装订线内 不要答题

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： 2017 年 4 月 12 日 **大班教师:**__________ **小班教师:**__________

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

北京大学考场纪律

1、考生进入考场后，按照监考老师安排隔位就座，将学生证放在桌面上。无学生证者不能参加考试；迟到超过 15 分钟不得入场。在考试开始 30 分钟后方可交卷出场。

2、除必要的文具和主考教师允许的工具书、参考书、计算器以外，其它所有物品（包括空白纸张、手机、或有存储、编程、查询功能的电子用品等）不得带入座位，已经带入考场的必须放在监考人员指定的位置。

3、考试使用的试题、答卷、草稿纸由监考人员统一发放，考试结束时收回，一律不准带出考场。若有试题印制问题请向监考教师提出，不得向其他考生询问。提前答完试卷，应举手示意请监考人员收卷后方可离开；交卷后不得在考场内逗留或在附近高声交谈。未交卷擅自离开考场，不得重新进入考场答卷。考试结束时间到，考生立即停止答卷，在座位上等待监考人员收卷清点后，方可离场。

4、考生要严格遵守考场规则，在规定时间内独立完成答卷。不准交头接耳，不准偷看、夹带、抄袭或者有意让他人抄袭答题内容，不准接传答案或者试卷等。凡有违纪作弊者，一经发现，当场取消其考试资格，并根据《北京大学本科考试工作与学术规范条例》及相关规定严肃处理。

5、考生须确认自己填写的个人信息真实、准确，并承担信息填写错误带来的一切责任与后果。

学校倡议所有考生以北京大学学生的荣誉与诚信答卷，共同维护北京大学的学术声誉。

答题要求：解答算法设计题目时，请先用一段话描述算法思想。若用动态规划算法，请写出递推方程、边界条件、标记函数等设计要素；贪心法需给出证明；回溯法需给出解向量、搜索树等、约束条件；各种算法需分析时间复杂度。阅卷时会根据算法的正确性和效率评分。

得分

一、（15 分）在下表中填入“是”或者“否”，其中 k>0 和 c>1 是常数。

$f(n)$	$g(n)$	$f(n)=O(g(n))$	$f(n)=o(g(n))$	$f(n)=\Theta(g(n))$
n^2	$\log(c^n)$	N	N	N
$\sqrt{n}$	$n^{\sin n}$	N	N	N
$n2^n$	2^{2n}	Y	Y	N
$n^{\log\log n}$	$c^{\log n}$	N	N	N
$\log(n!)$	$\log(n^n)$	Y	N	Y

得分

二、（10 分）求解递推方程。要求给出求解过程或依据。

1) $T(n) = T(\mathrm{n}-1) + \log(5^n)$

$T(1) = 1$

2) $T(n) = 2T\left(\frac{n}{2}\right) + n^2\log n$

$T(1) = 1$

解答

（1）

迭代法归纳，$\mathrm{T}(n) = \Theta(n^2)$

（2）

主定理第三种情况，取$\mathrm{c} = 3/4$，则$af\left(\frac{n}{b}\right) \leq cf(n)$， $\mathrm{T(n)} = \Theta(n^2\log n)$

得分

三、（15 分）下面是算法 Determine 的伪码，输入 S 是 n 个不等的递减排好序的正整数的数组，a 是一个给定的正整数。请说明该算法的输出是什么？该算法在最坏情况下做多少次比较运算？

```
Determine(S,a)
1.  i←n
2.  y←S[i]
3.  while i ≥2 do
4.      x ← a*y
5.      在 S[1..i-1] 中二分检索 x
6.      if 存在 x
7.      then   return   x, y
8.      else   {i←i–1;   goto 2；}
9.  return "No"
```

