

主管
领导
审核
签字

哈尔滨工业大学 2021 学年春季学期

数据库系统 试题 A

题号	一	二	三	四	五	六	七	八	九	十	总分
得分											
阅卷人											

片纸鉴心 诚信不败

授课教师

姓名

学号

院系

一、单选题（5分，每题1分）

- 下列存储介质中属于按块访问的是（ ）。
A. 寄存器 B. 高速缓存 C. 主存 D. 硬盘
- 已知关系 $R(A, B)$ 和 $S(B, C)$ 上的域关系演算表达式 $\{(a, b, c) | (a, b) \in R \wedge (b, c) \in S\}$ 。该表达式与下面的关系代数表达式（ ）等价。
A. $R - S$ B. $R \cap S$ C. $R \cup S$ D. $R \bowtie S$
- 关于关系模式范式之间的关系，下列说法正确的是（ ）。
A. $1NF \subseteq 2NF \subseteq 3NF \subseteq BCNF$ B. $BCNF \subseteq 1NF \subseteq 2NF \subseteq 3NF$
C. $BCNF \subseteq 3NF \subseteq 2NF \subseteq 1NF$ D. $3NF \subseteq 2NF \subseteq 1NF \subseteq BCNF$
- 当使用 2PL 进行并发控制时，如果事务 T1 已经获得了对象 A 上的共享锁且尚未释放该锁，那么当事务 T2 准备写 A 时，T2 会请求对 A 加（ ），该加锁请求会被（ ）。
A. 共享锁，拒绝 B. 共享锁，批准 C. 互斥锁，拒绝 D. 互斥锁，批准
- 下列属于冲突的是（ ）。
A. 事务 T1 读对象 A，事务 T2 读对象 A
B. 事务 T1 读对象 A，事务 T2 写对象 B
C. 事务 T1 读对象 A，事务 T2 写对象 A
D. 事务 T1 写对象 A，事务 T2 写对象 B

二、填空题（10分，每空1分）

- 数据库的物理独立性是指当数据库的（ ）模式发生变化时，基于数据库的外模式（视图）编写的应用程序无需进行修改。
- 关系数据库的完整性约束分为实体完整性约束、（ ）和用户定义完整性约束。
- 设关系 R 中包含 r 个元组，关系 S 中包含 s 个元组，且 $r > s$ 。关系代数表达式 $R - S$ 的结果中至少包含（ ）个元组，最多包含（ ）个元组。
- 在数据库设计的 3 个阶段中，进行数据库存储与索引设计的阶段是（ ）。
- 关系模式 R 上的 2 个函数依赖集 F 和 G 等价当且仅当 $F^+ =$ （ ）。
- 如果关系 R 上一个索引的键是 R 的主键，则该索引是 R 上的（ ）索引。
- 一个线性哈希表中共有 6 个桶，桶号分别为 0, 1, 2, 3, 4, 5。设键值 K 的哈希值 $hash(K)$ 等于 7，则键值 K 属于（ ）号桶。
- 在基于代价的查询优化方法中，我们用一个查询计划的（ ）来估计它的代价。
- 一组并发事务产生了死锁当且仅当这组事务的等待图中存在（ ）。

三、查询题（20分）

已知一个关系数据库的模式如下：

- 关系 Emp(eid, ename, age, salary)记录员工数据，其中 eid 是员工工号（整型），ename 是员工姓名（字符串型），age 是员工年龄（整型），salary 是员工薪资（浮点型）；
- 关系 Dept(did, budget, managerid)记录部门数据，其中 did 是部门编号（整型），budget 是部门的经费数额（浮点型），managerid 是部门经理的工号（整型）；
- 关系 Works(eid, did, pct_time)记录员工在部门中工作的数据。一个员工可以在多个部门工作，pct_time 记录一个员工在一个部门的工作时间所占的比例（整型）。如果比例为40%，则 pct_time 为40。

