The Fundamentals of C++

**Preface**

Welcome to yet another of my many computer engineering “documents” or whatever you’d like to refer to them as. This one will be on the fundamentals of C++. Please note that this will not really be an “introduction” to C++. This is not a language that I believe you should start with. In fact, I’d highly recommend that you learn *at least* one other computer programming before learning C++, and preferably one which uses the OOP parasigm. Why, you ask? C++ builds on the principles of C, but adds classes, making it object-oriented. I have already written extensive notes for Java and C, which are two languages that will be very beneficial for better understanding C++. Now, if you absolutely only want to learn C++ and nothing else, I suggest you check out codecademy, and perhaps the C++ series of YouTube videos by “The Cherno” along with some other resources on the language. This document assumes that you’re at least familiar with C and will be split into 2 main sections: 1) The differences between C and C++, as well as additions to the language that C++ adds. 2) A more in-depth look at some of the functions, classes, and libraries available to you as a C++ programmer, as well as conventions for good C++ code. I really do enjoy C++ as a language, but perhaps its biggest flaw is that it has so many features, to the point of being a fault. By that, I mean there are so many ways to accomplish one task, that it is often difficult to be confident in your decisions as to how you go about solving specific problems. Not only that, but because there are so many things that you are able to do, the syntax often gets muddy, as certain convensions are reused, making it a bit of a difficult language to read. If, at the very least, you understand the core of C, and how it behaves, you should have a much easier time understanding C++, although I would also hope that you understand OOP principles as well, as those will certainly come into play.

**Introduction**

C++ was created in 1979 by a gentleman named Bjarne Stroustrup. Stroustrup, born in 1950, worked for Bell Laboratories (not to be confused with Bell Canada, the internet/telecommunications company). Stroutrup wanted to modernize the very popular language, C, written by Dennis Richie and Ken Thompson. C++ was originally called “C with Classes” and sought to take the fundamental principles and syntax of C, and simply wrap an object-oriented veil over it. C++ was not readily available for public use until about 1985. As of 2021, C is 49 years old, and while it is still taught in order to help better understand how computers interface with software, and how higher level languages are abstracted, it is not very practical for most modern applications. In many ways C++ simply “beats” C, because it essentially **is** C, but with more features. C++ is known for being very lightweight and fast, as was C. C++ is a very general purpose programming language that can create a wide range of applications. Examples might be game engines, operating systems, web backends, etc. With an idea of what C++ actually is, we can begin to look at some of the differences between it, and C.

**Basic Syntactical Differences**

I’ll just jump straight into this so that we don’t waste any time. Let’s begin with printing text to the console. In C, we generally used printf, which would format our variables into a character array, and print it to stdout. You will notice that since C++ uses classes, we often will see something known as namespaces used much more. The syntax for printing text in C++ is: std::cout << “text”; std is a namespace and the two colons :: are used to denote scope. std is C++’s standard library namespace, meaning that it is the namespace which encapsulates the definitions for functions in the standard library. cout and cin will need to be included by including <iostream> using, you guessed it, the #include directive. The function cout stands for character output stream. You’ll likely recognize the arrows << as being the redirection operator. Therefor, std::cout << “text”; can be read as ‘redirect the string “text” to the character output stream, defined in the std namespace‘. Without std, C++ would search what is called the global namespace, where cout is not defined, and a compile time error would be thrown. We might as well look at C++’s way of character input as well, while we’re discussing streams. std::cin >> variable; is how we take input from the user. stdin will await the enter key and will then redirect the character buffer to whichever variable we place afterwards. Two things to note here: 1) Notice that the redirection operator points towards the variable, whereas, with std::cout, the redirection operator pointed towards the stream. This is pretty intuitive but easily forgotten. 2) Unlike C, std::cin will cast whatever is in the character buffer to the appropriate datatype (which is quite intuitive, but perhaps sometimes frustrating if we require more controll), meaning we don’t need to worry about the datatype of the variable. When it comes to std::cout, we can additionally use std::endl (end line) which will do two things: insert a newline feed, as well as flush the stream. It is acceptable to just manually add a newline character to the end of your string, however, endl is generally considered cleaner, so I recommend you let any habits of appending \n die. eg. std::cout << “Hello world!” << std::endl;

Another minor difference is that we now have Booleans as a primitive type. These are declared with the keyword bool eg. bool isTrue = false; which is a welcomed convenience over using macros to create booleans in C. Other than that, loops, functions, conditionals, and logical operators are all identical.

**Strings and Vectors**

In C, Strings didn’t really exist as a data type, we would simply have to create a pointer to some block of memory where we could store characters, similar to an array. Arrays were also viewed as contiguous blocks of memory that carried data of a particular type and could be indexed due to the array acting as a pointer to the first block of memory in the sequence. As you know, in an object-oriented language like Java, we have instance variables/wrapper variables/object variables/reference variables (all names for the same thing). This applies in C++ as well, as Strings and vectors are considered to be instance variables since they are instances of their classes. A vector, in case you didn’t know, is mathematically defined as a list, however, you can just think of it as an array list in the context of C++. Whilst arrays do exist, vectors are common substitutes for arrays ie. you don’t see C style arrays all that often in C++. Unlike C, where we would have to #include <string.h> in order to use functions pertaining to strings, string is included in the std namespace of C++, and we can simply call functions pertaining to strings as we would for cout or cin. For example, I will define a string: std::string s = “This is a string”; and then I can perform actions on that string such as s.size(), s.clear(), s.replace, etc. Note that there is a header for string in C++ <string> if you require additional functions beyond those already included in std. Typically, C++ does not append the .h to the end of it’s libraries to differentiate between the classic C headers, and the redefined C++ ones. In other words, both <string.h> and <string> can be included into your C++ project; the <string.h> will contain the standard C definitions for it’s functions, and <string> will contain the C++ ones. Vectors are declared a bit differently. You may want to brush up on generic types from Java if you are not familiar with those, as vectors work in a similar fashion. A vector is declared as: std::vector<T> vec(size); where T is the datatype of the vector, and size is the size of the vector. Similar to most languages, if we know the data that we want to initialize the vector with, we can use curly braces and comma separated values eg. std::vector<float> prices = {9.8f, 3.4f, 6.2f, 0.98f}; Functions which apply to vectors are also stored in the std namespace, and so we can easily call those as well (in fact, string and vector share many function definitions). To add items to the end of a vector, we usually use the function push\_back(). To remove this element, we can use pop\_back(), although note that this function has no return value unlike most other programming languages. Arrays still exist in C++ and this might confuse you, however, you may just think of std::vector as you would an array list in Java. They are dynamically sized, and include many useful functions, but are slower and allocate more memory than might be necesarry.

**Classes and Objects**

This section of the document is where we are going to begin getting a bit more in-depth on the OOP aspects of C++. The basic syntax differences should hopefully be fairly easy if you’re coming from C, but classes and objects will get more confusing, even if you’re familiar with a language like Java. I’d like to review how we use header files, source files, and object files, as well as the compilation process before we begin looking at classes and objects. In our header files in C (.h files) we would declare our function prototypes (as well as maybe structs, macros, or extern variables) that needed to be accessible by multiple source files. The source files (.c files) would #include our header files, in order to define the bodies of the functions we declared in our headers. Then, in one of our source files, we would declare main() and our program would begin execution there. Typically, we would call most of our functions from main(). Object files (.o files) were created during the compilation process, and we’re essentially the machine code/object code of our .c files. Finally, during the linking process, the linker would combine all of the object files, as well as all the necessary libraries, and spit out an executable which we could run. This may sound off topic from classes and objects, but I assure you that they are quite entangled. In C++, header files can either remain as .h files, or can use the .hpp extension (I prefer .h, but you could make a case that .hpp is the better choice since it might lead someone to believe that it is compatible with C, when it may not be). Source files, however, must adopt the .cpp extension. Generally, as a convention, we declare our class outline and method declarations in the .h file, while the .cpp files essentially act the same as they did in C (but it is important to note that classes and function definitions can be declared within a .cpp file if you wish). A class can be created as follows:

class Animal {

private:

// Member variables

bool fur;

double run\_speed;

std::string name;

public:

// Member functions

void set\_run\_speed(double new\_run\_speed);

double get\_run\_speed();

std::string get\_name();

};

This class might reside in a file called animal.h (although it could just as easily reside in a source file as well, as mentioned). Instance variables are defined at the top as they usually are, in order to be used as the objects attributes/fields (often called “members” in C++). Notice the two access specifiers, public and private. The private specifier is assumed if no specifier is given (something we will discuss more shortly), however, I personally think it’s still good to explicitely declare your variables as being private because it takes no time and improves readability. We might use private for variables or methods which should only be accessible to the class itself, as well as subclasses/child classes, whereas public can be accessed from any file or class. In animal.cpp, we would then define the actual code block/body for each of the declared functions in the animal.h’s Animal class.

