

The Beauty and Joy of Computing¹

Lab Exercise 9: Problem self-similarity and recursion

Objectives

By completing this lab exercise, you should learn to

- Recognize simple self-similar problems which are amenable to recursive solutions;
- Identify components of recursion: base case and recursive case;
- Build simple recursive functions in guided settings; and
- Give rough arguments as to why recursive solutions work.

Background (Pre-Lab Reading)

You should read this section before coming to the lab. It describes how various things work in BYOB and provides pictures to show what they look like in BYOB. You only need to read this material to get familiar with it. There is no need to actually do the actions described in this section, but you certainly can do them if it would help you follow along with the examples.

A **recursive algorithm** is an algorithm that implements a function by *calling itself* in a controlled way. The general process of a function calling itself is called **recursion**. Recursion is an extremely powerful computing technique, but can be confusing for beginning programmers (and sometimes even for experienced programmers!). However, once you learn to “think recursively” there are a huge number of problems that have natural recursive algorithms or recursive ways of expressing a solution. In computer science there are problems for which recursion leads to highly efficient algorithms; there are also problems for which recursion is the *only* readily apparent way to solve the problem; there are important data structures that are recursively built; and there are powerful algorithmic techniques like dynamic programming in which the first step is to recursively characterize the solution to the problem. All this is just to point out how important and central recursion is to computer science - while you may not be particularly comfortable with recursion after the light coverage in this class, if you take more computer science classes this concept will keep coming up over and over. You will eventually (hopefully!) gain a deeper understanding of how to use recursion, and in this lab you will start your journey to understanding recursion by doing some basic and carefully guided examples.

We start by demonstrating the concepts of recursion and the techniques for designing a recursive algorithm with a simple example: computing the factorial function. As you hopefully remember from math classes, n factorial (written $n!$) is the product of the first n positive integers:

$$n! = 1 * 2 * 3 * 4 * \dots * (n-1) * n$$

So looking at a few examples, $2!$ is 2, $3!$ is 6, and $5!$ is 120. The factorial function exhibits a property that we call self-similarity: the solution for $n!$ “contains” solutions for smaller versions of the same problem. Look at the formula above - if we simply leave out the last term (the last

¹ Lab Exercises for “The Beauty and Joy of Computing” [This is the final version for the Fall 2012 class]

Copyright © 2012 by Stephen R. Tate - Creative Commons License

See <http://www.uncg.edu/cmp/faculty/srtate/csc100labs> for more information

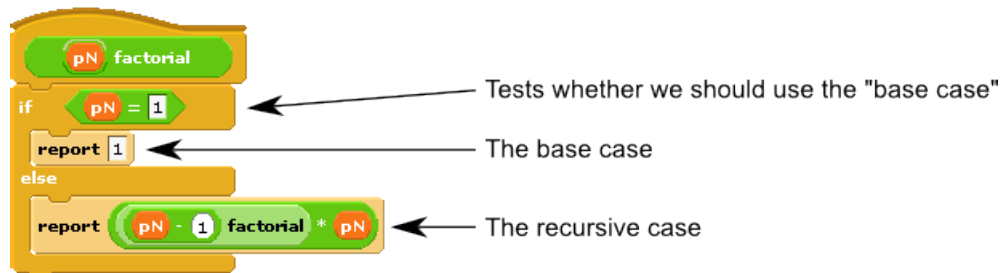
factor of n), then we have $(n-1)!$. In fact, this shows us that for large enough n (“large enough” so that there are enough terms in the formula to split it like this), we see that $n! = (n-1)! * n$.

So for “large enough” n , we can write the solution for $n!$ using the solution for $(n-1)!$ and a small amount of additional work (one additional multiplication). What does “large enough” mean? In this case, it just means that n is at least 2 (so there is something to the left of the last term). If n is not “large enough”, then we can just write down the answer. In this case, if $n=1$, we can just say that $n!$ is 1.

