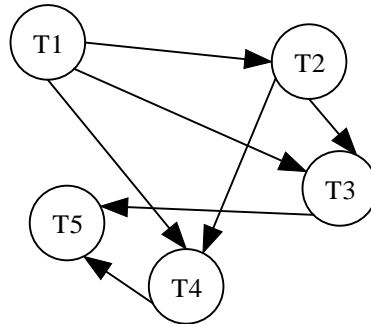
评分细则：本题 4 分，回答不完全者酌情给分

6、对于下列调度，判断其是否为冲突可串行化？如果是冲突可串行化的，给出与之等价的一个串行调度序列；如果不是，请说明理由。

T_1	T_2	T_3	T_4	T_5
	Read(A)			
Read(Q)				
				Read(C)
Write(Q)				
			Read(A)	
	Write(Q)			
Write(B)				
		Read(Q)		
		Read(M)		
			Read(Q)	
			Read(M)	
				Read(M)

				Write(M)
--	--	--	--	----------

答：画出优先图：



图中没有环，所以是可串行化的。

等价的串行调度为： $T_1 T_2 T_3 T_4 T_5$ 或 $T_1 T_2 T_4 T_3 T_5$

评分细则：本题 4 分，可串行化判断和调度序列各占 2 分，回答不完全者酌情给分

二、综合题（21 分）

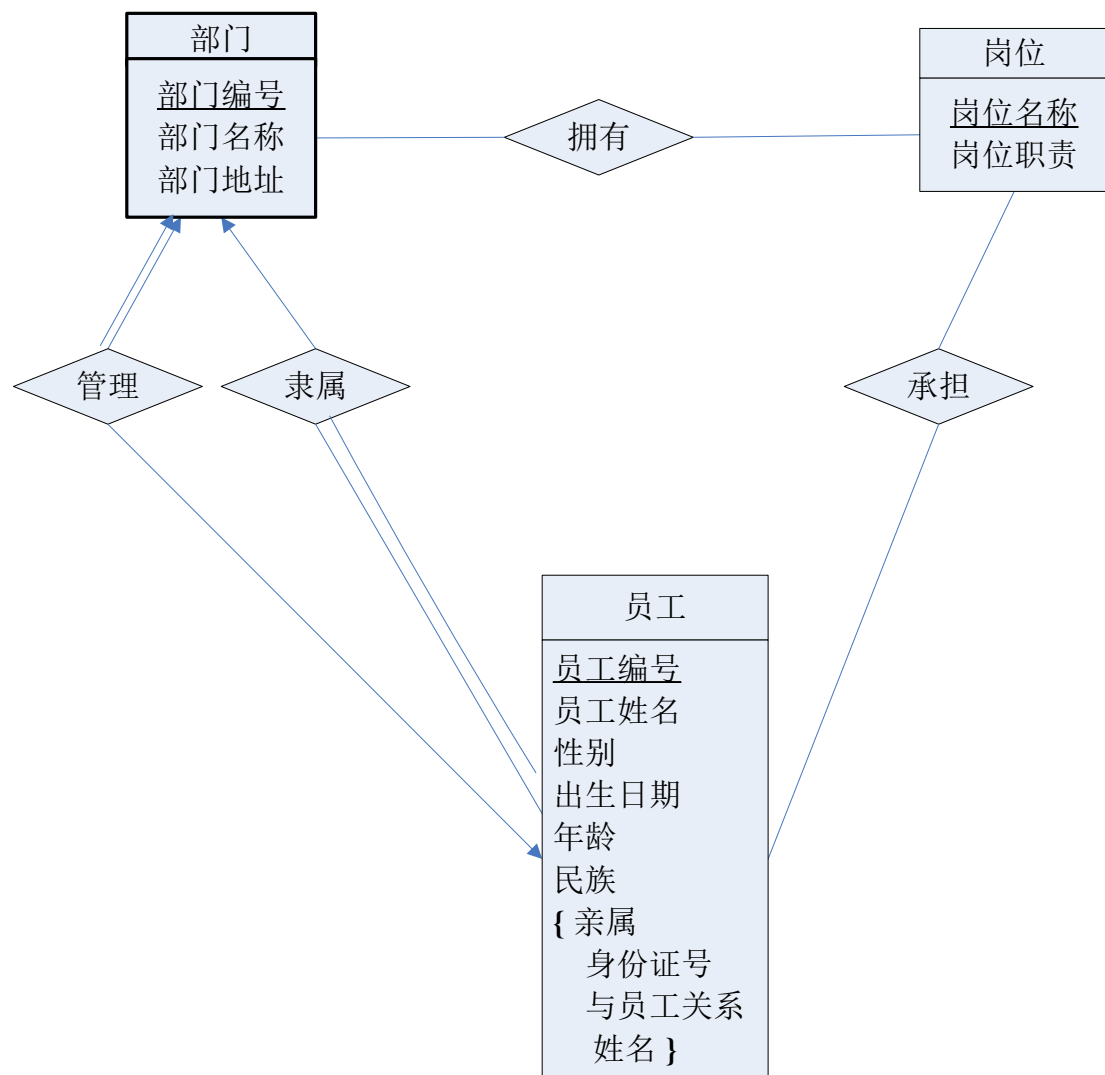
（一）数据库设计（9 分）

某企业的人力资源管理需求如下：

企业有员工，员工需要关注的信息包括：员工号、员工姓名、性别、所在部门、出生日期、年龄、民族、亲属信息；企业有若干部门，部门需要关注的信息包括：部门编号、部门名称、部门地址；企业有若干岗位，岗位需要关注的信息包括：岗位名称、岗位职责；每个员工都隶属于一个部门，一个部门有多个员工，一个部门有一个负责人；一个员工可以承担多个岗位的工作职责，一个岗位可以有多个员工承担；一个部门可以拥有多个岗位，一个岗位也可以在多个部门中存在。员工的亲属信息关注身份证号码、与员工的关系和姓名，员工有多位亲属。

要求：

- 1、用 E-R 图表示上述内容。
- 2、将 E-R 图转换成相应的关系模式。



关系模式：

部门（部门编号，部门名称，部门地址，负责人员工编号）

岗位（岗位名称，岗位职责）

员工（员工编号，员工姓名，性别、出生日期、年龄、民族、所在部门编号）

员工亲属（员工编号，身份证号，与员工的关系，姓名）

拥有（部门编号，岗位名称）

承担（员工编号，岗位名称）

评分细则：本题 9 分，第 1 题 5 分，第 2 题 4 分，回答不完全者酌情给分

（二）规范化（9 分）

给定关系模式 $R(P, W, X, Y, Z)$ ，其上的函数依赖集 $F = \{ Z \rightarrow X, X \rightarrow P, XY \rightarrow WP, XYP \rightarrow ZW \}$ ，

要求：1、计算 $(Z)^+$ 。 2、求 F 的正则覆盖 F_c 。

