

《算法设计与分析》期末考试

考试形式：闭卷

考试时间：两小时

| 问题 | 一 | 二 | 三 | 四 | 五 | 六 | 附加题 | 总分 |
|-------|----|----|----|----|----|----|-----|-----|
| 分数 | 15 | 15 | 15 | 20 | 15 | 20 | 0 | 100 |
| 附加题分数 | 0 | 0 | 0 | 0 | 0 | 0 | 15 | 15 |
| 得分 | | | | | | | | |

第一题.....15 分

给定一个包含 n 个不同整数的数组 $A[1 \cdots n]$ ，假设数组中的元素已经按照升序排列。请设计一个算法能够在 $O(n)$ 时间内判断数组中是否存在一个元素 $A[i]$ 使得 $A[i] = i$ 成立。要求简述算法设计思路，给出伪代码，并简要说明算法时间复杂度。

第二题.....15 分

请设计一个支持以下两个操作的数据结构：

- $\text{Insert}(x)$ ：向数据结构中插入一个正整数元素 x 。
- $\text{Median}()$ ：返回当前所有已插入元素的中位数。

设计目标是，当数据结构在包含 n 个元素时，上述每个操作的时间复杂度为 $O(\log n)$ 。你可以设计一个全新的数据结构，也可以在课程中学过的数据结构的基础上进行改进。要求简述 Insert 与 Median 操作的实现方法，并简要说明两个操作的时间复杂度确实为 $O(\log n)$ 。在描述 Insert 与 Median 操作的实现方法的过程中，可以借助伪代码以方便表述。

第三题.....15 分

单调栈是一种特殊的栈，它支持如下两种操作：

- $\text{ORDEREDPUSH}(x)$ ：将栈顶部所有小于 x 的元素弹出，然后将 x 添加到栈顶；
- $\text{POP}()$ ：将栈顶的元素从栈中弹出并返回。

假设使用双向链表来实现单调栈。请证明，如果从一个空的单调栈开始执行一系列 ORDEREDPUSH 和 POP 操作，则每个操作的均摊开销均为 $O(1)$ 。

第四题.....20 分

从下面两题中任选一题进行解答。如果两题你都尝试给出解答，请明确指出改卷时以哪道题目的解答记分；否则改卷时默认以你第一题的解答记分。

(1) 考虑一个用邻接表存储的连通无向图 $G = (V, E)$ ，一个边集的子集 E' 是一个 **feedback edge set**。当且仅当其满足如下条件： E' 与图 G 中的每一个环均相交。也即，将 E' 中的边从图中去除后，图将变为无环图。假设图 G 也是一个带权图：每条边 $e \in E$ 有一个正整数权重 w_e 。

- 请简要描述一个算法用于计算图 G 的最大生成树（如自然语言难以清晰描述，可配合给出伪代码），并说明其时间复杂度。（10 分）
- 请设计一个算法找出图 G 的最小 **feedback edge set**，也即找出一个 $E' \subseteq E$ 使其为一个 **feedback edge set**，且 $\sum_{e \in E'} w_e$ 尽量小。要求简述算法设计思路（如自然语言难以清晰描述，可配合给出伪代码），满分解答的时间复杂度应为 $O(|E| \cdot \log |V|)$ 。（10 分）

(2) 考虑一个连通无向图 $G = (V, E)$ ，其中每条边 $e \in E$ 均有一个非负权重 w_e 。

- 请写出单源点最短路径算法 **Dijkstra** 算法的伪代码。（10 分）
- 请说明如何修改 **Dijkstra** 算法，使得给定起点 u 和终点 v ，能够给出一条从 u 到 v 的包含边数最少的最短路径。要求给出改进后的算法的伪代码，并说明算法的时间复杂度。（10 分）