#include “animal.h”

void Animal::set\_run\_speed(double new\_run\_speed) {

run\_speed = new\_run\_speed;

}

double Animal::get\_run\_speed() {

return run\_speed;

}

std::string Animal::get\_name() {

return name;

}

and then in Main.cpp, we might have main() defined, and call these functions. Note how each method definition in the animal.cpp file starts with the “Animal” namespace. This is because, even though we #include “animal.h”, C++ always looks in the global namespace by default. If you recall, we received a compilation error for the same reason when it came to cout and cin. Namespaces actually very similar to packages in Java, only that namespaces do not represent physical folders like packages. One issue that we still have is that we have not created any instances of Animal i.e. no Animal objects! In order to do this, we need to create a constructor. Here is the revision to Animal.h:

class Animal {

private:

// Member variables

bool fur;

double run\_speed;

std::string name;

public:

// Constructor

Animal(bool new\_fur, std::string new\_name);

// Member function decls

void set\_run\_speed(double new\_run\_speed);

double get\_run\_speed();

std::string get\_name();

};

Here we add a constructor that will instantiate fur and name when an Animal object is created. This way we can initialize objects of type animal and set their attributes on creation. Note that if we don’t add our own constructor, there is a default constructor similar to Java, meaning that we can still create Animal objects without adding a constructor, but we won’t be able to modify any of their attributes without using setter methods. Let’s see what we need to add to animal.cpp:

#include “animal.h”

Animal::Animal(bool new\_fur, std::string new\_name) {

fur = new\_fur;

name = new\_name;

}

void Animal::set\_run\_speed(double new\_run\_speed) {

run\_speed = new\_run\_speed;

}

double Animal::get\_run\_speed() {

return run\_speed;

}

std::string Animal::get\_name() {

return name;

}

Now that we have a constructor, let’s take a look inside Main.cpp where we will declare an Animal object, initialize it’s instance variables, set it’s speed, and print the speed:

#include <iostream>

#include “animal.h”

int main() {

Animal cat(true, “cat”);

cat.set\_run\_speed(23.4);

//prints “cats can run 23.4 km/hour”

std::cout << cat.get\_name() << “s can run ” << cat.get\_run\_speed() << “ km/hour” << std::endl;

return 0;

}

As I mentioned, member variables within classes are private by default. This means that if in main(), I had tried to access one of cat’s attributes without using a public getter method, I would have received a compile time error. We’ve discussed constructors, but let’s discuss something a bit unique to C++. Unlike Java, C++ has no garbage collection (after all, it is still a somewhat old language). And in C, we would need to use the .free() method to deallocate memory and allow it to be overwritten. Objects in C++ may need to be deallocated and this is where destructors come into play. A destructor is similar to a constructor in that you are able to run whatever code you’d like before the actual destruction of the object occurs. This is typically used for any extra cleanup that needs to be done. Notice that while we can easily overload object constructors with different signatures, destructors are not the same. Destructors may *not* take any parameters. Since they cannot take any parameters, we cannot overload them with a different signature, therefore, there may only be one destructor per object.Here is the syntax for a destructor:

class Animal {

private:

//attribute/instance variables

bool fur;

double run\_speed;

std::string name;

public:

//Object constructor

Animal(bool new\_fur, std::string new\_name);

//Object destructor

~Animal();

//method declarations

void set\_run\_speed(double new\_run\_speed);

double get\_run\_speed();

std::string get\_name();

};

That’s right, it’s essentially like the constructor but without the parameters and it uses a tilde to differentiate itself. Similar to the default constructor, if you do not explicitely declare a destructor, C++ will call the default destructor which is hidden, however it is always a good idea to explicitely declare the destructor in your code. I don’t think it is necessary for me to show you that.

Hopefully, that gave you a very basic idea of how creating classes and objects works in C++. The Animal example was certainly not my finest, however, I aimed to get the core concepts in your head, and the rest will come with practice and viewing other examples as we continue throughout this paper. The basic idea is that .h/.hpp files generally house our classes (though not always; they can appear within our .cpp files as well), our functions/variables that we want to be accessible by multiple files should be under the public: scope identifier, and those functions require definitions in the respective .cpp file (with some exceptions that we are about to look into).

**Function Nuances in C++**

The next topic I’d like to cover is default arguments. I really like this feature in C++. Essentially if you have a function that has a particular parameter which will have the same value passed in most of the time, you can set that argument to have a default value so that if no argument is passed when calling the function, it presumes the default value. This is quite intuitive, we just set the parameter = to the value we want eg. void func\_foo(int param1, int param2 = 5) {...} Now if we call func\_foo() with only one argument eg. func\_foo(3), param2 will be 5, but if we call func\_foo() with both arguments, param2 will use the one we provided. Similar to var args (variable arguments), default arguments need to be placed last in the set of parameters. ie. you cannot have a default parameter followed by a non-default parameter.

A third addition which C did not have is function/method overloading. Obviously in C, we could sort of recreate the same principle by setting the parameters to void pointers and then casting the pointer to whichever type was necessary, as well as using var args, however, this is not actually the same as function overloading. If the overloaded functions signature differs by either having a different number of parameters or changing at least one of the datatypes of at least one parameter (can be achieved by rearranging the order of datatypes in the parameter list as well) then we can create a separate definition and overload that function.

Another feature that I’d like to mention is that we now have the ability to create function templates. If you recall back when we were discussing vectors, and how they utilize the same principle of generic types in Java in order to have dynamic typing, that is thanks to function templates. Function templates can be used on functions or classes, and allow the user the pass in a data type as a parameter which can then be used to dynamically change the data type of certain members, return types, or parameter types. This works in the same manner as inline methods. Instead of the keyword “inline”, we use “template <typename T>” followed by the function definition. Anytime ‘T’ is referenced, C++ interprets it according to it’s context similar to how dynamic/interpreted languages function. So, If I defined a function this way eg.

template <typename T>

T useless\_function(T some\_value) {

return some\_value;

}

Then the return type would be of type T, and some\_value would also be of type T. In order to call such a function, we would simply do “useless\_function<float>(7.0f);”. We can define as many types as we want by separating each with commas. For example “template <class T, class R, class Y>”. Also, just like functions, templates can have default parameters as well eg. template <class A=int, class B=std::string>. Templates are considered to be so powerful that C++ developers have created a concept called template meta-programming – essentially a way of creating a meta language within C++ using templates, which we will cover later on.

**Pointers vs References**

Raw pointers have been carried over from C with the same syntax that you’re familiar with. This includes an asterisk after the datatype and setting the pointer to the address of another variable using the ampersand (&). While we certainly can still use raw pointers in C++, the idiomatic equivallent is to use references. In essence, references accomplish the same goal as pointers, however, there are differences between the two. For references, we use the ampersand, similar to how we would use the asterisk when declaring a pointer e.g. datatype &ref\_name; in order to differentiate the variable as a reference. Outside of variable declaration, the ampersand always refers to a variable’s address. In order to set a reference to the address of another variable, we no longer require the address operation as you would with raw pointers. Instead, we just assign the reference to the value we want it to point to, and the address of the rvalue is retrieved implicitly. For example: int &ptr = num; Since ptr is a reference type, the assignment operator implicitly assigns the address of num to ptr. Another difference between pointers and references is that pointers would need to be dereferenced to alter the value at the pointers address. Once again, dereferencing is done implicitly, so we can access member variables or functions with a reference variable the same way that you would access them normally. One distinct difference between pointers and references is that references cannot be of type void. This was an intentional decision made by Stroustroup to avoid dangerous type-punning operations. Take a look at this example code:

#include <iostream>

#include <vector>

void addFive(int &x)

{

x += 5;

}

int main()

{

std::vector<int> vec { 10, 20, 30, 40 };

for (int i = 0; i < vec.size(); ++i)

addFive(vec[i]);

for (std::vector<int>::iterator i = vec.begin(); i < vec.end(); ++i)

std::cout << "Element " << i - vec.begin() << ": " << \*i << std::endl;

}

Output: 15, 25, 35, 45

Aside from the iterator stuff at the end, which we’ll cover later, this code is fairly self-explanatory. We loop through the vector that we create on the stack and implicitly pass the address of each element in said vector to our addFive function, which accepts a reference to the element and increments it. This change persists precisely because we are modifying the actual underlying data that we pass to the function, as opposed to a local copy of it on the stack.

It is not uncommon within C++ programs to have to use a C library which uses raw pointers. Knowing how to convert between references and raw pointers is a skill which is acquired gradually. Below is some additional code examples to demonstrate the differences between references and raw pointers:

inline void takeRawPtr(int \*x) { \*x += 5; }

inline void takeRawDblPtr(int \*\*x) { \*\*x += 5; }

inline int \*retRawPtr(int \*x) { \*x += 5; return x; }

inline void takeRef(int &x) { x += 5; }

inline void takeRef2RawPtr(int \*&x) { x += 5; }

inline int &retRef(int &x) { x += 5; return x; }

int main()

{

int x = 10;

int \*px = &x;

int &rx = x;

takeRawPtr(px);

takeRawDblPtr(&px);

px = retRawPtr(px);

takeRef(\*px);

takeRef2RawPtr(px);

rx = retRef(\*px);

return EXIT\_SUCCESS;

}

It should be noted that references, unlike raw pointers, must be assigned when they are declared, and cannot be assigned as nullptr or NULL. Once again, this is a safety feature to avoid attempts at dereferencing null pointers, which is a common mistake made in C programs.

You are surely aware by now that const prevents a value from being altered during runtime. In order to save a bit of computational cost, if we happen to have a parameter that we do not want to be altered during the execution of a function, we can replace it with a constant reference and this will save the computer having to make a local copy of the parameter in stack space. For example, instead of having a function func(int const dontAlterMe) {...} we could instead use a reference: func(int const &dontAlterMe) and that would use the address of the argument rather than making dontAlterMe it’s own local variable on the stack.

