Red-Black Trees

Balanced Insertions and Deletions
Operations in $O(\lg n)$

Def.: Red-Black Tree

A red-black tree is a binary search tree $+$ 1 bit per node: an attribute color, which is either red or black.

All leaves are empty (nil) and colored black.

- We use a single sentinel, nil[T], for all the leaves of red-black tree $T$.
- color[nil[T]] is black.
- The root’s parent is also nil[T].

All other attributes of binary search trees are inherited by red-black trees (key, left, right, and $p$). We don’t care about the key in nil[T].

Red-black properties

1. Every node is either red or black.
2. The root is black.
3. Every leaf (nil[T]) is black.
4. If a node is red, then both its children are black. (Hence no two reds in a row on a simple path from the root to a leaf.)
5. For each node, all paths from the node to descendant leaves contain the same number of black nodes.

Example:
Red-Black Trees

- Height
  Height of a red-black tree
  - Height of a node $x$: $bh(x)$ is the number of black nodes (including $n(\emptyset)$) on the path from $x$ to a leaf, not counting $x$. By property 5, black-height is well-defined.

- Two sub-lemmata
  Claim
  Any node with height $h$ has black-height $\geq h/2$.
  Proof By property 4, $\geq h/2$ nodes on the path from the node to a leaf are red. Hence $\geq h/2$ are black. $\blacksquare$

  Claim
  The subtree rooted at any node $x$ contains $\geq 2^{bh(x)} - 1$ internal nodes.
  Proof By induction on height of $x$.
  Basis: Height of $x = 0 \Rightarrow x$ is a leaf $\Rightarrow bh(x) = 0$. The subtree rooted at $x$ has 0 internal nodes. $2^0 - 1 = 0$.
  Inductive step: Let the height of $x$ be $h$ and $bh(x) = b$. Any child of $x$ has height $h-1$ and black-height either $b$ (if the child is red) or $b-1$ (if the child is black). By the inductive hypothesis, each child has $\geq 2^{bh(x)} - 1$ internal nodes. Thus, the subtree rooted at $x$ contains $\geq 2 \cdot (2^{bh(x)-1}) + 1 = 2^{bh(x)} - 1$ internal nodes. (The +1 is for $x$ itself.) $\blacksquare$

Red-Black Trees

- Another one
  Lemma
  A red-black tree with $n$ internal nodes has height $\leq 2\log(n + 1)$.
  Proof Let $h$ and $b$ be the height and black-height of the root, respectively. By the above two claims,
  
  $n \geq 2^h - 1 \geq 2^{h/2} - 1$.

  Adding 1 to both sides and then taking logs gives $\log(n + 1) \geq h/2$, which implies that $h \leq 2\log(n + 1)$. $\blacksquare$

Red-Black Trees

- Insertion & Deletion
  If we insert, what color to make the new node?
  - Black? Might violate property 5.
  
  If we delete, thus removing a node, what color was the node that was removed?
  - Red? OK, since we won’t have changed any black-heights, nor will we have created two red nodes in a row. Also, cannot cause a violation of property 2, since if the removed node was red, it could not have been the root.
  - Black? Could cause there to be two reds in a row (violating property 4), and can also cause a violation of property 5. Could also cause a violation of property 2, if the removed node was the root and its child—which becomes the new root—was red.
Red-Black Trees

Rotations

- Won’t upset the binary-search-tree property.
- Have both left rotation and right rotation. They are inverses of each other.
- A rotation takes a red-black-tree and a node within the tree.

