COT 6405: Analysis of Algorithms

Giri NARASIMHAN

www.cs.fiu.edu/~giri/teach/6405F19.html

Data Structure Evolution

- Standard operations on data structures
  - Search
  - Insert
  - Delete
- Linear Lists
  - Implementation: Arrays (Unsorted and Sorted)
- Dynamic Linear Lists
  - Implementation: Linked Lists
- Dynamic Trees
  - Implementation: Binary Search Trees

Common Data Structures

	Search	Insert	Delete	Comments
Unsorted Arrays	O(N)	O(1)	O(N)	
Sorted Arrays	O(log N)	O(N)	O(N)	
Unsorted Linked Lists	O(N)	O(1)	O(N)	
Sorted Linked Lists	O(N)	O(N)	O(N)	
Binary Search Trees	O(H)	O(H)	O(H)	H = O(N)
Balanced BSTs	O(log N)	O(log N)	O(log N)	As H = O(log N)

Animations

- https://www.cs.usfca.edu/~galles/visualization/Algorithms.html
- https://visualgo.net/
- http://www.cs.armstrong.edu/liang/animation/animation.html
- http://www.cs.jhu.edu/~goodrich/dsa/trees/
- https://www.youtube.com/watch?v=Y-5ZodPvhmM
- http://www.algoanim.ide.sk/

Red-Black (RB) Trees

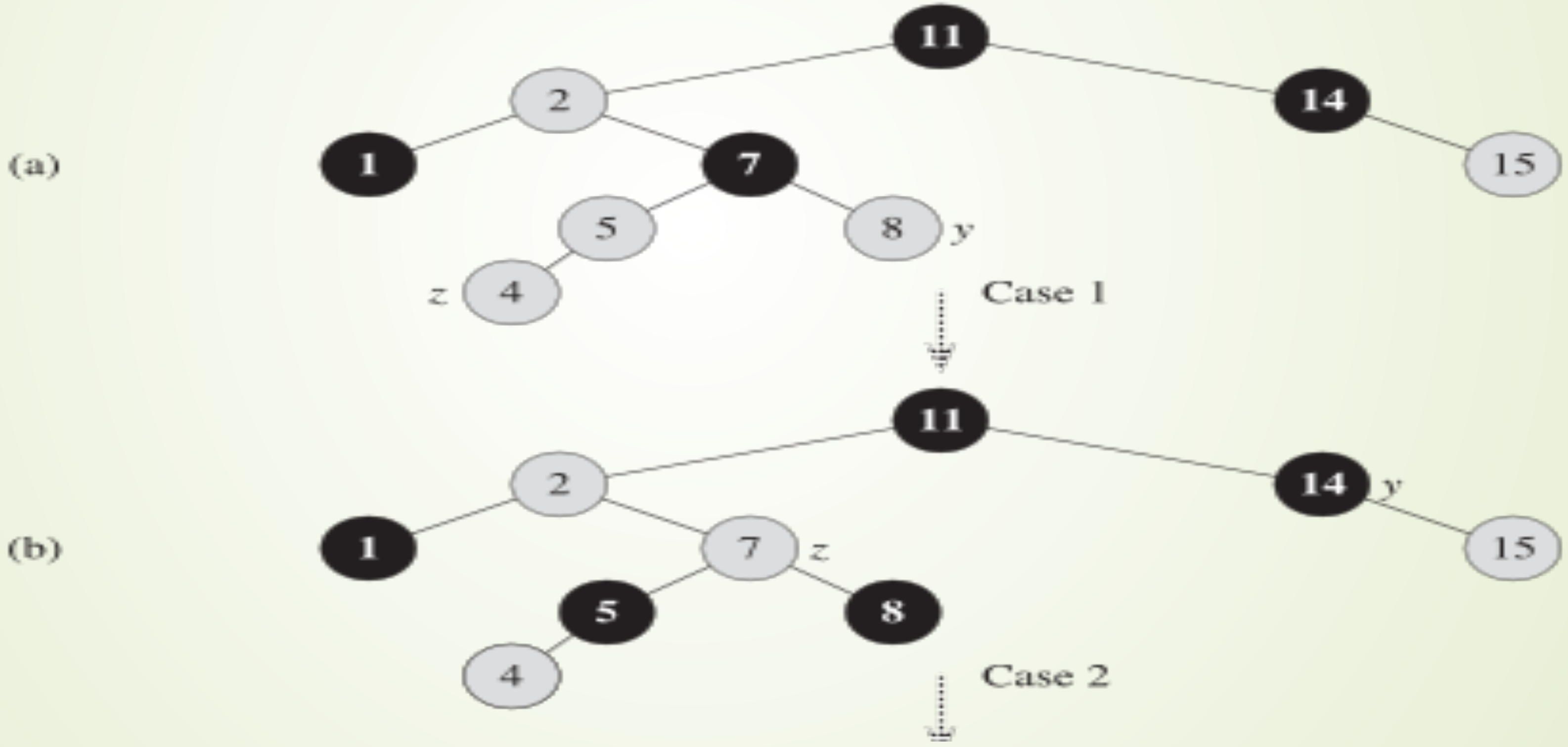
- Every node in a red-black tree is colored either red or black.
  - The root is always black.
  - Every path on the tree, from the root down to the leaf, has the same number of black nodes.
  - No red node has a red child.
  - Every NIL pointer points to a special node called NIL[T] and is colored black.
- Every RB-Tree with n nodes has black height at most logn
- Every RB-Tree with n nodes has height at most 2logn

