9.5 Efficiency of Algorithms II.

Binary Search.

$a[1], \ldots, a[n]$ is a sequence of distinct numbers arranged in ascending order.

Given a number x, we want to find if it appears in the sequence, and, if so, where is it.

```
index := 0,   top := n,   bot := 1.
while ( top ≥ bot  and  index = 0 )
        mid := ⌊(bot + top)/2⌋
        if a[mid] = x        then index := mid
        if a[mid] > x,       then top := mid - 1
                             else bot := mid + 1
end while
Output : index.
```

$a[bot] \ldots, a[mid-1] \quad a[mid] \quad a[mid+1] \ldots a[top]$

$x < a[mid]$ look here

$x = a[mid]$, $j = mid$, stop

$x > a[mid]$ look here

x

W_n = # of while loop in a worst case search in $a[1], \ldots, a[n]$.

1 loop : comparing x & $a\lfloor \frac{n+1}{2} \rfloor$.

then search x in $a[1], \ldots, a[\lfloor \frac{n+1}{2} \rfloor - 1]$ (L) $(x < a\lfloor \frac{n+1}{2} \rfloor)$

or $a[\lfloor \frac{n+1}{2} \rfloor + 1], \ldots, a[n]$ (R) $(x > a\lfloor \frac{n+1}{2} \rfloor)$

L has $\lfloor \frac{n+1}{2} \rfloor - 1 = \lfloor \frac{n-1}{2} \rfloor$ elements

R has $n - \lfloor \frac{n+1}{2} \rfloor = -(-n + \lfloor \frac{n+1}{2} \rfloor) = -\lfloor -\frac{n-1}{2} \rfloor = \lceil \frac{n-1}{2} \rceil = \lfloor \frac{n}{2} \rfloor$

R has more elements $\Rightarrow$ R is worse.

$\Rightarrow$ Worst case need another $W_{\lfloor \frac{n}{2} \rfloor}$ while loops.

$$\Rightarrow \begin{cases} W_n = 1 + W_{\lfloor \frac{n}{2} \rfloor}, & n \geq 2 \\ W_1 = 1 & \text{(obvious)} \end{cases}$$

Solve $\begin{cases} W_n = 1 + W_{\lfloor \frac{n}{2} \rfloor}, & n \geq 2 \\ W_1 = 1. \end{cases}$

$W_1 = 1$, $W_2 = W_3 = 2$, $W_4 = W_5 = W_6 = W_7 = 3$.

Conjecture: (1) $W_n = \lfloor \log_2 n \rfloor + 1 \quad n \geq 1$.

$\Leftrightarrow$ (1)' $W_n = k+1$ for $2^k \leq n < 2^{k+1}$.

Prove (1)' by induction on k.

(i) $k=0$, $2^0 \leq n < 2^1 \Rightarrow n = 1$.

$W_1 = 1 = 0+1$. So (1)' is true for $k=0$.

(ii) Assume $W_n = k+1$ for $2^k \leq n < 2^{k+1}$

For any m s.t. $2^{k+1} \leq m < 2^{k+2}$, We have $2^k \leq \lfloor \frac{m}{2} \rfloor < 2^{k+1}$.

$\Rightarrow W_{\lfloor \frac{m}{2} \rfloor} = k \Rightarrow W_m = 1 + W_{\lfloor \frac{m}{2} \rfloor} = 1 + k$.

So (1)' is true for $k+1$.

(i) & (ii) $\Rightarrow$ (1)' is true for $\forall k \geq 0$

$\Rightarrow$ (1) $W_n = \lfloor \log_2 n \rfloor + 1$, $n \geq 1$.

Clearly, $W_n = \lfloor \log_2 n \rfloor + 1 = \Theta(\log_2 n)$.

Merging sort:

Let $a[1], \ldots, a[l]$, and $b[1], \ldots, b[m]$ be two ascending sequences.

We want to merge these two sequences into an ascending sequence $c[1], \ldots, c[l+m]$.

