

主管  
领导  
审核  
签字

数据库系统（A）试题

考试时间：120 分钟                      试卷满分：100 分

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

注：本试卷为回忆版，在浏览本试卷前，请先阅读最后一页的备注。

姓名

学号

班号

院号

密

封

线

一、填空题（每小题 1 分，共 15 小题，满分 15 分）

1. 概念模式又称\_\_\_\_\_，将概念模式映射为内模式体现出数据库系统的\_\_\_\_\_性。
2. 层次模型的数据结构是\_\_\_\_\_结构，网状模型的数据结构是\_\_\_\_\_结构。
3. 关系 R 有三个属性 A、B、C，其中属性 A 和 C 的权限为 C，属性 B 的权限为 U，元组的权限 TC 为 C，则对于权限为 S 的用户来说\_\_\_\_\_读\_\_\_\_\_写（以上两空均填“可以”或“不可以”）。
4. 一个关系模式满足第三范式,那么它一定满足\_\_\_\_\_范式。
5. select 是\_\_\_\_\_（填“DML”或“DDL”）。
6. 关系是从\_\_\_\_\_抽象出来的。参与发生联系的实体的数目，称为联系的\_\_\_\_\_或\_\_\_\_\_。
7. RAID 的两种数据拆分方式是\_\_\_\_\_和\_\_\_\_\_。
8. 辅助索引\_\_\_\_\_（填“可以”或“不可以”）建立在主键上。
9. B+和 B 树本质的区别是\_\_\_\_\_。
10. 基于时间戳的并发控制\_\_\_\_\_（填“有”或“无”）冲突可串行性。
11. 故障恢复可以保证事务的\_\_\_\_\_性。
12. 三种典型的不一致现象中，\_\_\_\_\_最不重要。
13. Undo/Redo 结合型日志的数据结构的样子是\_\_\_\_\_。
14. 现有 table Student = {#S, SName, Adress} 以及视图 create mitview mit as select SName, Adress from Student，若使用 update mitview values ("张伟","红花岭 1 栋")进行更新，这是\_\_\_\_\_（填“合法”或“不合法”）的。

---

15. SQL 语句 `select stu_id from Student where age>18 ∨ age<22 ∧ department='计算机'` 的含义是\_\_\_\_\_。

二、选择题（每小题 1 分，共 15 小题，满分 15 分，每小题中给出的四个选项中只有一个是符合题目要求的，把所选项的字母填在题后的括号内）

16. 下列选项中，具有唯一性是

- A. 超键            B. 候选键            C. 主键            D. 外键

17. 下列关于外键的说法中，正确的是

- A. 外键删除时必须级联删除    B. 外键值不允许为空  
C. 外键可以包含多个属性        D. 缺失

18. 下列关于关于线性散列索引和聚簇索引的说法中，正确的是

- A. 缺失（四个选项问的都是桶的问题，其中有桶分裂相关的问题）  
B. 缺失  
C. 缺失  
D. 缺失

19. 在关系代数运算中，五种基本运算为

- A. 并、差、选择、投影、乘积    B. 并、交、选择、投影、连接  
C. 并、差、交、选择、投影        D. 并、交、选择、投影、笛卡尔积

20. 在关系数据库中,任何二元关系模式最高范式必定是

- A. 1NF            B. 2NF            C. 3NF            D. BCNF

21. 下列关于物理存储的说法中，正确的是

- A. 对磁盘的读写是以字节为单位  
B. 内存有电易失性  
C. 缺失，磁盘块相关的问题  
D. 缺失，选项中包含了“奇偶校验”

22. 下列关于稠密索引和非稠密索引的说法中，错误的是

- A. 稠密索引主文件必须按主键排序  
B. 缺失  
C. 缺失  
D. 缺失

23. Create Table 有三种功能，以下不是其中一项功能的是

- A. 定义安全性约束                B. 定义关系模式

### C. 定义完整性约束

#### D. 定义物理存储特性

24. 下列聚合函数中不忽略空值（null）的是

### A. sum(列名)

B. count(\*)

C. max(列名)

#### D. avg(列名)

25. 关系 R、S 分别有 M、N 个元组，则  $R \bowtie R, R \bowtie S$  的元组个数可能为

A.  $M \times N$ ,  $\max(M, N)$  B.  $M$ ,  $M \times N$

### C. $M^*N$ , $M^*N$

#### D. $M \times N, \max(M, N)$

26. 已知关系  $R(A, B, C, D, E, F, G)$ , 函数依赖集  $F$  为  $\{A \rightarrow B, B \rightarrow D, AD \rightarrow EF, AG \rightarrow C\}$ , 则  $A$  关于  $F$  的属性闭包为

