

北京大学信息科学技术学院考试试卷

考试科目： 算法设计与分析 姓名： _____ 学号： _____

考试时间： 2009 年 6 月 9 日 任课教师： _____

题号	一	二	三	四	五	六	七	八	总分
分数									
阅卷人									

考 场 纪 律

1. 请持学生证入场考试，并按指定座位就座；除必要的文具和教师指定的用具用书外，其他所有物品包括手机、呼机、MP3、电子词典、书籍、笔记、纸张等严禁带入座位，必须放在指定位置。凡有试题印制问题请向监考教师提出，不得向其他考生询问。
2. 认真、诚实、独立并在规定时间内完成答卷，严禁任何形式的违纪作弊行为；否则，本答卷成绩以 0 分记，并根据《北京大学本科考试工作与学术规范条例》给予纪律处分。
3. 提前交卷的考生不要在考场逗留，不要在门口、窗外大声喧哗。考试结束时间到，请停止答卷，在座位等候监考教师收卷并清点完毕，方可离开考场；考题和试卷不得带出考场。

以下为试题和答题纸，共 9 页。

装
订
线
内

不
要
答
题

一、填空题(选做 5 道, 10 分)

1. 用矩阵幂的方法求斐波那契数, 其运行时间为 ($\log n$)。
2. 对于一个可以用动态规划法求解的问题, 要求问题既要满足 (最优子结构) 的特性, 又要具有大量的 (重叠子问题)。
3. 对于一个可以用贪心法求解的问题, 不仅要求问题满足 (最优子结构) 的特性, 还应证明其贪心策略的 (正确性)。
4. 设有 n 个栈操作 (PUSH、POP、MULTIPOP) 的序列, 作用于初始为空的栈 S 。不区分三种操作, 则每个操作的最坏运行时间为 (n), 平摊运行时间为 (1)。
5. 三种平摊分析的方法分别为 (聚集分析)、(记账)、(势能分析)。
6. 四后问题的搜索空间为 (4叉树) 树; 0-1 背包问题的搜索空间为 (子集) 树; 巡回售货员问题的搜索空间为 (排列) 树。
7. (回溯) 法的求解目标是找出解空间树中满足约束条件的所有解, 而 (分支限界) 法的求解目标则是找出满足约束条件的一个解, 或是在满足约束条件的解中找出在某种意义下的最优解。
8. 回溯法一般以 (深度) 优先的方式搜索解空间树, 而分支限界法则一般以 (广度) 优先或以最小耗费优先的方式搜索解空间树。

二、单项选择题 (10 分)

1. 下列关于排序算法的叙述, 不正确的是? ()

- A) 堆排序的最差情形运行时间为 $\Theta(n \lg n)$
- B) 快速排序平均情形运行时间为 $\Theta(n \lg n)$
- C) 任何排序算法的最差情形运行时间都不可能比 $\Omega(n \lg n)$ 更小**
- D) 插入排序在最好情形下的运行时间为 $\Theta(n)$

2. 对于课堂讲解的线性时间内找第 i 小的元素的算法, ()
下列叙述中不正确的是?

- A) 算法第一步中可以按每五个元素一组找中位数;
- B) 算法第一步中可以按每七个元素一组找中位数;
- B) 算法第一步中不能按每三个元素一组找中位数;
- D) 如果要求的 n 个元素的中位数, 则中位数一定是第一步中找到的中位数中的某一个。**

3. 主方法可以求解满足形如下式的递推方程, ()

$$T(n) = aT\left(\frac{n}{b}\right) + f(n)$$

则下列关于方程中的约束中不准确的是?

- A) 对于系数 a , 必须满足 $a \geq 1$
- B) 对于系数 b , 必须满足 $b > 1$
- C) 若对于常数 $\varepsilon > 0$, $f(n) = O(n^{\log_b a - \varepsilon})$, 则 $T(n) = \Theta(n^{\log_b a})$
- D) 若 $f(n) = O(n^{\log_b a})$, 则 $T(n) = \Theta(n^{\log_b a} \log n)$**

4. 下列哪些问题不能用贪心法求解? ()

- A) 霍夫曼编码问题
- B) 单源最短路径问题
- C) 0-1 背包问题**
- D) 最小生成树问题

5. 可合并堆上可以不包含下列哪个操作? ()

- A) DECREASE-KEY(H, x, k) B) UNION(H_1, H_2)
C) INSERT(H, x) D) EXTRACT-MIN(H)

6. 不同堆上插入操作最差情形下的开销或平摊开销, ()

对二叉堆、二项堆和斐波那契堆, 下列选项中描述错误的是?

