

2020 年硕士研究生招生考试试题

考试科目:807 软件工程专业综合

请考生注意:

- (1) 所有答案(包括选择题和填空题)一律写在答题纸上, 否则不计成绩。
- (2) 允许使用计算器。
- (3) 本考题包括数据结构, 操作系统, 数据库三个科目。其中数据结构为必选。操作系统与数据库为二选一, 考生需选择其中一个科目答题并且注明所选科目的名称。

必选科目数据结构(90 分)

一、选择题(每小题 2 分, 共 20 分)

1. 考虑下面的程序段

```
void running(int n)
{
    int j = 0;      int k = 0;
    while(j < n) {  k = k+1;  j = j + k;  }
}
```

此段代码的时间复杂度为

- A. $O(\log n)$ B. $O(n^{1/2})$ C. $O(n)$ D. $O(n^2)$
2. 设 T 是高度为 h 的平衡二叉树(又称 AVL 树), 并且是高度为 h 的包含节点最少的平衡二叉树, 则 T 包含节点数目的数量级是?
- A. 1.41421^h B. 1.61803^h C. 2.71828^h D. 3.14159^h
3. 循环单链表的主要优点是
- A. 不再需要头节点指针
- B. 从表的任一节点出发都能够遍历整个链表
- C. 已知某个节点位置后能够容易找他其前趋
- D. 在进行插入删除操作时能够保证链表不断开
4. 将 n 阶对称矩阵 $A=[a_{j,k}]$ ($0 \leq j, k < n$) 的上三角元素按行优先压缩存储在数组 $b[0, N)$ 中, 则矩阵元素 $a_{j,k}$ ($j \leq k$) 在数组中对应的位置是
- A. $b[j*n-j*(j-1)/2 + k]$ B. $b[j*(j-1)/2 + k]$
- C. $b[j*n-j*(j+1)/2 + k-1]$ D. $b[j*(j+1)/2 + k-1]$
5. 对快速排序算法较为不利的情况是
- A. 数据量太大 B. 数据基本有序
- C. 数据中包含太多的相同键值 D. 数据量为奇数



6. n 个节点的二叉树大约有多少个不同的形状?

- A. 3.14159^n B. 2^n C. 2.71828^n D. 4^n

7. 在采用线性探测处理冲突的散列表上作查找操作。一次查找可能需要探测不止一个位置。则一次查找过程中所探测的这些位置上的键值

- A. 都相同
B. 散列值都相同
C. 散列值不一定相同
D. 散列值一定不相同

8. 二叉树中序遍历算法如下

```
void inorder(Node* root)
```

```
{  
    if(root!=0)  
    {    inorder(root->左子);    visit(root);    inorder(root->右子);    }  
}
```

关于它的时间复杂度，下面哪一个命题最准确?

- A. 是二叉树中节点数目的倍数
B. 是二叉树中节点路径长度之和的倍数
C. 与 A、B 中的两个因素都相关
D. 与树的形状相关

9. 下面哪一种算法可以求无向图的所有连通分量?

- A. 拓扑排序
B. 求关键路径算法
C. 求最短路径算法
D. 广度优先遍历

10. 假设模式串的长度为 m ，目标串的长度为 n ，则简单模式匹配算法的时间复杂度是

- A. $O(n+m)$ B. $O(m)$ C. $O(n)$ D. $O(m*n)$

二、简答题（每小题 10 分，共 40 分）

1. 从空的平衡二叉树开始，依次添加 9, 8, 7, 6, 5, 4, 3, 2, 1 九个关键字。画出添加每一个关键字后的平衡二叉树。

2. 已知一颗二叉树节点的中序序列是 B, F, D, A, E, G, C；后序序列是 F, D, B, G, E, C, A。画出此二叉树，并且画出此二叉树对应的森林。

3. 一篇文章中出现九个字符 a, b, c, d, e, f, g, h, i。它们在文章中出现的频率分别是 1, 2, 3, 4, 5, 6, 7, 8, 9。给出这些字符的 Huffman 编码并且求此篇文章用 Huffman 编码后的长度。