A. {A, B, C, D}

B. {A, B, D, E}

C. {A, B, D, E, F}

D.  $\{A, B, C, D, E, F, G\}$

27. 已知关系  $R(A, B, C, D, E, F, G)$ , 函数依赖集  $F$  为  $\{AB \rightarrow CF, AD \rightarrow CE, AG \rightarrow B, D \rightarrow C, B \rightarrow D\}$ , 则  $F$  的最小覆盖为

A.  $\{AB \rightarrow C, AB \rightarrow F, AD \rightarrow C, AD \rightarrow E, AG \rightarrow B, B \rightarrow D, D \rightarrow C\}$

B.  $\{AB \rightarrow F, AD \rightarrow E, AG \rightarrow B, D \rightarrow C, B \rightarrow D\}$

C.  $\{AB \rightarrow F, AD \rightarrow C, AG \rightarrow B, D \rightarrow C, B \rightarrow D\}$

D.  $\{A \rightarrow F, B \rightarrow F, AD \rightarrow E, A \rightarrow B, G \rightarrow B, D \rightarrow C, B \rightarrow D\}$

28. 在关系模式  $R(U, F)$  中, 如果  $X \rightarrow Y$ , 存在  $X$  的真子集  $X_1$ , 使  $X_1 \rightarrow Y$ , 则称函数依赖  $X \rightarrow Y$  为

### A. 平凡函数依赖

### B. 部分函数依赖

### C.完全函数依赖

### D. 传递函数依赖

29. 给定关系模式  $R(U, F)$ , 其中  $U=\{A_1, A_2, A_3, A_4, A_5, A_6\}$ , 给定函数依赖集合  $F=\{A_1 \rightarrow (A_2, A_3); A_3 \rightarrow A_4; (A_2, A_3) \rightarrow (A_5, A_6) ; A_5 \rightarrow A_1\}$ , 则分解  $\rho=\{R_1(A_1, A_2, A_3, A_4), R_2(A_2, A_3, A_5, A_6)\}$

A. 既具有无损连接性，又保持函数依赖

B. 不具有无损连接性，但保持函数依赖

C. 具有无损连接性，但不保持函数依赖

D. 既不具有无损连接性，又不保持函数依赖

30. DBMS 管理数据库缓冲区有四种策略: No Steal, Steal, No Force, Force。则效率较低但不会出现问题的策略组合是\_\_\_\_\_，而效率最高最常用但会出现问题的策略组合是\_\_\_\_\_。

### A. No Steal+No Force, Steal + Force

### B. No Steal+ Force, Steal + No Force

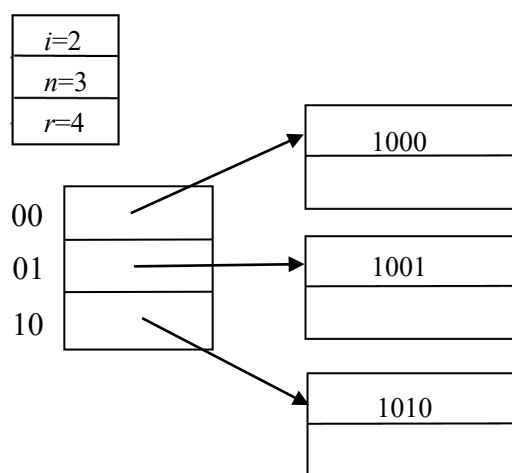
### C. Steal+No Force, No Steal + Force

#### D. Steal+ Force, No Steal + No Force

---

### 三、简答题（每小题 5 分，共 5 小题，满分 25 分）

31. 转储点和检查点是不是同一个概念？如果不是的话，他们的定义是什么？转储点是哪种故障提出的？
32. 假设内存有 4 块，而待排序的数据块有 50 块，请给出归并次数最少的算法，并且计算磁盘读写次数，给出过程。
33. 假设线性哈希表之中，一个桶不能超过其 80% 的容量。现插入两个值 1111, 1110，请你写出插入的过程。



