Sorting and Searching:

A Cookbook

Thomas Niemann
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Introduction</td>
<td>7</td>
</tr>
<tr>
<td>2. Sorting</td>
<td>11</td>
</tr>
<tr>
<td>2.1 Insertion Sort</td>
<td>11</td>
</tr>
<tr>
<td>2.2 Shell Sort</td>
<td>12</td>
</tr>
<tr>
<td>2.3 Quicksort</td>
<td>14</td>
</tr>
<tr>
<td>2.4 Comparison of Methods</td>
<td>17</td>
</tr>
<tr>
<td>3. Dictionaries</td>
<td>18</td>
</tr>
<tr>
<td>3.1 Hash Tables</td>
<td>18</td>
</tr>
<tr>
<td>3.2 Binary Search Trees</td>
<td>21</td>
</tr>
<tr>
<td>3.3 Red-Black Trees</td>
<td>24</td>
</tr>
<tr>
<td>3.4 Skip Lists</td>
<td>28</td>
</tr>
<tr>
<td>3.5 Comparison of Methods</td>
<td>29</td>
</tr>
<tr>
<td>4. Code Listings</td>
<td>32</td>
</tr>
<tr>
<td>4.1 Insertion Sort Code</td>
<td>32</td>
</tr>
<tr>
<td>4.2 Shell Sort Code</td>
<td>33</td>
</tr>
<tr>
<td>4.3 Quicksort Code</td>
<td>34</td>
</tr>
<tr>
<td>4.4 Qsort Code</td>
<td>36</td>
</tr>
<tr>
<td>4.5 Hash Table Code</td>
<td>37</td>
</tr>
<tr>
<td>4.6 Binary Search Tree Code</td>
<td>38</td>
</tr>
<tr>
<td>4.7 Red-Black Tree Code</td>
<td>41</td>
</tr>
<tr>
<td>4.8 Skip List Code</td>
<td>46</td>
</tr>
<tr>
<td>5. Bibliography</td>
<td>50</td>
</tr>
</tbody>
</table>
Preface

This booklet contains a collection of sorting and searching algorithms. While many books on data structures describe sorting and searching algorithms, most assume a background in calculus and probability theory. Although a formal presentation and proof of asymptotic behavior is important, a more intuitive explanation is often possible.

The way each algorithm works is described in easy-to-understand terms. It is assumed that you have the knowledge equivalent to an introductory course in C or Pascal. In particular, you should be familiar with arrays and have a basic understanding of pointers. The material is presented in an orderly fashion, beginning with easy concepts and progressing to more complex ideas. Even though this collection is intended for beginners, more advanced users may benefit from some of the insights offered. In particular, the sections on hash tables and skip lists should prove interesting.

Santa Cruz, California
Thomas Niemann
March, 1995
1. Introduction

Arrays and linked lists are two basic data structures used to store information. We may wish to search, insert or delete records in a database based on a key value. This section examines the performance of these operations on arrays and linked lists.

Arrays

Figure 1.1 shows an array, seven elements long, containing numeric values. To search the array sequentially, we may use the algorithm in Figure 1.2. The maximum number of comparisons is 7, and occurs when the key we are searching for is in A[6]. If the data is sorted, a binary search may be done (Figure 1.3). Variables Lb and Ub keep track of the lower bound and upper bound of the array, respectively. We begin by examining the middle element of the array. If the key we are searching for is less than the middle element, then it must reside in the top half of the array. Thus, we set Ub to (M ÷ 1). This restricts our next iteration through the loop to the top half of the array. In this way, each iteration halves the size of the array to be searched. For example, the first iteration will leave 3 items to test. After the second iteration, there will be 1 item left to test. Thus it takes only three iterations to find any number.

Figure 1.1: An Array

This is a powerful method. For example, if the array size is 1023, we can narrow the search to 511 items in one comparison. Another comparison, and we're looking at only 255 elements. In fact, only 10 comparisons are needed to search an array containing 1023 elements.

In addition to searching, we may wish to insert or delete entries. Unfortunately, an array is not a good arrangement for these operations. For example, to insert the number 18 in Figure 1.1, we would need to shift A[3]ÈA[6] down by one slot. Then we could copy number 18 into A[3]. A similar problem arises when deleting numbers. To improve the efficiency of insert and delete operations, linked lists may be used.

int function SequentialSearch (Array A , int Lb , int Ub , int Key );
begin
for i = Lb to Ub do
if A ( i ) =  Key then
return i ;
return Ñ1;
end;

Figure 1.2: Sequential Search

int function BinarySearch (Array A , int Lb , int Ub , int Key );
begin
do forever
M = ( Lb + Ub )/2;
if ( Key < A[M]) then
Ub = M Ñ 1;
else if (Key > A[M]) then
    Lb = M + 1;
else
    return M ;
if (Lb > Ub ) then
    return Ñ1;
end;

Figure 1.3: Binary Search

Linked Lists

In Figure 1.4 we have the same values stored in a linked list. Assuming pointers X and P, as shown in the figure, value 18 may be inserted as follows:

X->Next = P->Next;
P->Next = X;

Insertion (and deletion) are very efficient using linked lists. You may be wondering how P was set in the first place. Well, we had to search the list in a sequential fashion to find the insertion point for X. Thus, while we improved our insert and delete performance, it has been at the expense of search time.

Figure 1.4: A Linked List

Timing Estimates

Several methods may be used to compare the performance of algorithms. One way is simply to run several tests for each algorithm and compare the timings. Another way is to estimate the time required. For example, we may state that search time is O(n) (big-oh of n). This means that, for large n, search time is no greater than the number of items n in the list. The big-O notation does not describe the exact time that an algorithm takes, but only indicates an upper bound on execution time within a constant factor. If an algorithm takes O(n^2) time, then execution time grows no worse than the square of the size of the list. To see the effect this has, Table 1.1 illustrates growth rates for various functions. A growth rate of O(lg n) occurs for algorithms similar to the binary search. The lg (logarithm, base 2) function increases by one when n is doubled. Recall that we can search twice as many items with one more comparison in the binary search. Thus the binary search is a O(lg n) algorithm.

1.1: Growth Rates
If the values in Table 1.1 represented microseconds, then a $O(lg n)$ algorithm may take 20 microseconds to process 1,048,476 items, a $O(n^{1.25})$ algorithm might take 33 seconds, and a $O(n^2)$ algorithm might take up to 12 days! In the following chapters a timing estimate for each algorithm, using big-O notation, will be included. For a more formal derivation of these formulas you may wish to consult the references.

Summary

As we have seen, sorted arrays may be searched efficiently using a binary search. However, we must have a sorted array to start with. In the next section various ways to sort arrays will be examined. It turns out that this is computationally expensive, and considerable research has been done to make sorting algorithms as efficient as possible.

Linked lists improved the efficiency of insert and delete operations, but searches were sequential and time-consuming. Algorithms exist that do all three operations efficiently, and they will be the discussed in the section on dictionaries.

1. Sorting
   a) Insertion Sort

   One of the simplest methods to sort an array is sort by insertion. An example of an insertion sort occurs in everyday life while playing cards. To sort the cards in your hand you extract a card, shift the remaining cards, and then insert the extracted card in the correct place. This process is repeated until all the cards are in the correct sequence. Both average and worst-case time is $O(n^2)$. For further reading, consult Knuth[1].

Theory

In Figure 2.1(a) we extract the 3. Then the above elements are shifted down until we find the correct place to insert the 3. This process repeats in Figure 2.1(b) for the number 1. Finally, in Figure 2.1(c), we complete the sort by inserting 2 in the correct place. Assuming there are n elements in the array, we must index through n Ñ 1 entries. For each entry, we may need to examine and shift up to n Ñ 1 other entries. For this reason, sorting is a time-consuming process.
The insertion sort is an in-place sort. That is, we sort the array in-place. No extra memory is required. The insertion sort is also a stable sort. Stable sorts retain the original ordering of keys when identical keys are present in the input data.

