

Binary Search Tree

- nothing (NULL)
- BST

If we know how often we look for each item
what is the best BST?

Keys = 1, 2, ..., n

Frequencies $f[1..n] \leftarrow \text{INPUT}$

$f[i] = \# \text{ times we search for key } i.$

$$\text{Cost}(T, f) = \sum_{i=1}^n f[i] \cdot \text{depth}(i, T)$$

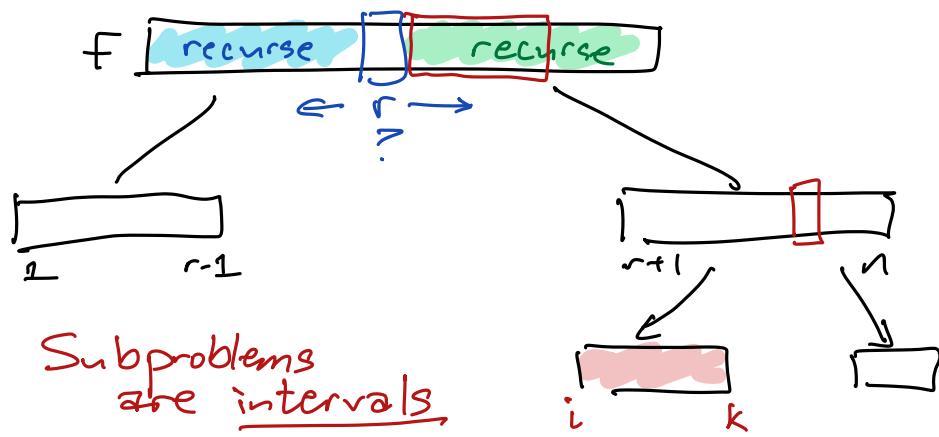
$$= \sum_i f[i] \cdot \#\text{ancestors of } i \text{ in } T$$

$$= \sum_i f[i] \cdot \begin{cases} 1 & (\#\text{ancestors of } f_i \text{ in left}(T)) \\ & + (\#\text{anc of } i \text{ in right}(T)) \end{cases}$$

$$= \sum_{i=1}^{r-1} f[i] \cdot \#\text{anc in left} + \sum_{i=1}^r f[i] + \sum_{i=r+1}^n f[i] \cdot \#\text{anc in right}$$

$$\begin{aligned} \text{Cost}(T, f[1..n]) &= \sum_{i=1}^n f[i] + \text{Cost}(\text{left}(T), f[1..r-1]) \\ &\quad + \text{Cost}(\text{right}(T), f[r+1..n]) \end{aligned}$$

$$r = \text{root}(T)$$



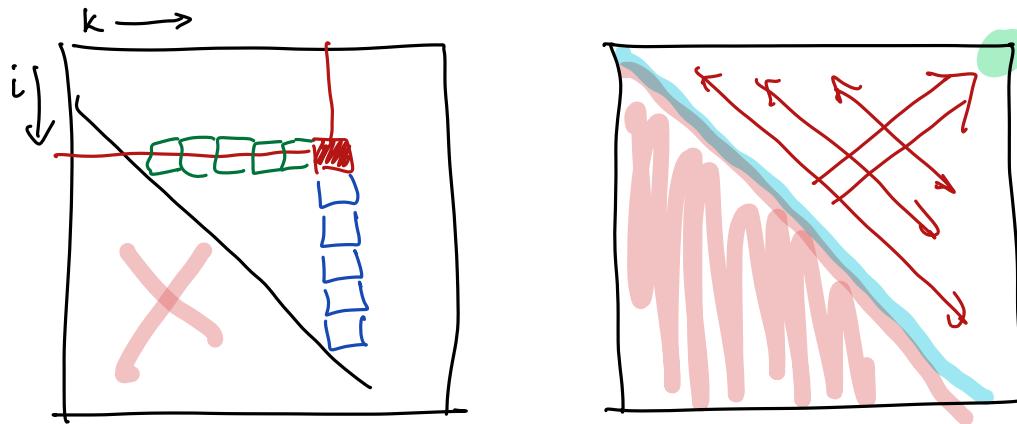
$\text{OptCost}[i..k]$ = total cost of opt. BST for frequencies $f[i..k]$

$$\text{OptCost}(i, k) = \begin{cases} 0 & \text{if } i > k \\ \sum_{j=i}^k f[j] + \min_{i \leq r \leq k} \left\{ \underbrace{\text{OptCost}(i, r-1)}_{\substack{\text{left subtree} \\ \#times we touch the root}} + \underbrace{\text{OptCost}(r+1, k)}_{\substack{\text{right subtree}}} \right\} & \text{otherwise} \end{cases}$$

