BINARY SEARCH TREES

Gustavo Carvalho (ghpc@cin.ufpe.br)

Universidade Federal de Pernambuco Centro de Informática, 50740-560, Brazil







Agenda

1 Trees

2 Dictionaries implemented as BSTs

3 Binary tree traversals

4 Bibliography

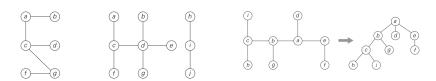




Trees¹

A free tree: a connected acyclic graph

- Forest: acyclic graph not necessarily connected
- Rooted tree (root typically on the top)



Applications: implement dictionaries, fault analysis, etc.





Trees: terminology for rooted trees

- Parent
- Siblings
- Subtree
- Internal vertices
- Ancestors / descendants (proper)





Trees: terminology for rooted trees

- Leaf: no children
- Depth/level of v: length of the unique path from the root to v
- Height: length of the longest unique path from the node to a leaf
- m-ary tree: every internal vertex has no more than m children
 - Complete *m*-ary tree: levels filled from top to bottom, left to right
 - Full *m*-ary tree: exactly *m* children
 - \blacksquare *m*-ary tree, where m=2: binary tree
- Ordered tree
 - Ordered binary tree: binary search tree (BST)
- Balanced tree

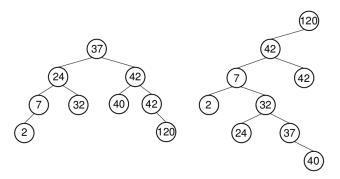




Binary search trees²

Left tree: insertion order = 37, 24, 42, 7, 2, 40, 42, 32, 120

Right tree: insertion order = 120, 42, 42, 7, 2, 32, 37, 24, 40







²Source: C. Shaffer. Data Structures and Algorithm Analysis. 2013.

Agenda

Dictionaries implemented as BSTs

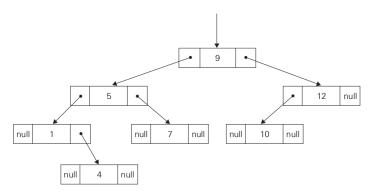
Binary tree traversals





Dictionaries implemented as BSTs³

Typical implementation: based on references (pointers)







8/49

Composite type (BSTNode):

- 1 Key key;
- 2 E element;
- 3 BSTNode left;
- 4 BSTNode right;

// left child // right child

Algorithm: BSTNode create_bstnode(Key k, E e)

- $n.key \leftarrow k$;
- 2 $n.element \leftarrow e$;
- 3 $n.left \leftarrow n.right \leftarrow NULL$;
- 4 return n;





Composite type (BST):

- 1 BSTNode root;
- 2 int nodecount;

// number of elements

Algorithm: BST create_bst()

- 1 $bst.root \leftarrow NULL$;
- 2 $bst.nodecount \leftarrow 0$;
- 3 return bst;





```
Algorithm: E find(BST bst, Key k)
```

1 return findhelp(bst.root, k);

Algorithm: E findhelp(BSTNode rt, Key k)

```
if rt = NULL then return NULL;
```

- 2 if rt.key > k then
- return findhelp(rt.left, k);
- 4 else if rt.key = k then
- 5 return rt.element;
- 6 else



4 D F 4 D F 4 D F 4 D



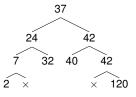
Algorithm: E find(BST bst, Key k)

return findhelp(bst.root, k);

Algorithm: E findhelp(BSTNode rt, Key k)

- if rt = NULL then return NULL;
- 2 if rt.key > k then
- return findhelp(rt.left, k);
- 4 else if rt.key = k then
- 5 return rt.element;
- 6 else

find(-,32)







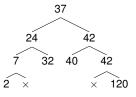
Algorithm: E find(BST bst, Key k)

return *findhelp*(*bst.root*, *k*);

Algorithm: E findhelp(BSTNode rt, Key k)

