

★ 12.4 Randomly built binary search trees

We have shown that all the basic operations on a binary search tree run in $O(h)$ time, where h is the height of the tree. The height of a binary search tree varies, however, as items are inserted and deleted. If, for example, the items are inserted in strictly increasing order, the tree will be a chain with height $n - 1$. On the other hand, Exercise B.5-4 shows that $h \geq \lfloor \lg n \rfloor$. As with quicksort, we can show that the behavior of the average case is much closer to the best case than the worst case.

Unfortunately, little is known about the average height of a binary search tree when both insertion and deletion are used to create it. When the tree is created by insertion alone, the analysis becomes more tractable. Let us therefore define a **randomly built binary search tree** on n keys as one that arises from inserting the keys in random order into an initially empty tree, where each of the $n!$ permutations of the input keys is equally likely. (Exercise 12.4-3 asks you to show that this notion is different from assuming that every binary search tree on n keys is equally likely.) In this section, we shall show that the expected height of a randomly built binary search tree on n keys is $O(\lg n)$. We assume that all keys are distinct.

We start by defining three random variables that help measure the height of a randomly built binary search tree. We denote the height of a randomly built binary search on n keys by X_n , and we define the **exponential height** $Y_n = 2^{X_n}$. When we build a binary search tree on n keys, we choose one key as that of the root, and we let R_n denote the random variable that holds this key's rank within the set of n keys. The value of R_n is equally likely to be any element of the set $\{1, 2, \dots, n\}$. If $R_n = i$, then the left subtree of the root is a randomly built binary search tree on $i - 1$ keys, and the right subtree is a randomly built binary search tree on $n - i$ keys. Because the height of a binary tree is one more than the larger of the heights of the two subtrees of the root, the exponential height of a binary tree is twice the larger of the exponential heights of the two subtrees of the root. If we know that $R_n = i$, we therefore have that

$$Y_n = 2 \cdot \max(Y_{i-1}, Y_{n-i}) .$$

As base cases, we have $Y_1 = 1$, because the exponential height of a tree with 1 node is $2^0 = 1$ and, for convenience, we define $Y_0 = 0$.

Next we define indicator random variables $Z_{n,1}, Z_{n,2}, \dots, Z_{n,n}$, where

$$Z_{n,i} = \mathbf{I}\{R_n = i\} .$$

Because R_n is equally likely to be any element of $\{1, 2, \dots, n\}$, we have that $\Pr\{R_n = i\} = 1/n$ for $i = 1, 2, \dots, n$, and hence, by Lemma 5.1,

$$\mathbf{E}[Z_{n,i}] = 1/n , \tag{12.1}$$

for $i = 1, 2, \dots, n$. Because exactly one value of $Z_{n,i}$ is 1 and all others are 0, we also have

$$Y_n = \sum_{i=1}^n Z_{n,i} (2 \cdot \max(Y_{i-1}, Y_{n-i})) .$$

We will show that $E[Y_n]$ is polynomial in n , which will ultimately imply that $E[X_n] = O(\lg n)$.

The indicator random variable $Z_{n,i} = I\{R_n = i\}$ is independent of the values of Y_{i-1} and Y_{n-i} . Having chosen $R_n = i$, the left subtree, whose exponential height is Y_{i-1} , is randomly built on the $i - 1$ keys whose ranks are less than i . This subtree is just like any other randomly built binary search tree on $i - 1$ keys. Other than the number of keys it contains, this subtree's structure is not affected at all by the choice of $R_n = i$; hence the random variables Y_{i-1} and $Z_{n,i}$ are independent. Likewise, the right subtree, whose exponential height is Y_{n-i} , is randomly built on the $n - i$ keys whose ranks are greater than i . Its structure is independent of the value of R_n , and so the random variables Y_{n-i} and $Z_{n,i}$ are independent. Hence,

$$\begin{aligned} E[Y_n] &= E \left[\sum_{i=1}^n Z_{n,i} (2 \cdot \max(Y_{i-1}, Y_{n-i})) \right] \\ &= \sum_{i=1}^n E[Z_{n,i} (2 \cdot \max(Y_{i-1}, Y_{n-i}))] \quad (\text{by linearity of expectation}) \\ &= \sum_{i=1}^n E[Z_{n,i}] E[2 \cdot \max(Y_{i-1}, Y_{n-i})] \quad (\text{by independence}) \\ &= \sum_{i=1}^n \frac{1}{n} \cdot E[2 \cdot \max(Y_{i-1}, Y_{n-i})] \quad (\text{by equation (12.1)}) \\ &= \frac{2}{n} \sum_{i=1}^n E[\max(Y_{i-1}, Y_{n-i})] \quad (\text{by equation (C.21)}) \\ &\leq \frac{2}{n} \sum_{i=1}^n (E[Y_{i-1}] + E[Y_{n-i}]) \quad (\text{by Exercise C.3-4}) . \end{aligned}$$

Each term $E[Y_0], E[Y_1], \dots, E[Y_{n-1}]$ appears twice in the last summation, once as $E[Y_{i-1}]$ and once as $E[Y_{n-i}]$, and so we have the recurrence

$$E[Y_n] \leq \frac{4}{n} \sum_{i=0}^{n-1} E[Y_i] . \tag{12.2}$$

Using the substitution method, we will show that for all positive integers n , the recurrence (12.2) has the solution

$$E[Y_n] \leq \frac{1}{4} \binom{n+3}{3} .$$

In doing so, we will use the identity

$$\sum_{i=0}^{n-1} \binom{i+3}{3} = \binom{n+3}{4}. \quad (12.3)$$

(Exercise 12.4-1 asks you to prove this identity.)

For the base case, we verify that the bound

$$1 = Y_1 = E[Y_1] \leq \frac{1}{4} \binom{1+3}{3} = 1$$

holds. For the substitution, we have that

$$\begin{aligned} E[Y_n] &\leq \frac{4}{n} \sum_{i=0}^{n-1} E[Y_i] \\ &\leq \frac{4}{n} \sum_{i=0}^{n-1} \frac{1}{4} \binom{i+3}{3} \quad (\text{by the inductive hypothesis}) \\ &= \frac{1}{n} \sum_{i=0}^{n-1} \binom{i+3}{3} \\ &= \frac{1}{n} \binom{n+3}{4} \quad (\text{by equation (12.3)}) \\ &= \frac{1}{n} \cdot \frac{(n+3)!}{4! (n-1)!} \\ &= \frac{1}{4} \cdot \frac{(n+3)!}{3! n!} \\ &= \frac{1}{4} \binom{n+3}{3}. \end{aligned}$$

We have bounded $E[Y_n]$, but our ultimate goal is to bound $E[X_n]$. As Exercise 12.4-4 asks you to show, the function $f(x) = 2^x$ is convex (see page 1109). Therefore, we can apply Jensen's inequality (C.25), which says that

$$2^{E[X_n]} \leq E[2^{X_n}] = E[Y_n],$$

to derive that

$$\begin{aligned} 2^{E[X_n]} &\leq \frac{1}{4} \binom{n+3}{3} \\ &= \frac{1}{4} \cdot \frac{(n+3)(n+2)(n+1)}{6} \\ &= \frac{n^3 + 6n^2 + 11n + 6}{24}. \end{aligned}$$

Taking logarithms of both sides gives $E[X_n] = O(\lg n)$. Thus, we have proven the following:

Theorem 12.4

The expected height of a randomly built binary search tree on n keys is $O(\lg n)$. ■

Exercises**12.4-1**

Prove equation (12.3).

12.4-2

Describe a binary search tree on n nodes such that the average depth of a node in the tree is $\Theta(\lg n)$ but the height of the tree is $\omega(\lg n)$. Give an asymptotic upper bound on the height of an n -node binary search tree in which the average depth of a node is $\Theta(\lg n)$.

12.4-3

Show that the notion of a randomly chosen binary search tree on n keys, where each binary search tree of n keys is equally likely to be chosen, is different from the notion of a randomly built binary search tree given in this section. (*Hint*: List the possibilities when $n = 3$.)

12.4-4

Show that the function $f(x) = 2^x$ is convex.

12.4-5 ★

Consider RANDOMIZED-QUICKSORT operating on a sequence of n input numbers. Prove that for any constant $k > 0$, all but $O(1/n^k)$ of the $n!$ input permutations yield an $O(n \lg n)$ running time.

Problems**12-1 Binary search trees with equal keys**

Equal keys pose a problem for the implementation of binary search trees.

- a. What is the asymptotic performance of TREE-INSERT when used to insert n items with identical keys into an initially empty binary search tree?

We propose to improve TREE-INSERT by testing before line 5 whether or not $key[z] = key[x]$ and by testing before line 11 whether or not $key[z] = key[y]$. If equality holds, we implement one of the following strategies. For each strategy, find the asymptotic performance of inserting n items with identical keys into an initially empty binary search tree. (The strategies are described for line 5, in which

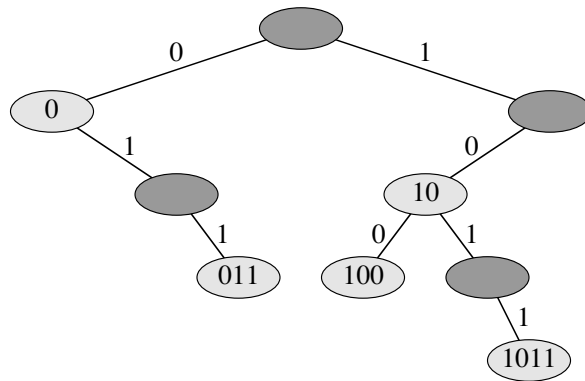


Figure 12.5 A radix tree storing the bit strings 1011, 10, 011, 100, and 0. Each node's key can be determined by traversing the path from the root to that node. There is no need, therefore, to store the keys in the nodes; the keys are shown here for illustrative purposes only. Nodes are heavily shaded if the keys corresponding to them are not in the tree; such nodes are present only to establish a path to other nodes.

we compare the keys of z and x . Substitute y for x to arrive at the strategies for line 11.)

- b.* Keep a boolean flag $b[x]$ at node x , and set x to either $left[x]$ or $right[x]$ based on the value of $b[x]$, which alternates between FALSE and TRUE each time x is visited during insertion of a node with the same key as x .
- c.* Keep a list of nodes with equal keys at x , and insert z into the list.
- d.* Randomly set x to either $left[x]$ or $right[x]$. (Give the worst-case performance and informally derive the average-case performance.)

12-2 Radix trees

Given two strings $a = a_0a_1 \dots a_p$ and $b = b_0b_1 \dots b_q$, where each a_i and each b_j is in some ordered set of characters, we say that string a is **lexicographically less than** string b if either

1. there exists an integer j , where $0 \leq j \leq \min(p, q)$, such that $a_i = b_i$ for all $i = 0, 1, \dots, j-1$ and $a_j < b_j$, or
2. $p < q$ and $a_i = b_i$ for all $i = 0, 1, \dots, p$.

For example, if a and b are bit strings, then $10100 < 10110$ by rule 1 (letting $j = 3$) and $10100 < 101000$ by rule 2. This is similar to the ordering used in English-language dictionaries.

The **radix tree** data structure shown in Figure 12.5 stores the bit strings 1011, 10, 011, 100, and 0. When searching for a key $a = a_0a_1 \dots a_p$, we go left at a node

of depth i if $a_i = 0$ and right if $a_i = 1$. Let S be a set of distinct binary strings whose lengths sum to n . Show how to use a radix tree to sort S lexicographically in $\Theta(n)$ time. For the example in Figure 12.5, the output of the sort should be the sequence 0, 011, 10, 100, 1011.

