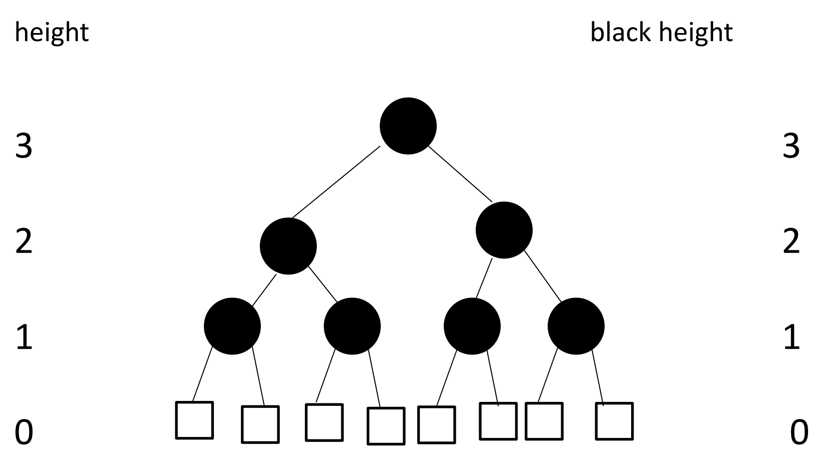
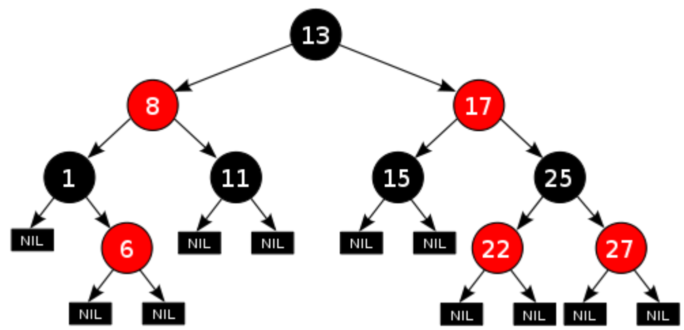
Proof of Log Height of RB Trees

- For the purpose of this proof, we will add null nodes to the tree.
- We will show that a red black tree which contains *n* internal nodes has a height of O(log(n)).
- An internal node is a non-null node.
- Definitions:
  - h(*v*) = height of subtree rooted at node *v*
  - bh(*v*) = the number of black nodes (not counting *v* if it is black) from *v* to any null leaf in the subtree (called the black-height).

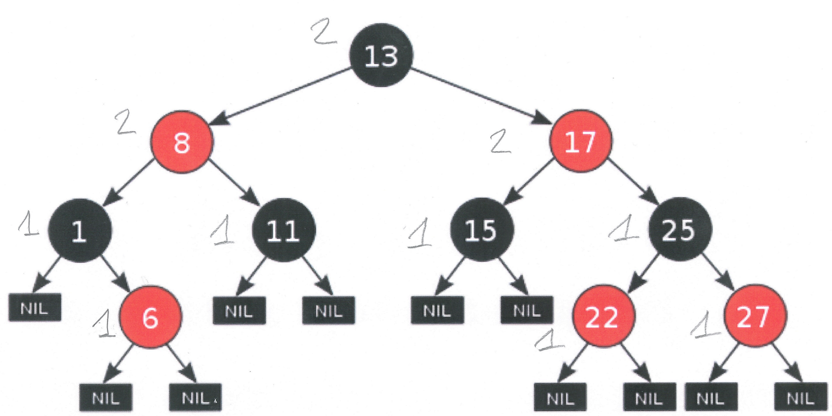
Examples of height and black height (assuming a tree with only black nodes) (Recall that we added null nodes)



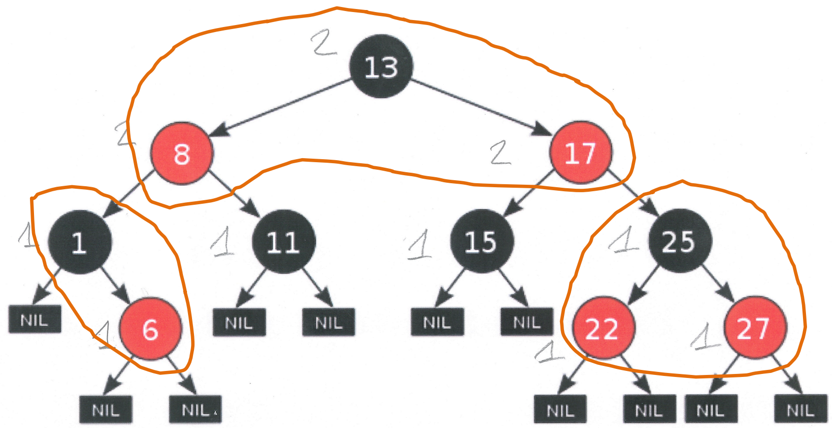
Class exercise of determining height and black height of nodes