参考答案：检查在 S 中是否存在两个不同的数 x 和 y，使得 $x/y=a$。

时间复杂度$O(n\log n)$。

得分

四、（15 分）Z 国有一条笔直的河流，恰好流经 Z 国的 n 个城市，按河水的流经顺序可以把 n 个城市分别编号为 1, 2, ⋯, n。我们把河流看成 x 轴，把 n 个城市看成 x 轴上的 n 个点，坐标分别记为 $x[1], x[2], \ldots, x[n]$（其中 $x[i] < x[i+1]$）。现在 Z 国的工程师需要在城市之间铺设线缆。

我们最多可以铺设 M 段线缆，每一段线缆可以从任意城市开始铺设，横跨若干连续的城市以后，在任意城市结束，线缆的长度任意（长度可以为 0，表示只在一个城市上建设一个网络，若长度大于 0，表示连续的多个城市在同一网络）。为了节省项目开支，国王希望给出一种铺设方案，使得这项技术惠及所有城市（即 M 条长度非负的线段覆盖 n 个点，形成 M 个连通块），但铺设线缆的总长度最小。

提示：若 $M \geq n$，每个城市的网络都是独立的，每段线缆长度均为 0，所以答案为 0。

参考答案：

假设方案 f 铺设了 K 段线缆，起始坐标和终止坐标分别为 s[1..K], t[1..K]。方案 f 满足条件（1）对任意线缆 i 有 s[i]<=t[i]；条件（2）对任意城市 j，存在线缆 i 使得 s[i]<=x[j]<=t[i]。方案 f 的铺设长度为

$$\mathrm{Len}(\mathrm{f}) = \sum_{i=1}^{K} (t[i] - s[i])$$

（a） 若 M>n，最优方案显然为：令 K=n，对任意线缆 i 有 s[i]=t[i]，则铺设长度为 0。

（b） 下面考虑 M<=n 的情况：

使用贪心法。

贪心策略：把相邻城市间距从小到大排序，先假想方案 f_0 由 n 段长度为 0 的线缆覆盖 n 个城市，每次在 f_{k-1} 的基础上把一对间距第 k 小的相邻城市(x[j], x[j+1])相连从而完成第 k 次连接两段线缆，得到方案 f_k，方案 f_k 仍满足条件（1）和条件（2），且线缆数目减小为 n-k，铺设长度增加了 x[j+1]-x[j]。任何铺设方案都可以由不同的合并策略获得。

贪心策略正确性证明使用归纳法。

显然 f_0 是使用 n 段线缆覆盖 n 个城市的最优方案。

假设 f_{k-1} 是使用 n-k+1 段线缆的最优方案，下面用反证法证明按贪心策略得到的方案 f_k 是使用 n-k 段线缆的最优方案。

假设 f_k 不是最优的，则存在最优方案 f_k'，使得 $Len(f_k')<Len(f_k)$。设从 f_{k-1} 到 f_k 转变中增加的那段覆盖区间为 $b=(x[j],x[j])$，则有 $Len(f_k)=Len(f_{k-1})+|b|$，所以 $Len(f_k')<Len(f_k)=Len(f_{k-1})+|b|$。

当 f_k' 覆盖了区间 b 时，在 f_k' 基础上去掉 b 得到 f_{k-1}'，有 $Len(f_{k-1}')=Len(f_k')-|b|<Len(f_{k-1})$，这与 f_{k-1} 是使用 n-k+1 段线缆覆盖 n 个城市的最优方案的假设矛盾。

当 f_k' 不覆盖区间 b 时，则 f_k' 必定覆盖了某个长度比的|b|大的区间 d，若交换 b 与 d 的覆盖状态，得到的 f_k'' 更优，这与 f_k' 是使用 n-k 段线缆的最优方案的假设矛盾。

综上所述，f_k 是使用 n-k 段线缆的最优方案。贪心策略正确性得证。

算法描述：

说明：我们最后只需要前 n-M 小的数字集合，所以实现时不一定要排序。

输入：n、M 及 x[1], x[2], ……, x[n]

输出：最小的铺设长度

（1） 特判 M>=n 的情况，直接输出 0

（2） 令 C[i]=x[i+1]-x[i]，i=1, 2, ……, n-1

（3） 使用 Select 算法选出数组 C 中第 n-M 小的城市间距，可同时得到前 n-M 小的集合 D

（4） 计算 Len=x[n]-x[1]-sum(D)