12-3 Average node depth in a randomly built binary search tree

In this problem, we prove that the average depth of a node in a randomly built binary search tree with n nodes is $O(\lg n)$. Although this result is weaker than that of Theorem 12.4, the technique we shall use reveals a surprising similarity between the building of a binary search tree and the running of RANDOMIZED-QUICKSORT from Section 7.3.

We define the **total path length** $P(T)$ of a binary tree T as the sum, over all nodes x in T , of the depth of node x , which we denote by $d(x, T)$.

a. Argue that the average depth of a node in T is

$$\frac{1}{n} \sum_{x \in T} d(x, T) = \frac{1}{n} P(T) .$$

Thus, we wish to show that the expected value of $P(T)$ is $O(n \lg n)$.

b. Let T_L and T_R denote the left and right subtrees of tree T , respectively. Argue that if T has n nodes, then

$$P(T) = P(T_L) + P(T_R) + n - 1 .$$

c. Let $P(n)$ denote the average total path length of a randomly built binary search tree with n nodes. Show that

$$P(n) = \frac{1}{n} \sum_{i=0}^{n-1} (P(i) + P(n-i-1) + n-1) .$$

d. Show that $P(n)$ can be rewritten as

$$P(n) = \frac{2}{n} \sum_{k=1}^{n-1} P(k) + \Theta(n) .$$

e. Recalling the alternative analysis of the randomized version of quicksort given in Problem 7-2, conclude that $P(n) = O(n \lg n)$.

At each recursive invocation of quicksort, we choose a random pivot element to partition the set of elements being sorted. Each node of a binary search tree partitions the set of elements that fall into the subtree rooted at that node.

- f.* Describe an implementation of quicksort in which the comparisons to sort a set of elements are exactly the same as the comparisons to insert the elements into a binary search tree. (The order in which comparisons are made may differ, but the same comparisons must be made.)

12-4 Number of different binary trees

Let b_n denote the number of different binary trees with n nodes. In this problem, you will find a formula for b_n , as well as an asymptotic estimate.

- a.* Show that $b_0 = 1$ and that, for $n \geq 1$,

$$b_n = \sum_{k=0}^{n-1} b_k b_{n-1-k} .$$

- b.* Referring to Problem 4-5 for the definition of a generating function, let $B(x)$ be the generating function

$$B(x) = \sum_{n=0}^{\infty} b_n x^n .$$

Show that $B(x) = xB(x)^2 + 1$, and hence one way to express $B(x)$ in closed form is

$$B(x) = \frac{1}{2x} (1 - \sqrt{1 - 4x}) .$$

The **Taylor expansion** of $f(x)$ around the point $x = a$ is given by

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} (x - a)^k ,$$

where $f^{(k)}(x)$ is the k th derivative of f evaluated at x .

- c.* Show that

$$b_n = \frac{1}{n+1} \binom{2n}{n}$$

(the n th **Catalan number**) by using the Taylor expansion of $\sqrt{1 - 4x}$ around $x = 0$. (If you wish, instead of using the Taylor expansion, you may use the generalization of the binomial expansion (C.4) to nonintegral exponents n , where for any real number n and for any integer k , we interpret $\binom{n}{k}$ to be $n(n-1) \cdots (n-k+1)/k!$ if $k \geq 0$, and 0 otherwise.)

d. Show that

$$b_n = \frac{4^n}{\sqrt{\pi} n^{3/2}} (1 + O(1/n)) .$$

Chapter notes

Knuth [185] contains a good discussion of simple binary search trees as well as many variations. Binary search trees seem to have been independently discovered by a number of people in the late 1950's. Radix trees are often called tries, which comes from the middle letters in the word retrieval. They are also discussed by Knuth [185].

Section 15.5 will show how to construct an optimal binary search tree when search frequencies are known prior to constructing the tree. That is, given the frequencies of searching for each key and the frequencies of searching for values that fall between keys in the tree, we construct a binary search tree for which a set of searches that follows these frequencies examines the minimum number of nodes.

The proof in Section 12.4 that bounds the expected height of a randomly built binary search tree is due to Aslam [23]. Martínez and Roura [211] give randomized algorithms for insertion into and deletion from binary search trees in which the result of either operation is a random binary search tree. Their definition of a random binary search tree differs slightly from that of a randomly built binary search tree in this chapter, however.

13 Red-Black Trees

Chapter 12 showed that a binary search tree of height h can implement any of the basic dynamic-set operations—such as SEARCH, PREDECESSOR, SUCCESSOR, MINIMUM, MAXIMUM, INSERT, and DELETE—in $O(h)$ time. Thus, the set operations are fast if the height of the search tree is small; but if its height is large, their performance may be no better than with a linked list. Red-black trees are one of many search-tree schemes that are “balanced” in order to guarantee that basic dynamic-set operations take $O(\lg n)$ time in the worst case.

13.1 Properties of red-black trees

A **red-black tree** is a binary search tree with one extra bit of storage per node: its *color*, which can be either RED or BLACK. By constraining the way nodes can be colored on any path from the root to a leaf, red-black trees ensure that no such path is more than twice as long as any other, so that the tree is approximately **balanced**.

Each node of the tree now contains the fields *color*, *key*, *left*, *right*, and *p*. If a child or the parent of a node does not exist, the corresponding pointer field of the node contains the value NIL. We shall regard these NIL’s as being pointers to external nodes (leaves) of the binary search tree and the normal, key-bearing nodes as being internal nodes of the tree.

A binary search tree is a red-black tree if it satisfies the following **red-black properties**:

1. Every node is either red or black.
2. The root is black.
3. Every leaf (NIL) is black.
4. If a node is red, then both its children are black.
5. For each node, all paths from the node to descendant leaves contain the same number of black nodes.

Figure 13.1(a) shows an example of a red-black tree.

As a matter of convenience in dealing with boundary conditions in red-black tree code, we use a single sentinel to represent NIL (see page 206). For a red-black tree T , the sentinel $nil[T]$ is an object with the same fields as an ordinary node in the tree. Its *color* field is BLACK, and its other fields—*p*, *left*, *right*, and *key*—can be set to arbitrary values. As Figure 13.1(b) shows, all pointers to NIL are replaced by pointers to the sentinel $nil[T]$.

We use the sentinel so that we can treat a NIL child of a node x as an ordinary node whose parent is x . Although we instead could add a distinct sentinel node for each NIL in the tree, so that the parent of each NIL is well defined, that approach would waste space. Instead, we use the one sentinel $nil[T]$ to represent all the NIL's—all leaves and the root's parent. The values of the fields *p*, *left*, *right*, and *key* of the sentinel are immaterial, although we may set them during the course of a procedure for our convenience.

We generally confine our interest to the internal nodes of a red-black tree, since they hold the key values. In the remainder of this chapter, we omit the leaves when we draw red-black trees, as shown in Figure 13.1(c).

We call the number of black nodes on any path from, but not including, a node x down to a leaf the **black-height** of the node, denoted $bh(x)$. By property 5, the notion of black-height is well defined, since all descending paths from the node have the same number of black nodes. We define the black-height of a red-black tree to be the black-height of its root.

The following lemma shows why red-black trees make good search trees.

Lemma 13.1

A red-black tree with n internal nodes has height at most $2 \lg(n + 1)$.

Proof We start by showing that the subtree rooted at any node x contains at least $2^{bh(x)} - 1$ internal nodes. We prove this claim by induction on the height of x . If the height of x is 0, then x must be a leaf ($nil[T]$), and the subtree rooted at x indeed contains at least $2^{bh(x)} - 1 = 2^0 - 1 = 0$ internal nodes. For the inductive step, consider a node x that has positive height and is an internal node with two children. Each child has a black-height of either $bh(x)$ or $bh(x) - 1$, depending on whether its color is red or black, respectively. Since the height of a child of x is less than the height of x itself, we can apply the inductive hypothesis to conclude that each child has at least $2^{bh(x)-1} - 1$ internal nodes. Thus, the subtree rooted at x contains at least $(2^{bh(x)-1} - 1) + (2^{bh(x)-1} - 1) + 1 = 2^{bh(x)} - 1$ internal nodes, which proves the claim.

To complete the proof of the lemma, let h be the height of the tree. According to property 4, at least half the nodes on any simple path from the root to a leaf, not

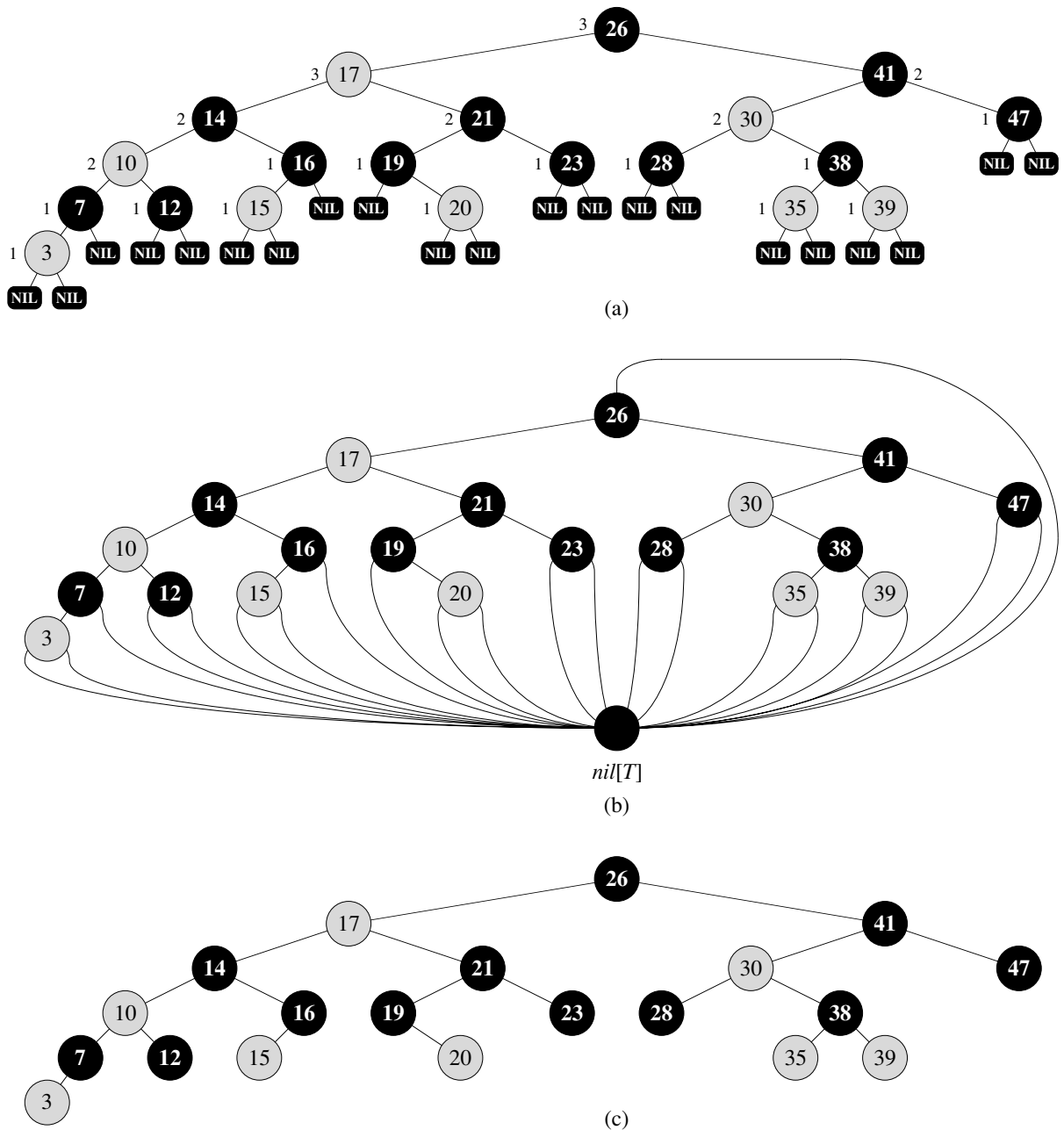


Figure 13.1 A red-black tree with black nodes darkened and red nodes shaded. Every node in a red-black tree is either red or black, the children of a red node are both black, and every simple path from a node to a descendant leaf contains the same number of black nodes. (a) Every leaf, shown as a NIL, is black. Each non-NIL node is marked with its black-height; NIL's have black-height 0. (b) The same red-black tree but with each NIL replaced by the single sentinel $nil[T]$, which is always black, and with black-heights omitted. The root's parent is also the sentinel. (c) The same red-black tree but with leaves and the root's parent omitted entirely. We shall use this drawing style in the remainder of this chapter.

