Search Trees

Madhavan Mukund

https://www.cmi.ac.in/~madhavan

Programming, Data Structures and Algorithms using Python

Week 6

Dynamic sorted data

- Sorting is useful for efficient searching

Dynamic sorted data

- Sorting is useful for efficient searching
- What if the data is changing dynamically?
  - Items are periodically inserted and deleted

Dynamic sorted data

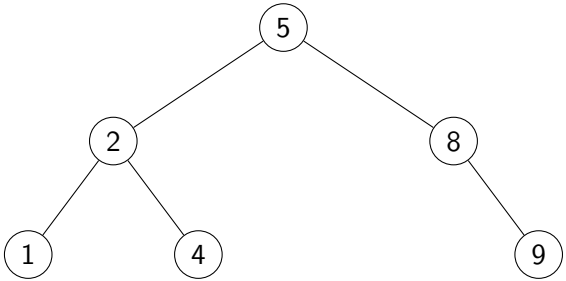
- Sorting is useful for efficient searching
- What if the data is changing dynamically?
  - Items are periodically inserted and deleted
- Insert/delete in a sorted list takes time $O(n)$

Dynamic sorted data

- Sorting is useful for efficient searching
- What if the data is changing dynamically?
  - Items are periodically inserted and deleted
- Insert/delete in a sorted list takes time $O(n)$
- Move to a tree structure, like heaps for priority queues

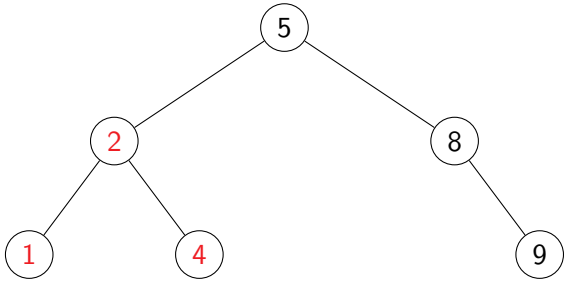
Binary search tree

- For each node with value v



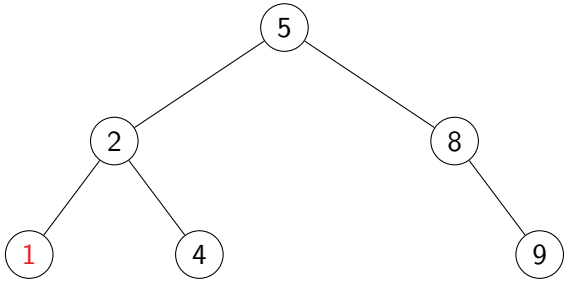
Binary search tree

- For each node with value v
  - All values in the left subtree are $< v$



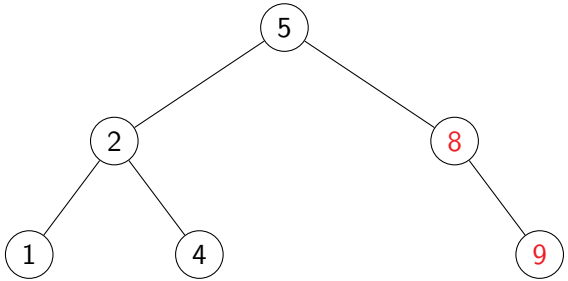
Binary search tree

- For each node with value v
  - All values in the left subtree are $< v$



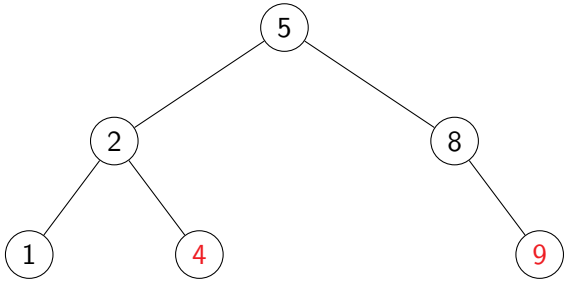
Binary search tree

- For each node with value v
  - All values in the left subtree are $< v$
  - All values in the left subtree are $> v$



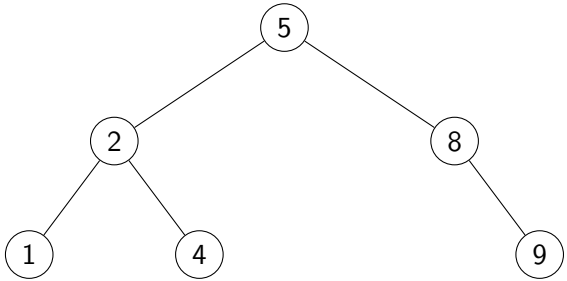
Binary search tree

- For each node with value v
  - All values in the left subtree are $< v$
  - All values in the left subtree are $> v$



