

数据结构与算法第二次作业

Shuxin Chen, 2015013229

2017 年 3 月 3 日

1. 证明：在过程 PERMUTE-BY-SORTING 的数组 P 中，所有元素都唯一的概率至少是 $1 - 1/n$ 。

设事件 X 表示所有元素都唯一，根据基本的排列组合，有

$$Pr(X) = \frac{P_{n^3}^n}{(n^3)^n}$$

通过不等式的放缩和二项式定理，我们有

$$\begin{aligned} Pr(X) &= \frac{P_{n^3}^n}{(n^3)^n} \\ &= \frac{(n^3)(n^3 - 1) \cdots (n^3 - n + 1)}{n^{3n}} \\ &\geq \frac{(n^3 - n + 1)^n}{n^{3n}} \\ &= \left(1 - \frac{n - 1}{n^3}\right)^n \\ &= \sum_{i=0}^n \binom{n}{i} \left(-\frac{n - 1}{n^3}\right)^i \\ &\geq \sum_{i=0}^1 \binom{n}{i} \left(-\frac{n - 1}{n^3}\right)^i \\ &= n\left(-\frac{n - 1}{n^3}\right) + 1 \\ &= 1 - \frac{n - 1}{n^2} \\ &\geq 1 - \frac{1}{n} \end{aligned}$$

证毕。

2. 在单位圆内给定 n 个点, $p_i = (x_i, y_i)$, 对所有 $i = 1, 2, \dots, n$, 有 $0 < x_i^2 + y_i^2 \leq 1$ 。假设所有的点服从均匀分布, 即在单位圆的任一区域内找到给定点的概率与该区域的面积成正比。请设计一个在平均情况下有 $\Theta(n)$ 时间代价的算法, 它能够按照点到原点之间的距离 $d_i = \sqrt{x_i^2 + y_i^2}$ 对这 n 个点进行排序。

令 $d^2 = x^2 + y^2$ 表示 (x, y) 到原点距离的平方。

设 B_1, B_2, \dots, B_n 为 n 个桶, 其中 $B_k = \{(x, y) | \frac{k-1}{n} < d^2 \leq \frac{k}{n}\}$ 。

设 S_1, S_2, \dots, S_n 表示对应的桶在单位圆上表示的面积, 那么当 $k = 1$ 时, 我们有

$$\begin{aligned} S_1 &= \int_{B_1} dS \\ &= \int_{0 < d^2 \leq \frac{1}{n}} dS \\ &= \frac{\pi}{n} \end{aligned}$$

当 $k > 1$ 时, 我们有

$$\begin{aligned} S_k &= \sum_{i=1}^k S_i - \sum_{i=1}^{k-1} S_i \\ &= \int_{0 < d^2 \leq \frac{k}{n}} dS - \int_{0 < d^2 \leq \frac{k-1}{n}} dS \\ &= \frac{k\pi}{n} - \frac{(k-1)\pi}{n} \\ &= \frac{\pi}{n} \end{aligned}$$

因此, $S_1 = S_2 = \dots = S_n$, 即每个桶的大小都是相等且互斥的, 即任意一个点都可以放在有且仅有的一个桶中, 这样桶的设计是合理的。根据桶排序时间复杂度的计算, 这个算法的平均时间复杂度为 $\Theta(n)$ 。