- A) 二叉堆为 $\Theta(\lg n)$ B) 二项堆为 $O(\lg n)$
C) 斐波那契堆为 $\Theta(1)$ D) 三种堆的开销都是 $\Theta(\lg n)$

7. 关于网络流的割, 下列选项中错误的是? ()

割(S, T)是流网络 $G=(V, E)$ 的一个划分, 其中 $s \in S, t \in T$ 。如果 f 是 G 上的流, 那么流经割的净流量为 $f(S, T)$, 割(S, T)上的容量定义为 $c(S, T)$ 。

- A) $|f| \leq c(S, T)$ B) $f(S, T) = |f|$
C) $f(s, V-s) = |f|$ **D) $f(S-s, V) = |f|$**

8. 下列随机算法一定有解但解不一定正确的是? ()

- A) Sherwood B) Las Vegas
C) Monte Carlo D) 三者都不是

9. 在快速排序算法中引入随机过程的主要目的是什么? ()

- A) 改善确定性算法的平均运行时间
B) 保证算法总能在 $O(n \lg n)$ 时间内结束
C) 避免了算法最坏情况下的发生
D) 改善了确定性算法最坏情形下的平均运行时间

10. 用 Monte Carlo 方法估计四后问题回溯算法的效率。 ()

五次实验结果分别为 $\langle 1, 4, 2 \rangle$ 、 $\langle 2, 4, 1, 3 \rangle$ 、 $\langle 4, 2 \rangle$ 、 $\langle 3, 1, 4, 2 \rangle$ 、 $\langle 1, 3 \rangle$, 则解空间中的结点数估计为?

- A) 16 **B) 16.2** C) 17 D) 16.5

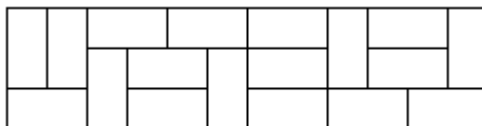
三、社会名流(20 分)

在 n 个人中，一个被所有人知道但却不知道别人的人，被定义为社会名流。现在的问题是如果存在，试找出该社会名流。你可以使用的唯一方式是询问：“请问你知道那个人吗？”请给出提问次数为 $O(n)$ 的算法，写出伪代码，分析算法的正确性，并给出算法运行时间的精确分析（即 $O(n)$ 中隐藏的系数）。

（提示：当你问 A 是否认识 B 时，如果 A 认识 B ，则 A 不是社会名流；如果 A 不认识 B ，则 B 不是社会名流）

四、地板覆盖 (20 分)

用 2×1 的地板块覆盖 $3 \times n$ 的地面有多少种方案？如下图是一个覆盖的例子，函数 fn 可用于求解这个问题，请说明 fn 算法的正确性，并说明算法运行时间的上界和下界。



```
int fn(int n) {
    if (n % 2 == 1) return 0;
    int [] f = new int[n+1];
    f[0] = 1;
    for (int i = 2; i <= n; i += 2) {
        f[i] = f[i-2]*3;
        for (int j = i-4; j >= 0; j -= 2)
            f[i] += f[j]*2;
    }
    return f[n];
}
```

五、田忌赛马（20 分）

你一定听过田忌赛马的故事吧？如果 3 匹马变成 n 匹，齐王仍然让他的马按从优到劣的顺序出赛，田忌可以按任意顺序选择他的赛马出赛。赢一局，田忌可以得到 200 两银子，输一局，田忌就要输掉 200 两银子。已知国王和田忌的所有马的奔跑速度，并且所有马奔跑的速度均不相同，现已经对两人的马分别从快到慢排好序，请设计一个算法，帮助田忌赢得最多的银子。写出伪代码，证明算法的正确性，并分析算法的复杂度。

（提示：可以设计一个贪心策略的算法，面对国王每匹顺序出场的马，如果田忌的马快，就派最快的出场；否则派最慢的马出场）

六、(20 分) 给出 n 项作业 $J_1, J_2 \dots J_n$ ，对应每项作业有一个运行时间 t_i ，在 m 个处理器上调度这些作业，使完成的时间最小。完成的时间定义为在所有的处理器中运行时间最长的处理器运行的时间。采用如下的近似算法：即，按照原始给定的作业顺序： $J_1, J_2 \dots J_n$ ，把每一项作业分配给当前情况下最近可用的那个处理器，使该作业尽可能早被处理（其它没有任何约束）

1. 试证明该算法的近似度为 $2 - 1/m$ 。(10 分)
2. 构造边界情况，说明这个界是紧的。(10 分)

(提示: $OPT \geq \frac{1}{m} \sum t_i, OPT \geq \max t_i$)