

2024 年秋 数据库系统 期末考试回忆版试题

编者：夏提雅

参与回忆者：一名不愿提供姓名的同学、夏提雅

Version: 1.0 （2024 年 12 月 7 日）

免责声明：本试题是在离开考场后，回忆出来的，不存在任何作弊行为；本试题题目部分不保证题干、选项与原题一致，但考察的中心思想一致。

（考试时间：2024 年 12 月 7 日；满分：100 分；时间：120 分钟）

一、填空题（共 15 小题，每小题 1 分，共 15 分）

1. 在数据库系统的两层映像中，_____映射支持实现数据概念视图向外部视图的转换，实现了_____数据的独立性。
2. 考虑在公共属性 A 上连接关系 R(A, B)和 S(A, C, D)。若可用内存总数为 M=3，关系 R 所占磁盘块数为 $B_R=4$ ，关系 S 所占磁盘块数为 $B_S=8$ ，则该连接所需最小 I/O 次数为_____（忽略最终结果的写代价）。
3. 已知关系表达式 E_1 和 E_2 满足并相容性，判断从 $\prod_{A_1, A_2, \dots, A_n}(E_1 - E_2)$ 到 $\prod_{A_1, A_2, \dots, A_n}(E_1) - \prod_{A_1, A_2, \dots, A_n}(E_2)$ 的变换是否等价_____（等价/不等价）。
4. 在强制安全机制里，用户与数据对象都进行了安全分级分别为 Level(S)和 Level(O)，那么当满足_____，用户可以写数据对象。
5. 当一个函数依赖的右端是其左端的子集时，称该函数依赖为_____。
6. 可扩展散列索引的指针数组每增长一次，桶的数目就_____；线性散列索引的指针数组每增长一次，桶的数目就_____。
7. 若 $A \rightarrow B$, $BC \rightarrow D$, $E \rightarrow A$ 为 $R = (A, B, C, D, E)$ 上的函数依赖，则 (CE) 的属性闭包为_____。
8. 关系 $R(stu_id, stu_name, age, department)$ ，那么关系代数表达式 $\sigma(age > 18 \wedge department = '机电' \vee age < 18 \wedge department = '计算机')$ 表达的含义为_____。
9. 稀疏索引文件的存在与否_____（改变/不改变）存储表的物理存储结构；目的为_____。
10. E-R 图向关系模型转换时，E-R 图中无属性的联系（1:m）应如何进行转换？_____。
11. 产生冲突可串行调度的方法有_____和_____。
12. 给定一个调度序列 $r_2(A); r_1(B); w_2(A); r_2(B); r_3(A); w_1(B); w_3(A); w_2(B)$ 。请判

断其是否为冲突可串行调度_____（是/否）。

13. 一个关系在满足 BCNF 范式的同时又满足_____范式。

14. 锁能否保证冲突可串行性？_____（是/否）。

15. 关系型数据库中的第三范式要求满足 2NF，同时消除_____依赖。

二、选择题（共 15 小题，每小题 1 分，共 15 分）

1. 在关系数据库中，关于主码说法**错误**的是（ ）。

- A. 可以唯一标识关系的每一个元组
- B. 主码的值具有唯一性
- C. 主码允许有空值
- D. 主码可以由该关系的若干个属性组成

2. 关于关系的说法**错误**的是（ ）。

- A. 关系的每一列中的分量来自同一域、是同一类型的数据
- B. 关系中不同属性的值域可以是相同的
- C. 关系具有行列无关性
- D. 关系中的属性可以再分

3. 如果两个表有多个相同的属性，如果只想把其中一个列值相等的连接起来，应该做（ ）连接。

- A. 自然
- B. 等值
- C. 非等值
- D. 以上选项均错误

4. 关于稀疏索引和稠密索引，下列说法不正确的是（ ）。

- A. 稠密索引的索引文件中，索引字段值是可以重复的
- B. 主索引可以是稀疏索引，辅助索引一定是稠密索引
- C. 如果一个搜索列的值在稠密索引中不存在，则在主文件中对应该搜索的值的记录也不存在
- D. 如果一个搜索列的值在稠密索引中存在一个，则在主文件中对应该搜索的值的记录也只存在一个

