ساختمان داده‌ها (۲۲۸۲۲)

مدرس: حسین بومری

[زمستان ۹۹]

سوال ۵: مرتب سازی خطی

نگارنده: آئیریا محمدی

دایره مورد نظر را به n بخش با مساحت‌های مساوی تقسیم می‌کنیم. مساحت هر بخش برابر با $\pi(۱^۲)/n = \pi/n$ خواهد بود. برای دایره مرکزی خواهیم داشت

$$\pi r^۲ = \pi/n \to r = \sqrt{۱/n} \tag{۱}$$

و شعاع دایره ناحیه دوم برابر خواهد بود با

$$\pi R^۲ - \pi r^۲ = \pi R^۲ - \pi/n = \pi/n \to R = \sqrt{۲/n} \tag{۲}$$

و برای هر ناحیه شماره i داریم $(r = r_{i-۱}, R = r_i)$

$$\pi R^۲ - \pi r^۲ = \pi/n \to R = \sqrt{\frac{۱}{n} + r^۲} \tag{۳}$$

تعریف:

$r_i =$ شعاع دایره i ام

ادعا: شعاع دایره محیطی ناحیه i برابر است با $\sqrt{\frac{i}{n}}$

اثبات با استقرا:

برای پایه استقرا، صحت فرض را به ازای $i = ۲$ در بالا نشان دادیم، و فرض را برای $i - ۱$ دایره اول درست فرض می‌کنیم.

اثبات برای i

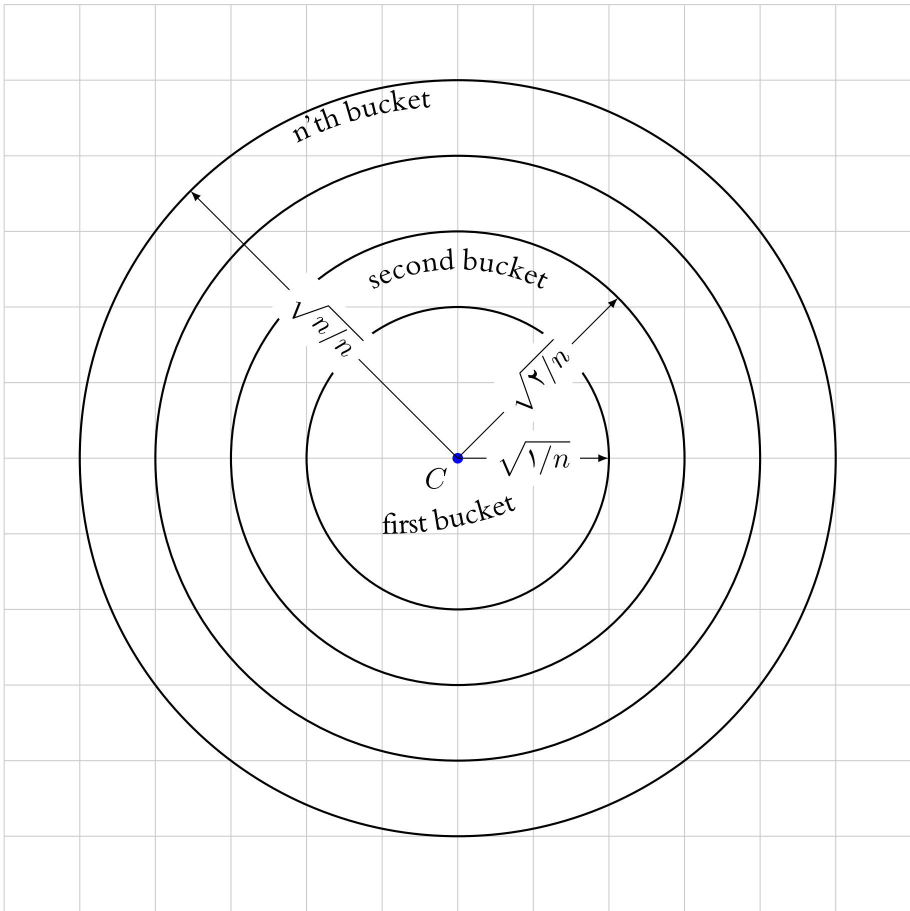
$$R = \sqrt{\frac{۱}{n} + r^۲} = \sqrt{\frac{۱}{n} + (\sqrt{\frac{i-۱}{n}})^۲} = \sqrt{\frac{۱}{n} + \frac{i-۱}{n}} = \sqrt{\frac{i}{n}} \tag{۴}$$

حال مسئله اصلی را حل می‌کنیم.

با توجه به فرض سوال، از آن‌جایی که متوسط تعداد نقاط در هر ناحیه متناسب با مساحت آن ناحیه است اگر دایره را به ناحیه‌های هم مساحت تقسیم کنیم به طور میانگین در هر ناحیه یک نقطه خواهیم داشت.

اگر از الگوریتم bucket sort استفاده کنیم:

$$T(n) = \Theta(n) + \sum_{i=۱}^{n} (n_i)^۲ \tag{۵}$$



شکل ۱: دایره را به n ناحیه با مساحت مساوی تقسیم کردیم

که n_i تعداد نقاط در ناحیه i است.

و اگر امید ریاضی مرتبه زمانی را حساب کنیم

$$E[T(n)] = E[\Theta(n) + \sum_{i=1}^{n} (n_i)^2] \tag{6}$$

بعد دو مرحله استفاده از خاصیت خطی‌بودن امیدریاضی

$$E[T(n)] = E[\Theta(n)] + \sum_{i=1}^{n} E[n_i^2] \tag{7}$$

اگر تعریف کنیم

$$X_{ij} = \begin{cases} 1 & \text{point i being in area j} \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

خواهیم داشت

$$n_j = \sum_{i=0}^{n} X_{ij} \tag{9}$$

پس

$$E[n_i^2] = E[(\sum_{j=0}^{n} X_{ji})^2] = E[\sum_{j=0}^{n} X_{ji}^2] + E[\sum_{j=0}^{n}\sum_{k=0}^{n} 2X_{ji}X_{ki}] \tag{10}$$

برای بخش اول از آن‌جایی که احتمال وجود نقطه در هر ناحیط برابر و برابر با $\frac{1}{n}$ است داریم

$$E[X_{ji}^2] = (1 - \frac{1}{n}) * 0^2 + \frac{1}{n} * 1^2 = \frac{1}{n} \tag{11}$$

$$E[\sum_{j=0}^{n} X_{ji}^2] = \sum_{j=0}^{n} E[X_{ji}^2] = n \cdot \frac{1}{n} = 1 \tag{12}$$

و برای بخش دوم با توجه با استقلال X_{ji} و X_{ki} داریم

$$E[\sum_{j=0}^{n}\sum_{k=0}^{n} 2X_{ji}X_{ki}] = \sum_{j=0}^{n}\sum_{k=0}^{n} 2E[X_{ji}]E[X_{ki}] = \binom{n}{2} \cdot 2 \cdot \frac{1}{n} \cdot \frac{1}{n} = \frac{n(n-1)}{n^2} = 1 - \frac{1}{n} \tag{13}$$

و در نتیجه

$$E[n_i^2] = 2 - \frac{1}{n} \tag{14}$$

در نهایت خواهیم داشت

$$E[T(n)] = E[\Theta(n)] + \sum_{i=1}^{n} 2 - \frac{1}{n} = \Theta(n) + 2n - n(1/n) = \Theta(n) + O(n) \tag{15}$$