4. 已知无向图的弧边及其权值为

$E = \{ (A, B, 1), (A, C, 8), (A, D, 4), (B, D, 9), (C, D, 5), (B, E, 3), (D, E, 7), (E, F, 1), (D, F, 3), (C, G, 2), (F, H, 5), (D, G, 6), (G, H, 7), (E, H, 2), (F, G, 6) \};$

1) 在答题纸上画出此图;

2) 在答题纸上画出此图从节点 A 到达 E, F, G, H 四个节点的最短路径。

三、算法应用题（每小题 15 分，共 30 分）

1. 给定二叉树节点的数据结构是

```
struct Node {
    T value;
    Node* lc;    //左子
    Node* rc;    //右子
};
```

编写函数 `bool is_search_tree(Node* root)`, 它判别以 `root` 为根的二叉树是否为单调增的排序二叉树。

2. 给定单链表节点的数据结构是

```
struct Node {
    T value;
    Node* next;
};
```

编写函数 `Node* merge_sort(Node* first)`, 它对不带头节点的单链表进行归并排序。函数返回指向排序后单链表第一个节点的指针。首先用文字简要描述一下算法。

二选一科目数据库（60 分）:

四、选择题（每小题 2 分，共 20 分）

1. ____是数据标准化存储的一种方法。

- A. MIS
- B. 结构化编程
- C. CODASYL 规范
- D. 以上均不对

2. 对于关系中的每个属性，都有一组允许的值，称为该属性的____。

- A. 域
- B. 关系
- C. 取值
- D. 模式

3. 能够唯一定义表中每一行的属性（或属性集）称为____。

- A. 标识符
- B. 索引
- C. 主键
- D. 对称密钥



4. SQL 语句 `select substr ('123456789', instr ('abcabcb', 'b'), 4) from emp` 的输出结果是_____。
- A. 6789 B. 2345
C. 1234 D. 456789
5. 包含了多个属性的索引键称为_____索引键。
- A. 简单 B. 复合
C. 多个 D. 二级
6. 属性年龄可以通过出生日期计算出来, 属性年龄是_____。
- A. 单值属性 B. 多值属性
C. 复合属性 D. 导出属性
7. _____是选择数据库的数据存储和数据访问特性的过程。
- A. 逻辑数据库设计 B. 物理数据库设计
C. 测试和性能调整 D. 评估和选择
8. _____可以帮助我们发现不良的 E-R 设计。
- A. 数据库设计过程 B. E-R 设计过程
C. 关系方案 D. 函数依赖
9. 为了减少基于日志的恢复造成的开销, 引入了_____。
- A. 检查点 B. 指数
C. 死锁 D. 锁
10. 允许并发事务访问同一表的不同行的锁称为_____。
- A. 数据库级锁 B. 表级锁
C. 页级锁 D. 行级锁

五、简答题（每小题 5 分，共 15 分）

1. 叙述自然连接、左连接、右连接、完全外连接、半连接的含义与区别。
2. 试述实现数据库安全性控制的常用方法。



3. 数据库运行中可能产生的故障有哪几类? 哪些故障影响事务的正常执行? 哪些故障破坏数据库数据?

六、计算题 (每小题 3 分, 共 15 分)

某研究所数据库由四张表组成, 表的信息如下:

研究室表 `departs`(`did`, `dname`, `address`), `did` 是主键, `dname` 是研究室名称, `address` 是地址;

项目表 `projects`(`pid`, `pname`, `total`, `did`), `pid` 是主键, `pname` 是项目名称, `total` 是项目经费额, `did` 是外键, 参照研究室表的 `did`;

职工表 `employeees`(`eid`, `ename`, `sex`, `age`, `job`, `did`), `eid` 为主键, `ename` 是姓名, `sex` 是性别, `age` 是年龄, `job` 是职务, `did` 是外键, 参照研究室表的 `did`;

参加表 `works`(`pid`, `eid`, `hour`), `pid`、`eid` 是外键, 分别参照项目表的 `pid` 和职工表的 `eid`, `hour` 是工作时长。

1. 找出参加“天宫 1 号”项目的科研人员职工号、姓名、性别, 写出关系代数表达式。
2. 找出既参加“天宫 1 号”项目, 又参加“嫦娥 1 号”项目的科研人员职工号, 写出关系代数表达式。
3. 画出上述问题 1 的关系代数表达式的语法树。
4. 查询科研人员“李明”参加的科研项目号、项目名、经费额、工作时长, 按照工作量降序排列, 写出 SQL 语句。
5. 向科研项目表中追加一条纪录(“K16204”, “神州飞船”, 1000), 将科研项目名称为“神州飞船”的科研经费增加 2000, 写出对应的 SQL 语句。