**Namespaces**

You may find as we write code in C++, that it becomes increasingly tiresome to explicitly state that our strings or vectors are from the standard C++ namespace. We can remedy this to some degree by telling C++ to search a particular namespace when it cannot find the definition of the function in the global namespace. By using the “using” directive, we can specify a namespace for C++ to search if namespace is explicitly provided. You will likely often see people add the line “using namespace std;” This tells C++ to search the standard namespace which means that we no longer have to type std::string, std::vector<datatype T>, std::cout, std::cin and so-forth. You may vaguely remember namespaces in C if you read through my notes, however, we didn’t discuss them much because aside from using the typedef keyword, they weren’t a large part of the syntax. As I briefly went over at some earlier point, namespaces essentially help us with scope. If we have a function foo() in library x and function foo() in library y, C++ needs to know which foo() we are referring to. We can also declare our own namespaces by wrapping our methods in a namespace wrapper. The general syntax is as follows:

namespace namespace\_name {

method() {

//code block

}

}

namespace is a keyword stating that we are declaring a new namespace, namespace\_name is the name that we want to call our namespace, and anything else within our namespace wrapper must now be prepended with namespace\_name eg. mynamespace::method(); This can get somewhat confusing since we can have nested namespaces, where we have namespaces within other namespaces by wrapping them within each other. Here’s an example:

namespace namespace1 {

foo() {

//code block

}

namespace namespace2 {

foo() {

//code block

}

}

}

Now, in order to call foo() from namespace1, we would call it as: namespace1::foo(); or if we wanted to call foo() from namespace2: namespace1::namespace2::foo(); We can use the “using” directive multiple times in our code to tell C++ to look in multiple namespaces, but this can become dangerous. For example, adding both “using namespace namespace1” and “using namespace namespace2” would cause ambiguity once again, which ironically, was supposed to be the goal of separating the 2 definitions of foo() into their own namespaces.

Because of the possibility of reintroducing ambiguity with the using directive, it is somewhat of a controversial topic within the C++ community. I think I’m inclined to agree that it should generally be discouraged in production code/professional code, however, if you know what you are doing then it may be acceptable in personal projects. Just note that when we “using namespace std;” we are sort of ignoring why namespaces exist in the first place so proceed with caution.

**The Unexpected Difference Between classes and structs**

If you’ve seen any C++ code being written, or have done a bit yourself, you have probably questioned what the difference between a struct and a class is. There is 1 and only 1 difference between a struct and a class in C++, and that is that classes are private by default, and structs are public by default. That is legitimately the only difference, and if you’re coming from C, you may be wondering how that could be the case. After all, structs don’t have constructors right? Well, in C++, a struct can indeed have constructors. So for example, I might have a class that appears like the following:

class MyClass {

std::string color;

int len;

bool setValue;

public:

MyClass(std::string color, int len, bool setValue);

~MyClass();

};

Notice how by default, the instance variables are private, and therefor outside the scope of other classes, hence why we need to explicitly state that the constructor/destructor are public methods. But if we simply take the exact same code and replace ‘class’ with ‘struct’:

struct MyClass {

std::string color;

int len;

bool setValue;

MyClass(std::string color, int len, bool setValue);

~MyClass();

};

Now it is no longer necessary to make the methods public, although without a private modifier, the instance variables are also public now. I will include another code snippet down below to demonstrate an actual implementation of this struct in a program:

#include <iostream>

struct MyClass {

*std*::*string* color;

int len;

bool setValue;

MyClass(*std*::*string* color, int len, bool setValue);

~MyClass();

};

MyClass::MyClass(*std*::*string* color, int len, bool setValue)

{

this->color = color;

this->len = len;

this->setValue = setValue;

}

MyClass::~MyClass() {

}

int main(void) {

struct MyClass myclass1("red", 5, true);

*std*::*cout* << myclass1.color << *std*::*endl*;

*std*::*cin*;

}

Here you can see that I’m creating an instance of a struct, not a class, and because my constructor is public by default, a can print out the myclass1’s color value. If I were to change struct to class however, then I get a compile time error “cannot access private member declared in class ‘MyClass’”.

So that then begs the question... If struct and class are essentially identical, why do we even need both? The simple answer to this is to maintain backwards compatibility with C, similar to why C++ keeps pointers. You could go your entire C++ career without ever creating a struct. However, it is up to you if or how you implement structs in your code, and this might be an advantage to you. For instance, I tend to use structs when I’m creating a block of data that is related, but doesn’t necesarilly constitute as a real life object. I’d use a class for more complex logic, if I want to be inheriting from other classes, or creating copy constructors, or creating member functions, etc. You do not have to do it my way, but I figured it’d be useful to share how you *could* use structs if you so choose to.

**Inheritance and Polymorphism:**

One of the most confusing aspects about C++ is in how it deals with OOP. You see, C++ was really one of the first languages to use the OOP paradigm, and as such, many oddities resulted. Let’s discuss a few of these oddities. The first oddity, which was intentially designed by Stroustrup, is that there is no universal *root class*. In Java, and most other modern OOP languages, there is a root class (in the case of Java, this would be the Object class), from which all classes implicitly derive from. Because there is no root class in C++, there is a much heavier reliance on templates and multiple inheritance. It also leads to other strange behaviour, such as type casting errors when trying to cast one class to another. In Java, if you try casting Foo to Bar when Bar does not inherit Foo, an exception will be thrown. In C++, this behavior is undefined. Multiple inheritance is perhaps the greatest sin commited by C++. In Java, a class may only extend one base class, but in C++, a class may derive from multiple classes. This leads to all sorts of confusion and disarray. A classic absurdity that arises from multiple inheritance is the infamous diamond problem. This occurs when one class inherits from 2 base classes, which themselves share the same base class. This relationship creates the shape of a diamond if layed out in a heirarchical view. The issue is that child classes automatically inherit all properties of their parents. This includes both member variables and member functions. If class A exists at the top of the diamond, and has a public member function named foo(), then its children B and C necessarily contain foo() as well. You may study this example for reference:

#include <iostream>

using namespace std;

class A

{

public:

void foo()

{

cout << "foo" << endl;

}

};

class B : public A

{};

class C : public A

{};

class D : public B, public C

{};

int main()

{

D d;

d.foo(); // Compile-time error: Member ‘foo’ found in multiple base-class subobjects of type ‘A’

return 0;

}

This can be resolved by either casting d to be of type B or C: ((C)d).foo(), using the scope resolution operator; or by using something known as virtual inheritance, which we will discuss in a bit. In general, it is recommended that you avoid multiple-inheritance.

I don’t want to gloss over what I said earlier about child classes inheriting all member variables and member functions from their parents, as I believe that this is crucial to your understanding of how inheritance works in C++. Since a derived class always inherits *all* properties of its parent, it necessarily contains at the very least, those same properties, plus, optionally, any additional properties which it declares. Very important to note, however, is that just because it inherits all properties of its parent, this does not mean that it inherits all *access* to those properties. For instance, a child class may not access private members which belong to its parent. Additionally, a child class cannot initialize members of the base class using an initializer list in its constructor.

**Virtual Functions / Virtual Inheritance:**

I mentioned that we would discuss using virtual inheritance as a means of bypassing the diamond problem, so that’s what we will now discuss. Basically, you can think of virtual inheritance as being the same as the @Override annotation in Java. Virtual inheritance does have subtle differences though, but the end result is the same. In order to understand how virtual inheritance works, you must understand something known as the V-table (virtual table). When a member function is declared virtual, the class which declared the virtual function immediately receives a pointer to a dyamically allocated v-table. The v-table is just a table of function pointers, which contains a pointer to each implementation of the virtual function. For instance, if class Foo has a virtual member function named foo(), it will recieve a pointer to a v-table that contains all current implementations of foo(). If a class named Bar extends Foo, since child classes automatically inherit all properties of its parent, Bar too, will inherit the same v-table. Then if Bar overloads the virtual function foo(), the v-table will now contain two function pointers: one to the version of foo() belonging to class Foo, and the other version of foo(), belonging to class Bar. The v-table will continue to dynamically grow in size as more and more classes overload the function foo().

So, what are the rules for virtual functions in C++? Basically, the base class must place the virtual keyword in front of any member function which it wishes to allow its children to overload. The child classes do *not* necessarily need to put the virtual keyword in front of the function overload, however, I strongly implore you to consider adding it, so that readers of your code understand that it is being overloaded! But technically speaking, at the bare minimum, the base class is the only class which must place the virtual keyword in front of the function.

**Pure Virtual Functions:**

In C++, there is no such thing as the *implements* keyword. You may be asking yourself then, “how do I implement interfaces in C++”? Interfaces as we’re used to them do not exist in C++, and therefore, cannot be implemented technically speaking. We can, however, use a class filled with functions known as *pure virtual functions*. This essentially turns the class into an abstract class, meaning that it cannot be instantiated. Other classes may extend this “abstract” class (remember, abstract classes don’t really exist in C++, but for all intents and purposes, classes which have pure virtual functions are abstract) and must then create a local definition for the pure virtual function, similar to an interface. So how do we mark a function as being “pure” virtual rather than just virtual? Well, the syntax is a bit odd, but we essentially set the function equal to 0 like so:

class Shape

{

public:

// Pure virtual function

virtual void calculateArea() = 0;

};

Now anything that derives from shape must implement its own definition for calculateArea(), otherwise a compile-time error will be thrown. Note that pure virtual function inheritance is very similar to virtual inheritance, with the key difference being that the parent class gets to have its own definition for a member function with normal virtual inheritance (as do its children), but with pure virtual inheritance, only the child classes get to have their own definitions.