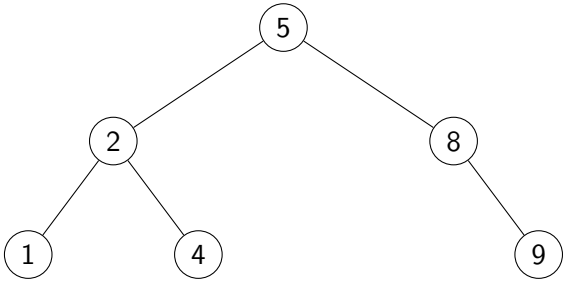
Binary search tree

- For each node with value v
  - All values in the left subtree are $< v$
  - All values in the left subtree are $> v$
- No duplicate values



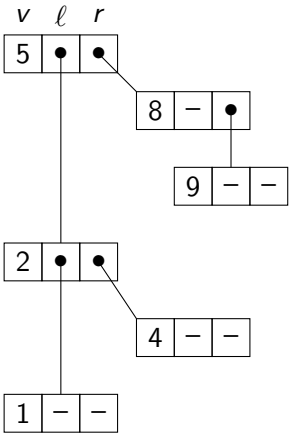
Implementing a binary search tree

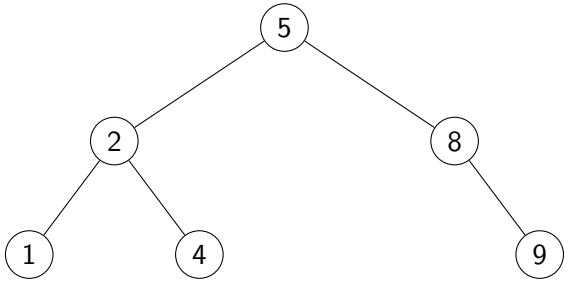
- Each node has a value and pointers to its children



Implementing a binary search tree

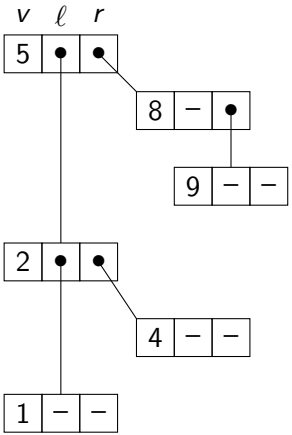
- Each node has a value and pointers to its children

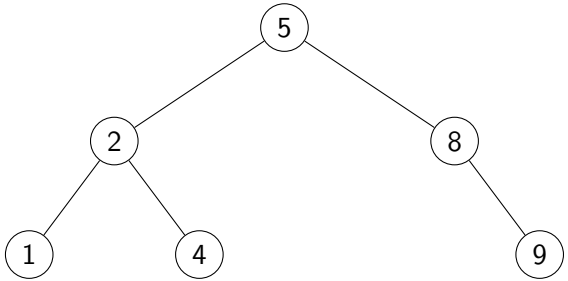




Implementing a binary search tree

- Each node has a value and pointers to its children

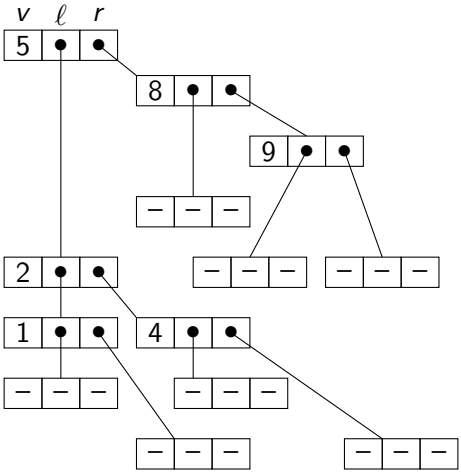


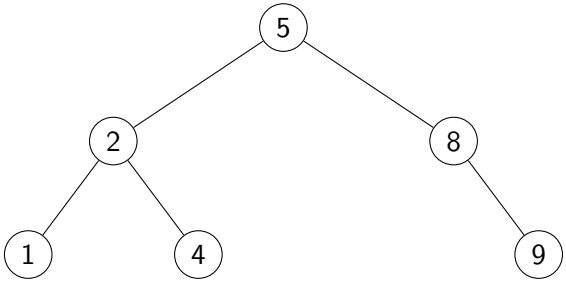


- Add a frontier with empty nodes, all fields –
  - Empty tree is single empty node
  - Leaf node points to empty nodes