including the root, must be black. Consequently, the black-height of the root must be at least $h/2$; thus,

$$n \geq 2^{h/2} - 1.$$

Moving the 1 to the left-hand side and taking logarithms on both sides yields $\lg(n + 1) \geq h/2$, or $h \leq 2 \lg(n + 1)$. ■

An immediate consequence of this lemma is that the dynamic-set operations SEARCH, MINIMUM, MAXIMUM, SUCCESSOR, and PREDECESSOR can be implemented in $O(\lg n)$ time on red-black trees, since they can be made to run in $O(h)$ time on a search tree of height h (as shown in Chapter 12) and any red-black tree on n nodes is a search tree with height $O(\lg n)$. (Of course, references to NIL in the algorithms of Chapter 12 would have to be replaced by $nil[T]$.) Although the algorithms TREE-INSERT and TREE-DELETE from Chapter 12 run in $O(\lg n)$ time when given a red-black tree as input, they do not directly support the dynamic-set operations INSERT and DELETE, since they do not guarantee that the modified binary search tree will be a red-black tree. We shall see in Sections 13.3 and 13.4, however, that these two operations can indeed be supported in $O(\lg n)$ time.

Exercises

13.1-1

In the style of Figure 13.1(a), draw the complete binary search tree of height 3 on the keys $\{1, 2, \dots, 15\}$. Add the NIL leaves and color the nodes in three different ways such that the black-heights of the resulting red-black trees are 2, 3, and 4.

13.1-2

Draw the red-black tree that results after TREE-INSERT is called on the tree in Figure 13.1 with key 36. If the inserted node is colored red, is the resulting tree a red-black tree? What if it is colored black?

13.1-3

Let us define a **relaxed red-black tree** as a binary search tree that satisfies red-black properties 1, 3, 4, and 5. In other words, the root may be either red or black. Consider a relaxed red-black tree T whose root is red. If we color the root of T black but make no other changes to T , is the resulting tree a red-black tree?

13.1-4

Suppose that we “absorb” every red node in a red-black tree into its black parent, so that the children of the red node become children of the black parent. (Ignore what happens to the keys.) What are the possible degrees of a black node after all its red children are absorbed? What can you say about the depths of the leaves of the resulting tree?

13.1-5

Show that the longest simple path from a node x in a red-black tree to a descendant leaf has length at most twice that of the shortest simple path from node x to a descendant leaf.

13.1-6

What is the largest possible number of internal nodes in a red-black tree with black-height k ? What is the smallest possible number?

13.1-7

Describe a red-black tree on n keys that realizes the largest possible ratio of red internal nodes to black internal nodes. What is this ratio? What tree has the smallest possible ratio, and what is the ratio?

13.2 Rotations

The search-tree operations TREE-INSERT and TREE-DELETE, when run on a red-black tree with n keys, take $O(\lg n)$ time. Because they modify the tree, the result may violate the red-black properties enumerated in Section 13.1. To restore these properties, we must change the colors of some of the nodes in the tree and also change the pointer structure.

We change the pointer structure through *rotation*, which is a local operation in a search tree that preserves the binary-search-tree property. Figure 13.2 shows the two kinds of rotations: left rotations and right rotations. When we do a left rotation on a node x , we assume that its right child y is not $nil[T]$; x may be any node in the tree whose right child is not $nil[T]$. The left rotation “pivots” around the link from x to y . It makes y the new root of the subtree, with x as y ’s left child and y ’s left child as x ’s right child.

The pseudocode for LEFT-ROTATE assumes that $right[x] \neq nil[T]$ and that the root’s parent is $nil[T]$.

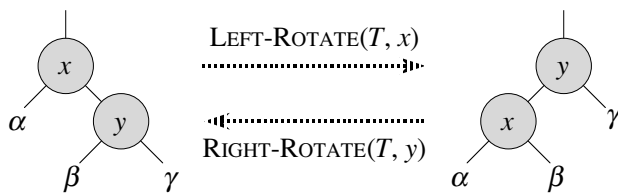


Figure 13.2 The rotation operations on a binary search tree. The operation $\text{LEFT-ROTATE}(T, x)$ transforms the configuration of the two nodes on the left into the configuration on the right by changing a constant number of pointers. The configuration on the right can be transformed into the configuration on the left by the inverse operation $\text{RIGHT-ROTATE}(T, y)$. The letters α , β , and γ represent arbitrary subtrees. A rotation operation preserves the binary-search-tree property: the keys in α precede $\text{key}[x]$, which precedes the keys in β , which precedes $\text{key}[y]$, which precedes the keys in γ .

$\text{LEFT-ROTATE}(T, x)$

```

1   $y \leftarrow \text{right}[x]$            ▷ Set  $y$ .
2   $\text{right}[x] \leftarrow \text{left}[y]$     ▷ Turn  $y$ 's left subtree into  $x$ 's right subtree.
3  if  $\text{left}[y] \neq \text{nil}[T]$ 
4      then  $p[\text{left}[y]] \leftarrow x$ 
5   $p[y] \leftarrow p[x]$            ▷ Link  $x$ 's parent to  $y$ .
6  if  $p[x] = \text{nil}[T]$ 
7      then  $\text{root}[T] \leftarrow y$ 
8      else if  $x = \text{left}[p[x]]$ 
9          then  $\text{left}[p[x]] \leftarrow y$ 
10         else  $\text{right}[p[x]] \leftarrow y$ 
11  $\text{left}[y] \leftarrow x$            ▷ Put  $x$  on  $y$ 's left.
12  $p[x] \leftarrow y$ 
```

Figure 13.3 shows how LEFT-ROTATE operates. The code for RIGHT-ROTATE is symmetric. Both LEFT-ROTATE and RIGHT-ROTATE run in $O(1)$ time. Only pointers are changed by a rotation; all other fields in a node remain the same.

Exercises

13.2-1

Write pseudocode for RIGHT-ROTATE .

13.2-2

Argue that in every n -node binary search tree, there are exactly $n - 1$ possible rotations.

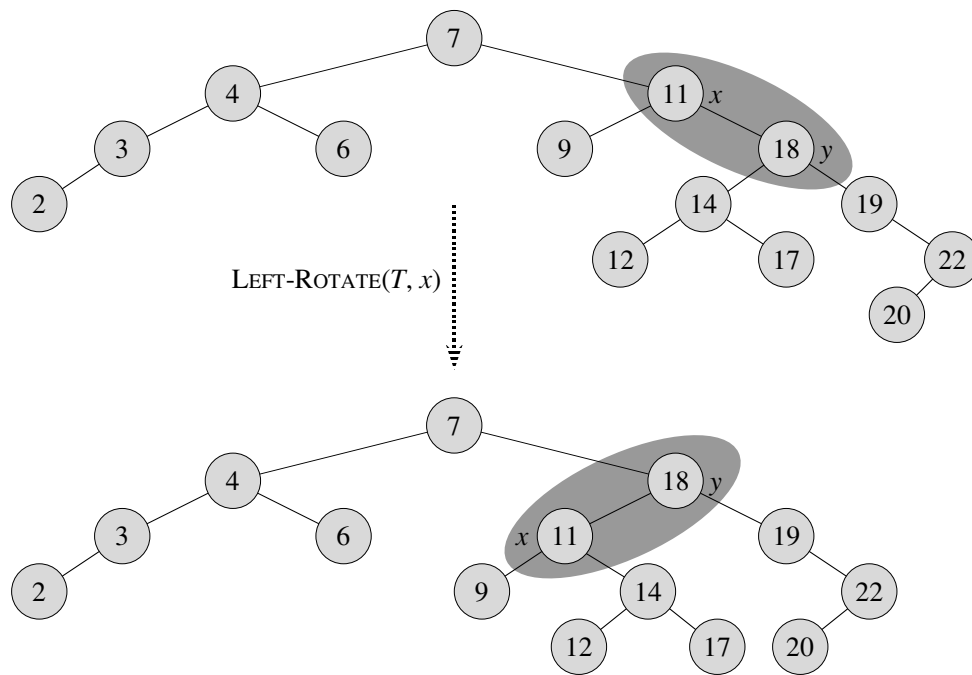


Figure 13.3 An example of how the procedure $\text{LEFT-ROTATE}(T, x)$ modifies a binary search tree. Inorder tree walks of the input tree and the modified tree produce the same listing of key values.

13.2-3

Let a , b , and c be arbitrary nodes in subtrees α , β , and γ , respectively, in the left tree of Figure 13.2. How do the depths of a , b , and c change when a left rotation is performed on node x in the figure?

13.2-4

Show that any arbitrary n -node binary search tree can be transformed into any other arbitrary n -node binary search tree using $O(n)$ rotations. (*Hint*: First show that at most $n - 1$ right rotations suffice to transform the tree into a right-going chain.)

13.2-5 ★

We say that a binary search tree T_1 can be **right-converted** to binary search tree T_2 if it is possible to obtain T_2 from T_1 via a series of calls to RIGHT-ROTATE . Give an example of two trees T_1 and T_2 such that T_1 cannot be right-converted to T_2 . Then show that if a tree T_1 can be right-converted to T_2 , it can be right-converted using $O(n^2)$ calls to RIGHT-ROTATE .

13.3 Insertion

Insertion of a node into an n -node red-black tree can be accomplished in $O(\lg n)$ time. We use a slightly modified version of the TREE-INSERT procedure (Section 12.3) to insert node z into the tree T as if it were an ordinary binary search tree, and then we color z red. To guarantee that the red-black properties are preserved, we then call an auxiliary procedure RB-INSERT-FIXUP to recolor nodes and perform rotations. The call RB-INSERT(T, z) inserts node z , whose *key* field is assumed to have already been filled in, into the red-black tree T .

```

RB-INSERT( $T, z$ )
1   $y \leftarrow nil[T]$ 
2   $x \leftarrow root[T]$ 
3  while  $x \neq nil[T]$ 
4      do  $y \leftarrow x$ 
5          if  $key[z] < key[x]$ 
6              then  $x \leftarrow left[x]$ 
7              else  $x \leftarrow right[x]$ 
8   $p[z] \leftarrow y$ 
9  if  $y = nil[T]$ 
10     then  $root[T] \leftarrow z$ 
11     else if  $key[z] < key[y]$ 
12         then  $left[y] \leftarrow z$ 
13         else  $right[y] \leftarrow z$ 
14   $left[z] \leftarrow nil[T]$ 
15   $right[z] \leftarrow nil[T]$ 
16   $color[z] \leftarrow RED$ 
17  RB-INSERT-FIXUP( $T, z$ )

```

There are four differences between the procedures TREE-INSERT and RB-INSERT. First, all instances of NIL in TREE-INSERT are replaced by $nil[T]$. Second, we set $left[z]$ and $right[z]$ to $nil[T]$ in lines 14–15 of RB-INSERT, in order to maintain the proper tree structure. Third, we color z red in line 16. Fourth, because coloring z red may cause a violation of one of the red-black properties, we call RB-INSERT-FIXUP(T, z) in line 17 of RB-INSERT to restore the red-black properties.

