CSCI-1200 Data Structures — Fall 2025
Lecture 1 — Introduction to C++, STL, & Strings

Instructor
Professor Barb Cutler, `cutler@cs.rpi.edu`
Lally 302, 518-276-3274

Instructional Support Coordinator
Meredith Widman, `widmam@rpi.edu`

Personal/confidential matters – contact both of us: `ds_instructors@cs.rpi.edu`

For general course questions, please use the Discussion Forum:
https://submitty.cs.rpi.edu/courses/f25/csci1200/forum

Today

- Brief Discussion of Website & Syllabus
  http://www.cs.rpi.edu/academics/courses/fall25/csci1200/
- Introduction to C++ *Many students will be transitioning from prior courses with Python or Java*
- Crash Course in C++ Syntax
  http://www.cs.rpi.edu/academics/courses/fall25/csci1200/crash_course_cpp_syntax.php
- STL Strings are "smart" & easy-to-use (vs. C-style `char` arrays)

1.1 Transitioning from Python to C++ (from CSCI-1100 Computer Science 1)

- Python is a great language to learn the power and flexibility of programming and computational problem solving. This semester we will work in C++ and study lower level programming concepts, focusing on details including efficiency and memory usage.
- Outside of this class, when working on large programming projects, you will find it is not uncommon to use a mix of programming languages and libraries. The individual advantages of Python and C++ (and Java, and Perl, and C, and bash, ...) can be combined into an elegant (or terrifyingly complex) masterpiece.

1.2 Compiled Languages vs. Interpreted Languages

- C/C++ is a *compiled language*, which means your code is processed (compiled & linked) to produce a low-level machine language executable that can be run on your specific hardware. You must re-compile & re-link after you edit any of the files – although a smart development environment or `Makefile` will figure out what portions need to be recompiled and save some time (especially on large programming projects with many lines of code and many files). Also, if you move your code to a different computer you will usually need to recompile. Generally the extra work of compilation produces an efficient and optimized executable that will run fast.
- In contrast, many newer languages including Python, Java, & Perl are *interpreted languages*, that favor incremental development where you can make changes to your code and immediately run all or some of your code without waiting for compilation. However, an interpreted program will almost always run slower than a compiled program.
- These days, the process of compilation is almost instantaneous for simple programs, and in this course we encourage you to follow the same incremental editing & frequent testing development strategy that is employed with interpreted languages.
- Finally, many interpreted languages have a Just-In-Time-Compiler (JIT) that can run an interpreted programming language and perform optimization on-the-fly resulting in program performance that rivals optimized compiled code. Thus, the differences between compiled and interpreted languages are somewhat blurry.
- You will practice the cycle of coding & compilation & testing during Lab 1 & Homework 0. You are encouraged to try out different development environments (code editor & compiler) and quickly settle on one that allows you to be most productive. Ask the your lab TAs & mentors about their favorite programming environments! The course website includes many helpful links as well.
- As you see in today's handout, C++ has more required punctuation than Python, and the syntax is more restrictive. The compiler will proofread your code in detail and complain about any mistakes you make. Even long-time C++ programmers make mistakes in syntax, and with practice you will become familiar with the compiler's error messages and how to correct your code.