请写出下列查询：

1. （4分）使用关系代数表达式查询“部门1中年龄在40岁以上（含40岁）的员工的工号、姓名及年龄”。

2. （4分）使用关系代数表达式查询“既在部门1工作又在部门2工作的员工的工号”。

3. （4分）使用 SQL 查询“薪资最高的员工的工号、姓名及工资”。

4. （4分）使用 SQL 查询“全职员工数量在10人以上（含10人）的部门编号”。

5. （4分）补全下面的嵌套 SQL 查询语句中的子查询，使之完成查询“哪个经理管理着经费最多的部门？列出其工号及姓名”。

```
SELECT eid, ename FROM Emp WHERE NOT EXISTS (
```

```
);
```

.....
授课教师
.....
姓名
.....
学号
.....
院系

四、计算与证明题 (20分)
已知关系模式 $R(A, B, C, D)$ 上的函数依赖集合 $F = \{AD \rightarrow BC, BC \rightarrow A, B \rightarrow C\}$, 回答下列问题 (要求给出具体的证明或计算过程):
1. (4分) 使用 Armstrong 公理证明 F 逻辑蕴含 $B \rightarrow A$.
.....
2. (4分) 计算属性集合 BD 关于 F 的闭包, 并根据该闭包证明 F 逻辑蕴含 $BD \rightarrow AC$.
.....
3. (4分) 计算 F 的最小覆盖。
.....
4. (4分) 判断 R 最高属于第几范式, 并给出证明。
.....
5. (4分) 将 R 分解为 2 个关系模式 $R_1(A, B, D)$ 和 $R_2(B, C)$, 判断该分解是否为无损连接分解, 并给出证明。
.....

五、设计题（15分，从题目 A 和题目 B 中任选一道作答）

题目 A. 设计某医院住院处的数据库，需求如下：

- 记录每个科室的名称、地址、主任、全体医生、全体护士、所有病房。每个科室只有 1 名医生担任主任。假设每个病房只住 1 名患者。
- 记录每名医生的职工号、姓名、职称、科室、主治的患者。每名医生只在 1 个科室工作。
- 记录每名护士的职工号、姓名、科室、主管的病房。每名护士只在 1 个科室工作，每名护士每天可以管理多个病房。病房的主管护士会轮换，但每个病房每天只有 1 名主管护士。
- 记录每个病房的病房号、住在该病房的患者、每天的主管护士等信息。病房号的取值为正整数，不同科室可能都有 1 号病房。假设每个病房只住 1 名患者。
- 记录每名住院患者的身份证号、姓名、性别、年龄、科室、主治医生、住院时间、出院时间、病房号等信息。患者可能会多次住院，每次住院分配唯一的住院号，每次住院只有 1 位主治医生。

根据上述需求完成下列数据库设计任务：

1. （10分）绘制实体-联系图（E-R 图）表示该数据库的概念模型。要求标出每个联系型的基数比（1:1、1:N 或 M:N）。实体型、联系型及属性使用中文命名。
2. （5分）将你绘制的 E-R 图转换为关系数据库模式。要求使用下划线标出每个关系的主键。如果关系的某个外键与其参照的主键不同名，需要明确写出该外键参照的主键。

题目 B. 任选一个你熟悉的系统，正确分析该系统的需求，绘制相应的 IDEF1X 图。要求：(1) 简要叙述系统的需求。注意：IDEF1X 图的完备性和正确性评价以次需求为基本判断依据。(2) 实体型数量不得少于 10 个。(3) 不得出现分类联系与分类实体。(4) 将此 IDEF1X 图转换为关系模式。