- if rt = NULL then return NULL:
- if rt.key > k then
- **return** findhelp(rt.left, k); 3
- else if rt.key = k then
- return rt.element; 5
- else
- **return** *findhelp*(*rt.right*, *k*);

find(-.32)







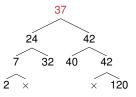
Algorithm: E find(BST bst, Key k)

return findhelp(bst.root, k);

Algorithm: E findhelp(BSTNode rt, Key k)

- if rt = NULL then return NULL;
- if rt.key > k then
- return findhelp(rt.left, k);
- 4 else if rt.key = k then
- 5 return rt.element;
- 6 else

find(-,32)







Algorithm: E find(BST bst, Key k)

return findhelp(bst.root, k);

Algorithm: E findhelp(BSTNode rt, Key k)

- if rt = NULL then return NULL;
- 2 if rt.key > k then
- return findhelp(rt.left, k);
- 4 else if rt.key = k then
- 5 return rt.element;
- 6 else

find(-,32)







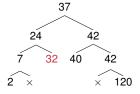
Algorithm: E find(BST bst, Key k)

return *findhelp*(*bst.root*, *k*);

Algorithm: E findhelp(BSTNode rt, Key k)

- if rt = NULL then return NULL;
- if rt.key > k then
- **return** findhelp(rt.left, k); 3
- else if rt.key = k then
- return rt.element: 5
- else
- **return** *findhelp*(*rt.right*, *k*);

find(-.32)



Returning element associated with

$$k = 32$$





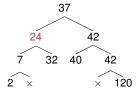
Algorithm: E find(BST bst, Key k)

return findhelp(bst.root, k);

Algorithm: E findhelp(BSTNode rt, Key k)

- if rt = NULL then return NULL;
- 2 if rt.key > k then
- return findhelp(rt.left, k);
- 4 else if rt.key = k then
- 5 return rt.element;
- 6 else
- return findhelp(rt.right, k);

$find(_-,32)$



Returning element associated with

$$k = 32$$





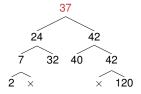
Algorithm: E find(BST bst, Key k)

return findhelp(bst.root, k);

Algorithm: E findhelp(BSTNode rt, Key k)

- 1 if rt = NULL then return NULL;
- 2 if rt.key > k then
- return findhelp(rt.left, k);
- 4 else if rt.key = k then
- 5 return rt.element;
- 6 else

find(-,32)



Returning element associated with

$$k = 32$$





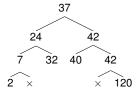
Algorithm: E find(BST bst, Key k)

return *findhelp*(*bst.root*, *k*);

Algorithm: E findhelp(BSTNode rt, Key k)

- if rt = NULL then return NULL;
- if rt.key > k then
- **return** findhelp(rt.left, k); 3
- else if rt.key = k then
- return rt.element; 5
- else
- **return** *findhelp*(*rt.right*, *k*);

find(-.32)



Returning element associated with

$$k = 32$$





Algorithm: void insert(BST bst, Key k, E e)

- 1 $bst.root \leftarrow inserthelp(bst.root, k, e);$
- 2 bst.nodecount++;

Algorithm: BSTNode inserthelp(BSTNode rt, Key k, E e)

```
if rt = NULL then return create\_bstnode(k, e);
```

- 2 if rt.key > k then
- $rt.left \leftarrow inserthelp(rt.left, k, e);$
- 4 else
- $rt.right \leftarrow inserthelp(rt.right, k, e);$
- 6 return rt;

Important: repeated keys go to the right subtree





Algorithm: void

insert(BST bst, Key k, E e)

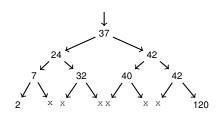
- 1 bst.root ←
 inserthelp(bst.root, k, e);
- bst.nodecount++;

Algorithm: BSTNode

inserthelp(BSTNode rt, Key k, E e)