3、求 R 的一个候选码。

4、将 R 分解，使其满足 BCNF 且分解具有无损连接性。

5、将 R 分解，使其满足 3NF 且分解具有无损连接性与保持依赖性。

答：1、 $\{Z, X, P\}$

2、 $\{ Z \rightarrow X, X \rightarrow P, XY \rightarrow WZ \}$

3、 (X, Y) 或 (Y, Z)

4、 $(X, Z) (Z, P) (W, Y, Z)$ 或 $(X, P) (X, Z) (W, Y, Z)$

5、 $(X, Z) (X, P) (X, Y, W, Z)$ 或 $(X, P) (X, Y, W, Z)$

评分细则：本题 9 分，第 1 题 1 分，第 2 至 5 题每题 2 分，回答不完全者酌情给分

（三）证明题（3 分）

给定关系模式 R ，其上的函数依赖集 F ，及其分解 $\{R_1, R_2\}$ ，证明如果 $R_1 \cap R_2 \rightarrow R_1$ 属于 F^+ ，则 R_1 和 R_2 是 R 的无损分解。

答：假设 $\forall r \in R$ ，分解得到 r_1, r_2 ，只需证明 $r = r_1 \bowtie r_2$ 。

容易知道， $\forall t \in r, t \in r_1 \bowtie r_2$ ，所以 $r \subseteq r_1 \bowtie r_2$ 。

令 $\alpha = R_1 \cap R_2$ ， $\beta = R_1 - \alpha$ ， $\gamma = R_2 - \alpha$

$\forall t \in r_1 \bowtie r_2$ ，得 $t[\alpha\beta] \in r_1$ ， $t[\alpha\gamma] \in r_2$ ；

设 $t[\alpha\beta]$ 在 r 里面对应元组为 t_1 ，则 $t_1[\alpha] = t[\alpha]$ ， $t_1[\beta] = t[\beta]$ ；

设 $t[\alpha\gamma]$ 在 r 里面对应元组为 t_2 ，则 $t_2[\alpha] = t[\alpha]$ ， $t_2[\gamma] = t[\gamma]$ ；

由 $\alpha \rightarrow R_1$ 得 $\alpha \rightarrow \beta$ ，又因为 $t_1[\alpha] = t_2[\alpha] = t[\alpha]$ ，得 $t_1[\beta] = t_2[\beta]$ ，所以 $t_2[\beta] = t[\beta]$ ，所以 $t_2 = t$ ，所以 $t \in r$ ；