$a[l]$ $\quad\quad$ $b[m]$

$\vdots$ $\quad\quad$ $\vdots$

$a[1]$ $\quad\quad$ $b[1]$

Every time, move the smaller of the two elements on the bottom to the first available spot in the sequence c. (If one column is exausted, then just move the elements of the other column to the available spots in c).

We only need $\leq l+m-1$ comparisons to merge a & b.

Recursive programming.

F_1. Input $a[r]$

Output $a[r]$

F_n: Input $a[r], \cdots, a[r+n-1]$,

mid $= r + \lfloor \frac{n}{2} \rfloor - 1$

$F_{\lfloor \frac{n}{2} \rfloor}$ $\{ a[r], \cdots, a[\text{mid}] \}$ (i)

(output $a[r], \cdots, a[\text{mid}]$)

$F_{\lceil \frac{n}{2} \rceil}$ $\{ a[\text{mid}+1], \cdots, a[r+n-1] \}$ (ii)

(output $a[\text{mid}+1], \cdots, a[r+n-1]$)

merging sort $\{ a[r], \cdots, a[\text{mid}] \}$ & $\{ a[\text{mid}+1], \cdots, a[r+n-1] \}$ (iii)

output $a[r], \cdots, a[r+n-1]$.

$m_n :=$ maximal # of comparisons used in F_n.

$$\begin{cases} m_n = \underset{\text{(iii)}}{n-1} + \underset{\text{(i)}}{m_{\lfloor \frac{n}{2} \rfloor}} + \underset{\text{(ii)}}{m_{\lceil \frac{n}{2} \rceil}}, & n \geq 2. \\ m_1 = 0 \end{cases}$$

THM. $\{m_n\}_{n=1}^{\infty}$ satisfies $\begin{cases} m_n = (n-1) + m_{\lfloor \frac{n}{2} \rfloor} + m_{\lceil \frac{n}{2} \rceil}, & n \geq 2 \\ m_1 = 0. \end{cases}$

Then $\frac{1}{2} n \log_2 n \leq m_n \leq n \log_2 n$, $\forall n \geq 1$.

(The Text claims $n\log_2 n \le m_n$, which is wrong.
when $n=2$, l.h.s. $= 2\log_2 2 = 2$, r.h.s. $= m_2 = 1$)

Proof. (A) $\frac{1}{2} n\log_2 n \le m_n$ (1)

Consider the function $f(x) = \frac{1}{2} x\log_2 x = \frac{1}{2\ln 2}(x\ln x)$.

$f'(x) = \frac{1}{2\ln 2}(\ln x + 1)$, $f''(x) = \frac{1}{2\ln 2}\cdot\frac{1}{x} > 0 \quad \forall x>0$.

So $f(x)$ is convex on $\mathbb{R}^+$, which means for any $x_1, x_2 \in \mathbb{R}^+$, we have $\frac{f(x_1)+f(x_2)}{2} \ge f\left(\frac{x_1+x_2}{2}\right)$.

Now induct on n to prove $\frac{1}{2} n\log_2 n \le m_n$.

(i) $n=1$. $m_1 = 0$. $\frac{1}{2}\,1\cdot\log_2 1 = 0$.

$\Rightarrow$ (1) is true for $n=1$.

(ii) Assume $\frac{1}{2} n\log_2 n \le m_n$ for $\forall n\le k$, where $k\ge 1$

Consider m_{k+1}.

$$m_{k+1} = k + m_{\lfloor\frac{k+1}{2}\rfloor} + m_{\lceil\frac{k+1}{2}\rceil} \ge k + \frac{1}{2}\left\lfloor\frac{k+1}{2}\right\rfloor\log_2\left\lfloor\frac{k+1}{2}\right\rfloor + \frac{1}{2}\left\lceil\frac{k+1}{2}\right\rceil\log_2\left\lceil\frac{k+1}{2}\right\rceil$$

$$= k + f\left(\left\lfloor\tfrac{k+1}{2}\right\rfloor\right) + f\left(\left\lceil\tfrac{k+1}{2}\right\rceil\right) \ge k + 2f\left(\frac{\lfloor\frac{k+1}{2}\rfloor + \lceil\frac{k+1}{2}\rceil}{2}\right)$$