- if rt = NULL then return
 create_bstnode(k, e);
- 2 if rt.key > k then
- $rt.left \leftarrow inserthelp(rt.left, k, e);$
- 4 else
- 6 return rt;

insert (_, 25,_)







Algorithm: void

insert(BST bst, Key k, E e)

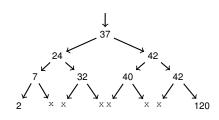
- 1 $bst.root \leftarrow inserthelp(bst.root, k, e);$
- bst.nodecount++;

Algorithm: BSTNode

inserthelp(BSTNode rt, Key k, E e)

- if rt = NULL then return $create_bstnode(k, e)$;
- 2 if rt.key > k then
- $rt.left \leftarrow inserthelp(rt.left, k, e);$
- 4 else
- $rt.right \leftarrow inserthelp(rt.right, k, e);$
- 6 return rt;

insert(_,25,_)







Algorithm: void

insert(BST bst, Key k, E e)

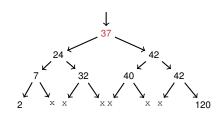
- 1 bst.root ←
 inserthelp(bst.root, k, e);
- 2 bst.nodecount++;

Algorithm: BSTNode

inserthelp(BSTNode rt, Key k, E e)

- 2 if rt.key > k then
- $rt.left \leftarrow inserthelp(rt.left, k, e);$
- 4 else
- $rt.right \leftarrow \\ inserthelp(rt.right, k, e);$
- 6 return rt;

insert (_, 25,_)







Algorithm: void

insert(BST bst, Key k, E e)

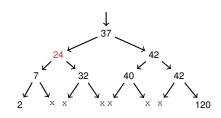
- $bst.root \leftarrow$ inserthelp(bst.root, k, e);
- bst.nodecount++:

Algorithm: BSTNode

inserthelp(BSTNode rt, Key k, E e)

- if rt = NUIII then return create_bstnode(k, e); if rt.key > k then
- $rt.left \leftarrow inserthelp(rt.left, k, e);$
- else
- $rt.right \leftarrow$ 5 inserthelp(rt.right, k, e);
- return rt;

insert(-.25.-)







Algorithm: void

insert(BST bst, Key k, E e)

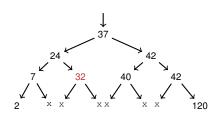
- 1 bst.root ←
 inserthelp(bst.root, k, e);
- 2 bst.nodecount++;

Algorithm: BSTNode

inserthelp(BSTNode rt, Key k, E e)

- if rt = NULL then return
 create_bstnode(k, e);
- 2 if rt.key > k then
- $rt.left \leftarrow inserthelp(rt.left, k, e);$
- 4 else
- $rt.right \leftarrow \\ inserthelp(rt.right, k, e);$
- 6 return rt;

insert(_, 25,_)







Algorithm: void

insert(BST bst, Key k, E e)

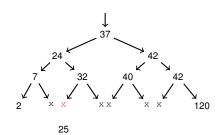
- 1 bst.root ←
 inserthelp(bst.root, k, e);
- bst.nodecount++;

Algorithm: BSTNode

inserthelp(BSTNode rt, Key k, E e)

- if rt = NULL then return $create_bstnode(k, e)$;
- 2 if rt.key > k then
- $rt.left \leftarrow inserthelp(rt.left, k, e);$
- 4 else
- $rt.right \leftarrow \\ inserthelp(rt.right, k, e);$
- 6 return rt;

insert(_,25,_)



Returning the reference to the new node





Algorithm: void

insert(BST bst, Key k, E e)

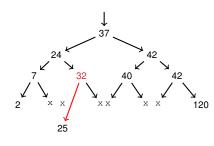
- 1 bst.root ←
 inserthelp(bst.root, k, e);
- 2 bst.nodecount++;

Algorithm: BSTNode

inserthelp(BSTNode rt, Key k, E e)

