

## Balanced Search: The Problem

Search trees important?

Insertion/deletion fast (on every operation, unlike hash table, to expand from time to time).

Range queries, sorting (unlike hash tables)

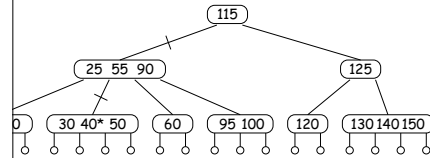
Performance from binary search tree requires remaining balanced  $\approx$  by some constant  $> 1$  at each node.

Worst case, that tree be "bushy"

Leaf nodes (most inner nodes with one child) perform like linked lists

Height difference between heights of any two subtrees of a node always differ by at most a constant factor  $K$ .

## Example Order 4 B-tree ((2,4) Tree)



Search paths show path when finding 40.

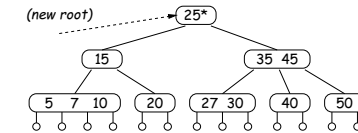
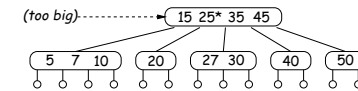
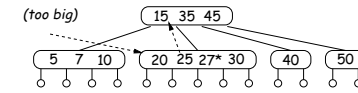
Search path: root, left child, second pointer, leaf.

Each node must have at least 2 children, and all leaves (little circles) are present, so height must be  $O(\lg N)$ .

B-tree, order typically much bigger

Order related to size of disk sector, page, or other convenient unit

## Inserting in B-Tree (Splitting)



## CS61B Lecture #29

Search structures (DS(IJ), Chapter 9)

Prime Numbers (DS(IJ), Chapter 11)

## Example of Direct Approach: B-Trees

Search grows/shrinks only at root, then two sides always have similar heights

A B-tree is an  $M$ -ary search tree,  $M > 2$ .

Each node, except root, has from  $\lceil M/2 \rceil$  to  $M$  children, and one key between each two children.

Each leaf has from 2 to  $M$  children (in non-empty tree).

The bottom of tree are all empty (don't really exist) and the top from root.

B-tree property:

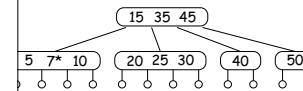
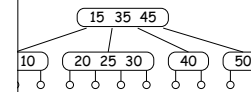
Keys are sorted in each node.

Search paths to subtrees to left of a key,  $K$ , are  $< K$ , and all to right are  $> K$ .

B-tree is a simple generalization of binary search.

Insertion: add just above bottom; split overfull nodes as needed, propagate up to parent.

## Inserting in B-tree (Simple Case)



## Red-Black Trees

Tree is a binary search tree with additional constraints which prevent it from being unbalanced if it can be.

Insertion is always  $O(\lg N)$ .

Java's TreeSet and TreeMap types.

When nodes are inserted or deleted, tree is rotated and recolored to restore balance.

Each node is (conceptually) colored red or black.

Root is black.

If a node contains no data (as for B-trees) and is black.

Every leaf node has same number of black ancestors.

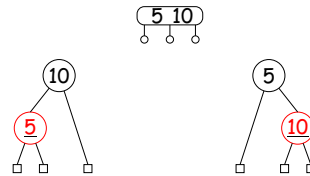
Internal node has two children.

Each internal node has two black children.

These constraints guarantee  $O(\lg N)$  searches.

## Left-Leaning Red-Black Trees

A (2,4) or (2,3) tree with three children may be represented in two different ways in a red-black tree:



This considerably simplifies insertion and deletion in a red-black tree by choosing the option on the left.

There is a one-to-one relationship between (2,4) trees and red-black trees.

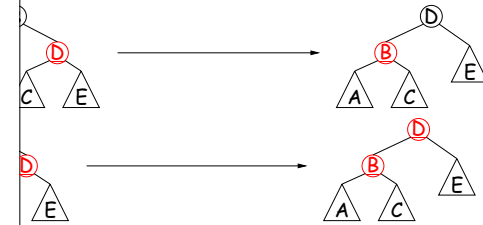
Left-leaning red-black trees are called *left-leaning red-black trees*.