Figure 2.1: Insertion Sort

Implementation

An ANSI-C implementation for insertion sort may be found in Section 4.1 (page ). Typedef T and comparison operator CompGT should be altered to reflect the data stored in the table. Pointer arithmetic was used, rather than array references, for efficiency.

a) Shell Sort

Shell sort, developed by Donald L. Shell, is a non-stable in-place sort. Shell sort improves on the efficiency of insertion sort by quickly shifting values to their destination. Average sort time is $O(n^{1.25})$, while worst-case time is $O(n^{1.5})$. For further reading, consult Knuth[1].

Theory

In Figure 2.2(a) we have an example of sorting by insertion. First we extract 1, shift 3 and 5 down one slot, and then insert the 1. Thus, two shifts were required. In the next frame, two shifts are required before we can insert the 2. The process continues until the last frame, where a total of $2 + 2 + 1 = 5$ shifts have been made.

In Figure 2.2(b) an example of shell sort is illustrated. We begin by doing an insertion sort using a spacing of two. In the first frame we examine numbers 3-1. Extracting 1, we shift 3 down one slot for a shift count of 1. Next we examine numbers 5-2. We extract 2, shift 5 down, and then insert the 2. After sorting with a spacing of two, a final pass is made with a spacing of one. This is simply the traditional insertion sort. The total shift count using shell sort is $1+1+1 = 3$. By using an initial spacing larger than one, we were able to quickly shift values to their proper destination.

Figure 2.2: Shell Sort

To implement shell sort, various spacings may be used. Typically the array is sorted with a large spacing, the spacing reduced, and the array sorted again. On the final sort, spacing is one. Although shell sort is easy to comprehend, formal analysis is difficult. In particular, optimal spacing values elude theoreticians. Knuth[1] has experimented with several values and recommends that spacing ($h$) for an array of size $N$ be based on the following formula:
Thus, values of $h$ are computed as follows:

To sort 100 items we first find an $h_5$ such that $h_5 \geq 100$. For 100
items, $h_5$ is selected. Our final value ($h_t$) is two steps lower, or $h_3$.
Thus, our sequence of $h$ values will be 13-4-1. Once the initial $h$ value
has been determined, subsequent values may be calculated using the
formula

Implementation

An ANSI-C implementation of shell sort may be found in Section 4.2 (page
).

Typedef $T$ and comparison operator CompGT should be altered to reflect the
data stored in the array. When computing $h$, care must be taken to avoid
underflows or overflows. The central portion of the algorithm is an
insertion sort with a spacing of $h$. To terminate the inner loop
correctly, it is necessary to compare $J$ before decrement. Otherwise, pointer values
may wrap through zero, resulting in unexpected behavior.
b) Quicksort

While the shell sort algorithm is significantly better than insertion sort, there is still room for improvement. One of the most popular sorting algorithms is quicksort. Quicksort executes in $O(n \log n)$ on average, and $O(n^2)$ in the worst-case. However, with proper precautions, worst-case behavior is very unlikely. Quicksort is a non-stable sort. It is not an in-place sort as stack space is required. For further reading, consult Cormen[2].

Theory

The quicksort algorithm works by partitioning the array to be sorted, then recursively sorting each partition. In Partition (Figure 2.3), one of the array elements is selected as a pivot value. Values smaller than the pivot value are placed to the left of the pivot, while larger values are placed to the right.

```
int function Partition (Array A, int Lb, int Ub);
begin
select a pivot from A[Lb]ÉA[Ub];
reorder A[Lb]ÉA[Ub] such that:
all values to the left of the pivot are £ pivot
all values to the right of the pivot are £ pivot
return pivot position;
end;
```

```
procedure QuickSort (Array A, int Lb, int Ub);
begin
if Lb Ub then
  M = Partition (A, Lb, Ub);
  QuickSort (A, Lb, M Ñ 1);
  QuickSort (A, M + 1, Ub);
end;
```

Figure 2.3: Quicksort Algorithm

In Figure 2.4(a), the pivot selected is 3. Indices are run starting at both ends of the array. Index i starts on the left and selects an element that is larger than the pivot, while index j starts on the right and selects an element that is smaller than the pivot. These elements are then exchanged, as is shown in Figure 2.4(b). QuickSort recursively sorts the two subarrays, resulting in the array shown in Figure 2.4(c).

Figure 2.4: Quicksort Example

As the process proceeds, it may be necessary to move the pivot so that correct ordering is maintained. In this manner, QuickSort succeeds in sorting the array. If we're lucky the pivot selected will be the median of all values, thus equally
dividing the array. For a moment, let's assume that this is the case. Since the array is split in half at each step, and Partition must eventually examine all n elements, the run time is \( O(n \log n) \).

To find a pivot value, Partition could simply select the first element \( A[Lb] \). All other values would be compared to the pivot value, and placed either to the left or right of the pivot as appropriate. However, there is one case that fails miserably. Suppose the array was originally in order. Partition would always select the lowest value as a pivot and split the array with one element in the left partition, and \( Ub \leq Lb \) elements in the other. Each recursive call to quicksort would only diminish the size of the array to be sorted by one. Thus, \( n \) recursive calls would be required to do the sort, resulting in a \( O(n^2) \) run time. One solution to this problem is to randomly select an item as a pivot. This would make it extremely unlikely that worst-case behavior would occur.

**Implementation**

An ANSI-C implementation of the quicksort algorithm may be found in Section 4.3 (page\( \text{Ê} \)). Typedef \( T \) and comparison operator \( \text{CompGT} \) should be altered to reflect the data stored in the array. Several enhancements have been made to the basic quicksort algorithm:

- The center element is selected as a pivot in Partition. If the list is partially ordered, this will be a good choice. Worst-case behavior occurs when the center element happens to be the largest or smallest element each time Partition is invoked.
- For short arrays, InsertSort is called. Due to recursion and other overhead, quicksort is not an efficient algorithm to use on small arrays. Thus, any array with fewer than 12 elements is sorted using an insertion sort. The optimal cutoff value is not critical and varies based on the quality of generated code.
- Tail recursion occurs when the last statement in a function is a call to the function itself. Tail recursion may be replaced by iteration which results in a better utilization of stack space. This has been done
the second call to QuickSort in Figure 2.3.

- After an array is partitioned, the smallest partition is sorted first. This results in a better utilization of stack space, as short partitions are quickly sorted and dispensed with.
- Pointer arithmetic, rather than array indices, is used for efficient execution.

Also included is a listing for qsort (Section 4.4, page ), an ANSI-C standard library function usually implemented with quicksort. For this implementation, recursive calls were replaced by explicit stack operations. Table 2.1 shows timing statistics and stack utilization before and after the enhancements were applied.

b) Comparison of Methods

In this section we will compare the sorting algorithms covered: insertion sort, shell sort and quicksort. There are several factors that influence the choice of a sorting algorithm:

- Stable sort. Recall that a stable sort will leave identical keys in the same relative position in the sorted output. Insertion sort is the only algorithm covered that is stable.

- Space. An in-place sort does not require any extra space to accomplish its task. Both insertion sort and shell sort are in-place sorts. Quicksort requires stack space for recursion, and thus is not an in-place sort. However, the amount required was considerably reduced by tinkering with the algorithm.

- Time. The time required to sort a dataset can easily become astronomical (Table 1.1). Table 2.2 shows the relative timings for each method. Actual timing tests are described below.

