

南京大学 2018 年硕士研究生入学考试初试试题
(A 卷)(三小时)

科目代码：845 科目名称：数据结构与算法、操作系统、
计算机系统基础、计算机网络 满分：150 分

适用专业：计算机科学与技术、软件工程、网络空间安全、计算机技术(专硕)

注意：①所有答案必须写在答题纸或答题卡上，写在本试题纸或草稿纸上均无效；
②本科目不允许使用计算器；③本试题纸须随答题纸一起装入试题袋中交回！

一、单项选择题：第 1~20 小题，每小题 2 分，共 40 分。下列每题给出的四个选项中，只有一个选项最符合题目要求。

1. 设某数据结构 A 的二元组形式表示为 $A=(D, R)$ ，其中 $D=\{a, b, c, d, e, f, g\}$ ， $R=\{\langle a, b \rangle, \langle a, c \rangle, \langle b, e \rangle, \langle c, d \rangle, \langle d, e \rangle, \langle d, g \rangle, \langle e, f \rangle, \langle f, g \rangle\}$ ，则数据结构 A 是 ()
A. 线性结构
B. 树型结构
C. 图型结构
D. 物理结构
2. 设输入序列依次为 a、b、c、d、e，则通过栈的作用后可以得到的输出序列为 ()
A. e、c、d、a、b
B. c、b、e、d、a
C. c、a、b、e、d
D. a、e、d、b、c
3. 设一棵二叉树中有 a 个度数为 2 的结点，b 个度数为 1 的结点，则该二叉树中共有结点个数为 ()
A. $a+2b$
B. $2a+b$
C. $a+b+2$
D. $2a+b+1$
4. 设有一组初始记录关键码序列为(30, 18, 25, 50, 60, 70)，则将它们调整成堆只需交换两个数的位置，即 ()
A. 30 与 18
B. 30 与 25
C. 50 与 18
D. 60 与 70
5. 已知 $f = 2^{2^n}$ ， $g = 2^{n^2}$ ，以下正确的是 ()
A. $g \in O(f)$
B. $f \in O(g)$
C. $g \in \Omega(f)$
D. $f \in \Theta(g)$

6. 以下关于 P, NP, NPC, NP-hard 问题说法正确的是 ()
A. 如果一个问题为 NP 问题, 则它一定是 NPC 问题
B. 一个问题可以既是 P 问题, 又是 NP 问题
C. 如果一个问题被证明是 NP 问题, 则它肯定不是 NP-hard 问题
D. 不存在既是 NP, 又是 NP-hard 的问题
7. 使用链表来实现优先队列, 在当前元素个数为 n 时, 插入一个元素操作的最坏情况下时间复杂度为 ()
A. $O(n)$ B. $O(1)$
C. $O(\log n)$ D. $O(n^2)$
8. 以下指令中, 可以在用户态执行的是 ()
A. 请求系统服务 B. 设置时钟
C. 启动磁盘设备 D. 设置 CPU 模式
9. 系统形成死锁的必要条件, 不包括以下的 ()
A. 进程互斥地使用资源
B. 申请新资源时不释放已占有资源
C. 系统拥有较少资源
D. 存在一组进程循环等待资源
10. 在作业调度时, 为避免“饥饿”, 应采用的调度算法是 ()
A. 优先级调度 B. 先来先服务
C. 最短作业优先 D. 最短剩余时间优先
11. 在分页虚拟存储管理系统中, 缺页中断的含义是 ()
A. 访问地址的内存损坏
B. 运行中的程序已被损坏
C. 访问的页面不在内存中
D. 系统中内存的配置太小
12. 系统中加入新设备时, 通常要为其配备的是 ()
A. 系统调用程序 B. 设备驱动程序
C. 文件系统 D. 库函数
13. 以下关于冯 诺依曼结构计算机的叙述中, 错误的是 ()
A. 由运算器、控制器、存储器和输入、输出设备组成
B. 程序代码由指令组成, 计算机按指定顺序自动执行指令
C. 程序中所包含的指令和数据都以二进制形式表示
D. 程序被启动运行后, 将被控制从外设直接读入 CPU 执行

14. 若 x 为 float 型变量, 赋值语句为 " $x=126.5;$ ", 则 x 的机器数为 ()
- A. 42FD0000H
B. C2FD0000H
C. 437D0000H
D. C37D0000H

15. 以下是在 Linux 系统中启动并加载可执行目标文件过程中 shell 命令行解释程序所做的部分操作:

