2022年算法设计与分析期中考试试卷

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、约束条件、优化算法等；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。

一、（10分）求解递推方程，要求给出求解过程。

（1） $\mathrm{T}(n) = 9T\left(\frac{n}{3}\right) + n$，$T(1) = 1$

使用主定理，$a=9, b=3, f(n)=n$，因此

$$T(n) = \Theta(n^{\log_3 9}) = \Theta(n^2)$$

（2） $\mathrm{T}(n) = T(n-1) + \log 3^n$，$T(1) = 1$

迭代归纳.

$$T(n) = n\log 3 + (n-1)\log 3 + \cdots + 2\log 3 + T(1)$$
$$= \log 3[n + (n-1) + \cdots + 2] + 1 = \Theta(n^2)$$

评分标准：每小题5分，

- 如果最终复杂度正确，中间过程正确，得满分；
- 如果最终复杂度正确，中间过程没有或者不正确，每小题扣2-3分。
- 如果最终复杂度不正确，中间过程有一定道理，每小题最多给3分。
- 没有写出Θ或者O符号表示的函数渐近的界，扣1分。

二、（10 分）逆序对计数

给定 1 到n的一个任意排列$x_1, x_2, \ldots, x_n$,请设计一个算法计数排列中逆序对的个数。逆序对的定义是：$x_i > x_j$ 并且$i < j$。

2.28 算法的主要思想是：在二分归并排序算法中附加计数逆序的工作．在递归调用算法分别对子数组 L_1 与 L_2 排序时，分别计数每个子数组内部的逆序；在归并排好序的子数组 L_1 与 L_2 的过程中，附带计数 L_1 的元素与 L_2 的元素之间产生的逆序．假设 L_1 是前半个数组，L_2 是后半个数组．如果 L_1 的最小元素 x 大于 L_2 的最小元素 y，那么算法将从 L_2 中取走 y．这时 L_1 中的每个元素都和 y 构成逆序，所增加的逆序数恰好等于此刻 L_1 中的元素总数．相反，如果 L_1 的最小元素 x 小于 L_2 的最小元素 y，那么算法将从 L_1 中取走 x．这时 L_2 中的每个元素都不会和 x 构成逆序，因此不必改变逆序总数．在算法运行中，每次把这样增加的逆序数加到逆序总数上．

初始逆序数 $N=0$，进入递归算法，算法的主要步骤是：

1. 将数组 L 从中间划分成前后两个子数组 L_1 和 L_2
2. 递归处理 L_1
3. 递归处理 L_2
4. 在归并 L_1 与 L_2 时计数 L_1 与 L_2 的元素产生的逆序数 m，并将 m 加到 N 上

该算法的第 2 步和第 3 步是递归调用，每个子问题的规模都是原问题规模的 1/2，第 4 步每次比较至少拿走 1 个元素，因此元素之间的比较次数与二分归并排序算法的比较次数一样，是 $n-1$ 次．如果以比较运算为基本运算，对于输入规模为 n 的数组所做的比较次数为 $T(n)$，那么有

$$\begin{cases} T(n) = 2T(n/2) + n - 1 \\ T(1) = 0 \end{cases}$$

这就是二分归并排序算法的递推方程，因此得到

$$T(n) = n\log n - n + 1$$

评分标准：

1. 算法设计思想正确，描述清楚+5 分
2. 时间复杂度递推方程描述正确，+3 分
3. T(n)最终求解正确，+2 分。

如果所设计算法正确，但是时间复杂度高于 nlogn，可以酌情给分。

例如，基于所有数字的成对比较，计数逆序对，复杂度 $O(n^2)$，可以得 5 分。

复杂度越高，得分越低。

三、（10 分）找坏硬币

在n枚硬币中有一枚重量过轻的硬币，其余$n-1$枚硬币重量相同。有一个天平可以用来称重，天平没有砝码，但天平两侧称重的硬币数没有限制。请设计一个算法，用较少的称重次数找出这枚过轻的硬币。

参考答案：

三分硬币，秤前两份：如果重量相同，递归处理第三份；否则处理轻的那份。

$T(n) \leqslant 1+T(n/3)=O(\log_3 n)$。

评分标准：

算法 5 分、分析 5 分。$O(\log_2 n)$算法只扣 1 分。

四、（10 分）最低收益最大化

某著名企业有一笔 100 亿的闲置资金，为了保值增值，需要通过投资 3 种基金进行组合投资。咨询机构为其预测了如下表所示的 4 种可能的年收益率（%）。

基金	可能性			
	1	2	3	4
1	$a_{11}=6$	$a_{21}=9$	$a_{31}=30$	$a_{41}=-12$
2	$a_{12}=10$	$a_{22}=3$	$a_{32}=-2$	$a_{42}=6$
3	$a_{13}=27$	$a_{23}=13$	$a_{33}=6$	$a_{43}=-7$