For simplification, let's restrict ourselves to red-black trees that correspond to (2,3) trees (whose nodes have no more than two children), so that no red-black node has two red children.

## Rotations and Recolorings

In these cases, we'll augment the general rotation algorithms with recoloring.

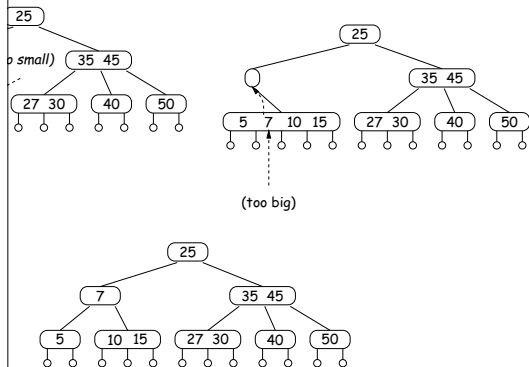
When we rotate, we pass the color from the original root to the new root, and color the root red. Examples:



These changes change the number of black nodes along any path from the root and the leaves.

## Deleting Keys from B-tree

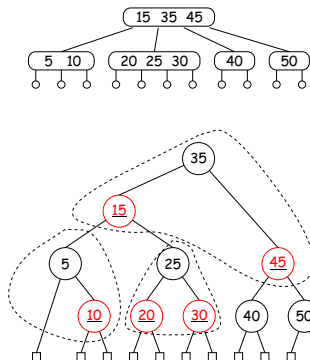
From last tree.



## Sample Red-Black Tree

A B-tree corresponds to a (2,4) tree, and the operations performed on the B-tree correspond to those on the (2,4) tree.

A (2,4) tree corresponds to a cluster of 1-3 red-black trees, where the top node is black and any others are red.

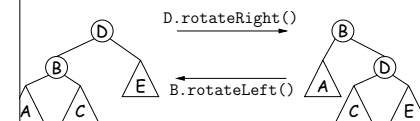


## Red-Black Insertion and Rotations

Insertion is performed just as for binary tree (color red except when tree is empty).

When a node has two red children, we rotate (and recolor) to restore red-black property, and thus rebalance the tree.

These operations preserve binary tree property, but change balance.



## The Algorithm (Sedgwick)

Binary-tree type RBTREE: basically ordinary BST nodes

the same as for ordinary BSTs, but we add some fixups to ensure the red-black properties.

```

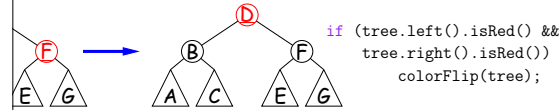
insert(RBTREE tree, KeyType key) {
    if (key == null)
        return null;
    return new RBTREE(key, null, null, RED);
    cmp = key.compareTo(tree.label());
    if (cmp < 0) tree.setLeft(insert(tree.left(), key));
    else tree.setRight(insert(tree.right(), key));
}

```

fixup(tree); // Only line that's all new!

## Fixing Up the Tree (II)

pack up 4-nodes into 3-nodes or 2-nodes.



```

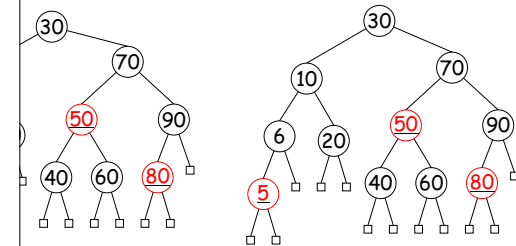
if (tree.left().isRed() &&
    tree.right().isRed())
    colorFlip(tree);

```

As a result of other fixups, or of insertion into the empty tree, it may end up red, so color the root black after the rest of the fixups are finished. (Not part of the fixup function; the end).

## Insertion Example (II)

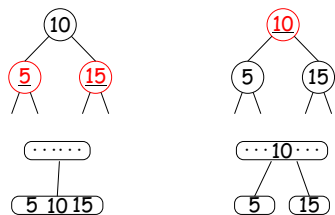
..., let's insert 6, leading to the tree on the left. This is not a red-black tree, so apply Fixup 1:



## Splitting by Recoloring

Insertions will temporarily create nodes with too many children, so we pack them up.