- if rt = NULL then return $create_bstnode(k, e)$;
- 2 if rt.key > k then
- $rt.left \leftarrow inserthelp(rt.left, k, e);$
- 4 else
- 6 **return** *rt*;

insert(_,25,_)



Returning the reference to the current node





Algorithm: void

insert(BST bst, Key k, E e)

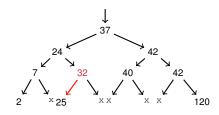
- 1 bst.root ←
 inserthelp(bst.root, k, e);
- 2 bst.nodecount++;

Algorithm: BSTNode

inserthelp(BSTNode rt, Key k, E e)

- if rt = NULL then return
 create_bstnode(k, e);
- 2 if rt.key > k then
- $rt.left \leftarrow inserthelp(rt.left, k, e);$
- 4 else
- 6 **return** *rt*;

insert(_,25,_)



Better presented this way

Returning the reference to the current node





Algorithm: void

insert(BST bst, Key k, E e)

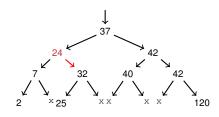
- 1 bst.root ←
 inserthelp(bst.root, k, e);
- 2 bst.nodecount++;

Algorithm: BSTNode

inserthelp(BSTNode rt, Key k, E e)

- if rt = NULL then return $create_bstnode(k, e)$;
- 2 if rt.key > k then
- $rt.left \leftarrow inserthelp(rt.left, k, e);$
- 4 else
- $rt.right \leftarrow \\ inserthelp(rt.right, k, e);$
- 6 **return** *rt*;

insert(_,25,_)



Returning the reference to the current node





Algorithm: void

insert(BST bst, Key k, E e)

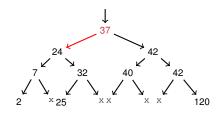
- 1 bst.root ←
 inserthelp(bst.root, k, e);
- 2 bst.nodecount++;

Algorithm: BSTNode

inserthelp(BSTNode rt, Key k, E e)

- if rt = NULL then return
 create_bstnode(k, e);
- 2 if rt.key > k then
- $rt.left \leftarrow inserthelp(rt.left, k, e);$
- 4 else
- 6 **return** *rt*;

insert(_,25,_)



Returning the reference to the current node (root)





Algorithm: void

insert(BST bst, Key k, E e)

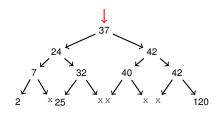
- 1 bst.root ←
 inserthelp(bst.root, k, e);
- bst.nodecount++;

Algorithm: BSTNode

inserthelp(BSTNode rt, Key k, E e)

- if rt = NULL then return
 create_bstnode(k, e);
- 2 if rt.key > k then
- $rt.left \leftarrow inserthelp(rt.left, k, e);$
- 4 else
- $rt.right \leftarrow inserthelp(rt.right, k, e);$
- 6 return rt;

insert(_,25,_)



Updating the reference to the root node





Algorithm: E remove(BST bst, Key k)

```
1 E temp ← findhelp(bst.root, k);
2 if temp ≠ NULL then
3 bst.root ← removehelp(bst.root, k);
4 bst.nodecount--;
```





return temp;

Algorithm: BSTNode removehelp(BSTNode rt, Key k)

```
if rt = NULL then return NULL:
    if rt.key > k then
         rt.left \leftarrow removehelp(rt.left, k);
3
    else if rt.key < k then
         rt.right \leftarrow removehelp(rt.right, k);
5
    else
        if rt.left = NULL then return rt.right;
7
        else if rt.right = NULL then return rt.left;
8
        else
9
             BSTNode temp \leftarrow getmin(rt.right);
10
             rt.element \leftarrow temp.element;
11
             rt.key \leftarrow temp.key;
12
             rt.right \leftarrow deletemin(rt.right);
                                                                Centro de
13
                                                                Informática
```