RB-INSERT-FIXUP(T, z)

```

1  while  $color[p[z]] = \text{RED}$ 
2      do if  $p[z] = \text{left}[p[p[z]]]$ 
3          then  $y \leftarrow \text{right}[p[p[z]]]$ 
4              if  $color[y] = \text{RED}$ 
5                  then  $color[p[z]] \leftarrow \text{BLACK}$                                 ▷ Case 1
6                       $color[y] \leftarrow \text{BLACK}$                                 ▷ Case 1
7                       $color[p[p[z]]] \leftarrow \text{RED}$                             ▷ Case 1
8                       $z \leftarrow p[p[z]]$                                     ▷ Case 1
9                  else if  $z = \text{right}[p[z]]$ 
10                     then  $z \leftarrow p[z]$                                 ▷ Case 2
11                     LEFT-ROTATE( $T, z$ )                                ▷ Case 2
12                      $color[p[z]] \leftarrow \text{BLACK}$                             ▷ Case 3
13                      $color[p[p[z]]] \leftarrow \text{RED}$                             ▷ Case 3
14                     RIGHT-ROTATE( $T, p[p[z]]$ )                            ▷ Case 3
15                 else (same as then clause
                        with “right” and “left” exchanged)
16   $color[\text{root}[T]] \leftarrow \text{BLACK}$ 

```

To understand how RB-INSERT-FIXUP works, we shall break our examination of the code into three major steps. First, we shall determine what violations of the red-black properties are introduced in RB-INSERT when the node z is inserted and colored red. Second, we shall examine the overall goal of the **while** loop in lines 1–15. Finally, we shall explore each of the three cases¹ into which the **while** loop is broken and see how they accomplish the goal. Figure 13.4 shows how RB-INSERT-FIXUP operates on a sample red-black tree.

Which of the red-black properties can be violated upon the call to RB-INSERT-FIXUP? Property 1 certainly continues to hold, as does property 3, since both children of the newly inserted red node are the sentinel $nil[T]$. Property 5, which says that the number of black nodes is the same on every path from a given node, is satisfied as well, because node z replaces the (black) sentinel, and node z is red with sentinel children. Thus, the only properties that might be violated are property 2, which requires the root to be black, and property 4, which says that a red node cannot have a red child. Both possible violations are due to z being colored red. Property 2 is violated if z is the root, and property 4 is violated if z 's parent is red. Figure 13.4(a) shows a violation of property 4 after the node z has been inserted.

The **while** loop in lines 1–15 maintains the following three-part invariant:

¹Case 2 falls through into case 3, and so these two cases are not mutually exclusive.

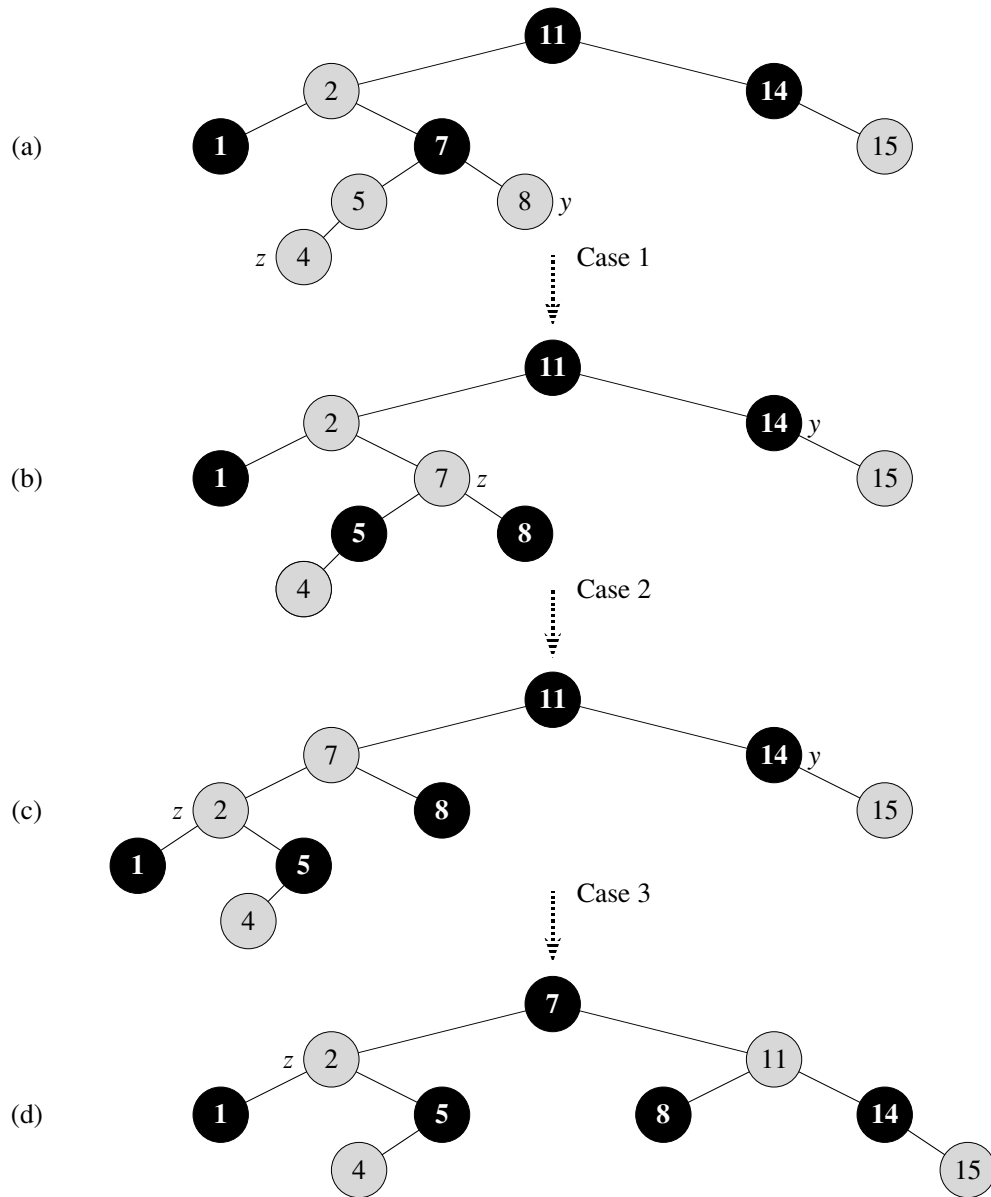


Figure 13.4 The operation of RB-INSERT-FIXUP. (a) A node z after insertion. Since z and its parent $p[z]$ are both red, a violation of property 4 occurs. Since z 's uncle y is red, case 1 in the code can be applied. Nodes are recolored and the pointer z is moved up the tree, resulting in the tree shown in (b). Once again, z and its parent are both red, but z 's uncle y is black. Since z is the right child of $p[z]$, case 2 can be applied. A left rotation is performed, and the tree that results is shown in (c). Now z is the left child of its parent, and case 3 can be applied. A right rotation yields the tree in (d), which is a legal red-black tree.

At the start of each iteration of the loop,

- a. Node z is red.
- b. If $p[z]$ is the root, then $p[z]$ is black.
- c. If there is a violation of the red-black properties, there is at most one violation, and it is of either property 2 or property 4. If there is a violation of property 2, it occurs because z is the root and is red. If there is a violation of property 4, it occurs because both z and $p[z]$ are red.

Part (c), which deals with violations of red-black properties, is more central to showing that RB-INSERT-FIXUP restores the red-black properties than parts (a) and (b), which we use along the way to understand situations in the code. Because we will be focusing on node z and nodes near it in the tree, it is helpful to know from part (a) that z is red. We shall use part (b) to show that the node $p[p[z]]$ exists when we reference it in lines 2, 3, 7, 8, 13, and 14.

Recall that we need to show that a loop invariant is true prior to the first iteration of the loop, that each iteration maintains the loop invariant, and that the loop invariant gives us a useful property at loop termination.

We start with the initialization and termination arguments. Then, as we examine how the body of the loop works in more detail, we shall argue that the loop maintains the invariant upon each iteration. Along the way, we will also demonstrate that there are two possible outcomes of each iteration of the loop: the pointer z moves up the tree, or some rotations are performed and the loop terminates.

Initialization: Prior to the first iteration of the loop, we started with a red-black tree with no violations, and we added a red node z . We show that each part of the invariant holds at the time RB-INSERT-FIXUP is called:

- a. When RB-INSERT-FIXUP is called, z is the red node that was added.
- b. If $p[z]$ is the root, then $p[z]$ started out black and did not change prior to the call of RB-INSERT-FIXUP.
- c. We have already seen that properties 1, 3, and 5 hold when RB-INSERT-FIXUP is called.

If there is a violation of property 2, then the red root must be the newly added node z , which is the only internal node in the tree. Because the parent and both children of z are the sentinel, which is black, there is not also a violation of property 4. Thus, this violation of property 2 is the only violation of red-black properties in the entire tree.

If there is a violation of property 4, then because the children of node z are black sentinels and the tree had no other violations prior to z being added, the violation must be because both z and $p[z]$ are red. Moreover, there are no other violations of red-black properties.

Termination: When the loop terminates, it does so because $p[z]$ is black. (If z is the root, then $p[z]$ is the sentinel $nil[T]$, which is black.) Thus, there is no violation of property 4 at loop termination. By the loop invariant, the only property that might fail to hold is property 2. Line 16 restores this property, too, so that when RB-INSERT-FIXUP terminates, all the red-black properties hold.

Maintenance: There are actually six cases to consider in the **while** loop, but three of them are symmetric to the other three, depending on whether z 's parent $p[z]$ is a left child or a right child of z 's grandparent $p[p[z]]$, which is determined in line 2. We have given the code only for the situation in which $p[z]$ is a left child. The node $p[p[z]]$ exists, since by part (b) of the loop invariant, if $p[z]$ is the root, then $p[z]$ is black. Since we enter a loop iteration only if $p[z]$ is red, we know that $p[z]$ cannot be the root. Hence, $p[p[z]]$ exists.

Case 1 is distinguished from cases 2 and 3 by the color of z 's parent's sibling, or "uncle." Line 3 makes y point to z 's uncle $right[p[p[z]]]$, and a test is made in line 4. If y is red, then case 1 is executed. Otherwise, control passes to cases 2 and 3. In all three cases, z 's grandparent $p[p[z]]$ is black, since its parent $p[z]$ is red, and property 4 is violated only between z and $p[z]$.

Case 1: z 's uncle y is red

Figure 13.5 shows the situation for case 1 (lines 5–8). Case 1 is executed when both $p[z]$ and y are red. Since $p[p[z]]$ is black, we can color both $p[z]$ and y black, thereby fixing the problem of z and $p[z]$ both being red, and color $p[p[z]]$ red, thereby maintaining property 5. We then repeat the **while** loop with $p[p[z]]$ as the new node z . The pointer z moves up two levels in the tree.

