Splay Trees

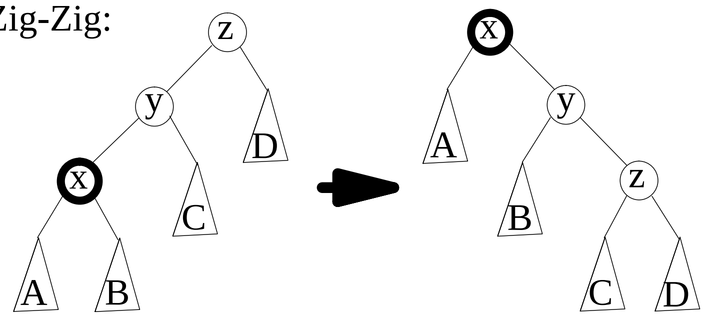
Self-adjusting counterpart to AVL and red-black trees

Advantages - 1) no balance bits, 2) some help with locality of reference, 3) amortized complexity is same as AVL and red-black trees

Disadvantage - worst-case for operation is O(n)

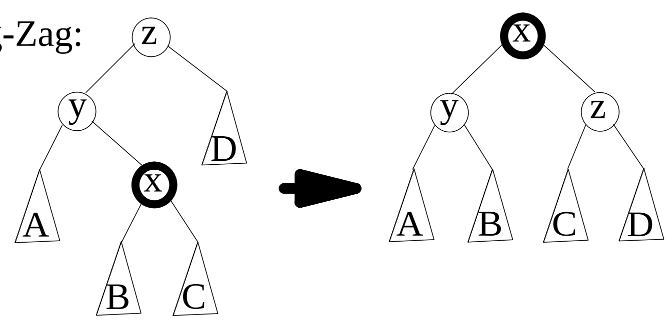
Algorithms are based on use of rotations to *splay* the last node processed (x) to root position.

Zig-Zig:



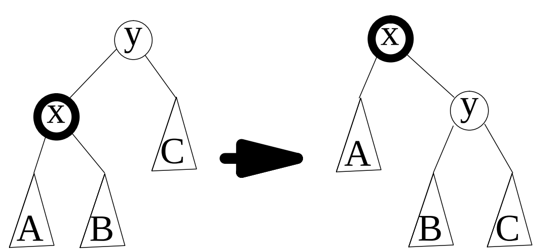
1. Single right rotation at z.
2. Single right rotation at y.
   (+ symmetric case)

Zig-Zag:



Double right rotation at z.
(+ symmetric case)

Zig: Applies ONLY at the root

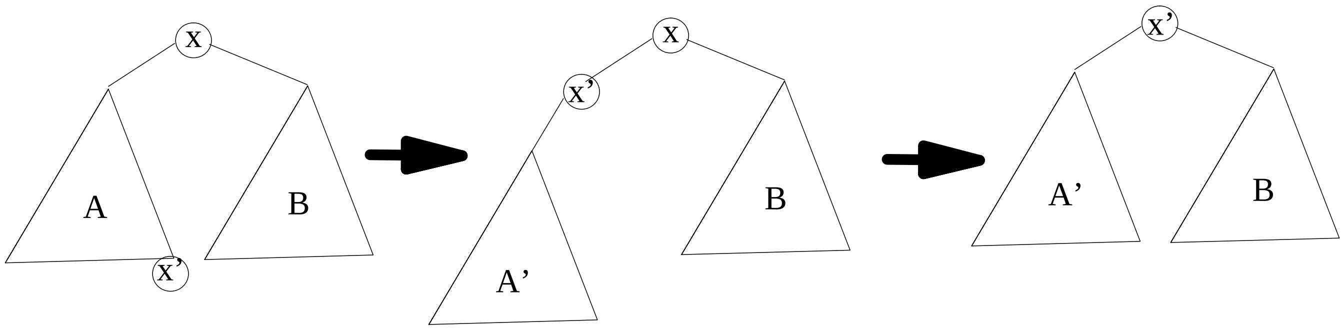


Single right rotation at y.
(+ symmetric case)

Insertion: Attach new leaf and then splay to root.