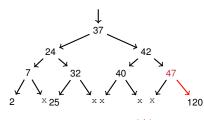


```
Algorithm: BSTNode
removehelp(BSTNode rt, Key k)
     if rt = NULL then return NULL:
     if rt.kev > k then
         rt.left \leftarrow removehelp(rt.left, k);
     else if rt.key < k then
         rt.riaht \leftarrow
          removehelp(rt.right, k);
     else
         if rt.left = NULL then return
          rt.riaht:
         else if rt.right = NULL then
 8
          return rt.left;
         else
             BSTNode
 10
               temp \leftarrow getmin(rt.right);
              rt.element \leftarrow temp.element;
 11
              rt.kev \leftarrow temp.kev;
 12
              rt.riaht \leftarrow
 13
               deletemin(rt.right);
```

Let *rt* be the node to be removed, three cases:

- (i) rt.left = NULL,
- (ii) rt.right = NULL, and
- (iii) otherwise.

remove (_, 47)



4 A B 4 A B 4



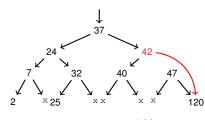


```
Algorithm: BSTNode
removehelp(BSTNode rt, Key k)
     if rt = NULL then return NULL:
     if rt.kev > k then
         rt.left \leftarrow removehelp(rt.left, k);
     else if rt.key < k then
         rt.riaht \leftarrow
          removehelp(rt.right, k);
     else
         if rt.left = NULL then return
          rt.right;
         else if rt.right = NULL then
 8
          return rt.left;
         else
             BSTNode
 10
               temp \leftarrow getmin(rt.right);
              rt.element \leftarrow temp.element;
 11
              rt.kev \leftarrow temp.kev;
 12
              rt.riaht \leftarrow
 13
               deletemin(rt.right);
```

Let *rt* be the node to be removed, three cases:

- (i) rt.left = NULL,
- (ii) rt.right = NULL, and
- (iii) otherwise.

remove (_, 47)



4 A B 4 A B 4



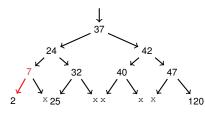


```
removehelp(BSTNode rt, Key k)
     if rt = NULL then return NULL:
     if rt.kev > k then
         rt.left \leftarrow removehelp(rt.left, k);
     else if rt.key < k then
         rt.riaht \leftarrow
           removehelp(rt.right, k);
     else
         if rt.left = NULL then return
           rt.right;
         else if rt.right = NULL then
 8
           return rt.left;
         else
              BSTNode
 10
               temp \leftarrow getmin(rt.right);
              rt.element \leftarrow temp.element;
 11
              rt.kev \leftarrow temp.kev;
 12
              rt.riaht \leftarrow
 13
               deletemin(rt.right);
```

Let rt be the node to be removed. three cases:

- (i) rt.left = NULL,
- (ii) rt.right = NULL, and
- (iii) otherwise.

remove (-,7)



4 A B 4 A B 4





LINIVERSIDADE FEDERAL DE PERNAMBLICO

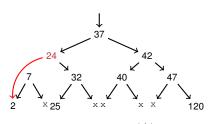
Algorithm: BSTNode

```
Algorithm: BSTNode
removehelp(BSTNode rt, Key k)
     if rt = NULL then return NULL:
     if rt.kev > k then
         rt.left \leftarrow removehelp(rt.left, k);
     else if rt.key < k then
         rt.riaht \leftarrow
          removehelp(rt.right, k);
     else
         if rt.left = NULL then return
          rt.right;
         else if rt.right = NULL then
 8
          return rt.left;
         else
             BSTNode
 10
               temp \leftarrow getmin(rt.right);
              rt.element \leftarrow temp.element;
 11
              rt.kev \leftarrow temp.kev;
 12
              rt.riaht \leftarrow
 13
               deletemin(rt.right);
```

Let *rt* be the node to be removed, three cases:

- (i) rt.left = NULL,
- (ii) rt.right = NULL, and
- (iii) otherwise.

remove (-, 7)



