6.006 Recitation
Build 2008.7
Outline

• Basic concepts review
• AVL algorithms
• Python implementation for AVLs
BST Invariants

- Binary rooted tree
- All left descendants have keys < node’s key
- All right descendants have keys > node’s key
Node Height

- Leaves: height = 0
- Inner nodes: height = max(children height) + 1
- Null tree: height = -1
- Rationale:
  - a subtree operation takes $O(h)$ time
Node Height

- Leaves: height = 0
- Inner nodes: height = max(children height) + 1
- Null tree: height = -1
- Rationale:
  - A subtree operation takes O(h) time
Node Height

• Leaves: height = 0

• Inner nodes: height = max(children height) + 1

• Null tree: height = -1

• Rationale:
  • a subtree operation takes $O(h)$ time
Node Height

- Leaves: height = 0
- Inner nodes: height = max(children height) + 1
- Null tree: height = -1
- Rationale:
  - a subtree operation takes $O(h)$ time
Node Height

- Leaves: height = 0
- Inner nodes: height = max(children height) + 1
- Null tree: height = -1

Rationale:
- a subtree operation takes $O(h)$ time
Node Height

- Leaves: height = 0
- Inner nodes: height = max(children height) + 1
- Null tree: height = -1
- Rationale:
  - a subtree operation takes $O(h)$ time
Balanced Trees

- Small tree height means fast operations
- Pack many nodes in trees with low heights
- Perfectly balanced tree: $2^{h+1} - 1$ nodes
- We only care about asymptotic notation
- Nodes = $f(\text{height})$ must be exponential
AVL Trees

- Regular BST with extra invariants:
  - absolute value(left child height - right child height) <= 1
  - Each subtree is AVL
Least dense AVL
Least dense AVL
Least dense AVL
Least dense AVL
Least dense AVL
Least dense AVL
Least dense AVL
Least dense structure

- \( \text{Nodes}(-1) = 0 \)
- \( \text{Nodes}(0) = 1 \)
- \( \text{Nodes}(h) = 1 + \text{Nodes}(h-1) + \text{Nodes}(h-2) \)

- Looks like Fibonacci, must be exponential
Pwnage with AVLs 101

• Goals
  • Reuse the code we wrote before
  • Start with an AVL, end up with an AVL
• Managerial Input (the ‘doh’ words)
  • Insert and delete like it’s a BST
  • Patch to make it an AVL again
Key Observation
Key Observation

Adding or removing a node only upsets the heights on a single path to the root.
Pwnage with AVLs 201

- Will obviously have to move nodes around
- But must keep track of
  - Height
  - Augmented data
  - Invariants for AVL, BST
- Need a tool that preserves most structure
Uberpoke (rotations)
Huh? Do that again?
Huh? Do that again?
Huh? Do that again?
Huh? Do that again?
Huh? Do that again?
Huh? Do that again?
Huh? Do that again?
Huh? Do that again?
Huh? Do that again? 

\[
\begin{array}{c}
\text{P} \\
\text{x} \\
\text{A} \\
\text{B} \\
\text{C}
\end{array}
\quad = 
\begin{array}{c}
\text{P} \\
\text{y} \\
\text{A} \\
\text{B} \\
\text{C}
\end{array}
\]
Rebalancing

• Rotations are quite teh uberpoke

• Need a master plan for using them

• Managerial Input: call it ‘rebalancing’

• Divide and conquer: start from the bottom, fix up the tree level by level
Rebalancing: easy

Rebalancing:

\[
\begin{align*}
&:x \\
&\downarrow \quad \downarrow \quad \downarrow \\
&k-1 \quad y \quad k+1 \\
&\downarrow \quad \downarrow \quad \downarrow \\
&A \quad B \quad C \\
&\downarrow \quad \downarrow \quad \downarrow \\
k-1 \quad k \quad k
\end{align*}
\]
Rebalancing: easy

Rotate left

:(
Rebalancing: easy

Rotate left

\[(k-1, k, k+1)\] or \[(k, k+1, k-1)\]
Rebalancing: easy?
Rebalancing: easy?

Rotate left

:(

k-1
A

x

k+1
y

B

k

C

k-1
Rebalancing: easy?

Rotate left

:(

\[
\begin{align*}
  x & \quad k+1 \\
  \quad y & \quad k+1 \\
  A & \quad k-1 \\
  B & \quad k \\
  C & \quad k-1
\end{align*}
\]

\[
\begin{align*}
  \quad y & \quad k+1 \\
  \quad x & \quad k+1 \\
  C & \quad k-1 \\
  A & \quad k-1 \\
  B & \quad k
\end{align*}
\]
WTF?