1.3 A Sample C++ Program: Find the Roots of a Quadratic Polynomial

```
#include <iostream>   // library for reading & writing from the console/keyboard
#include <cmath>      // library with the square root function & absolute value
#include <cstdlib>    // library with the exit function


// Returns true if the candidate root is indeed a root of the polynomial a*x*x + b*x + c = 0
bool check_root(int a, int b, int c, float root) {
  // plug the value into the formula
  float check = a * root * root + b * root + c;
  // see if the absolute value is zero (within a small tolerance)
  if (fabs(check) > 0.0001) {
    std::cerr << "ERROR:  " << root << " is not a root of this formula." << std::endl;
    return false;
  } else {
    return true;
  }
}

/* Use the quadratic formula to find the two real roots of polynomial.   Returns
true if the roots are real, returns false if the roots are imaginary.  If the roots
are real, they are returned through the reference parameters root_pos and root_neg. */
bool find_roots(int a, int b, int c, float &root_pos, float &root_neg) {
  // compute the quantity under the radical of the quadratic formula
  int radical = b*b - 4*a*c;
  // if the radical is negative, the roots are imaginary
  if (radical < 0) {
    std::cerr << "ERROR:  Imaginary roots" << std::endl;
    return false;
  }
  float sqrt_radical = sqrt(radical);
  // compute the two roots
  root_pos = (-b + sqrt_radical) / float(2*a);
  root_neg = (-b - sqrt_radical) / float(2*a);
  return true;
}

int main() {
  // We will loop until we are given a polynomial with real roots
  while (true) {
    std::cout << "Enter 3 integer coefficients to a quadratic function: a*x*x + b*x + c = 0" << std::endl;
    int my_a, my_b, my_c;
    std::cin >> my_a >> my_b >> my_c;
    // create a place to store the roots
    float root_1, root_2;
    bool success = find_roots(my_a,my_b,my_c, root_1,root_2);
    // If the polynomial has imaginary roots, skip the rest of this loop and start over
    if (!success) continue;
    std::cout << "The roots are: " << root_1 << " and " << root_2 << std::endl;
    // Check our work...
    if (check_root(my_a,my_b,my_c, root_1) && check_root(my_a,my_b,my_c, root_2)) {
      // Verified roots, break out of the while loop
      break;
    } else {
      std::cerr << "ERROR:  Unable to verify one or both roots." << std::endl;
      // if the program has an error, we choose to exit with a
      // non-zero error code
      exit(1);
    }
  }
  // by convention, main should return zero when the program finishes normally
  return 0;
}
```

1.4 Some Basic C++ Syntax

Crash Course in C++: Lesson #3

- Comments are indicated using `//` for single line comments and `/*` and `*/` for multi-line comments.
- `#include` asks the compiler for parts of the standard library and other code that we wish to use (e.g. the input/output stream function `std::cout`).
- `int main()` is a necessary component of all C++ programs; it returns a value (integer in this case) **and** it may have parameters.
- `{ }`: the curly braces indicate to C++ to treat *everything* between them as a unit.

1.5 The C++ Standard Library, a.k.a. "STL"

Crash Course in C++: Lesson #4

- The standard library contains types and functions that are important extensions to the core C++ language. We will use the standard library to such a great extent that it will feel like part of the C++ core language. `std` is a *namespace* that contains the standard library.
- I/O streams are the first component of the standard library that we see. `std::cout` ("console output") and `std::endl` ("end line") are defined in the standard library header file, `iostream`

1.6 A few notes on C++ vs. Java

- In Java, everything is an object and everything "inherits" from `java.lang.Object`. In C++, functions can exist outside of classes. In particular, the `main` function is never part of a class.
- Source code file organization in C++ does not need to be related to class organization as it does in Java. On the other hand, creating one C++ class (when we get to classes) per file is the *preferred* organization, with the *main* function in a separate file on its own or with a few helper functions.

1.7 Variables and Types

Crash Course in C++: Lesson #1

- A variable is an object with a name. A name is C++ identifier such as "`a`", "`root_1`", or "`success`".
- An object is computer memory that has a type. A type (e.g., `int`, `float`, and `bool`) is a structure to memory and a set of operations.
- For example, a `float` is an object and each `float` variable is assigned to 4 bytes of memory, and this memory is formatted according IEEE floating point standards for what represents the exponent and mantissa. There are many operations defined on floats, including addition, subtraction, printing to the screen, etc.
- In C++ and Java the programmer must specify the data type when a new variable is declared. The C++ compiler enforces type checking (a.k.a. *static typing*). In contrast, the programmer does not specify the type of variables in Python and Perl. These languages are *dynamically-typed* — the interpreter will deduce the data type at runtime.

1.8 Expressions, Assignments, and Statements

Crash Course in C++: Lesson #2 & #3

Consider the *statement*: `root_pos = (-b + sqrt_radical) / float(2*a);`

- The calculation on the right hand side of the `=` is an expression. You should review the definition of C++ arithmetic expressions and operator precedence from any reference textbook. The rules are pretty much the same in C++ and Java and Python.
- The value of this expression is assigned to the memory location of the float variable `root_pos`. Note also that if all expression values are type `int` we need a *cast* from `int` to `float` to prevent the truncation of integer division.

1.9 Conditionals and IF statements

Crash Course in C++: Lesson #5 & #7

- The general form of an if-else statement is

```
if (conditional-expression)
   statement;
else
   statement;
```

- Each `statement` may be a single statement, such as the `cout` statement above, or multiple statements delimited by curly braces `{ ... }` (a.k.a. scope).