Deletion:

1. Access node x to delete, including splay to root.



2. Access predecessor x' in left subtree A and then splay to root of left subtree.

3. Take right subtree of x and make it the right subtree of x'

Amortized Analysis of Splaying for Retrieval:

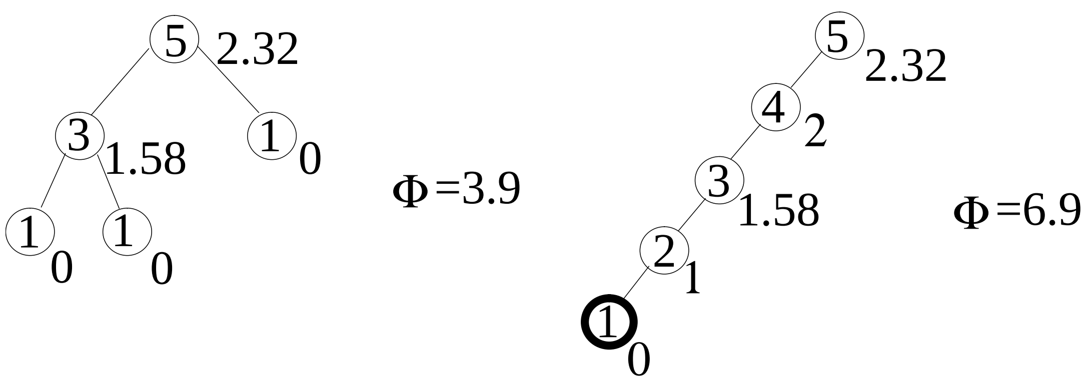
Actual cost (rotations) is 2 for zig-zig and zig-zag, but 1 for zig.

S(x) = number of nodes in subtree with x as root (''size'')

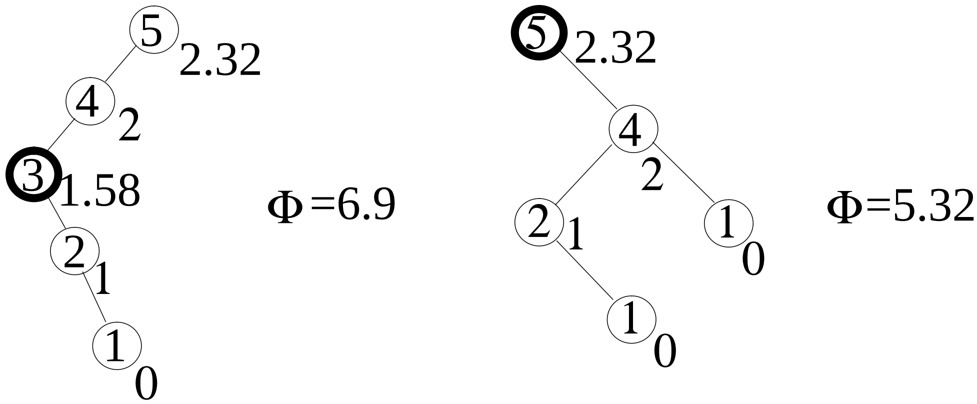
r(x) = $\log_2$ S(x) (''rank'')

$$\Phi(T) = \sum_{x \in T} r(x)$$

Examples:

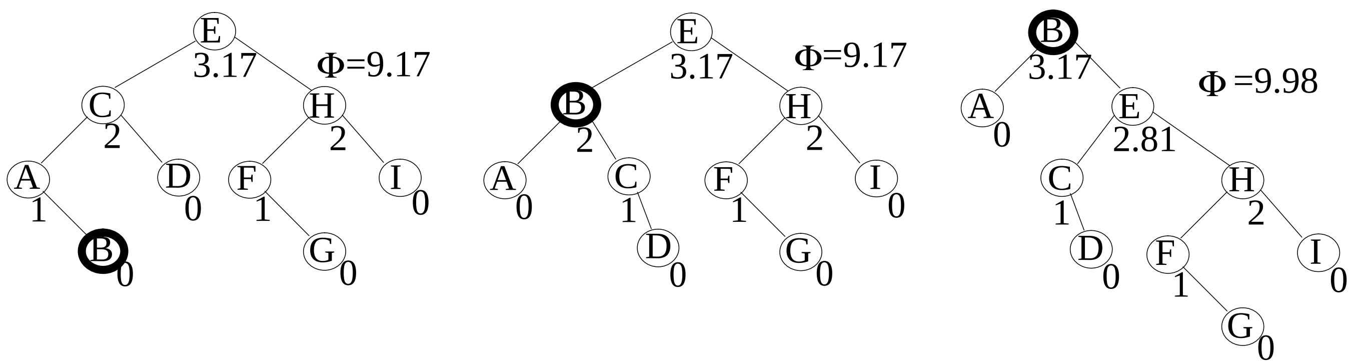


Now suppose that the leaf in the second example is retrieved. Two zig-zigs occur.



