

1、给定数轴 X 上 n 个不同点的集合 $\{x_1, x_2, \dots, x_n\}$, 其中 $x_1 < x_2 < \dots < x_n$ 。现在用若干个长度为 1 的闭区间来覆盖这些点。设计一个算法找到最少的闭区间个数和位置。

- (1) 请简要叙述贪心算法的策略；(2 分)
- (2) 请证明你的贪心策略能够给出最优解(4 分)；
- (3) 请分析算法的空间和时间复杂度(4 分)。

(1)、如下图，从第一个点开始，每个长度为 1 的区间起始处刚好覆盖前一个区间没有覆盖的下一个点，自该点开始依次往后判断，若当前点在覆盖区间内，则跳过，否则继续添加下一个覆盖区间。直到第 n 个点为止。



(2)、正确性证明（归纳法）

命题：对于 n 个点，算法能得到最优解。

归纳基础：对于只有一个点的情况，该算法显然是最优解。

归纳步骤：

假设对于 n 个点 (x_1, x_2, \dots, x_n) ，算法能得到最优解

考虑 $n+1$ 个点 $(x_1, x_2, \dots, x_n, x_{n+1})$ ，由归纳假设：

对于 n 个点 x_2, x_3, \dots, x_{n+1} ，算法能得到最优解 I'

令 $I = I' \cup \{1\}$

I 是 $n+1$ 个点 $x_1, x_2, \dots, x_n, x_{n+1}$ 的最优解

若不然，存在 $n+1$ 个点的最优解 I^* （必定包括覆盖第 1 个点的区间）

则 $I^* - \{1\}$ 是 n 个点 x_2, x_3, \dots, x_{n+1} 的最优解

$\because I^* < I \quad \therefore I^* - \{1\} < I - \{1\} = I'$

则 n 个点 x_2, x_3, \dots, x_{n+1} 的最优解为 $I^* - \{1\}$ ，与最优解为 I' 矛盾

所以 $I = I' \cup \{1\}$ 是 $n+1$ 个点 $x_1, x_2, \dots, x_n, x_{n+1}$ 的最优解

综上所述，对于 n 个点，算法能得到最优解。

(3)、时空复杂度分析

可以定义一个数组存储这些覆盖区间的位置，一个计数器累加存储覆盖区间的个数，最坏情况下每个点对应一个覆盖区间，空间复杂度为 $O(n)$

从第一个点依次往后判断，直到最后一个点为止，时间复杂度为 $O(n)$

2. 一个公司需要购买 n 个密码软件的许可证，按规定每个月至多可得到一个软件许可证，每个许可证当前售价都是 1000 元，但是第 i 个许可证的售价将按照 $r_i > 1$ 的指数因子增长， $i = 1, 2, \dots, n$ 。例如，第 i 个许可证的售价在 1 个月后将是 $r_i * 1000$ 元，2 月后将是 $r_i^2 * 1000$ 元， k 个月后将是 $r_i^k * 1000$ 元，假设 r_1, r_2, \dots, r_n 是给定正整数，试给出一个购买许可证的顺序，以使得话费的总钱数最少。

1. 令问题的解 I 是 i_1, i_2, \dots, i_n ，其中 i_1, i_2, \dots, i_n 是自然数 1 到 n 的排列。问：按照这个解给出的顺序，当月购买第一个许可证，以后每个月恰好购买一个软件许可证，总的花费是多少元？

2. 设计一个贪心算法求解这个问题，用一段简短的话说明该算法的贪心策略。

3. 证明算法的正确性。

4. 求出算法在最坏情况下的时间复杂度。

(1)、

总的花费 $S = 1000 + r_{i_2} * 1000 + r_{i_3}^2 * 1000 + \dots + r_{i_n}^{n-1} * 1000$

$$= 1000(1 + r_{i_2} + r_{i_3}^2 + \dots + r_{i_n}^{n-1})$$

(2)、

按照贪心算法的思路，首先买 r_i 比较大的软件许可证。

所以，可以先将 r 由大到小排序，再按照这个排列购买。

(3)、正确性证明（反证法）

当只有一个软件许可证时，显然是最优解

当有 n 个软件许可证时 ($n>1$)，将 r 从大到小排序

假设得到的序列为 $i_1 i_2 \dots i_n$ ，对应的花费 S 为最优解

若不然，则在该序列中必存在 r_p 和 r_q ($1 \leq p, q \leq n, p < q, r_p < r_q$)，将其对换，得到对应的

花费应当为 $S' < S$ ，然而根据问题 1 的式子，因为兑换后 r_q 在 r_p 前面， $S' > S$ ，两者相矛盾

所以将 r 从大到小排序后得到的序列 $i_1 i_2 \dots i_n$ 对应的花费 S 为最优解

(4)、时间复杂度

最坏情况下，排序的时间复杂度为 $O(n \log n)$ ，计算的时间复杂度为 $O(n)$

所以最坏情况下的时间复杂度为 $O(n \log n)$ 。