Recursion

TOPICS

19.1	Introduction to Recursion	19.6	Focus on Problem Solving and		
19.2	Solving Problems with Recursion		Program Design: A Recursive		
19.3	Focus on Problem Solving and		Binary Search Function		
	Program Design: The Recursive	19.7	The Towers of Hanoi		
	gcd Function	19.8	Focus on Problem Solving		
19.4	Focus on Problem Solving and		and Program Design:		
	Program Design: Solving Recursively		The QuickSort Algorithm		
	Defined Problems	19.9	Exhaustive Algorithms		
19.5	Focus on Problem Solving and	19.10	Focus on Software Engineering:		
	Program Design: Recursive Linked		Recursion vs. Iteration		
	List Operations				

19.1 Introduction to Recursion

CONCEPT: A recursive function is one that calls itself.

You have seen instances of functions calling other functions. Function A can call function B, which can then call function C. It's also possible for a function to call itself. A function that calls itself is a *recursive function*. Look at this message function:

```
void message()
{
    cout << "This is a recursive function.\n";
    message();
}</pre>
```

This function displays the string "This is a recursive function.\n", and then calls itself. Each time it calls itself, the cycle is repeated. Can you see a problem with the function? There's no way to stop the recursive calls. This function is like an infinite loop because there is no code to stop it from repeating.



NOTE: The function example message will eventually cause the program to crash. Do you remember learning in Chapter 18 that the system stores temporary data on a stack each time a function is called? Eventually, these recursive function calls will use up all available stack memory and cause it to overflow.

Like a loop, a recursive function must have some method to control the number of times it repeats. The following is a modification of the message function. It passes an integer argument, that holds the number of times the function is to call itself.

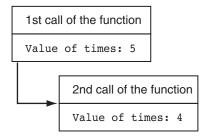
```
void message(int times)
{
    if (times > 0)
    {
       cout << "This is a recursive function.\n";
       message(times - 1);
    }
}</pre>
```

This function contains an if statement that controls the repetition. As long as the times argument is greater than zero, it will display the message and call itself again. Each time it calls itself, it passes times - 1 as the argument. For example, let's say a program calls the function with the following statement:

```
message(5);
```

The argument, 5, will cause the function to call itself five times. The first time the function is called, the if statement will display the message and then call itself with 4 as the argument. Figure 19-1 illustrates this:

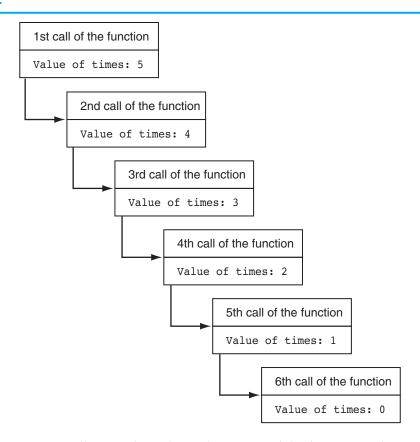
Figure 19-1



The diagram in Figure 19-1 illustrates two separate calls of the message function. Each time the function is called, a new instance of the times parameter is created in memory. The first time the function is called, the times parameter is set to 5. When the function calls itself, a new instance of times is created, and the value 4 is passed into it. This cycle repeats until, finally, zero is passed to the function. This is illustrated in Figure 19-2.

As you can see from Figure 19-2 the function is called a total of six times. The first time it is called from main, and the other five times it calls itself, so the *depth of recursion* is five. When the function reaches its sixth call, the times parameter will be set to 0. At that point, the if statement's conditional expression will be false, so the function will return. Control

Figure 19-2



of the program will return from the sixth instance of the function to the point in the fifth instance directly after the recursive function call:

Because there are no more statements to be executed after the function call, the fifth instance of the function returns control of the program back to the fourth instance. This repeats until all instances of the function return. Program 19-1 demonstrates the recursive message function.

```
// This program demonstrates a simple recursive function.
include <iostream>
using namespace std;
```

```
Program 19-1
               (continued)
   // Function prototype
   void message(int);
 7
8
   int main()
9
   {
10
       message(5);
11
      return 0;
12
  }
13
14 //******************************
15 // Definition of function Message. If the value in times is *
16 // greater than 0, the message is displayed and the
   // function is recursively called with the argument
18
   //*******************
19
20
21 void message(int times)
22
23
       if (times > 0)
24
       {
25
           cout << "This is a recursive function.\n";</pre>
26
          message(times - 1);
27
       }
```

Program Output

28 }

```
This is a recursive function.
```

To further illustrate the inner workings of this recursive function, let's look at another version of the program. In Program 19-2, a message is displayed each time the function is entered, and another message is displayed just before the function returns.

```
// This program demonstrates a simple recursive function.
    #include <iostream>
    using namespace std;
 5 // Function prototype
 6
   void message(int);
 7
8
   int main()
9
   {
10
        message(5);
11
       return 0;
12
   }
13
```

```
//*********************
   // Definition of function message. If the value in times is *
   // greater than 0, the message is displayed and the function *
   // is recursively called with the argument times - 1.
17
   //*******************
18
19
20
   void message(int times)
21
22
       cout << "message called with " << times << " in times.\n";</pre>
23
24
       if (times > 0)
25
26
           cout << "This is a recursive function.\n";</pre>
           message(times - 1);
2.8
       }
29
30
       cout << "message returning with " << times;</pre>
31
       cout << " in times.\n";</pre>
32 }
```

Program Output

```
message called with 5 in times.
This is a recursive function.
message called with 4 in times.
This is a recursive function.
message called with 3 in times.
This is a recursive function.
message called with 2 in times.
This is a recursive function.
message called with 1 in times.
This is a recursive function.
message called with 0 in times.
message returning with 0 in times.
message returning with 1 in times.
message returning with 2 in times.
message returning with 3 in times.
message returning with 4 in times.
message returning with 5 in times.
```

19.2 Solving Problems with Recursion

CONCEPT: A problem can be solved with recursion if it can be broken down into successive smaller problems that are identical to the overall problem.

Programs 19-1 and 19-2 in the previous section show simple demonstrations of *how* a recursive function works. But these examples don't show us *why* we would want to write a recursive function. Recursion can be a powerful tool for solving repetitive problems and is an important topic in upper-level computer science courses. What might not be clear to you yet is how to use recursion to solve a problem.

First, it should be noted that recursion is never absolutely required to solve a problem. Any problem that can be solved recursively can also be solved iteratively, with a loop. In fact, recursive algorithms are usually less efficient than iterative algorithms. This is because a function call requires several actions to be performed by the C++ runtime system. These actions include allocating memory for parameters and local variables and storing the address of the program location where control returns after the function terminates. These actions, which are sometimes referred to as *overhead*, take place with each function call. Such overhead is not necessary with a loop.