5. 关于 B+树，下列说法正确的是（ ）。
- A. 如果会发生合并，不一定会减少索引存储块的数目
 - B. 如果会发生分裂，不一定会增加索引存储块的数目
 - C. 只有 B+树中所有节点的索引项，不能覆盖主文件的完整索引
 - D. 以上选项均不对
6. 已知一个存储块可存放主文件的 5 条记录、可存放索引文件的 20 个索引项目。已知主文件有 n 条记录，在无重复值、无序前提下创建稠密索引和无重复值、有序前提下创建稀疏索引，各需要存储块数目为（ ）。
- A. $n, n/20$
 - B. $n/100, n/100$
 - C. $n/20, n/100$
 - D. $n/20, n/5$
7. 名为“选课”的关系，其关系模式如下：选课（学号，课号，学期，授课老师，成绩）。若每个学生同一学期可以选多门不同的课，每个学期每门课可以有多个学生选；每个学生不同学期可以重复选同一门课且每次都会获得一个成绩；同一门课各个学期可能由不同教师讲授，则该关系（ ）。
- A. 属于 1NF 但不属于 2NF
 - B. 属于 2NF 但不属于 3NF
 - C. 属于 3NF 但不属于 BCNF
 - D. 属于 BCNF
8. 下列表述不正确的是（ ）。
- A. 若关系模式 $R \in 1NF$ ，且 R 的每个候选键都只包含单个属性，则必有 $R \in 2NF$
 - B. 若关系模式 $R \in 1NF$ ，且 R 的唯一候选键是所有属性的集合 U ，则必有 $R \in BCNF$
 - C. 若关系模式 $R \in 1NF$ ，且 R 上任何非平凡函数依赖的左端都是 R 的一个主属性，则必有 $R \in BCNF$
 - D. 若关系模式 $R \in 1NF$ ，且 R 上任何非平凡函数依赖的右端都不包含 R 的主属性，则必有 $R \in BCNF$

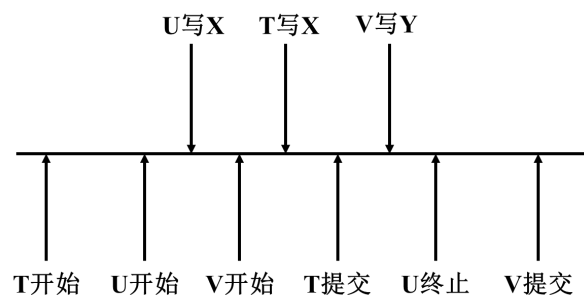
9. 若 E-R 图中包含两个实体，且两个实体之间存在多对多（ $m:n$ ）的联系，则该 E-R 图转换成关系模型后，包含（ ）个关系模式。
- A. 2
B. 3
C. 4
D. 1
10. 已知内存共有 100 块，若要排序有 10,000 块的数据集，其中趟数指的是对数据集的扫描次数，归并阶段内存需留出一块作为输出块，则下列说法正确的是（ ）。
- A. 该数据集可以在两趟内实现排序，最小磁盘读写次数为 40000 次
B. 该数据集不能在两趟内实现排序，最小磁盘读写次数为 40400 次
C. 该数据集可以在两趟内实现排序，最小磁盘读写次数为 40400 次
D. 该数据集不能在两趟内实现排序，最小磁盘读写次数为 60000 次
11. 基于时间戳的并发控制，当发生冲突时需要撤销并重启事务以解决冲突。已知 T_1 , T_2 , T_3 三个事务，记 $w_i(A)$ 为事务 T_i 写数据对象 A 、 $r_i(A)$ 为事务 T_i 读数据对象 A 。 T_1 , T_2 , T_3 三个事务的时间戳为 180, 120 和 140，三个事务的操作依下列次序进行，被撤销事务自动加到队列尾部，请判断被撤销的事务为（ ）。
- $r_1(B); r_2(A); r_3(C); w_1(B); w_1(A); w_2(C); w_3(A);$
- A. T_1 被撤销/重启
B. T_2 被撤销/重启
C. T_3 被撤销/重启
D. 三个事务均被撤销/重启
12. 关于线性散列和可扩展散列索引，下列说法正确的是（ ）。
- A. 线性散列桶满时桶增加
B. 可扩展散列桶满时桶不一定增加
C. 线性散列和可扩展散列都是桶满时 i 增加
D. 以上选项均不对
13. 用 B+ 树建立主索引，下列说法正确的是（ ）。