1.10 `for` & `while` Loops

Crash Course in C++: Lesson #6 & #7

- Here is the basic form of a `for` loop:

```
for (expr1; expr2; expr3)
   statement;
```

  - `expr1` is the initial expression executed at the start before the loop iterations begin;
  - `expr2` is the test applied before the beginning of each loop iteration, the loop ends when this expression evaluates to `false` or `0`;
  - `expr3` is evaluated at the very end of each iteration;
  - `statement` is the "loop body"

- Here is the basic form of a `while` loop:

```
while (expr)
   statement;
```

  `expr` is checked before entering the loop and after each iteration. If `expr` evaluates the false the loop is finished.

1.11 Functions and Arguments

Crash Course in C++: Lesson #11

- Functions are used to:
  - Break code up into modules for ease of programming and testing, and for ease of reading by other people (never, ever, under-estimate the importance of this!).
  - Create code that is reusable at several places in one program and by several programs.

- Each function has a sequence of parameters and a return type. The function *prototype* below has a return type of `bool` and five parameters.

```
bool find_roots(int a, int b, int c, float &root_pos, float &root_neg);
```

- The order and types of the parameters in the calling function (the main function in this example) must match the order and types of the parameters in the function prototype.

1.12 Value Parameters and Reference Parameters

Crash Course in C++: Lesson #12

- What's with the `&` symbol on the 4th and 5th parameters in the `find_roots` function prototype?

- Note that when we call this function, we haven't yet stored anything in those two root variables.

```
float root_1, root_2;
bool success = find_roots(my_a,my_b,my_c, root_1,root_2);
```

- The first first three parameters to this function are *value parameters*.
  - These are essentially local variables (in the function) whose initial values are *copies* of the values of the corresponding argument in the function call.
  - Thus, the value of `my_a` from the main function is used to initialize `a` in function `find_roots`.
  - Changes to value parameters within the called function do NOT change the corresponding argument in the calling function.

- The final two parameters are *reference parameters*, as indicated by the `&`.
  - Reference parameters are just aliases for their corresponding arguments. No new objects are created.
  - As a result, changes to reference parameters are changes to the corresponding variables (arguments) in the calling function.
- In general, the "Rules of Thumb" for using value and reference parameters:
  - When a function (e.g., `check_root`) needs to provide just one result, make that result the return value of the function and pass other parameters by value.
  - When a function needs to provide more than one result (e.g., `find_roots`, these results should be returned using multiple reference parameters.
- We'll see more examples of the importance of value vs. reference parameters as the semester continues.

1.13 C-style Arrays

Crash Course in C++: Lesson #8

- An array is a fixed-length, consecutive sequence of objects all of the same type. The following declares an array with space for 15 double values. Note the spots in the array are currently *uninitialized*.

```
double a[15];
```

- The values are accessed through subscripting operations. The following code assigns the value 3.14159 to location `i=5` of the array. Here `i` is the *subscript* or *index*.

```
int i = 5;
a[i] = 3.14159;
```

- In C/C++, array indexing starts at 0.
- Arrays are fixed size, and each array knows *NOTHING* about its own size. The programmer must keep track of the size of each array (often storing it in an additional helper variable).

1.14 Character Arrays and String Literals (a.k.a., "C-style strings")

Crash Course in C++: Lesson #9

- In the line below `"Hello!"` is a *string literal*. The type of this value is a character array, which can be written as `char*` or `char[]`.

```
cout << "Hello!" << endl;
```

- A char array variable can be initialized as:

```
char h1[] = {'H', 'e', 'l', 'l', 'o', '!', '\0'};
char h2[] = "Hello!";
char* h3 = "Hello!";
```

  In all 3 examples, the variable stores 7 characters, the last one being the special null character, '`\0`'.

- The C language provides many functions for manipulating these "C-style strings". We won't cover them much in this course because "C++ style" STL string library is much more logical and easier to use.
- We will use `char` arrays for file names and command-line arguments, which you will use in Homework 0.

1.15 Editing `char` arrays & L-Values vs. R-Values

Crash Course in C++: Lesson #12

- Consider the simple code below. The `char` array `a` becomes `"Tim"`. No big deal, right?

```
char* a = "Kim";
char* b = "Tom";
a[0] = b[0];
```

- Let's look more closely at the line: `a[0] = b[0];` and think about what happens.