Black heights of nodes



Black heights of nodes

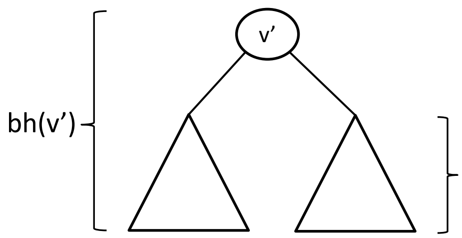


Proof of Log Height of RB Tree with **ONLY** black nodes

- Assume an all black tree for now!
- We perform an induction over the height of a tree
- Lemma: A sub-tree rooted at node *v* has at least $2^{bh(v)}$ - 1 internal (i.e. non-null) nodes.
- Base case: A tree of height h=0.
  - Such a tree consists of a null node: ☐
  - bh(v) = 0
  - 2^0- 1 = 0

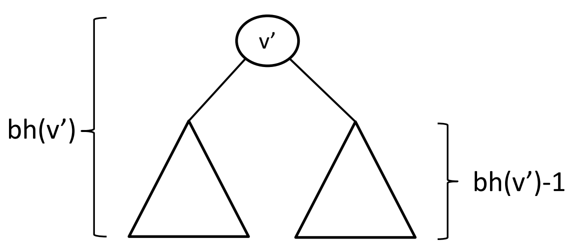
Proof of Log Height of RB Tree with **ONLY** black nodes

- Inductive Hypothesis: Assume that a tree rooted at node *v* with height k has at least $2^{bh(v)} - 1$ internal nodes.
- Inductive step: Show that a tree of height k + 1 has at least $2^{bh(v')} - 1$ internal nodes, where v' is the new root.
  - A red-black tree with height k+1 consisting of only black nodes can only be built like this:

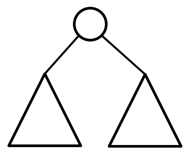


Proof of Log Height of RB Tree with **ONLY** black nodes

- Inductive Hypothesis: Assume that a tree rooted at node *v* with height k has at least $2^{bh(v)} - 1$ internal nodes.
- Inductive step: Show that a tree of height k + 1 has at least $2^{bh(v')} - 1$ internal nodes, where v' is the new root.
  - A red-black tree with height k+1 consisting of only black nodes can only be built like this:



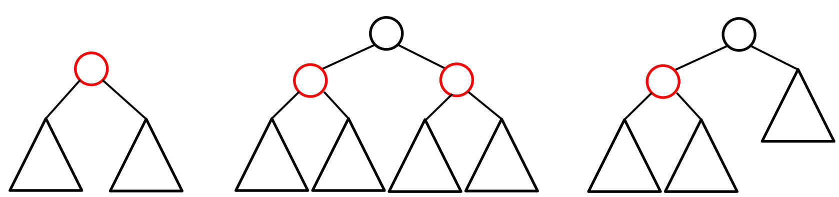
Proof of Log Height of RB Trees with **ONLY** black nodes



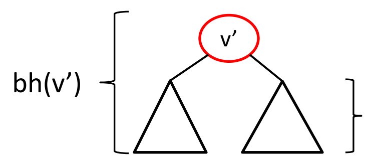
- By the inductive hypothesis, each sub-tree has at least $2^{bh(v')-1} - 1$ internal nodes.
- This gives an overall total of:
  $2 * (2^{bh(v')-1} - 1) + 1$ nodes, i.e. $2^{bh(v')} - 1$

Proof of Log Height of RB Trees

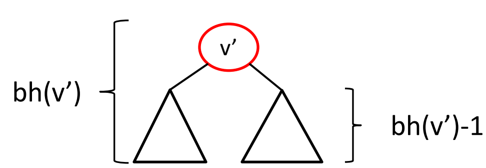
- Now assume that we have red and black nodes
- The base case and inductive hypothesis remain the same.
- For the inductive step we have three additional cases to consider:



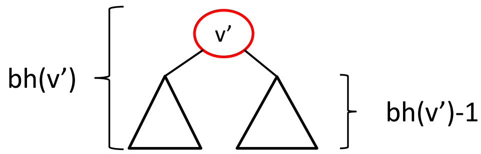
Case 1: New red root, Black sub roots



Case 1: New red root, Black sub roots

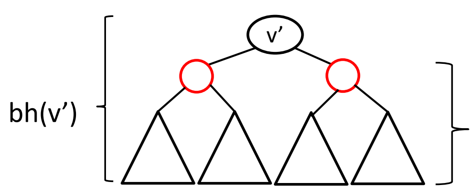


Case 1: New red root, Black sub roots

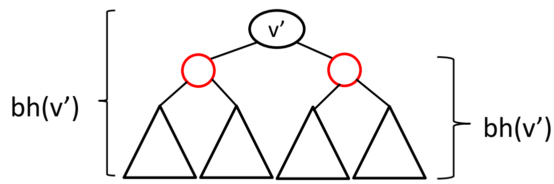


– Since the new root is red, the sub-roots must be black
– By the inductive hypothesis, each sub-tree has at least $2^{bh(v')-1} - 1$ internal nodes.
– This gives an overall total of:
$2 * (2^{bh(v')-1} - 1) + 1$ nodes, i.e. $2^{bh(v')} - 1$

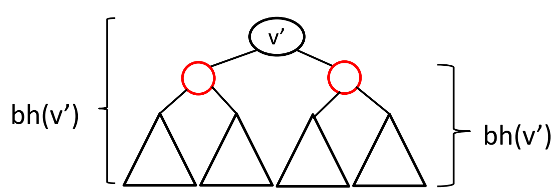
Case 2: New black root, Red sub-roots



Case 2: New black root, Red sub-roots

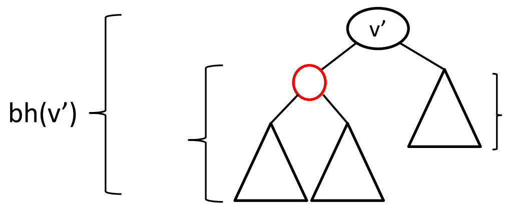


Case 2: New black root, Red sub-roots

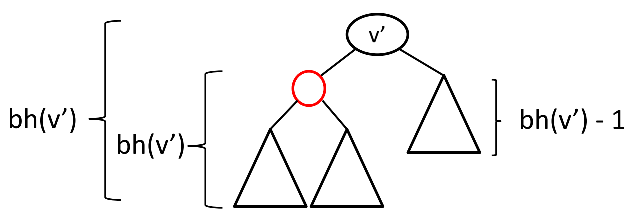


- The black height of the new tree does not increase.
- By the inductive hypothesis, each sub-tree has at least $2^{bh(v')} - 1$ internal nodes.
- This gives an overall total of:
  $2 * (2^{bh(v')} - 1) + 1$ nodes, i.e. $2^{bh(v')+1} - 1 > 2^{bh(v')} - 1$

Case 3: New black root, red and black sub-roots



Case 3: New black root, red and black sub-roots

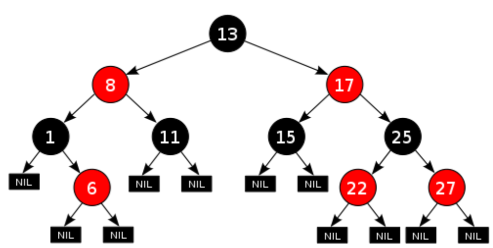


- Based on the left sub-tree alone, we have: $(2^{bh(v')} - 1)$ nodes
- If you want to by fussy, we have the new root as well as $(2^{bh(v')-1} - 1)$ additional nodes

Proof of Log Height of RB Trees

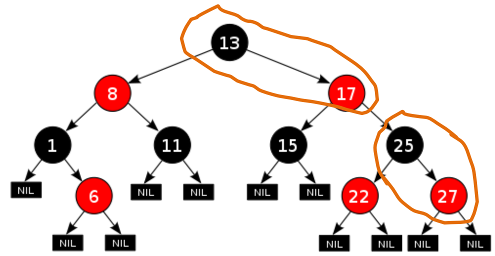
- In all four cases, the new tree has at least $2^{bh(v')} - 1$ internal nodes
- The lemma holds.

Proof of Log Height of RB Trees



- Let's look at the root, we denote its height by h_{root}
- We know that on a path from the root to a null node, at most half the nodes can be red.

Proof of Log Height of RB Trees



- Let's look at the root, we denote its height by h_{root}
- We know that on a path from the root to a null node, at most half the nodes can be red.
- As such, the black height of the root is at least $h_{root} / 2$

Proof of Log Height of RB Trees

- As proven by the lemma, the number of nodes *n* in a RB Tree is >= $2^{bhroot} – 1 >= 2^{hroot/2} – 1$
- Adding 1 on both sides: $n + 1 >= 2^{hroot/2}$
- Taking the log: $\log_2(n+1) >= h_{root}/2$
- In other words: $h <= 2* \log_2(n+1)$