As programming projects become more complex, we need techniques for making projects manageable. In general, the probability that we can write all the code of the project at once and expect it to just work is very low. And when things do not work, it becomes increasingly difficult to isolate the problem, and in fact, it may be the case that the error is in the interaction of two or more parts of the project that seem to work fine individually. This handout is a work-in-progress intended to give the reader some basic tips on managing the complexity of a large programming project. It is neither exemplary nor exhaustive, but if you follow the tips, it should make managing any project easier.

0. Serious analysis and design up-front reduces implementation time
This may seem obvious, but it is the most important tip. Resist the urge to write any code until you have thought about what you are trying to do and how to get there. The more you know where you are going, the less time it will take to get there. The goal of this step is to break down the problem into small, well-defined pieces.

An exception to this guideline is sometimes you do want to write some prototype code just to see how something works. That's fine as long as you expect to "discard" the actual code and do not try to evolve it into your entire project without serious thought as to how it fits in.

1. Implement the pieces of your design a few at a time
This also may seem obvious. Resist the urge to write all the code at once, and then try to debug it. Solve the pieces a few at a time, making sure that they work by writing one or more driver programs that test the pieces individually. Start with the foundation pieces, then move on to the more refined pieces. Always test the interactions between what you have already written and the new piece by rerunning your previous tests. For example, when writing a new class, start with the constructors, so that you can build objects, then the accessors and output operations that let you see what you have built. Write a driver program that uses just these operations. If the project requires more than one class, start with the simpler one.

Sometimes it is advantageous to write a skeleton of a function without any of the actual implementation. That is, you write the prototype and header of the function, but only a minimal body, usually just a message saying the function isn't implemented, yet, and/or returning some default value. These skeleton functions are called stubs and should be used whenever you know you want a particular function, but want to defer actually implementing it until a later time. This allows you to compile and run a program that uses the function, though of course, the results will not be valid, allowing you to see if the other parts of your program work.

2. Debugging statements should contain real information
A debugging message such as "Entering operator=" or "operator+ : rhs operand is empty" is much more informative than messages like "Got here" or "Here 2". If you are going to write debugging statements, they might as well say something useful.

One way to manage these statements is to use debug flags and a PrintDebug function:

```cpp
const bool DBG_PHASE1 = false;
const bool DBG_PHASE2 = true;

void PrintDebug (bool debug, string message) {
```
if (debug)
    cerr << message << endl;
} // end PrintDebug

The PrintDebug function just prints a message if debugging is "on" (represented by the true value) during the call. The flags are used to turn on and off groups of debugging messages. (Otherwise, you would have to find each one and change the call argument from true to false or vice versa.) For example, a default constructor body might have:

    PrintDebug (DBG_PHASE1, "Entering default constructor");
    // Statements to create default object
    PrintDebug (DBG_PHASE1, "Exitting default constructor");

Another tip is to "indent" the debugging messages if they are inside nested control structures. For example,

    PrintDebug (DBG_PHASE2, "Entering operator+");
    :
    if (rightPolyPtr == 0)
    {
        PrintDebug (DBG_PHASE2, " operator+: rhs operand is done");
    :
    }

Now when the program runs, the debugging statements line up as:

    Entering operator+
    operator+: rhs operand is done

and you can see the control structure in the output.

When you are finished with your project, you can set all of the debugging flags to false. A good optimizing compiler should be able to detect that the conditions will be false and that the function will not have any other code executed and optimize away the function calls and perhaps the entire function itself. Later, if there is a newly-discovered problem, or you want to make an enhancement, you can turn the debugging messages back on easily.

3. Start with the simplest test cases: expected and boundaries
   This also may seem obvious, but resist the urge to use a complex test case the first time you run your program. Instead start with the simplest expected test case that will demonstrate the use of the piece you are working on. Once that test passes, go onto to the next simplest test case, etc.

   Generally, the number of interesting test cases can be reduced to the phrase "0/empty, 1, more than 1, n/full." For example, this means that one should test for the case of no input, one input, and expected input. In some cases, there will also be a maximum size input that needs to be tested. A concrete example is inserting into a doubly-linked circular list. The algorithm is different for lists of 0, 1, and more than 1 elements, so you should test for each case.

4. If things get too confusing, start over
   Sometimes the right thing to do when you cannot figure out where to start debugging is to just start over writing
the program. Usually this happens after you have made a major changes to the design of the project and have forgotten to propagate the changes through all of the existent code. Do this a few times, and you are likely to forget which constructs reflect the latest change and which ones still need to be changed. Throw in some old comments, and you have a mess.

Although, it may seem like a "waste" to discard the work, it really is not. By having written the discarded code, you have figured out what you really want to do. So when you start over, you will be able to complete the project faster, more efficiently, and it will be cleaner than trying to "save" the old work.