企业需要按保守策略进行投资，要求可能的最低收益率最大。请问该如何确定 3 种基金的投资比例？试建立该问题的数学模型。

标准答案：

变量y表示最低收益率，变量x_1,x_2,x_3分别是三种基金的投资比例或投资额

$$\max y$$

$$\text{st. } a_{11}x_1+a_{12}x_2+a_{13}x_3 \geq y$$

$$a_{21}x_1+a_{22}x_2+a_{23}x_3 \geq y$$

$$a_{31}x_1+a_{32}x_2+a_{33}x_3 \geq y$$

$$a_{41}x_1+a_{42}x_2+a_{43}x_3 \geq y$$

$$x_1+x_2+x_3=1$$

$$x_1,x_2,x_3 \geq 0$$

评分标准

$\max\{\min\{\dots\}\}$ 形式，扣 3 分

$x_1,x_2,x_3 \geq 0$ 缺失，扣 2 分

$x_1+x_2+x_3=1$ 或 $x_1+x_2+x_3=100$ 或 $x_1+x_2+x_3 \leq 100$缺失，扣 5 分

$y \geq 0$ 加了这个条件不加分也不扣分

考察要点：

能否构造出线性规划形式的数学模型。其中用变量$y \leq min\{\dots\}$和$\max y$替代 $\max\{\min\{\dots\}\}$这种非线性规划形式的目标函数是要点

五、（15 分）最小生成树的唯一性

如果一个无向带权连通图$G=(V,E,W)$，任意两个不同的边的权值均不同，即

$$\forall e_i,e_j \in E, e_i \neq e_j \iff W(e_i) \neq W(e_j),$$

则G的最小生成树是唯一的（即无论是用 Prim 算法还是 Kruskal 算法，计算的结果都是一样的）。试证明上述结论。

标准答案：

类似 Kruskal 算法的证明。

首先证明图G最短边e_0一定在任意一个最小生成树T中，否则，将e_0加入T构成回路，用e_0替换回路上的任意其他边，得到权值和更小的生成树T^*，与T是最小生成树矛盾。

然后收缩e_0得到结点数更少的图G'，和对应T的G'的生成树T'，T'是G'的最小生成树，否则G'的最小生成树T''（补回边e_0）还原为G中的生成树T'''的权值和比T小，与T'是最小生成树矛盾。如果T不唯一，则T'也不唯一，与归纳假设结点数小于n时命题成立矛盾。

归纳基础，$n=1$时命题显然成立。

评分标准：

归纳法证明，少归纳基础扣 2 分。

根据推导过程的合理性来评分。

其他证明方法要点：

1、 反证法

假设图G存在两个不同的最小生成树T_1和T_2，则根据所有边长不等，存在权值最小的边$e_1 \in T_1 \otimes T_2$，不失一般性，设$e_1 \in T_1$。T_2+e_1中存在环R，$e_1 \in R$，且存在$e_2 \in (R-T_1)$，使得$w(e_2)>w(e_1)$，$T_2+e_1-e_2$是一个比T_2权值更小的生成树，与T_2是最小生成树矛盾。

「要点：选权值最小的边$e_1 \in T_1 \otimes T_2$在另外一个最小生成树中构成环」

「普遍问题：没有保证边$e_1 \in T_1 \otimes T_2$的权值最小。此时需要分$w(e_2)>w(e_1)$和$w(e_2)<w(e_1)$两种情况讨论，当$w(e_2)<w(e_1)$时，需要把e_2当做新的e_1不断循环讨论下去，最终保证e_1的权值最小，才能保证$T_2+e_1-e_2$是一个比T_2权值更小的生成树」

2、 普遍问题

（即无论是用 Prim 算法还是 Kruskal 算法，计算的结果都是一样的）这段话是辅助解释，它是的最小生成树唯一的必要但不充分的条件。对应证明 Prim 算法的结果和 Kruskal 算法结果一样且唯一的正确证明，我们没扣分，但这样的证明是存在逻辑性错误的。

