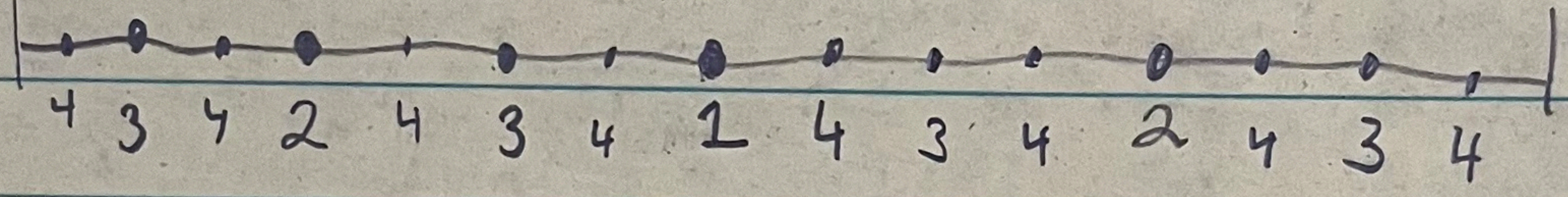
Average Cost of Binary Search

- Assume $n = 2^k = 2^{\log_2 n}$
- Maximum number of comparisons: $k + 1 = \log(n) + 1$

- Items found with 1 comparison: $1 = 2^0$
- Items found with 2 comparisons: $2 = 2^1$
- Items found with 3 comparisons: 4 items $= 2^2$
- Items found with 4 comparisons: 8 items $= 2^3$



- Items found with $\log(n) = k$ comparisons $= 2^{k-1} = \frac{n}{2}$

- Items found with $\leq k$ comparisons is

$$1 + 2 + 4 + 8 + \dots + 2^{k-1} = \sum_{j=0}^{k-1} 2^j = \frac{1 - 2^k}{1 - 2} = \frac{1 - 2^k}{-1} = 2^k - 1 = n - 1$$

is this right?

Try $n = 4$ so $k = 2$ max comparisons is 3

$[a_1, a_2, a_3, a_4]$
↑ ↑ ↑ ↑
2 1 2 3

1 comparison: a_2
2 comps: a_1, a_3 ⇒
3 comps: a_4

≤ 2 comparisons $\rightarrow 3 = n-1$
3 comparisons $\rightarrow 1$

$n = 8$ $k = 3$, max $= 4$

$[a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8]$
↑ ↑ ↑ ↑ ↑ ↑ ↑ ↑
3 2 3 1 3 2 3 4

$7 = n - 1$ items ≤ 3 comparisons
1 item with 4 comparisons

N = # of comparisons

$P(N=1) = \frac{1}{n} = \frac{1}{2^K}$ $\quad P(N=2) = \frac{2}{n} = \frac{1}{2^{K-1}}$ $\quad P(N=3) = \frac{4}{n} = \frac{1}{2^{K-2}}$

$\cdots P(N=K) = \frac{n/2}{n} = \frac{1}{2} = \frac{1}{2^1}$ $\quad P(N=K+1) = \frac{1}{2^K}$

so for $1 \leq j \leq K$ $\quad P(N=j) = \frac{1}{2^{K-(j-1)}} = \frac{1}{2^{K-j+1}} = \frac{1}{2 \cdot 2^{K-j}} = \frac{2^j}{2^{K+1}}$

and $P(N=K+1) = \frac{1}{2^K}$

$$E[N] = \sum_{j=1}^{K+1} j \cdot P(N=j) = \frac{K+1}{2^K} + \sum_{j=1}^{K} \frac{j \cdot 2^j}{2^{K+1}} = \frac{K+1}{2^K} + \frac{1}{2^{K+1}} \sum_{j=1}^{K} j \cdot 2^j$$

$$= \frac{1}{2^K}\left[K+1 + \frac{1}{2}\sum_{j=1}^{K} j \cdot 2^j\right] = \frac{1}{2^K}\left[K+1 + \sum_{j=1}^{K} j \cdot 2^{j-1}\right]$$

Let $f(x) = \sum_{j=0}^{K} x^j = 1 + x + x^2 + \cdots + x^K = \frac{1-x^{K+1}}{1-x}$ $\quad x \neq 1$

$$f'(x) = 1 + 2x + 3x^2 + \cdots + Kx^{K-1} = \sum_{j=1}^{K} j \cdot x^{j-1} = \frac{d}{dx}\left[\frac{1-x^{K+1}}{1-x}\right]$$

So $E[N] = \frac{1}{2^K}\left[K+1 + f'(2)\right]$

$$f'(x) = \frac{(1-x)\cdot(-1)(K+1)x^K - (1-x^{K+1})\cdot(-1)}{(1-x)^2} = \frac{(x-1)(K+1)x^K + 1 - x^{K+1}}{(1-x)^2}$$

$f'(2) = (K+1)2^K + 1 - 2^{K+1}$

$$E[N] = \frac{1}{2^K}\left[K+1 + (K+1)2^K + 1 - 2 \cdot 2^K\right]$$

$K = \log(n) \quad 2^K = n$

$$E[N] = \frac{1}{n}\left[\log(n) + n\cdot\log(n) - 2\cdot n + 2\right]$$

$$= \frac{\log(n)}{n} + \log(n) - 2 + \frac{2}{n}$$

$$\lim_{n\to\infty} \frac{E[N]}{\log(n)} = \lim_{n\to\infty}\left[\frac{1}{n} + 1 - \frac{2}{\log(n)} + \frac{2}{n\cdot\log(n)}\right] = 1$$

So $E[N] \sim \log(n)$

Can also show $E[N] = \Theta(\log n)$

$\forall \varepsilon > 0 \ \exists\ M = m(\varepsilon)$ s.t. $n \geq m(\varepsilon) \Rightarrow \left|\frac{E[N]}{\log(n)} - 1\right| < \varepsilon$

Take $\varepsilon = \frac{1}{2}$ then for $n \geq M$ $\quad -\frac{1}{2} < \frac{E[N]}{\log(n)} - 1 < \frac{1}{2}$

$$\Rightarrow \frac{1}{2} < \frac{E[N]}{\log(n)} < \frac{3}{2}$$

$E[N] = \Omega(\log(n))$ $\qquad$ $E[N] = O(\log(n))$

$$E[N] = \Theta(\log(n))$$

Corollary: if $f(n) \sim g(n)$ then $f(n) = \Theta(g(n))$

(but $f(n) = \Theta(g(n)) \not\Rightarrow f(n) \sim g(n)$)