Some repetitive problems, however, are more easily solved with recursion than with iteration. Where an iterative algorithm might result in faster execution time, the programmer might be able to design a recursive algorithm faster.

In general, a recursive function works like this:

- If the problem can be solved now, without recursion, then the function solves it and returns.
- If the problem cannot be solved now, then the function reduces it to a smaller but similar problem and calls itself to solve the smaller problem.

In order to apply this approach, we first identify at least one case in which the problem can be solved without recursion. This is known as the *base case*. Second, we determine a way to solve the problem in all other circumstances using recursion. This is called the *recursive case*. In the recursive case, we must always reduce the problem to a smaller version of the original problem. By reducing the problem with each recursive call, the base case will eventually be reached and the recursion will stop.

Example: Using Recursion to Calculate the Factorial of a Number

Let's take an example from mathematics to examine an application of recursion. In mathematics, the notation n! represents the factorial of the number n. The factorial of a nonnegative number can be defined by the following rules:

```
If n = 0 then n! = 1
If n > 0 then n! = 1 \times 2 \times 3 \times ... \times n
```

Let's replace the notation n! with factorial(n), which looks a bit more like computer code, and rewrite these rules as

```
If n = 0 then factorial(n) = 1
If n > 0 then factorial(n) = 1 \times 2 \times 3 \times ... \times n
```

These rules state that when n is 0, its factorial is 1. When n is greater than 0, its factorial is the product of all the positive integers from 1 up to n. For instance, factorial(6) is calculated as $1 \times 2 \times 3 \times 4 \times 5 \times 6$.

When designing a recursive algorithm to calculate the factorial of any number, we first identify the base case, which is the part of the calculation that we can solve without recursion. That is the case where n is equal to 0:

```
If n = 0 then factorial(n) = 1
```



This tells how to solve the problem when n is equal to 0, but what do we do when n is greater than 0? That is the recursive case, or the part of the problem that we use recursion to solve. This is how we express it:

```
If n > 0 then factorial(n) = n \times factorial(n-1)
```

This states that if n is greater than 0, the factorial of n is n times the factorial of n-1. Notice how the recursive call works on a reduced version of the problem, n-1. So, our recursive rule for calculating the factorial of a number might look like this:

```
If n = 0 then factorial(n) = 1
If n > 0 then factorial(n) = n \times factorial(n - 1)
```

The following pseudocode shows how we might implement the factorial algorithm as a recursive function:

```
factorial(n)
    If n is 0 then
        return 1.
    else
        return n times the factorial of n - 1.
end factorial.
```

Here is the C++ code for such a function:

Program 19-3 demonstrates the recursive factorial function.

Program 19-3

```
1  // This program demonstrates a recursive function to
2  // calculate the factorial of a number.
3  #include <iostream>
4  using namespace std;
5  // Function prototype
7  int factorial(int);
8  
9  int main()
10  {
11   int number;
12
```

(program continues)

Program 19-3 (continued)

```
13
        // Get a number from the user.
14
        cout << "Enter an integer value and I will display\n";</pre>
15
        cout << "its factorial: ";</pre>
16
        cin >> number;
17
18
        // Display the factorial of the number.
19
        cout << "The factorial of " << number << " is ";</pre>
20
       cout << factorial(number) << endl;</pre>
21
        return 0;
22
   }
23
   //********************
24
25
   // Definition of factorial. A recursive function to calculate *
26
   // the factorial of the parameter n.
    //****************
27
28
29
   int factorial(int n)
31
        if (n == 0)
32
                                       // Base case
            return 1;
33
        else
34
           return n * factorial(n - 1); // Recursive case
35
   }
```

Program Output with Example Input Shown in Bold

```
Enter an integer value and I will display its factorial: 4 [Enter]
The factorial of 4 is 24
```

In the example run of the program, the factorial function is called with the argument 4 passed into n. Because n is not equal to 0, the if statement's else clause executes the following statement, in line 34:

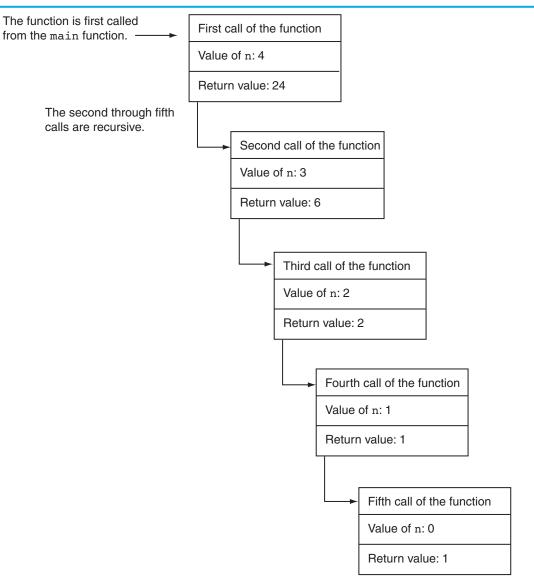
```
return n * factorial(n - 1);
```

Although this is a return statement, it does not immediately return. Before the return value can be determined, the value of factorial (num - 1) must be determined. The factorial function is called recursively until the fifth call, in which the n parameter will be set to zero. The diagram in Figure 19-3 illustrates the value of n and the return value during each call of the function.

This diagram illustrates why a recursive algorithm must reduce the problem with each recursive call. Eventually the recursion has to stop in order for a solution to be reached. If each recursive call works on a smaller version of the problem, then the recursive calls work toward the base case. The base case does not require recursion, so it stops the chain of recursive calls.

Usually, a problem is reduced by making the value of one or more parameters smaller with each recursive call. In our factorial function, the value of the parameter n gets closer to 0 with each recursive call. When the parameter reaches 0, the function returns a value without making another recursive call.

Figure 19-3



Example: Using Recursion to Count Characters

Let's look at another simple example of recursion. The following function counts the number of times a specific character appears in a string. The line numbers are from Program 19-4, which we will examine momentarily.

```
36
        else if (str[subscript] == search)
37
38
             // Recursive case: A matching character was found.
             // Return 1 plus the number of times the search
39
40
             // character appears in the rest of the string.
41
             return 1 + numChars(search, str, subscript+1);
42
        }
43
        else
        {
45
             // Recursive case: A character that does not match the
             // search character was found. Return the number of times
46
             // the search character appears in the rest of the string.
47
48
             return numChars(search, str, subscript+1);
49
        }
50
    }
```

The function's parameters are

- search: The character to be searched for and counted
- str: a string object containing the string to be searched
- subscript: The starting subscript for the search

When this function examines a character in the string, three possibilities exist:

- The end of the string has been reached. This is the base case because there are no more characters to search.
- A character that matches the search character is found. This is a recursive case because we still have to search the rest of the string.
- A character that does not match the search character is found. This is also a recursive case because we still have to search the rest of the string.