Implementing a binary search tree

- Each node has a value and pointers to its children

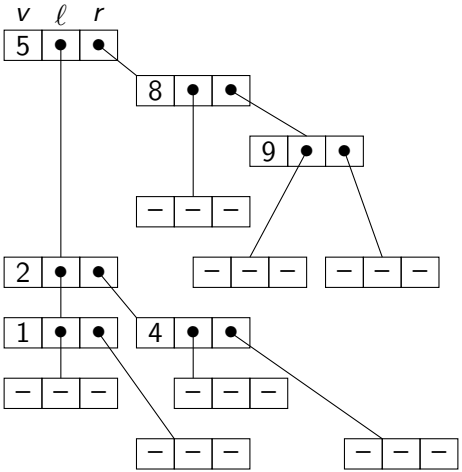


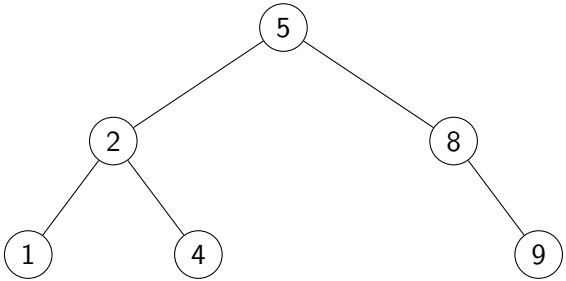


- Add a frontier with empty nodes, all fields –
  - Empty tree is single empty node
  - Leaf node points to empty nodes

Implementing a binary search tree

- Each node has a value and pointers to its children





- Add a frontier with empty nodes, all fields –
  - Empty tree is single empty node
  - Leaf node points to empty nodes
- Easier to implement operations recursively

The class Tree

- Three local fields, `value`, `left`, `right`
- Value `None` for empty value –
- Empty true has all fields `None`
- Leaf has a nonempty `value` and empty `left` and `right`

```
class Tree:

    # Constructor:
    def __init__(self,initval=None):
        self.value = initval
        if self.value:
            self.left = Tree()
            self.right = Tree()
        else:
            self.left = None
            self.right = None
        return

    # Only empty node has value None
    def isempty(self):
        return (self.value == None)

    # Leaf nodes have both children empty
    def isleaf(self):
        return (self.value != None and
                self.left.isempty() and
                self.right.isempty())
```

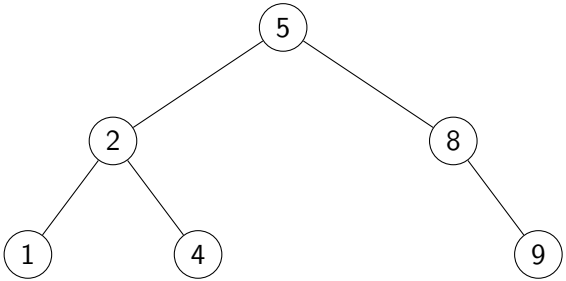
Inorder traversal

- List the left subtree, then the current node, then the right subtree
- Lists values in sorted order
- Use to print the tree

```
class Tree:
    ...
    # Inorder traversal
    def inorder(self):
        if self.isempty():
            return([])
        else:
            return(self.left.inorder()+
                   [self.value]+
                   self.right.inorder())

    # Display Tree as a string
    def __str__(self):
        return(str(self.inorder()))
```

Inorder traversal

- List the left subtree, then the current node, then the right subtree
- Lists values in sorted order
- Use to print the tree




```
class Tree:
    ...
    # Inorder traversal
    def inorder(self):
        if self.isempty():
            return([])
        else:
            return(self.left.inorder()+
                   [self.value]+
                   self.right.inorder())

    # Display Tree as a string
    def __str__(self):
        return(str(self.inorder()))
```

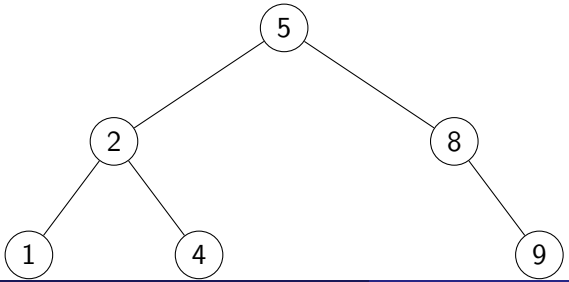
Find a value v

- Check value at current node
- If v smaller than current node, go left
- If v smaller than current node, go right
- Natural generalization of binary search

```
class Tree:
    ...
    # Check if value v occurs in tree
    def find(self,v):
        if self.isempty():
            return(False)

        if self.value == v:
            return(True)

        if v < self.value:
            return(self.left.find(v))

        if v > self.value:
            return(self.right.find(v))
```

Find a value v

- Check value at current node
- If v smaller than current node, go left
- If v smaller than current node, go right
- Natural generalization of binary search




```
class Tree:
    ...
    # Check if value v occurs in tree
    def find(self,v):
        if self.isempty():
            return(False)

        if self.value == v:
            return(True)

        if v < self.value:
            return(self.left.find(v))

        if v > self.value:
            return(self.right.find(v))
```