- ① 构造参数 $argv$ 和 $envp$
- ② 调用 $fork()$ 系统调用函数
- ③ 调用 $execve()$ 系统调用函数
- ④ 读入命令 (可执行文件名) 及参数

启动并加载可执行目标文件的正确步骤是 ()

- A. ①→②→③→④
B. ②→④→①→③
C. ④→①→②→③
D. ④→①→③→②

16. 某计算机主存地址 32 位, 按字节编址, L1 data cache 和 L1 code cache 采用 8-路组相联方式, 主存块大小 64B, 采用回写 (Write Back) 方式和随机替换策略, 两种 cache 的数据区都是 32KB, 问 L1 cache 总容量至少有 ()

- A. 530K 位
B. 531K 位
C. 533K 位
D. 534K 位

17. 以下选项中, 会引起 CPU 自动查询有无中断请求进而可能进入中断响应周期的是 ()

- A. 一条指令执行结束
B. 一次 I/O 传输结束
C. 一次总线操作结束
D. 一次 DMA 传输结束

18. 在开放系统互联 (OSI) 七层协议体系中, 超文本标记语言 HTML 处于下列哪一层 ()

- A. 表示层
B. 应用层
C. 传输层
D. 会话层

19. IPv4 地址标记 192.218.36.0/24 所定义的子网包含可用的 IP 单机地址数为 ()

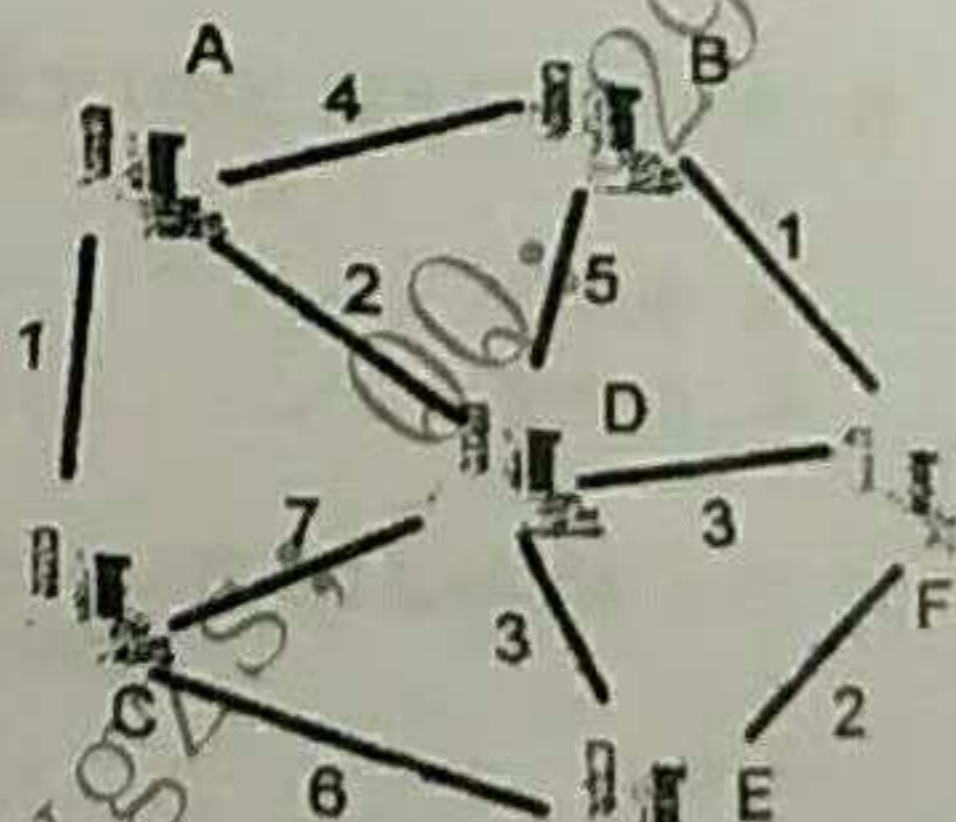
- A. 24
B. 254
C. 255
D. 256

20. 下列交换设备完成数据帧的存储转发, 在过程中不需要修改帧数据的设备是 ()