Red-Black Tree Insert

```
RB-Insert (T,z)          // pg 315
        // Insert node z in tree T
        y = NIL[T]
        x = root[T]
        while (x ≠ NIL[T]) do
              y = x
              if (key[z] < key[x])
                     x = left[x]
                     x = right[x]
        p[z] = y
        if (y == NIL[T])
              root[T] = z
        else if (key[z] < key[y])
              left[y] = z
        else right[y] = z
        // new stuff
        left[z] = NIL[T]
        right[z] = NIL[T]
        color[z] = RED
        RB-Insert-Fixup (T,z)
```

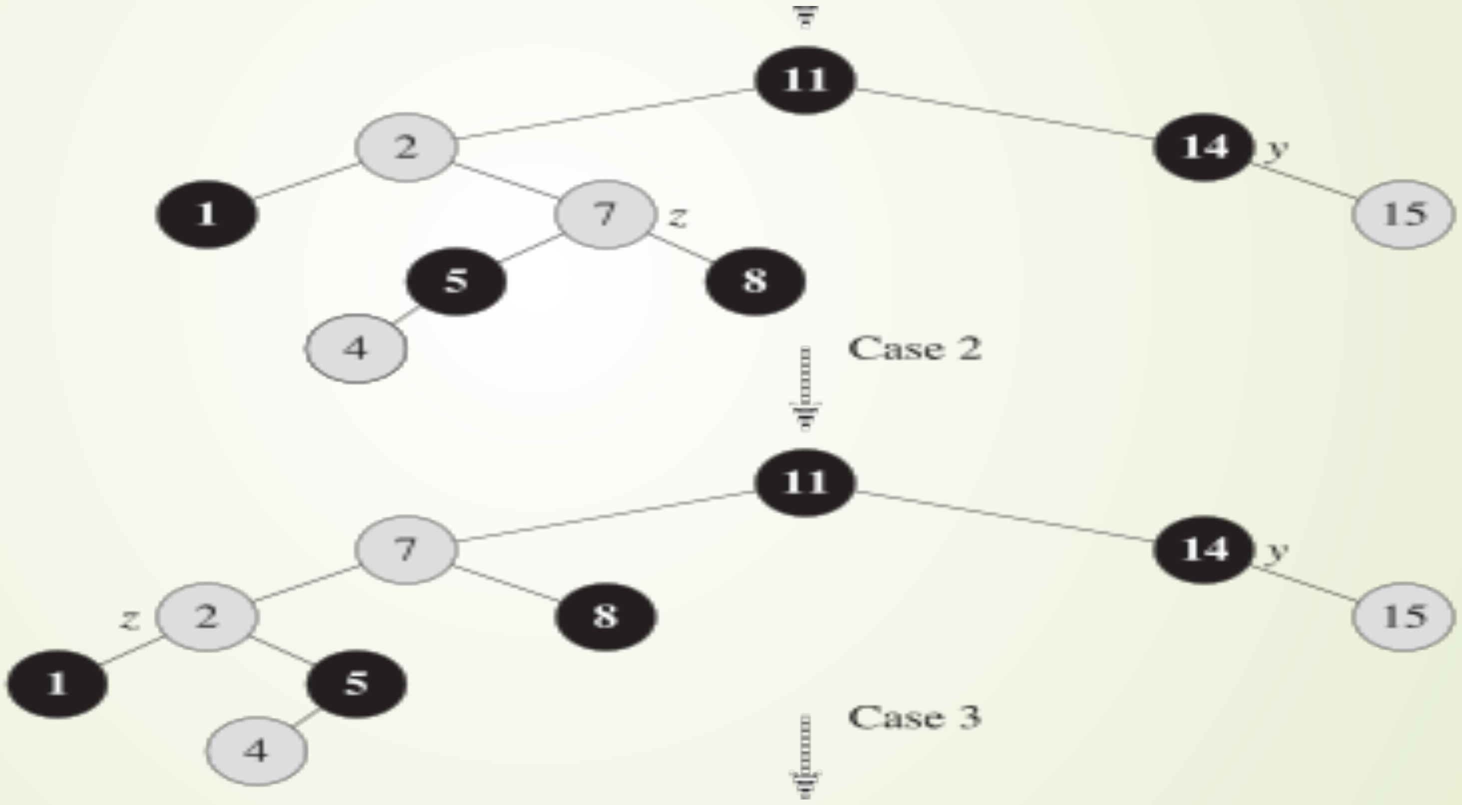
```
RB-Insert-Fixup (T,z)
    while (color[p[z]] == RED) do
      if (p[z] = left[p[p[z]]]) then
          y = right[p[p[z]]]
          if (color[y] == RED) then            // C-1
            color[p[z]] = BLACK
            color[y] = BLACK
            z = p[p[z]]
            color[z] = RED
          else    if (z == right[p[z]]) then // C-2
                z = p[z]
                LeftRotate(T,z)
            color[p[z]] = BLACK       // C-3
            color[p[p[z]]] = RED
            RightRotate(T,p[p[z]])
      else
          // Symmetric code: "right" ↔ "left"
          . . .
    color[root[T]] = BLACK
```