**Virtual vs Override vs Final:**

C++ has both a virtual keyword and an override identifier. This can be confusing because new C++ programmers aren’t sure when it is necessary to use virtual or when to use override. The answer is that it really comes down to preference. The virtual keyword must be present before a member function declaration in the base class for it (and by extension, further overrides to it in descendent classes) to be placed in a vtable. Children of the base class may optionally include the virtual keyword, although it is not a requirement for the compiler. Similarly, the override identifier is entirely optional. If placed after a function declaration, the override keyword ensures that the function’s signature matches that of the base class’. The override identifier may only be used if the virtual keyword is also explicitly used. We can use the final identifier rather than, or in tandem with the override identifier, to suggest that any classes which inherrit from the current class may not override the member function, effectively preventing further additions to the vtable. Here is an example:  
  
class A {

virtual void foo() {

...

}

};

class B : public A {

virtual void foo() override {

...

}

};

class C : public B {

virtual void foo() final {

...

}

}

class D : public C {

virtual void foo() { // Error: Declaration of ‘foo’ overrides a ‘final’ function

...

}

};

**Setting Constructor Values**

If you’ve been paying attention closely, you may have noticed that in my example about struct vs. class, that I initialized the instance variables using the this-> keyword, similar to what you’d do in Java. We’ve also seen that you can simply name the instance variables one name, and then name the parameters another name, and then the compiler won’t get confused. Neither of these methods are ideal, however. Likely the most preferred method is to use an initializer list. The syntax for an initializer list is just a colon followed by each instance variable (comma separated) and input parameters next to each instance variable. For example, say I have a constructor:

private:

int instance1;

double instance2;

float instance3;

public:

//Constructor

Initializer(int inst1);

Then in the definition, we will use an initializer list to fill out the values for each instance variable:

Initializer::Initializer(int inst1) : instance1(inst1), instance2(67.54), instance3(10.4f) { ... }

For values that are to be set at the time of object creation (instance1 in this example), the parameter for the constructor should still be given a different name to avoid the compiler accidentally getting confused, however, for instance variables that will have a default value, they can be set directly in the initializer list. Initializer lists are also the method used to chain constructors. In C++ chaining constructors is known as “constructor delegation”. This was introduced in C++11 so previous versions will not support constructor delegation. Take the following example code:

class DelegationExample {

private:

int number;

public:

DelegationExample();

DelegationExample(int n);

void setNumber(int n);

};

DelegationExample::DelegationExample() : DelegationExample(5) {}

DelegationExample::DelegationExample(int n) {

setNumber(n);

}

void DelegationExample::setNumber(int n) {

number = n;

}

int main() {

DelegationExample delInstance = new DelegationExample();

return 0;

}

In this example, when delInstance gets created, the default constructor is called. Because DelegationExample is in our initializer list, however, before the code within the default constructor is called (not that there is any in this example), the constructor with the int n parameter is called, which calls setNumber and sets number = 5. The call stack will then begin to return and the stack pointer will go back to run any code within the default constructor since it didn’t get a chance to once the initalizer list took over. It should be noted that initializer lists can also work for objects but there can be a lot of confusion about that. The following code is from stack overflow and demonstrates the different ways in which you could create an object with an initializer list and default constructor and the outcome of each method:

class NewFoo

{

int x;

int y;

};

// Version 1:

class Bar1

{

NewFoo f;

};

// Version 2:

class Bar2

{

NewFoo f;

public:

Bar2() {} // f not in list.

};

// Version 3:

class Bar3

{

NewFoo f;

public:

Bar3() : f() {}

};

int main()

{

Bar1 b1a; // x and y not initialized.

Bar1 b1b = Bar1(); // x and y zero initialized.

Bar2 b2a; // x and y not initialized.

Bar2 b2b = Bar2(); // x and y not initialized.

Bar3 b3a; // x and y zero initialized.

Bar3 b3b = Bar3(); // x and y zero initialized.

}

<https://stackoverflow.com/questions/13238234/initializer-list-for-objects-with-default-constructor>

**Auto-generated functions:**

C++ generates special functions for us when we create a new class. These are obviously implicitly declared, as we do not need to create them ourselves. As we’ve seen, the default constructor is one such member function which is automatically generated for us if we do not create it explicitly. This makes sense, as without any constructor, we would not be able to instantiate the object (which we presumably want to do). C++ auto-generates other member functions, namely, a copy constructor, copy-assignment operator, and destructor. As you know, the copy constructor allows us to insert an instance of a class as an argument to the constructor when instantiating another instance of the same class. This is effectively the same as the copy-assignment operator, with the only difference being that one is a constructor, and the other is an operator.

**Deleting member functions:**

As of C++ 11, we have the ability to delete member functions, including special functions such as the ones listed in the previous section. We can do this using the delete keyword. We can also specify which constructor the compiler ought to treat as the default constructor with the default keyword. Here is an example taken from MSDN:

struct noncopyable

{

noncopyable() = default;

noncopyable(const noncopyable&) = delete;

noncopyable& operator=(const noncopyable&) = delete;

};

In this example, we set noncopyable() as the default constructor using the default keyword. This prevents the compiler from auto-generating the default constructor, though simply defining any constructor with no parameters would do prevent its generation in the same manner, even without use of the default keyword. We delete the auto-generated copy constructor, as well as the auto-generated copy operator, making it so that the instance cannot be copied. The deletion of member function applies to non-special functions as well, though its use is more nuanced. One reason you might use this is to avoid unwanted type promotion. For example, we can prevent type promotion from float to double from succeeding like so:

void callWithTrueDoubleOnly(float) = delete;

void callWithTrueDoubleOnly(double param) { return; }

In the example above, we prevent the call to callWithTrueDoubleOnly() from succeeding if the argument is a float (normally it would be promoted to a double to work with the function). Note, however, that an int will still work in this case. Rather than create another delete function for int, we can use templates to make this process cleaner:

template<typename T>

void callWithTrueDoubleOnly(T) = delete;

void callWithTrueDoubleOnly(double param) { return; }

**Entity Qualifiers**

static, const, and auto make a return in C++, however, with some slight modifications in comparison to C. As there are subtle differences with C++, it would be good for you to review them.

static: This keyword has not really changed since C. The static keyword still indicates to the compiler that the function or variable which proceeds will only be visible within the current translation unit. Of course, this means that we can still get away with things like defining the same function name twice, although this is still not recommended, as there are much better ways in C++ to do this (e.g., function overloading or namespaces). When a member function is declared static within a struct or class, it essentially makes the function no longer a member function of that class. In fact, a member function declared with the static keyword works exactly the same whether it is declared inside or outside of the class body. This is because all instances of a class have access to the same static member i.e., it is not reproduced for each instance of the class. The advantage to declaring a static member within a class is that we can access it using the class’ namespace e.g., Foo::staticMember;

const: const, as you probably know, discourages modifications to a variable’s memory or value. There happen to be ways to get around this, however, using pointers. Because it is fairly easy to change a const value, it is typically considered more of a promise, rather than an absolutely sure-fire way of protecting the data from being modified. The const keyword takes on a different meaning when qualifying a member function. As you can see in the code snippet below, placing const after the function signature prevents modifiication of a class’ member variables. This only applies to member functions and not static or global functions, for obvious reasons.   
  
Class Foo {

private:

int i;

public:

void mutate\_i() const {

this->i = 4; /\* Cannot assign to non-static data member within const member function ‘mutate\_i’ \*/

}

};

int main() {

Foo f;

f.mutate\_i();

}

auto: In C, auto was the implicit default for all stack-allocated variables. There was essentially never any reason to actually use it. I suppose this is why C++ decided to repurpose it for something entirely different. In C++, auto now tells the compiler to infer the type of a variable. Note that this is still type-safe since the type is inferred from the rvalue of the expression. Once inferred, its type cannot be dynamically altered unlike an interpreted language such as JavaScript or Python. For example, we cannot do:

auto a;

a = 5;

Rather, we must do auto a = 5;. a’s type will be determined by the compiler by looking at the rvalue (in this case, 5), which it knows is an int, and therefore, will make a of type int.

