

BST RUNNING TIME ANALYSIS

Problem Solving with Computers-II

C++

```
#include <iostream>
using namespace std;

int main(){
    cout<<"Hola Facebook!";
    return 0;
}
```

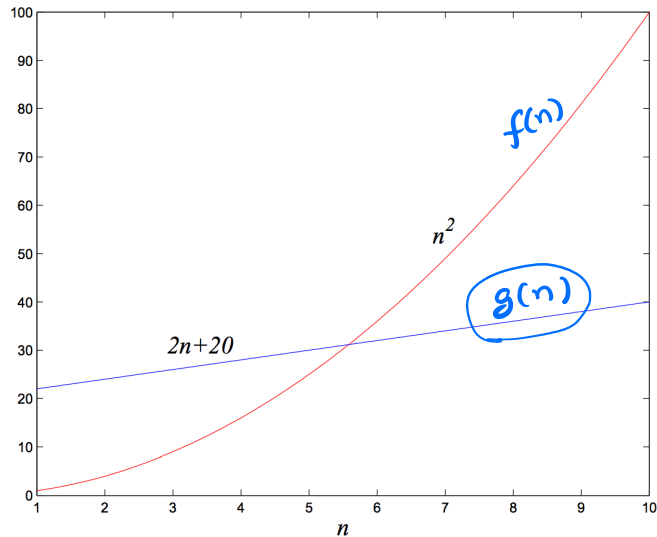
Big-Omega

- $f(n)$ and $g(n)$ map positive integer inputs to positive reals.

We say $f = \Omega(g)$ if there are constants $c > 0, k > 0$ such that $c \cdot g(n) \leq f(n)$ for $n \geq k$

$$f = \Omega(g)$$

means that “f grows at least as fast as g”

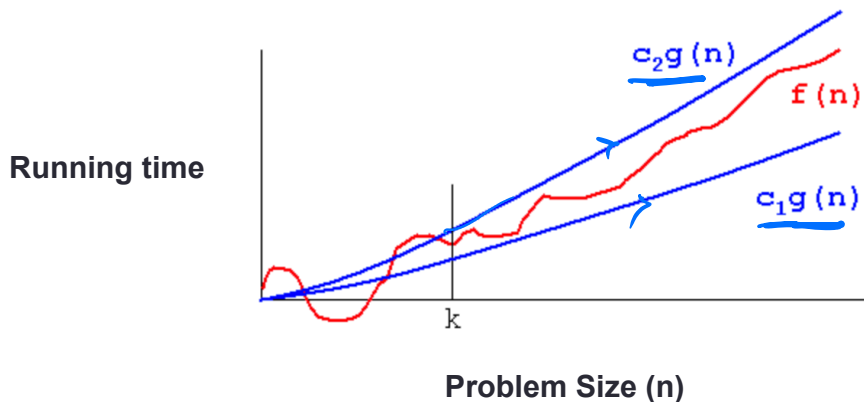


Big-Theta





- $f(n)$ and $g(n)$ map positive integer inputs to positive reals.

We say $f = \Theta(g)$ if there are constants c_1, c_2, k such that

$$0 \leq c_1 g(n) \leq f(n) \leq c_2 g(n), \text{ for } n \geq k$$



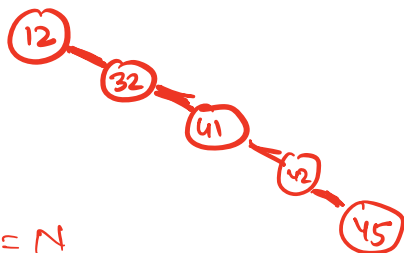
Binary Search Trees

- WHAT are the operations supported?

- HOW do we implement them?

- WHAT are the (worst case) running times of each operation?
 



- Path – a sequence of nodes and edges connecting a node with another node.
- A path starts from a node and ends at another node or a leaf
- Height of node – The height of a node is the number of edges on the longest downward path between that node and a leaf.

1.

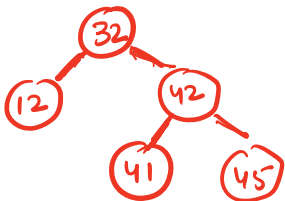


$$H = N$$

path: $12 \rightarrow 32 \rightarrow 41 \rightarrow 42$

Height: 4

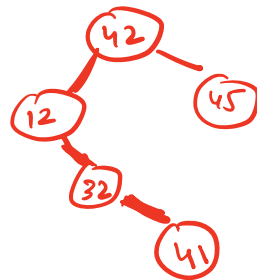
2.