- A. 叶结点某索引项 X 的左侧指针, 指向键值 $= X$ 记录所在的索引文件存储块
 - B. 非叶结点某索引项 X 的左侧指针, 指向键值 $< X$ 记录所在的索引文件存储块
 - C. 叶结点某索引项 X 的左侧指针, 指向键值 $< X$ 记录所在的主文件存储块
 - D. 非叶结点某索引项 X 的左侧指针, 指向键值 $< X$ 记录所在的主文件存储块
14. 给定关系模式 $R = (U, F)$, $U = (A, B, C, D)$, F 为 R 上的函数依赖集合, 若 $(AB)^+_F = \{A, B, C, D\}$, 则下列说法一定成立的是 ()。
- A. $A^+_F = \{A, B, C, D\}$
 - B. (A, B) 为 R 的候选键
 - C. $(AB \rightarrow D) \in F$
 - D. $(ABC)^+_F = \{A, B, C, D\}$
15. 给定关系模式 $R = (A, B, C, D)$, 以及两个函数依赖集合: $F_1 = \{AB \rightarrow C, C \rightarrow D, A \rightarrow D\}$, $F_2 = \{A \rightarrow C, BC \rightarrow D, C \rightarrow D\}$, 下列说法正确的是 ()。
- A. F_1 覆盖 F_2 , 但 F_2 不覆盖 F_1
 - B. F_2 覆盖 F_1 , 但 F_1 不覆盖 F_2
 - C. F_1 与 F_2 等价
 - D. 以上选项均不对

三、简答题（共 4 小题，每小题 5 分，共 20 分）

1. 给定 $Course(\underline{cno}, cname, tno)$, $Teacher(\underline{tno}, tname, dno)$, $Dept(\underline{dno}, dname)$, 写 SQL 查询找出计算机系至少讲过一门课程的教师总数。

2. 针对并发控制常见的三种不一致问题，如果应用程序只需要解决丢失修改和脏读问题，请简述你将使用哪种封锁协议，并画图示意。
3. 现有一个空的可扩展散列索引，每个桶最多可以装两个数据，往其中插入 11110, 11101, 11111, 01111, 01101, 01011 这六个索引数据，试述插入完成后需要使用桶数组的个数及最终结果图（可扩展散列索引的序号是从高位开始取的）。
4. 有下列三个事务 T, U, V，从左往右代表事务的执行顺序，为放行不影响并发结果的写操作，请问在该图中是否可以使用托马斯规则，请说明原因。



四、（10 分）

在教师管理数据库中，假设有如下实体：

教师：教师号、姓名、性别、职称

课程：课程号、课程名、任课教师

学院：学院代码、学院名称、联系电话

且以上实体存在如下联系：

（1）一个教师可以讲授多门课程，一门课程可由多位教师讲授。

（2）一个学院可有多位教师，一位教师只属于一个学院。

请用 Chen 方法设计教师管理的 E-R 图（5 分），并标明参与各方的基数（不要求标最小基数）（5 分）。

五、（10 分）

对学生课程数据库有如下查询：

```
select s#, sname from Student, SC where c# = '001' and Student.s# = SC.s#
```

1. 写出该查询的关系代数表达式；（2 分）
2. 根据关系代数表达式画出语法树；（3 分）
3. 用关系代数表达式优化算法对原始的语法树进行优化处理，并画出优化后的语法树。（5 分）

六、（15 分）

设有关系模式 $R = (U, F)$ ， $U = (A, B, C, D, E, G, H)$ ，函数依赖集合 $F = \{BCD \rightarrow A, BC \rightarrow E, A \rightarrow G, G \rightarrow H, C \rightarrow D, A \rightarrow H\}$ ，请回答下列问题。

1. 请找到 R 的所有候选码。（5 分）
2. 给出与 F 等价的最小覆盖集，给出简要求解步骤。（5 分）
3. 对 R 进行无损连接且保持依赖分解，保证结果均符合 3NF，同时指出分解后的每个关系模式是否也属于 BCNF。（5 分）

七、（15 分）

假设有三个关系：读者 $Readers(\underline{rid}, name, address)$ ，图书 $Book(\underline{bid}, title, author, publication)$ 和借阅关系 $Borrow(\underline{rid}, \underline{bid}, \underline{date})$ ，如果想查询“借阅过 rid 为 20240207 号读者所有书的读者 rid ”，请回答下列问题。

1. 请用基本关系代数书写该查询。（5 分）
2. 请用 SQL 语句书写该查询。（5 分）
3. 请描述上述 SQL 语句的执行过程。（5 分）