Now we show that case 1 maintains the loop invariant at the start of the next iteration. We use z to denote node z in the current iteration, and $z' = p[p[z]]$ to denote the node z at the test in line 1 upon the next iteration.

- Because this iteration colors $p[p[z]]$ red, node z' is red at the start of the next iteration.
- The node $p[z']$ is $p[p[p[z]]]$ in this iteration, and the color of this node does not change. If this node is the root, it was black prior to this iteration, and it remains black at the start of the next iteration.
- We have already argued that case 1 maintains property 5, and it clearly does not introduce a violation of properties 1 or 3.

If node z' is the root at the start of the next iteration, then case 1 corrected the lone violation of property 4 in this iteration. Since z' is red and it is the root, property 2 becomes the only one that is violated, and this violation is due to z' .

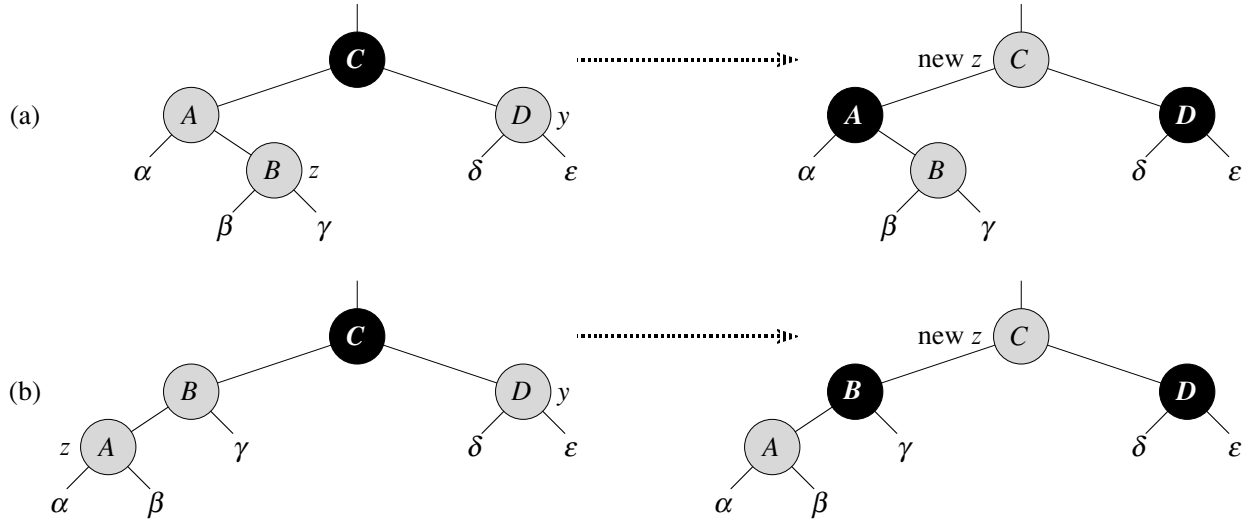


Figure 13.5 Case 1 of the procedure RB-INSERT. Property 4 is violated, since z and its parent $p[z]$ are both red. The same action is taken whether (a) z is a right child or (b) z is a left child. Each of the subtrees α , β , γ , δ , and ϵ has a black root, and each has the same black-height. The code for case 1 changes the colors of some nodes, preserving property 5: all downward paths from a node to a leaf have the same number of blacks. The **while** loop continues with node z 's grandparent $p[p[z]]$ as the new z . Any violation of property 4 can now occur only between the new z , which is red, and its parent, if it is red as well.

If node z' is not the root at the start of the next iteration, then case 1 has not created a violation of property 2. Case 1 corrected the lone violation of property 4 that existed at the start of this iteration. It then made z' red and left $p[z']$ alone. If $p[z']$ was black, there is no violation of property 4. If $p[z']$ was red, coloring z' red created one violation of property 4 between z' and $p[z']$.

Case 2: z 's uncle y is black and z is a right child

Case 3: z 's uncle y is black and z is a left child

In cases 2 and 3, the color of z 's uncle y is black. The two cases are distinguished by whether z is a right or left child of $p[z]$. Lines 10–11 constitute case 2, which is shown in Figure 13.6 together with case 3. In case 2, node z is a right child of its parent. We immediately use a left rotation to transform the situation into case 3 (lines 12–14), in which node z is a left child. Because both z and $p[z]$ are red, the rotation affects neither the black-height of nodes nor property 5. Whether we enter case 3 directly or through case 2, z 's uncle y is black, since otherwise we would have executed case 1. Additionally, the node $p[p[z]]$ exists, since we have argued that this node existed at the time that

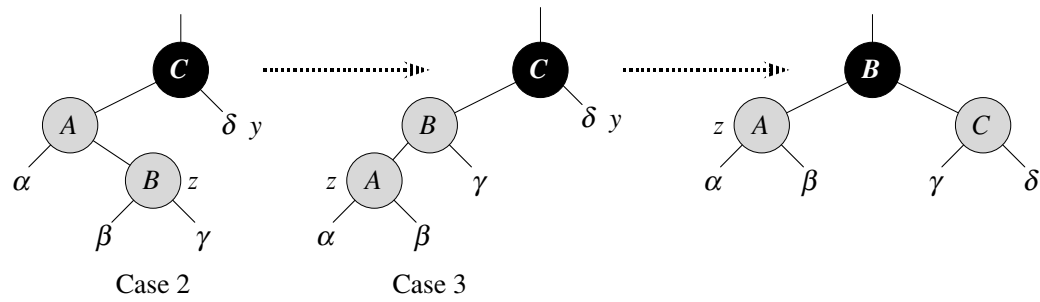


Figure 13.6 Cases 2 and 3 of the procedure RB-INSERT. As in case 1, property 4 is violated in either case 2 or case 3 because z and its parent $p[z]$ are both red. Each of the subtrees α , β , γ , and δ has a black root (α , β , and γ from property 4, and δ because otherwise we would be in case 1), and each has the same black-height. Case 2 is transformed into case 3 by a left rotation, which preserves property 5: all downward paths from a node to a leaf have the same number of blacks. Case 3 causes some color changes and a right rotation, which also preserve property 5. The **while** loop then terminates, because property 4 is satisfied: there are no longer two red nodes in a row.

lines 2 and 3 were executed, and after moving z up one level in line 10 and then down one level in line 11, the identity of $p[p[z]]$ remains unchanged. In case 3, we execute some color changes and a right rotation, which preserve property 5, and then, since we no longer have two red nodes in a row, we are done. The body of the **while** loop is not executed another time, since $p[z]$ is now black.

Now we show that cases 2 and 3 maintain the loop invariant. (As we have just argued, $p[z]$ will be black upon the next test in line 1, and the loop body will not execute again.)

- Case 2 makes z point to $p[z]$, which is red. No further change to z or its color occurs in cases 2 and 3.
- Case 3 makes $p[z]$ black, so that if $p[z]$ is the root at the start of the next iteration, it is black.
- As in case 1, properties 1, 3, and 5 are maintained in cases 2 and 3.

Since node z is not the root in cases 2 and 3, we know that there is no violation of property 2. Cases 2 and 3 do not introduce a violation of property 2, since the only node that is made red becomes a child of a black node by the rotation in case 3.

Cases 2 and 3 correct the lone violation of property 4, and they do not introduce another violation.

Having shown that each iteration of the loop maintains the invariant, we have shown that RB-INSERT-FIXUP correctly restores the red-black properties.

Analysis

What is the running time of RB-INSERT? Since the height of a red-black tree on n nodes is $O(\lg n)$, lines 1–16 of RB-INSERT take $O(\lg n)$ time. In RB-INSERT-FIXUP, the **while** loop repeats only if case 1 is executed, and then the pointer z moves two levels up the tree. The total number of times the **while** loop can be executed is therefore $O(\lg n)$. Thus, RB-INSERT takes a total of $O(\lg n)$ time. Interestingly, it never performs more than two rotations, since the **while** loop terminates if case 2 or case 3 is executed.

Exercises

13.3-1

In line 16 of RB-INSERT, we set the color of the newly inserted node z to red. Notice that if we had chosen to set z 's color to black, then property 4 of a red-black tree would not be violated. Why didn't we choose to set z 's color to black?

13.3-2

Show the red-black trees that result after successively inserting the keys 41, 38, 31, 12, 19, 8 into an initially empty red-black tree.

13.3-3

Suppose that the black-height of each of the subtrees $\alpha, \beta, \gamma, \delta, \varepsilon$ in Figures 13.5 and 13.6 is k . Label each node in each figure with its black-height to verify that property 5 is preserved by the indicated transformation.

13.3-4

Professor Teach is concerned that RB-INSERT-FIXUP might set $color[nil[T]]$ to RED, in which case the test in line 1 would not cause the loop to terminate when z is the root. Show that the professor's concern is unfounded by arguing that RB-INSERT-FIXUP never sets $color[nil[T]]$ to RED.

13.3-5

Consider a red-black tree formed by inserting n nodes with RB-INSERT. Argue that if $n > 1$, the tree has at least one red node.

13.3-6

Suggest how to implement RB-INSERT efficiently if the representation for red-black trees includes no storage for parent pointers.

13.4 Deletion

Like the other basic operations on an n -node red-black tree, deletion of a node takes time $O(\lg n)$. Deleting a node from a red-black tree is only slightly more complicated than inserting a node.

The procedure RB-DELETE is a minor modification of the TREE-DELETE procedure (Section 12.3). After splicing out a node, it calls an auxiliary procedure RB-DELETE-FIXUP that changes colors and performs rotations to restore the red-black properties.

RB-DELETE(T, z)

```

1  if  $\text{left}[z] = \text{nil}[T]$  or  $\text{right}[z] = \text{nil}[T]$ 
2      then  $y \leftarrow z$ 
3      else  $y \leftarrow \text{TREE-SUCCESSOR}(z)$ 
4  if  $\text{left}[y] \neq \text{nil}[T]$ 
5      then  $x \leftarrow \text{left}[y]$ 
6      else  $x \leftarrow \text{right}[y]$ 
7   $p[x] \leftarrow p[y]$ 
8  if  $p[y] = \text{nil}[T]$ 
9      then  $\text{root}[T] \leftarrow x$ 
10     else if  $y = \text{left}[p[y]]$ 
11         then  $\text{left}[p[y]] \leftarrow x$ 
12         else  $\text{right}[p[y]] \leftarrow x$ 
13 if  $y \neq z$ 
14     then  $\text{key}[z] \leftarrow \text{key}[y]$ 
15     copy  $y$ 's satellite data into  $z$ 
16 if  $\text{color}[y] = \text{BLACK}$ 
17     then RB-DELETE-FIXUP( $T, x$ )
18 return  $y$ 
```

There are three differences between the procedures TREE-DELETE and RB-DELETE. First, all references to NIL in TREE-DELETE are replaced by references to the sentinel $\text{nil}[T]$ in RB-DELETE. Second, the test for whether x is NIL in line 7 of TREE-DELETE is removed, and the assignment $p[x] \leftarrow p[y]$ is performed unconditionally in line 7 of RB-DELETE. Thus, if x is the sentinel $\text{nil}[T]$, its parent pointer points to the parent of the spliced-out node y . Third, a call to RB-DELETE-FIXUP is made in lines 16–17 if y is black. If y is red, the red-black properties still hold when y is spliced out, for the following reasons:

- no black-heights in the tree have changed,
- no red nodes have been made adjacent, and

- since y could not have been the root if it was red, the root remains black.

