作业里无关紧要的小错误也掰扯清楚了，否则考试又扣分

喜欢考作业题，尤其是体现点想法又有点难的

南京大学人工智能学院2021至2022学年秋季学期

《数据结构与算法分析》期中考试

考试形式：闭卷　　考试时间：两小时

问题	一	二	三	四	五	六	七	总分
分数	5	20	15	15	15	15	15	100
得分								

- 请将答案写在答题纸上。
- 请在答题纸的每一面上写上学号与姓名。
- 越后面的题目不一定越难！不要一直卡在一道题目上。
- 看清题目要求！例如题目未要求提供正确性证明时则无需说明正确性。
- 如无法提供题目要求的算法，给出一个性能较差的算法有时也可得到部分分数。

第一题 （5分） 下列每个问题的答案都是以下选项中的一个。对每个问题，给出答案所对应选项的字母即可，你无需证明你的结论。

(a) $\Theta(1)$　(b) $\Theta(\log n)$　(c) $\Theta(n)$　(d) $\Theta(n\log n)$　(e) $\Theta(n^2)$　(f) 以上答案均不对

c （1） 递推式$T(n) = T(\sqrt{n}) + n$的解。

d （2） 递推式$T(n) = T(n-1) + \lg n$的解。

d （3） 递推式$T(n) = 2T\left(\lceil \frac{n+27}{2} \rceil\right) + 5n - 8\sqrt{\lg n} + \frac{2021}{n}$的解。

d （4） 堆排序n个元素最坏情况下的时间复杂度。

e （5） 使用随机化选择算法（randomized selection）计算n个元素的中位数最坏情况下的时间复杂度。

第二题 （20分） 假设有k个已经排好序的数组$A_1, A_2, \cdots, A_k$，每个数组包含n个整数。我们的目标是将这些数组合并成为一个包含kn个整数的有序数组。请回答下列问题：

（1） 考虑所有基于比较的算法，给出并证明这些算法的一个时间复杂性下界。

（2） 设计一个基于比较的算法，要求给出算法描述并分析算法时间复杂性。描述过程中，可以借助伪代码以方便表述。

满分解答中，算法复杂性上界应与问题复杂性下界匹配。提示：$n! \approx \sqrt{2\pi n}\cdot(n/e)^n$。

第三题 （15分） 图1中的算法（参见第2面）是一个正确有效的排序算法：

（1） 分析该算法的时间复杂度。

（2） 简要证明该算法的正确性，即证明其可以正确的排序输入数组A。

第四题 （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。

```
StrangeSort(A[0, ··· , n − 1])
if (n == 2 and A[0] > A[1]) then
    Swap(A[0], A[1]).
else if (n > 2) then
    m ← ⌈2n/3⌉.
    StrangeSort(A[0, ··· , m − 1]).
    StrangeSort(A[n − m, ··· , n − 1]).
    StrangeSort(A[0, ··· , m − 1]).
```

图 1: StrangeSort排序算法

第五题 （15分） 请设计一个支持以下两个操作的数据结构：认为不相同元素

- Insert(x)：向数据结构中插入一个正整数元素x。
- Median()：返回当前所有已插入元素的中位数。$\lceil \frac{n}{2} \rceil$

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。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。

第六题 （15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$n \le m/2$。假设均匀散列（uniform hashing）条件成立。对任意$1 \le i \le n$，令X_i代表在进行第i次插入时进行的探查（probe）数目。令$X = \max_{1 \le i \le n}\{X_i\}$代表最长的探查序列的长度。请回答下列问题：

（1） 证明$\Pr[X_i > k] \le 1/2^k$对任意非负整数k成立。

（2） 证明$\Pr[X > 2\lg n] \le 1/n$。

（3） 证明$\mathrm{E}[X] = O(\log n)$。

第七题 （15分） 考虑一串数字$\langle a_1, a_2, \cdots, a_m \rangle$，我们称其为震荡的当且仅当：对所有奇数$i$，满足$a_i < a_{i+1}$；对所有偶数$i$，满足$a_i > a_{i+1}$。（即$a_1 < a_2 > a_3 < a_4 > \cdots$。）现在给定一个包含$n$个整数的数组$A[1 \cdots n]$，请设计一个算法计算出其中最长的震荡子数组的长度。（算法只需给出长度，无需返回具体的震荡子数组。）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。满分解答的算法时间复杂度应为$O(n)$。 （子数组连续）