- A. 路由器
B. 无线 Hub
C. 网桥
D. ATM 交换机

二、综合应用题：第 21~29 题，共 110 分。

21. (10 分) 将关键码序列 $\{20, 13, 9, 5, 8, 2, 3\}$ 依次插入一棵初始为空的 AVL 树中，
- 1) 请画出每插入一个关键码后的 AVL 树；(5 分)
 - 2) 请用 C/C++ 语言写出在 AVL 树中搜索关键码 X 的递归算法；(3 分)
 - 3) 假设 AVL 树中的结点数为 n ，分析(2)中算法的时间复杂度(用大 O 表示法)。(2 分)
22. (10 分) 散列表的地址区间为 0-12，散列函数为 $H(K) = K \bmod 13$ ，采用开散列法(链地址法)处理冲突，请将关键码序列 $\{20, 26, 7, 38, 8, 12, 40\}$ 依次存储到散列表中。
- 1) 请画出所有关键码存储完后散列表的结构；(6 分)
 - 2) 基于 1) 中的散列表，以相同概率从 $\{20, 26, 7, 38, 8, 12, 40\}$ 中抽取关键码进行搜索，计算搜索成功的平均搜索长度；(2 分)
 - 3) 基于 1) 中的散列表，假设每个散列地址被搜索的概率相同，计算搜索失败的平均搜索长度。(2 分)
23. (10 分) 一个 Bitonic 数组定义为一个先增后降的整数序列。Bitonic 搜索问题是：给定一个由 N 个互不相同的正整数组成的 Bitonic 数组，输入一个整数 key ，判断该整数是否在该数组中。
- 例如，给定 Bitonic 数组 $A = [2, 3, 5, 7, 8, 6, 4]$ ，
- 输入： $key = 5$ ，返回：true
- 输入： $key = 9$ ，返回：false
- 1) 写一个算法实现 Bitonic 搜索。要求比较的次数越少越好。请用自然语言描述你的算法思想。(7 分)
 - 2) 分析在最坏情况下，你的算法的比较次数。(3 分)
24. (16 分) 如图所示，一个电信公司要在 6 个城市之间铺设光缆，使得城市之间可以相互通信。各城市之间铺设光缆的费用不同，其代价如图上的数字表示。电信公司希望寻找一种造价最低的最优铺设方案。请回答以下问题：
- 1) 用图模型对该问题进行建模，请给出该图的邻接矩阵表示形式。(2 分)
 - 2) 请针对该图给出一种最小代价铺设方案，并计算该方案的总代价。(2 分)
 - 3) 请用伪代码描述最优铺设方案的算法，并分析该算法的时间复杂度(用大 O 表示法)。(5 分)
 - 4) 假如在图中删除一条边，会使得最优铺设方案的代价增加，则该边称为“关键边”。请用自然语言描述一个算法思想，对给定的图 G ，输出所有关键边，并分析你的算法的时间复杂度(用大 O 表示法)。(7 分)



25. (7分) 简述虚拟存储管理的基本原理。举一个简单的例子说明：为什么可以用较小的物理内存配置，支持多个并发进程使用较大的运行空间？
26. (7分) 某一系统中，共有4种资源：A、B、C、D，资源总量分别为5、6、5、4个。假定系统中有5个进程（P0、P1、P2、P3和P4）运行，对4种资源的最大需求分别为（5，2，1，3）、（4，1，1，3）、（3，1，0，0）、（2，1，2，0）和（1，0，2，1）。如果系统开始后，进程P0、P2、P4依次产生如下资源请求： $R_0 = (3, 2, 0, 0)$ ， $R_1 = (2, 1, 0, 0)$ ， $R_2 = (1, 0, 1, 1)$ 。为避免死锁，系统应如何处理这3个请求，并给出理由。
27. (11分) 一条南北走向的大河上，有一座独木桥。东岸和西岸的汽车可分别通行：为保证安全，东岸或者西岸的汽车需要过桥时，如果桥面上无车，则允许一方的汽车过桥；当一方的汽车全部过完后，另一方的汽车才可以过桥。而且，桥面上最多可以承载12辆汽车同时通行。试用信号量和PV操作，写出东岸、西岸的汽车过独木桥的并发程序（需要说明信号量和变量的作用）。
28. (25分) 一个C语言程序有两个源文件：main.c和test.c，它们的内容如图所示。

```
/* main.c */
1  #include <stdio.h>
2
3  int sum();
4  int buf[4] = { -259, -126, -1, 60 };
5  extern int s;
6  void main()
7  {
8      s = sum();
9      printf("sum=%d\n", s);
10 }
```

```
/* test.c */
1  #define N 4
2  extern int buf[];
3  int s = 0;
4  int sum()
5  {
6      int i;
7      for (i = 0; i < N; i++)
8          s += buf[i];
9      return s;
10 }
```