This is how recursion always works: we consider two cases, depending on whether the input is large enough to express the answer in terms of solutions to smaller problems, or whether it is so small that we can’t break it down but can write the solution out directly. We call these the **recursive case** and the **base case**, and identifying these two cases is the first step in designing a recursive algorithm. Writing this out explicitly, for the factorial problem we might start like this:

- Recursive case: When $n \geq 2$, we know that $n! = (n-1)! * n$
- Base case: When $n = 1$, we know that $n! = 1$

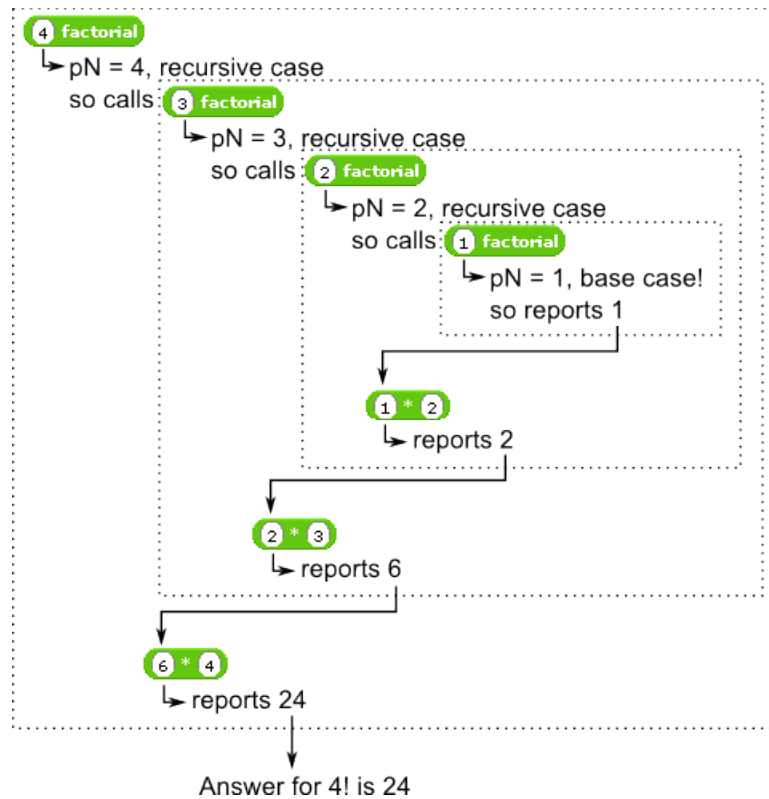
Now comes the magic part: we can write out code based on these observations. First, we use an if/else block to test whether we are in the base case or in the recursive case, and then put in code for each case. While we described the recursive case first above, typical practice in programming is to test for and handle the base case first, and so the result is the following BYOB block definition to compute the factorial function:



Look carefully at the formula in the last “report” block (in the “else” part of the if/else). The innermost block is $(pN-1)$, which is supplied as the argument to the “factorial” function. Assuming our recursive call magically works (and it does!), it gives $(pN-1)!$ which we then multiply by pN and report that value.

It’s worthwhile to trace out how this works, but it should be stressed that tracing out recursive functions is *not the right way to think about recursion!* We’ll do this once, for this one example, and once you are reassured that this actually works then never do it again! The right way to think about recursion is just building the base case and the recursive case. Thinking about the recursive case, just think about one step - don’t keep tracing beyond a single step to a smaller problem, because if you think across multiple steps then you’re really missing the whole point of recursion. That said, here’s how the factorial function defined above works - first, calling factorial for argument 4, this is the recursive case so we end up calculating $pN-1$ (which is 3) and calling factorial with argument 3. This calls factorial of 2, and finally that calls factorial of 1. We stop making recursive calls at this point since we’re in the base case. Our functions now return, and

at each step multiplying the result of the recursive call with the value of pN at that level. We can visualize this as follows:



Let's reason about why this works. We start off by asking, if I call "factorial" with an argument of 4, does it give the right answer? Well, we compute factorial of 3 and multiply that by 4, so as long as factorial of 3 is computed correctly then we get the right answer. Do we compute factorial of 3 correctly? Well, to compute factorial of 3 we compute factorial of 2 and multiply that by 3, so as long as we compute factorial of 2 correctly then factorial of 3 is correct (which will make factorial of 4 correct!). Do we compute factorial of 2 correctly? Well, to compute factorial of 2 we compute factorial of 1 and multiply that by 2, so as long as we compute factorial of 1 correctly then factorial of 2 is correct (which will make factorial of 3 and hence factorial of 4 both correct). Finally, do we compute factorial of 1 correctly? Yes! That's our base case, so the correct value is simply filled in - so all intermediate results are correct, giving us the correct value for 4 factorial. That was a somewhat tedious argument, but if you think about how the correctness builds from smaller cases to bigger cases it should make sense. In later computer science classes you'll see a way to formalize this kind of argument in what is called a **proof by induction** - a powerful way of reasoning about the correctness of statements and algorithms.

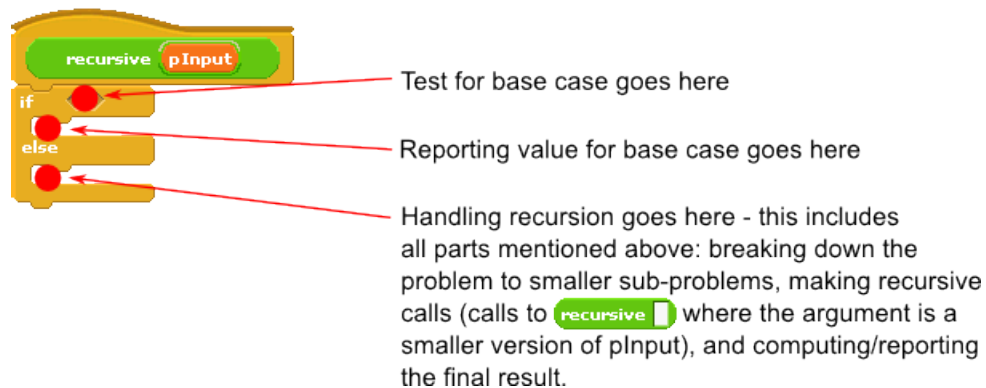
Let's go back and look at this specific example and think about the design process for coming up with this recursive algorithm. In designing a simple recursive function, we need to design the two specific parts of the recursion, which requires answering the following questions:

1. *What is the recursive step?* The recursive step will do three things: break down the problem to one or more smaller versions of the same problem, called the subproblem(s)

(we've only seen a single subproblem in our example, but there might be more, like in Activity 3 below); make the appropriate recursive call(s); and process the results of the recursive calls to produce our answer. In the factorial case, the only computation that had to be done to "break down" the problem of computing n factorial to a smaller subproblem was to compute $n-1$; then we made the recursive call; then we did the post-recursion processing, which in this case was just taking the result of the recursive call and multiplying by n . In deciding on the recursive step, it is best to picture in your mind a value of n that is not very small - for example, don't think about recursion when n is 2 - think about how the recursion works when $n=10$ or $n=100$.

2. *What is the base case?* The recursive step should result in a recursive call with a smaller argument than the original argument - for example, in the factorial problem we make a recursive call with argument $n-1$. Each recursive step makes the input smaller and smaller, and so at some point you reach the smallest possible problem, which you must solve directly. This is the base case, and there are two parts to the base case. First, what defines the base case and how can you test if your input falls into the base case rather than the recursive case? In the factorial example, we are writing a factorial function that handles inputs n for all $n \geq 1$, so the smallest possible input is $n = 1$, which defines the base case (and gives the test for being in the base case). The second thing we must figure out is what the value of the function is in the base case. For the factorial problem, when $n = 1$, we know that $n!$ is 1, so that's our answer.