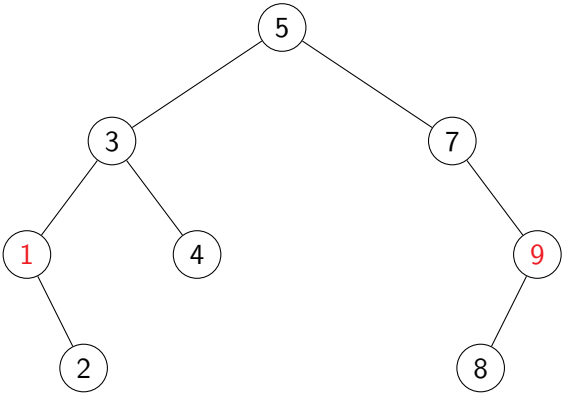
Minimum and maximum

- Minimum is left most node in the tree
- Maximum is right most node in the tree

```
class Tree:
    ...
    def minval(self):
        if self.left.isempty():
            return(self.value)
        else:
            return(self.left.minval())


    def maxval(self):
        if self.right.isempty():
            return(self.value)
        else:
            return(self.right.maxval())
```

Minimum and maximum

- Minimum is left most node in the tree
- Maximum is right most node in the tree




```
class Tree:
    ...
    def minval(self):
        if self.left.isempty():
            return(self.value)
        else:
            return(self.left.minval())


    def maxval(self):
        if self.right.isempty():
            return(self.value)
        else:
            return(self.right.maxval())
```

Insert a value v

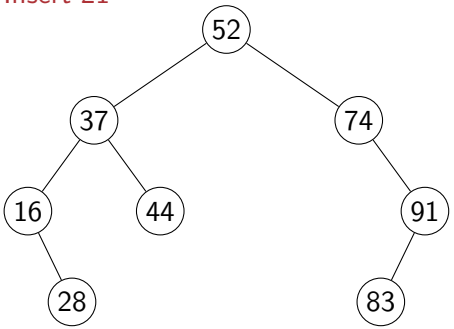
- Try to find `v`
- Insert at the position where `find` fails

```
class Tree:
    ...
    def insert(self,v):
        if self.isempty():
            self.value = v
            self.left = Tree()
            self.right = Tree()

        if self.value == v:
            return

        if v < self.value:
            self.left.insert(v)
            return

        if v > self.value:
            self.right.insert(v)
            return
```

Insert a value v

- Try to find `v`
- Insert at the position where `find` fails

Insert 21

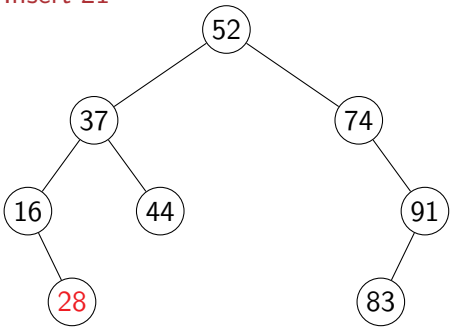



```
class Tree:
    ...
    def insert(self,v):
        if self.isempty():
            self.value = v
            self.left = Tree()
            self.right = Tree()

        if self.value == v:
            return

        if v < self.value:
            self.left.insert(v)
            return

        if v > self.value:
            self.right.insert(v)
            return
```

Insert a value v

- Try to find `v`
- Insert at the position where `find` fails

Insert 21

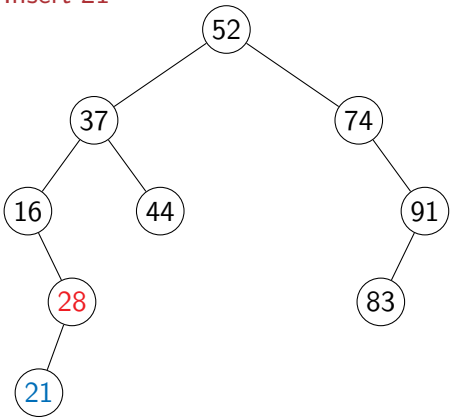



```
class Tree:
    ...
    def insert(self,v):
        if self.isempty():
            self.value = v
            self.left = Tree()
            self.right = Tree()

        if self.value == v:
            return

        if v < self.value:
            self.left.insert(v)
            return

        if v > self.value:
            self.right.insert(v)
            return
```