Rotate left

Before:

\[ k-1 \quad k \quad k+1 \]

\[ \begin{array}{c}
    A \\
    B \\
    C
\end{array} \]

\[ (\) \]

After:

\[ k+1 \quad k \quad k-1 \]

\[ \begin{array}{c}
    k-1 \\
    A \\
    B \\
    C
\end{array} \]

\[ (\) \]
WTF?

Rotate left

:( x

k-1

A

y

k

B

k+1

C

:( y

k+1

x

k-1

C

k-1

A

B

k
B cannot be taller than C
Rebalancing: hack it up

\[ x \rightarrow y \rightarrow \begin{array}{c} k-1 \\ A \end{array} \]

\[ z \rightarrow \begin{array}{c} k \\ C \end{array} \rightarrow \begin{array}{c} k-1 \\ k-1 \\ k-2 \\ D \end{array} \]

\[ \text{or} \]

\[ \begin{array}{c} k-1 \\ k-2 \\ E \end{array} \rightarrow \begin{array}{c} k-1 \\ k-2 \\ \end{array} \]
Rebalancing: hack it up

Rotate right

(x

A

y

k+1

z

k

k-1

D

k-1 or k-2

E

k-1 or k-2

C

k-1

:((
Rebalancing: hack it up

Rotate right

(k-1 or k-2)
(k-1 or k-2)

(k-1)
(k-1)

(k+1)
(k+1)
and in the end it's right
and in the end it's right

Rotate left

:(

x

A

k-1

k-1 or k-2

D

k-1 or k-2

E

k-1

k

z

k+1

y

C
and in the end it's right

Rotate left

):( x

k-1 A

k-1 or k-2

k-1 or k-2

k-1

z

k+1

y

k

k-1

D

E

C

k

k-1

x

k

k-1

A

D

E

k-1 or k-2

C

k-1

z

k+1

y

k

k-1

D

E

k-1 or k-2

C

k-1
Rebalancing one level

• AVL violation at current node
• Right is than left?
  • Right.left taller than right.right?
  • Rotate right to the right
• Either way, rotate current to the left
• Left is heavier than right: symmetry
Rebalancing wrap-up

• Know how to fix one level, use that to fix everything along the path to the root

• Must recompute height on-the-go

• If recomputing for all nodes along the path on each rotation, $O(\log^2(h))$

• Why is rebalancing $O(\log(h))$?
Python Code

‘cause you can’t live on bubbles and lines
AVL Design

- BST
  - incorporate the deletion hack
- AVL
  - inherited from BST, uses AVLnode
- AVLnode
  - does all the heavy lifting
Return values matter!

- **insert**: returns the newly inserted node
- **delete**: returns the deleted node (its parent link still indicates where it was hanging)
class BST(object):
    def __init__(self, NodeType=BSTnode):
        self.root = None
        self.NodeType = NodeType
        self.psroot = self.NodeType(None, None)

    def reroot(self):
        self.root = self.psroot.left

    def insert(self, t):
        if self.root is None:
            self.psroot.left = self.NodeType(self.psroot, t)
            self.reroot()
            return self.root
        else:
            return self.root.insert(t, self.NodeType)
class AVL(BST):
    def __init__(self):
        BST.__init__(self, AVLNode)

    def insert(self, t):
        node = BST.insert(self, t)
        node.rebalance()
        self.reroot()

    def delete(self):
        node = BST.delete(self)
        node.parent.rebalance()
        self.reroot()
def height(node):
    if node is None:
        return -1
    else:
        return node.height

class AVLnode(BSTnode):
    def update_stats(self):
        self.height = max(height(self.left), height(self.right)) + 1
        BSTnode.update_stats(self)
class AVLnode(BSTnode):
    def left_rotate(self):
        x = self; y = x.right
        y.parent = x.parent
        if y.parent.left is x:
            y.parent.left = y
        elif y.parent.right is x:
            y.parent.right = y
        x.right = y.left
        if x.right is not None:
            x.right.parent = x
        y.left = x
        x.parent = y
        x.update_stats()
        y.update_stats()
        return y
class AVLnode(BSTnode):

    def rebalance(self):
        if self.key is None: return

        self.update_height()

        if height(self.left) >= 2 + height(self.right):
            if height(self.left.left) < height(self.left.right):
                self.left.left_rotate()
                self.right_rotate()
            else:
                self.right.right_rotate()
                self.left_rotate()
        else:
            if height(self.right.right) < height(self.right.left):
                self.right.right_rotate()
                self.left_rotate()
            else:
                self.left.left_rotate()

        self.parent.rebalance()
And we’re done!

- costan@mit.edu
- (617) 230-9694, no voicemail
- AIM: victorcostan
- Google Talk: costan@gmail.com
- 32G-8th Floor
v. Next

• Don’t need to reroot(), can just use root() as a function that returns psroot.left

• Put the [1, 2... 8] exercise on the slides so it’s remembered

• Quote relevant CLRS chapters on the wiki