Just like iteration over lists, there is a pattern we can use for basic recursion, which looks like this:



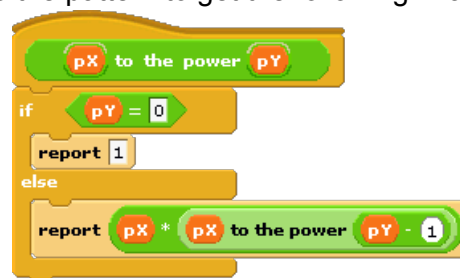
Lets do one more numerical example of recursion, which will be very similar to what you will do in activities 1 and 2 below. BYOB does not have a block to compute powers, and we'd like to be able to raise numbers to powers that are non-negative integers. In other words, we are interested in being able to compute things like 2^3 , 3^2 , 2^9 , or even $(2.5)^3$ or $(1.25)^8$, but not something with non-integer exponents like $5^{1.5}$. Expressed as a general formula, we want to compute x^y , where y is an integer. Let's ask and answer the questions given above, and then we can fill in our code pattern to get the final recursive algorithm.

What is the recursive step? One thing that makes this problem a little more difficult is that we have two numbers that are in the input, so we have to decide whether to make one of them smaller or both of them smaller in the recursion (some recursive solutions are even more

complicated than this, but you would never need anything complex for this class). In this case, there is a clear answer if you think about powering as repeated multiplication (so $x^4 = x * x * x * x$): we will do one less multiplication for the sub-problem, so we reduce the exponent. Now we need to write out x^y in terms of $x^{\text{something smaller than } y}$. There's actually more than one way to do this, but for now we'll just think of "peeling off" one multiplication so that we use $x^y = x * x^{y-1}$. This has defined our recursive step! The pre-recursion processing is computing $y-1$, the recursive call is computing x^{y-1} , and the post-recursion processing is multiplying x by the result of the recursive call.

What is the base case? In this situation, we are computing powers in which the exponent y gets smaller and smaller, so we ask "What is the smallest value of y for which this problem makes sense?" Since we said above that the power will be a non-negative integer, the answer in this case is $y = 0$ - we might ask what x^0 is for some x , and we can in fact easily answer that since x^0 is 1 for all values of x (let's ignore the slightly strange case of 0^0). So now we have answered this question: we can tell if we're in the base case by testing if $y = 0$, and we can handle the base case by reporting 1.

Now we plug our answers into the pattern to get the following final recursive algorithm definition:



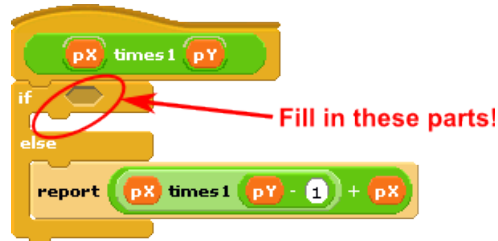
This is our final recursive function to compute powers. Testing it out with some simple values, we see that it works quite well!

Activities (In-Lab Work)

Activity 1: Early microprocessors in 1970's and early 1980's, like the Intel 8080, Motorola 6800, and early Sun SPARC processors, did not include hardware or instructions to multiply integers, so programmers supplied functions to do multiplication using an algorithm made out of simpler operations. While things are much simpler today since microprocessors are more powerful, this problem illustrates some important aspects of recursion, so we'll look at it in this activity. Practically identical issues arise in modern code for doing large modular exponentiations in cryptography, but we'll focus on the simpler setting of multiplication for this lab.