$32 \rightarrow 42 \rightarrow 41$

Height: 2

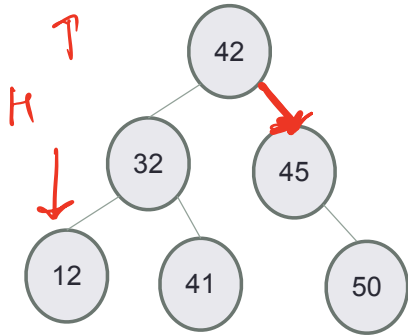
3.



Height: 3

BSTs of different heights are possible with the same set of keys
 Examples for keys: 12, 32, 41, 42, 45

Worst case Big-O of search, insert, min, max : $O(H)$



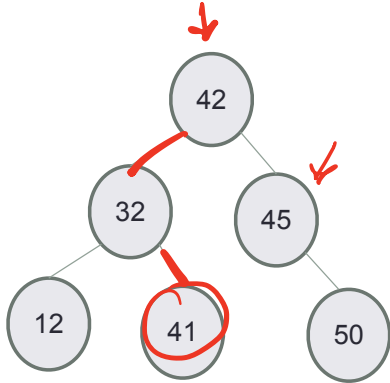
Given a BST of height H with N nodes, what is the worst case complexity of searching for a key?

- A. $O(1)$
- B. $O(\log H)$
- C. $O(H)$ Worst case.
- D. $O(H \cdot \log H)$
- E. $O(N)$

Best case : Searching for the key of root
 $O(1)$

Worst case : Search for a leaf that is H edges away.

Worst case Big-O of predecessor / successor



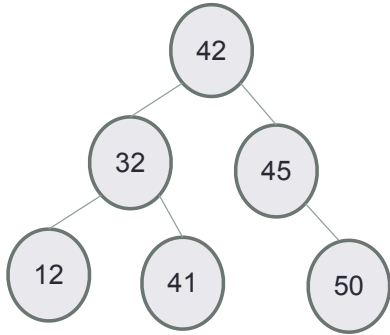
Best case: $O(1)$

Worst case: $O(H)$

Given a BST of height H and N nodes, what is the worst case complexity of finding the predecessor or successor key?

- A. $O(1)$
- B. $O(\log H)$
- C. $O(H)$**
- D. $O(H \cdot \log H)$
- E. $O(N)$

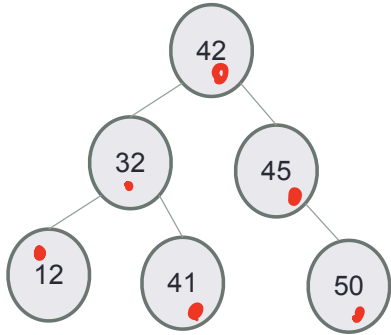
Worst case Big-O of delete



Given a BST of height H and N nodes, what is the worst case complexity of deleting a node?

- A. $O(1)$
- B. $O(\log H)$
- C. $O(H)$
- D. $O(H \cdot \log H)$
- E. $O(N)$

Big O of traversals



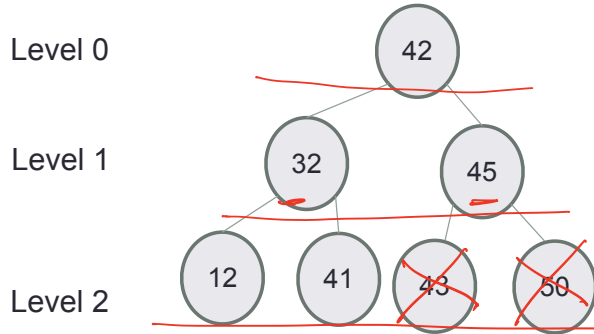
In Order: $O(n)$

Pre Order: $O(n)$

Post Order: $O(n)$

All operations except traversals are $O(H)$

Types of BSTs



Balanced BST:

Height of the tree

$$H = O(\log_2 N)$$

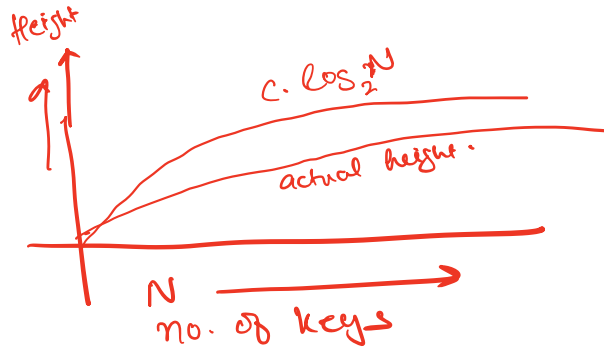
e.g. AVL, Red-Black Trees

Full Binary Tree: Every node other than the leaves has two children.