$F[i..k] = \sum_{j=i}^k f[j] = \begin{cases} 0 & \text{if } i > k \\ F[i..k-1] + f(k) & \text{o/w} \end{cases}$

<pre> INITF($f[1..n]$): for $i \leftarrow 1$ to n $F[i, i-1] \leftarrow 0$ for $k \leftarrow i$ to n $F[i, k] \leftarrow F[i, k-1] + f[k]$ </pre>

$$\text{OptCost}(i, k) = \begin{cases} 0 & \text{if } i > k \\ F[i, k] + \min_{i \leq r \leq k} \left\{ \text{OptCost}(i, r-1) + \text{OptCost}(r+1, k) \right\} & \text{otherwise} \end{cases}$$



COMPUTEOPTCOST(i, k):

```

 $OptCost[i, k] \leftarrow \infty$ 
for  $r \leftarrow i$  to  $k$ 
     $tmp \leftarrow OptCost[i, r - 1] + OptCost[r + 1, k]$ 
    if  $OptCost[i, k] > tmp$ 
         $OptCost[i, k] \leftarrow tmp$ 
 $OptCost[i, k] \leftarrow OptCost[i, k] + F[i, k]$ 

```

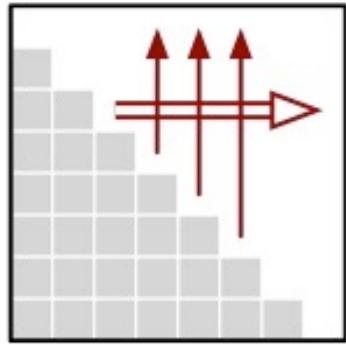
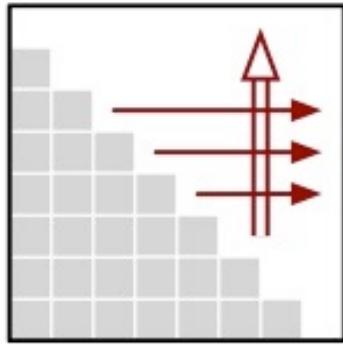
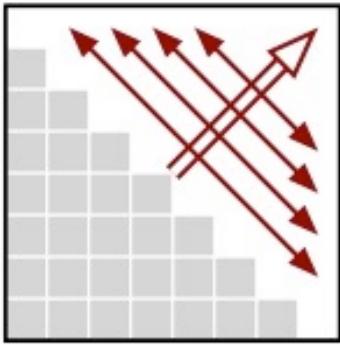
OPTIMALBST($f[1..n]$):

```

INITF( $f[1..n]$ )  $\leftarrow O(n^2)$  time
for  $i \leftarrow 1$  to  $n + 1$ 
     $OptCost[i, i - 1] \leftarrow 0$ 
for  $d \leftarrow 0$  to  $n - 1$ 
    for  $i \leftarrow 1$  to  $n - d$       «... or whatever »
        COMPUTEOPTCOST( $i, i + d$ )  $\leftarrow O(n)$  time
return  $OptCost[1, n]$ 

```

$O(n^3)$ time



```
OPTIMALBST( $f[1..n]$ ):
    INITF( $f[1..n]$ )
    for  $i \leftarrow 1$  to  $n + 1$ 
         $OptCost[i, i - 1] \leftarrow 0$ 
    for  $d \leftarrow 0$  to  $n - 1$ 
        for  $i \leftarrow 1$  to  $n - d$       {... or whatever}
            COMPUTE $OptCost(i, i + d)$ 
    return  $OptCost[1, n]$ 
```

```
OPTIMALBST2( $f[1..n]$ ):
    INITF( $f[1..n]$ )
    for  $i \leftarrow n + 1$  downto 1
         $OptCost[i, i - 1] \leftarrow 0$ 
    for  $j \leftarrow i$  to  $n$ 
        COMPUTE $OptCost(i, j)$ 
    return  $OptCost[1, n]$ 
```

```
OPTIMALBST3( $f[1..n]$ ):
    INITF( $f[1..n]$ )
    for  $j \leftarrow 0$  to  $n + 1$ 
         $OptCost[j + 1, j] \leftarrow 0$ 
    for  $i \leftarrow j$  downto 1
        COMPUTE $OptCost(i, j)$ 
    return  $OptCost[1, n]$ 
```

$$OptCost(i, k) = \begin{cases} 0 & \text{if } i > k \\ F[i, k] + \min_{1 \leq r \leq k} \left\{ OptCost(i, r - 1) + OptCost(r + 1, k) \right\} & \text{otherwise} \end{cases}$$

↑

2 input params $\Rightarrow O(n^2)$ space

3 vars on right $\Rightarrow O(n^3)$ time