- Simplicity. The number of statements required for each algorithm may be found in Table 2.2. Simpler algorithms result in fewer programming errors.

The time required to sort a randomly ordered dataset is shown in Table 2.3.

1. Dictionaries

   a) Hash Tables
A dictionary requires that search, insert and delete operations be supported. One of the most effective ways to implement a dictionary is through the use of hash tables. Average time to search for an element is \( O(1) \), while worst-case time is \( O(n) \). Cormen[2] and Knuth[1] both contain excellent discussions on hashing. In case you decide to read more material on this topic, you may want to know some terminology. The technique presented here is chaining, also known as open hashing[3]. An alternative technique, known as closed hashing[3], or open addressing[1], is not presented. Got that?

Theory
A hash table is simply an array that is addressed via a hash function. For example, in Figure 3.1, HashTable is an array with 8 elements. Each element is a pointer to a linked list of numeric data. The hash function for this example simply divides the data key by 8, and uses the remainder as an index into the table. This yields a number from 0 to 7. Since the range of indices for HashTable is 0 to 7, we are guaranteed that the index is valid.

Figure 3.1: A Hash Table

To insert a new item in the table, we hash the key to determine which list the item goes on, and then insert the item at the beginning of the list. For example, to insert 11, we divide 11 by 8 giving a remainder of 3. Thus, 11 goes on the list starting at HashTable[3]. To find a number, we hash the number and chain down the correct list to see if it is in the table. To delete a number, we find the number and remove the node from the linked list.

If the hash function is uniform, or equally distributes the data keys among the hash table indices, then hashing effectively subdivides the list to be searched. Worst-case behavior occurs when all keys hash to the same index. Then we simply have a single linked list that must be sequentially scanned. Consequently, it is important to choose a good hash function. Several methods may be used to hash key values. To illustrate the techniques, I will assume unsigned char is 8-bits, unsigned short int is 16-bits and unsigned long int is 32-bits.

- Division method (tablesize = prime). This technique was used in the preceding example. A HashValue, from 0 to (HashTableSize - 1), is computed by dividing the key value by the size of the hash table and taking the remainder. For example:

```c
typedef int HashIndexType;
HashIndexType Hash(int Key) {
    return Key % HashTableSize;
}
```

- Selecting an appropriate HashTableSize is important to the success of this method. For example, a HashTableSize of two would yield even hash values for even Keys, and odd hash values for odd Keys. This is an undesirable property, as all keys would hash to the same value if they happened to be even. If HashTableSize is a power of two, then the hash function simply selects a subset of the Key bits as the table index. To obtain a more random scattering, HashTableSize should be a prime number not too close to a power of two.

- Multiplication method (tablesize = 2n). The multiplication method may be used for a HashTableSize that is a power of 2. The Key is multiplied by a constant, and then the necessary bits are extracted to index into the table. Knuth[1] recommends using the golden ratio, or \( \frac{1 + \sqrt{5}}{2} \), as the constant. The following definitions may be used for the multiplication method:

```c
/* 8-bit index */
typedef unsigned char HashIndexType;
static const HashIndexType K = 158;

/* 16-bit index */
typedef unsigned short int HashIndexType;
static const HashIndexType K = 40503;
```
typedef unsigned long int HashIndexType;
static const HashIndexType K = 2654435769;

/* w=bitwidth(HashIndexType), size of table=2**m */
static const int S = w - m;
HashIndexType HashValue = (HashIndexType)(K * Key) >> S;
For example, if HashTableSize is 1024 (2^10), then a 16-bit index is sufficient
and S would be assigned a value of 16 ÷ 10 = 6. Thus, we have:
typedef unsigned short int HashIndexType;

HashIndexType Hash(int Key) {
  static const HashIndexType K = 40503;
  static const int S = 6;
  return (HashIndexType)(K * Key) >> S;
}

· Variable string addition method (tablesize = 256). To hash a variable-
  length string, each character is added, modulo 256, to a total. A HashValue,
  range 0-255, is computed.
typedef unsigned char HashIndexType;

HashIndexType Hash(char *str) {
  HashIndexType h = 0;
  while (*str) h += *str++;
  return h;
}

· Variable string exclusive-or method (tablesize = 256). This method is
  similar to the addition method, but successfully distinguishes similar words and
  anagrams. To obtain a hash value in the range 0-255, all bytes in the string
  are exclusive-or'd together. However, in the process of doing each exclusive-
  or, a random component is introduced.
typedef unsigned char HashIndexType;
unsigned char Rand8[256];

HashIndexType Hash(char *str) {
  unsigned char h = 0;
  while (*str) h = Rand8[h ^ *str++];
  return h;
}

Rand8 is a table of 256 8-bit unique random numbers. The exact ordering is not
critical. The exclusive-or method has its basis in cryptography, and is quite
effective[4].

· Variable string exclusive-or method (tablesize £ 65536). If we hash the
  string twice, we may derive a hash value for an arbitrary table size up to
  65536. The second time the string is hashed, one is added to the first
  character. Then the two 8-bit hash values are concatenated together to form a
  16-bit hash value.
typedef unsigned short int HashIndexType;
unsigned char Rand8[256];

HashIndexType Hash(char *str) {
  HashIndexType h;
  unsigned char h1, h2;
  if (*str == 0) return 0;
  h1 = *str; h2 = *str + 1;
  str++;
  while (*str) {
    h1 = Rand8[h1 ^ *str];
    h2 = Rand8[h2 ^ *str];
  }
Assuming \( n \) data items, the hash table size should be large enough to accommodate a reasonable number of entries. As seen in Table 3.1, a small table size substantially increases the average time to find a key. A hash table may be viewed as a collection of linked lists. As the table becomes larger, the number of lists increases, and the average number of nodes on each list decreases. If the table size is 1, then the table is really a single linked list of length \( n \). Assuming a perfect hash function, a table size of 2 has two lists of length \( n/2 \). If the table size is 100, then we have 100 lists of length \( n/100 \). This considerably reduces the length of the list to be searched. As we see in Table 3.1, there is much leeway in the choice of table size.

![Table 3.1: HashTableSize vs. Average Search Time (ms), 4096 entries](image)

**Implementation**

An ANSI-C implementation of a hash table may be found in Section 4.5 (page ). Typedef \( T \) and comparison operator \( \text{CompEQ} \) should be altered to reflect the data stored in the table. \( \text{HashTableSize} \) must be determined and the \( \text{HashTable} \) allocated. The division method was used in the Hash function. \( \text{InsertNode} \) allocates a new node and inserts it in the table. \( \text{DeleteNode} \) deletes and frees a node from the table. \( \text{FindNode} \) searches the table for a particular value.

**a) Binary Search Trees**

In Section 1 we used the binary search algorithm to find data stored in an array. This method is very effective, as each iteration reduced the number of items to search by one-half. However, since data was stored in an array, insertions and deletions were not efficient. Binary search trees store data in nodes that are linked in a tree-like fashion. For randomly inserted data, search time is \( O(\lg n) \). Worst-case behavior occurs when ordered data is inserted. In this case the search time is \( O(n) \). See Cormen[2] for a more detailed description.

**Theory**

A binary search tree is a tree where each node has a left and right child. Either child, or both children, may be missing. Figure 3.2 illustrates a binary search tree. Assuming Key represents the value of a given node, then a binary search tree also has the following property: all children to the left of the node have values smaller than Key, and all children to the right of the node have values larger than Key. The top of a tree is known as the root, and the exposed nodes at the bottom are known as leaves. In Figure 3.2, the root is node 20 and the leaves are nodes 4, 16, 37 and 43. The height of a tree is the length of the longest path from root to leaf. For this example the tree height is 2.