The node x passed to RB-DELETE-FIXUP is one of two nodes: either the node that was y 's sole child before y was spliced out if y had a child that was not the sentinel $nil[T]$, or, if y had no children, x is the sentinel $nil[T]$. In the latter case, the unconditional assignment in line 7 guarantees that x 's parent is now the node that was previously y 's parent, whether x is a key-bearing internal node or the sentinel $nil[T]$.

We can now examine how the procedure RB-DELETE-FIXUP restores the red-black properties to the search tree.

RB-DELETE-FIXUP(T, x)

```

1  while  $x \neq \text{root}[T]$  and  $\text{color}[x] = \text{BLACK}$ 
2      do if  $x = \text{left}[p[x]]$ 
3          then  $w \leftarrow \text{right}[p[x]]$ 
4              if  $\text{color}[w] = \text{RED}$ 
5                  then  $\text{color}[w] \leftarrow \text{BLACK}$                                 ▷ Case 1
6                       $\text{color}[p[x]] \leftarrow \text{RED}$                                 ▷ Case 1
7                      LEFT-ROTATE( $T, p[x]$ )                                ▷ Case 1
8                       $w \leftarrow \text{right}[p[x]]$                                 ▷ Case 1
9              if  $\text{color}[\text{left}[w]] = \text{BLACK}$  and  $\text{color}[\text{right}[w]] = \text{BLACK}$ 
10                 then  $\text{color}[w] \leftarrow \text{RED}$                                 ▷ Case 2
11                      $x \leftarrow p[x]$                                 ▷ Case 2
12                 else if  $\text{color}[\text{right}[w]] = \text{BLACK}$ 
13                     then  $\text{color}[\text{left}[w]] \leftarrow \text{BLACK}$                                 ▷ Case 3
14                          $\text{color}[w] \leftarrow \text{RED}$                                 ▷ Case 3
15                         RIGHT-ROTATE( $T, w$ )                                ▷ Case 3
16                          $w \leftarrow \text{right}[p[x]]$                                 ▷ Case 3
17                      $\text{color}[w] \leftarrow \text{color}[p[x]]$                                 ▷ Case 4
18                      $\text{color}[p[x]] \leftarrow \text{BLACK}$                                 ▷ Case 4
19                      $\text{color}[\text{right}[w]] \leftarrow \text{BLACK}$                                 ▷ Case 4
20                     LEFT-ROTATE( $T, p[x]$ )                                ▷ Case 4
21                      $x \leftarrow \text{root}[T]$                                 ▷ Case 4
22                 else (same as then clause with “right” and “left” exchanged)
23      $\text{color}[x] \leftarrow \text{BLACK}$ 

```

If the spliced-out node y in RB-DELETE is black, three problems may arise. First, if y had been the root and a red child of y becomes the new root, we have violated property 2. Second, if both x and $p[y]$ (which is now also $p[x]$) were red, then we have violated property 4. Third, y 's removal causes any path that previously contained y to have one fewer black node. Thus, property 5 is now violated by any ancestor of y in the tree. We can correct this problem by saying

that node x has an “extra” black. That is, if we add 1 to the count of black nodes on any path that contains x , then under this interpretation, property 5 holds. When we splice out the black node y , we “push” its blackness onto its child. The problem is that now node x is neither red nor black, thereby violating property 1. Instead, node x is either “doubly black” or “red-and-black,” and it contributes either 2 or 1, respectively, to the count of black nodes on paths containing x . The *color* attribute of x will still be either RED (if x is red-and-black) or BLACK (if x is doubly black). In other words, the extra black on a node is reflected in x ’s pointing to the node rather than in the *color* attribute.

The procedure RB-DELETE-FIXUP restores properties 1, 2, and 4. Exercises 13.4-1 and 13.4-2 ask you to show that the procedure restores properties 2 and 4, and so in the remainder of this section, we shall focus on property 1. The goal of the **while** loop in lines 1–22 is to move the extra black up the tree until

1. x points to a red-and-black node, in which case we color x (singly) black in line 23,
2. x points to the root, in which case the extra black can be simply “removed,” or
3. suitable rotations and recolorings can be performed.

Within the **while** loop, x always points to a nonroot doubly black node. We determine in line 2 whether x is a left child or a right child of its parent $p[x]$. (We have given the code for the situation in which x is a left child; the situation in which x is a right child—line 22—is symmetric.) We maintain a pointer w to the sibling of x . Since node x is doubly black, node w cannot be $nil[T]$; otherwise, the number of blacks on the path from $p[x]$ to the (singly black) leaf w would be smaller than the number on the path from $p[x]$ to x .

The four cases² in the code are illustrated in Figure 13.7. Before examining each case in detail, let’s look more generally at how we can verify that the transformation in each of the cases preserves property 5. The key idea is that in each case the number of black nodes (including x ’s extra black) from (and including) the root of the subtree shown to each of the subtrees $\alpha, \beta, \dots, \zeta$ is preserved by the transformation. Thus, if property 5 holds prior to the transformation, it continues to hold afterward. For example, in Figure 13.7(a), which illustrates case 1, the number of black nodes from the root to either subtree α or β is 3, both before and after the transformation. (Again, remember that node x adds an extra black.) Similarly, the number of black nodes from the root to any of γ, δ, ϵ , and ζ is 2, both before and after the transformation. In Figure 13.7(b), the counting must involve the value c of the *color* attribute of the root of the subtree shown, which can be either RED or BLACK. If we define $\text{count}(\text{RED}) = 0$ and $\text{count}(\text{BLACK}) = 1$, then the num-

²As in RB-INSERT-FIXUP, the cases in RB-DELETE-FIXUP are not mutually exclusive.

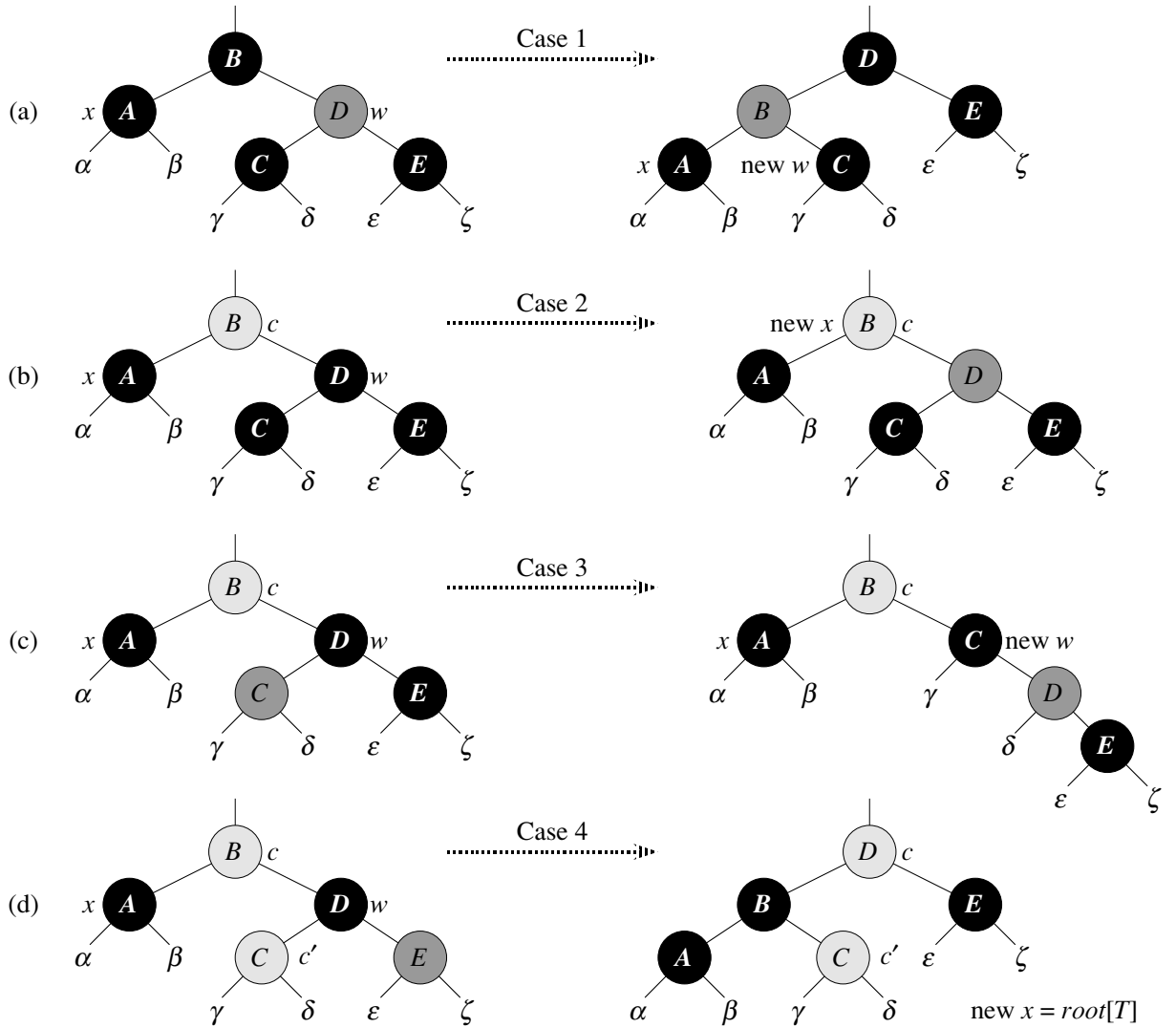


Figure 13.7 The cases in the **while** loop of the procedure RB-DELETE-FIXUP. Darkened nodes have *color* attributes BLACK, heavily shaded nodes have *color* attributes RED, and lightly shaded nodes have *color* attributes represented by c and c' , which may be either RED or BLACK. The letters $\alpha, \beta, \dots, \zeta$ represent arbitrary subtrees. In each case, the configuration on the left is transformed into the configuration on the right by changing some colors and/or performing a rotation. Any node pointed to by x has an extra black and is either doubly black or red-and-black. The only case that causes the loop to repeat is case 2. (a) Case 1 is transformed to case 2, 3, or 4 by exchanging the colors of nodes B and D and performing a left rotation. (b) In case 2, the extra black represented by the pointer x is moved up the tree by coloring node D red and setting x to point to node B . If we enter case 2 through case 1, the **while** loop terminates because the new node x is red-and-black, and therefore the value c of its *color* attribute is RED. (c) Case 3 is transformed to case 4 by exchanging the colors of nodes C and D and performing a right rotation. (d) In case 4, the extra black represented by x can be removed by changing some colors and performing a left rotation (without violating the red-black properties), and the loop terminates.

ber of black nodes from the root to α is $2 + \text{count}(c)$, both before and after the transformation. In this case, after the transformation, the new node x has *color* attribute c , but this node is really either red-and-black (if $c = \text{RED}$) or doubly black (if $c = \text{BLACK}$). The other cases can be verified similarly (see Exercise 13.4-5).

Case 1: x 's sibling w is red

Case 1 (lines 5–8 of RB-DELETE-FIXUP and Figure 13.7(a)) occurs when node w , the sibling of node x , is red. Since w must have black children, we can switch the colors of w and $p[x]$ and then perform a left-rotation on $p[x]$ without violating any of the red-black properties. The new sibling of x , which is one of w 's children prior to the rotation, is now black, and thus we have converted case 1 into case 2, 3, or 4.

Cases 2, 3, and 4 occur when node w is black; they are distinguished by the colors of w 's children.