![Rotation Diagram]

The Rotation Algorithm in Pseudo-Code

LEFT-ROTATE(T, x)

- Set y.
- right[y] ← left[y]
- Then right[y] = nil[T]
- Then p[y] = p[x]
- Then p[x] = y
- Then root[T] ← y
- Else if x = left[p[x]]
- Then left[p[x]] ← y
- Else right[p[x]] ← y
- left[x] ← z
- Then p[z] = y
- Else if y = nil[T]
- Then root[T] ← z
- Else if key[z] < key[y]
- Then left[z] ← z
- Else right[z] ← z
- left[z] = nil[T]
- right[z] = nil[T]
- Then color[z] = RED
- Then RB-INSERT-FIXUP(T, z)

Rotation Example:

![Rotation Example Diagram]

Insertion Pseudo-Code

RB-INSERT(T, z)

- Let key[z] = a
- Let p[z] = a
- Let root[T] = a
- Then color[z] = red
Red-Black Trees

Insertion

- RB-INSERT ends by coloring the new node \( z \) red.
- Then it calls RB-INSERT-FIXUP because we could have violated a red-black property.

Which property might be violated?
1. OK.
2. If \( z \) is the root, then there’s a violation. Otherwise, OK.
3. OK.
4. If \( p[z] \) is red, there’s a violation: both \( z \) and \( p[z] \) are red.
5. OK.

Remove the violation by calling RB-INSERT-FIXUP.

Fix the Problems

RedBlack Trees

Fix the Problems

Red-Black Trees

Fix the Problems

Red-Black Trees

Fix the Problems

Red-Black Trees

Insert another

Red-Black Trees

Insert another
**Red-Black Trees**

- **Fix it**

  ```
  RB-INSERT-FIXUP(T, z)
  while color[p[z]] = RED
    do if p[z] = left[p[p[z]]]
      then y ← right[p[z]]
      if color[y] = RED
        then color[y] ← BLACK
          color[p[z]] ← BLACK
          color[p[p[z]]] ← RED
          z ← p[z]
        else z ← p[z]
          LEFT-ROTATE(T, z)
          color[p[z]] ← BLACK
          color[p[p[z]]] ← RED
    else if z = left[p[z]]
      then z ← p[z]
        Right-Rotate(T, p[z])
        color[p[z]] ← BLACK
        color[p[p[z]]] ← RED
        Left-Rotate(T, p[z])
    else (same as then clause)
    color[root[T]] ← BLACK
  Nothing to do - root already black
  ```

- **Insert another**

  ```
  RB-INSERT(T, z)
  y ← nil[T]
  z ← root[T]
  while y ≠ nil[T]
    do if x = z
      then z ← x
        y ← right[x]
      else x ← right[x]
      p[z] ← y
      if y = nil[T]
        then root[T] ← z
        else if key[y] = key[z]
          then b[y] ← z
            else b[y] ← z
      if y = nil[T]
        then root[T] ← z
        else if key[y] = key[z]
          then b[y] ← z
            else b[y] ← z
      b[z] ← y
      if color[z] = RED
        then color[z] ← BLACK
        color[y] ← BLACK
        color[p[z]] ← RED
    else if z = left[p[z]]
      then z ← p[z]
        Right-Rotate(T, p[z])
        color[z] ← BLACK
        color[p[z]] ← RED
        Left-Rotate(T, p[z])
    else (same as then clause)
    if color[z] = RED
      then color[z] ← BLACK
        color[p[z]] ← RED
  color[root[T]] ← BLACK
  ```

Note: y is BLACK and z is NOT a left child of its parent, so we color p[z] BLACK, p[p[z]] RED and Left-Rotate on p[p[z]] and finish by coloring p[z] (the new root) BLACK.
Red-Black Trees

Here is the sequence: color \( p[z] \) BLACK, \( p[p[z]] \) RED and Left-Rotate on \( p[p[z]] \) and finish by coloring \( p[z] \) (the new root) BLACK.

\[
\begin{align*}
&\text{Insert } e. \text{ The problem, at this point, is that the number of black nodes along each path must change. Look back at the code: does this apply?} \\
&\text{Since } y \text{ is RED, set } color[p[z]] \text{ to BLACK, } color[y] \text{ to BLACK, color} \\
&\text{ } p[p[z]] \text{ to RED, } z \text{ to } p[z]: \text{ since now } color[p[z]] = color[p[root[T]]] = BLACK, \text{ the while loop ends. Set } z = \\
&\text{root[T]} \text{ to BLACK.}
\end{align*}
\]
Add f: Do we use the left or right code? We start using the left code.