In the pre-lab reading we treated powering as repeated multiplication, and in this activity we will do something very similar but treating multiplication as repeated addition. For example, if we want to compute x^4 , we can write this as $x+x+x+x$, accomplishing the multiplication with four additions. In general, if we are interested in computing $x*y$, where y is a non-negative integer, we can write $x*y = x*(y-1) + x$. This defines our recursive step. What you are to do in this activity is write a "times1" reporter block that implements this recursive algorithm (note the 1 in the

name - we'll make a "times2" in Activity 2, so naming the blocks like this will keep them separate). The following picture shows the block definition with the recursive part filled in:



You need to think about the base case: what is the appropriate base case (what value of y), and what should you report in the base case? Once you have written this code, test it on some simple values and convince yourself that it works. Save your final code in a file named Lab9-Activity1.

Activity 2: In this activity, we'll consider a different way to handle the recursive case. The time complexity of a recursive algorithm depends strongly on how much we reduce the problem at each recursive step, which determines how fast we can reach the base case, and just decreasing it by one is not a particularly fast way to get to the base case. For multiplication, it turns out that we can do a recursive case that roughly halves the parameter y at each step, rather than just reducing it by 1. So, for example, if we were to call this with $y = 10,000$, then our solution in activity 1 will make a recursive call with $y=9,999$ and then $9,998$, and then $9,997$ and so on; however, this new technique will make a recursive call with $y=5,000$ and then $2,500$, and then $1,250$ and so on. Note that it looks like we're multiplying and dividing in the following algorithm, even though we said we couldn't multiply - isn't that a violation of our restrictions? Not in this case: all multiplications are doing multiplication by 2 and divisions are dividing by 2 - these are really simple operations on binary numbers, just like multiplying and dividing by 10 are trivial in decimal. Multiplying a binary number by two just shifts all the bits left one position, and dividing by two shifts all the bits to the right, so these operations are allowed.

The key to this is to think about what you can do if y is even. In that case, we notice that $x*y = 2*(x * (y/2))$. So if we made a recursive call to compute $x * (y/2)$ and saved that result in a variable, we could multiply that by two (or add that variable to itself) to get $2 * (x * (y/2)) = x*y$. Note that if you add the recursive call to itself, it is extremely important to use a variable - if you were to make two recursive calls to add together, rather than using a variable, this would be slower. Really, really, *really* slower. Take my word for it.

What if y is odd? In that case, we can reduce y by 1 so that it's even, pull the same trick to get $2*(x * (y-1)/2) = x*(y-1) = x*y - x$, where $x * (y-1)/2$ is a recursive call, and add x to the result to get the answer that we need. In other words, when y is odd, we can compute $2*(x * (y-1)/2) + x = x*y$. If we write this multiplication algorithm out as $\text{mult}(x,y)$, then the way to express this as a mathematical expression is as follows:

$$\begin{aligned} \text{mult}(x,y) &= 2 * \text{mult}(x, y/2) && \text{if } y \text{ is even;} \\ \text{mult}(x,y) &= 2 * \text{mult}(x, (y-1)/2) + x && \text{if } y \text{ is odd.} \end{aligned}$$

Your job is to take that description and formula, and fill in the missing pieces in this code for “times2”:

```

times2(pX, pY)
if [ ]
else
  script variables sSubAnswer sAnswer
  if [pY mod 2 = 0]
    set sSubAnswer to [ ]
    set sAnswer to [ ]
  else
    set sSubAnswer to [ ]
    set sAnswer to [ ]
  report sAnswer

```

Use same base case as Activity 1

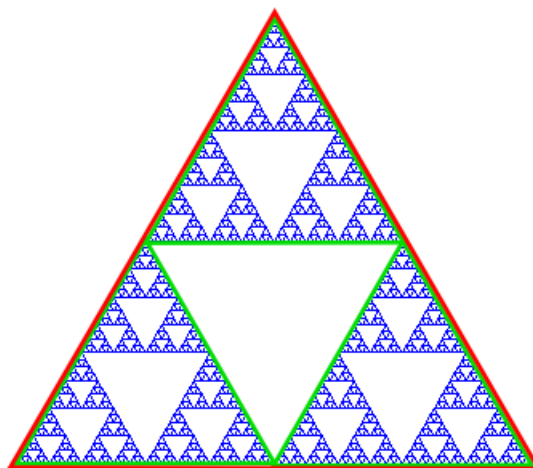
Make appropriate recursive calls (to "times2") and save result