Let's take a closer look at the code. The first if statement, in line 31, determines whether the end of the string has been reached:

```
if (subscript >= str.length())
```

Reaching the end of the string is the base case of the problem. If the end of the string has been reached, the function returns 0, indicating that 0 matching characters were found. Otherwise, the following else if clause, in lines 36 through 42, is executed:

```
else if (str[subscript] == search)
{
    // Recursive case: A matching character was found.
    // Return 1 plus the number of times the search
    // character appears in the rest of the string.
    return 1 + numChars(search, str, subscript+1);
}
```

If str[subscript] contains the search character, then we have found one matching character. But because we have not reached the end of the string, we must continue to search the rest of the string for more matching characters. So, at this point the function performs a recursive call. The return statement returns 1 plus the number of times the search character appears in the string, starting at subscript+1. In essence, this statement returns 1 plus the number of times the search character appears in the rest of the string.

Finally, if str[subscript] does not contain the search character, the following else clause in lines 43 through 49 is executed:

```
else
{
    // Recursive case: A character that does not match the
    // search character was found. Return the number of times
    // the search character appears in the rest of the string.
    return numChars(search, str, subscript+1);
}
```

The return statement in line 48 makes a recursive call to search the remainder of the string. In essence, this code returns the number of times the search character appears in the rest of the string. Program 19-4 demonstrates the function.

```
// This program demonstrates a recursive function for counting
   // the number of times a character appears in a string.
   #include <iostream>
   #include <string>
 4
   using namespace std;
 7
   // Function prototype
 8
   int numChars(char, string, int);
 9
10
   int main()
11
12
        string str = "abcddddef";
13
14
        // Display the number of times the letter
15
        // 'd' appears in the string.
        cout << "The letter d appears "</pre>
            << numChars('d', str, 0) << " times.\n";
17
18
19
        return 0;
20
   }
21
   //*************
22
   // Function numChars. This recursive function
   // counts the number of times the character
   // search appears in the string str. The search *
    // begins at the subscript stored in subscript. *
    //*************
27
28
29
   int numChars(char search, string str, int subscript)
30
31
        if (subscript >= str.length())
32
33
            // Base case: The end of the string is reached.
34
            return 0;
35
        }
36
        else if (str[subscript] == search)
                                                           (program continues)
```

Program 19-4 (continued)

```
37
        {
38
             // Recursive case: A matching character was found.
39
             // Return 1 plus the number of times the search
40
             // character appears in the rest of the string.
41
             return 1 + numChars(search, str, subscript+1);
42
        }
43
        else
44
        {
45
             // Recursive case: A character that does not match the
46
             // search character was found. Return the number of times
47
             // the search character appears in the rest of the string.
48
             return numChars(search, str, subscript+1);
49
        }
50
```

Program Output

The letter d appears 4 times.

Direct and Indirect Recursion

The examples we have discussed so far show recursive functions that directly call themselves. This is known as *direct recursion*. There is also the possibility of creating *indirect recursion* in a program. This occurs when function A calls function B, which in turn calls function A. There can even be several functions involved in the recursion. For example, function A could call function B, which could call function C, which calls function A.



Checkpoint

- 19.1 What happens if a recursive function never returns?
- 19.2 What is a recursive function's base case?
- 19.3 What will the following program display?

```
#include <iostream>
using namespace std;
// Function prototype
void showMe(int arg);
int main()
{
    int num = 0;
    showMe(num);
    return 0;
}
void showMe(int arg)
{
    if (arg < 10)
         showMe(++arg);
    else
         cout << arg << endl;
```

19.4 What is the difference between direct and indirect recursion?



19.3 Focus on Problem Solving and Program Design: The Recursive gcd Function

CONCEPT: The gcd function uses recursion to find the greatest common divisor (GCD) of two numbers.

Our next example of recursion is the calculation of the greatest common divisor, or GCD, of two numbers. Using Euclid's algorithm, the GCD of two positive integers, x and y, is:

```
if y divides x evenly
gcd(x, y) = y;
gcd(y, remainder of x/y);
                               otherwise
```

The definition above states that the GCD of x and y is y if x/y has no remainder. Otherwise, the answer is the GCD of y and the remainder of x/y. Program 19-5 shows the recursive C++ implementation:

Program 19-5

```
// This program demonstrates a recursive function to calculate
   // the greatest common divisor (gcd) of two numbers.
   #include <iostream>
 4 using namespace std;
   // Function prototype
 7
   int gcd(int, int);
 8
 9
   int main()
10
11
       int num1, num2;
12
       // Get two numbers.
14
      cout << "Enter two integers: ";</pre>
15
      cin >> num1 >> num2;
16
17
      // Display the GCD of the numbers.
       cout << "The greatest common divisor of " << num1;</pre>
       cout << " and " << num2 << " is ";
19
       cout << gcd(num1, num2) << endl;</pre>
20
21
       return 0;
22
   }
23
   //****************
24
  // Definition of gcd. This function uses recursion to
   // calculate the greatest common divisor of two integers, *
   // passed into the parameters x and y.
   //********************
28
29
   int gcd(int x, int y)
```

(program continues)

Program 19-5

(continued)

```
31
32
        if (x % y == 0)
33
            return y;
                                   // Base case
34
        else
            return gcd(y, x % y); // Recursive case
35
36 }
```

Program Output with Example Input Shown in Bold

```
Enter two integers: 49 28 [Enter]
The greatest common divisor of 49 and 28 is 7
```

19.4 Focus on Problem Solving and Program Design: **Solving Recursively Defined Problems**

CONCEPT: Some mathematical problems are designed for a recursive solution.