Insert a value v

- Try to find `v`
- Insert at the position where `find` fails

Insert 21

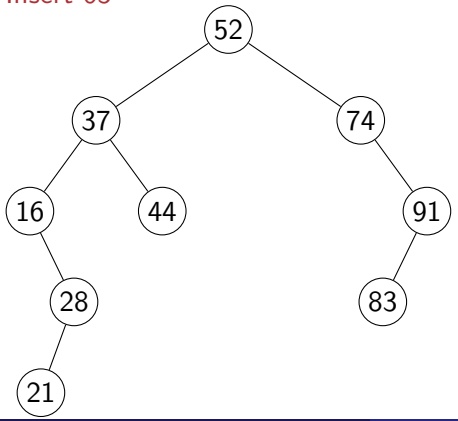



```
class Tree:
    ...
    def insert(self,v):
        if self.isempty():
            self.value = v
            self.left = Tree()
            self.right = Tree()

        if self.value == v:
            return

        if v < self.value:
            self.left.insert(v)
            return

        if v > self.value:
            self.right.insert(v)
            return
```

Insert a value v

- Try to find `v`
- Insert at the position where `find` fails

Insert 65

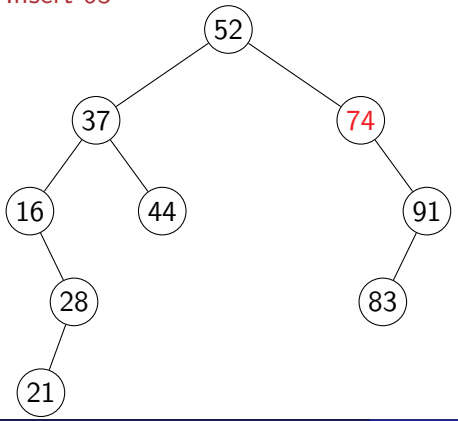



```
class Tree:
    ...
    def insert(self,v):
        if self.isempty():
            self.value = v
            self.left = Tree()
            self.right = Tree()

        if self.value == v:
            return

        if v < self.value:
            self.left.insert(v)
            return

        if v > self.value:
            self.right.insert(v)
            return
```

Insert a value v

- Try to find `v`
- Insert at the position where `find` fails

Insert 65

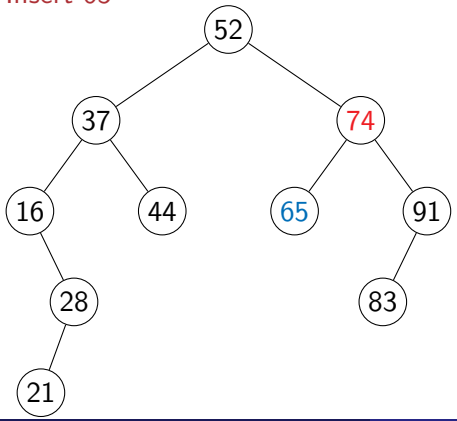



```
class Tree:
    ...
    def insert(self,v):
        if self.isempty():
            self.value = v
            self.left = Tree()
            self.right = Tree()

        if self.value == v:
            return

        if v < self.value:
            self.left.insert(v)
            return

        if v > self.value:
            self.right.insert(v)
            return
```

Insert a value v

- Try to find `v`
- Insert at the position where `find` fails

Insert 65

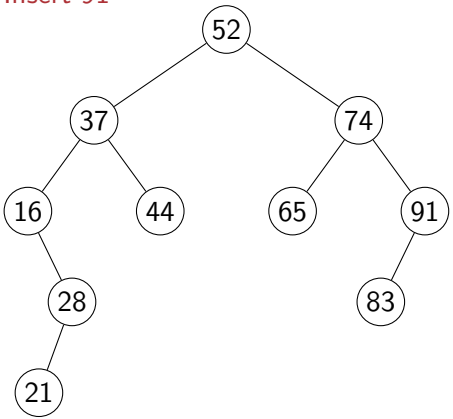



```
class Tree:
    ...
    def insert(self,v):
        if self.isempty():
            self.value = v
            self.left = Tree()
            self.right = Tree()

        if self.value == v:
            return

        if v < self.value:
            self.left.insert(v)
            return

        if v > self.value:
            self.right.insert(v)
            return
```

Insert a value v

- Try to find `v`
- Insert at the position where `find` fails

Insert 91

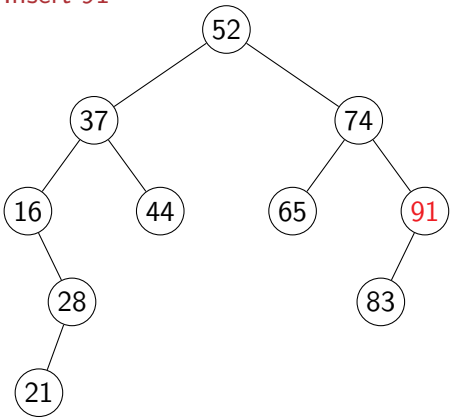



```
class Tree:
    ...
    def insert(self,v):
        if self.isempty():
            self.value = v
            self.left = Tree()
            self.right = Tree()

        if self.value == v:
            return

        if v < self.value:
            self.left.insert(v)
            return

        if v > self.value:
            self.right.insert(v)
            return
```

