Computational Projects

Lecture 8 part 2 **: Ideas about programming**

Dr Rob Jack, DAMTP

http://www.maths.cam.ac.uk/undergrad/catam/part-ia-lectures

What makes a good program?

A good program should do a specific job, reliably and efficiently.

> ideally, it should be flexible: it should work in general cases, not only for simple cases. It should provide output that is easy to understand.

For any given task, there are would be many possible programs that can accomplish this

> *When you write a program, think of it primarily as a work of literature. You're trying to write something that human beings are going to read. Don't think of it primarily as something a computer is going to follow. The more effective you are at making your program readable, the more effective it's going to be: You'll understand it today, you'll understand it next week, and your successors who are going to maintain and modify it will understand it.*
>
> Donald Knuth

Standard advice

Give names to your variables and functions that make your program easy to read

> In the old days there were reasons to use short variable names and people often used single letters, `i,j,k` etc
>
> *... no longer a good idea...*

Include plenty of comments

Use functions sensibly, avoid copy-pasting similar sections of code...

Don't be tempted to do "something clever" unless you are really sure that it makes your program more efficient *(... keep it simple)*

something clever: Recursion

A recursive function is a function whose definition includes the function itself

```
function [ y ] = recursiveFactorial( x )
%recursiveFactorial compute x factorial

  % check x>=0, else we can get stuck in an infinite loop
  if x < 0
      error('factorials of positive numbers only')
  end

  if x <= 1
      y = 1;
  else
      y = x * recursiveFactorial( x-1 );
  end

end
```

compute the factorial of x-1

"to understand recursion, you must first understand recursion"

something clever(?): Recursion

... keep it simple

```
function [ y ] = simpleFactorial( x )
%simpleFactorial compute x factorial

  if x < 0
      error('factorials of positive numbers only')
  end

  y=1;
  for k=1:x
      y = y*k;
  end

end
```

... easier to understand, just as efficient

(Also, recursive functions are often harder to debug and test; they can be useful in some situations but caution is advised...)

Debugging

As you practise programming, you will find that sometimes your programs don't work as you hoped...

The process of fixing the problems ("bugs") is called debugging

Most people agree that debugging is the hardest and most frustrating aspect of programming...

The best way to help with debugging is:
write programs with fewer bugs in the first place

... but there are other useful strategies...

Avoiding bugs

Keep it simple

Add comments to your program *as you write it*, don't think of this as something to do at the end (... once it's working)

Break the problem into functions and test each function separately (make sure that you know what will happen if a function gets an unexpected input... consider adding a check)

Sometimes you can solve the problem by breaking it into several parts and combining them together at the end...

... when combining, it will help a lot if you have used sensible variable names and you have included comments.

Debugging

There is no perfect debugging method but a good plan is:

Try to guess what might be wrong, and *test this hypothesis*

A classical method is to print out the values of variables that might be wrong, and check them

Eg for LU decomposition, maybe you want to find a simple example, print out all the matrices $A^{(k)}$, and check that the first k rows and columns are indeed full of zeros

Once you find one thing that is wrong, try to trace things backwards: where did the first mistake appear?

Debugging

... in a similar vein:

If you are doing something complicated, try to gradually make it simpler, and see if the problem survives

... if you deal with matrices, is the problem still there if you use a 3x3 matrix, where you can check everything by hand?

... if you do something iteratively, is the problem already there in the first iteration? At what iteration does it appear?

... can you tie down the problem to one specific function that is getting 'correct' input but producing 'incorrect' output?
(can you make a *minimal, complete, and verifiable* example?)

Breakpoints, stepping, etc

MATLAB can help you with debugging

The workspace window will tell you the current values of all variables

You can use "breakpoints" to make scripts stop in the middle, so that you can check the values of variables

You can "step" through programs, one line at a time, to see what is happening

In other programming languages you can often get this functionality through an IDE (integrated development environment)

Summary

If you can come up with easy-to-read programs that do computations in a clear and logical way then they will be easier to develop, easier to use, and easier to modify

This sounds easy but the only way to be good at it is to practise!

Bugs are just mistakes, we all make them, the trick is to find them early, when they are easy to fix...