![Figure 3.2: A Binary Search Tree](image)

To search a tree for a given value, we start at the root and work down. For example, to search for 16, we first note that 16 < 20 and we traverse to the left child. The second comparison finds that 16 > 7, so we traverse to the right child. On the third comparison, we succeed.
Each comparison results in reducing the number of items to inspect by one-half. In this respect, the algorithm is similar to a binary search on an array. However, this is true only if the tree is balanced. For example, Figure 3.3 shows another tree containing the same values. While it is a binary search tree, its behavior is more like that of a linked list, with search time increasing proportional to the number of elements stored.

Figure 3.3: An Unbalanced Binary Search Tree
Insertion and Deletion

Let us examine insertions in a binary search tree to determine the conditions that can cause an unbalanced tree. To insert an 18 in the tree in Figure 3.2, we first search for that number. This causes us to arrive at node 16 with nowhere to go. Since 18 > 16, we simply add node 18 to the right child of node 16 (Figure 3.4).

Now we can see how an unbalanced tree can occur. If the data is presented in an ascending sequence, each node will be added to the right of the previous node. This will create one long chain, or linked list. However, if data is presented for insertion in a random order, then a more balanced tree is possible.

Deletions are similar, but require that the binary search tree property be maintained. For example, if node 20 in Figure 3.4 is removed, it must be replaced by node 37. This results in the tree shown in Figure 3.5. The rationale for this choice is as follows. The successor for node 20 must be chosen such that all nodes to the right are larger. Thus, we need to select the smallest valued node to the right of node 20. To make the selection, chain once to the right (node 38), and then chain to the left until the last node is found (node 37). This is the successor for node 20.

Figure 3.4: Binary Tree After Adding Node 18

Figure 3.5: Binary Tree After Deleting Node 20

Implementation

An ANSI-C implementation of a binary search tree may be found in Section 4.6 (page E). Typedef T and comparison operators CompLT and CompEQ should be altered to reflect the data stored in the tree. Each Node consists of Left, Right and Parent pointers designating each child and the parent. Data is stored in the Data field. The tree is based at Root, and is initially NULL. InsertNode allocates a new node and inserts it in the tree. DeleteNode deletes and frees a node from the tree. FindNode searches the tree for a particular value.

a) Red-Black Trees

Binary search trees work best when they are balanced or the path length from root to any leaf is within some bounds. The red-black tree algorithm is a method for balancing trees. The name derives from the fact that each node is colored red or black, and the color of the node is instrumental in determining the balance of the tree. During insert and delete operations, nodes may be rotated to maintain tree balance. Both average and worst-case search time is O(lg n).

This is, perhaps, the most difficult section in the book. If you get glassy-eyed looking at tree rotations, try skipping to skip lists, the next section. For further reading, Cormen[2] has an excellent section on red-black trees.

Theory

A red-black tree is a balanced binary search tree with the following properties[2]:
1. Every node is colored red or black.
2. Every leaf is a NIL node, and is colored black.
3. If a node is red, then both its children are black.
4. Every simple path from a node to a descendant leaf contains the same number of black nodes.

The number of black nodes on a path from root to leaf is known as the black height of a tree. These properties guarantee that any path from the root to a leaf is no more than twice as long as any other. To see why this is true, consider a tree with a black height of two. The shortest distance from root to leaf is two, where both nodes are black. The longest distance from root to leaf is four, where the nodes are colored (root to leaf): red, black, red, black. It is not possible to insert more black nodes as this would violate property 4, the black-height requirement. Since red nodes must have black children (property
3), having two red nodes in a row is not allowed. Thus, the largest path we can construct consists of an alternation of red-black nodes, or twice the length of a path containing only black nodes. All operations on the tree must maintain the properties listed above. In particular, operations which insert or delete items from the tree must abide by these rules.
Insertion
To insert a node, we search the tree for an insertion point, and add the node to the tree. A new node will always be inserted as a leaf node at the bottom of the tree. After insertion, the node is colored red. Then the parent of the node is examined to determine if the red-black tree properties have been violated. If necessary, we recolor the node and do rotations to balance the tree.

By inserting a red node, we have preserved black-height property (property 4). However, property 3 may be violated. This property states that both children of a red node must be black. While both children of the new node are black (they're NIL), consider the case where the parent of the new node is red. Inserting a red node under a red parent would violate this property. There are two cases to consider:

- Red parent, red uncle: Figure 3.6 illustrates a red-red violation. Node X is the newly inserted node, with both parent and uncle colored red. A simple recoloring removes the red-red violation. After recoloring, the grandparent (node B) must be checked for validity, as its parent may be red. Note that this has the effect of propagating a red node up the tree. On completion, the root of the tree is marked black. If it was originally red, then this has the effect of increasing the black-height of the tree.

- Red parent, black uncle: Figure 3.7 illustrates a red-red violation, where the uncle is colored black. Here the nodes may be rotated, with the subtrees adjusted as shown. At this point the algorithm may terminate as there are no red-red conflicts and the top of the subtree (node A) is colored black. Note that if node X was originally a right child, a left rotation would be done first, making the node a left child.

Each adjustment made while inserting a node causes us to travel up the tree one step. At most 1 rotation (2 if the node is a right child) will be done, as the algorithm terminates in this case. The technique for deletion is similar.

Figure 3.6: Insertion Ñ Red Parent, Red Uncle

Figure 3.7: Insertion Ñ Red Parent, Black Uncle
Implementation
An ANSI-C implementation of a red-black tree may be found in Section 4.7 (pageÊ ). Typedef T and comparison operators CompLT and CompEQ should be altered to reflect the data stored in the tree. Each Node consists of Left, Right and Parent pointers designating each child and the parent. The node color is stored in Color, and is either Red or Black. The data is stored in the Data field. All leaf nodes of the tree are Sentinel nodes, to simplify coding. The tree is based at Root, and initially is a Sentinel node.

InsertNode allocates a new node and inserts it in the tree. Subsequently, it calls InsertFixup to ensure that the red-black tree properties are maintained. DeleteNode deletes a node from the tree. To maintain red-black tree properties, DeleteFixup is called. FindNode searches the tree for a particular value.

Skip Lists
Skip lists are linked lists that allow you to skip to the correct node. Thus the performance bottleneck inherent in a sequential scan is avoided, while insertion and deletion remain relatively efficient. Average search time is \( O(\lg n) \). Worst-case search time is \( O(n) \), but is extremely unlikely. An excellent reference for skip lists is Pugh[5].

Theory
The indexing scheme employed in skip lists is similar in nature to the method used to lookup names in an address book. To lookup a name, you index to the tab representing the first character of the desired entry. In Figure 3.8, for example, the top-most list represents a simple linked list with no tabs. Adding tabs (middle figure) facilitates the search. In this case, level-1 pointers are traversed. Once the correct segment of the list is found, level-0 pointers are traversed to find the specific entry.

The indexing scheme may be extended as shown in the bottom figure, where we now have an index to the index. To locate an item, level-2 pointers are traversed until the correct segment of the list is identified. Subsequently, level-1 and level-0 pointers are traversed.

During insertion the number of pointers required for a new node must be determined. This is easily resolved using a probabilistic technique. A random number generator is used to toss a computer coin. When inserting a new node, the coin is tossed to determine if it should be level-1. If you win, the coin is tossed again to determine if the node should be level-2. Another win, and the coin is tossed to determine if the node should be level-3. This process repeats until you lose.