**Constexpr:**

The C++ keyword constexpr, pronounced “const expression” and introduced in C++ 11, refers to an expression which cannot be modified at runtime, and which is computed at compile time where possible. The constexpr keyword can be applied to variables, functions, and classes. For example, an arithmetic operation such as float z = exp(5, 3); would normally be computed at runtime, however, placing constexpr before the statement would cause the result to be computed at compile time. Note that const expressions will fail to evaluate if they contain variables or functions which are not also const expressions. For example, the following is an error:

int j = 0;

constexpr int k = j + 1; // Fails since j was not qualified with the constexpr keyword

**Minor Class differences between C++ and Java**

I just wanted to briefly discuss differences between classes in C++ and classes in Java. Firstly, a .java file must share the same name with the class that is defined within that file. The reasoning for this is so that you know which file contains which class. In C++ however, we can declare multiple classes within a single file, and because the file can be named anything, this makes it difficult to track down certain classes if we don’t stick to good practices. Another difference is that Java has a root object called Object. Any new class that you create extends Object by default. This can obviously be changed if you extend another class but ultimately the root class is always Object ie. everything is a child of Object.But why did Java make this (perhaps seemingly arbitrary) change? The answer is to avoid multiple inherritance. As you know, if you’ve programmed in Java, Java classes can only inherrit one parent class. However, C++, being one of the first OOP languages, allows for multiple inherritance. This is yet another controversial topic with C++ because multiple inherritance can lead to problems, specifically, the diamond problem. Here is a basic example of multiple inherritance being implemented, taken from <https://www.geeksforgeeks.org/multiple-inheritance-in-c/>:

#include <iostream>

using namespace std;

class A

{

public:

A() { cout << "A's constructor called" << endl; }

};

class B

{

public:

B() { cout << "B's constructor called" << endl; }

};

class C: public B, public A // Note the order

{

public:

C() { cout << "C's constructor called" << endl; }

};

int main()

{

C c;

return 0;

}

Output:

B’s constructor called

A’s constructor called

C’s constructor called

The object ‘c’ first extends/inherrits from B class, then A class. The ordering here determines the sequence of constructors that get called. Now the primary issue with multiple inherritance is called “The Diamond Problem”. This occurs when a child class extends two parent classes that both extend the same grandparent class. This creates a diamond where you have the child class extending two parents, and then those two parents extending one common class. The child will then recieve duplicate attributes since it’s parents will both inherrit the same properties from the grandparent class. This can be avoided using the virtual keyword which acts in sort of the same manner as #pragma once, only allowing the constructor of the grandparent class to be called once. Here is an example from the same website:

#include <iostream>

using namespace std;

class Person {

public:

Person(int x) { cout << "Person::Person(int ) called" << endl; }

Person() { cout << "Person::Person() called" << endl; }

};

class Faculty : virtual public Person {

public:

Faculty(int x):Person(x) {

cout<<"Faculty::Faculty(int ) called"<< endl;

}

};

class Student : virtual public Person {

public:

Student(int x):Person(x) {

cout<<"Student::Student(int ) called"<< endl;

}

};

class TA : public Faculty, public Student {

public:

TA(int x):Student(x), Faculty(x) {

cout<<"TA::TA(int ) called"<< endl;

}

};

int main() {

TA ta1(30);

}

**The new and delete keywords**

One of the primary differentiators between C and C++ is the addition of the new and delete keywords. New and delete are simply wrappers for the allocator and deallocator. The new keyword, when used for creating an instance of an object, will first allocate the appropriate amount of memory on the heap required for storing the object. Once allocated, new will automatically invoke the object’s constructor. Likewise the delete keyword will first invoke the object’s destructor, followed by a call to free(). Aside from instantiating objects, it is advised that you use new and delete for anything that is allocated in heap memory. We advise using new and delete over malloc() and free() for the simple reason that object’s may not be properly instantiated or destroyed since malloc() and free() ignore the object’s constructor and destructor. For creating heap allocated variables with primitive types, we still recommend new and delete simply for idiomaticity.

**The “this” keyword:**

As is the case with most OOP languages, we can use the this keyword to refer to the callee of a function. Whenever a member function is invoked, a pointer to its callee is implicitly passed to the function in the form of the *this* keyword. The this keyword is really just an alias for T\* const. For example, if we have a class named Foo, we can do Foo\* const foo = this; and that would be perfectly valid. Due to the fact that the this keyword is a pointer, it means that in order to access members of the instance using it, we must do so via the arrow operator. Something interesting is that the type of this changes within a scope marked as const. For example, if the member function foo() is marked const (i.e., we’re not allowed to modify member variables belonging to the class) then the type of this changes from Foo\* const to Foo const \* const. Note the additional const preventing mutation of the value that ‘this’ points to (our instance).

**C++ Casts vs C Casts**

Casting in C++ is actually quite the conundrum, believe it or not. To maintain backwards compatibility with C, C++ implements C-style casts (using brackets), but the implementation of C-style casts function quite differently than they do in C (more on that later). The typical way of casting in C was simply to use the round brackets with the target datatype in between to indicate that a value should be cast to that type before variable assignment. This would look something like the following: int asciiToDec = (int)’a’; Here the character ‘a’ would be cast to int, and become 97 before being assigned to variable asciiToDec. However, I also mentioned that C++ was different in that casting a void pointer must be explicit. C++ has 4 standard types of cast that must be declared explicitly. Note that C++ has a really odd feature that C didn’t have whereby we can swap the bracket position for the variable and the type being cast to in a C-style cast. Here is an example:  
  
char a = ‘a’;

// Regular method:

// int b = (int)a;

// Alternative method:

int b = int(a);

I say this now so that hopefully the syntax for C++ casts makes more sense as we look at them. Without further ado, we’ll begin with the first type of C++ cast: const\_cast.

- **const\_cast <type> (expr):** A const\_cast will actually alter whether or not a variable is read-only. Ex:

int main() {

const int a = 9;

const\_case<int&>(a) = 4; // This is a valid reassignment!

}

- **static\_cast <type> (expr):** The static\_cast performs a non-polymorphic cast. For example, it can cast a base class pointer to a derived class pointer. Unlike dynamic cast, static cast occurs at compile time. In C++, we do not need to explicitly use static cast to downcast from a base pointer to another class type which inherits from the base class, however, it is advised due to it being idiomatic. By using static cast, we differentiate between const, dynamic, and reinterpret casts, unlike C-style casts, which have no visual method of differentiating between each type. Ex:

class A {};

class B : public A {};

int main() {

A a;

B \*b = static\_cast<B\*>(&a);

return 0;

}

- **reinterpret\_cast <type> (expr**): The reinterpret\_cast is essentially the same as static cast, but whereas static cast has some compile time checks to ensure that the conversions of types make intuitive sense (e.g. casting a double to an int fails), reinterpret cast has no cares about bit twidling to fit a data type that does not make sense for the actual underlying data. This is essentially equivallent to storing some data of one type in a void pointer in C, and then reinterpreting that data as another type which may or may not make sense. This practice is known as type punning.

Ex:

#include <iostream>

struct A {

int m = 5;

};

struct B {

int m = 10;

};

int main() {

A a;

B \*b = reinterpret\_cast<B\*>(&a);

std::cout << b->m << std::endl; // Prints 5

return 0;

}

- **dynamic\_cast <type> (expr):** The dynamic cast performs a polymorphic cast. They operate at runtime and will return nullptr if the cast cannot succeed. The primary use of this is to downcast a pointer or reference to a base class into another class type which inherits from the base class. This is only possible if the base class has at least one virtual function. Ex:

class A {

public:

int i;

A() : i(7) {}

virtual void foo() {}

};

class B : public A {

public:

int i;

B() : i(3) {}

virtual void foo() override {}

};

int main() {

B \*b = new B();

auto b2a = dynamic\_cast<A\*>(b);

if (b2a) {

cout << b2a->i << endl; // Prints 7

}

}

**How C-Style Casts Work in C++**

As mentioned earlier, C-style casts have a different implementation within C++, which is obfuscated from you, the programmer, and which makes it rather dangerous to use. When performing a C-style cast, it will attempt the following C++ casts under the hood, in order, until one succeeds:

1. const\_cast

2. static\_cast

3. static\_cast\*, then const\_cast

4. reinterpret\_cast

5. reinterpret\_cast, then const\_cast

This is pretty dangerous right of the bat, because lets say that you have a const variable, and then you perform a C-style cast to change its type, however you forget to add the const qualifier in front of the type you’re casting to. The compiler will attempt the first C++ cast, being const\_cast. Since we’re changing the type of the variable, the compiler will not accept this and continue to step 2. Static cast will also fail, due to the fact that we accidentally failed to add the const qualifer to the type being cast to. The compiler will then try 3, which will succeed, since it will change the variables type and remove the const qualifier silently, without warning us! This is even scarier if we perform a C-style cast from one type to another when those two types have no relationship with one another. A static\_cast will at least catch this and throw an error, but a C-style cast will just end up doing a reinterpret\_cast whether you like it or not. This is why it is so heavily discouraged to use C-style casts in C++, and why its best to stick to the primary C++ casts even though they can be mildly more annoying to type out.

**Operator Overloading**

Another feature that I am actually quite fond of in C++ (and I think many other people like as well), is operator overloading. This is pretty unique to C++ (though I’m sure other languages have this ability), a sort of C++ staple if you will. Have you ever wished to be able to check if 2 objects are equivellant using ==, or perhaps wanted to increment every variable in an array using ++? This is possible in C++ because we can overload the definitions of operators and create our own definitions. In C++, operators are treated as functions, where the statements to the left and right of the operator are parameters. For example, let’s say that we wanted to set an object of a parent class A, equal to an instance of the derived class, B (A a = B;). This is possible in C++ even without operator overloading, but my point is that the argument on the left of the = sign is treated as a parameter, as is B on the right. If = was the name of the function, this would sort of be like calling “=(a, B);” although this is not completely accurate. Anyhoo, parameters are implicitly passed for unary operators such as ++, --, &, etc. And are explicitely passed for binary operators such as <<, >>, ==, =, +, -, /, \*, %, etc. So how do we actually overload an operator. Well, let me preface by saying that operator overloading can only be performed on classes. Within a class, we can declare a function using the “operator” keyword followed by the operator we want to overload.