时间复杂度分析：

上述算法中每一步时间复杂度均为 O(n)，故总时间复杂度为 O(n)。

得分

五、（15 分）Z 国有一条笔直的河流，恰好流经 Z 国的 n 个城市，按河水的流经顺序可以把 n 个城市分别编号为 $1, 2, \ldots, n$。我们把河流看成 x 轴，把 n 个城市看成 x 轴上的 n 个点，坐标分别记为 $x[1], x[2], ..., x[n]$（其中 $x[i] < x[i+1]$）。现在 Z 国的国王需要从 n 个城市中选一些城市来建设通信基站。

每个基站的覆盖范围为半径为 R 的圆。如果城市 i 在基站 j 的覆盖范围内，则城市 i 的居民可以利用基站 j 进行通信。如果两个基站的覆盖范围有重叠部分（也即两个圆不是相离的，注意是不相离！！！），则认为这两个基站是互通的，且基站间的互通具有传递性。现在国王想知道，给定城市个数 n，城市坐标 $x[1], x[2], ..., x[n]$，以及基站覆盖半径 R，至少需要建设多少个基站才能使城市 1 与城市 n 可以互相通信。

提示：每个基站都必须建在城市上，但不是每个城市上都建基站。

参考答案：使用贪心法。

贪心策略：（1）首先选出最大的 p_1 使得 $x[p_1]-x[1]<=R$；（2）然后依次选出尽量大的 p_k，满足 $x[p_k]-x[p_{k-1}]<=2R$；（3）重复第（2）步直到存在 p_m 满足 $x[n]-x[p_m]<=R$。

贪心策略正确性证明（反证法）：

设按上述贪心策略选出的基站选址序列为 $p_1, p_2,\cdots\cdots, p_m$，假设该序列不是最优解，则存在最优解 $q_1, q_2,\cdots\cdots, q_t$，其中 $t<m$。

必定存在 k 使得 $p_i= q_i$, $i=1,2,\cdots\cdots,k-1$，且 $p_k\neq q_k$（否则上述贪心策略在找到 p_t 时就会停下来）。且因为上述贪心策略每次选出尽量大的 p_k 满足 $x[p_k]-x[p_{k-1}]<=2R$，所以有 $q_k< p_k$。于是有 $p_i>=q_i$, $i=k+1, k+2, \cdots\cdots, t$，又因为 q_t 为最优解的最后一个基站选址，所以有 $x[n]-x[q_t]<=R$，于是有 $x[n]-x[p_t]<=R$，进一步有 $t=m$，这与 $t<m$ 的假设矛盾。

所以上述贪心策略选出的基站选址序列为最优解。

算法描述：

（1）首先选出最大的 p_1 使得 $x[p_1]-x[1]<=R$

（2）然后依次选出尽量大的 p_k，满足 $x[p_k]-x[p_{k-1}]<=2R$

（3）重复第（2）步直到存在 p_m 满足 $x[n]-x[p_m]<=R$。

时间复杂度分析：

序列 p 具有单调性，所以扫描一遍原坐标序列即可找出基准选址序列，时间复杂度为 O(n)。

得分

六、（15 分）Z 国需要通过进口原油来补充满足国内旺季需求。Z 国的某海港建有储油库存储原油，存储总容量为 L 吨，每吨原油在储油库每存储一天都有一个固定的费用 C 元。原油通过远洋油轮运来，一艘远洋油轮的运输量非常大，足够装满整个储油库。远洋油轮的运输费用按次计算，每次运费固定为 P 元。每年的旺季长 n 天，基于以往的经验，Z 国知道旺季的每一天原油缺口数量为 $g[i]$吨，缺口原油每天从储油库运走。为了在原油需求旺季结束那天恰好清空原油库存，Z 国需要制定一个原油采购运输计划，以保证总体的运营成本（运输+存储）最小。请问 Z 国都需要在旺季的哪些天让采购的原油由远洋油轮运到港口入库？假设旺季开始前储油库也是空的，每天到港的原油当天就可以入库并使用，题目中所有的数值均为正整数。