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

In case 2 (lines 10–11 of RB-DELETE-FIXUP and Figure 13.7(b)), both of w 's children are black. Since w is also black, we take one black off both x and w , leaving x with only one black and leaving w red. To compensate for removing one black from x and w , we would like to add an extra black to $p[x]$, which was originally either red or black. We do so by repeating the **while** loop with $p[x]$ as the new node x . Observe that if we enter case 2 through case 1, the new node x is red-and-black, since the original $p[x]$ was red. Hence, the value c of the *color* attribute of the new node x is RED, and the loop terminates when it tests the loop condition. The new node x is then colored (singly) black in line 23.

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

Case 3 (lines 13–16 and Figure 13.7(c)) occurs when w is black, its left child is red, and its right child is black. We can switch the colors of w and its left child $\text{left}[w]$ and then perform a right rotation on w without violating any of the red-black properties. The new sibling w of x is now a black node with a red right child, and thus we have transformed case 3 into case 4.

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

Case 4 (lines 17–21 and Figure 13.7(d)) occurs when node x 's sibling w is black and w 's right child is red. By making some color changes and performing a left rotation on $p[x]$, we can remove the extra black on x , making it singly black, without violating any of the red-black properties. Setting x to be the root causes the **while** loop to terminate when it tests the loop condition.

Analysis

What is the running time of RB-DELETE? Since the height of a red-black tree of n nodes is $O(\lg n)$, the total cost of the procedure without the call to RB-DELETE-FIXUP takes $O(\lg n)$ time. Within RB-DELETE-FIXUP, cases 1, 3, and 4 each terminate after performing a constant number of color changes and at most three rotations. Case 2 is the only case in which the **while** loop can be repeated, and then the pointer x moves up the tree at most $O(\lg n)$ times and no rotations are performed. Thus, the procedure RB-DELETE-FIXUP takes $O(\lg n)$ time and performs at most three rotations, and the overall time for RB-DELETE is therefore also $O(\lg n)$.

Exercises

13.4-1

Argue that after executing RB-DELETE-FIXUP, the root of the tree must be black.

13.4-2

Argue that if in RB-DELETE both x and $p[y]$ are red, then property 4 is restored by the call RB-DELETE-FIXUP(T, x).

13.4-3

In Exercise 13.3-2, you found the red-black tree that results from successively inserting the keys 41, 38, 31, 12, 19, 8 into an initially empty tree. Now show the red-black trees that result from the successive deletion of the keys in the order 8, 12, 19, 31, 38, 41.

13.4-4

In which lines of the code for RB-DELETE-FIXUP might we examine or modify the sentinel $nil[T]$?

13.4-5

In each of the cases of Figure 13.7, give the count of black nodes from the root of the subtree shown to each of the subtrees $\alpha, \beta, \dots, \zeta$, and verify that each count remains the same after the transformation. When a node has a *color* attribute c or c' , use the notation $\text{count}(c)$ or $\text{count}(c')$ symbolically in your count.

13.4-6

Professors Skelton and Baron are concerned that at the start of case 1 of RB-DELETE-FIXUP, the node $p[x]$ might not be black. If the professors are correct, then lines 5–6 are wrong. Show that $p[x]$ must be black at the start of case 1, so that the professors have nothing to worry about.

13.4-7

Suppose that a node x is inserted into a red-black tree with RB-INSERT and then immediately deleted with RB-DELETE. Is the resulting red-black tree the same as the initial red-black tree? Justify your answer.

Problems
13-1 Persistent dynamic sets

During the course of an algorithm, we sometimes find that we need to maintain past versions of a dynamic set as it is updated. Such a set is called *persistent*. One way to implement a persistent set is to copy the entire set whenever it is modified, but this approach can slow down a program and also consume much space. Sometimes, we can do much better.

Consider a persistent set S with the operations INSERT, DELETE, and SEARCH, which we implement using binary search trees as shown in Figure 13.8(a). We maintain a separate root for every version of the set. In order to insert the key 5 into the set, we create a new node with key 5. This node becomes the left child of a new node with key 7, since we cannot modify the existing node with key 7. Similarly, the new node with key 7 becomes the left child of a new node with key 8 whose right child is the existing node with key 10. The new node with key 8 becomes, in turn, the right child of a new root r' with key 4 whose left child is the existing node with key 3. We thus copy only part of the tree and share some of the nodes with the original tree, as shown in Figure 13.8(b).

Assume that each tree node has the fields *key*, *left*, and *right* but no parent field. (See also Exercise 13.3-6.)

- a.* For a general persistent binary search tree, identify the nodes that need to be changed to insert a key k or delete a node y .
- b.* Write a procedure PERSISTENT-TREE-INSERT that, given a persistent tree T and a key k to insert, returns a new persistent tree T' that is the result of inserting k into T .
- c.* If the height of the persistent binary search tree T is h , what are the time and space requirements of your implementation of PERSISTENT-TREE-INSERT? (The space requirement is proportional to the number of new nodes allocated.)
- d.* Suppose that we had included the parent field in each node. In this case, PERSISTENT-TREE-INSERT would need to perform additional copying. Prove that PERSISTENT-TREE-INSERT would then require $\Omega(n)$ time and space, where n is the number of nodes in the tree.

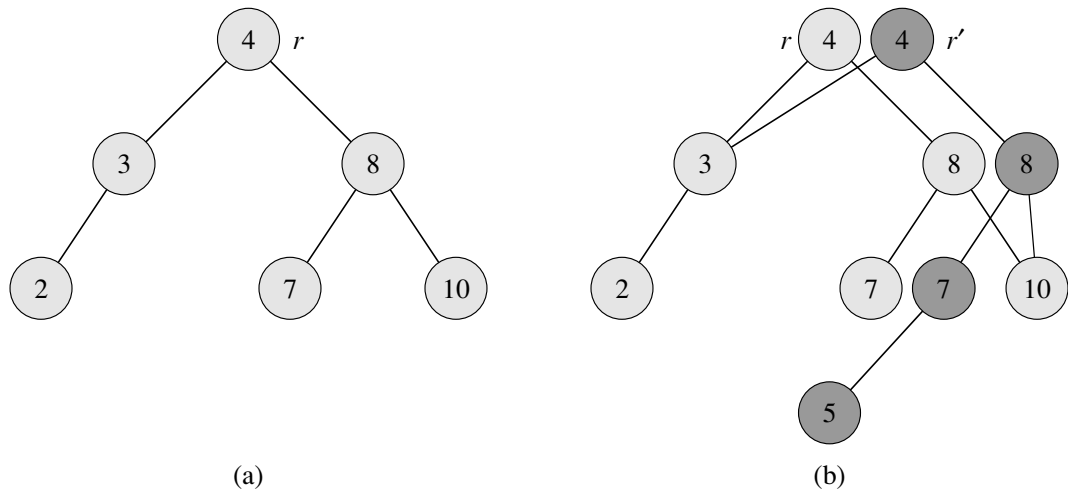


Figure 13.8 (a) A binary search tree with keys 2, 3, 4, 7, 8, 10. (b) The persistent binary search tree that results from the insertion of key 5. The most recent version of the set consists of the nodes reachable from the root r' , and the previous version consists of the nodes reachable from r . Heavily shaded nodes are added when key 5 is inserted.

- e. Show how to use red-black trees to guarantee that the worst-case running time and space are $O(\lg n)$ per insertion or deletion.

13-2 Join operation on red-black trees

The *join* operation takes two dynamic sets S_1 and S_2 and an element x such that for any $x_1 \in S_1$ and $x_2 \in S_2$, we have $\text{key}[x_1] \leq \text{key}[x] \leq \text{key}[x_2]$. It returns a set $S = S_1 \cup \{x\} \cup S_2$. In this problem, we investigate how to implement the join operation on red-black trees.

- a. Given a red-black tree T , we store its black-height as the field $bh[T]$. Argue that this field can be maintained by RB-INSERT and RB-DELETE without requiring extra storage in the nodes of the tree and without increasing the asymptotic running times. Show that while descending through T , we can determine the black-height of each node we visit in $O(1)$ time per node visited.

We wish to implement the operation $\text{RB-JOIN}(T_1, x, T_2)$, which destroys T_1 and T_2 and returns a red-black tree $T = T_1 \cup \{x\} \cup T_2$. Let n be the total number of nodes in T_1 and T_2 .

- b. Assume that $bh[T_1] \geq bh[T_2]$. Describe an $O(\lg n)$ -time algorithm that finds a black node y in T_1 with the largest key from among those nodes whose black-height is $bh[T_2]$.

- c. Let T_y be the subtree rooted at y . Describe how $T_y \cup \{x\} \cup T_2$ can replace T_y in $O(1)$ time without destroying the binary-search-tree property.
- d. What color should we make x so that red-black properties 1, 3, and 5 are maintained? Describe how properties 2 and 4 can be enforced in $O(\lg n)$ time.
- e. Argue that no generality is lost by making the assumption in part (b). Describe the symmetric situation that arises when $bh[T_1] \leq bh[T_2]$.
- f. Argue that the running time of RB-JOIN is $O(\lg n)$.

13-3 AVL trees

An **AVL tree** is a binary search tree that is **height balanced**: for each node x , the heights of the left and right subtrees of x differ by at most 1. To implement an AVL tree, we maintain an extra field in each node: $h[x]$ is the height of node x . As for any other binary search tree T , we assume that $root[T]$ points to the root node.

- a. Prove that an AVL tree with n nodes has height $O(\lg n)$. (*Hint*: Prove that in an AVL tree of height h , there are at least F_h nodes, where F_h is the h th Fibonacci number.)
- b. To insert into an AVL tree, a node is first placed in the appropriate place in binary search tree order. After this insertion, the tree may no longer be height balanced. Specifically, the heights of the left and right children of some node may differ by 2. Describe a procedure $BALANCE(x)$, which takes a subtree rooted at x whose left and right children are height balanced and have heights that differ by at most 2, i.e., $|h[right[x]] - h[left[x]]| \leq 2$, and alters the subtree rooted at x to be height balanced. (*Hint*: Use rotations.)
- c. Using part (b), describe a recursive procedure $AVL-INSERT(x, z)$, which takes a node x within an AVL tree and a newly created node z (whose key has already been filled in), and adds z to the subtree rooted at x , maintaining the property that x is the root of an AVL tree. As in $TREE-INSERT$ from Section 12.3, assume that $key[z]$ has already been filled in and that $left[z] = NIL$ and $right[z] = NIL$; also assume that $h[z] = 0$. Thus, to insert the node z into the AVL tree T , we call $AVL-INSERT(root[T], z)$.
- d. Give an example of an n -node AVL tree in which an $AVL-INSERT$ operation causes $\Omega(\lg n)$ rotations to be performed.

13-4 Treaps

If we insert a set of n items into a binary search tree, the resulting tree may be horribly unbalanced, leading to long search times. As we saw in Section 12.4, however,

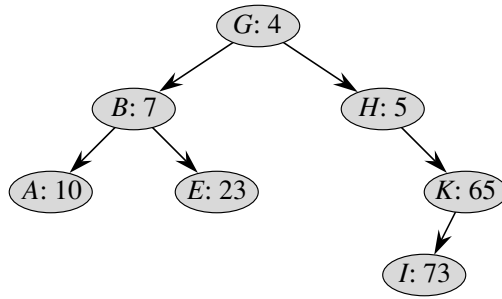


Figure 13.9 A treap. Each node x is labeled with $key[x] : priority[x]$. For example, the root has key G and priority 4.

randomly built binary search trees tend to be balanced. Therefore, a strategy that, on average, builds a balanced tree for a fixed set of items is to randomly permute the items and then insert them in that order into the tree.