Let’s first say that we have a class called OpOverload which contains a member variable called “number”. We want to be able to use the pre and post increment operators on instances of this class. For example, we want to be able to do the following:

OpOverload op;

++op; // pre increment

op++; // post increment

where the pre and post increment operators will increment the “number” associated with the instance of the class. I will define the class, and the definitions for pre and post increment and explain afterwards:

#include <iostream>

class OpOverload {

private:

int number;

public:

OpOverload() {}

~OpOverload() {}

// Pre-increment

OpOverload& operator++() {

this->number++;

return \*this;

}

// Post-increment

OpOverload& operator++(int) {

OpOverload& temp = \*this;

temp.number = this->number;

this->number++;

return temp;

}

};

int main() {

OpOverload op;

++op;

op++;

}

Alright, I’m sure this is quite confusing, which is to be expected because the syntax isn’t exactly pretty. Let’s look at pre-increment first. Recall that when we say ++op; we are sort of calling the ++ function with op as a parameter. Because the pre-increment operator is unary, op is passed implicitely. Notice that we also return a reference of the class. This is because whichever object we are pre-incrementing will essentially get set equal to the return value of the operator function. In other words, our instance, op, will get set = to the reference which is returned from our operator overload. So, we call ++(); then, we acces the member variable “number” which belongs to “this”, which is the implicit argument that we passed, op. But hold on, why are we using post-increment and not pre-increment on number? Well in this case it doesn’t matter if we increment the number using pre or post increment because all that matters is what gets returned. Either way, number gets incremented before we return the reference to OpOverload, so no harm is done. As for the return value, we must dereference “this” before returning it since it is a pointer and we need to convert it to a reference. I hope you followed that, but I understand that it’s confusing.

As for the post increment, well this is perhaps even more confusing. With post increment, we are also calling ++(); but then that begs the question: what differentiates the pre and post increment overloads? How does C++ know which you are overloading? Well this is why we placed that ominous int parameter: in order for C++ to recognize it as a post increment operator. You couldv’e given this parameter a name like int i, however, since it is useless and never gets used except to check that this is a post increment overload, I did not include the variable name. Moving onto the code, I create a temporary reference called temp which gets set equal to the implit parameter, op (once again, dereference to convert pointer to reference). We set the member variable “number” of the temporary reference = to the current value of “number” in “this”. Then, we increment this->number. Notice though that we return temp, and not this. Recall that I said only what we return matters. Now op will be set to the reference which was returned, which was temp, so it’s member, “number” will remain the same, until we access number on the next line, where it will be incremented since we said this->number++; It’s all very confusing I know. This stuff simply comes with practice.

Now let’s say that we want to create a sort of “toString()” function like we have in Java, and overload the redirection operator << to print the member variable “number” of the object like so std::cout << objectInstance << std::endl; Here is the code for that:

#include <iostream>

class OpOverload {

private:

int number;

public:

OpOverload() {}

~OpOverload() {}

//Pre-increment

OpOverload& operator++() {

this->number++;

return \*this;

}

//Post-increment

OpOverload& operator++(int) {

OpOverload& temp = \*this;

temp.number = this->number;

this->number++;

return temp;

}

friend std::ostream& operator<<(std::ostream& os, const OpOverload& op) {

os << "Object’s number variable = " << op.number;

return os;

}

};

int main() {

OpOverload op;

++op;

op++;

std::cout << op << std::endl;

}

Output: Object’s number variable = 2

Notice the use of a new keyword for this one: friend. friend allows external functions or objects to access our classes private and public members. Since ostream must access op.number, we must declare this function as a friend function, essentially granting ostream permission to access our private/public members in OpOverload. ostream is simply short for output stream and must be returned when overloading the output redirection operator. The const keyword is used for our reference of OpOverload since it’s data is not modified in this function.

Here is another slightly more complex example of operator overloading where we overload the post increment operator to act upon each member of a collection simultaneously (in this case std::Vector):

#include <vector>

#include <iostream>

template <class T>

class Vec : public std::vector<T> {

public:

std::vector<T> m\_vec;

Vec(std::initializer\_list<T> l) : m\_vec(l) {}

~Vec() {}

void operator++(int) {

for (auto& i : this->m\_vec) {

i++;

}

}

};

int main() {

Vec<float> vec{1.0f, 1.0f, 1.0f};

vec++;

for (auto i : vec.m\_vec)

std::cout << i << " ";

std::cout << std::endl;

return 0;

}

So, let me begin to explain myself here. Since we cannot access the vector class directly, the way in which we overload it is to extend it. This is why I create a new class called Vec : public std::vector<T>. You’ll notice that the class has been templated, i.e. it accepts any data type from the user when we create an instance of it, and anywhere where T arises, we can replace it with the data type that was passed. Within the constructor, we take in one parameter, std::initializer\_list<T> l. Initializer list is an object which accepts a list of variables enclosed in {} (an initializer list). This way, when we create an instance of the class, we can pass in a list of elements that we want to add to the vector. The class contains a member variable, m\_vec, which is the vector that we are actually modifying. The constructor also initializes m\_vec with the initializer list that we pass in. Finally, we get to the overload of the post increment operator. This time we don’t need to return anything since the instance of the object is not being modified, only m\_vec is. An advanced for loop says that for every element in this->m\_vec (remember that this is the implicit lefthand parameter; in our case vec), we want to create a reference called i (a reference to the current next element in m\_vec). The auto keyword is common in advanced for loops, and in this case makes our life easier since the data type could vary. Alternatively, we could’ve said for(T& i : this->m\_vec). Anyways, we simply increment each element with i++;

And that’s pretty much it. The second for loop simply accesses m\_vec and prints every element to standard out. Notice that this even works for characters, so if we created the vector Vec<char> vec{‘a’, ‘b’, ‘c’};

vec++;

Output: b c d //Pretty neat!

**Enums**

Enums are something that we discussed in C coding. In C it was as basic as just creating a new enum eg. enum faveGames = {Mario=1, Zelda, Pokemon}; and those would be assigned their appropriate numeric values. This works much the same in C++. Enums in C++ are very similar to enums in C, however, now we can create variables that are of type “enum type”. Sort of like typedef in C. These are really just integers, but are enforced as enums:

enum Color {

color\_red;

color\_blue;

};

Color color = color\_red;

Here we create a Color enum called color and assign it to color\_red. This is actually just treated as an int when making comparisons, however, the type of the color variable is treated as a Color enum. Additionally, enums can be set to any integer type. By default, enums are set to int, however, you may change this default to an unsigned char for example with the following syntax:

enum Characters : unsigned char {

A, B, C

};

It should be noted that enums are not the same as classes even though they share some of the same syntax. There are such things as enum classes, and these are essentially the same, but create scope for the enum values. So “class enum Characters” would make it so that A, B, C were within the enum Characters class and they would need to be referrenced as Characters::A for example. This is slightly better than defining regular enums since regular enums exist in the public namespace and can potentially cause collisions.

**Smart Pointers:**

For the most part, the C++ community has strayed away from raw pointers in favor of something called smart pointers in C++. Introduced in C++, smart pointers act in similar fashion to a Garbage Collector. The difference between the two is that a GC is a separate program which monitors the heap an performs deallocation when it notices that the memory can no longer be accessed, whereas smart pointers are individual instances which act as wrappers for raw pointers and then perform deallocation on their owned resource once they go out of scope. This works because the smart pointer’s destructor contains the call to delete, which then invokes delete on its owned resource i.e., the raw pointer. Smart pointers are very useful, as they alleviate the issue of memory leaks, which occur when we forget to clean up a resource after it’s done being used.

There are 3 primary types of smart pointer in C++:

**- unique\_ptr:** Unique pointer is kind of like specifying the restrict keyword in C. It specifies that the underlying raw pointer may only have one owner, which is the unique pointer that encapsulates it. A runtime exception will be thrown if another pointer attempts access at the underlying raw pointer owned by a unique\_ptr. Because shared pointers typically ought to be avoided, unique\_ptr is a good default choice for POCOs.

**- shared\_ptr:** A shared pointer allows for multiple owners of a single resource. It maintains a reference count and does not delete the underlying raw pointer until all owners have gone out of scope or have given up ownership (i.e. ref count becomes 0). Shared pointers actually occupy double the amount of bytes as a normal pointer, since the shared pointer object must contain a reference to the resource being pointed to as well as a reference to the number of shared pointers that currently hold ownership over the same resource. To showcase this, we can run the following code:

#include <iostream>

#include <memory>

using namespace std;

class A {};

int main()

{

unique\_ptr<A> uniq(new A);

cout << sizeof(uniq) << endl;

shared\_ptr<A> shared(new A);

cout << sizeof(shared) << endl;

return 0;

}

Output:

8

16

**- weak\_ptr:** Finally, weak pointer is a special-case pointer which is used in tandem with shared\_ptr. Weak pointers provide access to an object which is owned by one or more shared\_ptr instances, but does not participate in incrementing the reference count. This is useful when you want to observe an object, but don’t care if it remains alive or not. The most practical application for weak pointers is to avoid a problem known as reference cycles. An issue arises when two or more shared pointers create a cyclic loop by pointing towards each other. For example, if we wanted to create a circular linked list where node A points to B, which points to C, which then points back to A. This will pose an issue with shared pointers specifically, because so long as the reference count is not 0, they will not be destroyed (even after main() exits)! This is a memory leak, and the way that it can be resolved is by making one of those shared pointers a weak pointer. Because weak pointers don’t maintain a reference count of their own, they will be destroyed when they go out of scope.