  In particular, what is the difference between the use of `a[0]` on the left hand side of the assignment statement and `b[0]` on the right hand side?

- Syntactically, they look the same. But,
  - The expression `b[0]` gets the char value, `'T'`, from location 0 in `b`. This is an *r-value*.
  - The expression `a[0]` gets a reference to the memory location associated with location 0 in `a`. This is an *l-value*.
  - The assignment operator stores the value in the referenced memory location.

  The difference between an *r-value* and an *l-value* will be especially significant when we get to writing our own operators later in the semester

- What's wrong with this code?

```
char* foo = "hello";
foo[2] = 'X';
cout << foo;
'X' = foo[3];
cout << foo;
```

  Your C++ compiler will complain with something like: "`non-lvalue in assignment`"

- Note: Strings in Python are immutable, and there is no difference between a string and a char in Python. Thus, `'a'` and `"a"` are both strings in Python, not individual characters.

  In C++ & Java, single quotes create a character type (exactly one character) and double quotes create a string of 0, 1, 2, or more characters.

1.16 About STL String Objects

Crash Course in C++: Lesson #9

- A `string` is an object type defined in the standard library to contain a sequence of characters.
- The string type, like all types (including `int`, `double`, `char`, `float`), defines an interface, which includes construction (initialization), operations, functions (methods), and even other types(!).
- When an object is created, a special function is run called a "constructor", whose job it is to initialize the object. There are several ways of constructing string objects:
  - By default to create an empty string:
    `std::string my_string_var;`
  - With a specified number of instances of a single char:
    `std::string my_string_var2(10, ' ');`
  - From another string:
    `std::string my_string_var3(my_string_var2);`
- The notation `my_string_var.size()` is a call to a function `size` that is defined as a **member function** of the `string` class. There is an equivalent member function called `length`.
- Input to string objects through streams (e.g. reading from the keyboard or a file) includes the following steps:
  1. The computer inputs and discards white-space characters, one at a time, until a non-white-space character is found.
  2. A sequence of non-white-space characters is input and stored in the string. This overwrites anything that was already in the string.
  3. Reading stops either at the end of the input or upon reaching the next white-space character (without reading it in).
- The (overloaded) operator '+' is defined on strings. It concatenates two strings to create a third string, without changing either of the original two strings.
- The assignment operation '=' on strings overwrites the current contents of the string.
- The individual characters of a string can be accessed using the subscript operator `[]` (similar to arrays).
  - Subscript 0 corresponds to the first character.
  - For example, given `std::string a = "Susan";` Then `a[0] == 'S'` and `a[1] == 'u'` and `a[4] == 'n'`.

- Strings define a special type `string::size_type`, which is the type returned by the string function `size()` (and `length()`).
  - The `::` notation means that `size_type` is defined within the scope of the `string` type.
  - `string::size_type` is generally equivalent to `unsigned int`.
  - You may see have compiler warnings and potential compatibility problems if you compare an `int` variable to `a.size()`.
- We regularly convert/cast between C-style & C++-style (STL) strings. For example:

```
std::string s1( "Hello!" );
char* h = "Hello!";
std::string s2( h );
std::string s3 = std::string(h);
```

  You can obtain the C-style string from a standard string using the member function `c_str`, as in `s1.c_str()`.

This seems like a lot to remember. Do I need to memorize this? Where can I find all the details on `string` *objects?*

1.17 Note about Strings in Java & the `new` Keyword

- Standard C++ library `std::string` objects behave like a combination of Java `String` and `StringBuffer` objects. If you aren't sure of how a `std::string` member function (or operator) will behave, check its semantics or try it on small examples (or both, which is preferable).
- Java objects must be created using `new`, as in:

```
String name = new String("Chris");
```

  This is not necessary in C++. The C++ (approximate) equivalent to this example is:

```
std::string name("Chris");
```

  Note: There is a `new` operator in C++ and its behavior is somewhat similar to the `new` operation in Java. We will study it in a couple weeks.

Whew! That was alot to speed through for one lecture!

We have just reviewed a wide range of C++ Syntax, which we will continue to practice in the Crash Course in C++ exercises, Lab 1, and Homework 0.