Some mathematical problems are designed to be solved recursively. One well-known example is the calculation of *Fibonacci numbers*. The Fibonacci numbers, named after the Italian mathematician Leonardo Fibonacci (born circa 1170), are the following sequence:

```
0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, \dots
```

Notice that after the second number, each number in the series is the sum of the two previous numbers. The Fibonacci series can be defined as

```
F_0 = 0
F_1 = 1
F_N = F_{N-1} + F_{N-2} for N \ge 2.
```

A recursive C++ function to calculate the *n*th number in the Fibonacci series is shown here:

```
int fib(int n)
    if (n \le 0)
        return 0;
    else if (n == 1)
         return 1;
    else
         return fib(n - 1) + fib(n - 2);
}
```

The function is demonstrated in Program 19-6, which displays the first 10 numbers in the Fibonacci series.

```
// This program demonstrates a recursive function
// that calculates Fibonacci numbers.
#include <iostream>
using namespace std;
```

```
// Function prototype
 7
   int fib(int);
 8
9
   int main()
10
11
        cout << "The first 10 Fibonacci numbers are:\n";</pre>
12
        for (int x = 0; x < 10; x++)
           cout << fib(x) << " ";
13
14
      cout << endl;
15
       return 0;
   }
16
17
18
   // Function fib. Accepts an int argument *
   // in n. This function returns the nth
   // Fibonacci number.
22
   //************
23
   int fib(int n)
24
25
26
        if (n \ll 0)
                                            // Base case
27
           return 0;
        else if (n == 1)
28
29
            return 1;
                                             // Base case
30
        else
           return fib(n - 1) + fib(n - 2); // Recursive case
31
32 }
```

Program Output

```
The first 10 Fibonacci numbers are: 0 1 1 2 3 5 8 13 21 34
```

Another such example is Ackermann's function. A Programming Challenge at the end of this chapter asks you to write a recursive function that calculates Ackermann's function.

19.5 Focus on Problem Solving and Program Design:
Recursive Linked List Operations

CONCEPT: Recursion can be used to traverse the nodes in a linked list.

Recall that in Chapter 17 we discussed a class named NumberList that holds a linked list of double values. In this section we will modify the class by adding recursive member functions. The functions will use recursion to traverse the linked list and perform the following operations:

• Count the number of nodes in the list.

To count the number of nodes in the list by recursion, we introduce two new member functions: numNodes and countNodes. countNodes is a private member function that uses recursion, and numNodes is the public interface that calls it.

• Display the value of the list nodes in reverse order.

To display the nodes in the list in reverse order, we introduce two new member functions: displayBackwards and showReverse. showReverse is a private member function that uses recursion, and displayBackwards is the public interface that calls it.

The class declaration, which is saved in NumberList.h, is shown here:

```
1 // Specification file for the NumberList class
 2 #ifndef NUMBERLIST H
 3 #define NUMBERLIST H
 5 class NumberList
 6 {
 7 private:
       // Declare a structure for the list
 9
       struct ListNode
10
11
            double value;
12
            struct ListNode *next;
13
       };
14
15
       ListNode *head;
                            // List head pointer
16
       // Private member functions
18
       int countNodes(ListNode *) const;
19
       void showReverse(ListNode *) const;
20
21 public:
22
       // Constructor
23
       NumberList()
24
            { head = nullptr; }
25
       // Destructor
26
       ~NumberList();
27
2.8
29
       // Linked List Operations
3.0
       void appendNode(double);
31
       void insertNode(double);
32
       void deleteNode(double);
33
       void displayList() const;
34
       int numNodes() const
35
            { return countNodes(head); }
36
       void displayBackwards() const
37
            { showReverse(head); }
38 };
39 #endif
```

Counting the Nodes in the List

The numNodes function is declared inline. It simply calls the countNodes function and passes the head pointer as an argument. (Because the head pointer, which is private, must be passed to countNodes, the numNodes function is needed as an interface.)

The function definition for countNodes is shown here:

```
173 int NumberList::countNodes(ListNode *nodePtr) const
174 {
175          if (nodePtr != nullptr)
176              return 1 + countNodes(nodePtr->next);
177          else
178              return 0;
179 }
```

The function's recursive logic can be expressed as:

```
If the current node has a value
Return 1 + the number of the remaining nodes.

Else
Return 0.

End If.
```

Program 19-7 demonstrates the function.

Program 19-7

```
// This program counts the nodes in a list.
    #include <iostream>
    #include "NumberList.h"
   using namespace std;
 4
 6
   int main()
 7
 8
        const int MAX = 10; // Maximum number of values
 9
        // Define a NumberList object.
10
11
        NumberList list;
12
        // Build the list with a series of numbers.
13
14
        for (int x = 0; x < MAX; x++)
15
             list.insertNode(x);
16
17
        // Display the number of nodes in the list.
        cout << "The number of nodes is "
19
              << list.numNodes() << endl;
20
        return 0;
21
    }
```

Program Output

The number of nodes is 10

Displaying List Nodes in Reverse Order

The technique for displaying the list nodes in reverse order is designed like the node counting procedure: A public member function, which serves as an interface, passes the head pointer to a private member function. The public displayBackwards function, declared

inline, is the interface. It calls the showReverse function and passes the head pointer as an argument. The function definition for showReverse is shown here:

The base case for the function is nodePtr being set to nullptr. When this is true, the function has reached the last node in the list, so it returns. It is not until this happens that any instances of the cout statement execute. The instance of the function whose nodePtr variable points to the last node in the list will be the first to execute the cout statement. It will then return, and the previous instance of the function will execute its cout statement. This repeats until all the instances of the function have returned.

The modified class declaration is stored in NumberList.h, and its member function implementation is in NumberList.cpp. The remainder of the class implementation is unchanged from Chapter 17, so it is not shown here. Program 19-8 demonstrates the function.

```
// This program demonstrates the recursive function
    // for displaying the list's nodes in reverse.
    #include <iostream>
    #include "NumberList.h"
    using namespace std;
 6
 7
    int main()
 8
 9
         const double MAX = 10.0; // Upper limit of values
10
         // Create a NumberList object.
11
12
        NumberList list;
13
14
         // Add a series of numbers to the list.
15
         for (double x = 1.5; x < MAX; x += 1.1)
16
             list.appendNode(x);
17
18
         // Display the values in the list.
19
         cout << "Here are the values in the list:\n";</pre>
20
         list.displayList();
21
22
         // Display the values in reverse order.
2.3
         cout << "Here are the values in reverse order:\n";</pre>
24
        list.displayBackwards();
25
        return 0;
26 }
```

Program Output Here are the values in the list: 1.5 2.6 3.7 4.8 5.9 7 8.1 9.2 Here are the values in reverse order: 9.2 8.1 7 5.9 4.8 3.7 2.6 1.5

19.6 Focus on Problem Solving and Program Design: **A Recursive Binary Search Function**

CONCEPT: The binary search algorithm can be defined as a recursive function.

In Chapter 8 you learned about the binary search algorithm and saw an iterative example written in C++. The binary search algorithm can also be implemented recursively. For example, the procedure can be expressed as

```
If array[middle] equals the search value, then the value is found.
```

Else, if array[middle] is less than the search value, perform a binary search on the upper half of the array.

Else, if array[middle] is greater than the search value, perform a binary search on the lower half of the array.