Here is a program that produces a reference cycle:

**STL Containers:**

**Iterators:**

Iterators are a behavioral design pattern, which are becoming more and more common in modern programming. Iterators do as they suggest – they iterate over a collection or container in a linear (O(n)) way. You may ask yourself, why do we use iterators if we can just use a for loop? Well the argument typically goes as follows: For loops assume that the container uses indices when it may, in fact, not use them. For instance, if we iterate over a linked list using a for loop and stopping when i == ll.size, this works fine, but the end user may be inclined to assume that the ith element of the linked list can be accessed using the index operator, as suggested by the variable i e.g. ll[i]. As we know, linked lists cannot be accessed in this manner. An iterator, on the other hand, is agnostic about which type of collection it is iterating over. That being said, although I’ve heard the argument that iterators should be used whenever possible for consistency, I do not personally believe that they ought to replace for loops entirely. If using an interpreted language such as JS, for loops can provide a big performance boost. For compiled languages, your milleage may vary.

In C++, iterators can be implemented by including the iterator header file. Each STL container has its own namespace, which we can use to access that type’s implementation of iterator e.g., *vector<int>::iterator* or *list<string>::iterator.* Notice that this gives us a consistent interface for declaring iterators, which abstracts away the implementation details of iterator for each type of container. Once we have an iterator, we set it equal to the element that we want to begin iterating at. Since this is usually the first element, we can use the begin() function which is implemented by each STL container to access its first element. We can then check if the iterator is smaller than container.end(), and if so increment the iterator. Here is some example code:

#include <iostream>

#include <iterator>

#include <vector>

using namespace std;

int main()

{

vector<int> ar = { 1, 2, 3, 4, 5 };

vector<int>::iterator ptr;

cout << “The vector elements are: “;

for (ptr = ar.begin(); ptr < ar.end(); ptr++)

cout << \*ptr << “ “;

return 0;

}

Note that the iterator must necessarily override the default pre and post increment/decrement operators. For containers which are allocated in a contiguous memory region, the iterator may simply take advantage of pointer arithmetic and jump forward by the size of it’s type and then check the address that the iterator is pointing to to see if it matches with the final element’s address. For containers which may allocate memory anywhere on the heap, such as a linked list, it would naturally need to override the increment/decrement operators so that they point the iterator towards the next or previous element in the list. If what I just said doesn’t make sense, that’s perfectly fine. Treating iterators as a black box until you gain a better grasp on pointers and data structures is not such a bad thing.

**The Boost Library:**

**Structure Alignment:**

Perhaps a more niche topic is structure alignment. This is closely related to bit packing which we discussed in my C notes. This is heavily dependent on hardware, which is why it is important to understand what kind of architecture you’re developing for. Depending on the architecture that you’re working with, data types in C/C++ will have a certain byte alignment requirement. On a 32-bit machine, the byte alignment is 4 bytes. This means that the CPU reads 4 bytes at a time when accessing variables. This actually means that the order in which we declare our variables makes a difference for performance. For example, if we had a struct where we had a short int (2 bytes), followed by an int (4 bytes), the compiler would create 2 bytes of padding (just a bunch of NO-OP instructions/null values) after the short so that the integer could occupy a full byte alignment. This means that everytime we traverse the struct, the CPU is reading two null bytes of data. Therefore, we want to pack our bytes according to our alignment for maximal efficiency where possible.

If you recall bit fields/bit packing, you’ll remember that we can explicitely tell C/C++ to reduce the number of bits required for our variables. For example, we can declare multiple ints, but have them only occupy a certain number of bits like so:

int a;

unsigned int b:1, c:1, d:30;

and this would only take up 4 bytes which meets the byte alignment requirement. Notice, however, that if we were to declare our variables like so:

unsigned int a:1;

int b;

unsigned int c:1, d:30;

Then we would actually be occupying 12 bytes when we could’ve packed it into 8 since there will be 31 bits of padding after a, and 1 bit of padding after d (ie. We wasted 4 bytes!).

**File I/O:**

Arguably one of the more convoluted aspects of C++ is how it handles file I/O. If you’re primarily used to C, as I am, you know that there are really only two ways of creating a file, which is either to use the FILE struct and fopen to return a pointer to a file stream, or to use the open() system call to return a file descriptor. In C++, akin to Java, we have much more options when it comes to handling I/O. You are already familiar with the redirection operator (<< or >>), as it is used when we print statements to stdout using cout. Similarly, cin reads from stdin. As you are probably aware, these are both file streams, and their C++ implementation is contained within iostream. In order to handle file streams, we need to also include the fstream header file. This provides access to an object called ifstream. Well, technically speaking, ifstream is a typedef for basic\_ifstream<char>. If we provide a filepath as an argument to the ifstream’s constructor, the constructor will implicitly invoke basic\_ifstream’s implementation of open(). Alternatively, we can simply instantiate the ifstream and then call its open() member function later, when we are ready to actually open the file. The open() function signature is as follows: void open(const char \*filename, ios\_base::openmode mode = ios\_base::in);. Note the default argument “mode”, which is set to ios\_base::in if we do not set it manually. This parameter is actually an object of the bitmask member type openmode, which contains the constants in, out, binary, ate, app, and trunc. The flag bits can be OR’d together e.g., ios\_base::in | ios\_base::binary if we wanted to read the file as a binary file instead of text. Rather than checking if the call to open() returned nullptr or something like that (as we would do in C), we instead call the is\_open() member function, which will return a boolean indicating whether or not the file is currently open. At this point, there are multiple functions that we can invoke to read the contents of said file before eventually calling close() (note that the destructor for ifstream calls close() implicitly).

The easiest way to read the contents of a file as a string is to use the redirection operator on the ifstream object. For example:

#include <iostream>

#include <fstream>

int main()

{

std::string contents;

std::ifstream fs(“todo.txt”);

if (fs.is\_open())

{

while (fs.good())

{

fs >> contents;

}

}

fs.close();

std::cout << contents;

return 0;

}

Note the use of the good() member function. The redirection operator will read text until it encounters a white space character or newline character, meaning that if we have multiple lines we need to use a while loop in order to read each one. The good() member function essentially checks to see if we’ve reached the EOF, an error occurred, or if there was a R/W error. There is also the bad() member function, which only checks if a R/W error occurred. Alternatively, we could have just done while (fs), which is equivallent to while (fs.good())

Similar to ftell() in C, we can use the tellg() member function (stands for tell get) to return the current position of the file pointer as it moves through the stream. Once the entire file has been traversed, this function returns -1.

Another way we can read files is by using the getline() member function. This is pretty much just a wrapper around the redirection operator which returns the line that it read, not including the newline character.

I’ve also seen people use iterators to read an entire file as a string, which I have no opinion on in terms efficiency or readability or any of that. None-the-less, it’s still good to be able to know that there are other ways of doing the same operation. Here is an example of this:

std::string contents((std::istreambuf\_iterator<char>(ifs)),

(std::istreambuf\_iterator<char>()));

What’s happening here is likely not obvious to new C++ programmers. Essentially, std::string has a constructor which accepts two parameters, both of which are iterators. Here is the definition for said constructor: template <class InputIterator> string (InputIterator first, InputIterator last); This constructor actually just invokes the assign() member function which has the same method signature. What assign() does is replace the current contents of a string within the range specified by InputIterator first and InputIterator last. Recall that std::ifstream is just a typedef for basic\_ifstream<char>, and istreambuf\_iterator<char> is an input iterator that can read successive elements from a stream buffer. If you look up the constructor for istreambuf\_iterator on cppreference, you’ll see that the empty constructor constructs an end-of-stream iterator, and the constructor which accepts a std::basic\_istream initializes the iterator and stores the value of is.rdbuf(). I know that this is perhaps starting to become confusing, but is.rdbuf() is actually what tellg() returns (the file pointer’s position). So in other words, the empty iterator returns and end-of-stream iterator and the non-empty one returns the current position of the file pointer (which is pressumably 0 if we just opened the file). Then we construct a string from these two iterators using std::string’s assing() member function!

**String Streams:**

Continuing off of file streams, I wanted to also discuss string streams. String streams are kind of like StringBuilder in Java. They allow us to construct and parse strings. To use string streams, we include the <sstream> header. A Stringstream object can optionally accept an initial string as a constructor parameter. Since a Stringstream is a stream, similar to an ifstream, we can use the redirection operator to append strings or emit strings. Recall that the output redirection operator will emit up until the next whitespace character or newline is encountered.