Insert a value v

- Try to find `v`
- Insert at the position where `find` fails

Insert 91




```
class Tree:
    ...
    def insert(self,v):
        if self.isempty():
            self.value = v
            self.left = Tree()
            self.right = Tree()

        if self.value == v:
            return

        if v < self.value:
            self.left.insert(v)
            return

        if v > self.value:
            self.right.insert(v)
            return
```

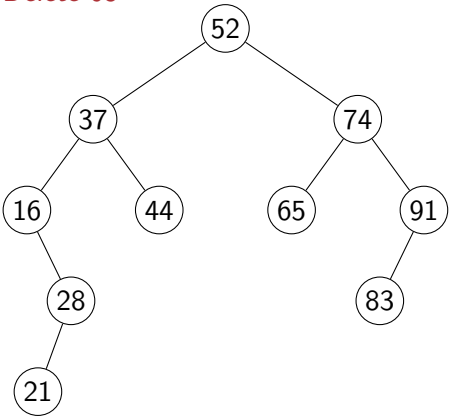
Delete a value v

- If `v` is present, delete
- Leaf node? No problem
- If only one child, promote that subtree
- Otherwise, replace `v` with `self.left.maxval()` and delete `self.left.maxval()`
  - `self.left.maxval()` has no right child

```
class Tree:
    ...
    def delete(self,v):
        if self.isempty():
            return
        if v < self.value:
            self.left.delete(v)
            return
        if v > self.value:
            self.right.delete(v)
            return
        if v == self.value:
            if self.isleaf():
                self.makeempty()
            elif self.left.isempty():
                self.copyright()
            elif self.right.isempty():
                self.copyleft()
            else:
                self.value = self.left.maxval()
                self.left.delete(self.left.maxval())
            return
```

Delete a value v

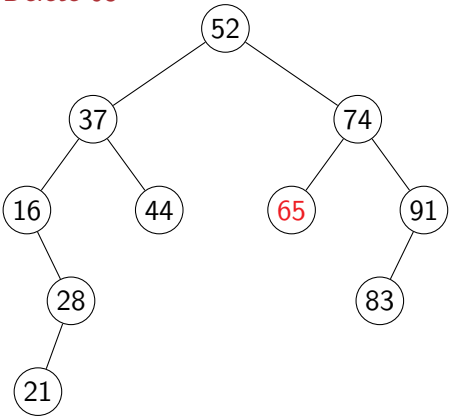
Delete 65




```
class Tree:
    ...
    def delete(self,v):
        if self.isempty():
            return
        if v < self.value:
            self.left.delete(v)
            return
        if v > self.value:
            self.right.delete(v)
            return
        if v == self.value:
            if self.isleaf():
                self.makeempty()
            elif self.left.isempty():
                self.copyright()
            elif self.right.isempty():
                self.copyleft()
            else:
                self.value = self.left.maxval()
                self.left.delete(self.left.maxval())
            return
```

Delete a value v

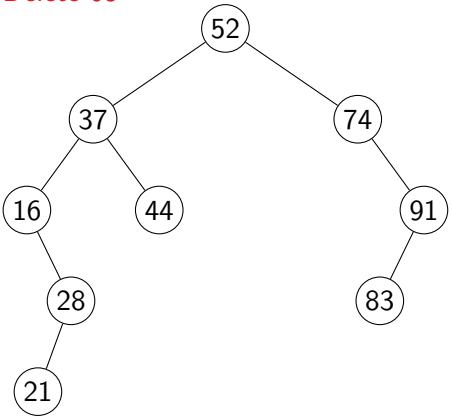
Delete 65




```
class Tree:
    ...
    def delete(self,v):
        if self.isempty():
            return
        if v < self.value:
            self.left.delete(v)
            return
        if v > self.value:
            self.right.delete(v)
            return
        if v == self.value:
            if self.isleaf():
                self.makeempty()
            elif self.left.isempty():
                self.copyright()
            elif self.right.isempty():
                self.copyleft()
            else:
                self.value = self.left.maxval()
                self.left.delete(self.left.maxval())
            return
```

Delete a value v

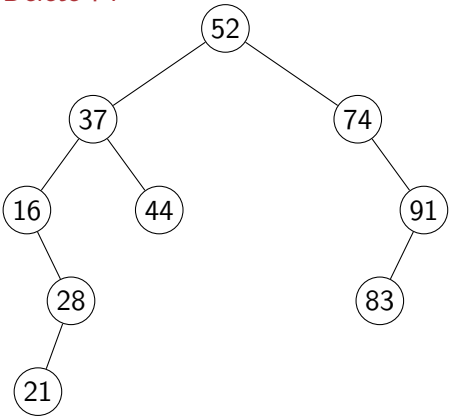
Delete 65




```
class Tree:
    ...
    def delete(self,v):
        if self.isempty():
            return
        if v < self.value:
            self.left.delete(v)
            return
        if v > self.value:
            self.right.delete(v)
            return
        if v == self.value:
            if self.isleaf():
                self.makeempty()
            elif self.left.isempty():
                self.copyright()
            elif self.right.isempty():
                self.copyleft()
            else:
                self.value = self.left.maxval()
                self.left.delete(self.left.maxval())
            return
```