七、设计题 (每小题 5 分, 共 10 分)

1. 图书馆向借书人提供书籍。每本书都按书名、版本和出版年份进行描述, 并使用 ISBN 进行唯一标识。每个借书人都用其姓名和地址进行描述, 并使用借书人编号进行唯一标识。图书馆提供每本书的一个或多个副本, 每个副本都使用一个副本编号、指示该书是否可供借阅的状态以及给定副本的允许借阅期限。借书人可以借出一本或多本书, 并记录借出和归还的日期, 借阅编号唯一标识每个帐簿借阅。
设计该数据库的 E-R 图 (画出 E-R 图), 将该 E-R 图转化为关系模式结构, 并指明主键和外键。



2. 下表显示了演员在电影中所扮演角色的详细信息。

电影号	电影名	导演号	导演名	演员号	演员名	角色	出境时长
F1100	马路	D101	张三	A1020	王五	甲	15. 45
				A1222	赵六	乙	25. 38
				A1020	王五	丙	22. 56
F1109	小白兔	D076	李四	A1567	钱七	丁	19. 56
				A1222	孙八	戊	10. 44

将关系规范化，使其达到 3NF，并清楚地描述 1NF、2NF 和 3NF 的各个阶段。

二选一科目操作系统（60 分）：

八、判断题（每小题 1 分，共 10 分）

请考生在答题纸上标记所给出判断题题目的题，并在题号后给出正确与否的选择，标注方式采用“对”/“错”，或者“True”/“False”，或者“T”/“F”形式。

1. 分时操作系统是指在一台主机上连接了多个联机终端，并允许多个用户通过终端以交互的方式使用主计算机，共享主机资源的系统。
2. 在每个进程中，访问临界资源的那段程序称为临界区。
3. 在操作系统中，当一个进程进入临界区使用临界资源时，另一个进程必须等待，当占用临界资源的进程退出临界区后，另一进程才允许去访问此临界资源，我们称进程之间的这种相互制约关系为进程同步。
4. 要想防止死锁的发生，只需破坏死锁产生的四个必要条件之一即可。比如，可以通过破坏互斥条件来防止死锁的发生。
5. 在操作系统中引入线程概念的主要目的是减少程序并发执行时所付出的时空开销，提高程序执行的并发度。
6. 进程调度是按照给定的原则和策略，将处于外存对换区中的又具备运行条件的就绪进程调入内存，或将处于内存就绪状态或内存阻塞状态的进程交换到外存对换区中。
7. 动态重定位是在程序执行过程中，每当访问指令或数据时，将要访问的程序或数据的逻辑地址转换成物理地址。
8. 实现虚拟存储器的目的是扩充辅存容量。



9. 文件目录的作用是实现在文件名到物理地址的转换。

10. 若磁头的当前位置为 100 磁道, 磁头正向磁道号增加方向移动。现有一个磁盘读写请求队列: 23、376、205、132、19、61、190、398、29、4、18、40。若采用最短寻道时间优先算法, 移动磁道数总数为 700。

九、选择题 (每小题 2 分, 共 20 分)

- 所谓_____是指将一个以上的作业放入内存, 并且同时处于运行状态, 这些作业共享处理机的时间和外围设备等其他资源。
 - 多重处理
 - 多道程序设计
 - 实时处理
 - 并行执行
- 分配到必要的资源并获得处理机时的进程状态是_____。
 - 就绪状态
 - 阻塞状态
 - 运行状态
 - 撤消状态
- 为多道程序提供的可共享资源不足时, 可能出现死锁。但是, 不适当的_____也可能产生死锁。
 - 进程优先权
 - 资源的线性分配
 - 分配队列优先权
 - 进程推进顺序
- 进程调度算法采用等时间片轮转法时, 时间片过大, 就会使轮转法转化为_____调度算法。
 - 先来先服务
 - 多级反馈队列
 - 抢占式最短作业优先
 - 非抢占式最短作业优先
- _____是指从作业提交给系统到作业完成的时间间隔。
 - 周转时间
 - 响应时间
 - 等待时间
 - 运行时间
- 采用_____不会产生内部碎片。