**Lambda Expressions:**

Lambda expressions are very useful for things like event handlers or threads, where we essentially want to pass in an “inlined” function to the handler or thread dispatcher. They have a bit of a unique oddity in C++, which is that we need to “capture” variables which are outside of the scope of the lambda expression, but which we also require access to. We are able to specify whether we want to capture these variables by value or by reference. For example, assume we have a variable x, which we want to be able to modify in a lambda expression. This would like something like the following:

int x;

giveMeALambdaFunc([x]() { // Capture by value

// Do something with x

});

giveMeALambdaFunc([&x]() { // Capture by reference

// Do something with x

});

Allow me to mention that if we just want to capture all variables which are available within the enclosing scope, we can do [=] to capture all by value or [&] to capture all by reference. An empty capture clause indicates that no variables within the enclosing scope can be accessed. An alternative syntax for capturing x in the example above is shown below:

[&, x]{}; // Capture x by reference

[=, x]{}; // Error: = is the default

Note that the parameter list is optional. Here are the actual order of specifiers that we can include in a lambda expression:

1. capture clause

2. parameter list (optional)

3. The ‘mutable’ specifier (optional)

4. An exception specification e.g., throw() or noexcept (optional)

5. Arrow operator followed by trailing return type (optional)

6. The lambda body

A lambda expression using all 6 of these specifiers would look something like the following:

[&](int a, int b) mutable noexcept -> int {

return (a \* b);

}

**Templates:**

You may or may not have used templates in other programming langauges before. In case you have never used templates before, I will summarize their utility. A template typically allows us to omit type information for a given variable or function parameter so that we can polymorphically reuse code by substituting the type that we want when the class is instantiated or function is invoked. Due to the fact that we can substitute any type that we like by default, we refer to the placeholder value as ‘generics’. In order to create a template in C++, we simply use the ‘template’ keyword, followed by opening and closing karats (<>), which will enclose one of two additional keywords (I’ll demonstrate what these are in a moment), as well as the name of the generic or generics. Assume that we have a function which adds two values called add(). Now assume that we want to be able to have three versions of the add() function: one for ints, one for floats, and one for doubles. Your instinct may jump to using function overloads. Overloads can become cumbersome, however, since we must type out the same function three times with the only difference being in the types that we assign to our parameters. This is solved with templates with the following code:

template <typename T>

T add(T a, T b)

{

return a + b;

}

When we want to invoke the add() function, we must now do so as add<T>(), where ‘T’ is the type that we want to pass in to our template as the generic. Here is the full code snippet:

#include <iostream>

using namespace std;

template <typename T>

T add(T a, T b)

{

return a + b;

}

int main()

{

double result = add<float>(2.5f, 0.9f);

cout << result << endl;

return 0;

}

If you’re wondering why I’m using double for the result, it’s simply because that is the widest type for floating point, so whether we use int, float, or double, result can store the return value. Anyhow, notice that when we specified float as the generic, the return type as well as both parameters (a and b) became floats. The beauty of this approach is that the compiler will generate a unique add() function at compile time for each type variant that it is invoked with. So for instance, if I were to invoke add() twice, once with int and once with float, it would generate two variants of the function. It is essentially doing function overloading for us behind the scenes. The downside to templates is that we can really pass in any type that we like as the generic and the compiler may or may not complain. For example, heres the same code, but now I’ve changed the function arguments and type for result:

#include <iostream>

using namespace std;

template <typename T>

static T add(T a, T b)

{

return a + b;

}

int main()

{

string result = add<string>("a", "b");

cout << result << endl;

return 0;

}

As you can guess, this is still valid code, but using the add() function to concatonate strings was likely not the programmer’s intention. In order to prevent certain types from being substituted as the generic, we use the static\_assert() function. As you may or may not know, static\_assert(), unlike assert(), performs a compile-time assert. Since generics are substituted at compile-time, static\_assert() can check for types that we want to blacklist or filter out. Something clever that most people do is to use the ‘using’ statement to create a type alias. This can be a bit confusing to follow, especially since I haven’t covered some of the functions being used here, but I’ll throw out the example anyways:

#include <iostream>

#include <type\_traits>

using namespace std;

template <typename T>

using \_T = typename enable\_if<

is\_same<int, T>::value ||

is\_same<float, T>::value ||

is\_same<double, T>::value,

T

>::type;

template <typename T>

\_T<T> add(\_T<T> a, \_T<T> b) {

return a + b;

}

int main() {

float result = add<float>(1.0f, 1.0f);

cout << result << endl;

return 0;

}

Here, \_T is an alias that only exists if T is either an int, float, or double, otherwise the type is undefined and the compilation fails. This works due to enable\_if<T, U>::type, which I’ll let you read up on. The type that we substitute the generic type T for when calling add<T>() will be wrapped by the enable\_if<T, U>::type check, aliased as \_T.

I mentioned two keywords that we could use in the template declaration. The first, we’ve already seen, which is typename. The second is the class keyword. When simply defining templates, these two keywords may be used interchangeably. There are cases, however

**Error Handling**

**std::optional**

Similar to modern languages e.g. TypeScript or Rust, C++ attempts to mitigate the use of nullptr by providing us with the std::optional<T> template. This template, introduced in C++17, allows us to wrap some data in the std::optional type. An std::optional can either contain no data, or some data. In order to perform a check, we can use the std::optional’s member function has\_value(), which returns true if the std::optional does contain some data, or false otherwise. The value() member function is essentially a getter for the value if the std::optional does contain something. Rather than tediously checking has\_value() followed by value() each time we want to access an std::optional, we can use the value\_or() member function, which either accesses the underlying data if it exists, or returns the value that we provide as an argument if empty. Here is a code snippet example:

#include <iostream>

#include <optional>

using namespace std;

int main() {

optional<int> opt;

cout << "opt: " << opt.value\_or(5) << endl; // Prints 5

opt = 10; // This works because std::optional overloads the assignment operator

cout << "opt: " << opt.value\_or(5) << endl; // Prints 10

return 0;

}

**Async:**

In C++11 futures and promises were introduced for asynchronous programming. If you’ve not read my POSIX programming notes, that’s okay, I will quickly summarize the difference between asynchronous operations and parallel operations. Parallel operations can only take place if the system supports multiple threads. A task can be subdivided into smaller units of work which can execute concurrently (at the same time) on separate threads i.e., in parallel. Asynchronous programming is built upon coroutines. Coroutines are a fairly simple concept with a fancy name. In essence, each subdivision of the task is assigned to a coroutine (a function). Whenever the main thread encounters a statement of code that may block (halt) execution for an extended period of time (this is explicitly declared by the programmer), then it can execute one of the pending coroutines as a callback function while it waits for the original statement to finish. In this architecture, the thread is never blocking, but rather continuously completing tasks. The completion of tasks may occur out of order, hence why this is called asynchronous programming. At the end of the day, the thread may still have to block. For example, if it is waiting for the result of some lengthy math operation so that it can print it to stdout, and there are no pending coroutines, then the thread may remain idle. To get a bit more technical, multithreading involves the system scheduler, wherby a single timeslice is indicative of when context switching between threads occurs. With coroutines/async, the thread performs context switching when the programmer indicates that it should, thus providing less overhead, and more control. This has the added benefit of avoiding locking on shared data structures.

With that explanation out of the way, what are promises and futures? A promise is used by the producer/writer of the asynchronous operation. It is the equivallent of using the async keyword in languages such as JavaScript. It essentially says “I promise to return the result of this operation to you at a later point in time. Until then, you are free to perform context switching and handle another coroutine”. The future is used by consumers/readers, and its equivallent to the await keyword of JavaScript. It essentially says “The following statement can only be completed after receiving the promise that the producer/writer vowed to give me”. There are three methods of retrieving an std::future:

1. A call to std::async

2. Through a std::packaged\_task

3. Through a std::future

The latter two provide a member function get\_future() for this purpose, whereas a function marked with the async keyword implicitly returns a std::future.

**Multi-threading:**

**The Standard Library:**

If you have experience in C, but not a lot of experience with C++, as was my case, you may get confused by the standard library. As you hopefully know, the POSIX specification defines the implemenation details for the C standard library, and then your compiler is typically what actually fufills those implementation details and determines what happens in the case of behavior that is undefined in the POSIX spec and so forth. My point being, on Linux, we use the GNU version of libc (glibc) with compilers such as gcc, or we can use another implementation such as llvm-libc if using Clang, which is part of the LLVM project. This is not so different for C++. If using gcc/g++ as your compiler, you ought to use GNU’s implementation of the standard C++ library, called libstdc++. If using LLVM/Clang, you ought to use libc++. Even when compiling with Clang, libstdc++ is often still used on Linux. One of the main reasons that libc++ exists if because libstdc++ is under a GPL3 license, meaning that Apple cannot ship it, whereas libc++ is under a custom BSD-based license.

**The Standard Template Library (STL):**

Hopefully I’ve cleared up some of the confusion pertaining to the standard library, however, more confusion arises when discussing the Standard Template Library (STL). Many people assume that the standard C++ library and the STL library can be used interchangeably, or that the standard C++ library includes the STL library, however, neither of these are accurate statements. Let’s first discuss what the STL even is. The STL is a library which implements basic data structures and algorithms, such as vectors, maps, iterators, as well as algorithms for sorting, searching, and data manipulation. The STL was written sometime in the 80s, before the C++ language was even standardized. When C++ was eventually standardized by the ISO/IEC, the language committee decided to model parts of the C++ standard library after the STL. In other words, the standard C++ library implements its own version of the STL, which is similar, and may even look identical in some cases to the end user, though the implementation details may differ internally. Because they are so similar, some people (even those who understand the differences between the standard C++ library and the STL) will claim that the standard C++ library *is* the STL, when in fact, it is not.

**GCC vs G++:**

You may be confused as to when we use gcc and when we use g++. It’s simple really – a quick read of the man page for g++ will reveal that it is near identical to gcc, except built specifically with C++ in mind. It states that you ought to use g++ instead of gcc when compiling C++ projects, primarily due to potential conflicts with flags and so forth.