What if we do not have all the items at once? If we receive the items one at a time, can we still randomly build a binary search tree out of them?

We will examine a data structure that answers this question in the affirmative. A **treap** is a binary search tree with a modified way of ordering the nodes. Figure 13.9 shows an example. As usual, each node x in the tree has a key value $key[x]$. In addition, we assign $priority[x]$, which is a random number chosen independently for each node. We assume that all priorities are distinct and also that all keys are distinct. The nodes of the treap are ordered so that the keys obey the binary-search-tree property and the priorities obey the min-heap order property:

- If v is a left child of u , then $key[v] < key[u]$.
- If v is a right child of u , then $key[v] > key[u]$.
- If v is a child of u , then $priority[v] > priority[u]$.

(This combination of properties is why the tree is called a “treap;” it has features of both a binary search tree and a heap.)

It helps to think of treaps in the following way. Suppose that we insert nodes x_1, x_2, \dots, x_n , with associated keys, into a treap. Then the resulting treap is the tree that would have been formed if the nodes had been inserted into a normal binary search tree in the order given by their (randomly chosen) priorities, i.e., $priority[x_i] < priority[x_j]$ means that x_i was inserted before x_j .

- Show that given a set of nodes x_1, x_2, \dots, x_n , with associated keys and priorities (all distinct), there is a unique treap associated with these nodes.

- b.** Show that the expected height of a treap is $\Theta(\lg n)$, and hence the time to search for a value in the treap is $\Theta(\lg n)$.

Let us see how to insert a new node into an existing treap. The first thing we do is assign to the new node a random priority. Then we call the insertion algorithm, which we call TREAP-INSERT, whose operation is illustrated in Figure 13.10.

- c.** Explain how TREAP-INSERT works. Explain the idea in English and give pseudocode. (*Hint:* Execute the usual binary-search-tree insertion procedure and then perform rotations to restore the min-heap order property.)
- d.** Show that the expected running time of TREAP-INSERT is $\Theta(\lg n)$.

TREAP-INSERT performs a search and then a sequence of rotations. Although these two operations have the same expected running time, they have different costs in practice. A search reads information from the treap without modifying it. In contrast, a rotation changes parent and child pointers within the treap. On most computers, read operations are much faster than write operations. Thus we would like TREAP-INSERT to perform few rotations. We will show that the expected number of rotations performed is bounded by a constant.

In order to do so, we will need some definitions, which are illustrated in Figure 13.11. The **left spine** of a binary search tree T is the path from the root to the node with the smallest key. In other words, the left spine is the path from the root that consists of only left edges. Symmetrically, the **right spine** of T is the path from the root consisting of only right edges. The **length** of a spine is the number of nodes it contains.

- e.** Consider the treap T immediately after x is inserted using TREAP-INSERT. Let C be the length of the right spine of the left subtree of x . Let D be the length of the left spine of the right subtree of x . Prove that the total number of rotations that were performed during the insertion of x is equal to $C + D$.

We will now calculate the expected values of C and D . Without loss of generality, we assume that the keys are $1, 2, \dots, n$, since we are comparing them only to one another.

For nodes x and y , where $y \neq x$, let $k = \text{key}[x]$ and $i = \text{key}[y]$. We define indicator random variables

$$X_{i,k} = \mathbf{I}\{y \text{ is in the right spine of the left subtree of } x \text{ (in } T)\}.$$

- f.** Show that $X_{i,k} = 1$ if and only if $\text{priority}[y] > \text{priority}[x]$, $\text{key}[y] < \text{key}[x]$, and, for every z such that $\text{key}[y] < \text{key}[z] < \text{key}[x]$, we have $\text{priority}[y] < \text{priority}[z]$.

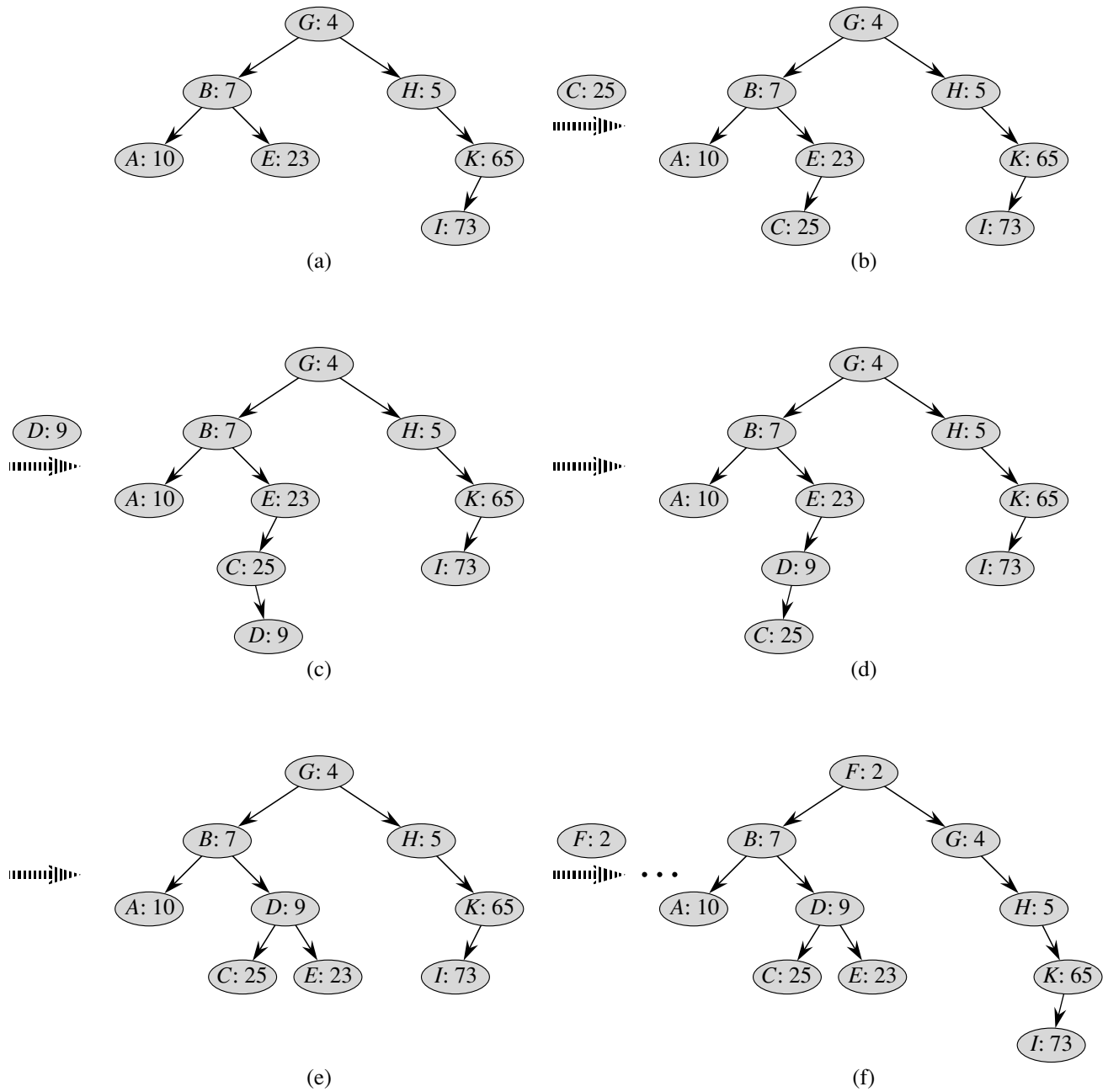


Figure 13.10 The operation of TREAP-INSERT. (a) The original treap, prior to insertion. (b) The treap after inserting a node with key *C* and priority 25. (c)–(d) Intermediate stages when inserting a node with key *D* and priority 9. (e) The treap after the insertion of parts (c) and (d) is done. (f) The treap after inserting a node with key *F* and priority 2.

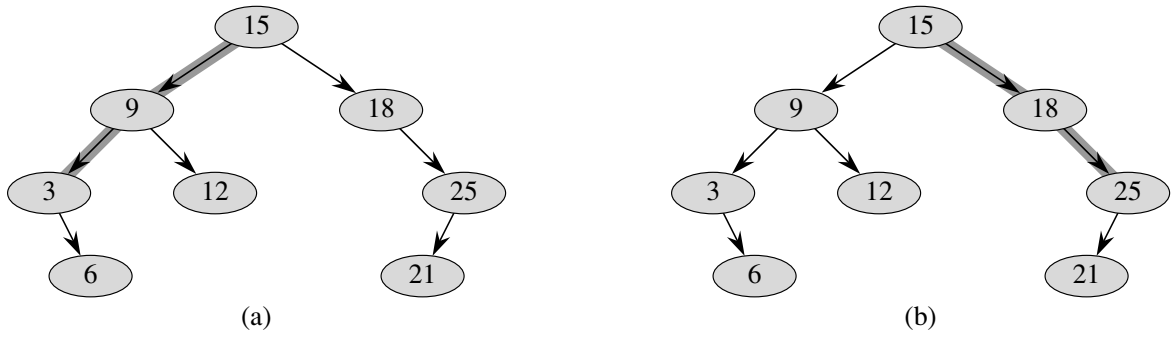


Figure 13.11 Spines of a binary search tree. The left spine is shaded in (a), and the right spine is shaded in (b).

g. Show that

$$\Pr\{X_{i,k} = 1\} = \frac{(k-i-1)!}{(k-i+1)!} = \frac{1}{(k-i+1)(k-i)}.$$

h. Show that

$$\mathbb{E}[C] = \sum_{j=1}^{k-1} \frac{1}{j(j+1)} = 1 - \frac{1}{k}.$$

i. Use a symmetry argument to show that

$$\mathbb{E}[D] = 1 - \frac{1}{n-k+1}.$$

j. Conclude that the expected number of rotations performed when inserting a node into a treap is less than 2.

Chapter notes

The idea of balancing a search tree is due to Adel'son-Vel'skiĭ and Landis [2], who introduced a class of balanced search trees called “AVL trees” in 1962, described in Problem 13-3. Another class of search trees, called “2-3 trees,” was introduced by J. E. Hopcroft (unpublished) in 1970. Balance is maintained in a 2-3 tree by manipulating the degrees of nodes in the tree. A generalization of 2-3 trees introduced by Bayer and McCreight [32], called B-trees, is the topic of Chapter 18.

Red-black trees were invented by Bayer [31] under the name “symmetric binary B-trees.” Guibas and Sedgwick [135] studied their properties at length and introduced the red/black color convention. Andersson [15] gives a simpler-to-code

variant of red-black trees. Weiss [311] calls this variant AA-trees. An AA-tree is similar to a red-black tree except that left children may never be red.

Treaps were proposed by Seidel and Aragon [271]. They are the default implementation of a dictionary in LEDA, which is a well-implemented collection of data structures and algorithms.

There are many other variations on balanced binary trees, including weight-balanced trees [230], k -neighbor trees [213], and scapegoat trees [108]. Perhaps the most intriguing are the “splay trees” introduced by Sleator and Tarjan [281], which are “self-adjusting.” (A good description of splay trees is given by Tarjan [292].) Splay trees maintain balance without any explicit balance condition such as color. Instead, “splay operations” (which involve rotations) are performed within the tree every time an access is made. The amortized cost (see Chapter 17) of each operation on an n -node tree is $O(\lg n)$.

Skip lists [251] are an alternative to balanced binary trees. A skip list is a linked list that is augmented with a number of additional pointers. Each dictionary operation runs in expected time $O(\lg n)$ on a skip list of n items.