The recursive binary search algorithm is an example of breaking a problem down into smaller pieces until it is solved. A recursive binary search function is shown here:

```
int binarySearch(int array[], int first, int last, int value)
{
   int middle;
                  // Midpoint of search
   if (first > last)
       return -1;
   middle = (first + last) / 2;
   if (array[middle] == value)
       return middle;
   if (array[middle] < value)</pre>
       return binarySearch(array, middle+1,last,value);
   else
       return binarySearch(array, first,middle-1,value);
```

The first parameter, array, is the array to be searched. The next parameter, first, holds the subscript of the first element in the search range (the portion of the array to be searched). The next parameter, last, holds the subscript of the last element in the search range. The last parameter, value, holds the value to be searched for. Like the iterative version, this function returns the subscript of the value if it is found, or -1 if the value is not found. Program 19-9 demonstrates the function.

```
// This program demonstrates the recursive binarySearch function.
    #include <iostream>
    using namespace std;
 5
   // Function prototype
    int binarySearch(int [], int, int, int);
 7
 8
    const int SIZE = 20; // Array size
 9
10
   int main()
11
12
        // Define an array of employee ID numbers
13
        int tests[SIZE] = {101, 142, 147, 189, 199, 207, 222,
14
                           234, 289, 296, 310, 319, 388, 394,
15
                           417, 429, 447, 521, 536, 600};
16
                      // To hold an ID number
        int empID;
17
        int results; // To hold the search results
18
19
        // Get an employee ID number to search for.
20
        cout << "Enter the Employee ID you wish to search for: ";
21
        cin >> empID;
22
23
        // Search for the ID number in the array.
24
        results = binarySearch(tests, 0, SIZE - 1, empID);
25
26
        // Display the results of the search.
27
        if (results == -1)
            cout << "That number does not exist in the array.\n";</pre>
28
29
        else
30
31
            cout << "That ID is found at element " << results;</pre>
32
            cout << " in the array\n";</pre>
33
34
        return 0;
35
    }
36
    //********************
37
   // The binarySearch function performs a recursive binary search *
38
   // on a range of elements of an integer array passed into the
39
   // parameter array. The parameter first holds the subscript of
40
   // the range's starting element, and last holds the subscript
42
   // of the range's last element. The parameter value holds the
   // search value. If the search value is found, its array
43
   // subscript is returned. Otherwise, -1 is returned indicating
    // the value was not in the array.
45
    //*********************
46
47
48
    int binarySearch(int array[], int first, int last, int value)
49
50
        int middle; // Midpoint of search
51
52
        if (first > last)
53
            return -1;
54
        middle = (first + last)/2;
```

```
if (array[middle]==value)
    return middle;
if (array[middle]<value)
    return binarySearch(array, middle+1,last,value);
else
    return binarySearch(array, first,middle-1,value);
}</pre>
```

Program Output with Example Input Shown in Bold

```
Enter the Employee ID you wish to search for: 521 [Enter] That ID is found at element 17 in the array
```

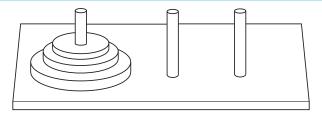
19.7

The Towers of Hanoi

CONCEPT: The repetitive steps involved in solving the Towers of Hanoi game can be easily implemented in a recursive algorithm.

The Towers of Hanoi is a mathematical game that is often used in computer science textbooks to illustrate the power of recursion. The game uses three pegs and a set of discs with holes through their centers. The discs are stacked on one of the pegs as shown in Figure 19-4.

Figure 19-4 The pegs and discs in the Towers of Hanoi game



Notice that the discs are stacked on the leftmost peg, in order of size with the largest disc at the bottom. The game is based on a legend in which a group of monks in a temple in Hanoi have a similar set of pegs with 64 discs. The job of the monks is to move the discs from the first peg to the third peg. The middle peg can be used as a temporary holder. Furthermore, the monks must follow these rules while moving the discs:

- Only one disc may be moved at a time.
- A disc cannot be placed on top of a smaller disc.
- All discs must be stored on a peg except while being moved.

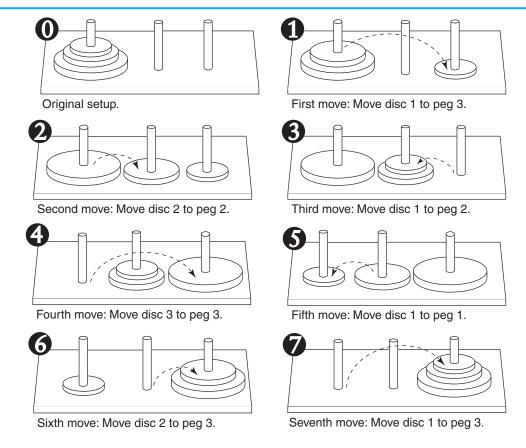
According to the legend, when the monks have moved all of the discs from the first peg to the last peg, the world will come to an end.

To play the game, you must move all of the discs from the first peg to the third peg, following the same rules as the monks. Let's look at some example solutions to this game, for different numbers of discs. If you only have one disc, the solution to the game is simple: move the disc from peg 1 to peg 3. If you have two discs, the solution requires three moves:

- Move disc 1 to peg 2.
- Move disc 2 to peg 3.
- Move disc 1 to peg 3.

Notice that this approach uses peg 2 as a temporary location. The complexity of the moves continues to increase as the number of discs increases. To move three discs requires the seven moves shown in Figure 19-5.

Figure 19-5



The following statement describes the overall solution to the problem:

Move n discs from peg 1 to peg 3 using peg 2 as a temporary peg.

The following algorithm can be used as the basis of a recursive function that simulates the solution to the game. Notice that in this algorithm we use the variables A, B, and C to hold peg numbers.

To move n discs from peg A to peg C, using peg B as a temporary peg: If n > 0 Then

Move n-1 discs from peg A to peg B, using peg C as a temporary peg.

Move the remaining disc from the peg A to peg C.