$$\sum \hat{}C_i = \sum C_i + \Phi(\text{After}) - \Phi(\text{Before}) = 4 + 5.32 - 6.9 = 2.42 \le 1 + 3\log_2 n$$

Another example of splaying. There will be a zig-zag and a zig.



$$\sum \hat{}C_i = \sum C_i + \Phi(\text{After}) - \Phi(\text{Before}) = 3 + 9.98 - 9.17 = 3.81 \le 1 + 3\log_2 n = 10.51$$

Compute amortized complexity of individual steps and then complete splaying sequence:

Lemma: If $\alpha > 0$, $\beta > 0$, $\alpha + \beta \le 1$, then $\log_2 \alpha + \log_2 \beta \le -2$.

Proof: $\log_2 \alpha + \log_2 \beta = \log_2 \alpha\beta$. $\alpha\beta$ is maximized when $\alpha = \beta = 1/2$, so $\max(\log_2 \alpha + \log_2 \beta) = -2$.

Access Lemma:

Suppose 1) x is node being splayed

2) subtree rooted by x has

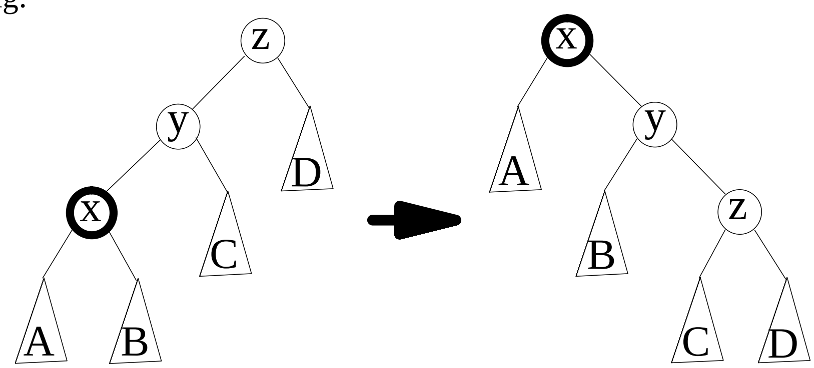
$S_{i-1}(x)$ and $r_{i-1}(x)$ before ith step

$S_i(x)$ and $r_i(x)$ after ith step

then $\hat{}C_i \le 3r_i(x) - 3r_{i-1}(x)$, except last step which has $\hat{}C_i \le 1 + 3r_i(x) - 3r_{i-1}(x)$

Proof: Proceeds by considering each of the three cases for splaying:

Zig-Zig:



$$\hat{C}_i = C_i + \Phi(T_i) - \Phi(T_{i-1})$$

$= 2 + r_i(x) + r_i(y) + r_i(z) - r_{i-1}(x) - r_{i-1}(y) - r_{i-1}(z)$ Potential only changes in this subtree

$= 2 + r_i(y) + r_i(z) - r_{i-1}(x) - r_{i-1}(y)$ $r_i(x) = r_{i-1}(z)$

$\leq 2 + r_i(y) + r_i(z) - 2r_{i-1}(x)$ $r_{i-1}(x) \leq r_{i-1}(y)$

(*) $\leq 2 + r_i(x) + r_i(z) - 2r_{i-1}(x)$ $r_i(y) \leq r_i(x)$

Let $\alpha = \dfrac{S_{i-1}(x)}{S_i(x)}$, $\beta = \dfrac{S_i(z)}{S_i(x)}$. $\alpha > 0$, $\beta > 0$. $\alpha + \beta = \dfrac{S_{i-1}(x) + S_i(z)}{S_i(x)} \leq 1$. (y is absent from numerator)