第五题.....15 分

课堂中我们对“活动选择”(activity selection)问题给出了一个贪心算法：挑选结束时间最早的活动，剔除与结束时间最早的活动相斥的活动，然后对剩下的活动递归。课堂中我们证明了这个贪心算法是正确的。本题中我们考察“带权的活动选择”问题：给定 n 个活动的开始时间 $s[1, \dots, n]$ 、结束时间 $f[1, \dots, n]$ 、权重 $w[1, \dots, n]$ ，我们希望找出一系列不相互冲突的活动，使得这些活动的总权重最大。李雷想使用贪心策略来解决带权的活动选择问题，他提出了三种策略：

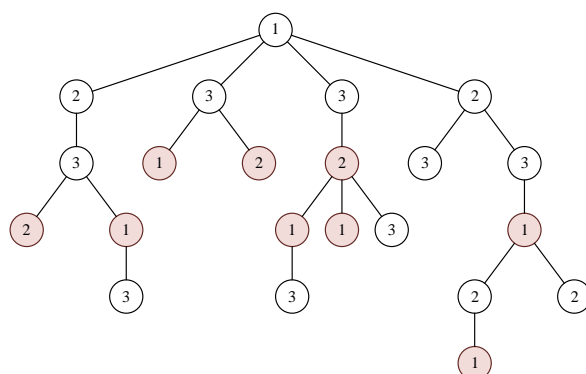
- 挑选结束时间最早的活动，剔除与该活动相斥的活动，然后对剩下的活动递归。
- 挑选权重最大的活动，剔除与该活动相斥的活动，然后对剩下的活动递归。
- 挑选与其他活动冲突数量最少的活动，剔除与该活动相斥的活动，然后对剩下的活动递归。

对于每一种策略，请分析所得的贪心算法是否正确。如果正确，请证明其正确性，否则请给出反例。(提示：三种策略中至少有一种是错误的。)

第六题.....20 分

从下面两题中任选一题进行解答。如果两题你都尝试给出解答，请明确指出改卷时以哪道题目的解答记分；否则改卷时默认以你第一题的解答记分。

- (1) 给定任意一个正整数 n ，我们总是可以找到一系列完全平方数(即 $1, 4, 9, 16, \dots$)使得它们的和为 n 。请设计一个算法，在输入正整数 n 后，计算出至少需要多少个完全平方数求和即可得到 n 。例如，如果输入12，则算法输出为3，即 $12 = 4 + 4 + 4$ 。又如，如果输入13，则算法输出为2，即 $13 = 4 + 9$ 。要求简述算法设计思路，给出伪代码，并说明算法的时间复杂度。注意如果你给出了一个贪心算法，则必须证明其正确性。
- (2) 假设你将要为一家公司组织一场年会。这家公司的层次森严，星树形结构：总经理是根节点，除总经理外每个员工都有一个直接上级。年会上，每个员工都会得到一件礼品。礼品有三种：特等礼品、一般礼品和安慰礼品。原则上，每位员工应该依据自己的考评成绩获得对应的礼品。但实际上，公司要求每位员工和自己的直接上级不能获得同样的礼品。并且，如果一个员工获得的礼品比自己的直接上级要好，那么在来年其可能因为上级的嫉妒心而被开除。例如下图中。节点内的数字代表了员工获得的礼品种类(1为特等礼品，3为安慰礼品)，加粗的节点为可能被开除的员工，有9位。



请设计一个算法决定年会上给每位员工发放什么礼品，以使得来年可能被开除的员工数目最少。注意，你并不需要知道每位员工的考评成绩。要求简述算法设计思路，给出伪代码，并说明算法时间复杂度。

【附加题】.....20 分

考虑一个使用开放寻址法(open addressing)构建的大小为 m 的哈希表。表中存储了 n 个元素，而且 $n \leq m/2$ 。假设均匀散列(uniform hashing)条件成立。对任意 $1 \leq i \leq n$ ，令随机变量 X_i 代表在进行第 i 次插入时进行的探查(probe)数目。令随机变量 $X = \max_{1 \leq i \leq n} \{X_i\}$ 代表最长的探查序列的长度，请证明 $E[X] = O(\log n)$ 。