Case 1: Non-elbow; sibling of parent (y) red



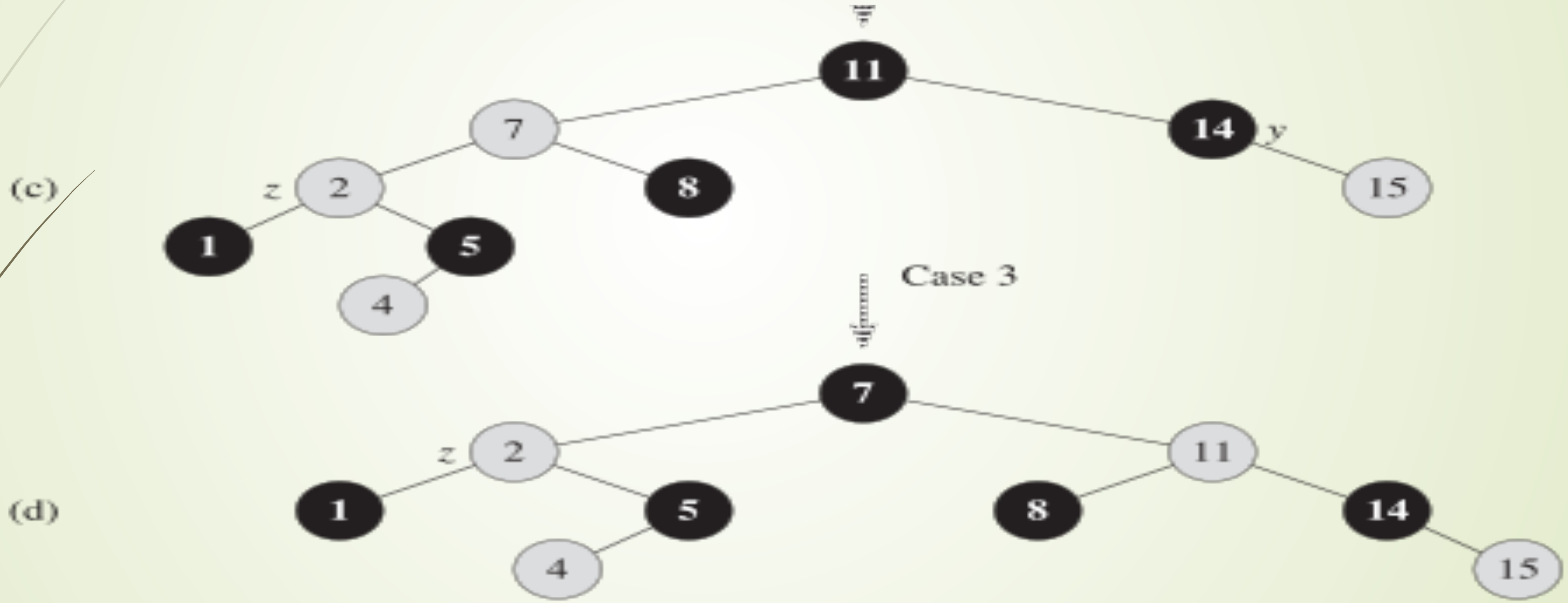
Case 2: Elbow case

(b)