得 $r_1 \bowtie r_2 \subseteq r$ 。

因此 $r = r_1 \bowtie r_2$ 得证。

评分细则：本题 3 分，回答不完全者，酌情给分

三、运算题（共 45 分）

（一）针对上述实例，给出下列表达式的结果（只写出结果即可，无需计算过程，每小题 3 分，共 24 分）

1、 $\Pi_{\text{sname}} (\sigma_{\text{sex}='M'} (S))$

Sname
赵晨
许亮
王勤勤

评分细则：结果集关系模式错误得 0 分，结果集关系模式正确得 1 分，完全正确得 3 分

2、 $\text{snoGcount}(\text{cno})(SC)$

Sno	count(cno)
S1	3
S2	1
S3	2

评分细则：结果集关系模式错误得 0 分，结果集关系模式正确得 1 分，完全正确得 3 分

3、 $\Pi_{\text{dname}}(D) - \Pi_{\text{dname}}(S \bowtie D)$

Dname
生命科学学院

评分细则：结果集关系模式错误得 0 分，结果集关系模式正确得 1 分，完全正确得 3 分

4、 $\Pi_{\text{ctitle}}((\sigma_{\text{sno}='s1' \wedge \text{score}>60}(SC)) \bowtie (C))$

Ctitle
计算机组成原理
离散数学

评分细则：结果集关系模式错误得 0 分，结果集关系模式正确得 1 分，完全正确得 3 分

5、 $\Pi_{a.sno}(\sigma_{a.score < b.score \wedge b.sno='s2' \wedge a.cno='c2' \wedge b.cno='c2'}(\rho_a(SC) \times \rho_b(SC)))$

a. sno
S3

评分细则：结果集关系模式错误得 0 分，结果集关系模式正确得 1 分，完全正确得 3 分

6、 $\Pi_{sname} (\Pi_{sno, cno}(SC) \div \Pi_{cno}(\sigma_{sno='s3'}(SC))) \bowtie S$

Sname
赵晨
肖莹

评分细则：结果集关系模式错误得 0 分，结果集关系模式正确得 1 分，完全正确得 3 分

7、 $\{ t \mid \exists r \in C (t[ctitle]=r[ctitle] \wedge r[cno]='c1') \}$

ctitle
计算机组成原理

评分细则：结果集关系模式错误得 0 分，结果集关系模式正确得 1 分，完全正确得 3 分

8、 $\{ t \mid \exists r \in D (t[dno] = r[dno]) \wedge \neg \exists v \in S (v[dno] = t[dno]) \}$

dno
d4

评分细则：结果集关系模式错误 0 分，结果集关系模式正确 1 分，元组正确 2 分

(二) 对于上述关系模式，用 SQL 完成下列操作（只写出 SQL 即可，不需要执行结果，每小题 3 分，共 21 分）

1、查询化学学院学生姓名，按照学生姓名降序排序。

Select sname

From s,d

Where d.dno = s.dno and dname = '化学学院'

Order by sname desc;

评分细则：需要 2 个表作连接，连接正确得 1 分，选择条件正确得 1 分，排序正确得 1 分

2、查询没有选修“c1”课程的学生学号。

Select sno

From s

Where sno not in (select sno

from sc

where cno = 'c1');

评分细则：子查询正确的得 2 分。完全正确的得 3 分。

3、查询与自己的班长住在同一个宿舍的学生学号。

Select a.sno

From s a, s b

Where a.monitor=b.sno and a.dorm=b.dorm;

评分细则：需要 2 个相同的表作连接，连接正确得 2 分，选择条件正确得 1 分。

4、查询获得“计算机网络”课程最高分的学生学号。

Select sno

From sc natural join c

Where ctitle= '计算机网络' and

score =(select max(score) from sc, c

where sc.cno=c.cno and ctitle= '计算机网络')

评分细则：三个查询条件，每个条件 1 分。

5、查询学生人数最少的院系名称。

Select dname

From d

Where dno in (select dno from s

group by dno

having count(*) >= all (select count(*) from
s group by dno))

评分细则：group by 和 having 各 1 分，>=all 得 1 分

6、查询被所有学生都选修的课程名称。

Select ctitle

From c

```

Where not exists ( select *
                    from s
                    where not exists ( select *
                                      from sc
                                      where      s.sno=sc.sno      and
sc.cno=c.cno))

```

或

```

Select ctitle
  From c
  Where not exists ( (select sno
                     From s)
                  Except
                  (select sno
                   From sc
                   Where sc.cno=c.cno))

```

或

```

Select ctitle
  From c
  Where cno in (select cno
                from sc
                group by cno
                having count(*) = (select count(*) from s))

```

评分细则：基本结构正确的得 2 分；完全正确的得 3 分。

7、如果学生某门课的考试成绩低于该门课的平均成绩，将该学生该门课的考试成绩提高 5%。

```

Update  sc

```

```

Set score=score*1.05

```

```

Where score < (select avg(score) from sc sc1 where sc1.cno=sc.cno)

```

评分细则：update/set 正确的得 1 分，子查询正确的得 2 分。