The skip list algorithm has a probabilistic component, and thus has a probabilistic bounds on the time required to execute. However, these bounds are quite tight in normal circumstances. For example, to search a list containing 1000 items, the probability that search time will be 5 times the average is about \( \frac{1}{1,000,000,000,000,000,000} \).[5]

Figure 3.8: Skip List Construction
Implementation
An ANSI-C implementation of a skip list may be found in Section 4.8 (pageÊ ). Typedef T and comparison operators CompLT and CompEQ should be altered to reflect the data stored in the list. In addition, MAXLEVEL should be set based on the maximum size of the dataset.

To initialize, InitList is called. The list header is allocated and initialized. To indicate an empty list, all levels are set to point to the header. InsertNode allocates a new node and inserts it in the list. InsertNode first searches for the correct insertion point. While searching, the update array maintains pointers to the upper-level nodes encountered. This information is subsequently used to establish correct links for the newly inserted node. NewLevel is determined using a random number generator, and the node allocated. The forward links are then established using information from the update array.
DeleteNode deletes and frees a node, and is implemented in a similar manner. FindNode searches the list for a particular value.

a) Comparison of Methods

We have seen several ways to construct dictionaries: hash tables, unbalanced binary search trees, red-black trees and skip lists. There are several factors that influence the choice of an algorithm:

- **Sorted output.** If sorted output is required, then hash tables are not a viable alternative. Entries are stored in the table based on their hashed value, with no other ordering. For binary trees, the story is different. An in-order tree walk will produce a sorted list. For example:

```c
void WalkTree(Node *P) {
    if (P == NIL) return;
    WalkTree(P->Left);
    /* examine P->Data here */
    WalkTree(P->Right);
}
WalkTree(Root);
```

- **To examine skip list nodes in order, simply chain through the level-0 pointers.** For example:

```c
Node *P = List.Hdr->Forward[0];
while (P != NIL) {
    /* examine P->Data here */
    P = P->Forward[0];
}
```

- **Space.** The amount of memory required to store a value should be minimized. This is especially true if many small nodes are to be allocated.
  - For hash tables, only one forward pointer per node is required. In addition, the hash table itself must be allocated.
  - For red-black trees, each node has a left, right and parent pointer. In addition, the color of each node must be recorded. Although this requires only one bit, more space may be allocated to ensure that the size of the structure is properly aligned. Thus, each node in a red-black tree requires enough space for 3-4 pointers.
  - For skip lists, each node has a level-0 forward pointer. The probability of having a level-1 pointer is $\frac{1}{2}$. The probability of having a level-2 pointer is $\frac{1}{4}$. In general, the number of forward pointers per node is 4.

- **Time.** The algorithm should be efficient. This is especially true if a large dataset is expected. Table 3.2 compares the search time for each algorithm. Note that worst-case behavior for hash tables and skip lists is extremely unlikely. Actual timing tests are described below.

  - **Simplicity.** If the algorithm is short and easy to understand, fewer mistakes may be made. This not only makes your life easy, but the maintenance programmer entrusted with the task of making repairs will appreciate any efforts you make in this area. The number of statements required for each algorithm is listed in Table 3.2.

```
method statements average timeworst-case time hash
table 26 O(1) O(n) unbalanced tree 41 O(lg n) O(n) red-black tree 120 O(lg n) O(lg n) O(n) skip list 55 O(lg n) O(n)
```

Table 3.2: Comparison of Dictionaries

Average time for insert, search and delete operations on a database of 65,536 (216) randomly input items may be found in Table 3.3. For this test the hash table size was 10,009 and 16 index levels were allowed for the skip list. While there is some variation in the timings for the four methods, they are close enough so that other considerations should come into play when selecting an algorithm.
Table 3.4 shows the average search time for two sets of data: a random set, where all values are unique, and an ordered set, where values are in ascending order. Ordered input creates a worst-case scenario for unbalanced tree algorithms, as the tree ends up being a simple linked list. The times shown are for a single search operation. If we were to search for all items in a database of 65,536 values, a red-black tree algorithm would take .6 seconds, while an unbalanced tree algorithm would take 1 hour.

```
typedef int T;
typedef int TblIndex;

#define CompGT(a,b) (a > b)

void InsertSort(T *Lb, T *Ub) {

    Jmin = Lb - 1;
    for (I = Lb + 1; I <= Ub; I++) {
        V = *I;
        /* Shift elements down until insertion point found. */
        for (J = I-1; J != Jmin && CompGT(*J, V); J--)
            *(J+1) = *J;
        *(J+1) = V;
    }
}
```
typedef int T;
typedef int TblIndex;
#define CompGT(a,b) (a > b)

void ShellSort(T *Lb, T *Ub) {
    TblIndex H, N;

    /**************************************************************************
    * Sort array A[Lb..Ub] *
    **************************************************************************/

    /* compute largest increment */
    N = Ub - Lb + 1;
    H = 1;
    if (N < 14)
        H = 1;
    else if (sizeof(TblIndex) == 2 && N > 29524)
        H = 3280;
    else {
        while (H < N) H = 3*H + 1;
        H /= 3;
        H /= 3;
    }

    while (H > 0) {

        /* sort-by-insertion in increments of H */
        /* Care must be taken for pointers that */
        /* wrap through zero. */
        Min = Lb + H;
        for (I = Min; I <= Ub; I++) {
            V = *I;
            for (J = I-H; CompGT(*J, V); J -= H) {
                *(J+H) = *J;
                if (J <= Min) { J -= H; break; }
            }
            *(J+H) = V;
        }

        /* compute next increment */
        H /= 3;
    }
}
a) Quicksort Code

typedef int T;
typedef int TblIndex;

#define CompGT(a,b) (a > b)

T *Partition(T *Lb, T *Ub) {
    unsigned int Offset;

    *****************************/
    * partition Array[Lb..Ub] *
    *****************************/

    /* select pivot and exchange with 1st element */
    Offset = (Ub - Lb)>>1;
    P = Lb + Offset;
    Pivot = *P;
    *P = *Lb;
    I = Lb + 1;
    J = Ub;
    while (I) {
        while (I < J && CompGT(Pivot, *I)) I++;
        while (J >= I && CompGT(*J, Pivot)) J--;
        if (I >= J) break;
        V = *I;
        *I = *J;
        *J = V;
        J--; I++;
    }

    /* pivot belongs in A[j] */
    *Lb = *J;
    *J = Pivot;
    return J;
}

void QuickSort(T *Lb, T *Ub) {
    T *M;

    *****************************/
    * Sort array A[Lb..Ub] *
    *****************************/

    while (Lb < Ub) {
        /* quickly sort short lists */
        if (Ub - Lb <= 12) {
            InsertSort(Lb, Ub);
            return;
        }

        /* partition into two segments */
        M = Partition (Lb, Ub);

        /* sort the smallest partition */
        /* to minimize stack requirements */
if (M - Lb <= Ub - M) {
    QuickSort(Lb, M - 1);
    Lb = M + 1;
} else {
    QuickSort(M + 1, Ub);
    Ub = M - 1;
}
}
a) Qsort Code

#include <limits.h>
#define MAXSTACK (sizeof(size_t) * CHAR_BIT)

static void Exchange(void *a, void *b, size_t size) {
    size_t i;

    for (i = sizeof(int); i <= size; i += sizeof(int)) {
        int t = (*((int *)a));
        (*((int *)a)++) = (*((int *)b));
        (*((int *)b)++) = t;
    }
    for (i = i - sizeof(int) + 1; i <= size; i++) {
        char t = (*((char *)a));
        (*((char *)a)++) = (*((char *)b));
        (*((char *)b)++) = t;
    }
}

void qsort(void *base, size_t nmemb, size_t size, 
int (*compar)(const void *, const void *)) {
    void *LbStack[MAXSTACK], *UbStack[MAXSTACK];
    int sp;
    unsigned int Offset;

