

\*作业请在超星平台提交，可以手写拍照，或者提交 word、pdf，截止时间为：10 月 24 日。

\*\*所有题目均需给出时间复杂度分析。

## 第 1 题 交替硬币游戏

将一系列面值不同的硬币排成一排，其中硬币  $i$  的价值为  $v(i)$ 。你和你朋友轮流从这一排硬币的任一端取出一枚硬币。取出所有硬币后，拥有总硬币面值最大的人获胜。假设你们每次都会选择当前最优策略（注意：并非简单取最大），请根据硬币排列决定你应该是先手还是后手。

## 第 2 题 编辑距离

剽窃检测器需要检测两个文本（字符串 A 和字符串 B）之间的相似性。编辑距离是一种相似性度量标准，它是指将字符串 A 转换为字符串 B 的最小编辑次数。编辑可能是以下三种操作之一：删除 A 的一个字符；用另一个字符替换 A 的一个字符；在 A 的两个字符之间插入一个字符。描述一个  $O(|A||B|)$  时间算法来计算 A 和 B 之间的编辑距离。

### 第 3 题 哈夫曼编码

哈夫曼编码是一种前缀码的编码方式。给定  $n$  个字符出现的频率  $(w_1, \dots, w_n)$ , 哈夫曼编码能够使得所有字符平均码长最短, 即  $L = \sum_{i=1}^n w_i \ell_i$  最小, 其中  $\ell_i$  是第  $i$  个字符的编码长度。哈夫曼编码的基本思路为:

- 用每一个字符按照出现的频率  $w_i$  初始化一个队列  $Q$ ;
- 对优先队列  $Q$  用贪心选择: 取出具有最小频率的 2 棵树  $X, Y$ , 并将这 2 棵树合并为新树  $Z$ , 其频率为合并的 2 棵树的频率之和, 并将新树插入优先队列  $Q$ ;
- 作  $n - 1$  次类似的合并。优先队列中只剩下一棵树, 即所要求的树  $T$ 。

按照树  $T$  进行编码, 便得到一个前缀码, 且每个字符编码的长度为该字符对应  $T$  中叶子节点的深度。请使用数学归纳法证明上述编码方式所得的平均码长是最优的。