Chapter 16: Recursion

Starting Out with Java: From Control Structures through Data Structures

Second Edition

by Tony Gaddis and Godfrey Muganda

Chapter Topics

Chapter 15 discusses the following main topics:

- Introduction to Recursion
- Solving Problems with Recursion
- Examples of Recursive Methods
- The Towers of Hanoi

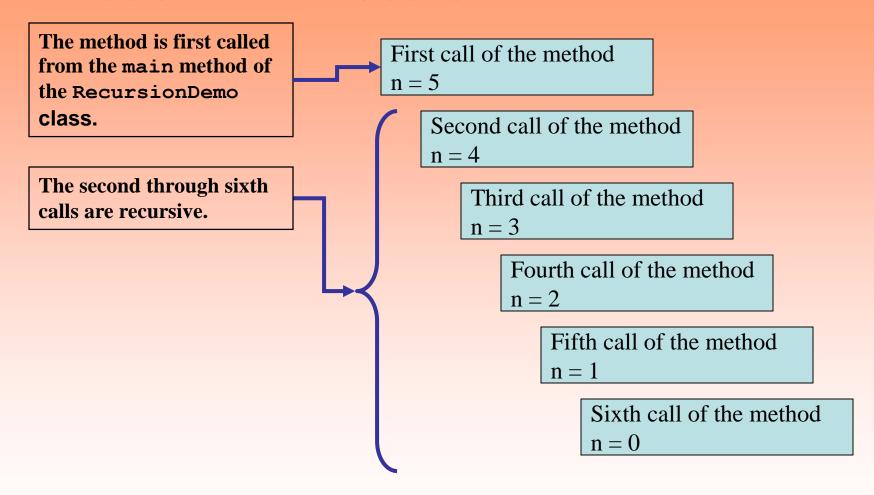
Introduction to Recursion

- We have been calling other methods from a method.
- It's also possible for a method to call itself.
- A method that calls itself is a recursive method.
- Example: EndlessRecursion.java

Introduction to Recursion

- This method in the example displays the string "This is a recursive method.", and then calls itself.
- Each time it calls itself, the cycle is repeated endlessly.
- Like a loop, a recursive method must have some way to control the number of times it repeats.
- Example: Recursive.java, RecursionDemo.java

Introduction to Recursion



- Recursion can be a powerful tool for solving repetitive problems.
- 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 many cases, recursive algorithms are less efficient than iterative algorithms.

- Recursive solutions repetitively:
 - allocate memory for parameters and local variables, and
 - store the address of where control returns after the method terminates.
- These actions are called *overhead* and take place with each method call.
- This overhead does not occur with a loop.
- Some repetitive problems are more easily solved with recursion than with iteration.
 - Iterative algorithms might execute faster; however,
 - a recursive algorithm might be designed faster.

- Recursion works like this:
 - A base case is established.
 - If matched, the method solves it and returns.
 - If the base case cannot be solved now:
 - the method reduces it to a smaller problem (recursive case) and calls itself to solve the smaller problem.
- By reducing the problem with each recursive call, the base case will eventually be reached and the recursion will stop.
- 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, and
 - when n greater than 0, its factorial is the product of all the positive integers from 1 up to n.
- Factorial(6) is calculated as
 - $-1 \times 2 \times 3 \times 4 \times 5 \times 6$.
- The base case is where *n* is equal to 0:

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

- The recursive case, or the part of the problem that we use recursion to solve is:
 - if n > 0 then factorial(n) = $n \times factorial(n 1)$

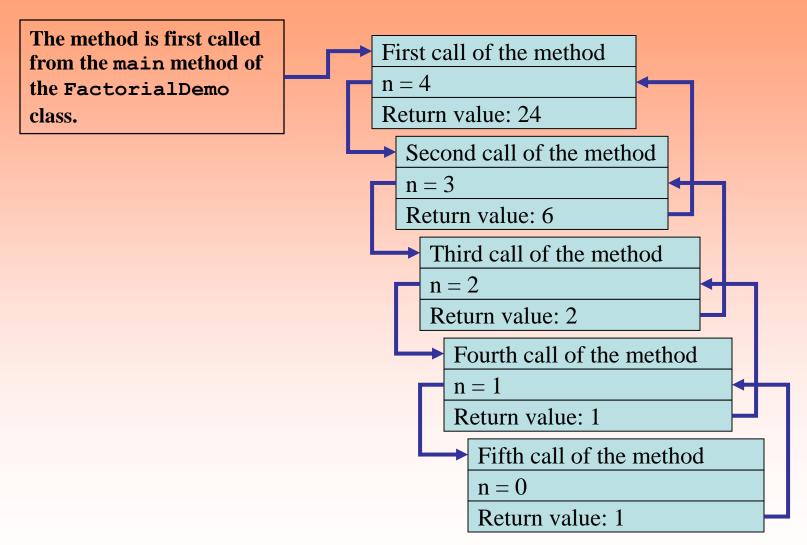
- The recursive call works on a reduced version of the problem, n-1.
- The recursive rule for calculating the factorial:

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

A Java based solution:

```
private static int factorial(int n)
{
  if (n == 0) return 1; // Base case
  else return n * factorial(n - 1);
}
```

• Example: <u>FactorialDemo.java</u>



Direct and Indirect Recursion

- When recursive methods directly call themselves it is known as direct recursion.
- *Indirect recursion* is when method A calls method B, which in turn calls method A.
- There can even be several methods involved in the recursion.
- Example, method A could call method B, which could call method C, which calls method A.
- Care must be used in indirect recursion to ensure that the proper base cases and return values are handled.

Greatest Common Divisor (GCD)

• The definition of the gcd method:

```
public static int gcd(int x, int y)
{
    if (x % y == 0)
        return y;
    else
        return gcd(y, x % y);
}
```

• Example: GCDdemo.java

The Towers of Hanoi

- The Towers of Hanoi is a mathematical game that uses:
 - three pegs and
 - a set of discs with holes through their centers.
- The discs are stacked on the leftmost peg, in order of size with the largest disc at the bottom.
- The object of the game is to move the pegs from the left peg to the right peg by these rules:
 - Only one disk may be moved at a time.
 - A disk cannot be placed on top of a smaller disc.
 - All discs must be stored on a peg except while being moved.

The Towers of Hanoi

- The overall solution to the problem is to move n discs from peg 1 to peg 3 using peg 2 as a temporary peg.
- This algorithm solves the game.

```
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.
- Example: <u>Hanoi.java</u>, <u>HanoiDemo.java</u>