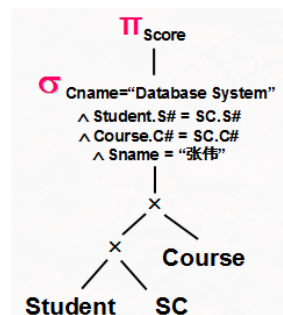
34. 针对堆文件和聚簇文件这两种文件组织方式，请分别给出他们的优缺点（3 分），并简单举例说明他们适合的应用场景。（2 分）
35. 基于时间戳的并发控制，需要冲突检测，当发生冲突的时候需要撤销事务、重启事务以解决冲突。现有三个事务， $T_1, T_2, T_3$ ，他们的时间戳分别为 180, 140, 160。三个事务的操作按照下列次序进行：

$r_1(B); r_2(A); r_3(C); w_1(A); w_2(C); w_3(A)$

在忽略重启的情况下，请分析因为冲突而最可能被撤销的事务。

#### 四、(共 1 小题, 满分 10 分)

36. 已知关系  $Student(S\#, Sname, Sage, Sclass)$ ,  $Course(C\#, Cname, Credit, Cteacher)$ ,  $SC(S\#, C\#, Score)$ 。给定 SQL 语句如下: “SELECT Score FROM Student, Course, SC WHERE Cname= ‘Database System’ and Sname= ‘张伟’ and Student.S#= SC.S# and Course.C#=SC.C#”。SQL 语句对应的语法树如下图所示, 请你给出语法树优化的中间过程并用语法树给出最终的优化结果。



#### 五、(共 1 小题, 满分 10 分)

37. 针对高并发的需求, 对于火车站订票系统来说, 有三种基本操作, 查询、下订单、2 分钟内付款。要求效率 (题干回忆不全, 这句话只复述了大概意思)。
- (1) 针对具体用户, 请你设计事务 (有几种事务), 并说明理由。(5 分)
- (2) 针对这些事务, 请确定他们的缓冲区策略, 并说明理由。(5 分)

#### 六、(共 1 小题, 满分 10 分)

38. 给定关系模式  $R(U, F)$ , 其中  $U=\{A, B, C, D, E\}$ , 给定函数依赖集合  $F=\{A \rightarrow C; C \rightarrow D; B \rightarrow C; DE \rightarrow C; CE \rightarrow A\}$
- (1) 写出  $R$  的候选码 (2 分)
- (2) 将  $R$  无损连接分解使其满足 BCNF (3 分)
- (3) 判断  $R_1=\{A, D\}$ ,  $R_2=\{A, B\}$ ,  $R_3=\{B, C\}$ ,  $R_4=\{C, D, E\}$ ,  $R_5=\{A, E\}$  是否为无损连接分解 (5 分)

#### 七、(共 1 小题, 满分 15 分)

39. 现有需求 “选出那些选择了全部计算机部门的老师的开设的课程的学生”, 数据库内的关系如下:  $Student = \{ \#S, SName \}$ ,  $SC = \{ \#S, \#C \}$ ,  $Teacher = \{ \#T, \#D \}$ ,  $Teach = \{ \#T, \#C \}$ ,  $Department = \{ \#D, DName \}$
- (1) 请给出这个需求的关系代数表达式 (5 分)
- (2) 请将这个表达式转化为 SQL 语句 (5 分)
- (3) 请叙述这个 SQL 语句的执行过程 (5 分)

---

备注：

- 关于知识点：第 19、26、27、28、29、30、35、36 题为 MOOC 上的题（19 题选项有变化但正确选项没变，35 为类似题，其余均为原题），部分题目在作业题中有对应知识点，部分题目在 2018 级期末试题中有对应知识点。2020 级的学长提到 2020 级的数据库期末考试考了 RAID，不过不清楚具体是怎么考的。
- 关于题目：第 14 题没回忆出原卷的代码，只记得原卷地址是红花岭 1 栋以及原卷的考点，本 pdf 中用了群友自编的代码，但不影响知识点的考查。第 19、35 题由于原卷中的题目和 MOOC 相比有改动，本 pdf 凭印象改了一下，不保证和原卷一模一样。
- 关于回忆版试卷的排版与细节：原卷的填空题空的位置有点潦草，本回忆版试卷进行了优化。原卷每道大题中的小题的题号都是重新计数的，本回忆版试卷由于个人喜好没有重新计数。原卷四、五、六、七大题中每小问都直接用的“1.”“2.”，没有给整道大题一个小题号然后每小问用“(1)”“(2)”。
- 本回忆版试卷由几位自救群的同学回忆而成。