Move n-1 discs from peg B to peg C, using peg A as a temporary peg. End If

The base case for the algorithm is reached when there are no more discs to move. The following code is for a function that implements this algorithm. Note that the function does not actually move anything, but displays instructions indicating all of the disc moves to make.

```
void moveDiscs(int num, int fromPeg, int toPeg, int tempPeg)
   if (num > 0)
   {
       moveDiscs(num - 1, fromPeg, tempPeg, toPeg);
       cout << "Move a disc from peg " << fromPeg</pre>
             << " to peg " << toPeg << endl;
       moveDiscs(num - 1, tempPeg, toPeg, fromPeg);
   }
}
```

This function accepts arguments into the following three parameters:

The number of discs to move. num The peg to move the discs from. fromPeg toPeg The peg to move the discs to. The peg to use as a temporary peg. tempPeg

If num is greater than 0, then there are discs to move. The first recursive call is

```
moveDiscs(num - 1, fromPeg, tempPeg, toPeg);
```

This statement is an instruction to move all but one disc from from Peg to tempPeg, using toPeg as a temporary peg. The next statement is

```
cout << "Move a disc from peg " << fromPeg</pre>
     << " to peg " << toPeg << endl:
```

This simply displays a message indicating that a disc should be moved from frompeg to toPeg. Next, another recursive call is executed:

```
moveDiscs(num - 1, tempPeg, toPeg, fromPeg);
```

This statement is an instruction to move all but one disc from tempPeg to toPeg, using fromPeg as a temporary peg. Program 19-10 demonstrates this function.

```
// This program displays a solution to the Towers of
   // Hanoi game.
   #include <iostream>
   using namespace std;
 5
   // Function prototype
 7
   void moveDiscs(int, int, int, int);
 8
 9
   int main()
10
11
        const int NUM DISCS = 3; // Number of discs to move
       const int FROM PEG = 1; // Initial "from" peg
12
       13
14
15
16
       // Play the game.
17
       moveDiscs(NUM DISCS, FROM PEG, TO PEG, TEMP PEG);
18
        cout << "All the pegs are moved!\n";</pre>
                                                          (program continues)
```

Program 19-10 (continued) 19 return 0; 20 } 21 //************** 22 23 // The moveDiscs function displays a disc move in * 24 // the Towers of Hanoi game. 25 // The parameters are: 26 // num: The number of discs to move. 27 // fromPeg: The peg to move from. 28 // toPeq: The peq to move to. 29 // tempPeg: The temporary peg. //*************** 30 31 32 void moveDiscs(int num, int fromPeg, int toPeg, int tempPeg) 33 if (num > 0)34 35 36 moveDiscs(num - 1, fromPeg, tempPeg, toPeg); 37 cout << "Move a disc from peg " << fromPeg</pre> 38 << " to peg " << toPeg << endl; 39 moveDiscs(num - 1, tempPeg, toPeg, fromPeg); 40 } 41 }

Program Output

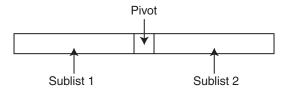
```
Move a disc from peg 1 to peg 3
Move a disc from peg 1 to peg 2
Move a disc from peg 3 to peg 2
Move a disc from peg 1 to peg 3
Move a disc from peg 2 to peg 1
Move a disc from peg 2 to peg 3
Move a disc from peg 1 to peg 3
All the pegs are moved!
```

19.8 Focus on Problem Solving and Program Design: The QuickSort Algorithm

CONCEPT: The QuickSort algorithm uses recursion to efficiently sort a list.

The QuickSort algorithm is a popular general-purpose sorting routine developed in 1960 by C.A.R. Hoare. It can be used to sort lists stored in arrays or linear linked lists. It sorts a list by dividing it into two sublists. Between the sublists is a selected value known as the pivot. This is illustrated in Figure 19-6.

Notice in the figure that sublist 1 is positioned to the left of (before) the pivot, and sublist 2 is positioned to the right of (after) the pivot. Once a pivot value has been selected, the algorithm exchanges the other values in the list until all the elements in sublist 1 are less than the pivot, and all the elements in sublist 2 are greater than the pivot.



Once this is done, the algorithm repeats the procedure on sublist 1, and then on sublist 2. The recursion stops when there is only one element in a sublist. At that point the original list is completely sorted.

The algorithm is coded primarily in two functions: quickSort and partition. quickSort is a recursive function. Its pseudocode is shown here:

```
quickSort:
   If Starting Index < Ending Index
       Partition the List around a Pivot.
       quickSort Sublist 1.
       quickSort Sublist 2.
   End If.
Here is the C++ code for the quickSort function:
   void quickSort(int set[], int start, int end)
   {
       int pivotPoint;
       if (start < end)
          // Get the pivot point.
          pivotPoint = partition(set, start, end);
          // Sort the first sublist.
          quickSort(set, start, pivotPoint - 1);
          // Sort the second sublist.
          quickSort(set, pivotPoint + 1, end);
       }
   }
```

This version of quickSort works with an array of integers. Its first argument is the array holding the list that is to be sorted. The second and third arguments are the starting and ending subscripts of the list.

The subscript of the pivot element is returned by the partition function. partition not only determines which element will be the pivot, but also controls the rearranging of the other values in the list. Our version of this function selects the element in the middle of the list as the pivot, then scans the remainder of the list searching for values less than the pivot.

The code for the partition function is shown here:

```
int partition(int set[], int start, int end)
{
   int pivotValue, pivotIndex, mid;
   mid = (start + end) / 2;
   swap(set[start], set[mid]);
   pivotIndex = start;
   pivotValue = set[start];
```



NOTE: The partition function does not initially sort the values into their final order. Its job is only to move the values that are less than the pivot to the pivot's left, and move the values that are greater than the pivot to the pivot's right. As long as that condition is met, they may appear in any order. The ultimate sorting order of the entire list is achieved cumulatively, though the recursive calls to quickSort.

There are many different ways of partitioning the list. As previously stated, the method shown in the function above selects the middle value as the pivot. That value is then moved to the beginning of the list (by exchanging it with the value stored there). This simplifies the next step, which is to scan the list.

A for loop scans the remainder of the list, and when an element is found whose value is less than the pivot, that value is moved to a location left of the pivot point.

A third function, swap, is used to swap the values found in any two elements of the list. The function is shown below.

```
void swap(int &value1, int &value2)
{
    int temp = value1;
    value1 = value2;
    value2 = temp;
}
```

Program 19-11 demonstrates the QuickSort algorithm shown here.

```
// This program demonstrates the QuickSort Algorithm.
    #include <iostream>
 3
   using namespace std;
   // Function prototypes
   void quickSort(int [], int, int);
7
    int partition(int [], int, int);
8
    void swap(int &, int &);
9
10
   int main()
11
        const int SIZE = 10; // Array size
12
13
        int count;
                             // Loop counter
14
        int array[SIZE] = \{7, 3, 9, 2, 0, 1, 8, 4, 6, 5\};
```

```
15
16
        // Display the array contents.
17
        for (count = 0; count < SIZE; count++)</pre>
            cout << array[count] << " ";</pre>
18
19
        cout << endl;</pre>
20
21
        // Sort the array.
22
        quickSort(array, 0, SIZE - 1);
23
        // Display the array contents.
24
        for (count = 0; count < SIZE; count++)</pre>
25
            cout << array[count] << " ";</pre>
26
        cout << endl;</pre>
27
28
        return 0;
29
30
31
   //************************
32
   // quickSort uses the quicksort algorithm to
33
   // sort set, from set[start] through set[end].
   //*************
34
35
36
   void quickSort(int set[], int start, int end)
37
38
        int pivotPoint;
39
40
        if (start < end)
41
42
            // Get the pivot point.
            pivotPoint = partition(set, start, end);
43
44
            // Sort the first sublist.
45
            quickSort(set, start, pivotPoint - 1);
46
            // Sort the second sublist.
47
            quickSort(set, pivotPoint + 1, end);
48
        }
49
50
   //***************
51
   // partition selects the value in the middle of the
   // array set as the pivot. The list is rearranged so
   // all the values less than the pivot are on its left
   // and all the values greater than pivot are on its right. *
56
   //********************
57
   int partition(int set[], int start, int end)
58
59
60
        int pivotValue, pivotIndex, mid;
61
62
        mid = (start + end) / 2;
63
        swap(set[start], set[mid]);
64
        pivotIndex = start;
65
        pivotValue = set[start];
        for (int scan = start + 1; scan <= end; scan++)</pre>
66
67
68
            if (set[scan] < pivotValue)</pre>
69
```