4 A B 4 B 4



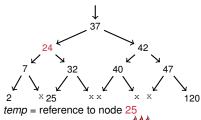


Universidade Federal de Pernambuco

```
Algorithm: BSTNode
removehelp(BSTNode rt, Key k)
     if rt = NULL then return NULL:
     if rt.kev > k then
         rt.left \leftarrow removehelp(rt.left, k);
     else if rt.key < k then
         rt.riaht \leftarrow
          removehelp(rt.right, k);
     else
         if rt.left = NULL then return
          rt.right;
         else if rt.right = NULL then
 8
          return rt.left;
         else
 9
             BSTNode
 10
               temp \leftarrow getmin(rt.right);
              rt.element \leftarrow temp.element:
 11
              rt.kev \leftarrow temp.kev;
 12
              rt.riaht \leftarrow
 13
               deletemin(rt.right);
```

```
Let rt be the node to be removed, three cases:
(i) rt.left = NULL,
(ii) rt.right = NULL, and
(iii) otherwise.
```

remove(_,24)





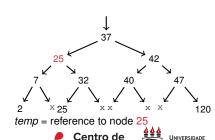
Universidade Federal de Pernambuco

```
Algorithm: BSTNode
removehelp(BSTNode rt, Key k)
     if rt = NULL then return NULL:
     if rt.kev > k then
         rt.left \leftarrow removehelp(rt.left, k);
     else if rt.key < k then
         rt.riaht \leftarrow
          removehelp(rt.right, k);
     else
         if rt.left = NULL then return
          rt.riaht:
         else if rt.right = NULL then
 8
          return rt.left;
         else
             BSTNode
 10
               temp \leftarrow getmin(rt.right);
             rt.element \leftarrow temp.element;
 11
             rt.kev \leftarrow temp.kev;
 12
              rt.riaht \leftarrow
 13
               deletemin(rt.right);
```

```
Let rt be the node to be removed.
three cases:
(i) rt.left = NULL,
(ii) rt.right = NULL, and
```

 $remove(_,24)$

(iii) otherwise.



Informática

DE PERNAMBLICO

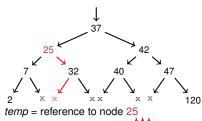
FEDERAL

```
Algorithm: BSTNode
removehelp(BSTNode rt, Key k)
     if rt = NULL then return NULL:
     if rt.kev > k then
         rt.left \leftarrow removehelp(rt.left, k);
     else if rt.key < k then
         rt.riaht \leftarrow
          removehelp(rt.right, k);
     else
         if rt.left = NULL then return
          rt.right;
         else if rt.right = NULL then
 8
          return rt.left;
         else
             BSTNode
 10
               temp \leftarrow getmin(rt.right);
              rt.element \leftarrow temp.element;
 11
              rt.kev \leftarrow temp.kev;
 12
              rt.riaht \leftarrow
 13
               deletemin(rt.right);
```

```
Let rt be the node to be removed, three cases:
(i) rt.left = NULL,
(ii) rt.right = NULL, and
```

remove (_, 24)

(iii) otherwise.



Centro de



Universidade Federal de Pernambuco

Algorithm: BSTNode getmin(BSTNode rt)

- if rt.left = NULL then return rt;
- 2 return getmin(rt.left);

Algorithm: BSTNode deletemin(BSTNode rt)

- if rt.left = NULL then return rt.right;
- *rt.left* \leftarrow *deletemin*(*rt.left*);
- з return rt;