在IA-32/Linux平台上用GCC编译驱动程序处理上述源程序，生成的可执行文件名为test，使用“objdump -d test”得到sum函数的反汇编结果如下。

```
00404448 <sum>:
00404448: 55                push    %ebp
00404449: 89 e5             mov     %esp, %ebp
0040444b: 83 ec 10         sub     $0x10, %esp
0040444e: c7 45 fc 00 00 00 00 movl    $0x0, -0x4(%ebp)
00404455: eb 1e            jmp     00404471 <sum+0x29>
00404457: 8b 45 fc         mov     -0x4(%ebp), %eax
0040445a: 8b 14 85 dc 95 00 mov     0x80495dc(, %eax, 4), %edx
0040445f: a1 f0 95 04 00   mov     0x80495f0, %eax
00404466: a3 f0 96 04 00   add     %edx, %eax
0040446d: 83 45 fc 01      mov     %eax, 0x80495f3
00404471: 83 7d fc 03      addl    $0x1, -0x4(%ebp)
00404475: 7e a9            cmpl    $0x3, -0x4(%ebp)
00404477: a1 f0 96 04 00   jle     00404457 <sum+0xf>
0040447c: c9               mov     0x80495f6, %eax
0040447d: c8               leave   %eax
ret
```

回答下列问题或完成下列任务。

- 1) 简要说明生成可执行文件test的过程。(4分)

- 2) 已知数组 buf 的首址为 0x80496dc，则 0x80496dc、0x80496de 这两个存储单元的内容分别是什么（用十六进制表示）？（2 分）
- 3) sum 函数机器代码占多少字节？哪几条是非顺序执行的跳转类指令？（4 分）
- 4) 地址 804845a 开始的 4 条指令实现了 sum 函数中那条语句的功能？其中，EDX 寄存器和存储单元 0x80496f0 中存放的分别是什么内容？（4 分）
- 5) 已知只读代码段和可读写数据段的起始地址分别为 0x8048000 和 0x8049000，符号 buf、s 和 sum 分别定义在虚拟地址空间的哪个段内？（3 分）
- 6) 若运行 test 过程中没有其他用户进程在执行，则在执行 sum 函数过程中，访问指令和访问数据各会发生几次缺页？为什么？（4 分）
- 7) 为什么在 main.c 中需要有 #include <stdio.h>？在 printf() 函数的执行过程中，如何从用户态陷入内核态执行？（4 分）

29. (14 分) 如图所示的网络，A 和 B 两台主机通过 TCP 协议通信，中间经过路由器 X 和 Y；其中主机 A 运行客户端程序，主机 B 运行服务器程序。



令：信号的链路传播速度为 $2 \times 10^8 \text{ m/s}$ ；

RTT：A、B 之间的往返传播时间；

W：A、B 采用的 TCP 窗口大小，以 Byte 为单位；

R：链路上传输带宽，以 bps (bit/s) 为单位；

- 1) 令 A 到 X、B 到 Y 的链路带宽 R_{AX} 和 R_{BY} 为 1G，X 到 Y 的骨干链路带宽 R_{XY} 为 10G， R_{XY} 平均分成 100 组信道，A 和 B 租用了其中一组，试计算 A 到 B 通信的可用带宽；（1 分）
- 2) 试计算 A 到 B 的往返传播时间；（1 分）
- 3) 简述 TCP 的滑动窗口机制和作用；（2 分）
- 4) 若只考虑服务器主机 B 向客户端主机 A 传送数据，A 的窗口大小 $W=64\text{K}$ ，则 A 和 B 之间的 TCP 传输能够达到的最大吞吐量是多少？注：采用 bps 表示，只计算 TCP 数据量；（2 分）
- 5) 令 IP 层和物理网的带宽消耗是 10%，如果要让 A、B 之间的 TCP 传输达到最大吞吐量，则 W 的值最小是多少？（1 分）
- 6) 如果 A 和 B 租用了 R_{XY} 的 10 组信道，则达到最大吞吐量的 W 值是多少？（1 分）
- 7) 简述 Jacobson 拥塞控制算法，即具有拥塞避免的慢启动算法（2 分）；
- 8) 假设 TCP 使用 Jacobson 拥塞控制算法，数据包大小为 1K Byte（窗口管理单位），若窗口 $W=2\text{M}$ ，确定当出现一个超时之后，再次达到该窗口所需要经过的时间，注：忽略传输时间；（2 分）
- 9) 根据计算结果，试说明 Jacobson 拥塞控制算法的缺点和改进方案。（2 分）