    LbStack[0] = (char *)base;
    UbStack[0] = (char *)base + (nmemb-1)*size;
    for (sp = 0; sp >= 0; sp--) {
        char *Lb, *Ub, *M;

        Lb = LbStack[sp];
        Ub = UbStack[sp];

        while (Lb < Ub) {
            /* select pivot and exchange with 1st element */
            Offset = (Ub - Lb) >> 1;
            P = Lb + Offset - Offset % size;
            Exchange (Lb, P, size);

            /* partition into two segments */
            I = Lb + size;
            J = Ub;
            while (I) {
                while (I < J && compar(Lb, I) > 0) I += size;
                while (J >= I && compar(J, Lb) > 0) J -= size;
                if (I >= J) break;
                Exchange (I, J, size);
                J -= size;
                I += size;
            }
        }
    }
}
/* pivot belongs in A[j] */
Exchange (Lb, J, size);
M = J;

/* keep processing smallest segment, and stack largest */
if (M - Lb <= Ub - M) {
    if (M + size < Ub) {
        LbStack[sp] = M + size;
        UbStack[sp++] = Ub;
    }
    Ub = M - size;
} else {
    if (M - size > Lb) {
        LbStack[sp] = Lb;
        UbStack[sp++] = M - size;
    }
    Lb = M + size;
}

a) Hash Table Code

#include <stdlib.h>
#include <stdio.h>

/* modify these lines to establish data type */
typedef int T;
#define CompEQ(a,b) (a == b)

typedef struct Node_ {
    struct Node_ *Next; /* next node */
    T Data; /* data stored in node */
} Node;

typedef int HashTableIndex;

Node **HashTable;
int HashTableSize;

HashTableIndex Hash(T Data) {
    /* division method */
    return (Data % HashTableSize);
}

Node *InsertNode(T Data) {
    Node *p, *p0;
    HashTableIndex bucket;

    /* insert node at beginning of list */
    bucket = Hash(Data);
    if ((p = malloc(sizeof(Node))) == 0) {
        fprintf (stderr, "out of memory (InsertNode)\n");
        exit(1);
    }
    p0 = HashTable[bucket];
    HashTable[bucket] = p;
    p->Next = p0;
p->Data = Data;
return p;
}

void DeleteNode(T Data) {
Node *p0, *p;
HashTableIndex bucket;

/* find node */
p0 = 0;
bucket = Hash(Data);
p = HashTable[bucket];
while (p && !CompEQ(p->Data, Data)) {
    p0 = p;
    p = p->Next;
}
if (!p) return;

/* p designates node to delete, remove it from list */
if (p0)
    /* not first node, p0 points to previous node */
p0->Next = p->Next;
else
    /* first node on chain */
    HashTable[bucket] = p->Next;

free (p);
}

Node *FindNode (T Data) {
Node *p;
p = HashTable[Hash(Data)];
while (p && !CompEQ(p->Data, Data))
    p = p->Next;
return p;
}

a) Binary Search Tree Code

#include <stdio.h>
#include <stdlib.h>

/* modify these lines to establish data type */
typedef int T;
#define CompLT(a,b) (a < b)
#define CompEQ(a,b) (a == b)

typedef struct Node_ {
    struct Node_ *Left;         /* left child */
    struct Node_ *Right;        /* right child */
    struct Node_ *Parent;       /* parent */
    T Data;                     /* data stored in node */
} Node;

Node *Root = NULL;               /* root of binary tree */

Node *InsertNode(T Data) {
Node *X, *Current, *Parent;

/******************************
 * allocate node for Data and insert in tree *
******************************/
/* setup new node */
if ((X = malloc (sizeof(*X))) == 0) {
    fprintf (stderr, "insufficient memory (InsertNode)\n");
    exit(1);
}
X->Data = Data;
X->Left = NULL;
X->Right = NULL;

/* find X's parent */
Current = Root;
Parent = 0;
while (Current) {
    if (CompEQ(X->Data, Current->Data)) return (Current);
    Parent = Current;
    Current = CompLT(X->Data, Current->Data)
    ?
        Current->Left : Current->Right;
}
X->Parent = Parent;

/* insert X in tree */
if(Parent)
    if(CompLT(X->Data, Parent->Data))
        Parent->Left = X;
    else
        Parent->Right = X;
else
    Root = X;
return(X);
}

void DeleteNode(Node *Z) {
    Node *X, *Y;

    /*****************************
    * delete node Z from tree *
    *****************************/

    /* Y will be removed from the parent chain */
    if (!Z || Z == NULL) return;

    /* find tree successor */
    if (Z->Left == NULL || Z->Right == NULL)
        Y = Z;
    else {
        Y = Z->Right;
        while (Y->Left != NULL) Y = Y->Left;
    }

    /* X is Y's only child */
    if (Y->Left != NULL)
        X = Y->Left;
    else
        X = Y->Right;

    /* remove Y from the parent chain */
    if (X) X->Parent = Y->Parent;
    if (Y->Parent)
        if (Y == Y->Parent->Left)
            Y->Parent->Left = X;
        else
            Y->Parent->Right = X;
    else
        Root = X;
else
    Y->Parent->Right = X;
else
    Root = X;
/* Y is the node we're removing */
/* Z is the data we're removing */
/* if Z and Y are not the same, replace Z with Y. */
if (Y != Z) {
    Y->Left = Z->Left;
    if (Y->Left) Y->Left->Parent = Y;
    Y->Right = Z->Right;
    if (Y->Right) Y->Right->Parent = Y;
    Y->Parent = Z->Parent;
    if (Z->Parent)
        if (Z == Z->Parent->Left)
            Z->Parent->Left = Y;
        else
            Z->Parent->Right = Y;
    else
        Root = Y;
    free (Z);
} else {
    free (Y);
}

Node *FindNode(T Data) {

    /**********************************************************
    * find node containing Data *
    **********************************************************/

    Node *Current = Root;
    while (Current != NULL)
        if (CompEQ(Data, Current->Data))
            return (Current);
        else
            Current = CompLT (Data, Current->Data) ?
                Current->Left : Current->Right;
    return(0);
}

a) Red-Black Tree Code

#include <stdlib.h>
#include <stdio.h>

/* modify these lines to establish data type */
typedef int T;
#define CompLT(a,b) (a < b)
#define CompEQ(a,b) (a == b)

/* red-black tree description */
typedef enum { Black, Red } NodeColor;
typedef struct Node_ {
    struct Node_ *Left; /* left child */
    struct Node_ *Right; /* right child */
    struct Node_ *Parent; /* parent */
    NodeColor Color; /* node color (black, red) */
}
T Data;                     /* data stored in node */
} Node;

#define NIL &Sentinel /* all leafs are sentinels */
Node Sentinel = { NIL, NIL, 0, Black, 0};

Node *Root = NIL;               /* root of red-black tree */

Node *InsertNode(T Data) {
    Node *Current, *Parent, *X;

    /***************************************************************/
    /* allocate node for Data and insert in tree */
    /*******************************************************************/

    /* setup new node */
    if ((X = malloc (sizeof(*X))) == 0) {
        printf("insufficient memory (InsertNode)\n");
        exit(1);
    }
    X->Data = Data;
    X->Left = NIL;
    X->Right = NIL;
    X->Parent = 0;
    X->Color = Red;

    /* find where node belongs */
    Current = Root;
    Parent = 0;
    while (Current != NIL) {
        if (CompEQ(X->Data, Current->Data)) return (Current);
        Parent = Current;
        Current = CompLT(X->Data, Current->Data) ?
            Current->Left : Current->Right;
    }

    /* insert node in tree */
    if(Parent) {
        if(CompLT(X->Data, Parent->Data))
            Parent->Left = X;
        else
            Parent->Right = X;
        X->Parent = Parent;
    } else
    Root = X;