解法：

引理 1：每次订购汽油时，为使得总体运营成本最低，前一天结束时的库存一定必为 0。否则，对于任意具有非零库存的最优解，我们可以将非零的库存汽油的等量汽油加在第二天订购时订购量中，则运费不变，而存储费用更低，矛盾。

推论 1：如果所有的订购日期确定下来，则订购量也相应的是确定的。

设$F(j)$表示在第j天结束时库存为空的一个最佳订购方案的最小运营费用。则

$$F(j) = \min_{1\le i<j,\sum_{k=i+1}^{j} g_k \le L} \left\{F(i) + P + C\sum_{k=i+1}^{j} g_k(k-i-1)\right\}$$

标记函数，对于$F(n)$取得最小值的i为最后一次订购的日期，订购数量根据推论 1 计算。其它订购日期依此类推。

动态规划公式正确 +5 分

复杂度分析：$O(n^2)$ +4 分 二分优化 $O(n\log n)$ +6 分 斜率优化$O(n)$ + 8 分

$O(nL)$ +2 分 （$F(i,j)$: 表示 i 天后油库剩余 j 的最小费用。）

标记函数含义正确 +2 分

得分

七、（15 分）随着社会的发展，Z 国的城市不再是沿河流分布。现在 Z 国的 n 个城市分布在一个二维平面上。

Z 国的工程师设计出新一代 xG 移动通信技术，推广这项技术需要在每个城市内部建设相应的基站来为该城市的居民提供服务，但为了避免不同城市的基站信号互相干扰的问题，相邻城市的基站需要调制到不同的发射频率上。

现给定城市个数 n 以及他们之间的相邻关系 C（C 以邻接矩阵的方式给出，若 $C[i][j]=1$ 表示城市 i 与城市 j 相邻，否则不相邻，数据保证构成平面图），问至少需要多少个不同的发射频段才能满足需求。

提示：可能存在城市不与其他城市相邻（四面环海）。

图例：

若 4 个城市的位置关系如下：

```
+---+---+
| 1 | 2 |
+---+---+
| 3 | 4 |
+---+---+
```

则邻接矩阵 C 如下：

```
0 1 1 0
1 0 0 1
1 0 0 1
0 1 1 0
```

一种可行的发射频段安排方案为：

```
+---+---+
| 1 | 2 |
+---+---+
| 2 | 1 |
+---+---+
```

其中数字相同表示在同一频段。

参考答案：

该题相当于地图着色问题，根据四色定理，所需频段数量不超过 4。易知可以在 O(n)时间内特判是否存在答案为 1 或 2 的频段分配方案，若不存在则通过搜索方法判断是否存在答案为 3 的频段分配方案，若仍不存在则可以直接确定答案为 4。下面给出判断是否存在答案为 3 的频段分配方案的搜索方法。

设解向量为 $<x_1, x_2, \cdots\cdots, x_n>$，其中 $x_i \in \{1, 2, 3\}$ 表示第 i 个基站的发射频段的编号。搜索空间是 3 叉树。部分向量 $<x_1, x_2, \cdots\cdots, x_k>$ 表示已经考虑前 k 个基站的频段分配，现考虑第 k+1 个基站的频段分配，结点分支的约束条件为：

$$x_{k+1} \in \{1,2,3\} - \{x_i \mid i <= k，C[i, k+1] = 1\}$$

进一步优化：由于频段分配的编号是无序的（即<1,2,1>与<2,1,2>等价），可以使其有序化，不妨设 C[1, 2]=1（若不为 1，可通过简单地调整城市编号使其满足），直接令 $x_1=1, x_2=2$，理论上能加快 9 倍。

时间复杂度分析：

共 n-2 层搜索结点，每个结点最多 3 个分支，故时间复杂度为 O(3^(n-2))。