Red-Black Trees

Why does this work?

Loop invariant:
At the start of each iteration of the while loop,

- z is red.
- There is at most one red-black violation.
  - Property 2: z is a red root, or
  - Property 4: z and p[z] are both red.

The Induction

Initialization: We’ve already seen why the loop invariant holds initially.
Termination: The loop terminates because p[z] is black. Hence, property 4 is OK. Only property 2 might be violated, and the last line fixes it.
Maintenance: We drop out when z is the root (since then p[z] is the sentinel nil[], which is black). When we start the loop body, the only violation is of property 4.

There are 6 cases, 3 of which are symmetric to the other 3. The cases are not mutually exclusive. We’ll consider cases in which p[z] is a left child.

Case 1: y is red

1. z is red.
2. z is a right child.
3. p[z] is red.
4. If z has a left child.
   - p[z] (z’s grandparent) must be black, since z and p[z] are both red and there are no other violations of property 4.
   - Make p[z] and y black. ⇒ now z and p[z] are not both red. But property 4 might now be violated.
   - Make p[z] red ⇒ restores property 5.
   - The next iteration has p[z] as the new z (i.e., z moves up 2 levels).
**Case 2**

Case 2: When $y$ is black, $z$ is a right child.

- Left rotate around $p(z)$ to make $z$ a left child, and both $z$ and $p(z)$ are red.
- Takes us immediately to case 3.

**Case 3**

Case 3: When $y$ is black, $z$ is a left child.

- Make $p(z)$ black and $p[p(z)]$ red.
- Then right rotate on $p[p(z)]$.
- No longer have 2 reds in a row.
- $p[z]$ is now black ⇒ no more iterations.

**Finish the last insertion (modified).**

Now add $g$:

- $nil[T]$
Now for the deletion: delete g.

Red-Black Trees

Chasing the code with the picture (or vice-versa, your pick), we end up just removing the node labeled g. Nothing else changes, other than the left pointer out of the node labeled h, which goes to nil[T], g just disappears without triggering any adjustments: it was a RED node, so the number of BLACK nodes along that path did not change. Notice that no labels or other data are copied, since the node we are deleting is the very last in a chain - the next nodes are the sentinel leaves.

Deleting any other node will trigger more complicated readjustments.

Remove f:

This is also easy, since removing f involves finding its successor (g), re-attaching the parent of g to a sentinel (rather than g), and copying the contents of g into f. Since the node actually removed (g) is RED, nothing needs to be done.

We now try to remove h - this is a BLACK node and, because it has only one child, it will actually be removed. This will finally trigger RB-Delete-Fixup(T, x).
Red-Black Trees

Before the call to RB-Delete-Fixup(T, x) We have:

Here is the pseudo-code

Red-Black Trees

Since x is neither the root, nor is it BLACK (it took the place of a BLACK node that was removed), the code tells us to just color it BLACK.

How about removing e from the original tree? Since the deletion itself does not worry about color, we just remove e, and x is the nil[T] sentinel. Note that the parent of the sentinel is now f.