Asymptotic efficiency of BSTs⁴

Average case: considering a balanced BST

Data Structure	Time Complexity								Space Complexity
	Average				Worst				Worst
	Access	Search	Insertion	Deletion	Access	Search	Insertion	Deletion	
Array	Θ(1)	O(n)	O(n)	O(n)	0(1)	0(n)	0(n)	0(n)	0(n)
Stack	Θ(n)	O(n)	0(1)	0(1)	0(n)	0(n)	0(1)	0(1)	0(n)
Queue	Θ(n)	O(n)	Θ(1)	Θ(1)	0(n)	0(n)	0(1)	0(1)	0(n)
Singly-Linked List	O(n)	O(n)	Θ(1)	0(1)	0(n)	0(n)	0(1)	0(1)	0(n)
Doubly-Linked List	O(n)	O(n)	0(1)	0(1)	0(n)	0(n)	0(1)	0(1)	0(n)
Skip List	O(log(n))	Θ(log(n))	Θ(log(n))	0(log(n))	0(n)	0(n)	0(n)	0(n)	0(n log(n))
Hash Table	N/A	0(1)	0(1)	0(1)	N/A	0(n)	0(n)	0(n)	0(n)
Binary Search Tree	0(log(n))	0(log(n))	0(log(n))	0(log(n))	0(n)	0(n)	0(n)	0(n)	0(n)
Cartesian Tree	N/A	Θ(log(n))	Θ(log(n))	O(log(n))	N/A	0(n)	0(n)	0(n)	0(n)
B-Tree	O(log(n))	O(log(n))	0(log(n))	O(log(n))	0(log(n))	0(log(n))	0(log(n))	0(log(n))	0(n)
Red-Black Tree	0(log(n))	Θ(log(n))	Θ(log(n))	0(log(n))	0(log(n))	0(log(n))	0(log(n))	0(log(n))	0(n)
Splay Tree	N/A	Θ(log(n))	Θ(log(n))	O(log(n))	N/A	0(log(n))	0(log(n))	0(log(n))	0(n)
AVL Tree	0(log(n))	O(log(n))	0(log(n))	0(log(n))	0(log(n))	0(log(n))	0(log(n))	0(log(n))	0(n)
KD Tree	Θ(log(n))	Θ(log(n))	Θ(log(n))	O(log(n))	0(n)	0(n)	0(n)	0(n)	0(n)





⁴Source: http://bigocheatsheet.com/

Agenda

- Binary tree traversals
- Bibliography



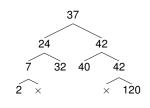


Binary tree traversals: preorder

The root is processed before the left and the right subtrees

Algorithm: void preorder(BSTNode rt)

```
if rt \neq NULL then
       // do something with rt
2
       preorder(rt.left);
       preorder(rt.right);
3
```



Let do something be printing the root's key

37, 24, 7, 2, 32, 42, 40, 42, 120



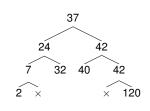


Binary tree traversals: inorder

The root is processed after the left subtree, but before the right subtree

Algorithm: void inorder(BSTNode rt)

```
if rt \neq NULL then
      inorder(rt.left);
2
       // do something with rt
       inorder(rt.right);
```



Let do something be printing the root's key

2, 7, 24, 32, 37, 40, 42, 42, 120

An inorder traversal visits the keys in a non-decreasing order



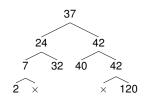


Binary tree traversals: posorder

The root is processed after the left and the right subtrees

Algorithm: void posorder(BSTNode rt)

```
if rt \neq NULL then
       posorder(rt.left);
2
       posorder(rt.right);
       // do something with rt
```



Let do something be printing the root's key

2, 7, 32, 24, 40, 120, 42, 42, 37





Agenda

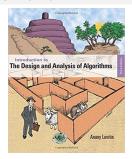
- Binary tree traversals
- Bibliography







Bibliography



Chapter 1 (pp. 31-35) Chapter 4 (pp. 163-164) Chapter 5 (pp. 182-185) Anany Levitin.

Introduction to the Design and Analysis of Algorithms. 3rd edition, Pearson, 2011.



Chapter 5 (pp. 145–170) Clifford Shaffer.

Data Structures and Algorithm Analysis. Dover, 2013.





BINARY SEARCH TREES

Gustavo Carvalho (ghpc@cin.ufpe.br)

Universidade Federal de Pernambuco Centro de Informática, 50740-560, Brazil