Use the result of the recursive calls to produce the answer

Make sure you test this thoroughly to make sure it works correctly for a variety of y values (at the very least, make sure you try some even and some odd values). Once you’ve got it working, you should compare the time required by the “time1” algorithm from Activity1 and the “time2” algorithm from Activity 2. Use the timing techniques from Lab 6 to see how long these algorithms take to multiply 1000 by 1000, and record these times so you can report them in the lab quiz questions.

Save your final project as Lab9-Activity2.

Activity 3: There are several cool looking self-similar shapes that people have come up with over the years, and recursion is a great way to draw these figures. This is more complex than the previous activities, but if you take it a small step at a time you should be able to do this. The shape you’re going to draw is a Sierpinski triangle, which looks like this (the red and green outlines are not part of the drawing, but are marking parts of it for this discussion):



So what exactly is a Sierpinski triangle? A Sierpinski triangle is a triangle made by drawing three smaller Sierpinski triangles and placing them next to each other as shown: two side by side and one on top of those two. In the diagram, the three smaller Sierpinski triangles are outlined in

green, and after combining they make a larger Sierpinski triangle, outlined in red. Think how strange this definition is: To draw a Sierpinski triangle, you draw three Sierpinski triangles. That's a recursive definition! To draw this, you will create a BYOB block that will be used something like this (these are in fact the exact parameters that created the drawing above):

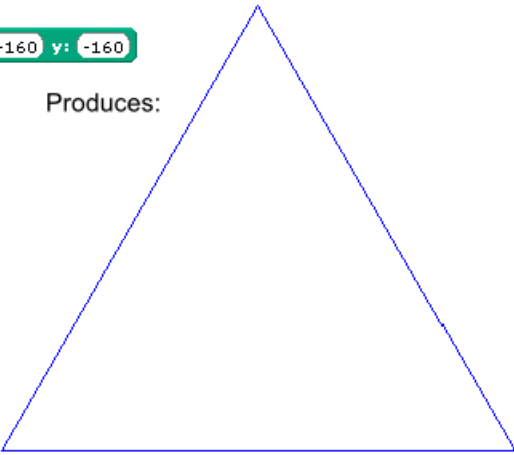
```
clear
draw sierpinski len: 320 levels: 7 at x: -160 y: -160
```

The parameters give the length of a side of the Sierpinski triangle (320 in this case) and the coordinates of the lower left corner of the triangle ((-160,-160) here). The remaining "levels" parameter says how many recursive calls to make drawing Sierpinski triangles - this one makes 7 recursive calls, making 7 levels of Sierpinski triangles.

The recursion will use the "levels" parameter, with the base case being a 1 level Sierpinski triangle, which looks like this:

```
draw sierpinski len: 320 levels: 1 at x: -160 y: -160
```

Produces:



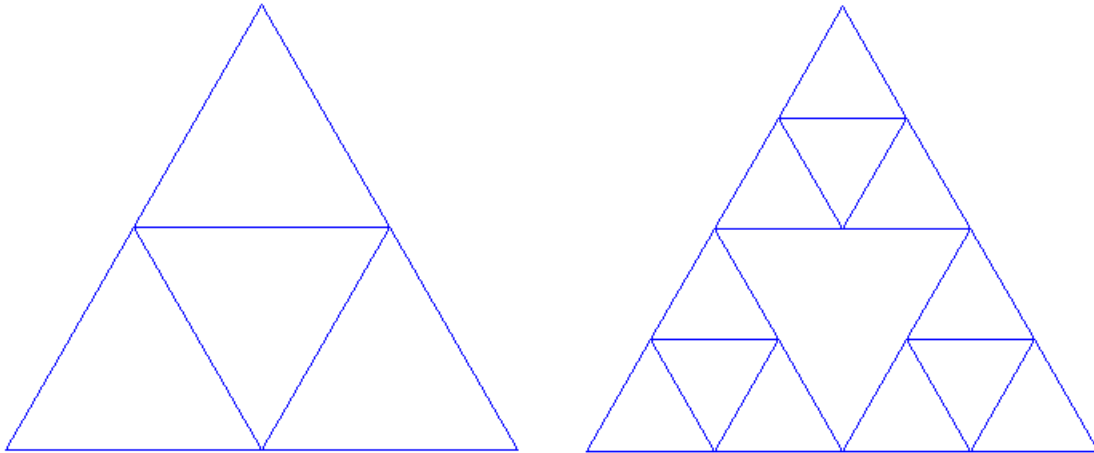
That's right, a 1-level Sierpinski triangle is just a regular equilateral triangle, "len" units per side. Your first task is to implement this simple base case. Here's what your code needs to do, written out in text (your job is to turn this into a BYOB script): raise the pen up, go to the specified (x,y) coordinates, put the pen down, draw by moving up and to the right (figure out the angle!) for "len" steps, turn 120 degrees and move another "len" steps, turn 120 degrees and move another "len" steps. Get this code working reliably before moving on: make sure you can draw at different sizes and different locations. This script is the base case of the "draw sierpinski" block, and the condition for the base case is "levels = 1" - make sure you put this in the basic recursion pattern that we have been using for all of our examples and activities.

Next, do the recursive case. The recursive case is exactly three recursive calls that each draw a Sierpinski triangle with side length "len/2" and with "levels-1" levels. You have to figure out the coordinates for each triangle, but here's a hint: the one on the lower right is drawn by this block:

```
draw sierpinski len: pLen / 2 levels: pLev - 1 at x:
pX + pLen / 2 y: pY
```

The only tricky coordinates are the one on top. You have to use some basic trigonometry to figure out this location - try to figure it out yourself, but if you can't figure it out look at the solution at the end of this handout.

Once you think you've got the three recursive calls correct, test it! The two pictures below show a 2-level Sierpinski triangle (on the left) and a 3-level triangle (on the right):



If you have any bugs in the recursive case, they'll show up in the 2-level triangle, which should just stack three simple triangles together in the proper locations.

Once you've got this working, save your project as Lab9-Activity3.

Activity 4: The following extra credit challenge does not need to be turned in during lab time. It can be turned in at any time before next week's lab, and is worth up to 40 points extra credit. It's not easy, but if you take it a step at a time you should be able to do it! Here's the problem: given a list of coin values, and an amount you need to make change for, how many different ways are there to make change for this amount? For example, the list of coin denominations for US coins up to a dollar would be the following list:



Your goal is to make a reporter block so that you could, for example, count the number of different ways you could make change for one dollar. Using this block could look something like this:

count ways to make 100 using first 6 of coins

The first argument is the amount you want to make change for (here 100 cents, or one dollar), and the last parameter is the list of coin denominations. The middle parameter is what controls the recursion, and essentially says that you are allowed to use any of the first 6 coin types (since there are exactly six coin types overall, you can use any coin type in this example). If you were to call this with a 1 for the middle argument instead of a 6, then it would be the number of different ways you could make change for a dollar using only pennies (and there's only one way to do that: you have to use 100 pennies!). If you called this with the middle parameter being 2,

then you can use pennies and/or nickels, and there are 21 different ways to do this (you can use 0 nickels, 1 nickel, or any number of nickels up to 20, making up the remainder of the dollar with pennies).

Hints: The middle parameter controls the recursion, so concentrate on this. Make a base case where you directly handle the case when the middle parameter is 1. Then figure out the recursive step: if you can use the first n coin types, you want to loop through every possibility for the number of coin type n can be used, and recursively call to make up the remainder of the amount with the first $n-1$ coin types. This is quite challenging for someone new to recursion, but is the kind of thing that is very natural once you gain some experience with this!