授课教师

姓名

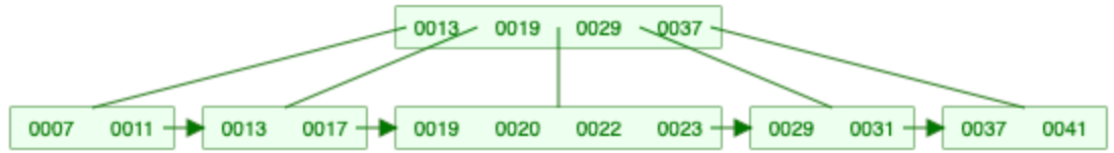
学号

院系

密
封
线

六、算法题（6分）

已知如下图所示的 B+树，其索引键的类型为 CHAR(4)。除根节点外，每个节点中最少可以存放 2 个索引键值，最多可以存放 4 个索引键值。绘制出向该 B+树中插入键值为“0021”的索引项后得到的 B+树。



七、分析题（6分）

已知一个关系数据库的模式如下：

- 关系 B(bno, bname, author)为图书表，其中 bno 为书号，bname 为书名，author 为作者；
- 关系 S(sno, sname, dept)为学生表，其中 sno 为学号，sname 为姓名，dept 为学生所在系；
- 关系 L(sno, bno, date)为借书表，其中 sno 为学号，bno 为书号，date 为借书时间。

回答下列问题：

1. （3分）绘制下面的 SQL 查询语句的逻辑查询计划树。

```
SELECT bname FROM B NATURAL JOIN S NATURAL JOIN L
WHERE date = '2021-06-04' AND college = 'CS';
```

2. （3分）使用启发式查询优化方法对上面的逻辑查询计划树进行优化，绘制优化后得到的逻辑查询计划树，具体说明你进行这些优化的理由。

八、分析题（6分）

已知2个关系 $R(\underline{A}, B)$ 和 $S(\underline{B}, C)$ ，其主键分别为 $R.A$ 和 $S.B$ 。设2个关系均采用聚簇存储，在外存中存储 R 和 S 分别占用 1000 和 2000 个块（页），且每个关系中的元组均已按照其主键值递增排序。现在要执行自然连接操作 $R \bowtie S$ 。设缓冲区中可用内存页数为 $M = 40$ 。回答下列问题：

- （3分）执行 $R \bowtie S$ 需要进行多少次 I/O？给出具体分析过程。

- （3分）设 $R.B$ 是关系 R 的外键，参照 $S.B$ 。如果 $R \bowtie S$ 的结果中元组的平均大小是 R 中元组平均大小的 1.2 倍， $R \bowtie S$ 的结果中元组的平均大小是 S 中元组平均大小的 2 倍，那么在外存中存储 $R \bowtie S$ 的结果需要占用多少个块（页）？给出具体分析过程。

九、分析题（6分）

已知某数据库采用即时更新方法（undo-redo 方法）记录 WAL 日志。设故障发生时 WAL 日志文件内容如下：

<T1, begin>
<T1, A, 23, 233>
<T2, begin>
<T1, B, “o”, “or”>
<T1, commit>
<T3, begin>
<T3, B, “or”, “orz”>
<T2, A, 233, 2333>

当系统重启后，DBMS 基于该 WAL 日志文件进行故障恢复。回答下列问题：

- （3分）当 DBMS 进行故障恢复时，需要对哪个事务进行 undo？对哪个事务进行 redo？给出具体理由。

- （3分）当故障恢复完成时，对象 A 和 B 的值分别是什么？描述故障恢复的具体过程。

授课教师

姓名

学号

院系

密
封
线

十、分析题（6分）

已知下面 2 个事务 T1 和 T2:

T1	T2
read(A)	read(B)
read(B)	read(A)
if A = 0 then B := B + 1	if B = 0 then A := A + 1
write(B)	write(A)

设对象 A 和 B 上的一致性约束条件为 A = 0 或 B = 0，A 和 B 的初始值均为 0。回答下列问题:

1. （3分）判断下面给出的调度是否为可串行化调度，并给出证明。

T1	T2
read(A)	
	read(B)
read(B)	
if A = 0 then B := B + 1	
write(B)	
	read(A)
	if B = 0 then A := A + 1
	write(A)

2. （3分）是否存在一种对 T1 和 T2 的非串行的可串行化调度？给出具体理由。