Delete a value v

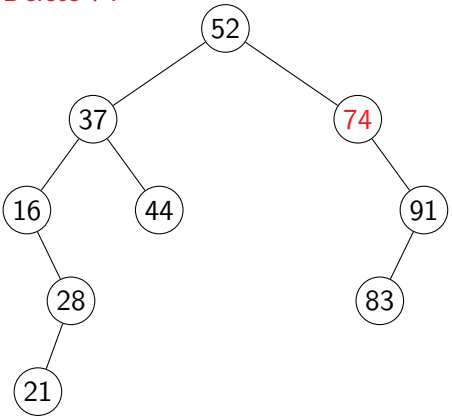
Delete 74




```
class Tree:
    ...
    def delete(self,v):
        if self.isempty():
            return
        if v < self.value:
            self.left.delete(v)
            return
        if v > self.value:
            self.right.delete(v)
            return
        if v == self.value:
            if self.isleaf():
                self.makeempty()
            elif self.left.isempty():
                self.copyright()
            elif self.right.isempty():
                self.copyleft()
            else:
                self.value = self.left.maxval()
                self.left.delete(self.left.maxval())
            return
```

Delete a value v

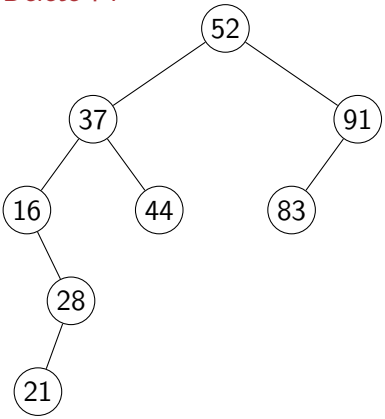
Delete 74




```
class Tree:
    ...
    def delete(self,v):
        if self.isempty():
            return
        if v < self.value:
            self.left.delete(v)
            return
        if v > self.value:
            self.right.delete(v)
            return
        if v == self.value:
            if self.isleaf():
                self.makeempty()
            elif self.left.isempty():
                self.copyright()
            elif self.right.isempty():
                self.copyleft()
            else:
                self.value = self.left.maxval()
                self.left.delete(self.left.maxval())
            return
```

Delete a value v

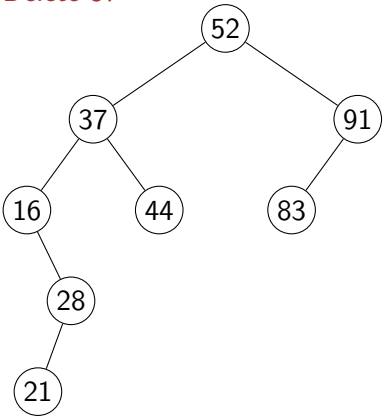
Delete 74




```
class Tree:
    ...
    def delete(self,v):
        if self.isempty():
            return
        if v < self.value:
            self.left.delete(v)
            return
        if v > self.value:
            self.right.delete(v)
            return
        if v == self.value:
            if self.isleaf():
                self.makeempty()
            elif self.left.isempty():
                self.copyright()
            elif self.right.isempty():
                self.copyleft()
            else:
                self.value = self.left.maxval()
                self.left.delete(self.left.maxval())
            return
```

Delete a value v

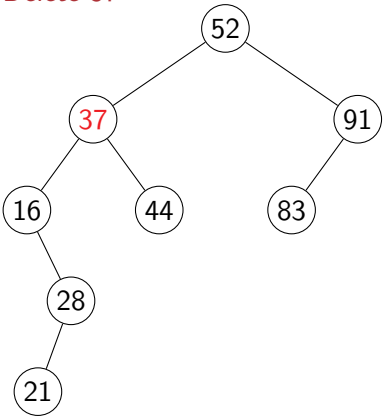
Delete 37




```
class Tree:
    ...
    def delete(self,v):
        if self.isempty():
            return
        if v < self.value:
            self.left.delete(v)
            return
        if v > self.value:
            self.right.delete(v)
            return
        if v == self.value:
            if self.isleaf():
                self.makeempty()
            elif self.left.isempty():
                self.copyright()
            elif self.right.isempty():
                self.copyleft()
            else:
                self.value = self.left.maxval()
                self.left.delete(self.left.maxval())
            return
```