Lemma conditions are satisfied , so $\log_2 \alpha + \log_2 \beta \leq -2$. Applying logs to α and β gives:

$$r_{i-1}(x) + r_i(z) - 2r_i(x) \leq -2$$

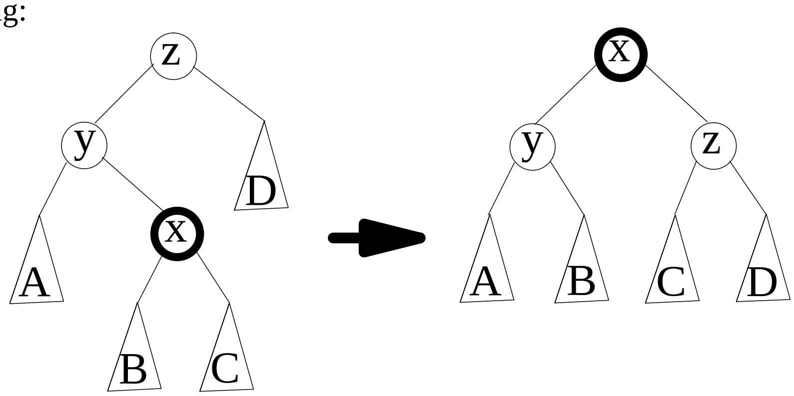
which may be rearranged as:

$$0 \leq 2r_i(x) - r_{i-1}(x) - r_i(z) - 2$$

Add this to (*) to obtain:

$$\hat{C}_i \leq 3r_i(x) - 3r_{i-1}(x)$$

Zig-Zag:



$\hat{C}_i = C_i + \Phi(T_i) - \Phi(T_{i-1})$

$= 2 + r_i(x) + r_i(y) + r_i(z) - r_{i-1}(x) - r_{i-1}(y) - r_{i-1}(z)$ Potential only changes in this subtree

$= 2 + r_i(y) + r_i(z) - r_{i-1}(x) - r_{i-1}(y)$ $r_i(x) = r_{i-1}(z)$

(**) $\leq 2 + r_i(y) + r_i(z) - 2r_{i-1}(x)$ $r_{i-1}(x) \leq r_{i-1}(y)$

Lemma may be applied by observing that $S_i(y) + S_i(z) \leq S_i(x)$ and thus

$$\frac{S_i(y)}{S_i(x)} + \frac{S_i(z)}{S_i(x)} \leq 1$$

By lemma, $\log_2\left(\frac{S_i(y)}{S_i(x)}\right) + \log_2\left(\frac{S_i(z)}{S_i(x)}\right) \leq -2$

$r_i(y) + r_i(z) - 2r_i(x) \leq -2$

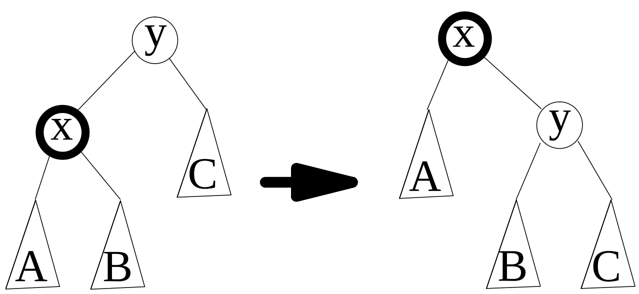
$r_i(y) + r_i(z) \leq 2r_i(x) - 2$ which can substitute into (**)

$\hat{C}_i \leq 2 + 2r_i(x) - 2 - 2r_{i-1}(x)$

$= 2r_i(x) - 2r_{i-1}(x)$

$\leq 3r_i(x) - 3r_{i-1}(x)$ Since $r_{i-1}(x) \leq r_i(x)$

Zig:



$\hat{C}_i = C_i + \Phi(T_i) - \Phi(T_{i-1})$

$= 1 + r_i(x) + r_i(y) - r_{i-1}(x) - r_{i-1}(y)$ Potential only changes in this subtree

$= 1 + r_i(y) - r_{i-1}(x)$ $r_{i-1}(y) = r_i(x)$

$\leq 1 + r_i(x) - r_{i-1}(x)$ $r_i(y) \leq r_i(x)$

$\leq 1 + 3r_i(x) - 3r_{i-1}(x)$ $r_{i-1}(x) \leq r_i(x)$

Total amortized cost for an entire splay sequence:

$$\sum_{i=1}^{m} \hat{C}_i = \sum_{i=1}^{m-1} \hat{C}_i + \hat{C}_m$$

$$\leq \sum_{i=1}^{m-1} \left(3r_i(x) - 3r_{i-1}(x)\right) + 1 + 3r_m(x) - 3r_{m-1}(x)$$

$$= 3r_{m-1}(x) - 3r_0(x) + 1 + 3r_m(x) - 3r_{m-1}(x)$$

$$= 1 + 3r_m(x) - 3r_0(x)$$

$$\leq 1 + 3r_m(x)$$

$$= 1 + 3\log_2 n$$

Since x is the root after final rotation.