    InsertFixup(X);
    return(X);
}

void InsertFixup(Node *X) {

    /***************************************************************
    * maintain red-black tree balance *
    * after inserting node X *
    ***************************************************************

    /* check red-black properties */
    while (X != Root && X->Parent->Color == Red) {
        /* we have a violation */

        /***************************************************************
        * check red-black properties *
        ***************************************************************
if (X->Parent == X->Parent->Parent->Left) {
    Node *Y = X->Parent->Parent->Right;
    if (Y->Color == Red) {
        /* uncle is red */
        X->Parent->Color = Black;
        Y->Color = Black;
        X->Parent->Parent->Color = Red;
        X = X->Parent->Parent;
    } else {
        /* uncle is black */
        if (X == X->Parent->Right) {
            /* make X a left child */
            X = X->Parent;
            RotateLeft(X);
        }
        /* recolor and rotate */
        X->Parent->Color = Black;
        X->Parent->Parent->Color = Red;
        RotateRight(X->Parent->Parent);
    }
} else {
    /* mirror image of above code */
    Node *Y = X->Parent->Parent->Left;
    if (Y->Color == Red) {
        /* uncle is red */
        X->Parent->Color = Black;
        Y->Color = Black;
        X->Parent->Parent->Color = Red;
        X = X->Parent->Parent;
    } else {
        /* uncle is black */
        if (X == X->Parent->Left) {
            /* make X a right child */
            X = X->Parent;
            RotateRight(X);
        }
        X->Parent->Color = Black;
        X->Parent->Parent->Color = Red;
        RotateLeft(X->Parent->Parent);
    }
}
Root->Color = Black;
}

void RotateLeft(Node *X) {
    /******************************************
    * rotate Node X to left *
    ******************************************/
    Node *Y = X->Right;
    /* establish X->Right link */
    X->Right = Y->Left;
    if (Y->Left != NIL) Y->Left->Parent = X;
/* establish Y->Parent link */
if (Y != NIL) Y->Parent = X->Parent;
if (X->Parent) {
    if (X == X->Parent->Left)
        X->Parent->Left = Y;
    else
        X->Parent->Right = Y;
} else {
    Root = Y;
}
/* link X and Y */
Y->Left = X;
if (X != NIL) X->Parent = Y;
}

void RotateRight(Node *X) {
    /*****************************************************************************
     * rotate Node X to right *
     *****************************************************************************
    Node *Y = X->Left;
    /* establish X->Left link */
    X->Left = Y->Right;
    if (Y->Right != NIL) Y->Right->Parent = X;
    /* establish Y->Parent link */
    if (Y != NIL) Y->Parent = X->Parent;
    if (X->Parent) {
        if (X == X->Parent->Right)
            X->Parent->Right = Y;
        else
            X->Parent->Left = Y;
    } else {
        Root = Y;
    }
    /* link X and Y */
    Y->Right = X;
    if (X != NIL) X->Parent = Y;
}

void DeleteNode(Node *Z) {
    Node *X, *Y;
    /*****************************************************************************
     * delete node Z from tree *
     *****************************************************************************
    if (!Z || Z == NIL) return;
    if (Z->Left == NIL || Z->Right == NIL) {
        /* Y has a NIL node as a child */
        Y = Z;
    } else {
        /* find tree successor with a NIL node as a child */
        Y = Z->Right;
        while (Y->Left != NIL) Y = Y->Left;
/* X is Y's only child */
if (Y->Left != NIL)
    X = Y->Left;
else
    X = Y->Right;

/* remove Y from the parent chain */
X->Parent = Y->Parent;
if (Y->Parent)
    if (Y == Y->Parent->Left)
        Y->Parent->Left = X;
    else
        Y->Parent->Right = X;
else
    Root = X;

if (Y != Z) Z->Data = Y->Data;
if (Y->Color == Black)
    DeleteFixup (X);
free (Y);

void DeleteFixup(Node *X) {

 /*****************************************************************************
 * maintain red-black tree balance *
 * after deleting node X *
 ******************************************************************************/

 while (X != Root && X->Color == Black) {
    if (X == X->Parent->Left) {
        Node *W = X->Parent->Right;
        if (W->Color == Red) {
            W->Color = Black;
            X->Parent->Color = Red;
            RotateLeft (X->Parent);
            W = X->Parent->Right;
        }
        if (W->Left->Color == Black && W->Right->Color == Black) {
            W->Color = Red;
            X = X->Parent;
        } else {
            if (W->Right->Color == Black) {
                W->Left->Color = Black;
                W->Color = Red;
                RotateRight (W);
                W = X->Parent->Right;
            }
            W->Color = X->Parent->Color;
            X->Parent->Color = Black;
            W->Right->Color = Black;
            RotateLeft (X->Parent);
            X = Root;
        }
    } else {
        Node *W = X->Parent->Left;
        if (W->Color == Red) {
            W->Color = Black;
            X->Parent->Color = Red;
        }
    }
}
```c
RotateRight (X->Parent);
W = X->Parent->Left;
}
if (W->Right->Color == Black && W->Left->Color == Black) {
    W->Color = Red;
    X = X->Parent;
} else {
    if (W->Left->Color == Black) {
        W->Right->Color = Black;
        W->Color = Red;
        RotateLeft (W);
        W = X->Parent->Left;
    }
    W->Color = X->Parent->Color;
    X->Parent->Color = Black;
    W->Left->Color = Black;
    RotateRight (X->Parent);
    X = Root;
    X->Color = Black;
} 

Node *FindNode(T Data) {
    /*******************************
    *  find node containing Data  *
    *******************************/
    Node *Current = Root;
    while(Current != NIL)
        if(CompEQ(Data, Current->Data))
            return (Current);
        else
            Current = CompLT (Data, Current->Data) ?
                Current->Left : Current->Right;
    return(0);
}
```

a) Skip List Code

```c
#include <stdio.h>
#include <stdlib.h>

/* define data-type and compare operators here */
typedef int T;
#define CompLT(a,b) (a < b)
#define CompEQ(a,b) (a == b)

/*
 * levels range from (0 .. MAXLEVEL)
 */
#define MAXLEVEL 15

typedef struct {
    T Data;                     /* user's data */
    struct Node_ *Forward[1];   /* skip list forward pointer */
} Node;
```
Node *Hdr;                     /* List Header */
int ListLevel;                 /* current level of list */
} SkipList;

SkipList List;                 /* skip list information */

#define NIL List.Hdr

void InitList() {
    int i;

    /**************************
    * initialize skip list *
    **************************/

    if ((List.Hdr = malloc(sizeof(Node) + MAXLEVEL*sizeof(Node *))) == 0) {
        printf ("insufficient memory (InitList)\n");
        exit(1);
    }
    for (i = 0; i <= MAXLEVEL; i++)
        List.Hdr->Forward[i] = NIL;
    List.ListLevel = 0;
}

Node *InsertNode(T Data) {
    int i, NewLevel;
    Node *update[MAXLEVEL+1];
    Node *X;

    /*********************************************************************************
    * allocate node for Data and insert in list *
    *********************************************************************************/

    /* find where data belongs */
    X = List.Hdr;
    for (i = List.ListLevel; i >= 0; i--) {
        while (X->Forward[i] != NIL
            && CompLT(X->Forward[i]->Data, Data))
            X = X->Forward[i];
        update[i] = X;
    }
    X = X->Forward[0];
    if (X != NIL && CompEQ(X->Data, Data)) return(X);