- A. 分页式存储管理
B. 固定分区式存储管理
C. 分段式存储管理
D. 段页式存储管理
7. 已知页面走向为 1、2、1、3、1、2、4、2、1、3、4，且开始执行时内存中没有页面。若只给该作业分配 2 个物理块，当采用 FIFO 页面淘汰算法时，缺页率为_____。
- A. 81.8%
B. 71.8%
C. 50%
D. 60.5%
8. 操作系统中的 SPooling 技术，实质是将_____转化为共享设备的技术。
- A. 虚拟设备
B. 独占设备
C. 脱机设备
D. 块设备
9. 为了解决不同用户文件的“命名冲突”问题，通常在文件系统中采用_____。
- A. 约定的方法
B. 路径
C. 多级目录
D. 索引
10. 某软盘有 40 个磁道，磁头从一个磁道移至另一个磁道需要 6ms。文件在磁盘上非连续存放，逻辑上相邻数据块的平均距离为 13 磁道，每块的旋转延迟时间及传输时间分别为 100ms、25ms，读取一个 100 块的文件需要_____。
- A. 203ms
B. 20300ms
C. 125ms
D. 2030ms

十、简答题（每小题 5 分，共 10 分）

1. 为什么对 CPU 调度来说，区分 I/O 限制的程序（I/O-bound programs）和 CPU 限制的程序（CPU-bound programs）是重要的？
2. 什么是分页？什么是分段？二者主要有何区别？

十一、计算或设计题（共 20 分）

1. （本题 8 分）考虑下面的页面访问串：
- 1、2、3、4、2、1、5、6、2、1、2、3、7、6、3、2、1、2、3、6、



- (1) 分别针对物理页块分配个数为 5 个和 6 个这两种情形, 采用 LRU (最近最久未使用算法) 页面替换算法, 计算会出现多少次缺页中断? (提示: 所给定的物理页块初始均为空, 因此, 首次访问一页时就会发生缺页中断。)
- (2) 分别针对物理页块分配个数为 5 个和 6 个这两种情形, 采用 FIFO (先进先出算法) 页面替换算法, 计算会出现多少次缺页中断? (提示: 所给定的物理页块初始均为空, 因此, 首次访问一页时就会发生缺页中断。)

2. (本题 8 分) 某大学教室 A 和教室 B 之间有一条弯曲的小路, 其中从 T 到 S 一段路每次只允许一辆自行车通过, 但其中有一个小的安全岛 M (同时允许两辆自行车停留), 可供两辆自行车已从两端进入小路情况下错车使用, 如图 1 所示。试设计同步算法使教室 A 和教室 B 之间来往的自行车均可顺利通过。(提示: 用 P、V 操作写出同步算法。)

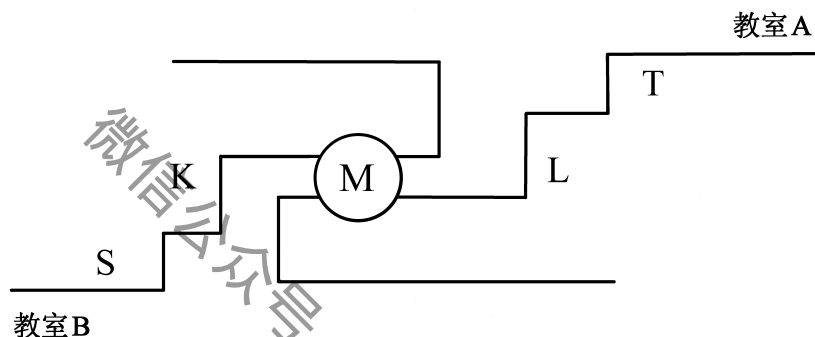


图 1 水路示意图

3. (本题 4 分) 有一页式系统, 其页表存放在内存中。
- (1) 如果对内存的一次存取需要 1.5 微秒, 问实现一次页面访问的存取时间是多少?
- (2) 如果系统增加有快表, 平均命中率为 85%, 当页表项在快表中时, 其查找时间忽略为 0, 问此时的存取时间为多少?