Program 19-11 (continued) 70 pivotIndex++; 71 swap(set[pivotIndex], set[scan]); 72 } 73 } 74 swap(set[start], set[pivotIndex]); 75 return pivotIndex; 76 } 77 //************ 78 79 // swap simply exchanges the contents of 80 // value1 and value2. //************* 81 82 void swap(int &value1, int &value2) 83 84 85 int temp = value1; 86 87 value1 = value2; value2 = temp; 88 89 }

Program Output

```
7 3 9 2 0 1 8 4 6 5
0 1 2 3 4 5 6 7 8 9
```

19.9

Exhaustive Algorithms

CONCEPT: An exhaustive algorithm is one that finds a best combination of items by looking at all the possible combinations.

Recursion is helpful if you need to examine many possible combinations and identify the best combination. For example, consider all the different ways you can make change for \$1.00 using our system of coins:

```
1 dollar piece, or
2 fifty-cent pieces, or
4 quarters, or
1 fifty-cent piece and 2 quarters, or
3 quarters, 2 dimes, and 1 nickel, or
... there are many more possibilities.
```

Although there are many ways to make change for \$1.00, some ways are better than others. For example, you would probably rather give a single dollar piece instead of 100 pennies.

An algorithm that looks at all the possible combinations of items in order to find the best combination of items is called an exhaustive algorithm. Program 19-12 presents a recursive function that exhaustively tries all the possible combinations of coins. The program then displays the total number of combinations that can be used to make the specified change and the best combination of coins.

```
1 // This program demonstrates a recursive function that exhaustively
 2 // searches through all possible combinations of coin values to find
 3 // the best way to make change for a specified amount.
 4 #include <iostream>
 5 using namespace std;
 7 // Constants
8 const int MAX COINS CHANGE = 100; // Max number of coins to give in change
9 const int MAX COIN VALUES = 6; // Max number of coin values
10 const int NO SOLUTION = INT MAX; // Indicates no solution
12 // Function prototype
13 void makeChange(int, int, int[], int);
15 // coinValues - global array of coin values to choose from
16 int coinValues[MAX COIN VALUES] = {100, 50, 25, 10, 5, 1 };
17
18 // bestCoins - global array of best coins to make change with
int bestCoins[MAX COINS CHANGE];
21 // Global variables
22 int numBestCoins - No_22

23 numSolutions = 0, // Number of ways co .......// Number of allowable coins
22 int numBestCoins = NO_SOLUTION, // Number of coins in bestCoins
                                    // Number of ways to make change
25
26
27 int main()
28 {
29
       int coinsUsed[MAX COINS CHANGE], // List of coins used
30
       numCoinsUsed = 0,
                                          // The number of coins used
                                          // The amount to make change for
31
       amount;
32
33
       // Display the possible coin values.
34
       cout << "Here are the valid coin values, in cents: ";</pre>
35
      for (int index = 0; index < 5; index++)</pre>
36
          cout << coinValues[index] << " ";</pre>
37
      cout << endl;
38
39
     // Get input from the user.
40
      cout << "Enter the amount of cents (as an integer) "</pre>
41
            << "to make change for: ";
     cin >> amount;
42
43
       cout << "What is the maximum number of coins to give as change? ";</pre>
44
       cin >> numCoins;
45
46
      // Call the recursive function.
47
      makeChange(numCoins, amount, coinsUsed, numCoinsUsed);
48
49
      // Display the results.
50
       cout << "Number of possible combinations: " << numSolutions << endl;</pre>
51
       cout << "Best combination of coins:\n";</pre>
52
       if (numBestCoins == NO SOLUTION)
53
            cout << "\tNo solution\n";</pre>
```

```
Program 19-12
                 (continued)
        else
 55
        {
 56
            for (int count = 0; count < numBestCoins; count++)</pre>
 57
                 cout << bestCoins[count] << " ";</pre>
 58
 59
        cout << endl;</pre>
 60
        return 0;
 61 }
 62
 63 //*******************************
 64 // Function makeChange. This function uses the following parameters:
 65 // coinsLeft - The number of coins left to choose from.
 66 // amount - The amount to make change for.
 67 // coinsUsed - An array that contains the coin values used so far.
 68 // numCoinsUsed - The number of values in the coinsUsed array.
 69 //
 70 // This recursive function finds all the possible ways to make change
 71 // for the value in amount. The best combination of coins is stored in *
 72 // the array bestCoins.
 73 //*********************************
 74
 75 void makeChange(int coinsLeft, int amount, int coinsUsed[],
                    int numCoinsUsed)
 77 {
 78
        int coinPos, // To calculate array position of coin being used
 79
            count; // Loop counter
 80
        if (coinsLeft == 0) // If no more coins are left
 81
 82
            return;
 83
        else if (amount < 0) // If amount to make change for is negative
 84
            return;
        else if (amount == 0) // If solution is found
 85
 86
 87
            // Store as bestCoins if best
 88
            if (numCoinsUsed < numBestCoins)</pre>
 89
            {
 90
                 for (count = 0; count < numCoinsUsed; count++)</pre>
 91
                     bestCoins[count] = coinsUsed[count];
 92
                 numBestCoins = numCoinsUsed;
 93
            }
 94
            numSolutions++;
 95
            return;
 96
        }
 97
 98
        // Find the other combinations using the coin
 99
        coinPos = numCoins - coinsLeft;
        coinsUsed[numCoinsUsed] = coinValues[coinPos];
101
        numCoinsUsed++;
102
        makeChange(coinsLeft, amount - coinValues[coinPos],
103
                   coinsUsed, numCoinsUsed);
104
105
        // Find the other combinations not using the coin.
106
        numCoinsUsed--;
107
        makeChange(coinsLeft - 1, amount, coinsUsed, numCoinsUsed);
108 }
```

Program Output with Example Input Shown in Bold

Here are the valid coin values, in cents: 100 50 25 10 5 1 Enter the amount of cents (as an integer) to make change for: 62 [Enter] What is the maximum number of coins to give as change? 6 [Enter] Number of possible combinations: 77 Best combination of coins: 50 10 1 1

19.10 Focus on Software Engineering: **Recursion vs. Iteration**

CONCEPT: Recursive algorithms can also be coded with iterative control structures. There are advantages and disadvantages to each approach.

