Binary Search

Małgorzata Dymek

1 Average case complexiy

$$A(n) = ave(X_n) = \sum_{k \geqslant 0}^{n} p_{k_n} \cdot k$$

The probability of finding our element in the first step is $\frac{1}{n}$, as we search in n elements. However, in the second step it's $\frac{2}{n}$, because we're searching in $\frac{n}{2}$ elements. Similarly for subsequent elements, so the probability of k'th element is $\frac{2^{k-1}}{n}$.

$$\sum_{k \geqslant 0}^{n} p_{k_n} \cdot k = \sum_{k=1}^{log_2 n} \frac{2^{k-1}}{n} \cdot k = \frac{1}{n} \cdot 1 + \frac{2}{n} \cdot 2 + \frac{4}{n} \cdot 3 + \ldots + \frac{2^{log_2 n-1}}{n} \cdot log_2 n$$

We can factor out $\frac{1}{n}$ and contiune:

$$1 \cdot 1 + 2 \cdot 2 + 4 \cdot 3 + \ldots + 2^{log_2 n-1} \cdot log_2 n = 2^0 \cdot 1 + 2^1 \cdot 2 + 2^2 \cdot 3 + \ldots + 2^{log_2 n-1} \cdot log_2$$

Let's take geometric series:

$$1 + x + x^2 + ... + x^m = \frac{x^{m+1} - 1}{x - 1} \cdot 1$$

And calculate the derivative of both sides of the equation:

$$0 + 1 + 2 \cdot x + 3 \cdot x^2 + ... + m \cdot x^{m-1} = \frac{(m+1) \cdot x^m \cdot (x-1) - (x^{m+1} - 1)}{(x-1)^2}$$

The result of the left side is similar to our sum.

$$\frac{(m+1) \cdot x^m \cdot (x-1) - (x^{m+1} - 1)}{(x-1)^2} = \frac{(m+1) \cdot x^{m+1} - (m+1)x^m - x^{m+1} + 1}{(x-1)^2} = \frac{m \cdot x^{m+1} - (m+1) \cdot x^m + 1}{(x-1)^2}$$

Let's do a replacement: $x = 2$, $m = log_2 n$.

$$\frac{log_2 \cdot 2^{log_2 n+1} - (log_2 n + 1) \cdot 2^{log_2 n} + 1}{(2-1)^2} = \frac{log_2 n \cdot 2^{log_2 n+1} - (log_2 n) \cdot 2^{log_2} - 2^{log_2 n} + 1}{1} =$$
$$= log_2 n \cdot 2^{log_2 n} - 2^{log_2 n} + 1 = 2^{log_2 n} \cdot (log_2 n - 1) + 1 = n \cdot (log_2 n - 1) + 1$$

Now we can contiune with our sum.

$$\sum_{k=1}^{log_2 n} \frac{2^{k-1}}{n} \cdot k = \frac{1}{n} \cdot (n \cdot (log_2 n - 1) + 1) = \frac{n \cdot (log_2 n - 1) + 1}{n} = log_2 n - 1 + \frac{1}{n}$$

So $log_2 n - 1 + \frac{1}{n}$ is the average case complexity, class $O(log_2 n)$.

2 Variancy of the average case complexity

$$var(X_n) = \sum_{k \geqslant 0}^{n} (k - ave(X_n))^2 \cdot p_{k_n}$$

We'll round our result to $log_2 n$ to simplify calculations.

$$\sum_{k \geqslant 0}^{n} (k - ave(X_n))^2 \cdot p_{k_n} = \sum_{k=1}^{log_2 n} (k - log_2 n)^2 \cdot \frac{2^{k-1}}{n} = \sum_{k=1}^{log_2 n} (k^2 + (log_2 n)^2 - 2 \cdot k \cdot log_2 n) \cdot \frac{2^{k-1}}{n}$$
$$= \frac{1}{n}(\sum_{k=1}^{log_2 n} k^2 \cdot 2^{k-1} + \sum_{k=1}^{log_2 n} (log_2 n)^2 \cdot 2^{k-1} - \sum_{k=1}^{log_2 n} 2^k \cdot k \cdot log_2 n)$$

$$\sum_{k=1}^{log_2 n} k^2 \cdot 2^{k-1} = n \cdot (log_2 n)^2 - 2 \cdot n \cdot log_2 n + 3 \cdot n - 3$$

$$(log_2 n)^2 \cdot \sum_{k=1}^{log_2 n} 2^{k-1} = n \cdot (log_2 n)^2 - \frac{1}{2} \cdot (log_2 n)^2$$

$$2 \cdot log_2 n \cdot \sum_{k=1}^{log_2 n} 2^k \cdot k = 4 \cdot n \cdot (log_2 n)^2 - 4 \cdot n \cdot log_2 n + 4 \cdot log_2 n$$

$$\sum_{k=1}^{log_2 n} (k - log_2 n)^2 \cdot \frac{2^{k-1}}{n} = \frac{1}{n} \cdot (n \cdot (log_2 n)^2 - 2 \cdot n \cdot log_2 n + 3 \cdot n - 3 + n \cdot (log_2 n)^2 - \frac{1}{2} \cdot (log_2 n)^2 -$$
$$- 4 \cdot n \cdot (log_2 n)^2 + 4 \cdot n \cdot log_2 n - 4 \cdot log_2 n) = \frac{1}{n}(-2 \cdot n \cdot (log_2 n)^2 + 2 \cdot n \cdot log_2 n - \frac{1}{2} \cdot (log_2 n)^2 - n \cdot log_2 n + 3 \cdot n - 3) =$$
$$= 2 \cdot log_2 n + 3 - (2 \cdot (log_2)^2 + \frac{1}{2n}(log_2 n)^2 - \frac{4}{n} \cdot log_2 n - \frac{3}{n})$$

So we received a result of class $O(log_2 n)$.