Show that a full BST

is a balanced BST

Complete Binary Tree: Every level, except possibly the last, is completely filled, and all nodes are as far left as possible



Relating H (height) and N (#nodes)

$$T(n) = c_1 n^2 + c_2 n$$

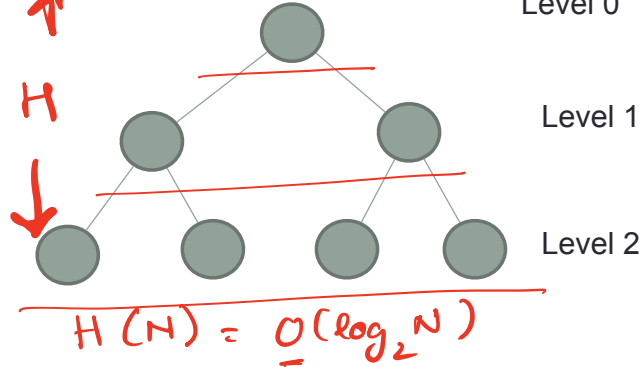
$$H(N) = ?$$

$$\sum_{i=0}^H 2^i = 2^{H+1} - 1$$

$$2^{H+1} - 1 = N$$

$$2^{H+1} = N + 1$$

$$H = \log_2(N+1) - 1$$




What is the height (exactly) of a full binary tree in terms of N ?

$$H(N) = \lceil \log_2(N+1) \rceil - 1 \quad \text{(Complete tree)}$$

$$H(N) \leq \log_2(N+1) - 1 + 1$$

$$H(N) = O(\log_2 N)$$

Balanced trees

- Balanced trees by definition have a height of $O(\log N)$
 - A completely filled tree is one example of a balanced tree
 - Other Balanced BSTs include AVL trees, red black trees and so on
 - Visualize operations on an AVL tree: <https://visualgo.net/bn/bst>
- 

Big-O analysis of iterative Fibonacci

```
function F(n) {  
  Create an array fib[1..n] ]  $O(1)$   
  fib[1] = 1  
  fib[2] = 1  
  for i = 3 to n:  $\leftarrow O(n)$   
    fib[i] = fib[i-1] + fib[i-2] ]  $O(1)$   
  return fib[n]  
}
```

$$T(n) = O(1) + O(n) \cdot O(1)$$
$$= O(n)$$

Big-O analysis of recursive Fibonacci

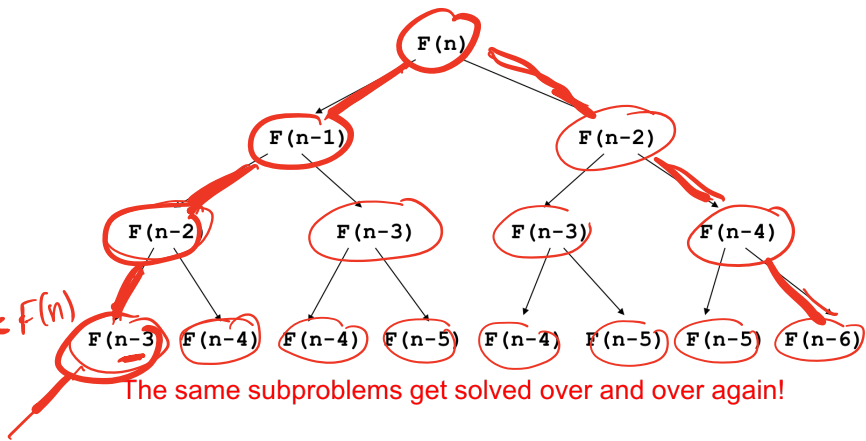
What takes so long? Let's unravel the recursion...

```
function F(n) {  
  if (n == 1) return 1  
  if (n == 2) return 1  
  return F(n-1) + F(n-2)  
}
```

$T(n)$: # of primitive steps to calculate $F(n)$

$$T(1) = 1 \quad T(2) = 2$$
$$T(n) = T(n-1) + T(n-2) + 5$$

Recurrence relation



$$T(n) = c \cdot \# \text{ of function calls}$$

We can obtain upper and lower bounds on the number of function calls by thinking about the minimum and maximum nodes in the call tree

$$T(n) \leq c(2^{n-1} - 1)$$

$$T(n) \geq c(2^{\lceil n/2 \rceil} - 1)$$

$$T(n) = O(2^n)$$

$$T(n) = \Omega(2^{\lceil n/2 \rceil})$$