有些同学的证明中，用到了任意最小生成树辅助证明 Prim 或 Kruskal 算法结果的唯一性，这里潜在的证明了最小生成树结果的唯一性，不算存在逻辑性错误。

另外一些同学的证明中，其实可以把另外一种算法的结果替换为任意的最小生成树，最终就能自然的证明最小生成树的唯一性了。

有些同学只分别证明了 Prim 算法和 Kruskal 算法结果的唯一性，但未说明两者结果的同一性，这在逻辑上无法推导出最小生成树的唯一性。

有些同学采用归纳法推导过程没有问题，但缺少对归纳基础的说明，这样的证明是有缺陷的。

六、（15 分）树的完美匹配

$2n$个顶点的无向图的完美匹配是大小为n的匹配（图的若干条没有公共顶点的边组成图的一个匹配；完美匹配是图的一个匹配，且图的每个顶点恰好在匹配中的一条边上）。请给出多项式时间算法，输入$2n$个顶点的树，输出树中的完美匹配或者判断完美匹配不存在。并证明算法的正确性。

参考答案：

贪心法，依次检查树中度数为 1 的顶点 u，并将覆盖该顶点的一条边(u,v)加入匹配集，删掉顶点 v、且删掉覆盖 u 的其余边。如果某度数为 1 的顶点没有余下的边能覆盖，输出“不存在”；否则处理完所有顶点后，输出覆盖集。复杂度 O(n)。

（对规模归纳）2 个顶点的树，算法能输出正确结果。假设 2n 个顶点的树，算法能输出正确结果。对于规模是 2(n+1)顶点的树 T，第一步去掉了与叶结点 u 相连的边(u,v)，若存在完美匹配，必须包含(u,v)，否则 u 无法被覆盖。可将 T 去掉(u,v)，转化成 2n 顶点的树 T'，继续使用贪心法，得到 T'的完美匹配，或者输出 T'不存在完美匹配。

算法 5 分、分析 5 分、证明 5 分。树 BFS 等线性算法同等分数。

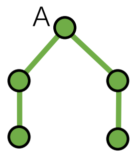
非线性的动规 3+3+3 分。回溯 2 至 3 分。

七、（30 分）广播协议

我们设计了一个用于超算中心数据广播的通讯协议，协议将所有n个超算结点组织成一个树形网络。设根结点有 1GB 的数据需要广播给所有其他结点，树中的每个结点每次只能向其一个直接子结点发送数据，发送 1GB 数据需要用时 1 秒，发送完毕后才能向下一个子结点发送数据。每个结点在收到完整的广播数据后，才能开始向其子结点转发数据。

（1）（20 分）在给定的树形网络上，从根节点向其第一个子结点发送数据开始，每个结点按怎样的顺序向其子结点转发数据，才能让广播过程的总时间（从开始到最后一个结点收到完整广播数据的时间）最短？请尝试用动态规划方法求解该问题，说明算法的正确性，并分析算法的时间复杂度。（提示：在如图所示的树形网络中根节点为 A，其广播过程的总时间最少为 3 秒。）

（2）（10 分）假设$n = 8$，请问如何组织树形网络才能让广播时间最短？画出其树形网络图来说明。



（1）

递推函数：

$Cost(c)$表示 c 结点广播完成的时间长度（标记函数的含义不清不正确扣 2 分）

$Cost(c) = \max\{i + Cost(c_{\pi(i)})\}$，（递推方程不正确，酌情扣 5~8 分）

$c_{\pi(i)}$按子结点的广播时间从大到小排序（没说从大到小的顺序，扣 5 分，没有提交换可以证明最优，扣 2 分）

没有标记函数，因为实际是贪心的问题。递推函数蕴含着贪心的归纳证明。

边界条件：

$$Cost(\text{leaf}) = 0$$

缺边界条件扣 2 分

算法复杂性：

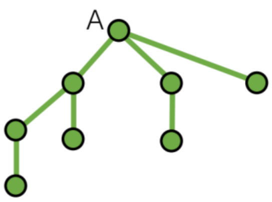
讨论了复杂性得 3 分。

最佳算法复杂度$O(n)$（采用桶分类处理方法或计数排序），

较差的也应该能限制在$O(n^2 \log n)$（每个结点处对子结点做比较排序），

一般应该不超过$O(n \log n)$（每轮对有多个子结点的结点的子结点总体排序）。

(2)



画出正确的图给 10 分。

如果图有错误，酌情扣分。

事实上，把n个结点划分成等分的两组，其中一组（有广播消息）的根结点向

另外一组的根结点连边（发送其第一个广播消息）。如此递归下去，即可构成一个轮次最少的广播树。