5. **When all else fails, use a symbolic debugger**
A symbolic debugger is a program that allows you to watch the state of your program as it runs and control how your program runs. It allows you to single-step through your code line by line and look at the values of all the variables. It never makes a mistake, so if you don't see the execution you were expecting or don't get the result values you were expecting, then you know something is amiss.

The rest of this handout is a very short tutorial on how to use **gdb**, a debugger under Linux for program written using **g++**. First, the source files in the project must be compiled with the **-g** option. This leaves the symbolic information in the executable. Usually, this is done by adding the option to the compiling command lines of a makefile. For example,

```
testdate.o : testdate.cpp date.h
<TAB> g++ -Wall -c -g testdate.cpp
```

(Normally, the symbolic information is stripped out of the executable to make it smaller, since the information is not needed by the program and run-time system. When you are done using the debugger, you should take the **-g** flag out of your makefile commands.)

To use gdb, invoke it with (only) the program's name as its only command-line argument. For example,

```
hwang@dh27pc:~/cs215/projects$ gdb testdate
```

This will start the debugger and load the program. Then the debugger waits for commands. The most common commands (in alphabetical order) are:

- **break <function>** - sets a breakpoint that causes the program to stop running and (re)enter the debugger whenever the function is called. You can also break on line numbers.
- **run** - start running the program. If the program expects command-line arguments, they are given here.
- **continue** - continue running the program after a breakpoint. Usually used if you just want to check something at the breakpoint and then go on rather than single-stepping.
- **next** - execute the next statement of code at the current level. If it is function call, this means to execute the function and return in one step.
- **print <expression>** - print the value of the expression. Usually used to look at the values of parameters and variables (including array and class elements). Can also be used as a calculator.
- **step** - step into the next statement of code. If it is a function call, this means the next shown step is the first line of the function. If it is not a function call, it has the same effect as next. (Note: many "basic" operations are actually function calls, for example, output statement cout << x, so make sure you want to step into these operator functions before you use step on these statements.)
- **backtrace** - see what the call stack looks like
- **quit** - exit the debugger. If the program is still running it will ask you if you really want to quit.
- show <topic> - the built-in help facility. show help will show instructions on how to use it.

There are many others. Consult the man page or using the built-in help facility to learn more.

Here is an example run using the test driver program from Project 1 for the Date class. Hardcopy of the program is attached to this handout. An on-line version is available on csserver. (Note that the tests for operator< are not complete.) Comments not part of the run are enclosed in square brackets [like this].

hwang@dh27pc:~cs215/project1$ gdb datetest
(Program greeting deleted)
(gdb) break main [often we just want to start at the beginning]
Breakpoint 1 at 0x8049c10: file datetest.cpp, line 23.
(gdb) run [datetest does not have any command line arguments]
    [if it did, they would be given here]
Starting program: /home/hwang/cs215/project1/datetest

Breakpoint 1, main () at datetest.cpp:23 [gdb displays the statement it is about to execute]
23        cerr << "Testing accessors and " << endl;
(gdb) next [execute the statement - output a message]
Testing accessors and [This is the output from the cerr statement]
24        cerr << " default constructed date" << endl;
(gdb) n [can abbreviate commands as long as they are unique]
    default constructed date
25        Date date1;
(gdb) [if no command uses the most recent previous on, n in this case]
26        int acctest = 0;
(gdb) print date1 [look at the value of date1, should be the default date]
$1 = {myMonth = 1, myDay = 1, myYear = 2008}
(gdb) n
27        int writetest = 0;
(gdb) acctest += TestAccessors (date1, 1, 1, 2008, "January 1, 2008", "Tuesday");
(gdb) step [step into TestAccessors function call, can be abbreviated s]
TestAccessors (date=0xbfd0410c, month=1, day=1, year=2008, name=0xbfd041fc, dofw=0xbfd041f8) at datetest.cpp:273
273        int errors = 0;
(gdb) n [usually want to go back to n, so that it doesn’t step into more function calls]
274        if (date.GetMonth() != month)
(gdb) [test is true so no body executed]
281        if (date.GetDay() != day)
(gdb)
289        if (date.GetYear() != year)
(gdb)
297        string s = date.StringName();
(gdb) s [step into StringName]
Date::StringName (this=0xbfd0410c) at date.cpp:38
38        ostringstream ostr;
(gdb) n

10/09/08
switch (myMonth)
    ostr << "January ";
break;
    ostr << myDay << ", " << myYear;
    return ostr.str();
} // end StringName
TestAccessors (date=@0xbfd0410c, month=1, day=1, year=2008,
    name=@0xbfd041fc,
    dofw=@0xbfd041f8) at datetest.cpp:298 [return from StringName]
    if (s != name)
    s = date.DayOfWeek();
    [n executes function call in one step]
    if (s != dofw)
    return errors;
} } // end TestAccessors
main () at datatest.cpp:29 [return from TestAccessors]
    writetest += TestWrite (date1, "1/1/2008");
(gdb) quit
The program is running. Exit anyway? (y or n) y
hwang@dh27pc:~/cs215/project1$