If you get this working, save it as Lab9-Bonus, and submit separately from the main lab assignment. In addition to writing the code, use your script to figure out how many ways there actually are to make change for a dollar using the US coin denominations. Then create a second list with European coin denominations (1, 2, 5, 10, 20, 50, and 100) and compute how many ways there are to make change for 1 Euro - you might be surprised by the answers!

Discussion (Post-Lab Follow-up)

Self-Similar Shapes: (Note - you're not responsible for this, but it's cool stuff nonetheless!)

There are lots of self-similar diagrams that have been defined over the years, with the two most famous being the Sierpinski diagram that you drew for Activity 3 and the Koch curve (sometimes called the Koch snowflake). Look this up on Wikipedia and you'll see a cool animation that zooms in repeatedly on a portion of the snowflake and you see an infinitely detailed snowflake emerge in this animation. The famous diagrams are all very regular and structured, but you can in fact make diagrams that have some irregularity but use the same rules on different scales, producing self-similar but irregular diagrams with infinite resolution. A lot of game engines use this kind of technique to generate things like game maps, where coastlines (or mountains) are curves that can be generated algorithmically, at whatever "zoom" level you want, and you always have as much resolution as you need.

Terminology

The following new words and phrases were used in this lab:

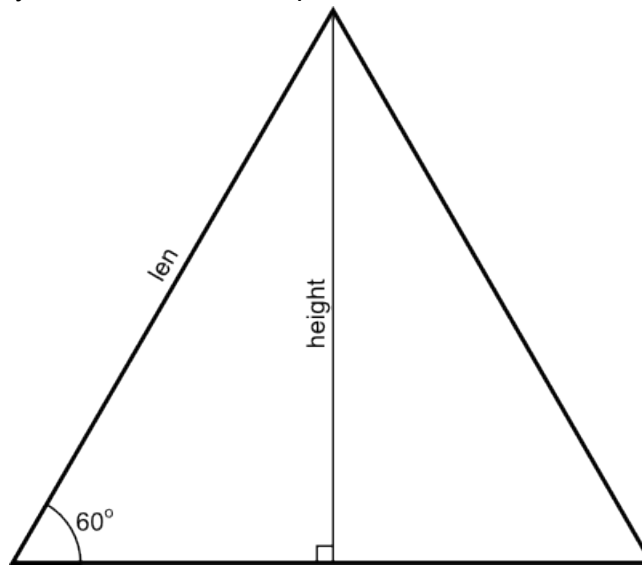
- ***base case***: the simplest and smallest case in a recursive algorithm, that is handled directly rather than making calls to smaller versions of the problem
- ***proof by induction***: a way of reasoning that builds up the truth of a statement (such as an algorithm working correctly) based on the truth of smaller statements that it depends on
- ***recursion***: the process of defining something in terms of itself
- ***recursive algorithm***: an algorithm that implements a function by calling itself
- ***recursive case***: the part of a recursive algorithm that calls itself on smaller inputs

Submission

In this lab, you should have saved the following files: Lab9-Activity1, Lab9-Activity2, and Lab9-Activity3. Turn these in using whatever submission mechanism your school has set up for you. If you do the bonus problem (Activity 4) then you will submit this separately.

Solution To Sierpinski Coordinates

In Activity 3, you had to figure out (x,y) coordinate so that you could place the three recursive Sierpinski triangles. The first step is to figure out the height of the triangle, given the “len” parameter - then the y location of the top sub-triangle is calculated by adding twice the height of the triangle to the base y location. Here’s the picture to consider:



Figuring out the height is a basic trigonometry problem, using a cosine: $\text{height} = \text{len} * \cos(60)$. Using a calculator to find the cosine, we find that the height is approximately $0.866 * \text{len}$. With this, you should be able to place all three recursive triangles.