    /* determine level */
    NewLevel = 0;
    while (rand() < RAND_MAX/2) NewLevel++;
    if (NewLevel > MAXLEVEL) NewLevel = MAXLEVEL;
    if (NewLevel > List.ListLevel) {
        for (i = List.ListLevel + 1; i <= NewLevel; i++)
            update[i] = NIL;
        List.ListLevel = NewLevel;
    }

    /* make new node */
    if (((X = malloc(sizeof(Node) +
        NewLevel*sizeof(Node *))) == 0) {
        printf ("insufficient memory (InsertNode)\n");
        exit(1);
    }
X->Data = Data;

    /* update forward links */
    for (i = 0; i <= NewLevel; i++) {
        X->Forward[i] = update[i]->Forward[i];
        update[i]->Forward[i] = X;
    }
    return(X);
}

void DeleteNode(T Data) {
    int i;
    Node *update[MAXLEVEL+1], *X;

    /******************************************
    * delete node containing Data from list *
    ******************************************/

    /* find where data belongs */
    X = List.Hdr;
    for (i = List.ListLevel; i >= 0; i--) {
        while (X->Forward[i] != NIL
             && CompLT(X->Forward[i]->Data, Data))
            X = X->Forward[i];
        update[i] = X;
    }
    X = X->Forward[0];
    if (X == NIL || !CompEQ(X->Data, Data)) return;

    /* adjust forward pointers */
    for (i = 0; i <= List.ListLevel; i++) {
        if (update[i]->Forward[i] != X) break;
        update[i]->Forward[i] = X->Forward[i];
    }
    free (X);

    /* adjust header level */
    while ((List.ListLevel > 0)
             && (List.Hdr->Forward[List.ListLevel] == NIL))
        List.ListLevel--;
}

Node *FindNode(T Data) {
    int i;
    Node *X = List.Hdr;

    /**************************************************
    * find node containing Data *
    **************************************************/

    for (i = List.ListLevel; i >= 0; i--) {
        while (X->Forward[i] != NIL
             && CompLT(X->Forward[i]->Data, Data))
            X = X->Forward[i];
    }
    X = X->Forward[0];
    if (X != NIL && CompEQ(X->Data, Data)) return (X);
    return(0);
}

1. Bibliography


PPNT Arial * 4 ; d PPNT " System ; d
y
Let }"
and stop w }=
hen with }? when } d PPNT Century Schoolbook (
4 + \text{ Symbol}
\text{ Century Schoolbook}
1 + 2 * 3 * 4 * 5
PPNT " Arial; d PPNT " System; d
PPNT " Arial " System
§ 2 \red{d} PPNT " Arial \)§ 2 \red{i} d PPNT " System \( t - Y, k | k \), \( k | Y | Y, k \), \( p - Y, k | k, k | Y | Y, k \), \( i d \)
( #
0 ; d PPNT " System ; d
PPNT
Arial
(System p A Ô W é A é W ò p A ã W W A Ÿ Ÿ P Æ 3 P Æ 3 B Æ 3 2 Æ 3 Ÿ)
???
3  ∈ – ??
∈ 3
 t ~  ! !
Å 0 Å - Å
Schoolbook

Schoolbook

Schoolbook

Schoolbook
[ ] a d h l p s v x
The text appears to be a code snippet or some form of programming language, possibly related to font or typographic settings. Without more context, it's challenging to provide a clear interpretation.
\[\delta \delta \ p - \delta\]
<table>
<thead>
<tr>
<th>Arial</th>
<th>d</th>
<th>PPNT</th>
<th>&quot;</th>
<th>System</th>
<th>d</th>
<th>PPNT</th>
<th>&quot;</th>
<th>System</th>
<th>p</th>
<th>5</th>
<th>5</th>
<th>Y</th>
<th>t</th>
<th>¾</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>6</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>t N</td>
<td>X</td>
<td>}</td>
<td>}</td>
<td>}</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>]</td>
<td>}</td>
<td>i d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
f \& f
f \delta \delta \quad p - f \delta
f
f d ṭ ṭ ṭ ṭ ṭ ṭ t - u ' f B f ' f B u B u ' f ' p - u ' f B f ' f B u B u ' f ' ī d
B²TP
0 0 n ε @ ε
B ô Vũ 0
ñ. ñia ñ{ ñ} ñ~ ñ—
ñè ñ° ñ» ñï ñ ñ( ñ) ññ ñ¿ ñ² ñè ñî ñï ñæ ñ ñ, ñ. ñ> ñ ñ! ñ§
ñ ñ) ñ ñ ñ± ñ
Body Text caption
SubHeading  Signature
endnote text

PseudoCode

InlineCode

List Bullet 4

List Bullet

List Number List List
<table>
<thead>
<tr>
<th>wyß</th>
<th>.</th>
<th>xýý</th>
<th>0</th>
<th>xýý</th>
<th>0ì</th>
<th>zýý</th>
<th>0à</th>
<th>[ýý</th>
<th>0ö</th>
</tr>
</thead>
<tbody>
<tr>
<td>\ýý</td>
<td>1</td>
<td>)ýý</td>
<td>1</td>
<td>_ýý</td>
<td>1D</td>
<td>^ýý</td>
<td>10</td>
<td>´ýý</td>
<td>B</td>
</tr>
<tr>
<td>aýý</td>
<td>Qx</td>
<td>býý</td>
<td>f</td>
<td>cýý</td>
<td>tl</td>
<td>dýý</td>
<td>w8</td>
<td>eýý</td>
<td>wT</td>
</tr>
<tr>
<td>fýý</td>
<td>wh</td>
<td>gýý</td>
<td>w</td>
<td>hýý</td>
<td>w</td>
<td>iýý</td>
<td>wM</td>
<td>kýý</td>
<td>wI</td>
</tr>
<tr>
<td>jýý</td>
<td>w.</td>
<td>lýý</td>
<td>&lt;</td>
<td>mýý</td>
<td>·X</td>
<td>nýý</td>
<td>½\</td>
<td>oýý</td>
<td>½p</td>
</tr>
<tr>
<td>pýý</td>
<td>¼\</td>
<td>qýý</td>
<td>¼-</td>
<td>rýý</td>
<td>¼-</td>
<td>sýý</td>
<td>¼A</td>
<td>uýý</td>
<td>¼è</td>
</tr>
<tr>
<td>týý</td>
<td>¼O</td>
<td>výý</td>
<td>!d</td>
<td>wýý</td>
<td>(</td>
<td>xýý</td>
<td>(§</td>
<td>yýý</td>
<td>(8</td>
</tr>
<tr>
<td>zýý</td>
<td>(L</td>
<td>(ýý</td>
<td>(‘</td>
<td>ŭý</td>
<td>(t</td>
<td>~ýý</td>
<td>(œ</td>
<td>)ýý</td>
<td>(’</td>
</tr>
<tr>
<td>ýý</td>
<td>6ì</td>
<td>cýý</td>
<td>&lt;</td>
<td>ýý</td>
<td>F</td>
<td>,ýý</td>
<td>F1</td>
<td>sýý</td>
<td>F€</td>
</tr>
<tr>
<td>&quot;ýý</td>
<td>F“</td>
<td>.ýý</td>
<td>F“</td>
<td>týý</td>
<td>F4</td>
<td>&quot;ýý</td>
<td>Fä</td>
<td>týý</td>
<td>Fö</td>
</tr>
<tr>
<td>hýý</td>
<td>G1</td>
<td>Šýý</td>
<td>G€</td>
<td>xýý</td>
<td>G“</td>
<td>Gëýý</td>
<td>G“</td>
<td>ýý</td>
<td>GÖ</td>
</tr>
<tr>
<td>Žýý</td>
<td>GD</td>
<td>ýý</td>
<td>Gä</td>
<td>‘ýý</td>
<td>H</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>