Delete a value v

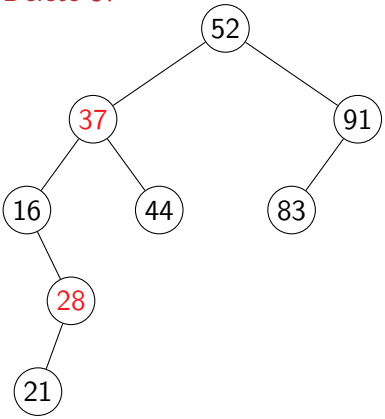
Delete 37




```
class Tree:
    ...
    def delete(self,v):
        if self.isempty():
            return
        if v < self.value:
            self.left.delete(v)
            return
        if v > self.value:
            self.right.delete(v)
            return
        if v == self.value:
            if self.isleaf():
                self.makeempty()
            elif self.left.isempty():
                self.copyright()
            elif self.right.isempty():
                self.copyleft()
            else:
                self.value = self.left.maxval()
                self.left.delete(self.left.maxval())
            return
```

Delete a value v

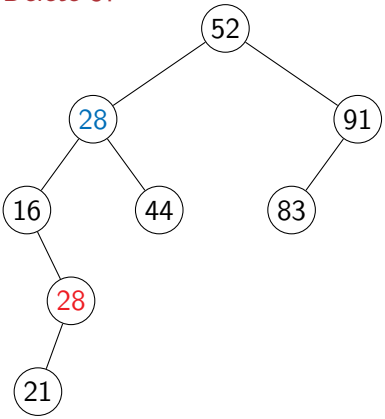
Delete 37




```
class Tree:
    ...
    def delete(self,v):
        if self.isempty():
            return
        if v < self.value:
            self.left.delete(v)
            return
        if v > self.value:
            self.right.delete(v)
            return
        if v == self.value:
            if self.isleaf():
                self.makeempty()
            elif self.left.isempty():
                self.copyright()
            elif self.right.isempty():
                self.copyleft()
            else:
                self.value = self.left.maxval()
                self.left.delete(self.left.maxval())
            return
```

Delete a value v

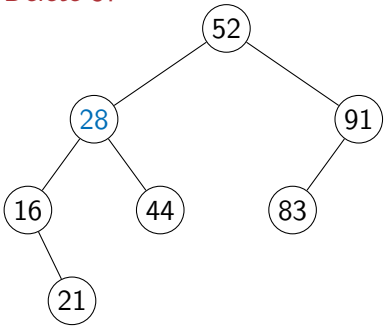
Delete 37




```
class Tree:
    ...
    def delete(self,v):
        if self.isempty():
            return
        if v < self.value:
            self.left.delete(v)
            return
        if v > self.value:
            self.right.delete(v)
            return
        if v == self.value:
            if self.isleaf():
                self.makeempty()
            elif self.left.isempty():
                self.copyright()
            elif self.right.isempty():
                self.copyleft()
            else:
                self.value = self.left.maxval()
                self.left.delete(self.left.maxval())
            return
```

Delete a value v

Delete 37




```
class Tree:
    ...
    def delete(self,v):
        if self.isempty():
            return
        if v < self.value:
            self.left.delete(v)
            return
        if v > self.value:
            self.right.delete(v)
            return
        if v == self.value:
            if self.isleaf():
                self.makeempty()
            elif self.left.isempty():
                self.copyright()
            elif self.right.isempty():
                self.copyleft()
            else:
                self.value = self.left.maxval()
                self.left.delete(self.left.maxval())
            return
```

Delete a value v

```
class Tree:
    ...
    def delete(self,v):
        if self.isempty():
            return
        if v < self.value:
            self.left.delete(v)
            return
        if v > self.value:
            self.right.delete(v)
            return
        if v == self.value:
            if self.isleaf():
                self.makeempty()
            elif self.left.isempty():
                self.copyright()
            elif self.right.isempty():
                self.copyleft()
            else:
                self.value = self.left.maxval()
                self.left.delete(self.left.maxval())
            return
```

```
# Convert leaf node to empty node
def makeempty(self):
    self.value = None
    self.left = None
    self.right = None
    return

# Promote left child
def copyleft(self):
    self.value = self.left.value
    self.right = self.left.right
    self.left = self.left.left
    return

# Promote right child
def copyright(self):
    self.value = self.right.value
    self.left = self.right.left
    self.right = self.right.right
    return
```

Complexity

- `find()`, `insert()` and `delete()` all walk down a single path
- Worst-case: height of the tree
- An unbalanced tree with n nodes may have height $O(n)$
- Balanced trees have height $O(\log n)$
- Will see how to keep a tree balanced to ensure all operations remain $O(\log n)$