(c)

Case 3: Non-elbow; sibling of parent black



Rotations

```
LeftRotate(T,x)    // pg 278
        // right child of x becomes x's parent.
        // Subtrees need to be readjusted.
        y = right[x]
        right[x] = left[y]    // y's left subtree becomes x's right
        p[left[y]] = x
        p[y] = p[x]
        if (p[x] == NIL[T]) then
              root[T] = y
        else if (x == left[p[x]]) then
              left[p[x]] = y
        else right[p[x]] = y
        left[y] = x
        p[x] = y
```

More Dynamic Operations

	Search	Insert	Delete	Comments
Unsorted Arrays	O(N)	O(1)	O(N)	
Sorted Arrays	O(log N)	O(N)	O(N)	
Unsorted Linked Lists	O(N)	O(1)	O(N)	
Sorted Linked Lists	O(N)	O(N)	O(N)	
Binary Search Trees	O(H)	O(H)	O(H)	H = O(N)
Balanced BSTs	O(log N)	O(log N)	O(log N)	As H = O(log N)

	Se/In/De	Rank	Select	Comments
Balanced BSTs	O(log N)	O(N)	O(N)	
Augmented BBSTs	O(log N)	O(log N)	O(log N)	

Operations on Dynamic RB Trees

- K-Selection
  - Select an item with a specified rank
- "Efficient" solution not possible without preprocessing
- Preprocessing - store additional information at nodes
- Inverse of K-Selection
  - Find rank of an item in the tree
- What information should be stored?
  - Rank
  - ??

OS-Rank

```
OS-RANK(x,y)
// Different from text (recursive version)
// Find the rank of x in the subtree rooted at y
  r = size[left[y]] + 1
  if x = y then return r
  else if ( key[x] < key[y] ) then
      return OS-RANK(x,left[y])
  else return r + OS-RANK(x,right[y] )
```

Time Complexity O(log n)

OS-Select

```
OS-SELECT(x,i) //page 304
// Select the node with rank i
// in the subtree rooted at x
1.  r = size[left[x]]+1
2.  if i = r then
3.          return x
4.  elseif  i < r then
5.          return OS-SELECT (left[x], i)
6.  else     return OS-SELECT (right[x], i-r)
```

Time Complexity O(log n)

RB-Tree Augmentation

- Augment x with **Size(x)**, where
  - Size(x) = size of subtree rooted at x
  - Size(NIL) = 0

Augmented Data Structures

- Why is it needed?
  - Because basic data structures not enough for all operations
  - storing extra information helps execute special operations more efficiently.
- Can any data structure be augmented?
  - Yes. Any data structure can be augmented.
- Can a data structure be augmented with any additional information?
  - Theoretically, yes.
- How to choose which additional information to store.
  - Only if we can maintain the additional information efficiently under all operations. That means, with additional information, we need to perform old and new operations efficiently maintain the additional information efficiently.

How to augment data structures

1. choose an underlying data structure
2. determine additional information to be maintained in the underlying data structure,
3. develop new operations,
4. verify that the additional information can be maintained for the modifying operations on the underlying data structure.

Augmenting RB-Trees

Theorem 14.1, page 309

Let f be a field that augments a red-black tree T with n nodes, and f(x) can be computed using only the information in nodes x, left[x], and right[x], including f[left[x]] and f[right[x]].

Then, we can maintain f(x) during insertion and deletion without asymptotically affecting the O(log n) performance of these operations.

For example,

size[x] = size[left[x]] + size[right[x]] + 1

rank[x] = ?

Augmenting information for RB-Trees

- Parent
- Height
- Any associative function on all previous values or all succeeding values.
- Next
- Previous

Reading for next class

- Red Black Trees
  - Properties
  - Invariants
  - Insert and Delete
- Mathematical Induction