(Note $\lceil\frac{k+1}{2}\rceil + \lfloor\frac{k+1}{2}\rfloor = k+1$) $= k + 2f\left(\frac{k+1}{2}\right) = k + 2\cdot\frac{1}{2}\cdot\frac{k+1}{2}\cdot\log_2\left(\frac{k+1}{2}\right)$

$$= k + \frac{1}{2}(k+1)(\log_2(k+1) - 1)$$

$$= \frac{1}{2}(k+1)\log_2(k+1) + \left(k - \frac{k+1}{2}\right)^{\ge 0}$$

$$\ge \frac{1}{2}(k+1)\log_2(k+1).$$

(i) & (ii) $\Rightarrow$ (1) is true for $\forall n\ge 1$.

(B) $m_n \le n\log_2 n$ (2).

We prove that " If $1\le n\le 2^{k+1}$, then $m_n \le kn+1$ (3)."

Induction on k. (i) $k=0$, $1\le n\le 2^{0+1} = 2$. $\Rightarrow n=1,2$.

$m_1 = 0 < 0\times 1+1$, $m_2 = 1 = 0\cdot 2+1$.

(ii) Assume $m_n \le kn+1$ for $\forall n\le 2^{k+1}$, for some $k\ge 0$.

Let N be any number with $N \le 2^{k+2}$. Then $\lfloor \frac{N}{2} \rfloor, \lceil \frac{N}{2} \rceil \le 2^{k+1}$, $\Rightarrow$

$$m_N = N-1 + m_{\lfloor \frac{N}{2} \rfloor} + m_{\lceil \frac{N}{2} \rceil} \le N-1 + k\lfloor \tfrac{N}{2} \rfloor + 1 + k\lceil \tfrac{N}{2} \rceil + 1$$
$$= N + k\left(\lfloor \tfrac{N}{2} \rfloor + \lceil \tfrac{N}{2} \rceil\right) + 1 = N + kN + 1$$
$$= (k+1)N + 1.$$

So (3) is true for $k+1$ too.

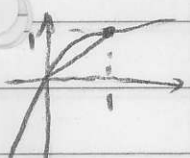
(i) & (ii) $\Rightarrow$ For $\forall k \ge 0$, if $1 \le n \le 2^{k+1}$, then $m_n \le kn+1$.

For any $n \ge 2$ $\exists!\ k \ge 0$, $1 < r \le 2^k$, s.t.,

$n = 2^k + r \le 2^{k+1}$. Then $m_n \le kn + 1$.

So
$$n\log_2 n - m_n \ge n\log_2 n - kn - 1$$
$$= n(\log_2 n - k) - 1 = n\log_2\left(\frac{n}{2^k}\right) - 1$$
$$= n\log_2\left(1 + \frac{r}{2^k}\right) - 1.$$

Let $g(x) = \log_2(1+x) = \ln(1+x) / \ln 2$ then $g''(x) = \frac{-1}{(\ln 2)(1+x)^2} < 0$.



So $g(x)$ is concave. $\Rightarrow$ for any $x \in [0,1]$,

$g(x) \ge x g(1) + (1-x) g(0) = x g(1) = x \log_2(1+1) = x$.

Thus
$$n\log_2 n - m_n \ge n\log_2\left(1 + \frac{r}{2^k}\right) - 1 \ge n \cdot \frac{r}{2^k} - 1$$
$$> r - 1 \ge 0.$$

$\Rightarrow$ $m_n < n\log_2 n$ $\forall\ n \ge 2$.

(Clearly, we have $m_1 = 0 = 1 \cdot \log_2 1$.)

Thus, $\frac{1}{2} n\log_2 n \le m_n \le n\log_2 n$ $\forall n \ge 1$.

$\Rightarrow$ $m_n = \Theta(n\log_2 n)$.