Any algorithm that can be coded with recursion can also be coded with an iterative control structure, such as a while loop. Both approaches achieve repetition, but which is best to use?

There are several reasons not to use recursion. Recursive algorithms are certainly less efficient than iterative algorithms. Each time a function is called, the system incurs overhead that is not necessary with a loop. Also, in many cases an iterative solution may be more evident than a recursive one. In fact, the majority of repetitive programming tasks are best done with loops.

Some problems, however, are more easily solved with recursion than with iteration. For example, the mathematical definition of the GCD formula is well-suited for a recursive approach. The QuickSort algorithm is also an example of a function that is easier to code with recursion than iteration.

The speed and amount of memory available to modern computers diminishes the performance impact of recursion so much that inefficiency is no longer a strong argument against it. Today, the choice of recursion or iteration is primarily a design decision. If a problem is more easily solved with a loop, that should be the approach you take. If recursion results in a better design, that is the choice you should make.

Review Questions and Exercises

Short Answer

- 1. What is the base case of each of the recursive functions listed in questions 12, 13, and 14?
- 2. What type of recursive function do you think would be more difficult to debug, one that uses direct recursion, or one that uses indirect recursion? Why?
- 3. Which repetition approach is less efficient, a loop or a recursive function? Why?
- 4. When should you choose a recursive algorithm over an iterative algorithm?
- 5. Explain what is likely to happen when a recursive function that has no way of stopping executes.

Fill-in-the-Blank

/	771	c	•		1 1	1 (11	• . 1	0
6	The	Ωt	recursion	10 f	the num	her of	fimes a	tunction	calle	1tsel	t
\cdot	1110	O1	1 CC UI 3I OII	10 1	and munin	DCI OI	tillico a	Tunction	Calls	11301	⊥.

7. A recursive function's solvable problem is known as its . This causes the recursion to stop.

- 8. _____ recursion is when a function explicitly calls itself.
- 9. _____ recursion is when function A calls function B, which in turns calls function A.

Algorithm Workbench

- 10. Write a recursive function to return the number of times a specified number occurs in an array.
- 11. Write a recursive function to return the largest value in an array.

Predict the Output

What is the output of the following programs?

```
12. #include <iostream>
   using namespace std;
   int function(int);
   int main()
   {
        int x = 10;
        cout << function(x) << endl;</pre>
        return 0;
   }
   int function(int num)
        if (num <= 0)
            return 0;
        else
            return function(num - 1) + num;
   }
13. #include <iostream>
   using namespace std;
   void function(int);
   int main()
        int x = 10;
        function(x);
        return 0;
   }
   void function(int num)
        if (num > 0)
        {
             for (int x = 0; x < num; x++)
                 cout << '*';
            cout << endl;</pre>
            function(num - 1);
        }
   }
```

```
14. #include <iostream>
   #include <string>
   using namespace std;
   void function(string, int, int);
   int main()
        string mystr = "Hello";
        cout << mystr << endl;</pre>
        function(mystr, 0, mystr.size());
        return 0;
   }
   void function(string str, int pos, int size)
        if (pos < size)
        {
            function(str, pos + 1, size);
             cout << str[pos];</pre>
        }
   }
```

Programming Challenges

1. Iterative Factorial

Write an iterative version (using a loop instead of recursion) of the factorial function shown in this chapter. Test it with a driver program.

2. Recursive Conversion

Convert the following function to one that uses recursion.

```
void sign(int n)
{
    while (n > 0)
        cout << "No Parking\n";
    n--;
}</pre>
```

Demonstrate the function with a driver program.

3. QuickSort Template

Create a template version of the QuickSort algorithm that will work with any data type. Demonstrate the template with a driver function.

4. Recursive Array Sum

Write a function that accepts an array of integers and a number indicating the number of elements as arguments. The function should recursively calculate the sum of all the numbers in the array. Demonstrate the function in a driver program.

5. Recursive Multiplication

Write a recursive function that accepts two arguments into the parameters x and y. The function should return the value of x times y. Remember, multiplication can be performed as repeated addition:

```
7 * 4 = 4 + 4 + 4 + 4 + 4 + 4 + 4
```



6. Recursive Power Function

Write a function that uses recursion to raise a number to a power. The function should accept two arguments: the number to be raised and the exponent. Assume that the exponent is a nonnegative integer. Demonstrate the function in a program.

7. Sum of Numbers

Write a function that accepts an integer argument and returns the sum of all the integers from 1 up to the number passed as an argument. For example, if 50 is passed as an argument, the function will return the sum of 1, 2, 3, 4, ... 50. Use recursion to calculate the sum. Demonstrate the function in a program.

8. isMember Function

Write a recursive Boolean function named isMember. The function should accept two arguments: an array and a value. The function should return true if the value is found in the array, or false if the value is not found in the array. Demonstrate the function in a driver program.

9. String Reverser

Write a recursive function that accepts a string object as its argument and prints the string in reverse order. Demonstrate the function in a driver program.

10. maxNode Function

Add a member function named maxNode to the NumberList class discussed in this chapter. The function should return the largest value stored in the list. Use recursion in the function to traverse the list. Demonstrate the function in a driver program.

11. Palindrome Detector

A palindrome is any word, phrase, or sentence that reads the same forward and backward. Here are some well-known palindromes:

```
Able was I, ere I saw Elba
A man, a plan, a canal, Panama
Desserts, I stressed
Kayak
```

Write a bool function that uses recursion to determine if a string argument is a palindrome. The function should return true if the argument reads the same forward and backward. Demonstrate the function in a program.

12. Ackermann's Function

Ackermann's Function is a recursive mathematical algorithm that can be used to test how well a computer performs recursion. Write a function A(m, n) that solves Ackermann's Function. Use the following logic in your function:

```
If m = 0 then return n + 1
If n = 0 then return A(m-1, 1)
            return A(m-1, A(m, n-1))
Otherwise,
```

Test your function in a driver program that displays the following values:

```
A(0, 0) A(0, 1) A(1, 1) A(1, 2) A(1, 3) A(2, 2) A(3, 2)
```