

1/22 Binary Search

Tuesday, January 23, 2018 9:54 PM

This will show an example of computing the average case runtime of an algorithm.

Consider the following,

Given a sorted array of $n = 2^k - 1$ values (for some positive integer, k). Find the index of a particular value x through binary search.

As an example let the array be

8	9	21	34	52	78	99
---	---	----	----	----	----	----

Thus $n = 7$ and $k = 3$

We will assume that each queried value will be in the array and each value in the array has an equally likely chance of being selected.

For our problem we will assume that the time taken for a query will be of the order of the number of recursive calls to the search function.

In our example

$$\begin{aligned}t(8) &= 3 \\ t(9) &= 2 \\ t(21) &= 3 \\ t(34) &= 1 \\ t(52) &= 3 \\ t(78) &= 2 \\ t(99) &= 3\end{aligned}$$

In general the following holds,

- one element of the array can be found in one recursive call
- two elements can be found in two recursive calls
- four elements can be found in three recursive calls
- 2^i elements can be found in $i+1$ recursive calls.

$$E(t(x)) = \frac{1 \cdot 1}{2^k - 1} + \frac{2 \cdot 2}{2^k - 1} + \frac{3 \cdot 4}{2^k - 1} + \dots \\ + \frac{i \cdot 2^{i-1}}{2^k - 1} + \dots + \frac{k \cdot 2^{k-1}}{2^k - 1}$$

$2^k - 1$ in the denominator is from the equally likely chance of a value being chosen.

$$\dots \frac{1}{2^k - 1} (1 \cdot 1 + 2 \cdot 2 + 3 \cdot 4 + \dots + k \cdot 2^{k-1})$$

Choose:

$$E(f(x)) = \frac{1}{2^k - 1} (1 \cdot 1 + 2 \cdot 2 + 3 \cdot 4 + \dots + k \cdot 2^{k-1})$$

$$\text{Let } S = E(f(x)) \cdot \frac{1}{2^k - 1}$$

then

$$S = 1 \cdot 1 + 2 \cdot 2 + 3 \cdot 4 + \dots + k \cdot 2^{k-1}$$

$$2S = \quad + 1 \cdot 2 + 2 \cdot 4 + \dots + (k-1)2^{k-1} + k2^k$$

$$S - 2S = 1 \cdot 1 + 1 \cdot 2 + 1 \cdot 4 + \dots + 1 \cdot 2^{k-1} + (-k \cdot 2^k)$$

$$S = k \cdot 2^k - (1 + 2 + 4 + \dots + 2^{k-1})$$

$$= k \cdot 2^k - (2^k - 1)$$

$$= (k-1)(2^k) + 1$$

So

$$E(f(x)) = \frac{(k-1)(2^k) + 1}{2^k - 1}$$

$$= 2(k-1) + \frac{1}{2^{k-1}}$$

$\approx 2k$ as k gets large

.....

Since $k \approx \lg(n)$

the expected time is about $\lg(n)$
which is also the worst case runtime