Coloring allows us to split nodes. We'll call it colorFlip:

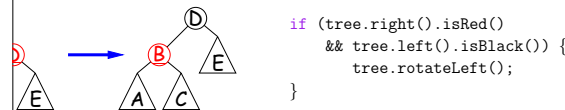


colorFlip joins the parent node, splitting the original.

## Fixing Up the Tree

As we walk back up the BST, we restore the left-leaning red-black tree and limit ourselves to red-black trees that correspond to the original by applying the following (in order) to each node:

Convert right-leaning trees to left-leaning:



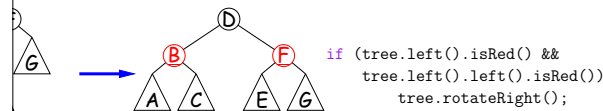
```

if (tree.right().isRed()
    && tree.left().isBlack()) {
    tree.rotateLeft();
}

```

Node B will be red, so that both B and D end up red. This is not a red-black tree.

Convert linked red nodes into a normal 4-node (temporarily).



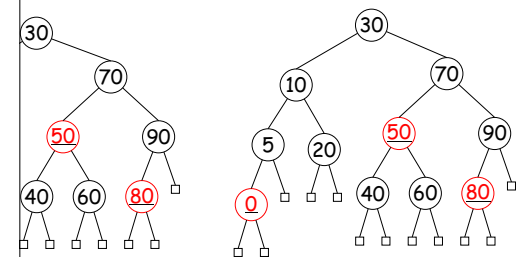
```

if (tree.left().isRed() &&
    tree.left().left().isRed())
    tree.rotateRight();

```

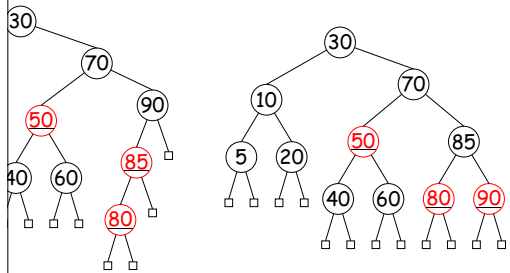
## Left-Leaning 2-3 Red-Black Insertion

..., initial tree on left. No fixups needed.



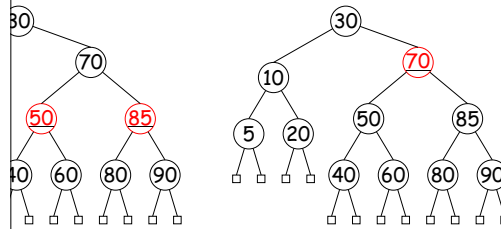
### Insertion Example (IIIa)

step 2.



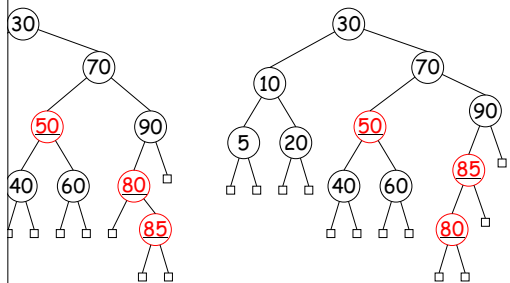
### Insertion Example (IIIc)

another 4-node, so apply fixup 3 again.



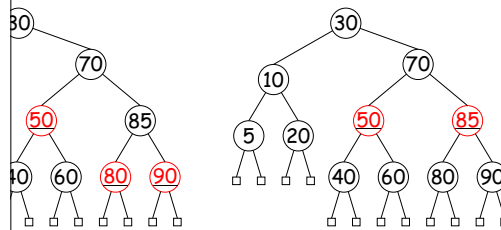
### Insertion Example (III)

for inserting 85. We need fixup 1 first.



### Insertion Example (IIIb)

is a 4-node, so apply fixup 3.



### Insertion Example (IIIId)

is a right-leaning tree, so apply fixup 1.