Furthermore, x is not the root, and its color is BLACK.
Red-Black Trees

Compare the code and the tree:

```c
Red-Black Trees

We have a number of cases to take care of - 8 (precisely but not mutually exclusive). We will look at 4 of them, leaving the 4 symmetric ones as exercises.

We observe that splicing out ANY RED node requires no fix-up: it does not alter the black height of any node; it does not introduce any pair of parent-child RED nodes; and it does not change the root.

The only case where fixing up will be needed is when the spliced out node is BLACK: that will alter the black height of its ancestors, and may violate the requirement that there be no pair of parent-child RED nodes.

How can the splicing out of a black node \( y \) effect the result?

1. \( y \) had been the root and a red child of \( y \) becomes (physically) the new root, violating property 2. This can occur only if we have the configuration on the left, or its symmetric counterpart.

   The fix is to just color the new root BLACK, but we will examine it in the context of RB-Delete-Fixup.

2. Both \( x \) and \( p[x] = p[y] \) are RED. Property 4 is violated (red has only black children).

3. If \( y \) was BLACK, its removal from any path will cause any path that contained it to have one fewer BLACK nodes. Property 5 (all paths from node down have same number of black nodes) is violated by any ancestor of \( y \) in the tree.

   How do we solve the problem? pretend that the node \( x \) has an extra BLACK. Then all is well (properties 4 & 5), and we have to figure out where to unload this extra black…
Red-Black Trees

Note that, in RB-Delete-Fixup(T, x), with y the node actually deleted, x is:
   a. y’s sole non-sentinel child before y was spliced out
   b. the sentinel itself, if y had no children.

Also, after deletion, \( p[x] = p[y] \)

Red-Black Trees

With y black, we could violate the properties:
1. Every node is either red or black. NO
2. The root is black: YES, if y is the root and x is red.
3. Every leaf is black. NO
4. If a node is red then both of its children are black: YES, if \( p[y] \) and x are both red (from the left hand example).
5. For each node, all paths from the node to descendant leaves contain the same number of black nodes: YES, any path that contained y now has one fewer black nodes.

We can fix 5 by giving x "an extra black" from its deleted parent - then the count of black nodes if “fixed”: we will push the extra black around until we can safely unload it...

Red-Black Trees

Note: we have violated Property 1, since we now have nodes that are neither red nor black: x is doubly-black if x was black; it is red&black if it was red. Note that color[x] is still just RED or BLACK - the extra black comes from pointing to it (which means we have to unload its extra blackness before we stop pointing to it).

IDEA: move the extra black up the tree until:
   - x points to a red&black node --> turn it into a red one
   - x points to the root --> just remove the extra black
   - perform rotations and recolorings

The first two points tell you when it’s safe to unload the extra black; the last one tells you how to move it up. That’s where we go now.

Red-Black Trees

We start by assuming x is the left child of its parent and that w is its sibling:

Case 1: w is red

- w must have black children.
- Make w black and \( p[w-1] \) red.
- Then left rotate on \( p[x] \).
- New sibling of x was a child of w before rotation \( \Rightarrow \) must be black.

The same black path conditions will be satisfied at the end of the recoloring and rotation as at the beginning.
Red-Black Trees

The tree rooted at A has one more black node than the trees rooted at B and C. We don't know the color of the children of C - assume they are both black (other cases later): A, B, D are, respectively, the old A, B, C.

Case 2: w is black and both of w's children are black

[- Diagram image -]

[Node with gray outline is of unknown color, denoted by c.]

- Take 1 black off x (⇒ singly black) and off w (⇒ red).
- Move that black to p[x].
- Do the next iteration with p[x] as the new x.
- If entered this case from case 1, then p[x] was red ⇒ new x is red & black ⇒ color attribute of new x is RED ⇒ loop terminates. Then new x is made black in the last line.

Next case: doesn't quite end, since we can't yet get rid of the extra black on A. You just move to the case where the colors of the children of w are swapped (γ is black)

Case 3: w is black, w's left child is red, and w's right child is black

[- Diagram image -]

- Make w red and w's left child black.
- Then right rotate on w.
- New sibling w of x is black with a red right child ⇒ case 4.

Finally:

Case 4: w is black, w's left child is black, and w's right child is red

[- Diagram image -]

[Now there are two nodes of unknown colors, denoted by c and c'.]

- Make w be p[x]'s color (c).
- Make p[x] black and w's right child black.
- Then left rotate on p[x].
- Remove the extra black on x (⇒ x is now singly black) without violating any red-black properties.
- All done. Setting x to root causes the loop to terminate.
Red-Black Trees